Springer Undergraduate Mathematics Series

S U M S

Nicolas Privault

Understanding Markov Chains

Examples and Applications

Second Edition

EXTRAS ONLINE

Springer

Springer Undergraduate Mathematics Series

More information about this series at http://www.springer.com/series/3423

Nicolas Privault

Understanding Markov Chains

Examples and Applications

Second Edition

Springer

Nicolas Privault
School of Physical and Mathematical
Sciences
Nanyang Technological University
Singapore
Singapore

Additional material to this book can be downloaded from http://extras.springer.com.

ISSN 1615-2085 ISSN 2197-4144 (electronic)
Springer Undergraduate Mathematics Series
ISBN 978-981-13-0658-7 ISBN 978-981-13-0659-4 (eBook)
https://doi.org/10.1007/978-981-13-0659-4

Library of Congress Control Number: 2018942179

Mathematics Subject Classification (2010): 60J10, 60J27, 60J28, 60J20

Printed on acid-free paper

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

Stochastic and Markovian modeling are of importance to many areas of science including physics, biology, engineering, as well as economics, finance, and social sciences. This text is an undergraduate-level introduction to the Markovian modeling of time-dependent randomness in discrete and continuous time, mostly on discrete state spaces, with an emphasis on the understanding of concepts by examples and elementary derivations. This second edition includes a revision of the main course content of the first edition, with additional illustrations and applications. In particular, the exercise sections have been considerable expanded and now contain 138 exercises and 11 longer problems.

The book is mostly self-contained except for its main prerequisites, which consist in a knowledge of basic probabilistic concepts. This includes random variables, discrete distributions (essentially binomial, geometric, and Poisson), continuous distributions (Gaussian and gamma), and their probability density functions, expectation, independence, and conditional probabilities, some of which are recalled in the first chapter. Such basic topics can be regarded as belonging to the field of "static" probability, i.e., probability without time dependence, as opposed to the contents of this text which is dealing with random evolution over time.

Our treatment of time-dependent randomness revolves around the important technique of first-step analysis for random walks, branching processes, and more generally for Markov chains in discrete and continuous time, with application to the computation of ruin probabilities and mean hitting times. In addition to the treatment of Markov chains, a brief introduction to martingales is given in discrete time. This provides a different way to recover the computations of ruin probabilities and mean hitting times which have been presented in the Markovian framework. Spatial Poisson processes on abstract spaces are also considered without any time ordering.

There already exist many textbooks on stochastic processes and Markov chains, including [BN96, Çin75, Dur99, GS01, JS01, KT81, Med10, Nor98, Ros96, Ste01]. In comparison with the existing literature, which is sometimes dealing with structural properties of stochastic processes via a more compact and abstract treatment, the present book tends to emphasize elementary and explicit calculations instead of quicker arguments that may shorten the path to the solution, while being sometimes difficult to reproduce by undergraduate students.

Some of the exercises have been influenced by [Çin75, JS01, KT81, Med10, Ros96] and other references, while a number of them are original, and their solutions have been derived independently. The *problems*, which are longer than the exercises, are based on various topics of application. This second edition only contains the answers to selected exercises, and the remaining solutions can be downloaded in a solution manual available from the publisher's Web site, together with Python and R codes. This text is also illustrated by 41 figures.

Some theorems whose proofs are technical, as in Chaps. 7 and 9, have been quoted from [BN96, KT81]. The contents of this book have benefited from numerous questions, comments, and suggestions from undergraduate students in Stochastic Processes at the Nanyang Technological University (NTU) in Singapore.

Singapore, Singapore
March 2018

Nicolas Privault

Contents

List of Figures

List of Tables

Introduction

A stochastic[1] process is a mathematical tool used for the modeling of time-dependent random phenomena. Here, the term "stochastic" means random and "process" refers to the time-evolving status of a given system. Stochastic processes have applications to multiple fields and can be useful anytime one recognizes the role of randomness and unpredictability of events that can occur at random times in, e.g., physical, biological, or financial system.

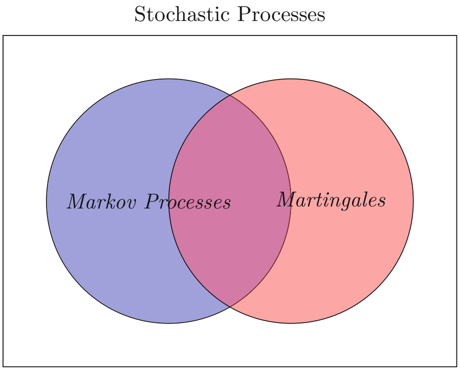
For example, in applications to physics one can mention phase transitions, atomic emission phenomena, etc. In biology, the time behavior of live beings is often subject to randomness, at least when the observer has only access to partial information. This latter point is of importance, as it links the notion of randomness to the concept of information: What appears random to an observer may not be random to another observer equipped with more information. Think, for example, of the observation of the apparent random behavior of cars turning at a crossroad versus the point of view of car drivers, each of whom are acting according to their own decisions. In finance, the importance of modeling time-dependent random phenomena is quite clear, as no one can make definite predictions for the future moves of risky assets. The concrete outcome of random modeling lies in the computation of *expectations* or *expected values*, which often turn out to be more useful than the probability values themselves. An average or expected lifetime, for example, can be easier to interpret than a (small) probability of default. The long-term statistical behavior of random systems, which also involves the estimation of expectations, is a related issue of interest.

Basically, a stochastic process is a time-dependent family $(X_t)_{t\in T}$ of random variables, where t is a time index belonging to a parameter set or timescale T. In other words, instead of considering a single random variable X, one considers a whole family of random variables $(X_t)_{t\in T}$, with the addition of another level of technical difficulty. The timescale T can be finite (e.g., $T = \{1, 2, \ldots, N\}$) or countably infinite (e.g. $T = \mathbb{N} = \{0, 1, 2, \ldots\}$) or even uncountable (e.g. $T = [0, 1]$,

[1]From the Greek "$\sigma\tau o'\chi o\varsigma$" (stokhos), meaning "guess", or "conjecture".

$T = \mathbb{R}_+$). The case of uncountable T corresponds to continuous-time stochastic processes, and this setting is the most theoretically difficult. A serious treatment of continuous-time processes would in fact require additional background in *measure theory*, which is outside of the scope of this text. Measure theory is the general study of measures on abstract spaces, including probability measures as a particular case, and allows for a rigorous treatment of integrals via integration in the Lebesgue sense. The Lebesgue integral is a powerful tool that allows one to integrate functions and random variables under minimal technical conditions. Here we mainly work in a discrete-time framework that mostly does not require the use of measure theory.

That being said, the definition of a stochastic process $(X_t)_{t \in T}$ remains vague at this stage since virtually *any* family of random variables could be called a stochastic process. In addition, working at such a level of generality without imposing any structure or properties on the processes under consideration could be of little practical use. As we will see later on, stochastic processes can be classified into two main families:



– *Markov Processes*
  Roughly speaking, a process is *Markov* when its statistical behavior after time t can be recovered from the value X_t of the process at time t. In particular, the values X_s of the process at times $s \in [0, t)$ have no influence on this behavior as long as the value of X_t is known.
– *Martingales*
  Originally, a martingale is a strategy designed to win repeatedly in a game of chance. In mathematics, a stochastic process $(X_t)_{t \in \mathbb{R}_+}$ is a martingale if the best possible estimate at time s of its future value X_t at time $t > s$ is simply given by X_s. This requires the careful definition of a "best possible estimate," and for this we need the tool of conditional expectation which relies on estimation in the mean square sense. Martingale are useful in physics and finance, where they are linked to the notion of equilibrium.

Time series of order greater than one form another class of stochastic processes that may have neither the Markov property nor the martingale property in general.

The outline of this text is as follows. After reviewing in Chap. 1 the probabilistic tools required in the analysis of Markov chains, we consider simple gambling problems in Chap. 2, due to their practical usefulness and to the fact that they only require a minimal theoretical background. Next, in Chap. 3, we turn to the study of discrete-time random walks with infinite state space, which can be defined as stochastic processes with independent increments, without requiring much abstract formalism. In Chap. 4, we introduce the general framework of Markov chains in discrete time, which includes the gambling process and the simple random walk of Chaps. 2 and 3 as particular cases. In the subsequent Chaps. 5, 6, and 7, Markov chains are considered from the point of view of first-step analysis, which is introduced in detail in Chap. 5. The classification of states is reviewed in Chap. 6, with application to the long-run behavior of Markov chains in Chap. 7, which also includes a short introduction to the Markov chain Monte Carlo method. Branching processes are other examples of discrete-time Markov processes which have important applications in life sciences, *e.g.*, for population dynamics or the control of disease outbreak, and they are considered in Chap. 8. Then in Chap. 9, we deal with Markov chains in continuous time, including Poisson and birth and death processes. Martingales are considered in Chap. 10, where they are used to recover in a simple and elegant way the main results of Chap. 2 on ruin probabilities and mean exit times for gambling processes. Spatial Poisson processes, which can be defined on an abstract space without requiring an ordered time index, are presented in Chap. 11. Reliability theory is an important engineering application of Markov chains, and it is reviewed in Chap. 12. All stochastic processes considered in this text have a discrete state space and discontinuous trajectories.

Chapter 1
Probability Background

In this chapter we review a number of basic probabilistic tools that will needed for the study of stochastic processes in the subsequent chapters. We refer the reader to e.g. [Dev03, JP00, Pit99] for additional background on probability theory.

1.1 Probability Spaces and Events

We will need the following notation coming from set theory. Given A and B to abstract sets, "$A \subset B$" means that A is contained in B, and in this case, $B \setminus A$ denotes the set of elements of B which do not belong to A. The property that the element ω belongs to the set A is denoted by "$\omega \in A$", and given two sets A and Ω such that $A \subset \Omega$, we let $A^c = \Omega \setminus A$ denote the *complement* of A in Ω. The finite set made of n elements $\omega_1, \ldots, \omega_n$ is denoted by $\{\omega_1, \ldots, \omega_n\}$, and we will usually make a distinction between the element ω and its associated singleton set $\{\omega\}$.

A probability space is an abstract set Ω that contains the possible outcomes of a random experiment.

Examples

(i) Coin tossing: $\Omega = \{H, T\}$.
(ii) Rolling one die: $\Omega = \{1, 2, 3, 4, 5, 6\}$.
(iii) Picking one card at random in a pack of 52: $\Omega = \{1, 2, 3, \ldots, 52\}$.
(iv) An integer-valued random outcome: $\Omega = \mathbb{N} = \{0, 1, 2, \ldots\}$.
In this case the outcome $\omega \in \mathbb{N}$ can be the random number of trials needed until some event occurs.
(v) A nonnegative, real-valued outcome: $\Omega = \mathbb{R}_+$.
In this case the outcome $\omega \in \mathbb{R}_+$ may represent the (nonnegative) value of a continuous random time.

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4_1

(vi) A random continuous parameter (such as time, weather, price or wealth, temperature,...): $\Omega = \mathbb{R}$.
(vii) Random choice of a continuous path in the space $\Omega = \mathcal{C}(\mathbb{R}_+)$ of all continuous functions on $\mathbb{R}_+$.
In this case, $\omega \in \Omega$ is a function $\omega : \mathbb{R}_+ \longrightarrow \mathbb{R}$ and a typical example is the graph $t \longmapsto \omega(t)$ of a stock price over time.

Product Spaces:

Probability spaces can be built as product spaces and used for the modeling of repeated random experiments.

(i) Rolling two dice: $\Omega = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$.

In this case a typical element of Ω is written as $\omega = (k, l)$ with $k, l \in \{1, 2, 3, 4, 5, 6\}$.
(ii) A finite number n of real-valued samples: $\Omega = \mathbb{R}^n$.

In this case the outcome ω is a vector $\omega = (x_1, \ldots, x_n) \in \mathbb{R}^n$ with n components.

Note that to some extent, the more complex Ω is, the better it fits a practical and useful situation, e.g. $\Omega = \{H, T\}$ corresponds to a simple coin tossing experiment while $\Omega = \mathcal{C}(\mathbb{R}_+)$ the space of continuous functions on $\mathbb{R}_+$ can be applied to the modeling of stock markets. On the other hand, in many cases and especially in the most complex situations, we will *not* attempt to specify Ω explicitly.

Events

An event is a collection of outcomes, which is represented by a subset of Ω.

The collections $\mathcal{G}$ of events that we will consider are called σ-algebras, and assumed to satisfy the following conditions.

(i) $\emptyset \in \mathcal{G}$,
(ii) For all countable sequences $A_n \in \mathcal{G}$, $n \geq 1$, we have $\bigcup_{n\geq 1} A_n \in \mathcal{G}$,
(iii) $A \in \mathcal{G} \Longrightarrow (\Omega \setminus A) \in \mathcal{G}$,

where $\Omega \setminus A := \{\omega \in \Omega \ : \ \omega \notin A\}$.

Note that Properties (ii) and (iii) above also imply

$$\bigcap_{n\geq 1} A_n = \left(\bigcup_{n\geq 1} A_n^c \right)^c \in \mathcal{G}, \tag{1.1.1}$$

for all countable sequences $A_n \in \mathcal{G}$, $n \geq 1$.

The collection of all events in Ω will often be denoted by $\mathcal{F}$. The empty set $\emptyset$ and the full space Ω are considered as events but they are of less importance because Ω corresponds to "any outcome may occur" while $\emptyset$ corresponds to an absence of outcome, or no experiment.

In the context of stochastic processes, two σ-algebras $\mathcal{F}$ and $\mathcal{G}$ such that $\mathcal{F} \subset \mathcal{G}$ will refer to two different amounts of information, the amount of information associated to $\mathcal{F}$ being here lower than the one associated to $\mathcal{G}$.
The formalism of σ-algebras helps in describing events in a short and precise way.

Examples

(i) $\Omega = \{1, 2, 3, 4, 5, 6\}$.
The event $A = \{2, 4, 6\}$ corresponds to

"the result of the experiment is an even number"

(ii) Taking again $\Omega = \{1, 2, 3, 4, 5, 6\}$,

$$\mathcal{F} := \{\Omega, \emptyset, \{2, 4, 6\}, \{1, 3, 5\}\}$$

defines a σ-algebra on Ω which corresponds to the knowledge of parity of an integer picked at random from 1 to 6.

Note that in the set-theoretic notation, an event A is a subset of Ω, i.e. $A \subset \Omega$, while it is an element of $\mathcal{F}$, i.e. $A \in \mathcal{F}$. For example, we have $\Omega \supset \{2, 4, 6\} \in \mathcal{F}$, while $\{\{2, 4, 6\}, \{1, 3, 5\}\} \subset \mathcal{F}$.

(iii) Taking

$$\mathcal{G} := \{\Omega, \emptyset, \{2, 4, 6\}, \{2, 4\}, \{6\}, \{1, 2, 3, 4, 5\}, \{1, 3, 5, 6\}, \{1, 3, 5\}\} \supset \mathcal{F},$$

defines a σ-algebra on Ω which is bigger than $\mathcal{F}$, and corresponds to the parity information contained in $\mathcal{F}$, completed by the knowledge of whether the outcome is equal to 6 or not.

(iv) Take

$$\Omega = \{H, T\} \times \{H, T\} = \{(H, H), (H.T), (T, H), (T, T)\}.$$

In this case, the collection $\mathcal{F}$ of all possible events is given by

$$\begin{aligned}\mathcal{F} = \{&\emptyset, \{(H, H)\}, \{(T, T)\}, \{(H, T)\}, \{(T, H)\}, \\ &\{(T, T), (H, H)\}, \{(H, T), (T, H)\}, \{(H, T), (T, T)\}, \\ &\{(T, H), (T, T)\}, \{(H, T), (H, H)\}, \{(T, H), (H, H)\}, \\ &\{(H, H), (T, T), (T, H)\}, \{(H, H), (T, T), (H, T)\}, \\ &\{(H, T), (T, H), (H, H)\}, \{(H, T), (T, H), (T, T)\}, \Omega\}.\end{aligned} \tag{1.1.2}$$

Note that the set $\mathcal{F}$ of all events considered in (1.1.2) above has altogether

$1 = \binom{n}{0}$ event of cardinality 0,

$4 = \binom{n}{1}$ events of cardinality 1,
$6 = \binom{n}{2}$ events of cardinality 2,
$4 = \binom{n}{3}$ events of cardinality 3,
$1 = \binom{n}{4}$ event of cardinality 4,

with $n = 4$, for a total of

$$16 = 2^n = \sum_{k=0}^{4} \binom{4}{k} = 1 + 4 + 6 + 4 + 1$$

events. The collection of events

$$\mathcal{G} := \{\emptyset, \{(T, T), (H, H)\}, \{(H, T), (T, H)\}, \Omega\}$$

defines a sub σ-algebra of $\mathcal{F}$, associated to the information "the results of two coin tossings are different".

Exercise: Write down the set of all events on $\Omega = \{H, T\}$.

Note also that (H, T) is different from (T, H), whereas $\{(H, T), (T, H)\}$ is equal to $\{(T, H), (H, T)\}$.

In addition, we will usually make a distinction between the *outcome* $\omega \in \Omega$ and its associated *event* $\{\omega\} \in \mathcal{F}$, which satisfies $\{\omega\} \subset \Omega$.

1.2 Probability Measures

A probability measure is a mapping $\mathbb{P} : \mathcal{F} \longrightarrow [0, 1]$ that assigns a probability $\mathbb{P}(A) \in [0, 1]$ to any event $A \in \mathcal{F}$, with the properties

(a) $\mathbb{P}(\Omega) = 1$, and
(b) $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$, whenever $A_k \cap A_l = \emptyset$, $k \neq l$.

Property (b) above is named the *law of total probability*. It states in particular that we have

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) = \mathbb{P}(A_1) + \cdots + \mathbb{P}(A_n)$$

when the subsets $A_1, \ldots, A_n$ of Ω are disjoints, and

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) \tag{1.2.1}$$

if $A \cap B = \emptyset$. We also have the *complement rule*

$$\mathbb{P}(A^c) = \mathbb{P}(\Omega \setminus A) = \mathbb{P}(\Omega) - \mathbb{P}(A) = 1 - \mathbb{P}(A).$$

When A and B are not necessarily disjoint we can write

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

The triple

$$(\Omega, \mathcal{F}, \mathbb{P}) \tag{1.2.2}$$

was introduced by A.N. Kolmogorov (1903–1987), and is generally referred to as the Kolmogorov framework.

A property or event is said to hold $\mathbb{P}$-*almost surely* (also written $\mathbb{P}$-*a.s.*) if it holds with probability equal to one.

Example

Take

$$\Omega = \big\{(T, T), (H, H), (H, T), (T, H)\big\}$$

and

$$\mathcal{F} = \{\emptyset, \{(T, T), (H, H)\}, \{(H, T), (T, H)\}, \Omega\}.$$

The *uniform* probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is given by setting

$$\mathbb{P}(\{(T, T), (H, H)\}) := \frac{1}{2} \quad \text{and} \quad \mathbb{P}(\{(H, T), (T, H)\}) := \frac{1}{2}.$$

In addition, we have the following convergence properties.

1. Let $(A_n)_{n\in\mathbb{N}}$ be a *nondecreasing* sequence of events, i.e. $A_n \subset A_{n+1}$, $n \in \mathbb{N}$. Then we have
$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n). \tag{1.2.3}$$

2. Let $(A_n)_{n\in\mathbb{N}}$ be a *nonincreasing* sequence of events, i.e. $A_{n+1} \subset A_n$, $n \in \mathbb{N}$. Then we have
$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n). \tag{1.2.4}$$

1.3 Conditional Probabilities and Independence

We start with an example.

Consider a population $\Omega = M \cup W$ made of a set M of men and a set W of women. Here the σ-algebra $\mathcal{F} = \{\Omega, \emptyset, W, M\}$ corresponds to the information given by gender. After polling the population, e.g. for a market survey, it turns out that a proportion $p \in [0, 1]$ of the population declares to like apples, while a proportion $1 - p$ declares to dislike apples. Let $A \subset \Omega$ denote the subset of individuals who like apples, while $A^c \subset \Omega$ denotes the subset individuals who dislike apples, with

$$p = \mathbb{P}(A) \quad \text{and} \quad 1 - p = \mathbb{P}(A^c),$$

e.g. $p = 60\%$ of the population likes apples. It may be interesting to get a more precise information and to determine

- the relative proportion $\dfrac{\mathbb{P}(A \cap W)}{\mathbb{P}(W)}$ of women who like apples, and
- the relative proportion $\dfrac{\mathbb{P}(A \cap M)}{\mathbb{P}(M)}$ of men who like apples.

Here, $\mathbb{P}(A \cap W)/\mathbb{P}(W)$ represents the probability that a randomly chosen woman in W likes apples, and $\mathbb{P}(A \cap M)/\mathbb{P}(M)$ represents the probability that a randomly chosen man in M likes apples. Those two ratios are interpreted as *conditional probabilities*, for example $\mathbb{P}(A \cap M)/\mathbb{P}(M)$ denotes the probability that an individual likes apples *given that* he is a man.

For another example, suppose that the population Ω is split as $\Omega = Y \cup O$ into a set Y of "young" people and another set O of "old" people, and denote by $A \subset \Omega$ the set of people who voted for candidate A in an election. Here it can be of interest to find out the relative proportion

$$\mathbb{P}(A \mid Y) = \frac{\mathbb{P}(Y \cap A)}{\mathbb{P}(Y)}$$

of young people who voted for candidate A.

More generally, given any two events $A, B \subset \Omega$ with $\mathbb{P}(B) \neq 0$, we call

$$\mathbb{P}(A \mid B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

the probability of A *given* B, or *conditionally to* B.

Remark 1.1 We note that if $\mathbb{P}(B) = 1$ we have $\mathbb{P}(A \cap B^c) \leq \mathbb{P}(B^c) = 0$, hence $\mathbb{P}(A \cap B^c) = 0$, which implies

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) = \mathbb{P}(A \cap B),$$

and $\mathbb{P}(A \mid B) = \mathbb{P}(A)$.

We also recall the following property:

$$\begin{aligned}\mathbb{P}\left(B \cap \bigcup_{n=1}^{\infty} A_n\right) &= \sum_{n=1}^{\infty} \mathbb{P}(B \cap A_n) \\ &= \sum_{n=1}^{\infty} \mathbb{P}(B \mid A_n)\mathbb{P}(A_n) \\ &= \sum_{n=1}^{\infty} \mathbb{P}(A_n \mid B)\mathbb{P}(B),\end{aligned}$$

for any family of disjoint events $(A_n)_{n\geq 1}$ with $A_i \cap A_j = \emptyset$, $i \neq j$, and $\mathbb{P}(B) > 0$, $n \geq 1$. This also shows that conditional probability measures are probability measures, in the sense that whenever $\mathbb{P}(B) > 0$ we have

(a) $\mathbb{P}(\Omega \mid B) = 1$, and
(b) $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n \middle| B\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n \mid B)$, whenever $A_k \cap A_l = \emptyset$, $k \neq l$.

In particular, if $\bigcup_{n=1}^{\infty} A_n = \Omega$, $(A_n)_{n\geq 1}$ becomes a *partition* of Ω and we get the *law of total probability*

$$\mathbb{P}(B) = \sum_{n=1}^{\infty} \mathbb{P}(B \cap A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n \mid B)\mathbb{P}(B) = \sum_{n=1}^{\infty} \mathbb{P}(B \mid A_n)\mathbb{P}(A_n), \quad (1.3.1)$$

provided that $A_i \cap A_j = \emptyset$, $i \neq j$, and $\mathbb{P}(B) > 0$, $n \geq 1$. However, we have in general

$$\mathbb{P}\left(A \middle| \bigcup_{n=1}^{\infty} B_n\right) \neq \sum_{n=1}^{\infty} \mathbb{P}(A \mid B_n),$$

even when $B_k \cap B_l = \emptyset$, $k \neq l$. Indeed, taking for example $A = \Omega = B_1 \cup B_2$ with $B_1 \cap B_2 = \emptyset$ and $\mathbb{P}(B_1) = \mathbb{P}(B_2) = 1/2$, we have

$$1 = \mathbb{P}(\Omega \mid B_1 \cup B_2) \neq \mathbb{P}(\Omega \mid B_1) + \mathbb{P}(\Omega \mid B_2) = 2.$$

Independent Events

Two events A and B are said to be *independent* if

$$\mathbb{P}(A \mid B) = \mathbb{P}(A),$$

which is equivalent to

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

In this case we find

$$\mathbb{P}(A \mid B) = \mathbb{P}(A).$$

1.4 Random Variables

A real-valued random variable is a mapping

$$\begin{aligned} X : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longmapsto X(\omega) \end{aligned}$$

from a probability space Ω into the state space $\mathbb{R}$. Given

$$X : \Omega \longrightarrow \mathbb{R}$$

a random variable and A a (measurable)[1] subset of $\mathbb{R}$, we denote by $\{X \in A\}$ the event

$$\{X \in A\} := \{\omega \in \Omega \ : \ X(\omega) \in A\}.$$

Given $\mathcal{G}$ a σ-algebra on Ω, the mapping $X : \Omega \longrightarrow \mathbb{R}$ is said to be $\mathcal{G}$-measurable if

$$\{X \leq x\} := \{\omega \in \Omega \ : \ X(\omega) \leq x\} \in \mathcal{G},$$

for all $x \in \mathbb{R}$. In this case we will also say that the knowledge of X depends only on the information contained in $\mathcal{G}$.

Examples

(i) Let $\Omega := \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$, and consider the mapping

$$\begin{aligned} X : \Omega &\longrightarrow \mathbb{R} \\ (k, l) &\longmapsto k + l. \end{aligned}$$

Then X is a random variable giving the sum of the two numbers appearing on each die.

(ii) the time needed everyday to travel from home to work or school is a random variable, as the precise value of this time may change from day to day under unexpected circumstances.

(iii) the price of a risky asset is a random variable.

In the sequel we will often use the notion of *indicator function* $\mathbb{1}_A$ of an event A. The indicator function $\mathbb{1}_A$ is the random variable

[1] Measurability of subsets of $\mathbb{R}$ refers to *Borel measurability*, a concept which will not be defined in this text.

$$\begin{aligned}\mathbb{1}_A : \Omega &\longrightarrow \{0, 1\}\\ \omega &\longmapsto \mathbb{1}_A(\omega)\end{aligned}$$

defined by

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \notin A, \end{cases}$$

with the property

$$\mathbb{1}_{A\cap B}(\omega) = \mathbb{1}_A(\omega)\mathbb{1}_B(\omega), \tag{1.4.1}$$

since

$$\begin{aligned}\omega \in A \cap B &\iff \{\omega \in A \text{ and } \omega \in B\}\\ &\iff \{\mathbb{1}_A(\omega) = 1 \text{ and } \mathbb{1}_B(\omega) = 1\}\\ &\iff \mathbb{1}_A(\omega)\mathbb{1}_B(\omega) = 1.\end{aligned}$$

We also have

$$\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A\cap B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A\mathbb{1}_B,$$

and

$$\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B, \tag{1.4.2}$$

if $A \cap B = \emptyset$.

For example, if $\Omega = \mathbb{N}$ and $A = \{k\}$, for all $l \in \mathbb{N}$ we have

$$\mathbb{1}_{\{k\}}(l) = \begin{cases} 1 & \text{if } k = l, \\ \\ 0 & \text{if } k \neq l. \end{cases}$$

Given X a random variable, we also let

$$\mathbb{1}_{\{X=n\}} = \begin{cases} 1 & \text{if } X = n, \\ \\ 0 & \text{if } X \neq n, \end{cases}$$

and

$$\mathbb{1}_{\{X<n\}} = \begin{cases} 1 & \text{if } X < n, \\ \\ 0 & \text{if } X \geq n. \end{cases}$$

1.5 Probability Distributions

The *probability distribution* of a random variable $X : \Omega \longrightarrow \mathbb{R}$ is the collection

$$\{\mathbb{P}(X \in A) \ : \ A \text{ is a measurable subset of } \mathbb{R}\}.$$

As the collection of *measurable* subsets of $\mathbb{R}$ coincides with the σ-algebra generated by the intervals in $\mathbb{R}$, the distribution of X can be reduced to the knowledge of either

$$\{\mathbb{P}(a < X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a) \ : \ a < b \in \mathbb{R}\},$$

or

$$\{\mathbb{P}(X \leq a) \ : \ a \in \mathbb{R}\}, \quad \text{or} \quad \{\mathbb{P}(X \geq a) \ : \ a \in \mathbb{R}\},$$

see e.g. Corollary 3.8 in [Çın11].

Two random variables X and Y are said to be independent under the probability $\mathbb{P}$ if their probability distributions satisfy

$$\mathbb{P}(X \in A \ , \ Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

for all (measurable) subsets A and B of $\mathbb{R}$.

Distributions Admitting a Density

We say that the distribution of X admits a probability *density* distribution function $f_X : \mathbb{R} \longrightarrow \mathbb{R}_+$ if, for all $a \leq b$, the probability $\mathbb{P}(a \leq X \leq b)$ can be written as

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)dx.$$

We also say that the distribution of X is absolutely continuous, or that X is an absolutely continuous random variable. This, however, does *not* imply that the density function $f_X : \mathbb{R} \longrightarrow \mathbb{R}_+$ is continuous.

In particular, we always have

$$\int_{-\infty}^{\infty} f_X(x)dx = \mathbb{P}(-\infty \leq X \leq \infty) = 1$$

for all probability density functions $f_X : \mathbb{R} \longrightarrow \mathbb{R}_+$.

Remark 1.2 Note that if the distribution of X admits a density then for all $a \in \mathbb{R}$, we have

$$\mathbb{P}(X = a) = \int_a^a f(x)dx = 0, \tag{1.5.1}$$

and this is not a contradiction.

In particular, Remark 1.2 shows that

$$\mathbb{P}(a \le X \le b) = \mathbb{P}(X = a) + \mathbb{P}(a < X \le b) = \mathbb{P}(a < X \le b) = \mathbb{P}(a < X < b),$$

for $a \le b$. Property (1.5.1) appears for example in the framework of lottery games with a large number of participants, in which a given number "a" selected in advance has a very low (almost zero) probability to be chosen.

The density f_X can be recovered from the cumulative distribution functions

$$x \longmapsto F_X(x) := \mathbb{P}(X \le x) = \int_{-\infty}^{x} f_X(s)ds,$$

and

$$x \longmapsto 1 - F_X(x) = \mathbb{P}(X \ge x) = \int_{x}^{\infty} f_X(s)ds,$$

as

$$f_X(x) = F_X(s) = \frac{\partial}{\partial x}\int_{-\infty}^{x} f_X(s)ds = -\frac{\partial}{\partial x}\int_{x}^{\infty} f_X(s)ds, \qquad x \in \mathbb{R}.$$

Examples

(i) The *uniform* distribution on an interval.

The probability density function of the uniform distribution on the interval $[a, b]$, $a < b$, is given by

$$f(x) = \frac{1}{b-a}\mathbb{1}_{[a,b]}(x), \qquad x \in \mathbb{R}.$$

(ii) The *Gaussian* distribution.

The probability density function of the standard normal distribution is given by

$$f(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}, \qquad x \in \mathbb{R}.$$

More generally, the probability density function of the Gaussian distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$ is given by

$$f(x) := \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-(x-\mu)^2/(2\sigma^2)}, \qquad x \in \mathbb{R}.$$

In this case, we write $X \simeq \mathcal{N}(\mu, \sigma^2)$.

(iii) The *exponential* distribution.

The probability density function of the exponential distribution with parameter $\lambda > 0$ is given by

$$f(x) := \lambda \mathbb{1}_{[0,\infty)}(x) e^{-\lambda x} = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ \\ 0, & x < 0. \end{cases} \tag{1.5.2}$$

We also have

$$\mathbb{P}(X > t) = e^{-\lambda t}, \qquad t \in \mathbb{R}_+. \tag{1.5.3}$$

(iv) The *gamma* distribution.

The probability density function of the gamma distribution is given by

$$f(x) := \frac{a^\lambda}{\Gamma(\lambda)} \mathbb{1}_{[0,\infty)}(x) x^{\lambda-1} e^{-ax} = \begin{cases} \dfrac{a^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-ax}, & x \geq 0 \\ \\ 0, & x < 0, \end{cases}$$

where $a > 0$ and $\lambda > 0$ are parameters, and

$$\Gamma(\lambda) := \int_0^\infty x^{\lambda-1} e^{-x} dx, \qquad \lambda > 0,$$

is the gamma function.

(v) The *Cauchy* distribution.

The probability density function of the Cauchy distribution is given by

$$f(x) := \frac{1}{\pi(1+x^2)}, \qquad x \in \mathbb{R}.$$

(vi) The *lognormal* distribution.

The probability density function of the lognormal distribution is given by

$$f(x) := \mathbb{1}_{[0,\infty)}(x) \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\mu-\log x)^2}{2\sigma^2}} = \begin{cases} \dfrac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\mu-\log x)^2}{2\sigma^2}}, & x \geq 0 \\ \\ 0, & x < 0. \end{cases}$$

Exercise: For each of the above probability density functions, check that the condition

$$\int_{-\infty}^\infty f(x) dx = 1$$

is satisfied.

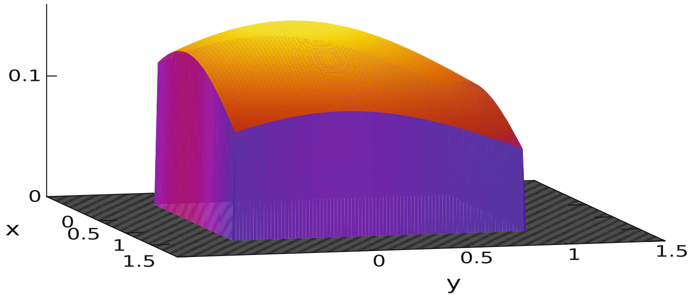


Fig. 1.1 Probability $\mathbb{P}((X, Y) \in [-0.5, 1] \times [-0.5, 1])$ computed as a volume integral

Joint Densities

Given two absolutely continuous random variables $X : \Omega \longrightarrow \mathbb{R}$ and $Y : \Omega \longrightarrow \mathbb{R}$ we can form the $\mathbb{R}^2$-valued random variable (X, Y) defined by

$$\begin{aligned}(X, Y) :\Omega &\longrightarrow \mathbb{R}^2\\ \omega &\longmapsto (X(\omega), Y(\omega)).\end{aligned}$$

We say that (X, Y) admits a joint probability density

$$f_{(X,Y)} : \mathbb{R}^2 \longrightarrow \mathbb{R}_+$$

when

$$\mathbb{P}((X, Y) \in A \times B) = \int_B \int_A f_{(X,Y)}(x, y)dxdy$$

for all *measurable* subsets A, B of $\mathbb{R}$, cf. Fig. 1.1.

The density $f_{(X,Y)}$ can be recovered from the joint cumulative distribution function

$$(x, y) \longmapsto F_{(X,Y)}(x, y) := \mathbb{P}(X \leq x \text{ and } Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,Y)}(s, t)dsdt,$$

and

$$(x, y) \longmapsto \mathbb{P}(X \geq x \text{ and } Y \geq y) = \int_x^{\infty} \int_y^{\infty} f_{(X,Y)}(s, t)dsdt,$$

as

$$\begin{aligned} f_{(X,Y)}(x, y) &= \frac{\partial^2}{\partial x \partial y} F_{(X,Y)}(x, y) \qquad (1.5.4)\\ &= \frac{\partial^2}{\partial x \partial y} \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,Y)}(s, t)dsdt \qquad (1.5.5)\end{aligned}$$

$$= \frac{\partial^2}{\partial x \partial y} \int_x^\infty \int_y^\infty f_{(X,Y)}(s,t) ds dt,$$

$x, y \in \mathbb{R}$.

The probability densities $f_X : \mathbb{R} \longrightarrow \mathbb{R}_+$ and $f_Y : \mathbb{R} \longrightarrow \mathbb{R}_+$ of $X : \Omega \longrightarrow \mathbb{R}$ and $Y : \Omega \longrightarrow \mathbb{R}$ are called the *marginal densities* of (X, Y) and are given by

$$f_X(x) = \int_{-\infty}^\infty f_{(X,Y)}(x,y) dy, \qquad x \in \mathbb{R}, \tag{1.5.6}$$

and

$$f_Y(y) = \int_{-\infty}^\infty f_{(X,Y)}(x,y) dx, \qquad y \in \mathbb{R}.$$

The conditional density $f_{X|Y=y} : \mathbb{R} \longrightarrow \mathbb{R}_+$ of X given $Y = y$ is defined by

$$f_{X|Y=y}(x) := \frac{f_{(X,Y)}(x,y)}{f_Y(y)}, \qquad x, y \in \mathbb{R}, \tag{1.5.7}$$

provided that $f_Y(y) > 0$. In particular, X and Y are independent if and only if $f_{X|Y=y}(x) = f_X(x)$, $x, y \in \mathbb{R}$, i.e.,

$$f_{(X,Y)}(x,y) = f_X(x) f_Y(y), \qquad x, y \in \mathbb{R}.$$

Example

If $X_1, \ldots, X_n$ are independent exponentially distributed random variables with parameters $\lambda_1, \ldots, \lambda_n$ we have

$$\begin{aligned} \mathbb{P}(\min(X_1, \ldots, X_n) > t) &= \mathbb{P}(X_1 > t, \ldots, X_n > t) \\ &= \mathbb{P}(X_1 > t) \cdots \mathbb{P}(X_n > t) \\ &= \mathrm{e}^{-t(\lambda_1 + \cdots + \lambda_n)}, \qquad t \in \mathbb{R}_+, \end{aligned} \tag{1.5.8}$$

hence $\min(X_1, \ldots, X_n)$ is an exponentially distributed random variable with parameter $\lambda_1 + \cdots + \lambda_n$.

Given the joint density of (X_1, X_2) given by

$$f_{(X_1,X_2)}(x,y) = f_{X_1}(x) f_{X_2}(y) = \lambda_1 \lambda_2 \mathrm{e}^{-\lambda_1 x - \lambda_2 y}, \qquad x, y \geq 0,$$

we can write

$$\begin{aligned}
\mathbb{P}(X_1 < X_2) &= \mathbb{P}(X_1 \le X_2) \\
&= \int_0^\infty \int_0^y f_{(X_1,X_2)}(x, y) dx dy \\
&= \lambda_1\lambda_2 \int_0^\infty \int_0^y \mathrm{e}^{-\lambda_1 x - \lambda_2 y} dx dy \\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2},
\end{aligned} \tag{1.5.9}$$

and we note that

$$\mathbb{P}(X_1 = X_2) = \lambda_1\lambda_2 \int_{\{(x,y)\in\mathbb{R}_+^2 \ : \ x=y\}} \mathrm{e}^{-\lambda_1 x - \lambda_2 y} dx dy = 0.$$

Discrete Distributions

We only consider integer-valued random variables, i.e. the distribution of X is given by the values of $\mathbb{P}(X = k)$, $k \in \mathbb{N}$.

Examples

(i) The *Bernoulli* distribution.

We have

$$\mathbb{P}(X = 1) = p \qquad \text{and} \qquad \mathbb{P}(X = 0) = 1 - p, \tag{1.5.10}$$

where $p \in [0, 1]$ is a parameter.

Note that any Bernoulli random variable $X : \Omega \longrightarrow \{0, 1\}$ can be written as the indicator function

$$X = \mathbb{1}_A$$

on Ω with $A = \{X = 1\} = \{\omega \in \Omega \ : \ X(\omega) = 1\}$.

(ii) The *binomial* distribution.

We have

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \qquad k = 0, 1, \ldots, n,$$

where $n \ge 1$ and $p \in [0, 1]$ are parameters.

(iii) The *geometric* distribution.

In this case, we have

$$\mathbb{P}(X = k) = (1-p)p^k, \qquad k \in \mathbb{N}, \tag{1.5.11}$$

where $p \in (0,1)$ is a parameter. For example, if $(X_k)_{k\in\mathbb{N}}$ is a sequence of independent Bernoulli random variables with distribution (1.5.10), then the random variable,[2]

$$T_0 := \inf\{k \in \mathbb{N} \ : \ X_k = 0\}$$

can denote the duration of a game until the time that the wealth X_k of a player reaches 0. The random variable T_0 has the geometric distribution (1.5.11) with parameter $p \in (0,1)$.

(iv) The *negative binomial* (or *Pascal*) distribution.

We have

$$\mathbb{P}(X = k) = \binom{k+r-1}{r-1}(1-p)^r p^k, \qquad k \in \mathbb{N}, \tag{1.5.12}$$

where $p \in (0,1)$ and $r \geq 1$ are parameters. Note that the sum of $r \geq 1$ independent geometric random variables with parameter p has a negative binomial distribution with parameter (r, p). In particular, the negative binomial distribution recovers the geometric distribution when $r = 1$.

(v) The *Poisson* distribution.

We have

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}, \qquad k \in \mathbb{N},$$

where $\lambda > 0$ is a parameter.

The probability that a discrete nonnegative random variable $X : \Omega \longrightarrow \mathbb{N} \cup \{+\infty\}$ is finite is given by

$$\mathbb{P}(X < \infty) = \sum_{k=0}^{\infty} \mathbb{P}(X = k), \tag{1.5.13}$$

and we have

$$1 = \mathbb{P}(X = \infty) + \mathbb{P}(X < \infty) = \mathbb{P}(X = \infty) + \sum_{k=0}^{\infty} \mathbb{P}(X = k).$$

Remark 1.3 The distribution of a discrete random variable cannot admit a density. If this were the case, by Remark 1.2 we would have $\mathbb{P}(X = k) = 0$ for all $k \in \mathbb{N}$ and

[2] The notation "inf" stands for "infimum", meaning the smallest $n \geq 0$ such that $X_n = 0$, if such an n exists.

$$1 = \mathbb{P}(X \in \mathbb{R}) = \mathbb{P}(X \in \mathbb{N}) = \sum_{k=0}^{\infty} \mathbb{P}(X = k) = 0,$$

which is a contradiction.

Given two discrete random variables X and Y, the conditional distribution of X given $Y = k$ is given by

$$\mathbb{P}(X = n \mid Y = k) = \frac{\mathbb{P}(X = n \text{ and } Y = k)}{\mathbb{P}(Y = k)}, \qquad n \in \mathbb{N},$$

provided that $\mathbb{P}(Y = k) > 0$, $k \in \mathbb{N}$.

1.6 Expectation of Random Variables

The *expectation*, or *expected value*, of a random variable X is the mean, or average value, of X. In practice, expectations can be even more useful than probabilities. For example, knowing that a given equipment (such as a bridge) has a failure probability of 1.78493 out of a billion can be of less practical use than knowing the expected lifetime (e.g. 200000 years) of that equipment.

For example, the time $T(\omega)$ to travel from home to work/school can be a random variable with a new outcome and value every day, however we usually refer to its expectation $\mathbb{E}[T]$ rather than to its sample values that may change from day to day.

Expected Value of a Bernoulli Random Variable

Any Bernoulli random variable $X : \Omega \longrightarrow \{0, 1\}$ can be written as the indicator function $X := \mathbb{1}_A$ where A is the event $A = \{X = 1\}$, and the parameter $p \in [0, 1]$ of X is given by

$$p = \mathbb{P}(X = 1) = \mathbb{P}(A) = \mathbb{E}[\mathbb{1}_A] = \mathbb{E}[X].$$

The expectation of a Bernoulli random variable with parameter p is defined as

$$\boxed{\mathbb{E}[\mathbb{1}_A] := 1 \times \mathbb{P}(A) + 0 \times \mathbb{P}(A^c) = \mathbb{P}(A).} \qquad (1.6.1)$$

Expected Value of a Discrete Random Variable

Next, let $X : \Omega \longrightarrow \mathbb{N}$ be a discrete random variable. The expectation $\mathbb{E}[X]$ of X is defined as the sum

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k\mathbb{P}(X = k), \qquad (1.6.2)$$

in which the possible values $k \in \mathbb{N}$ of X are weighted by their probabilities. More generally we have

$$\mathbb{E}[\phi(X)] = \sum_{k=0}^{\infty} \phi(k)\mathbb{P}(X = k),$$

for all sufficiently summable functions $\phi : \mathbb{N} \longrightarrow \mathbb{R}$.

The expectation of the indicator function $X = \mathbb{1}_A = \mathbb{1}_{\{X=1\}}$ can be recovered from (1.6.2) as

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{1}_A] = 0 \times \mathbb{P}(\Omega \setminus A) + 1 \times \mathbb{P}(A) = 0 \times \mathbb{P}(\Omega \setminus A) + 1 \times \mathbb{P}(A) = \mathbb{P}(A).$$

Note that the expectation is a *linear* operation, i.e. we have

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y], \qquad a, b \in \mathbb{R}, \tag{1.6.3}$$

provided that

$$\mathbb{E}[|X|] + \mathbb{E}[|Y|] < \infty.$$

Examples

(i) Expected value of a Poisson random variable with parameter $\lambda > 0$:

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k\mathbb{P}(X = k) = \mathrm{e}^{-\lambda} \sum_{k=1}^{\infty} k\frac{\lambda^k}{k!} = \lambda \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda, \tag{1.6.4}$$

where we used the exponential series (A.1).

(ii) Estimating the expected value of a Poisson random variable using R:

Taking $\lambda := 2$, we can use the following R code:

```
poisson_samples <- rpois(100000, lambda = 2)
poisson_samples
mean(poisson_samples)
```

Given $X : \Omega \longrightarrow \mathbb{N} \cup \{+\infty\}$ a discrete nonnegative random variable X, we have

$$\mathbb{P}(X < \infty) = \sum_{k=0}^{\infty} \mathbb{P}(X = k),$$

and

$$1 = \mathbb{P}(X = \infty) + \mathbb{P}(X < \infty) = \mathbb{P}(X = \infty) + \sum_{k=0}^{\infty} \mathbb{P}(X = k),$$

and in general

$$\mathbb{E}[X] = +\infty \times \mathbb{P}(X = \infty) + \sum_{k=0}^{\infty} k\mathbb{P}(X = k).$$

In particular, $\mathbb{P}(X = \infty) > 0$ implies $\mathbb{E}[X] = \infty$, and the finiteness $\mathbb{E}[X] < \infty$ condition implies $\mathbb{P}(X < \infty) = 1$, however the converse is *not true*.

Examples

(a) Assume that X has the geometric distribution

$$\mathbb{P}(X = k) := \frac{1}{2^{k+1}}, \qquad k \geq 0, \tag{1.6.5}$$

with parameter $p = 1/2$, and

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} \frac{k}{2^{k+1}} = \frac{1}{4} \sum_{k=1}^{\infty} \frac{k}{2^{k-1}} = \frac{1}{4} \frac{1}{(1 - 1/2)^2} = 1 < \infty,$$

by (A.4). Letting $\phi(X) := 2^X$, we have

$$\mathbb{P}(\phi(X) < \infty) = \mathbb{P}(X < \infty) = \sum_{k=0}^{\infty} \frac{1}{2^{k+1}} = 1,$$

and

$$\mathbb{E}[\phi(X)] = \sum_{k=0}^{\infty} \phi(k)\mathbb{P}(X = k) = \sum_{k=0}^{\infty} \frac{2^k}{2^{k+1}} = \sum_{k=0}^{\infty} \frac{1}{2} = +\infty,$$

hence the expectation $\mathbb{E}[\phi(X)]$ is *infinite* although $\phi(X)$ is *finite* with probability one.[3]

(b) The uniform random variable U on $[0, 1]$ satisfies $\mathbb{E}[U] = 1/2 < \infty$ and

$$\mathbb{P}(1/U < \infty) = \mathbb{P}(U > 0) = \mathbb{P}(U \in (0, 1]) = 1,$$

however we have

$$\mathbb{E}[1/U] = \int_0^1 \frac{dx}{x} = +\infty,$$

and $\mathbb{P}(1/U = +\infty) = \mathbb{P}(U = 0) = 0$.

(c) If the random variable X has an exponential distribution with parameter $\mu > 0$ we have

[3] This is the *St. Petersburg paradox*.

$$\mathbb{E}\big[\mathrm{e}^{\lambda X}\big] = \mu \int_0^\infty \mathrm{e}^{\lambda x}\mathrm{e}^{-\mu x}\,dx = \begin{cases} \dfrac{1}{\mu-\lambda} < \infty & \text{if } \mu > \lambda, \\ \\ +\infty, & \text{if } \mu \le \lambda. \end{cases}$$

Conditional Expectation

The notion of expectation takes its full meaning under conditioning. For example, the expected return of a random asset usually depends on information such as economic data, location, etc. In this case, replacing the expectation by a conditional expectation will provide a better estimate of the expected value.

For instance, *life expectancy* is a natural example of a conditional expectation since it typically depends on location, gender, and other parameters.

The *conditional expectation* of $X : \Omega \longrightarrow \mathbb{N}$ a finite random variable given an event A is defined by

$$\mathbb{E}[X \mid A] = \sum_{k=0}^{\infty} k\mathbb{P}(X = k \mid A) = \sum_{k=0}^{\infty} k\frac{\mathbb{P}(X = k \text{ and } A)}{\mathbb{P}(A)}.$$

Lemma 1.4 *Given an event A such that $\mathbb{P}(A) > 0$, we have*

$$\mathbb{E}[X \mid A] = \frac{1}{\mathbb{P}(A)}\mathbb{E}\left[X\mathbb{1}_A\right]. \tag{1.6.6}$$

Proof The proof is done only for $X : \Omega \longrightarrow \mathbb{N}$ a discrete random variable, however (1.6.6) is valid for general real-valued random variables. By Relation (1.4.1) we have

$$\begin{aligned}
\mathbb{E}[X \mid A] &= \frac{1}{\mathbb{P}(A)}\sum_{k=0}^{\infty} k\mathbb{P}(X = k \mid A) \\
&= \frac{1}{\mathbb{P}(A)}\sum_{k=0}^{\infty} k\mathbb{P}(\{X = k\} \cap A) = \frac{1}{\mathbb{P}(A)}\sum_{k=0}^{\infty} k\mathbb{E}\left[\mathbb{1}_{\{X=k\}\cap A}\right] \quad (1.6.7) \\
&= \frac{1}{\mathbb{P}(A)}\sum_{k=0}^{\infty} k\mathbb{E}\left[\mathbb{1}_{\{X=k\}}\mathbb{1}_A\right] = \frac{1}{\mathbb{P}(A)}\mathbb{E}\left[\mathbb{1}_A \sum_{k=0}^{\infty} k\mathbb{1}_{\{X=k\}}\right] \\
&= \frac{1}{\mathbb{P}(A)}\mathbb{E}\left[\mathbb{1}_A X\right], \quad (1.6.8)
\end{aligned}$$

where we used the relation

$$X = \sum_{k=0}^{\infty} k\mathbb{1}_{\{X=k\}}$$

which holds since X takes only integer values. $\square$

Example

(i) Consider $\Omega = \{1, 3, -1, -2, 5, 7\}$ with the uniform probability measure given by

$$\mathbb{P}(\{k\}) = 1/6, \quad k = 1, 3, -1, -2, 5, 7,$$

and the random variable

$$X : \Omega \longrightarrow \mathbb{Z}$$

given by

$$X(k) = k, \qquad k = 1, 3, -1, -2, 5, 7.$$

Then $\mathbb{E}[X \mid X > 0]$ denotes the expected value of X given

$$\{X > 0\} = \{1, 3, 5, 7\} \subset \Omega,$$

i.e. the mean value of X given that X is strictly positive. This conditional expectation can be computed as

$$\begin{aligned}\mathbb{E}[X \mid X > 0] &= \frac{1+3+5+7}{4} \\ &= \frac{1+3+5+7}{6}\frac{1}{4/6} \\ &= \frac{1}{\mathbb{P}(X > 0)}\mathbb{E}[X\mathbb{1}_{\{X>0\}}],\end{aligned}$$

where $\mathbb{P}(X > 0) = 4/6$ and the truncated expectation $\mathbb{E}[X\mathbb{1}_{\{X>0\}}]$ is given by $\mathbb{E}[X\mathbb{1}_{\{X>0\}}] = (1+3+5+7)/6$.

(ii) Estimating a conditional expectation using R:

```
geo_samples <- rgeom(100000, prob = 1/4)
mean(geo_samples)
mean(geo_samples[geo_samples<10])
```

Taking $p := 3/4$, by (A.4) we have

$$\mathbb{E}[X] = (1-p)\sum_{k=1}^{\infty} kp^k = \frac{p}{1-p} = 3,$$

and

$$\begin{aligned}\mathbb{E}[X \mid X < 10] &= \frac{1}{\mathbb{P}(X < 10)}\mathbb{E}\left[X\mathbb{1}_{\{X<10\}}\right] \\ &= \frac{1}{\mathbb{P}(X < 10)}\sum_{k=0}^{9} k\mathbb{P}(X = k)\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\sum_{k=0}^{9} p^k} \sum_{k=1}^{9} k p^k \\
&= \frac{p(1-p)}{1-p^{10}} \frac{\partial}{\partial p} \sum_{k=0}^{9} p^k \\
&= \frac{p(1-p)}{1-p^{10}} \frac{\partial}{\partial p} \left(\frac{1-p^{10}}{1-p} \right) \\
&= \frac{p(1-p^{10}-10(1-p)p^9)}{(1-p)(1-p^{10})} \\
&\simeq 2.4032603455.
\end{aligned}
$$

If the random variable $X : \Omega \longrightarrow \mathbb{N}$ is independent of the event A[4] we have

$$
\mathbb{E}[X \mathbb{1}_A] = \mathbb{E}[X]\mathbb{E}[\mathbb{1}_A] = \mathbb{E}[X]\mathbb{P}(A),
$$

and we naturally find

$$
\mathbb{E}[X \mid A] = \mathbb{E}[X]. \tag{1.6.9}
$$

Taking $X = \mathbb{1}_A$ with

$$
\begin{aligned}
\mathbb{1}_A : \Omega &\longrightarrow \{0, 1\} \\
\omega &\longmapsto \mathbb{1}_A := \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \notin A, \end{cases}
\end{aligned}
$$

shows that, in particular,

$$
\begin{aligned}
\mathbb{E}[\mathbb{1}_A \mid A] &= 0 \times \mathbb{P}(X = 0 \mid A) + 1 \times \mathbb{P}(X = 1 \mid A) \\
&= \mathbb{P}(X = 1 \mid A) \\
&= \mathbb{P}(A \mid A) \\
&= 1.
\end{aligned}
$$

One can also define the conditional expectation of X given $A = \{Y = k\}$, as

$$
\mathbb{E}[X \mid Y = k] = \sum_{n=0}^{\infty} n \mathbb{P}(X = n \mid Y = k),
$$

where $Y : \Omega \longrightarrow \mathbb{N}$ is a discrete random variable.

Proposition 1.5 *Given X a discrete random variable such that $\mathbb{E}[|X|] < \infty$, we have the relation*

[4] i.e., $\mathbb{P}(\{X = k\} \cap A) = \mathbb{P}(\{X = k\})\mathbb{P}(A)$ for all $k \in \mathbb{N}$.

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]], \tag{1.6.10}$$

which is sometimes referred to as the tower property.

Proof We have

$$\begin{aligned}
\mathbb{E}[\mathbb{E}[X \mid Y]] &= \sum_{k=0}^{\infty} \mathbb{E}[X \mid Y = k]\mathbb{P}(Y = k) \\
&= \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} n\mathbb{P}(X = n \mid Y = k)\mathbb{P}(Y = k) \\
&= \sum_{n=0}^{\infty} n \sum_{k=0}^{\infty} \mathbb{P}(X = n \text{ and } Y = k) \\
&= \sum_{n=0}^{\infty} n\mathbb{P}(X = n) = \mathbb{E}[X],
\end{aligned}$$

where we used the marginal distribution

$$\mathbb{P}(X = n) = \sum_{k=0}^{\infty} \mathbb{P}(X = n \text{ and } Y = k), \qquad n \in \mathbb{N},$$

that follows from the *law of total probability* (1.3.1) with $A_k = \{Y = k\}$, $k \geq 0$.

Taking

$$Y = \sum_{k=0}^{\infty} k \mathbb{1}_{A_k},$$

with $A_k := \{Y = k\}$, $k \in \mathbb{N}$, from (1.6.10) we also get the *law of total expectation*

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E}[\mathbb{E}[X \mid Y]] \qquad (1.6.11) \\
&= \sum_{k=0}^{\infty} \mathbb{E}[X \mid Y = k]\mathbb{P}(Y = k) \\
&= \sum_{k=0}^{\infty} \mathbb{E}[X \mid A_k]\mathbb{P}(A_k).
\end{aligned}$$

Example

Life expectancy in Singapore is $\mathbb{E}[T] = 80$ years overall, where T denotes the lifetime of a given individual chosen at random. Let $G \in \{m, w\}$ denote the gender of that individual. The statistics show that

$$\mathbb{E}[T \mid G = m] = 78 \quad \text{and} \quad \mathbb{E}[T \mid G = w] = 81.9,$$

and we have

$$\begin{aligned}
80 &= \mathbb{E}[T] \\
&= \mathbb{E}[\mathbb{E}[T|G]] \\
&= \mathbb{P}(G=w)\mathbb{E}[T \mid G=w] + \mathbb{P}(G=m)\mathbb{E}[T \mid G=m] \\
&= 81.9 \times \mathbb{P}(G=w) + 78 \times \mathbb{P}(G=m) \\
&= 81.9 \times (1-\mathbb{P}(G=m)) + 78 \times \mathbb{P}(G=m),
\end{aligned}$$

showing that

$$80 = 81.9 \times (1-\mathbb{P}(G=m)) + 78 \times \mathbb{P}(G=m),$$

i.e.

$$\mathbb{P}(G=m) = \frac{81.9-80}{81.9-78} = \frac{1.9}{3.9} = 0.487.$$

Variance

The *variance* of a random variable X is defined in general by

$$\mathrm{Var}[X] := \mathbb{E}[X^2] - (\mathbb{E}[X])^2,$$

provided that $\mathbb{E}[|X|^2] < \infty$. If $(X_k)_{k\in\mathbb{N}}$ is a sequence of independent random variables we have

$$\begin{aligned}
&\mathrm{Var}\left[\sum_{k=1}^n X_k\right] = \mathbb{E}\left[\left(\sum_{k=1}^n X_k\right)^2\right] - \left(\mathbb{E}\left[\sum_{k=1}^n X_k\right]\right)^2 \\
&= \mathbb{E}\left[\sum_{k=1}^n X_k \sum_{l=1}^n X_l\right] - \mathbb{E}\left[\sum_{k=1}^n X_k\right]\mathbb{E}\left[\sum_{l=1}^n X_l\right] \\
&= \mathbb{E}\left[\sum_{k=1}^n \sum_{l=1}^n X_k X_l\right] - \sum_{k=1}^n \sum_{l=1}^n \mathbb{E}[X_k]\mathbb{E}[X_l] \\
&= \sum_{k=1}^n \mathbb{E}[X_k^2] + \sum_{1\le k\ne l\le n} \mathbb{E}[X_k X_l] - \sum_{k=1}^n (\mathbb{E}[X_k])^2 - \sum_{1\le k\ne l\le n} \mathbb{E}[X_k]\mathbb{E}[X_l] \\
&= \sum_{k=1}^n (\mathbb{E}[X_k^2] - (\mathbb{E}[X_k])^2) \\
&= \sum_{k=1}^n \mathrm{Var}[X_k]. \qquad (1.6.12)
\end{aligned}$$

Random Sums

In the sequel we consider $Y : \Omega \longrightarrow \mathbb{N}$ an *a.s.* finite, integer-valued random variable, i.e. we have $\mathbb{P}(Y < \infty) = 1$ and $\mathbb{P}(Y = \infty) = 0$.

Based on the tower property or ordinary conditioning, the expectation of a random sum $\sum_{k=1}^{Y} X_k$, where $(X_k)_{k\in\mathbb{N}}$ is a sequence of random variables, can be computed from the *tower property* (1.6.10) or from the *law of total expectation* (1.6.11) as

$$\begin{aligned}\mathbb{E}\left[\sum_{k=1}^{Y} X_k\right] &= \mathbb{E}\left[\mathbb{E}\left[\sum_{k=1}^{Y} X_k \Big| Y\right]\right] \\ &= \sum_{n=0}^{\infty} \mathbb{E}\left[\sum_{k=1}^{Y} X_k \Big| Y = n\right] \mathbb{P}(Y = n) \\ &= \sum_{n=0}^{\infty} \mathbb{E}\left[\sum_{k=1}^{n} X_k \Big| Y = n\right] \mathbb{P}(Y = n),\end{aligned}$$

and if Y is (mutually) independent of the sequence $(X_k)_{k\in\mathbb{N}}$ this yields

$$\begin{aligned}\mathbb{E}\left[\sum_{k=1}^{Y} X_k\right] &= \sum_{n=0}^{\infty} \mathbb{E}\left[\sum_{k=1}^{n} X_k\right] \mathbb{P}(Y = n) \\ &= \sum_{n=0}^{\infty} \mathbb{P}(Y = n) \sum_{k=1}^{n} \mathbb{E}[X_k].\end{aligned}$$

Similarly, for a random product we will have, using the independence of Y with $(X_k)_{k\in\mathbb{N}}$,

$$\begin{aligned}\mathbb{E}\left[\prod_{k=1}^{Y} X_k\right] &= \sum_{n=0}^{\infty} \mathbb{E}\left[\prod_{k=1}^{n} X_k\right] \mathbb{P}(Y = n) \qquad (1.6.13) \\ &= \sum_{n=0}^{\infty} \mathbb{P}(Y = n) \prod_{k=1}^{n} \mathbb{E}[X_k],\end{aligned}$$

where the last equality requires the (mutual) independence of the random variables in the sequence $(X_k)_{k\geq 1}$.

Distributions Admitting a Density

Given a random variable X whose distribution admits a density $f_X : \mathbb{R} \longrightarrow \mathbb{R}_+$ we have

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx,$$

and more generally,

$$\mathbb{E}[\phi(X)] = \int_{-\infty}^{\infty} \phi(x) f_X(x) dx, \tag{1.6.14}$$

for all sufficiently integrable function ϕ on $\mathbb{R}$. For example, if X has a standard normal distribution we have

$$\mathbb{E}[\phi(X)] = \int_{-\infty}^{\infty} \phi(x) \mathrm{e}^{-x^2/2} \frac{dx}{\sqrt{2\pi}}.$$

In case X has a Gaussian distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$ we get

$$\mathbb{E}[\phi(X)] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \phi(x) \mathrm{e}^{-(x-\mu)^2/(2\sigma^2)} dx. \tag{1.6.15}$$

Exercise: In case $X \simeq \mathcal{N}(\mu, \sigma^2)$ has a Gaussian distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$, check that

$$\mu = \mathbb{E}[X] \quad \text{and} \quad \sigma^2 = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

When $(X, Y) : \Omega \longrightarrow \mathbb{R}^2$ is a $\mathbb{R}^2$-valued couple of random variables whose distribution admits a density $f_{X,Y} : \mathbb{R}^2 \longrightarrow \mathbb{R}_+$ we have

$$\mathbb{E}[\phi(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(x, y) f_{X,Y}(x, y) dx dy,$$

for all sufficiently integrable function ϕ on $\mathbb{R}^2$.
The expectation of an absolutely continuous random variable satisfies the same linearity property (1.6.3) as in the discrete case.
The conditional expectation of an absolutely continuous random variable can be defined as

$$\mathbb{E}[X \mid Y = y] = \int_{-\infty}^{\infty} x f_{X|Y=y}(x) dx$$

where the conditional density $f_{X|Y=y}(x)$ is defined in (1.5.7), with the relation

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]] \tag{1.6.16}$$

which is called the *tower property* and holds as in the discrete case, since

$$\begin{aligned}\mathbb{E}[\mathbb{E}[X \mid Y]] &= \int_{-\infty}^{\infty} \mathbb{E}[X \mid Y = y] f_Y(y) dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X|Y=y}(x) f_Y(y) dx dy\end{aligned}$$

$$= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{(X,Y)}(x, y) dy dx$$
$$= \int_{-\infty}^{\infty} x f_X(x) dx = \mathbb{E}[X],$$

where we used Relation (1.5.6) between the density of (X, Y) and its marginal X.

For example, an exponentially distributed random variable X with probability density function (1.5.2) has the expected value

$$\mathbb{E}[X] = \lambda \int_0^{\infty} x \mathrm{e}^{-\lambda x} dx = \frac{1}{\lambda}.$$

Conditional Expectation Revisited

The construction of conditional expectation given above for discrete and absolutely continuous random variables can be generalized to σ-algebras.

Definition 1.6 Given $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space and $p \geq 1$, we let $L^p(\Omega, \mathcal{F})$ denote the space of $\mathcal{F}$-measurable and p-integrable random variables, i.e.

$$L^p(\Omega, \mathcal{F}) := \big\{ F : \Omega \longrightarrow \mathbb{R} \ : \ \mathbb{E}[|F|^p] < \infty \big\}.$$

We define a *scalar product* $\langle \cdot, \cdot \rangle_{L^2(\Omega,\mathcal{F})}$ between elements of $L^2(\Omega, \mathcal{F})$, as

$$\langle F, G \rangle_{L^2(\Omega,\mathcal{F})} := \mathbb{E}[FG], \qquad F, G \in L^2(\Omega, \mathcal{F}).$$

This scalar product is associated to the norm $\| \cdot \|_{L^2(\Omega,\mathcal{F})}$ by the relation

$$\|F\| = \sqrt{\mathbb{E}[F^2]} = \sqrt{\langle F, F \rangle_{L^2(\Omega,\mathcal{F})}}, \qquad F \in L^2(\Omega, \mathcal{F}),$$

and it induces a notion of *orthogonality*, namely F is *orthogonal* to G in $L^2(\Omega, \mathcal{F})$ if and only if $\langle F, G \rangle_{L^2(\Omega,\mathcal{F})} = 0$.

Definition 1.7 Given $\mathcal{G} \subset \mathcal{F}$ a sub σ-algebra of $\mathcal{F}$ and $F \in L^2(\Omega, \mathcal{F})$, the conditional expectation of F given $\mathcal{G}$, and denoted

$$\mathbb{E}[F \mid \mathcal{G}],$$

is defined as the *orthogonal projection* of F onto $L^2(\Omega, \mathcal{G})$.

As a consequence of the uniqueness of the orthogonal projection onto the subspace $L^2(\Omega, \mathcal{G})$ of $L^2(\Omega, \mathcal{F})$, $\mathbb{E}[F \mid \mathcal{G}]$ is characterized by the relation

$$\langle G, F - \mathbb{E}[F \mid \mathcal{G}] \rangle_{L^2(\Omega,\mathcal{F})} = 0,$$

which rewrites as

$$\mathbb{E}[G(F - \mathbb{E}[F \mid \mathcal{G}])] = 0,$$

i.e.

$$\mathbb{E}[GF] = \mathbb{E}[G\mathbb{E}[F \mid \mathcal{G}]],$$

for all bounded and $\mathcal{G}$-measurable random variables G, where $\langle \cdot, \cdot \rangle_{L^2(\Omega,\mathcal{F})}$ denotes the inner product in $L^2(\Omega, \mathcal{F})$.
In addition, $\mathbb{E}[F \mid \mathcal{G}]$ realizes the minimum in mean square distance between F and $L^2(\Omega, \mathcal{G})$, i.e. we have

$$\|F - \mathbb{E}[F \mid \mathcal{G}]\|_{L^2(\Omega,\mathcal{F})} = \inf_{G \in L^2(\Omega,\mathcal{G})} \|F - G\|_{L^2(\Omega,\mathcal{F})}. \tag{1.6.17}$$

The following proposition will often be used as a characterization of $\mathbb{E}[F \mid \mathcal{G}]$.

Proposition 1.8 *Given $F \in L^2(\Omega, \mathcal{F})$, $X := \mathbb{E}[F \mid \mathcal{G}]$ is the unique random variable X in $L^2(\Omega, \mathcal{G})$ that satisfies the relation*

$$\mathbb{E}[GF] = \mathbb{E}[GX] \tag{1.6.18}$$

for all bounded and $\mathcal{G}$-measurable random variables G.

The conditional expectation operator has the following properties.

(i) $\mathbb{E}[FG \mid \mathcal{G}] = G\mathbb{E}[F \mid \mathcal{G}]$ if G depends only on the information contained in $\mathcal{G}$.

Proof: By the characterization (1.6.18) it suffices to show that

$$\mathbb{E}[HFG] = \mathbb{E}[HG\mathbb{E}[F|\mathcal{G}]], \tag{1.6.19}$$

for all bounded and $\mathcal{H}$-measurable random variables H, which implies $\mathbb{E}[FG \mid \mathcal{G}] = G\mathbb{E}[F \mid \mathcal{G}]$.

Relation (1.6.19) holds from (1.6.18) because the product HG is $\mathcal{G}$-measurable hence G in (1.6.18) can be replaced with HG.

(ii) $\mathbb{E}[G|\mathcal{G}] = G$ when G depends only on the information contained in $\mathcal{G}$.

Proof: This is a consequence of point (i) above by taking $F = 1$.

(iii) $\mathbb{E}[\mathbb{E}[F|\mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[F|\mathcal{H}]$ if $\mathcal{H} \subset \mathcal{G}$, called the *tower property*.

Proof: First we note that (iii) holds when $\mathcal{H} = \{\emptyset, \Omega\}$ because taking $G = 1$ in (1.6.18) yields

$$\mathbb{E}[F] = \mathbb{E}[\mathbb{E}[F \mid \mathcal{G}]]. \tag{1.6.20}$$

Next, by the characterization (1.6.18) it suffices to show that

$$\mathbb{E}[H\mathbb{E}[F|\mathcal{G}]] = \mathbb{E}[H\mathbb{E}[F|\mathcal{H}]], \tag{1.6.21}$$

for all bounded and $\mathcal{G}$-measurable random variables H, which will imply (iii) from (1.6.18).

In order to prove (1.6.21) we check that by point (i) above and (1.6.20) we have

$$\begin{aligned}\mathbb{E}[H\mathbb{E}[F|\mathcal{G}]] &= \mathbb{E}[\mathbb{E}[HF|\mathcal{G}]] = \mathbb{E}[HF]\\ &= \mathbb{E}[\mathbb{E}[HF|\mathcal{H}]] = \mathbb{E}[H\mathbb{E}[F|\mathcal{H}]],\end{aligned}$$

and we conclude by the characterization (1.6.18).

(iv) $\mathbb{E}[F|\mathcal{G}] = \mathbb{E}[F]$ when F "does not depend" on the information contained in $\mathcal{G}$ or, more precisely stated, when the random variable F is *independent* of the σ-algebra $\mathcal{G}$.

Proof: It suffices to note that for all bounded $\mathcal{G}$-measurable G we have

$$\mathbb{E}[FG] = \mathbb{E}[F]\mathbb{E}[G] = \mathbb{E}[G\mathbb{E}[F]],$$

and we conclude again by (1.6.18).

(v) If G depends only on $\mathcal{G}$ and F is independent of $\mathcal{G}$, then

$$\mathbb{E}[h(F,G)|\mathcal{G}] = \mathbb{E}[h(F,x)]_{x=G}. \tag{1.6.22}$$

Proof: This relation can be proved using the tower property, by noting that for any $K \in L^2(\Omega, \mathcal{G})$ we have

$$\begin{aligned}\mathbb{E}[K\mathbb{E}[h(x,F)]_{x=G}] &= \mathbb{E}[K\mathbb{E}[h(x,F) \mid \mathcal{G}]_{x=G}]\\ &= \mathbb{E}[K\mathbb{E}[h(G,F) \mid \mathcal{G}]]\\ &= \mathbb{E}[\mathbb{E}[Kh(G,F) \mid \mathcal{G}]]\\ &= \mathbb{E}[Kh(G,F)],\end{aligned}$$

which yields (1.6.22) by the characterization (1.6.18).

The notion of conditional expectation can be extended from square-integrable random variables in $L^2(\Omega, \mathcal{F})$ to integrable random variables in $L^1(\Omega, \mathcal{F})$, cf. e.g. [Kal02], Theorem 5.1.

When the σ-algebra $\mathcal{G} := \sigma(A_1, A_2, \ldots, A_n)$ is generated by n disjoint events $A_1, A_2, \ldots, A_n \in \mathcal{F}$, we have

$$\mathbb{E}[F \mid \mathcal{G}] = \sum_{k=1}^{n} \mathbb{1}_{A_k}\mathbb{E}[F \mid A_k] = \sum_{k=1}^{n} \mathbb{1}_{A_k}\frac{\mathbb{E}[F\mathbb{1}_{A_k}]}{\mathbb{P}(A_k)}.$$

1.7 Moment and Probability Generating Functions

Characteristic Functions

The *characteristic function* of a random variable X is the function $\Psi_X : \mathbb{R} \longrightarrow \mathbb{C}$ defined by

$$\Psi_X(t) = \mathbb{E}\big[e^{itX}\big], \qquad t \in \mathbb{R}.$$

The characteristic function Ψ_X of a random variable X with density $f : \mathbb{R} \longrightarrow \mathbb{R}_+$ satisfies

$$\Psi_X(t) = \int_{-\infty}^{\infty} e^{ixt} f(x)dx, \qquad t \in \mathbb{R}.$$

On the other hand, if $X : \Omega \longrightarrow \mathbb{N}$ is a discrete random variable we have

$$\Psi_X(t) = \sum_{n=0}^{\infty} e^{itn}\mathbb{P}(X = n), \qquad t \in \mathbb{R}.$$

The main applications of characteristic functions lie in the following theorems:

Theorem 1.9 *Two random variables $X : \Omega \longrightarrow \mathbb{R}$ and $Y : \Omega \longrightarrow \mathbb{R}$ have same distribution if and only if*

$$\Psi_X(t) = \Psi_Y(t), \qquad t \in \mathbb{R}.$$

Theorem 1.9 is used to identify or to determine the probability distribution of a random variable X, by comparison with the characteristic function Ψ_Y of a random variable Y whose distribution is known.

The characteristic function of a random vector (X, Y) is the function $\Psi_{X,Y} : \mathbb{R}^2 \longrightarrow \mathbb{C}$ defined by

$$\Psi_{X,Y}(s, t) = \mathbb{E}\big[e^{isX+itY}\big], \qquad s, t \in \mathbb{R}.$$

Theorem 1.10 *The random variables $X : \Omega \longrightarrow \mathbb{R}$ and $Y : \Omega \longrightarrow \mathbb{R}$ are* independent *if and only if*

$$\Psi_{X,Y}(s, t) = \Psi_X(s)\Psi_Y(t), \qquad s, t \in \mathbb{R}.$$

A random variable X is Gaussian with mean μ and variance σ^2 if and only if its characteristic function satisfies

$$\mathbb{E}\big[e^{i\alpha X}\big] = e^{i\alpha\mu - \alpha^2\sigma^2/2}, \qquad \alpha \in \mathbb{R}. \tag{1.7.1}$$

In terms of moment generating functions we have, replacing $i\alpha$ by α,

$$\mathbb{E}\big[e^{\alpha X}\big] = e^{\alpha\mu + \alpha^2\sigma^2/2}, \qquad \alpha \in \mathbb{R}. \tag{1.7.2}$$

From Theorems 1.9 and 1.10 we deduce the following proposition.

Proposition 1.11 *Let $X \simeq \mathcal{N}(\mu, \sigma_X^2)$ and $Y \simeq \mathcal{N}(\nu, \sigma_Y^2)$ be independent Gaussian random variables. Then $X + Y$ also has a Gaussian distribution*

$$X + Y \simeq \mathcal{N}(\mu + \nu, \sigma_X^2 + \sigma_Y^2).$$

Proof Since X and Y are independent, by Theorem 1.10 the characteristic function Ψ_{X+Y} of $X + Y$ is given by

$$\begin{aligned} \Phi_{X+Y}(t) &= \Phi_X(t)\Phi_Y(t) \\ &= \mathrm{e}^{it\mu - t^2\sigma_X^2/2}\mathrm{e}^{it\nu - t^2\sigma_Y^2/2} \\ &= \mathrm{e}^{it(\mu+\nu) - t^2(\sigma_X^2+\sigma_Y^2)/2}, \qquad t \in \mathbb{R}, \end{aligned}$$

where we used (1.7.1). Consequently, the characteristic function of $X + Y$ is that of a Gaussian random variable with mean $\mu + \nu$ and variance $\sigma_X^2 + \sigma_Y^2$ and we conclude by Theorem 1.9.

Moment Generating Functions

The *moment generating function* of a random variable X is the function $\Phi_X : \mathbb{R} \longrightarrow \mathbb{R}$ defined by

$$\Phi_X(t) = \mathbb{E}\big[\mathrm{e}^{tX}\big], \qquad t \in \mathbb{R},$$

provided that the expectation is finite. In particular, we have

$$\mathbb{E}[X^n] = \frac{\partial^n}{\partial t}\Phi_X(0), \qquad n \geq 1,$$

provided that $\mathbb{E}[|X|^n] < \infty$, and

$$\Phi_X(t) = \mathbb{E}\big[\mathrm{e}^{tX}\big] = \sum_{n=0}^{\infty} \frac{t^n}{n!}\mathbb{E}[X^n],$$

provided that $\mathbb{E}\big[\mathrm{e}^{t|X|}\big] < \infty$, $t \in \mathbb{R}$, and for this reason the moment generating function G_X characterizes the *moments* $\mathbb{E}[X^n]$ of $X : \Omega \longrightarrow \mathbb{N}$, $n \geq 0$.

The moment generating function Φ_X of a random variable X with density $f : \mathbb{R} \longrightarrow \mathbb{R}_+$ satisfies

$$\Phi_X(t) = \int_{-\infty}^{\infty} \mathrm{e}^{xt} f(x)dx, \qquad t \in \mathbb{R}.$$

Note that in probability we are using the *bilateral* moment generating function transform for which the integral is from $-\infty$ to $+\infty$.

Probability Generating Functions

Consider

$$X : \Omega \longrightarrow \mathbb{N} \cup \{+\infty\}$$

a *discrete* random variable possibly taking infinite values. The *probability generating function* of X is the *function*

$$\begin{aligned} G_X : [-1,1] &\longrightarrow \mathbb{R} \\ s &\longmapsto G_X(s) \end{aligned}$$

defined by

$$G_X(s) := \mathbb{E}\left[s^X \mathbb{1}_{\{X<\infty\}}\right] = \sum_{n=0}^{\infty} s^n \mathbb{P}(X=n), \qquad -1 \leq s \leq 1. \tag{1.7.3}$$

Note that the series summation in (1.7.3) is over the *finite* integers, which explains the presence of the truncating indicator $\mathbb{1}_{\{X<\infty\}}$ inside the expectation in (1.7.3). If the random variable $X : \Omega \longrightarrow \mathbb{N}$ is almost surely finite, i.e. $\mathbb{P}(X < \infty) = 1$, we simply have

$$G_X(s) = \mathbb{E}\left[s^X\right] = \sum_{n=0}^{\infty} s^n \mathbb{P}(X=n), \qquad -1 \leq s \leq 1,$$

and for this reason the probability generating function G_X characterizes the *probability distribution* $\mathbb{P}(X=n)$, $n \geq 0$, of $X : \Omega \longrightarrow \mathbb{N}$.

Examples

(i) Poisson distribution. Consider a random variable X with probability generating function

$$G_X(s) = \mathrm{e}^{\lambda(s-1)}, \qquad -1 \leq s \leq 1,$$

for some $\lambda > 0$. What is the distribution of X?

Using the exponential series (A.1) we have

$$G_X(s) = \mathrm{e}^{\lambda(s-1)} = \mathrm{e}^{-\lambda} \sum_{n=0}^{\infty} s^n \frac{\lambda^n}{n!}, \qquad -1 \leq s \leq 1, \tag{1.7.4}$$

hence by identification with (1.7.3) we find

$$\mathbb{P}(X=n) = \mathrm{e}^{-\lambda} \frac{\lambda^n}{n!}, \qquad n \in \mathbb{N},$$

i.e. X has the Poisson distribution with parameter λ.

(ii) Geometric distribution. Given X a random variable with geometric distribution $\mathbb{P}(X = n) = (1-p)p^n$, $n \in \mathbb{N}$, we have

$$G_X(s) = \sum_{n=0}^{\infty} s^n \mathbb{P}(X = n) = (1-p) \sum_{n=0}^{\infty} s^n p^n = \frac{1-p}{1-ps}, \qquad -1 < s < 1,$$

where we applied the geometric series (A.3).

We note that from (1.7.3) we can write

$$G_X(s) = \mathbb{E}[s^X], \qquad -1 < s < 1,$$

since $s^X = s^X \mathbb{1}_{\{X<\infty\}}$ when $-1 < s < 1$.

Properties of Probability Generating Functions

(i) Taking $s = 1$, we have

$$G_X(1) = \sum_{n=0}^{\infty} \mathbb{P}(X = n) = \mathbb{P}(X < \infty) = \mathbb{E}\left[\mathbb{1}_{\{X<\infty\}}\right],$$

hence

$$\boxed{G_X(1) = \mathbb{P}(X < \infty).}$$

(ii) Taking $s = 0$, we have

$$G_X(0) = \mathbb{E}[0^X] = \mathbb{E}[\mathbb{1}_{\{X=0\}}] = \mathbb{P}(X = 0),$$

since $0^0 = 1$ and $0^X = \mathbb{1}_{\{X=0\}}$, hence

$$\boxed{G_X(0) = \mathbb{P}(X = 0). \qquad (1.7.5)}$$

(iii) The *derivative* $G'_X(s)$ of $G_X(s)$ with respect to s satisfies

$$G'_X(s) = \sum_{n=1}^{\infty} n s^{n-1} \mathbb{P}(X = n), \qquad -1 < s < 1,$$

hence, taking $s := 1$ we have

$$\boxed{G'_X(1) = \mathbb{E}[X] = \sum_{k=0}^{\infty} k\mathbb{P}(X = k),}$$

provided that $\mathbb{E}[X] < \infty$.

(iv) By computing the second derivative

$$\begin{aligned} G''_X(s) &= \sum_{k=2}^{\infty} k(k-1)s^{k-2}\mathbb{P}(X = k) \\ &= \sum_{k=0}^{\infty} k(k-1)s^{k-2}\mathbb{P}(X = k) \\ &= \sum_{k=0}^{\infty} k^2 s^{k-2}\mathbb{P}(X = k) - \sum_{k=0}^{\infty} k s^{k-2}\mathbb{P}(X = k), \qquad -1 < s < 1, \end{aligned}$$

we similarly find

$$\begin{aligned} G''_X(1) &= \sum_{k=0}^{\infty} k(k-1)\mathbb{P}(X = k) \\ &= \sum_{k=0}^{\infty} k^2\mathbb{P}(X = k) - \sum_{k=0}^{\infty} k\mathbb{P}(X = k) \\ &= \mathbb{E}[X^2] - \mathbb{E}[X], \end{aligned}$$

hence

$$\boxed{\mathrm{Var}[X] = G''_X(1) + G'_X(1)(1 - G'_X(1)), \qquad (1.7.6)}$$

provided that $\mathbb{E}[X^2] < \infty$.

(v) When $X : \Omega \longrightarrow \mathbb{N}$ and $Y : \Omega \longrightarrow \mathbb{N}$ are two finite independent random variables we have

$$\boxed{G_{X+Y}(s) = \mathbb{E}\left[s^{X+Y}\right] = \mathbb{E}\left[s^X s^Y\right] = \mathbb{E}\left[s^X\right]\mathbb{E}\left[s^Y\right] = G_X(s)G_Y(s), \qquad (1.7.7)}$$

$-1 \leq s \leq 1$.

(vi) The probability generating function can also be used from (1.7.3) to recover the distribution of the discrete random variable X as

$$\boxed{\mathbb{P}(X = n) = \frac{1}{n!} \frac{\partial^n}{\partial s^n} G_X(s)_{|s=0}, \qquad n \in \mathbb{N},} \tag{1.7.8}$$

extending (1.7.5) to all $n \geq 0$.

Exercise: Show that the probability generating function of a Poisson random variable X with parameter $\lambda > 0$ is given by

$$G_X(s) = \mathrm{e}^{\lambda(s-1)}, \qquad -1 \leq s \leq 1.$$

From the generating function we also recover the mean

$$\mathbb{E}[X] = G'_X(1) = \lambda \mathrm{e}^{\lambda(s-1)}_{|s=1} = \lambda,$$

of the Poisson random variable X with parameter λ, and its variance

$$\begin{aligned}
\mathrm{Var}[X] &= G''_X(1) + G'_X(1) - (G'_X(1))^2 \\
&= \lambda^2 \mathrm{e}^{\lambda(s-1)}_{|s=1} + \lambda s \mathrm{e}^{\lambda(s-1)}_{|s=1} - \lambda^2 \\
&= \lambda^2 + \lambda - \lambda^2 = \lambda,
\end{aligned}$$

by (1.7.6).

Exercises

Exercise 1.1 Consider a random variable $X : \Omega \longrightarrow \mathbb{N} \cup \{\infty\}$ with distribution

$$\mathbb{P}(X = k) = qp^k, \qquad k \in \mathbb{N} = \{0, 1, 2, \ldots\},$$

where $q \in [0, 1 - p]$ and $0 \leq p < 1$.

(a) Compute $\mathbb{P}(X < \infty)$ and $\mathbb{P}(X = \infty)$ by considering two cases, and give the value of $\mathbb{E}[X]$ when $0 \leq q < 1 - p$.

(b) Assume that $q = 1 - p$ and consider the random variable $Y := r^X$ for some $r > 0$. Explain why $\mathbb{P}(Y < \infty) = 1$ and compute $\mathbb{E}[Y]$ by considering two cases depending on the value of $r > 0$.

Exercise 1.2 Let $N \in \{1, 2, 3, 4, 5, 6\}$ denote the integer random variable obtained by tossing a six faced die and by noting the number on the upper side of the die. Given the value of N, an *independent*, unbiased coin is thrown N times. We denote by Z the total number of heads that appear in the process of throwing the coin N times.

(a) Using conditioning on the value of $N \in \{1, 2, 3, 4, 5, 6\}$, compute the mean and the variance of the random variable Z.
(b) Determine the probability distribution of Z.
(c) Recover the result of Question (a) from the data of the probability distribution computed in Question (b).

Exercise 1.3 Thinning of Poisson random variables. Given a random sample N of a Poisson random variable with parameter λ, we perform a number N of *independent* $\{0, 1\}$-valued Bernoulli experiments *independent* of N, each of them with parameter $p \in (0, 1)$. We let Z denote the total number of $+1$ outcomes occuring in the N Bernoulli trials.

(a) Express Z as a random sum, and use this expression to compute the mean and variance of Z.
(b) Compute the probability distribution of Z.
(c) Recover the result of Question (a) from the data of the probability distribution computed in Question (b).

Exercise 1.4 Given X and Y two independent exponentially distributed random variables with parameters λ and μ, show the relation

$$\mathbb{E}[\min(X, Y) \mid X < Y] = \frac{1}{\lambda + \mu} = \mathbb{E}[\min(X, Y)]. \tag{1.7.9}$$

Exercise 1.5 Given a random sample L of a gamma random variable with density

$$f_L(x) = \mathbb{1}_{[0,\infty)} x e^{-x},$$

consider U a uniform random variable taking values in the interval $[0, L]$ and let $V = L - U$.
Compute the joint probability density function of the couple (U, V) of random variables.

Exercise 1.6 Let X and Y denote two independent Poisson random variables with parameters λ and μ.

(a) Show that the random variable $X + Y$ has the Poisson distribution with parameter $\lambda + \mu$.
(b) Compute the conditional distribution $\mathbb{P}(X = k \mid X + Y = n)$ given that $X + Y = n$, for all $k, n \in \mathbb{N}$.
(c) Assume that respective parameters of the distributions of X and Y are random, independent, and chosen according to an exponential distribution with parameter $\theta > 0$.

Give the probability distributions of X and Y, and compute the conditional distribution $\mathbb{P}(X = k \mid X + Y = n)$ given that $X + Y = n$, for all $k, n \in \mathbb{N}$.

(d) Assume now that X and Y have same random parameter represented by a single exponentially distributed random variable Λ with parameter $\theta > 0$, independent of X and Y.

Compute the conditional distribution $\mathbb{P}(X = k \mid X + Y = n)$ given that $X + Y = n$, for all $k, n \in \mathbb{N}$.

Exercise 1.7 A red pen and a green pen are put in a hat. A pen is chosen at random in the hat, and replaced inside after it color has been noted.

- In case the pen is of red color, then a supplementary red pen is placed in the hat.
- On the other hand if the pen color is green, then another green pen is added.

After this first part of the experiment is completed, a second pen is chosen at random.

Determine the probability that the first drawn pen was red, given that the color of the second pen chosen was red.

Exercise 1.8 A machine relies on the functioning of three parts, each of which having a probability $1 - p$ of being under failure, and a probability p of functioning correctly. All three parts are functioning independently of the others, and the machine is working if and only if two at least of the parts are operating.

(a) Compute the probability that the machine is functioning.
(b) Suppose that the machine itself is set in a random environment in which the value of the probability p becomes random. Precisely we assume that p is a uniform random variable taking real values between 0 and 1, independently of the state of the system.

Compute the probability that the machine operates in this random environment.

Chapter 2
Gambling Problems

This chapter consists in a detailed study of a fundamental example of random walk that can only evolve by going up of down by one unit within the finite state space $\{0, 1, \ldots, S\}$. This allows us in particular to have a first look at the technique of first step analysis that will be repeatedly used in the general framework of Markov chains, particularly in Chap. 5.

2.1 Constrained Random Walk

To begin, let us repeat that this chapter on "gambling problems" is not primarily designed to help a reader dealing with problem gambling, although some comments on this topic are made of the end of Sect. 2.3.

We consider an amount $\$S$ of S dollars which is to be shared between two players A and B. At each round, Player A may earn \$1 with probability $p \in (0, 1)$, and in this case Player B loses \$1. Conversely, Player A may lose \$1 with probability $q := 1 - p$, in which case Player B gains \$1, and the successive rounds are independent.

We let X_n represent the wealth of Player A at time $n \in \mathbb{N}$, while $S - X_n$ represents the wealth of Player B at time $n \in \mathbb{N}$.

The initial wealth X_0 of Player A could be negative, but for simplicity we will assume that it is comprised between 0 and S. Assuming that the value of X_n, $n \geq 0$, belongs to $\{1, 2, \ldots, S - 1\}$ at the time step n, at the next step $n + 1$ we will have

$$X_{n+1} = \begin{cases} X_n + 1 \text{ if Player A wins round } n + 1, \\ X_n - 1 \text{ if Player B wins round } n + 1. \end{cases}$$

Moreover, as soon as X_n hits one of the boundary points $\{0, S\}$, the process remains frozen at that state over time, i.e.

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4_2

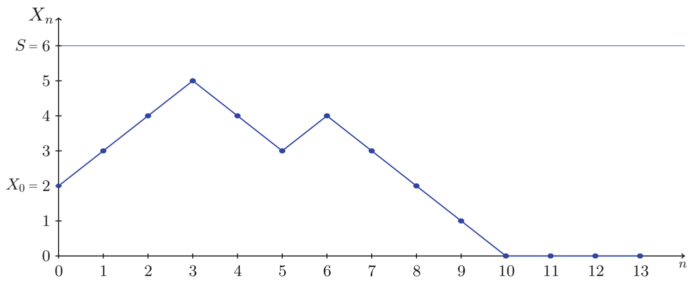


Fig. 2.1 Sample path of a gambling process $(X_n)_{n\in\mathbb{N}}$

$$X_n = 0 \Longrightarrow X_{n+1} = 0 \quad \text{and} \quad X_n = S \Longrightarrow X_{n+1} = S,$$

i.e.

$$\mathbb{P}(X_{n+1} = 0 \mid X_n = 0) = 1 \quad \text{and} \quad \mathbb{P}(X_{n+1} = S \mid X_n = S) = 1, \quad n \in \mathbb{N}.$$

In other words, the game ends whenever the wealth of any of the two players reaches \$0, in which case the other player's account contains \$$S$ (see Fig. 2.1).
Among the main issues of interest are:

- the probability that Player A (or B) gets eventually ruined,
- the mean duration of the game.

We will also be interested in the probability distribution of the random game duration T, i.e. in the knowledge of $\mathbb{P}(T = n)$, $n \geq 0$.

According to the above problem description, for all $n \in \mathbb{N}$ we have

$$\mathbb{P}(X_{n+1} = k + 1 \mid X_n = k) = p \quad \text{and} \quad \mathbb{P}(X_{n+1} = k - 1 \mid X_n = k) = q,$$

$k = 1, 2, \ldots, S - 1$, and in this case the chain is said to be *time homogeneous* since the transition probabilities do not depend on the time index n.

Since we do not focus on the behavior of the chain after it hits states 0 or S, the probability distribution of X_{n+1} given $\{X_n = 0\}$ or $\{X_n = S\}$ can be left unspecified.

The probability space Ω corresponding to this experiment could be taken as the (uncountable) set

$$\Omega := \{-1, +1\}^{\mathbb{N}} = \left\{\omega = (\omega_0, \omega_1, \ldots) \ : \ \omega_i = \pm 1, \ n \in \mathbb{N}\right\},$$

with any element $\omega \in \Omega$ represented by a countable sequence of $+1$ or -1, depending whether the process goes up or down at each time step. However, in the sequel we will not focus on this particular expression of Ω.

2.2 Ruin Probabilities

We are interested in the event

$$R_A = \text{"Player } A \text{ loses all his capital at some time"} = \bigcup_{n\in\mathbb{N}} \{X_n = 0\}, \tag{2.2.1}$$

and in computing the conditional probability

$$f_S(k) := \mathbb{P}(R_A \mid X_0 = k), \qquad k = 0, 1, \ldots, S. \tag{2.2.2}$$

Pathwise Analysis

First, let us note that the problem is easy to solve in the case $S = 1$, $S = 2$ and $S = 3$.

(i) $\underline{S = 1}$.

In this case the boundary $\{0, 1\}$ is reached from time 0 and we find

$$\begin{cases} f_1(0) = \mathbb{P}(R_A \mid X_0 = 0) = 1, \\ f_1(1) = \mathbb{P}(R_A \mid X_0 = 1) = 0. \end{cases} \tag{2.2.3}$$

(ii) $\underline{S = 2}$.

In this case we find

$$\begin{cases} f_2(0) = \mathbb{P}(R_A \mid X_0 = 0) = 1, \\ f_2(1) = \mathbb{P}(R_A \mid X_0 = 1) = q, \\ f_2(2) = \mathbb{P}(R_A \mid X_0 = 2) = 0. \end{cases} \tag{2.2.4}$$

(iii) $\underline{S = 3}$.

The value of $f_2(1) = \mathbb{P}(R_A \mid X_0 = 1)$ is computed by noting that starting from state ①, one can reach state ⓪ only by an odd number $2n + 1$ of step, $n \in \mathbb{N}$, and that every such path decomposes into $n + 1$ independent downwards steps, each of them having probability q, and n upwards steps, each of them with probability p. By summation over n using the geometric series identity (A.3), this yields

$$\begin{cases} f_3(0) = \mathbb{P}(R_A \mid X_0 = 0) = 1, \\ f_3(1) = \mathbb{P}(R_A \mid X_0 = 1) = q\displaystyle\sum_{n=0}^{\infty}(pq)^n = \frac{q}{1-pq}, \\ f_3(2) = \mathbb{P}(R_A \mid X_0 = 2) = q^2\displaystyle\sum_{n=0}^{\infty}(pq)^n = \frac{q^2}{1-pq}, \\ f_3(3) = \mathbb{P}(R_A \mid X_0 = 3) = 0. \end{cases} \tag{2.2.5}$$

The value of $f_3(2)$ is computed similarly by considering $n+2$ independent downwards steps, each of them with probability q, and n upwards steps, each of them with probability p. Clearly, things become quite complicated for $S \geq 4$, and increasingly difficult as S gets larger.

First Step Analysis

The general case will be solved by the method of *first step analysis*, which will be repeatedly applied to other Markov processes in Chaps. 3 and 5 and elsewhere.

Lemma 2.1 *For all* $k = 1, 2, \ldots, S-1$ *we have*

$$\mathbb{P}(R_A \mid X_0 = k) = p\mathbb{P}(R_A \mid X_0 = k+1) + q\mathbb{P}(R_A \mid X_0 = k-1).$$

Proof The idea is to apply conditioning given the first transition from X_0 to X_1. For all $k = 1, 2, \ldots, S-1$, by (1.2.1) we have

$$\begin{aligned}
&\mathbb{P}(R_A \mid X_0 = k)\\
&= \mathbb{P}(R_A \text{ and } X_1 = k+1 \mid X_0 = k) + \mathbb{P}(R_A \text{ and } X_1 = k-1 \mid X_0 = k)\\
&= \frac{\mathbb{P}(R_A \text{ and } X_1 = k+1 \text{ and } X_0 = k)}{\mathbb{P}(X_0 = k)} + \frac{\mathbb{P}(R_A \text{ and } X_1 = k-1 \text{ and } X_0 = k)}{\mathbb{P}(X_0 = k)}\\
&= \frac{\mathbb{P}(R_A \text{ and } X_1 = k+1 \text{ and } X_0 = k)}{\mathbb{P}(X_1 = k+1 \text{ and } X_0 = k)} \times \frac{\mathbb{P}(X_1 = k+1 \text{ and } X_0 = k)}{\mathbb{P}(X_0 = k)}\\
&\quad + \frac{\mathbb{P}(R_A \text{ and } X_1 = k-1 \text{ and } X_0 = k)}{\mathbb{P}(X_1 = k-1 \text{ and } X_0 = k)} \times \frac{\mathbb{P}(X_1 = k-1 \text{ and } X_0 = k)}{\mathbb{P}(X_0 = k)}\\
&= \mathbb{P}(R_A \mid X_1 = k+1 \text{ and } \ X_0 = k)\mathbb{P}(X_1 = k+1 \mid X_0 = k)\\
&\quad + \mathbb{P}(R_A \mid X_1 = k-1 \text{ and } \ X_0 = k)\mathbb{P}(X_1 = k-1 \mid X_0 = k)\\
&= p\mathbb{P}(R_A \mid X_1 = k+1 \text{ and } X_0 = k) + q\mathbb{P}(R_A \mid X_1 = k-1 \text{ and } X_0 = k)\\
&= p\mathbb{P}(R_A \mid X_0 = k+1) + q\mathbb{P}(R_A \mid X_0 = k-1),
\end{aligned}$$

where we used Lemma 2.2 below on the last step. □

In the case $S = 3$, Lemma 2.1 shows that

$$\begin{cases}
f_3(0) = \mathbb{P}(R_A \mid X_0 = 0) = 1,\\
f_3(1) = pf_3(2) + qf_3(0) = pf_3(2) + q = pqf_3(1) + q,\\
f_3(2) = pf_3(3) + qf_3(1) = qf_3(1) = pqf_3(2) + q^2,\\
f_3(3) = \mathbb{P}(R_A \mid X_0 = 3) = 0,
\end{cases}$$

which can be easily solved to recover the result of (2.2.5).

More generally, Lemma 2.1 shows that the function

$$f_S : \{0, 1, \ldots, S\} \longrightarrow [0, 1]$$

defined by (2.2.2) satisfies the linear equation[1]

$$f(k) = pf(k+1) + qf(k-1), \qquad k = 1, 2, \ldots, S-1, \quad (2.2.6)$$

subject to the *boundary conditions*

$$f_S(0) = \mathbb{P}(R_A \mid X_0 = 0) = 1, \quad (2.2.7)$$

and

$$f_S(S) = \mathbb{P}(R_A \mid X_0 = S) = 0, \quad (2.2.8)$$

for $k \in \{0, S\}$. It can be easily checked that the expressions (2.2.3), (2.2.4) and (2.2.5) do satisfy the above Eq. (2.2.6) and the boundary conditions (2.2.7) and (2.2.8).

Note that Lemma 2.1 is frequently stated without proof. The last step of the proof stated above rely on the following lemma, which shows that the data of X_1 entirely determines the probability of the ruin event R_A. In other words, the probability of ruin depends only on the initial amount k owned by the gambler when he enters the casino. Whether he enters the casino at time 1 with $X_1 = k \pm 1$ or at time 0 with $X_0 = k \pm 1$ makes no difference on the ruin probability.

Lemma 2.2 *For all* $k = 1, 2, \ldots, S-1$ *we have*

$$\mathbb{P}(R_A \mid X_1 = k \pm 1 \text{ and } X_0 = k) = \mathbb{P}(R_A \mid X_1 = k \pm 1) = \mathbb{P}(R_A \mid X_0 = k \pm 1).$$

In other words, the ruin probability depends on the data of the starting point and not on the starting time.

Proof This relation can be shown in various ways:

1. Descriptive proof (*preferred*): we note that given $X_1 = k+1$, the transition from X_0 to X_1 has no influence on the future of the process after time 1, and the probability of ruin starting at time 1 is the same as if the process is started at time 0.
2. Algebraic proof: first for $1 \leq k \leq S-1$ and $k \pm 1 \geq 1$, letting $\tilde{X}_0 := X_1 - Z$ where $Z \simeq X_1 - X_0$ has same distribution as $X_1 - X_0$ and is independent of X_1, by (2.2.1) we have

[1] Due to the relation $(f+g)(k) = f(k) + g(k)$ we can check that if f and g are two solutions of (2.2.6) then $f+g$ is also a solution of (2.2.6), hence the equation is *linear*.

$$\mathbb{P}(R_A \mid X_1 = k \pm 1 \text{ and } X_0 = k) = \mathbb{P}\left(\bigcup_{n=0}^{\infty}\{X_n = 0\}\,\middle|\, X_1 = k \pm 1,\ X_0 = k\right)$$
$$= \frac{\mathbb{P}\left(\left(\bigcup_{n=0}^{\infty}\{X_n = 0\}\right) \cap \{X_1 = k \pm 1\} \cap \{X_0 = k\}\right)}{\mathbb{P}\left(\{X_1 = k \pm 1\} \cap \{X_0 = k\}\right)}$$
$$= \frac{\mathbb{P}\left(\bigcup_{n=0}^{\infty}\left(\{X_n = 0\} \cap \{X_1 = k \pm 1\} \cap \{X_0 = k\}\right)\right)}{\mathbb{P}\left(\{X_1 = k \pm 1\} \cap \{X_0 = k\}\right)}$$
$$= \frac{\mathbb{P}\left(\bigcup_{n=2}^{\infty}\left(\{X_n = 0\} \cap \{X_1 = k \pm 1\} \cap \{X_0 = k\}\right)\right)}{\mathbb{P}\left(\{X_1 = k \pm 1\} \cap \{X_0 = k\}\right)}$$
$$= \frac{\mathbb{P}\left(\left(\bigcup_{n=2}^{\infty}\{X_n = 0\}\right) \cap \{X_1 = k \pm 1\} \cap \{X_0 = k\}\right)}{\mathbb{P}\left(\{X_1 = k \pm 1\} \cap \{X_0 = k\}\right)}$$

$$= \frac{\mathbb{P}\left(\left(\bigcup_{n=2}^{\infty}\{X_n = 0\}\right) \cap \{X_1 = k \pm 1\} \cap \{X_1 - \tilde{X}_0 = \pm 1\}\right)}{\mathbb{P}\left(\{X_1 = k \pm 1\} \cap \{X_1 - \tilde{X}_0 = \pm 1\}\right)} \tag{2.2.9}$$
$$= \frac{\mathbb{P}\left(\left(\bigcup_{n=2}^{\infty}\{X_n = 0\}\right) \cap \{X_1 = k \pm 1\}\right)\mathbb{P}\left(\{X_1 - \tilde{X}_0 = \pm 1\}\right)}{\mathbb{P}\left(\{X_1 = k \pm 1\}\right)\mathbb{P}\left(\{X_1 - \tilde{X}_0 = \pm 1\}\right)} \tag{2.2.10}$$

$$= \mathbb{P}\left(\bigcup_{n=2}^{\infty}\{X_n = 0\}\,\middle|\,\{X_1 = k \pm 1\}\right)$$
$$= \mathbb{P}\left(\bigcup_{n=1}^{\infty}\{X_n = 0\}\,\middle|\,\{X_0 = k \pm 1\}\right)$$
$$= \mathbb{P}(R_A \mid X_0 = k \pm 1),$$

otherwise if $k = 1$ we easily find that

$$\mathbb{P}(R_A \mid X_1 = 0 \text{ and } X_0 = 1) = 1 = \mathbb{P}(R_A \mid X_0 = 0),$$

since $\{X_1 = 0\} \subset R_A = \bigcup_{n\in\mathbb{N}}\{X_n = 0\}$. Note that when switching from (2.2.9) to (2.2.10), using $\tilde{X}_0 := X_1 - Z$ we regard the process increment starting from X_1 as run backward in time.

□

In the remaining of this section we will prove that in the non-symmetric case $p \neq q$ the solution of (2.26) is given by

$$\boxed{f_S(k) = \mathbb{P}(R_A \mid X_0 = k) = \frac{(q/p)^k - (q/p)^S}{1-(q/p)^S} = \frac{(p/q)^{S-k} - 1}{(p/q)^S - 1},} \tag{2.2.11}$$

$k = 0, 1, \ldots, S$, and that in the symmetric case $p = q = 1/2$ the solution of (2.2.6) is given by

$$\boxed{f_S(k) = \mathbb{P}(R_A \mid X_0 = k) = \frac{S-k}{S} = 1 - \frac{k}{S},} \tag{2.2.12}$$

$k = 0, 1, \ldots, S$, cf. also Exercise 2.3 for a different derivation.

Remark that (2.2.11) and (2.2.12) do satisfy both boundary conditions (2.2.7) and (2.2.8). When the number S of states becomes large we find that, for all $k \geq 0$,

$$f_\infty(k) := \lim_{S\to\infty} \mathbb{P}(R_A \mid X_0 = k) = \begin{cases} 1 & \text{if } q \geq p, \\ \left(\dfrac{q}{p}\right)^k & \text{if } p > q, \end{cases} \tag{2.2.13}$$

which represents the probability of hitting the origin starting from state ⓚ, cf. also 3.4.16 below and Exercise 3.2-(c) for a different derivation of this statement.

Exercise: Check that (2.2.11) agrees with (2.2.4) and (2.2.5) when $S = 2$ and $S = 3$.

In the graph of Fig. 2.2 the ruin probability (2.2.11) is plotted as a function of k for $p = 0.45$ and $q = 0.55$.

We now turn to the solution of (2.2.6), for which we develop two different approaches (called here the "standard solution" and the "direct solution") that both recover (2.2.11) and (2.2.12).

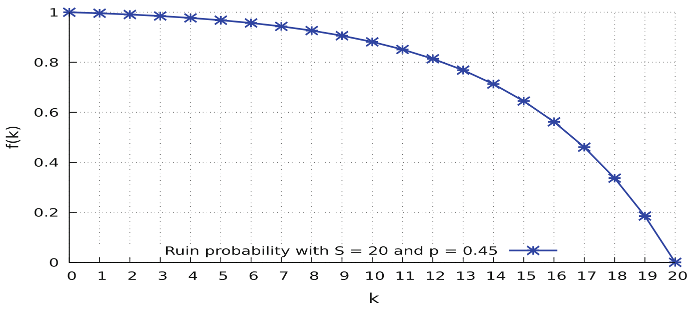


Fig. 2.2 Ruin probability $f_{20}(k)$ function of $X_0 = k \in [0, 20]$ for $S = 20$ and $p = 0.45$

Standard Solution Method

We decide to look for a solution of (2.2.6) of the form[2]

$$k \longmapsto f(k) = Ca^k, \tag{2.2.14}$$

where C and a are constants which will be determined from the boundary conditions and from the Eq. (2.2.6), respectively.

Substituting (2.2.14) into (2.2.6) when C is non-zero yields the *characteristic equation*

$$pa^2 - a + q = p(a-1)(a-q/p) = 0, \tag{2.2.15}$$

of degree 2 in the unknown a, and this equation admits in general two solutions a_1 and a_2 given by

$$\{a_1, a_2\} = \left\{ \frac{1+\sqrt{1-4pq}}{2p}, \frac{1-\sqrt{1-4pq}}{2p} \right\} = \left\{ 1, \frac{q}{p} \right\} = \{1, r\},$$

for all $p \in (0, 1]$, with

$$a_1 = 1 \quad \text{and} \quad a_2 = r = \frac{q}{p}.$$

Note that we have $a_1 = a_2 = 1$ in case $p = q = 1/2$.

Non-symmetric Case*: $p \neq q$ - *Proof of (2.2.11)

In this case we have $p \neq q$, i.e.[3] $r \neq 1$, and

$$f(k) = C_1 a_1^k = C_1 \qquad \text{and} \qquad f(k) = C_2 r^k$$

[2]Where did we get this idea? From intuition, experience, or empirically by multiple trials and errors.

[3]From the Latin "id est" meaning "that is".

are both solutions of (2.2.6). Since (2.2.6) is linear, the sum of two solutions remains a solution, hence the general solution of (2.2.6) is given by

$$f_S(k) = C_1 a_1^k + C_2 a_2^k = C_1 + C_2 r^k, \qquad k = 0, 1, \ldots, S, \tag{2.2.16}$$

where $r = q/p$ and C_1, C_2 are two constants to be determined from the boundary conditions.

From (2.2.7), (2.2.8) and (2.2.16) we have

$$\begin{cases} f_S(0) = 1 = C_1 + C_2, \\ \\ f_S(S) = 0 = C_1 + C_2 r^S, \end{cases} \tag{2.2.17}$$

and solving the system (2.2.17) of two equations we find

$$C_1 = -\frac{r^S}{1 - r^S} \text{ and } C_2 = \frac{1}{1 - r^S},$$

which yields (2.2.11) as by (2.2.16) we have

$$f_S(k) = C_1 + C_2 r^k = \frac{r^k - r^S}{1 - r^S} = \frac{(q/p)^k - (q/p)^S}{1 - (q/p)^S}, \quad k = 0, 1, \ldots, S.$$

Symmetric Case*: *p=q=1/2* - *Proof of (2.2.12)

In this case, Eq. (2.2.6) rewrites as

$$f(k) = \frac{1}{2} f(k+1) + \frac{1}{2} f(k-1), \qquad k = 1, 2, \ldots, S-1, \tag{2.2.18}$$

and we have $r = 1$ (fair game) and (2.2.15) reads

$$a^2 - 2a + 1 = (a-1)^2 = 0,$$

which has the unique solution $a = 1$, since the constant function $f(k) = C$ is solution of (2.2.6).

However this is not enough and we need to combine $f(k) = C_1$ with a second solution. Noting that $g(k) = C_2 k$ is also solution of (2.2.6), the general solution is found to have the form

$$f_S(k) = f(k) + g(k) = C_1 + C_2 k. \tag{2.2.19}$$

From (2.2.7), (2.2.8) and (2.2.19) we have

$$\begin{cases} f_S(0) = 1 = C_1, \\ \\ f_S(S) = 0 = C_1 + C_2 S, \end{cases} \tag{2.2.20}$$

and solving the system (2.2.20) of two equations we find

$$C_1 = 1 \text{ and } C_2 = -1/S,$$

which yields the quite intuitive solution

$$f_S(k) = \mathbb{P}(R_A \mid X_0 = k) = \frac{S-k}{S} = 1 - \frac{k}{S}, \quad k = 0, 1, \ldots, S. \tag{2.2.21}$$

Direct Solution Method

Noting that $p + q = 1$ and due to its special form, we can rewrite (2.2.6) as

$$(p+q) f_S(k) = p f_S(k+1) + q f_S(k-1),$$

$k = 1, 2, \ldots, S-1$, i.e. as the *difference equation*

$$p(f_S(k+1) - f_S(k)) - q(f_S(k) - f_S(k-1)) = 0, \tag{2.2.22}$$

$k = 1, 2, \ldots, S-1$, which rewrites as

$$f_S(k+1) - f_S(k) = \frac{q}{p}(f_S(k) - f_S(k-1)), \qquad k = 1, 2, \ldots, S-1,$$

hence for $k = 1$,

$$f_S(2) - f_S(1) = \frac{q}{p}(f_S(1) - f_S(0)),$$

and for $k = 2$,

$$f_S(3) - f_S(2) = \frac{q}{p}(f_S(2) - f_S(1)) = \left(\frac{q}{p}\right)^2 (f_S(1) - f_S(0)),$$

and following by induction on k we can show that

$$f_S(k+1) - f_S(k) = \left(\frac{q}{p}\right)^k (f_S(1) - f_S(0)), \tag{2.2.23}$$

$k = 0, 1, \ldots, S-1$. Next, by the telescoping sum

$$f_S(n) = f_S(0) + \sum_{k=0}^{n-1} (f_S(k+1) - f_S(k)),$$

Relation (2.2.23) implies

$$f_S(n) = f_S(0) + (f_S(1) - f_S(0)) \sum_{k=0}^{n-1} \left(\frac{q}{p}\right)^k, \tag{2.2.24}$$

$n = 1, 2, \ldots, S-1$. The remaining question is how to find $f_S(1) - f_S(0)$, knowing that $f_S(0) = 1$ by (2.2.7).

***Non-symmetric Case*: $p \neq q$**

In this case we have $r = q/p \neq 1$ and we get

$$\begin{aligned} f_S(n) &= f_S(0) + (f_S(1) - f_S(0)) \sum_{k=0}^{n-1} r^k \\ &= f_S(0) + \frac{1-r^n}{1-r}(f_S(1) - f_S(0)), \end{aligned} \tag{2.2.25}$$

$n = 1, 2, \ldots, S-1$, where we used (A.2).

Conditions (2.2.7) and (2.2.8) show that

$$0 = f_S(S) = 1 + \frac{1-r^S}{1-r}(f_S(1) - f_S(0)),$$

hence

$$f_S(1) - f_S(0) = -\frac{1-r}{1-r^S},$$

and combining this relation with (2.2.25) yields

$$f_S(n) = f_S(0) - \frac{1-r^n}{1-r^S} = 1 - \frac{1-r^n}{1-r^S} = \frac{r^n - r^S}{1-r^S},$$

$n = 0, 1, \ldots, S$, which recovers (2.2.11).

***Symmetric Case*: $p = q = 1/2$**

In this case we have $r = 1$ and in order to solve (2.2.18) we note that (2.2.22) simply becomes

$$f_S(k+1) - f_S(k) = f_S(1) - f_S(0), \qquad k = 0, 1, \ldots, S-1,$$

and (2.2.24) reads

$$f_S(n) = f_S(0) + n(f_S(1) - f_S(0)), \qquad n = 1, 2, \ldots, S-1,$$

which has the form (2.2.19). Then the conditions $f_S(0) = 1$ and $f_S(S) = 0$, cf. (2.2.7) and (2.2.8), yield

$$0 = f_S(S) = 1 + S(f_S(1) - f_S(0)), \quad \text{hence} \quad f_S(1) - f_S(0) = -\frac{1}{S},$$

and

$$f_S(n) = f_S(0) - \frac{n}{S} = 1 - \frac{n}{S} = \frac{S-n}{S},$$

$n = 0, 1, \ldots, S$, which coincides with (2.2.21).

Remark 2.3 Note that when $p = q = 1/2$, (2.2.22) can be read as a discretization of the continuous Laplace equation as

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(x) &\simeq \frac{\partial f}{\partial x}(x+1/2) - \frac{\partial f}{\partial x}(x-1/2) \\ &\simeq (f(x+1) - f(x) - (f(x) - f(x-1))) \\ &= 0, \qquad x \in \mathbb{R}, \end{aligned} \tag{2.2.26}$$

which admits a solution of the form

$$f(x) = f(0) + xf'(0) = f(0) + x(f(1) - f(0)), \qquad x \in \mathbb{R},$$

showing the intuition behind the linear form of (2.2.19).

In order to compute the probability of ruin of Player B given that $X_0 = k$ we only need to swap k to $S-k$ and to exchange p and q in (2.2.11). In other words, when $X_0 = k$ then Player B starts with an initial amount $S-k$ and a probability q of winning each round, which by (2.2.11) yields

$$\boxed{\mathbb{P}(R_B \mid X_0 = k) = \frac{(p/q)^{S-k} - (p/q)^S}{1-(p/q)^S} = \frac{1-(q/p)^k}{1-(q/p)^S} \quad \text{if } p \neq q,} \tag{2.2.27}$$

where

$$R_B := \text{“Player } B \text{ loses all his capital at some time”} = \bigcup_{n \in \mathbb{N}} \{X_n = S\}.$$

In the symmetric case $p = q = 1/2$ we similarly find

$$\boxed{\mathbb{P}(R_B \mid X_0 = k) = \frac{k}{S}, \quad k = 0, 1, \ldots, S, \quad \text{if } p = q = \frac{1}{2},} \tag{2.2.28}$$

cf. also Exercise 2.3 below.

Note that (2.2.27) and (2.2.28) satisfy the expected boundary conditions

$$\mathbb{P}(R_B \mid X_0 = 0) = 0 \quad \text{and} \quad \mathbb{P}(R_B \mid X_0 = S) = 1,$$

since X_0 represents the wealth of Player A at time 0.

By (2.2.11) and (2.2.27) we can check that[4]

$$
\begin{aligned}
\mathbb{P}(R_A \cup R_B \mid X_0 = k) &= \mathbb{P}(R_A \mid X_0 = k) + \mathbb{P}(R_B \mid X_0 = k) \\
&= \frac{(q/p)^k - (q/p)^S}{1 - (q/p)^S} + \frac{1 - (q/p)^k}{1 - (q/p)^S} \\
&= 1, \qquad k = 0, 1, \ldots, S,
\end{aligned} \tag{2.2.29}
$$

which means that eventually one of the two players has to lose the game. This means in particular that, with probability one, the game cannot continue endlessly.

In other words, we have

$$
\mathbb{P}(R_A^c \cap R_B^c \mid X_0 = k) = 0, \qquad k = 0, 1, \ldots, S.
$$

In the particular case $S = 3$ we can indeed check by (1.2.4) that, taking

$$
A_n := \bigcap_{k=1}^{n} \{X_{2k-1} = 1 \text{ and } X_{2k} = 2\}, \qquad n \geq 1,
$$

the sequence $(A_n)_{n\geq 1}$ is nonincreasing, hence

$$
\begin{aligned}
&\mathbb{P}\left(\bigcap_{n\geq 1} \{X_{2n-1} = 1 \text{ and } X_{2n} = 2\} \,\Big|\, X_0 = 2\right) \\
&= \mathbb{P}\left(\bigcap_{n\geq 1} \bigcap_{k=1}^{n} \{X_{2k-1} = 1 \text{ and } X_{2k} = 2\} \,\Big|\, X_0 = 2\right) \\
&= \lim_{n\to\infty} \mathbb{P}\left(\bigcap_{k=1}^{n} \{X_{2k-1} = 1 \text{ and } X_{2k} = 2\} \,\Big|\, X_0 = 2\right) \\
&= p \lim_{n\to\infty} (pq)^n = 0,
\end{aligned}
$$

since we always have $0 \leq pq < 1$, and where we used (1.2.4). However, this is a priori not completely obvious when $S \geq 4$.

When the number S of states becomes large, (2.2.7) also shows that for all $k \geq 0$ we have

$$
\lim_{S\to\infty} \mathbb{P}(R_B \mid X_0 = k) = \begin{cases} 0 & \text{if } p \leq q, \\[2ex] 1 - \left(\dfrac{q}{p}\right)^k & \text{if } p > q, \end{cases}
$$

which represents the complement of the probability (2.2.13) of hitting the origin starting from state Ⓚ, and is the probability that the process $(X_n)_{n\in\mathbb{N}}$ "escapes to infinity".

[4]Exercise: check by hand computation that the equality to 1 holds as stated.

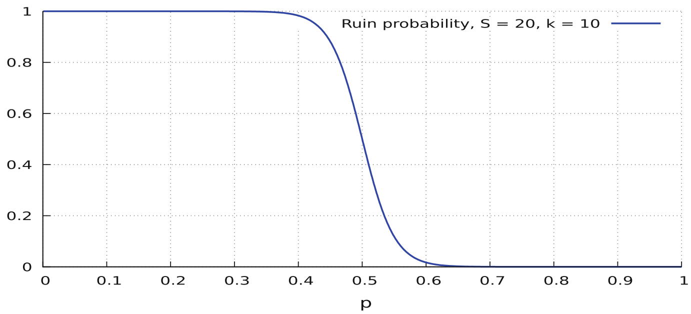


Fig. 2.3 Ruin probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$

In Fig. 2.3 above the ruin probability (2.2.11) is plotted as a function of p for $S = 20$ and $k = 10$.

Gambling machines in casinos are computer controlled and most countries permit by law a certain degree of "unfairness" (see the notions of "payout percentage" or "return to player") by taking $p < 1/2$ in order to allow the house to make an income.[5] Interestingly, we can note that taking e.g. $p = 0.45 < 1/2$ gives a ruin probability

$$\mathbb{P}(R_A \mid X_0 = 10) = 0.8815,$$

almost equal to 90%, which means that the slightly unfair probability $p = 0.45$ at the level of each round translates into a probability of only $0.1185 \simeq \%12$ of finally winning the game, i.e. a division by 4 from 0.45, although the average proportion of winning rounds is still 45%.

Hence a "slightly unfair" game on each round can become devastatingly unfair in the long run. Most (but not all) gamblers are aware that gambling machines are slightly unfair, however most people would intuitively believe that a small degree of unfairness on each round should only translate into a reasonably low degree of unfairness in the long run.

2.3 Mean Game Duration

Let now

$$T_{0,S} = \inf\{n \geq 0 \ : \ X_n = 0 \text{ or } X_n = S\}$$

denote the time[6] until any of the states ⓪ or Ⓢ are reached by $(X_n)_{n \in \mathbb{N}}$, with $T_{0,S} = \infty$ in case neither states are ever reached, i.e. when there exists no integer $n \geq 0$ such that $X_n = 0$ or $X_n = S$ (see Fig. 2.4).

[5] In this game, the payout is \$2 and the payout percentage is $2p$.

[6] The notation "inf" stands for "infimum", meaning the smallest $n \geq 0$ such that $X_n = 0$ or $X_n = S$, if such an n exists.

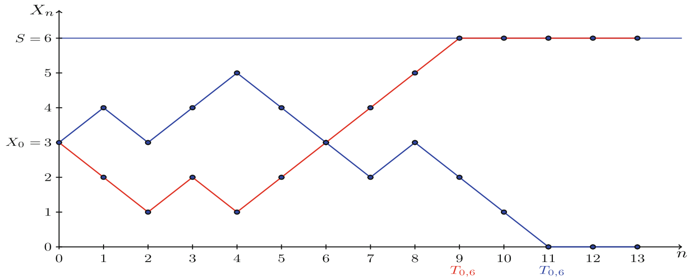


Fig. 2.4 Sample paths of a gambling process $(X_n)_{n\in\mathbb{N}}$

Note that by (2.2.29) we have

$$\mathbb{P}(T_{0,S} < \infty \mid X_0 = k) = \mathbb{P}(R_A \cup R_B \mid X_0 = k) = 1, \quad k = 0, 1, \ldots, S.$$

and therefore

$$\mathbb{P}(T_{0,S} = \infty \mid X_0 = k) = 0, \qquad k = 0, 1, \ldots, S.$$

We are now interested in computing the expected duration

$$h_S(k) := \mathbb{E}[T_{0,S} \mid X_0 = k]$$

of the game given that Player A starts with a wealth equal to $X_0 = k \in \{0, 1, \ldots, S\}$. Clearly, we have the boundary conditions

$$\begin{cases} h_S(0) = \mathbb{E}[T_{0,S} \mid X_0 = 0] = 0, & \text{(2.3.1a)} \\ h_S(S) = \mathbb{E}[T_{0,S} \mid X_0 = S] = 0. & \text{(2.3.1b)} \end{cases}$$

We start by considering the particular cases $S = 2$ and $S = 3$.

(i) $\underline{S = 2}$.

We have

$$T_{0,2} = \begin{cases} 0 \ \text{ if } X_0 = 0, \\ \\ 1 \ \text{ if } X_0 = 1, \\ \\ 0 \ \text{ if } X_0 = 2, \end{cases}$$

thus $T_{0,2}$ is *deterministic* given the value of X_0 and we simply have $h_2(1) = T_{0,2} = 1$ when $X_0 = 1$.

(ii) $\underline{S = 3}$.

In this case the probability distribution of $T_{0,3}$ given $X_0 \in \{0, 1, 2, 3\}$ can be determined explicitly and we find, when $X_0 = 1$,

$$\begin{cases} \mathbb{P}(T_{0,3} = 2k \mid X_0 = 1) = p^2(pq)^{k-1}, & k \geq 1, \\ \mathbb{P}(T_{0,3} = 2k + 1 \mid X_0 = 1) = q(pq)^k, & k \geq 0, \end{cases}$$

since in an even number $2k$ of steps we can only exit through state ③ after starting from ①, while in an odd number $2k + 1$ of steps we can only exit through state ⓪. By exchanging p with q in the above formulas we get, when $X_0 = 2$,

$$\begin{cases} \mathbb{P}(T_{0,3} = 2k \mid X_0 = 2) = q^2(pq)^{k-1}, & k \geq 1, \\ \mathbb{P}(T_{0,3} = 2k + 1 \mid X_0 = 2) = p(pq)^k, & k \geq 0, \end{cases}$$

whereas $T_{0,3} = 0$ whenever $X_0 = 0$ or $X_0 = 3$.
As a consequence, we can directly compute

$$\begin{aligned} h_3(2) = \mathbb{E}[T_{0,3} \mid X_0 = 2] &= 2\sum_{k=1}^{\infty} k\mathbb{P}(T_{0,3} = 2k \mid X_0 = 2) \\ &\quad + \sum_{k=0}^{\infty}(2k + 1)\mathbb{P}(T_{0,3} = 2k + 1 \mid X_0 = 2) \\ &= 2q^2 \sum_{k=1}^{\infty} k(pq)^{k-1} + p\sum_{k=0}^{\infty}(2k + 1)(pq)^k \\ &= \frac{2q^2}{(1 - pq)^2} + \frac{2p^2q}{(1 - pq)^2} + \frac{p}{1 - pq} \\ &= \frac{2q^2 + p + qp^2}{(1 - pq)^2} \\ &= \frac{1 + q}{1 - pq}, \end{aligned} \tag{2.3.2}$$

where we applied (A.4), and by exchanging p and q we get

$$h_3(1) = \mathbb{E}[T_{0,3} \mid X_0 = 1] = \frac{2p^2 + q + pq^2}{(1 - pq)^2} = \frac{1 + p}{1 - pq}. \tag{2.3.3}$$

Again, things can become quite complicated for $S \geq 4$, and increasingly difficult when S becomes larger.

In the general case $S \geq 4$ we will only compute the conditional expectation of $T_{0,S}$ and not its probability distribution. For this we rely again on *first step analysis*, as stated in the following lemma.

Lemma 2.4 *For all* $k = 1, 2, \ldots, S-1$ *we have*

$$\mathbb{E}[T_{0,S} \mid X_0 = k] = 1 + p\,\mathbb{E}[T_{0,S} \mid X_0 = k+1] + q\,\mathbb{E}[T_{0,S} \mid X_0 = k-1].$$

Proof We condition on the first transition from X_0 to X_1. Using the equality $\mathbb{1}_A = \mathbb{1}_{A\cap B} + \mathbb{1}_{A\cap B^c}$ under the form

$$\mathbb{1}_{\{X_0=k\}} = \mathbb{1}_{\{X_1=k+1, X_0=k\}} + \mathbb{1}_{\{X_1=k-1, X_0=k\}},$$

cf. (1.4.2), and conditional expectations we show, by first step analysis, that for all $k = 1, 2, \ldots, S-1$, applying Lemma 1.4 and (1.6.6) successively to $A = \{X_0 = k\}$, $A = \{X_0 = k-1\}$ and $A = \{X_0 = k+1\}$, we have

$$\begin{aligned}
\mathbb{E}[T_{0,S} \mid X_0 = k] &= \frac{1}{\mathbb{P}(X_0 = k)}\,\mathbb{E}\left[T_{0,S}\mathbb{1}_{\{X_0=k\}}\right] \\
&= \frac{1}{\mathbb{P}(X_0 = k)}\left(\mathbb{E}\left[T_{0,S}\mathbb{1}_{\{X_1=k+1,X_0=k\}}\right] + \mathbb{E}\left[T_{0,S}\mathbb{1}_{\{X_1=k-1,X_0=k\}}\right]\right) \\
&= \frac{\mathbb{P}(X_1 = k+1 \text{ and } X_0 = k)}{\mathbb{P}(X_0 = k)}\,\mathbb{E}[T_{0,S} \mid X_1 = k+1, X_0 = k] \\
&\quad + \frac{\mathbb{P}(X_1 = k-1 \text{ and } X_0 = k)}{\mathbb{P}(X_0 = k)}\,\mathbb{E}[T_{0,S} \mid X_1 = k-1, X_0 = k] \\
&= \mathbb{P}(X_1 = k+1 \mid X_0 = k)\,\mathbb{E}[T_{0,S} \mid X_1 = k+1, X_0 = k] \\
&\quad + \mathbb{P}(X_1 = k-1 \mid X_0 = k)\,\mathbb{E}[T_{0,S} \mid X_1 = k-1, X_0 = k] \\
&= p\,\mathbb{E}[T_{0,S} \mid X_1 = k+1, X_0 = k] + q\,\mathbb{E}[T_{0,S} \mid X_1 = k-1, X_0 = k] \qquad (2.3.4) \\
&= p\,\mathbb{E}[T_{0,S} + 1 \mid X_0 = k+1, X_{-1} = k] + q\,\mathbb{E}[T_{0,S} + 1 \mid X_0 = k-1, X_{-1} = k] \qquad (2.3.5) \\
&= p\,\mathbb{E}[T_{0,S} + 1 \mid X_0 = k+1] + q\,\mathbb{E}[T_{0,S} + 1 \mid X_0 = k-1] \\
&= p(1 + \mathbb{E}[T_{0,S} \mid X_0 = k+1]) + q(1 + \mathbb{E}[T_{0,S} \mid X_0 = k-1]) \\
&= p + q + p\,\mathbb{E}[T_{0,S} \mid X_0 = k+1] + q\,\mathbb{E}[T_{0,S} \mid X_0 = k-1] \\
&= 1 + p\,\mathbb{E}[T_{0,S} \mid X_0 = k+1] + q\,\mathbb{E}[T_{0,S} \mid X_0 = k-1].
\end{aligned}$$

From (2.3.4) to (2.3.5) we relabelled X_1 as X_0, which amounts to changing $T_{0,S} - 1$ into $T_{0,S}$, or equivalently changing $T_{0,S}$ into $T_{0,S} + 1$. □

In the case $S = 3$, Lemma 2.4 shows that

$$\begin{cases} h_3(0) = \mathbb{E}[T_{0,S} \mid X_0 = 0] = 0, \\ h_3(1) = 1 + ph_3(2) + qh_3(0) = 1 + ph_3(2) = 1 + p(1 + qh_3(1)) = 1 + p + pqh_3(1), \\ h_3(2) = 1 + ph_3(3) + qh_3(1) = 1 + qh_3(1) = 1 + q(1 + ph_3(2)) = 1 + q + pqh_3(2), \\ h_3(3) = \mathbb{E}[T_{0,S} \mid X_0 = 3] = 0, \end{cases}$$

which can be solved to recover the result of (2.3.2), (2.3.3).

More generally, defining the function $h_S : \{0, 1, \ldots, S\} \longrightarrow \mathbb{R}_+$ by

$$h_S(k) := \mathbb{E}[T_{0,S} \mid X_0 = k], \qquad k = 0, 1, \ldots, S,$$

Lemma 2.4 shows that

$$h_S(k) = p(1 + h_S(k+1)) + q(1 + h_S(k-1)) = 1 + ph_S(k+1) + qh_S(k-1),$$

$k = 1, 2, \ldots, S-1$, i.e. we have to solve the equation

$$\boxed{\begin{cases} h(k) = 1 + ph(k+1) + qh(k-1), \qquad k = 1, 2, \ldots, S-1, \\ h(0) = h(S) = 0, \end{cases} \qquad (2.3.6)}$$

for the function $h(k)$. Using the fact that $p + q = 1$, we can rewrite (2.3.6) as

$$(p+q)h(k) = 1 + ph(k+1) + qh(k-1), \qquad k = 1, 2, \ldots, S-1,$$

or as the *difference equation*

$$p(h(k+1) - h(k)) - q(h(k) - h(k-1)) = -1, \qquad k = 1, 2, \ldots, S-1, \tag{2.3.7}$$

under the *boundary conditions* (2.3.1a) and (2.3.1b).

The equation

$$p(f(k+1) - f(k)) - q(f(k) - f(k-1)) = 0, \qquad k = 1, 2, \ldots, S-1, \tag{2.3.8}$$

cf. (2.2.22), is called the *homogeneous equation* associated to (2.3.7).

We will use the following fact:

> The general solution to (2.3.7) can be written as the sum of a *homogeneous solution* of (2.3.8) and a *particular solution* of (2.3.7).

***Non-symmetric Case*: $p \neq q$**

By (2.2.16) we know that the *homogeneous solution* of (2.3.8) is of the form $C_1 + C_2 r^k$. Next, searching for a *particular solution* of (2.3.7) of the form $k \longmapsto Ck$ shows that C has to be equal to $C = 1/(q-p)$. Therefore, when $r = q/p \neq 1$, the general solution of (2.3.7) has the form

$$h_S(k) = C_1 + C_2 r^k + \frac{1}{q-p} k. \tag{2.3.9}$$

From the boundary conditions (2.3.1a) and (2.3.1b) and from (2.3.9) we have

$$\begin{cases} h_S(0) = 0 = C_1 + C_2, \\ \\ h_S(S) = 0 = C_1 + C_2 r^S + \dfrac{1}{q-p} S, \end{cases} \tag{2.3.10}$$

and solving the system (2.3.10) of two equations we find

$$C_1 = -\frac{S}{(q-p)(1-r^S)} \quad \text{and} \quad C_2 = \frac{S}{(q-p)(1-r^S)},$$

hence from (2.3.9) we get

$$\boxed{\begin{aligned} h_S(k) = \mathbb{E}[T_{0,S} \mid X_0 = k] &= \frac{1}{q-p}\left(k - S\frac{1-(q/p)^k}{1-(q/p)^S}\right) \qquad (2.3.11) \\ &= \frac{1}{q-p}\left(k - S\mathbb{P}(R_B \mid X_0 = k)\right), \quad k = 0, 1, 2, \ldots, S, \end{aligned}}$$

which does satisfy the boundary conditions (2.3.1a) and (2.3.1b). Note that changing k to $S-k$ and p to q does not modify (2.3.11), as it also represents the mean game duration for Player B.

When $p = 1$, i.e. $r = 0$, we can check easily that e.g. $h_S(k) = S - k$, $k = 0, 1, 2, \ldots, S$. On the other hand, when the number S of states becomes large, we find that for all $k \geq 1$,

$$h_\infty(k) := \lim_{S\to\infty} h_S(k) = \lim_{S\to\infty} \mathbb{E}[T_{0,S} \mid X_0 = k] = \begin{cases} \infty & \text{if } q \leq p, \\ \\ \dfrac{k}{q-p} & \text{if } q > p, \end{cases} \tag{2.3.12}$$

with $h_\infty(0) = 0$, cf. also the symmetric case treated below when $p = q = 1/2$. In particular, for $k \geq 1$ we have

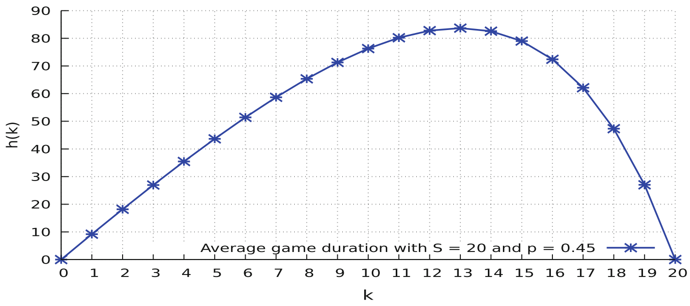


Fig. 2.5 Mean game duration $h_{20}(k)$ as a function of $X_0 = k \in [0, 20]$ for $p = 0.45$

$$\begin{cases} \mathbb{P}(T_0 < \infty \mid X_0 = k) < 1 \quad \text{and} \quad \mathbb{E}[T_0 = \infty \mid X_0 = k] = \infty, \quad p > q, \\ \\ \mathbb{P}(T_0 < \infty \mid X_0 = k) = 1 \quad \text{and} \quad \mathbb{E}[T_0 = \infty \mid X_0 = k] = \infty, \quad p = q = \dfrac{1}{2}, \\ \\ \mathbb{P}(T_0 < \infty \mid X_0 = k) = 1 \quad \text{and} \quad \mathbb{E}[T_0 = \infty \mid X_0 = k] < \infty, \quad p < q. \end{cases}$$

When $r = q/p \leq 1$, this yields an example of a random variable T_0 which is (almost surely[7]) finite, while its expectation is infinite.[8] This situation is similar to that of the St. Petersburg paradox as in (1.6.5).

It is easy to show that (2.3.11) yields $h_2(1) = 1$ when $S = 2$. When $S = 3$, (2.3.11) shows that, using the relation $p + q = 1$,[9]

$$\mathbb{E}[T_{0,3} \mid X_0 = 1] = \frac{1}{q-p}\left(1 - 3\frac{1 - q/p}{1 - (q/p)^3}\right) = \frac{1+p}{1-pq}, \tag{2.3.13}$$

and

$$\mathbb{E}[T_{0,3} \mid X_0 = 2] = \frac{1}{q-p}\left(2 - 3\frac{1 - (q/p)^2}{1 - (q/p)^3}\right) = \frac{1+q}{1-pq}, \tag{2.3.14}$$

however it takes more time to show that (2.3.13) and (2.3.14) agree respectively with (2.3.2) and (2.3.3). In Fig. 2.5 below the mean game duration (2.3.11) is plotted as a function of k for $p = 0.45$.

***Symmetric Case*: $p = q = 1/2$**

In this case (fair game) the homogeneous solution of (2.3.8) is $C_1 + C_2 k$, given by (2.2.19).

---

[7]"Almost surely" means "with probability 1".

[8]Recall that an infinite set of finite data values may have an infinite average.

[9]This point is left as exercise.

Since $r = 1$ we see that $k \longmapsto Ck$ can no longer be a particular solution of (2.3.7). However we can search for a particular solution of the form $k \longmapsto Ck^2$, in which case we find that C has to be equal to $C = -1$.

Therefore when $r = q/p = 1$ the general solution of (2.3.7) has the form

$$h_S(k) = C_1 + C_2 k - k^2, \qquad k = 0, 1, 2, \ldots, S. \tag{2.3.15}$$

From the boundary conditions (2.3.1a) and (2.3.1b) and from (2.3.15) we have

$$\begin{cases} h_S(0) = 0 = C_1, \\ \\ h_S(S) = 0 = C_1 + C_2 S - S^2, \end{cases} \tag{2.3.16}$$

and solving the above system (2.3.16) of two equations yields

$$C_1 = 0 \quad \text{and} \quad C_2 = S,$$

hence from (2.3.15) we get

$$\boxed{h_S(k) = \mathbb{E}[T_{0,S} \mid X_0 = k] = k(S-k), \qquad k = 0, 1, 2, \ldots, S, \tag{2.3.17}}$$

which does satisfy the boundary conditions (2.3.1a) and (2.3.1b) and coincides with (2.3.12) when S goes to infinity.

We note that for all values of p the expectation $\mathbb{E}[T_{0,S} \mid X_0 = k]$ has a *finite* value, which shows that the game duration $T_{0,S}$ is finite with probability one for all $k = 0, 1, \ldots, S$, i.e. $\mathbb{P}(T_{0,S} = \infty \mid X_0 = k) = 0$ for all $k = 0, 1, \ldots, S$.

Remark 2.5 When $r = 1$, by the same argument as in (2.2.26) we find that (2.3.7) is a discretization of the continuous Laplace equation

$$\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x) = -1, \qquad x \in \mathbb{R},$$

which has for solution

$$f(x) = f(0) + x f'(0) - x^2, \qquad x \in \mathbb{R}.$$

Note that (2.3.17) can also be recovered from (2.3.11) by letting p go to $1/2$. In the next Fig. 2.6 the expected game duration (2.3.11) is plotted as a function of p for $S = 20$ and $k = 10$.

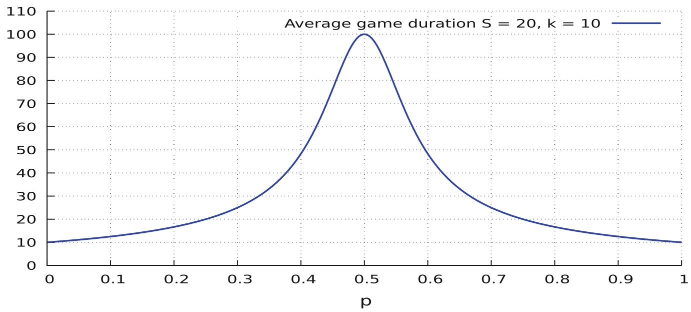


Fig. 2.6 Mean game duration as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$

As expected, the duration will be maximal in a fair game for $p = q = 1/2$. On the other hand, it always takes exactly $10 = S - k = k$ steps to end the game in case $p = 0$ or $p = 1$, in which case there is no randomness. When $p = 0.45$ the expected duration of the game becomes 76.3, which represents only a drop of 24% from the "fair" value 100, as opposed to the 73% drop noticed above in terms of winning probabilities. Thus, a game with $p = 0.45$ is only slightly shorter than a fair game, whereas the probability of winning the game drops down to 0.12.

Remark 2.6 In this chapter we have noticed an interesting connection between analysis and probability. That is, a probabilistic quantity such as $k \longmapsto \mathbb{P}(R_A \mid X_0 = k)$ or $k \longmapsto \mathbb{E}[T_{0,S} \mid X_0 = k]$ can be shown to satisfy a difference equation which is solved by analytic methods. This fact actually extends beyond the present simple framework, and in continuous time it yields other connections between probability and partial differential equations.

In the next chapter we will consider a family of simple random walks which can be seen as "unrestricted" gambling processes.

Exercises

Exercise 2.1 We consider a gambling problem with the possibility of a draw,[10] i.e. at time n the gain X_n of Player A can increase by one unit with probability $r \in (0, 1/2]$, decrease by one unit with probability r, or remain stable with probability $1 - 2r$. We let

$$f(k) := \mathbb{P}(R_A \mid X_0 = k)$$

denote the probability of ruin of Player A, and let

$$h(k) := \mathbb{E}[T_{0,S} \mid X_0 = k]$$

[10] Also called "lazy random walk".

denote the expectation of the game duration $T_{0,S}$ starting from $X_0 = k$, $k = 0, 1, \ldots, S$.

(a) Using first step analysis, write down the difference equation satisfied by $f(k)$ and its boundary conditions, $k = 0, 1, \ldots, S$. We refer to this equation as the *homogeneous equation*.
(b) Solve the homogeneous equation of Question (a) by your preferred method. Is this solution compatible with your intuition of the problem? Why?
(c) Using first step analysis, write down the difference equation satisfied by $h(k)$ and its boundary conditions, $k = 0, 1, \ldots, S$.
(d) Find a particular solution of the equation of Question (c).
(e) Solve the equation of Question (c).
Hint: recall that the general solution of the equation is the sum of a particular solution and a solution of the homogeneous equation.
(f) How does the mean duration $h(k)$ behave as r goes to zero? Is this solution compatible with your intuition of the problem? Why?

Exercise 2.2 Recall that for any standard gambling process $(Z_k)_{k\in\mathbb{N}}$ on a state space $\{a, a+1, \ldots, b-1, b\}$ with absorption at states ⓐ and ⓑ and probabilities $p \neq q$ of moving by ± 1, the probability of hitting state ⓐ before hitting state ⓑ after starting from state $Z_0 = k \in \{a, a+1, \ldots, b-1, b\}$ is given by

$$\frac{1-(p/q)^{b-k}}{1-(p/q)^{b-a}}. \tag{2.3.18}$$

In questions (a), (b), (c) below we consider a gambling process $(X_k)_{k\in\mathbb{N}}$ on the state space $\{0, 1, \ldots, S\}$ with absorption at ⓪ and Ⓢ and probabilities $p \neq q$ of moving by ± 1.

(a) Using Relation (2.3.18), give the probability of coming back in finite time to a given state $m \in \{1, 2, \ldots, S-1\}$ after starting from $X_0 = k \in \{m+1, \ldots, S\}$.
(b) Using Relation (2.3.18), give the probability of coming back in finite time to the given state $m \in \{1, 2, \ldots, S-1\}$ after starting from $X_0 = k \in \{0, 1, , \ldots, m-1\}$.
(c) Using first step analysis, give the probability of coming back to state ⓜ in finite time after starting from $X_0 = m$.
(d) Using first step analysis, compute the mean time to either come back to m of reach any of the two boundaries $\{0, S\}$, whichever comes first?
(e) Repeat the above questions (c), (d) with equal probabilities $p = q = 1/2$, in which case the probability of hitting state ⓐ before hitting state ⓑ after starting from state $Z_0 = k$ is given by

$$\frac{b-k}{b-a}, \qquad k = a, a+1, \ldots, b-1, b. \tag{2.3.19}$$

Exercise 2.3 Consider a gambling process $(X_n)_{n\in\mathbb{N}}$ on the state space $\mathbb{S} = \{0, 1, \ldots, S\}$, with probability p, resp. q, of moving up, resp. down, at each time step. For $k = 0, 1, \ldots, S$, let τ_k denote the first hitting time

$$\tau_k := \inf\{n \geq 0 \ : X_n = k\}.$$

of state k by the process $(X_n)_{n\in\mathbb{N}}$, and let

$$p_k := \mathbb{P}(\tau_{k+1} < \tau_0 \mid X_0 = k), \qquad k = 0, 1, \ldots, S-1,$$

denote the probability of hitting state $\boxed{k+1}$ before hitting state 0.

(a) Show that $p_k = \mathbb{P}(\tau_{k+1} < \tau_0 \mid X_0 = k)$ satisfies the recurrence equation

$$p_k = p + qp_{k-1}p_k, \qquad k = 1, 2, \ldots, S-1, \tag{2.3.20}$$

i.e.

$$p_k = \frac{p}{1 - qp_{k-1}}, \qquad k = 1, 2, \ldots, S-1.$$

(b) Check by substitution that the solution of (2.3.20) is given by

$$p_k = \frac{1-(q/p)^k}{1-(q/p)^{k+1}}, \qquad k = 0, 1, \ldots, S-1. \tag{2.3.21}$$

(c) Compute $\mathbb{P}(\tau_S < \tau_0 \mid X_0 = k)$ by a product formula and recover (2.2.11) and (2.2.27) based on the result of part (2.3.21).

(d) Show that (2.2.12) and (2.2.28) can be recovered in a similar way in the symmetric case $p = q = 1/2$ by trying the solution $p_k = k/(k+1)$, $k = 0, 1, \ldots, S-1$.

Exercise 2.4 Consider a gambling process $(X_n)_{n\in\mathbb{N}}$ on the state space $\{0, 1, 2\}$, with transition probabilities given by

$$\begin{bmatrix} \mathbb{P}(X_1 = 0 \mid X_0 = 0) & \mathbb{P}(X_1 = 1 \mid X_0 = 0) & \mathbb{P}(X_1 = 2 \mid X_0 = 0) \\ \mathbb{P}(X_1 = 0 \mid X_0 = 1) & \mathbb{P}(X_1 = 1 \mid X_0 = 1) & \mathbb{P}(X_1 = 2 \mid X_0 = 1) \\ \mathbb{P}(X_1 = 0 \mid X_0 = 2) & \mathbb{P}(X_1 = 1 \mid X_0 = 2) & \mathbb{P}(X_1 = 2 \mid X_0 = 2) \end{bmatrix}$$

$$= \begin{array}{c c} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{bmatrix} q & p & 0 \\ q & 0 & p \\ 0 & 0 & 1 \end{bmatrix} \end{array},$$

where $0 < p < 1$ and $q = 1 - p$. In this game, Player A is allowed to "rebound" from state ⓪ to state ① with probability p, and state ② is absorbing.

In order to be ruined, Player A has to visit state ⓪ *twice*. Let

$$f(k) := \mathbb{P}(R_A \mid X_0 = k), \qquad k = 0, 1, 2,$$

denote the probability of ruin of Player A starting from $k = 0, 1, 2$. Starting from ⓪ counts as one visit to ⓪.

(a) Compute the boundary condition $f(0)$ using pathwise analysis.
(b) Give the value of the boundary condition $f(2)$, and compute $f(1)$ by first step analysis.

Exercise 2.5 (a) Recover (2.3.17) from (2.3.11) by letting p go to $1/2$, i.e. when $r = q/p$ goes to 1.
(b) Recover (2.2.21) from (2.2.11) by letting p go to $1/2$, i.e. when $r = q/p$ goes to 1.

Exercise 2.6 Extend the setting of Exercise 2.1 to a non-symmetric gambling process with draw and respective probabilities $\alpha > 0$, $\beta > 0$, and $1 - \alpha - \beta > 0$ of increase, decrease, and draw. Compute the ruin probability $f(k)$ and the mean game duration $h(k)$ in this extended framework. Check that when $\alpha = \beta \in (0, 1/2)$ we recover the result of Exercise 2.1.

Problem 2.7 We consider a discrete-time process $(X_n)_{n\geq 0}$ that models the wealth of a gambler within $\{0, 1, \ldots, S\}$, with the transition probabilities

$$\begin{cases} \mathbb{P}(X_{n+1} = k + 1 \mid X_n = k) = p, & k = 0, 1, \ldots, S - 1, \\ \\ \mathbb{P}(X_{n+1} = k - 1 \mid X_n = k) = q, & k = 1, 2, \ldots, S, \end{cases}$$

and

$$\mathbb{P}(X_{n+1} = 0 \mid X_n = 0) = q,$$

for all $n \in \mathbb{N} = \{0, 1, 2, \ldots\}$, where $q = 1 - p$ and $p \in (0, 1]$. In this model the gambler is given a second chance, and may be allowed to "rebound" after reaching 0. Let

$$W = \bigcup_{n\in\mathbb{N}} \{X_n = S\}$$

denote the event that the player eventually wins the game.

(a) Let

$$g(k) := \mathbb{P}(W \mid X_0 = k)$$

denote the probability that the player eventually wins after starting from state $k \in \{0, 1, \ldots, S\}$. Using first step analysis, write down the difference equations satisfied by $g(k)$, $k = 0, 1, \ldots, S-1$, and their boundary condition(s), which may not be given in explicit form. This question is standard, however one has to pay attention to the special behavior of the process at state 0.

(b) Obtain $\mathbb{P}(W \mid X_0 = k)$ for all $k = 0, 1, \ldots, S$ as the unique solution to the system of equations stated in Question (a). The answer to this question is very simple and can be obtained through intuition. However, a (mathematical) proof is required.

(c) Let

$$T_S = \inf\{n \geq 0 \ : \ X_n = S\}$$

denote the first hitting time of S by the process $(X_n)_{n\geq 0}$. Let

$$h(k) := \mathbb{E}[T_S \mid X_0 = k]$$

denote the expected time until the gambler wins after starting from state $k \in \{0, 1, \ldots, S\}$. Using first step analysis, write down the difference equations satisfied by $h(k)$ for $k = 0, 1, \ldots, S-1$, and state the corresponding boundary condition(s). Again, one has to pay attention to the special behavior of the process at state 0, as the equation obtained by first step analysis for $h(0)$ will take a particular form and can be viewed as a second boundary condition.

(d) Compute $\mathbb{E}[T_S \mid X_0 = k]$ for all $k = 0, 1, \ldots, S$ by solving the equations of Question (c).

This question is more difficult than Question (b), and it could be skipped at first reading since its result is not used in the sequel. One can solve the homogeneous equation for $k = 1, 2, \ldots, S-1$ using the results of Sect. 2.3, and a particular solution can be found by observing that here we consider the time until Player A (not B) wins. As usual, the cases $p \neq q$ and $p = q = 1/2$ have to be considered separately at some point. The formula obtained for $p = 1$ should be quite intuitive and may help you check your result.

(e) Let now

$$T_0 = \inf\{n \geq 0 \ : \ X_n = 0\}$$

denote the first hitting time of 0 by the process $(X_n)_{n\geq 0}$. Using the results of Sect. 2.2 for the ruin of Player B, write down the value of

$$p_k := \mathbb{P}(T_S < T_0 \mid X_0 = k)$$

as a function of p, S, and $k = 0, 1, \ldots, S$.

Note that according to the notation of this chapter, $\{T_S < T_0\}$ denotes the event "Player A wins the game".

(f) Explain why the equality

$$\begin{aligned}\mathbb{P}(T_S < T_0 \mid X_1 = k+1 \text{ and } X_0 = k) &= \mathbb{P}(T_S < T_0 \mid X_1 = k+1) \\ &= \mathbb{P}(T_S < T_0 \mid X_0 = k+1). \qquad (2.3.22)\end{aligned}$$

holds for $k \in \{0, 1, \ldots, S-1\}$ (an explanation in words will be sufficient here).

(g) Using Relation (2.3.22), show that the probability

$$\mathbb{P}(X_1 = k+1 \mid X_0 = k \text{ and } T_S < T_0)$$

of an upward step given that state S is reached first, is equal to

$$\mathbb{P}(X_1 = k+1 \mid X_0 = k \text{ and } T_S < T_0) = p\frac{\mathbb{P}(T_S < T_0 \mid X_0 = k+1)}{\mathbb{P}(T_S < T_0 \mid X_0 = k)} = p\frac{p_{k+1}}{p_k}, \qquad (2.3.23)$$

$k = 1, 2, \ldots, S-1$, to be computed explicitly from the result of Question (e). How does this probability compare to the value of p?
No particular difficulty here, the proof should be a straightforward application of the definition of conditional probabilities.

(h) Compute the probability

$$\mathbb{P}(X_1 = k-1 \mid X_0 = k \text{ and } T_0 < T_S), \qquad k = 1, 2, \ldots, S,$$

of a downward step given that state 0 is reached first, using similar arguments to Question (g).

(i) Let

$$h(k) = \mathbb{E}[T_S \mid X_0 = k,\ T_S < T_0], \qquad k = 1, 2, \ldots, S,$$

denote the expected time until the player wins, given that state 0 is never reached. Using the transition probabilities (2.3.23), state the finite difference equations satisfied by $h(k)$, $k = 1, 2, \ldots, S-1$, and their boundary condition(s).
The derivation of the equation is standard, but you have to make a careful use of conditional transition probabilities given $\{T_S < T_0\}$. There is an issue on whether and how $h(0)$ should appear in the system of equations, but this point can actually be solved.

(j) Solve the equation of Question (i) when $p = 1/2$ and compute $h(k)$ for $k = 1, 2, \ldots, S$. What can be said of $h(0)$?

There is actually a way to transform this equation using an homogeneous equation already solved in Sect. 2.3.

Problem 2.8 Let $S \geq 1$. We consider a discrete-time process $(X_n)_{n\geq 0}$ that models the wealth of a gambler within $\{0, 1, \ldots, S\}$, with the transition probabilities

$$\mathbb{P}(X_{n+1} = k+2 \mid X_n = k) = p, \quad \mathbb{P}(X_{n+1} = k-1 \mid X_n = k) = 2p,$$

and

$$\mathbb{P}(X_{n+1} = k \mid X_n = k) = r, \quad k \in z,$$

for all $n \in \mathbb{N} = \{0, 1, 2, \ldots\}$, where $p > 0$, $r \geq 0$, and $3p + r = 1$. We let

$$\tau := \inf\{n \geq 0 \ : \ X_n \leq 0 \text{ or } X_n \geq S\}.$$

(a) Consider the probability

$$g(k) := \mathbb{P}(X_\tau \geq S \mid X_0 = k)$$

that the game ends with Player A winning the game, starting from $X_0 = k$. Give the values of $g(0)$, $g(S)$ and $g(S+1)$.

(b) Using first step analysis, write down the difference equation satisfied by $g(k)$, $k = 1, 2, \ldots, S-1$, and its boundary conditions, by taking *overshoot* into account. We refer to this equation as the *homogeneous equation*.

(c) Solve the equation of Question (b) from its characteristic equation as in (2.2.15).

(d) Does the answer to Question (c) depend on p? Why?

(e) Consider the expected time

$$h(k) := \mathbb{E}[\tau \mid X_0 = k], \qquad k = 0, 1, \ldots, S+1,$$

spent until the end of the game. Give the values of $h(0)$, $h(S)$ and $h(S+1)$.

(f) Using first step analysis, write down the difference equation satisfied by $h(k)$, $k = 1, 2, \ldots, S-1$, and its boundary conditions.

(g) Find a particular solution of the equation of Question (e)

(h) Solve the equation of Question (2.1 c).

Hint: the general solution of the equation is the sum of a particular solution and a solution of the homogeneous equation.

(i) How does the mean duration $h(k)$ behave as p goes to zero? Is this compatible with your intuition of the problem? Why?

(j) How do the values of $g(k)$ and $h(k)$ behave for fixed $k \in \{1, 2, \ldots, S-1\}$ as S tends to infinity?

Problem 2.9 Consider a gambling process $(X_n)_{n \geq 0}$ on the state space $\mathbb{S} = \{0, 1, \ldots, S\}$, with transition probabilities

$$\mathbb{P}(X_{n+1} = k+1 \mid X_n = k) = p, \qquad \mathbb{P}(X_{n+1} = k-1 \mid X_n = k) = q,$$

$k = 1, 2, \ldots, S-1$, with $p + q = 1$. Let

$$\tau := \inf\{n \geq 0 \ : \ X_n = 0 \text{ or } X_n = S\}$$

denote the time until the process hits either state 0 or state S, and consider the second moment

$$h(k) := \mathbb{E}\left[\tau^2 \mid X_0 = k\right],$$

of τ after starting from $k = 0, 1, 2, \ldots, S$.

(a) Give the values of $h(0)$ and $h(S)$.
(b) Using first step analysis, find an equation satisfied by $h(k)$ and involving $\mathbb{E}[\tau \mid X_0 = k+1]$ and $\mathbb{E}[\tau \mid X_0 = k-1]$, $k = 1, 2, \ldots, S-1$.
(c) From now on we take $\underline{p = q = 1/2}$. Recall that in this case we have

$$\mathbb{E}[\tau \mid X_0 = k] = (S-k)k, \qquad k = 0, 1, \ldots, S.$$

Show that the function $h(k)$ satisfies the finite difference equation

$$h(k) = -1 + 2(S-k)k + \frac{1}{2}h(k+1) + \frac{1}{2}h(k-1), \qquad k = 1, 2, \ldots, S-1. \tag{2.3.24}$$

(d) Knowing that

$$k \longmapsto \frac{2}{3}k^2 - \frac{2S}{3}k^3 + \frac{k^4}{3}$$

is a particular solution of the equation (2.3.24) of Question (c), and that the solution of the homogeneous equation

$$f(k) = \frac{1}{2}f(k+1) + \frac{1}{2}f(k-1), \qquad k = 1, 2, \ldots, S-1,$$

takes the form

$$f(k) = C_1 + C_2 k,$$

compute the value of the expectation $h(k)$ solution of (2.3.24) for all $k = 0, 1, \ldots, S$.
(e) Compute the variance

$$v(k) = \mathbb{E}\left[\tau^2 \mid X_0 = k\right] - \left(\mathbb{E}[\tau \mid X_0 = k]\right)^2$$

of the game duration starting from $k = 0, 1, \ldots, S$.
(f) Compute $v(1)$ when $\underline{S = 2}$ and explain why the result makes pathwise sense.

Chapter 3
Random Walks

In this chapter we consider our second important example of discrete-time stochastic process, which is a random walk allowed to evolve over the set $\mathbb{Z}$ of signed integers without any boundary restriction. Of particular importance are the probabilities of return to a given state in finite time, as well as the corresponding mean return time.

3.1 Unrestricted Random Walk

The simple unrestricted random walk $(S_n)_{n\geq 0}$, also called *Bernoulli* random walk, is defined by $S_0 = 0$ and

$$S_n = \sum_{k=1}^{n} X_k = X_1 + \cdots + X_n, \qquad n \geq 1,$$

where the random walk *increments* $(X_k)_{k\geq 1}$ form a family of independent, $\{-1, +1\}$-valued random variables.

We will assume in addition that the family $(X_k)_{k\geq 1}$ is *i.i.d.*, i.e. it is made of *independent and identically distributed* Bernoulli random variables, with distribution

$$\begin{cases} \mathbb{P}(X_k = +1) = p, \\ \\ \mathbb{P}(X_k = -1) = q, \qquad k \geq 1, \end{cases}$$

with $p + q = 1$.

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4_3

3.2 Mean and Variance

In this case the mean and variance of X_n are given by

$$\mathbb{E}[X_n] = -1 \times q + 1 \times p = 2p - 1 = p - q,$$

and

$$\begin{aligned}\operatorname{Var}[X_n] &= \mathbb{E}[X_n^2] - (\mathbb{E}[X_n])^2 \\ &= 1 \times q + 1 \times p - (2p-1)^2 \\ &= 4p(1-p) = 4pq.\end{aligned}$$

As a consequence, we find that

$$\mathbb{E}[S_n \mid S_0 = 0] = \mathbb{E}\left[\sum_{k=1}^{n} X_k\right] = \sum_{k=1}^{n} \mathbb{E}[X_k] = n(2p-1) = n(p-q),$$

and the variance can be computed by (1.6.12) as

$$\operatorname{Var}[S_n \mid S_0 = 0] = \operatorname{Var}\left[\sum_{k=1}^{n} X_k\right] = \sum_{k=1}^{n} \operatorname{Var}[X_k] = 4npq,$$

where we used (1.6.12).

3.3 Distribution

First we note that in an even number of time steps, $(S_n)_{n\in\mathbb{N}}$ can only reach an even state in $\mathbb{Z}$ starting from 0. Similarly, in an odd number of time steps, $(S_n)_{n\in\mathbb{N}}$ can only reach an odd state in $\mathbb{Z}$ starting from 0. Indeed, starting from $S_n = k$ the value of S_{n+2} after two time steps can only belong to $\{k-2, k, k+2\}$. Consequently, we have

$$\begin{cases} \mathbb{P}(S_{2n} = 2k+1 \mid S_0 = 0) = 0, & k \in \mathbb{Z},\ n \in \mathbb{N}, \\ \\ \mathbb{P}(S_{2n+1} = 2k \mid S_0 = 0) = 0, & k \in \mathbb{Z},\ n \in \mathbb{N}, \end{cases} \tag{3.3.1}$$

and

$$\mathbb{P}(S_n = k \mid S_0 = 0) = 0, \qquad \text{for } k < -n \text{ or } k > n, \tag{3.3.2}$$

since $S_0 = 0$. Next, let l denote the number of upwards steps between time 0 and time $2n$, whereas $2n - l$ will denote the number of downwards steps. If $S_{2n} = 2k$ we have

$$2k = l - (2n - l) = 2l - 2n,$$

hence there are $l = n + k$ upwards steps and $2n - l = n - k$ downwards steps, $-n \leq k \leq n$. The probability of a given paths having $l = n + k$ upwards steps and $2n - l = n - k$ downwards steps is

$$p^{n+k} q^{n-k}$$

and in order to find $\mathbb{P}(S_{2n} = 2k \mid S_0 = 0)$ we need to multiply this probability by the total number of paths leading from 0 to $\boxed{2k}$ in $2n$ steps. We find that this number of paths is

$$\binom{2n}{n+k} = \binom{2n}{n-k}$$

which represents the number of ways to arrange $n + k$ upwards steps (or $n - k$ downwards steps) within $2n$ time steps.
Hence we have

$$\boxed{\mathbb{P}(S_{2n} = 2k \mid S_0 = 0) = \binom{2n}{n+k} p^{n+k} q^{n-k}, \quad -n \leq k \leq n,} \qquad (3.3.3)$$

in addition to (3.3.1) and (3.3.2). Figure 3.1 shows one of the $120 = \binom{10}{7} = \binom{10}{2}$ possible paths corresponding to $n = 5$ and $k = 2$.

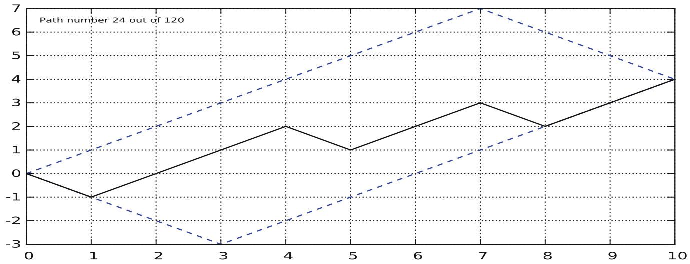


Fig. 3.1 Graph of $120 = \binom{10}{7} = \binom{10}{3}$ paths with $n = 5$ and $k = 2$

Exercises:

(i) Show by a similar analysis that

$$\mathbb{P}(S_{2n+1} = 2k+1 \mid S_0 = 0) = \binom{2n+1}{n+k+1} p^{n+k+1} q^{n-k}, \quad -n \leq k \leq n, \tag{3.3.4}$$

i.e. $(2n+1+S_{2n+1})/2$ is a binomial random variable with parameter $(2n+1, p)$, and

$$\begin{aligned}\mathbb{P}\left(\frac{2n+1+S_{2n+1}}{2} = k \,\Big|\, S_0 = 0\right) &= \mathbb{P}\left(S_{2n+1} = 2k-2n-1 \,\Big|\, S_0 = 0\right) \\ &= \binom{2n+1}{k} p^k q^{2n+1-k},\end{aligned}$$

$k = 0, 1, \ldots, 2n+1$.

(ii) Show that $n + S_{2n}/2$ is a binomial[1] random variable with parameter $(2n, p)$, i.e., show that

$$\begin{aligned}\mathbb{P}\left(n + \frac{S_{2n}}{2} = k \,\Big|\, S_0 = 0\right) &= \mathbb{P}\left(S_{2n} = 2k - 2n \,\Big|\, S_0 = 0\right) \\ &= \binom{2n}{k} p^k q^{2n-k}, \qquad k = 0, 1, \ldots, 2n.\end{aligned}$$

3.4 First Return to Zero

Let

$$T_0^r := \inf\{n \geq 1 \ : \ S_n = 0\}$$

denote the time of first return to (0) of the random walk started at (0), with the convention $\inf \emptyset = \infty$.[2] We are interested in particular in computing the mean time $\mathbb{E}[T_0^r \mid S_0 = 0]$ it takes to return to state (0) after starting from state (0) (see Fig. 3.2).

We are interested in computing the distribution

$$g(n) = \mathbb{P}(T_0^r = n \mid S_0 = 0), \qquad n \geq 1,$$

of the first return time T_0^r to (0). It is easy to show by pathwise analysis that T_0^r can only be even-valued starting from (0), hence $g(2k+1) = 0$ for all $k \in \mathbb{N}$, and in particular we have

[1] Note that S_{2n} is always an even number after we start from $S_0 = 0$.

[2] Recall that the notation "inf" stands for "infimum", meaning the smallest $n \geq 0$ such that $S_n = 0$, with $T_0^r = \infty$ if no such $n \geq 0$ exists.

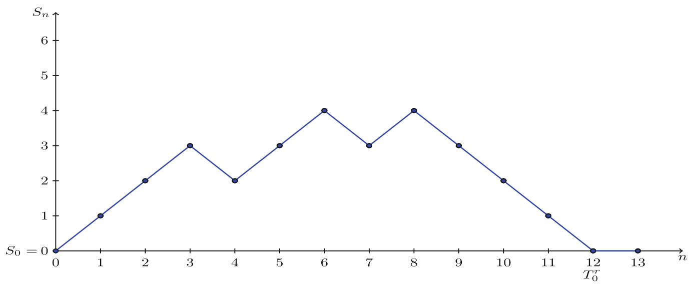


Fig. 3.2 Sample path of the random walk $(S_n)_{n\in\mathbb{N}}$

$$\mathbb{P}(T_0^r = 1 \mid S_0 = 0) = 0, \quad \mathbb{P}(T_0^r = 2 \mid S_0 = 0) = 2pq, \tag{3.4.1}$$

and

$$\mathbb{P}(T_0^r = 4 \mid S_0 = 0) = 2p^2q^2, \tag{3.4.2}$$

by considering the two paths leading from (0) to (0) in two steps and the only two paths leading from (0) to (0) in four steps without hitting (0). However the computation of $\mathbb{P}(T_0^r = 2n \mid S_0 = 0)$ by this method is difficult to extend to all $n \geq 3$.

In order to completely solve this problem we will rely on the computation of the probability generating function $G_{T_0^r}$ of T_0^r, cf. (3.4.9) below.

This computation will use the following tools:

- convolution equation, see Relation (3.4.3) below,
- Taylor expansions, see Relation (3.4.22) below,
- probability generating functions.

First, we will need the following Lemma 3.1 which will be used in the proof of Lemma 3.3 below.

Lemma 3.1 (Convolution equation). *The function*

$$\begin{aligned} g : \{1, 2, 3, \ldots\} &\longrightarrow [0, 1] \\ n &\longmapsto g(n) \end{aligned}$$

defined by

$$g(n) := \mathbb{P}(T_0^r = n \mid S_0 = 0), \qquad n \geq 1,$$

satisfies the convolution equation

$$h(n) = \sum_{k=0}^{n-2} g(n-k)h(k), \qquad n \geq 1, \tag{3.4.3}$$

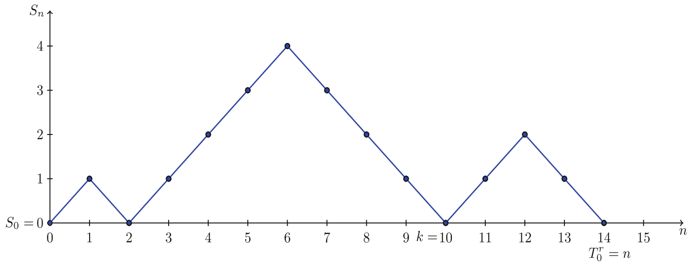


Fig. 3.3 Last return to state 0 at time $k = 10$

with the initial condition $g(1) = 0$*, where* $h(n) := \mathbb{P}(S_n = 0 \mid S_0 = 0)$ *is given from (3.3.3) by*

$$h(2n) = \binom{2n}{n} p^n q^n, \quad \textit{and} \quad h(2n+1) = 0, \qquad n \in \mathbb{N}. \tag{3.4.4}$$

Proof We first partition the event $\{S_n = 0\}$ into

$$\{S_n = 0\} = \bigcup_{k=0}^{n-2} \{S_k = 0,\ S_{k+1} \neq 0, \ldots, S_{n-1} \neq 0,\ S_n = 0\}, \qquad n \geq 1,$$

according to all possible times $k = 0, 1, \ldots, n-2$ of *last* return to state $\textcircled{0}$ *before* time n, with $\{S_1 = 0\} = \emptyset$ since we are starting from $S_0 = 0$ (see Fig. 3.3).

Then we have

$$\begin{aligned}
h(n) &:= \mathbb{P}(S_n = 0 \mid S_0 = 0) \\
&= \sum_{k=0}^{n-2} \mathbb{P}\big(S_k = 0,\ S_{k+1} \neq 0, \ldots, S_{n-1} \neq 0,\ S_n = 0 \mid S_0 = 0\big) \\
&= \sum_{k=0}^{n-2} \mathbb{P}\big(S_{k+1} \neq 0, \ldots, S_{n-1} \neq 0,\ S_n = 0 \mid S_k = 0,\ S_0 = 0\big) \mathbb{P}(S_k = 0 \mid S_0 = 0) \\
&= \sum_{k=0}^{n-2} \mathbb{P}\big(S_{k+1} \neq 0, \ldots, S_{n-1} \neq 0,\ S_n = 0 \mid S_k = 0\big) \mathbb{P}(S_k = 0 \mid S_0 = 0) \qquad (3.4.5) \\
&= \sum_{k=0}^{n-2} \mathbb{P}\big(S_1 \neq 0, \ldots, S_{n-k-1} \neq 0,\ S_{n-k} = 0 \mid S_0 = 0\big) \mathbb{P}(S_k = 0 \mid S_0 = 0) \qquad (3.4.6) \\
&= \sum_{k=0}^{n-2} \mathbb{P}(T_0^r = n - k \mid S_0 = 0) \mathbb{P}(S_k = 0 \mid S_0 = 0) \qquad (3.4.7) \\
&= \sum_{k=0}^{n-2} h(k) g(n-k), \qquad n \geq 1,
\end{aligned}$$

where from (3.4.5) to (3.4.6) we applied a shift of k steps in time, from time $k+1$ to time 1. □

We now need to solve the convolution equation (3.4.3) for $g(n) = \mathbb{P}(T_0^r = n \mid S_0 = 0)$, $n \geq 1$, knowing that $g(1) = 0$. For this we will derive a simple equation for the *probability generating function*

$$\begin{aligned} G_{T_0^r} \ : [-1, 1] &\longrightarrow \mathbb{R} \\ s &\longmapsto G_{T_0^r}(s) \end{aligned}$$

of the random variable T_0^r, defined by

$$G_{T_0^r}(s) := \mathbb{E}[s^{T_0^r} \mathbb{1}_{\{T_0^r < \infty\}} \mid S_0 = 0] = \sum_{n=0}^{\infty} s^n \mathbb{P}(T_0^r = n \mid S_0 = 0) = \sum_{n=0}^{\infty} s^n g(n),$$

$-1 \leq s \leq 1$, cf. (1.7.3).

Recall that the knowledge of $G_{T_0^r}(s)$ provides certain information on the distribution of T_0^r, such as the probability

$$\mathbb{P}(T_0^r < \infty \mid S_0 = 0) = \mathbb{E}[\mathbb{1}_{\{T_0^r < \infty\}} \mid S_0 = 0] = G_{T_0^r}(1)$$

and the expectation

$$\mathbb{E}[T_0^r \mathbb{1}_{\{T_0^r < \infty\}} \mid S_0 = 0] = \sum_{n=1}^{\infty} n \mathbb{P}(T_0^r = n \mid S_0 = 0) = G'_{T_0^r}(1).$$

In Lemma 3.3 below we will compute $G_{T_0^r}(s)$ for all $s \in [-1, 1]$. First, let the function

$$\begin{aligned} H \ : \mathbb{R} &\longrightarrow \mathbb{R} \\ s &\longmapsto H(s) \end{aligned}$$

be defined by

$$H(s) := \sum_{k=0}^{\infty} h(k) s^k = \sum_{k=0}^{\infty} s^k \mathbb{P}(S_k = 0 \mid S_0 = 0), \qquad -1 \leq s \leq 1.$$

In the following lemma we show that the function $H(s)$ can be computed in closed form.

Proposition 3.2 *We have*

$$H(s) = (1 - 4pqs^2)^{-1/2}, \qquad |s| < \frac{1}{2\sqrt{pq}}.$$

Proof By (3.4.4) and the fact that $\mathbb{P}(S_{2k+1} = 0 \mid S_0 = 0) = 0$, $k \in \mathbb{N}$, we have

$$\begin{aligned}
H(s) &= \sum_{k=0}^{\infty} s^k \mathbb{P}(S_k = 0 \mid S_0 = 0) \\
&= \sum_{k=0}^{\infty} s^{2k} \mathbb{P}(S_{2k} = 0 \mid S_0 = 0) = \sum_{k=0}^{\infty} s^{2k} \binom{2k}{k} p^k q^k \\
&= \sum_{k=0}^{\infty} (pq)^k s^{2k} \frac{(2k)(2k-1)(2k-2)(2k-3) \times \cdots \times 4 \times 3 \times 2 \times 1}{(k(k-1) \times \cdots \times 2 \times 1)^2} \\
&= \sum_{k=0}^{\infty} (4pq)^k s^{2k} \frac{k(k-1/2)(k-2/2)(k-3/2) \times \cdots \times (4/2) \times (3/2) \times (2/2) \times (1/2)}{(k(k-1) \times \cdots \times 2 \times 1)^2} \\
&= \sum_{k=0}^{\infty} (4pq)^k s^{2k} \frac{(k-1/2)(k-3/2) \times \cdots \times (3/2) \times (1/2)}{k(k-1) \times \cdots \times 2 \times 1} \\
&= \sum_{k=0}^{\infty} (-1)^k (4pq)^k s^{2k} \frac{(-1/2-(k-1))(3/2-k) \times \cdots \times (-3/2) \times (-1/2)}{k(k-1) \times \cdots \times 2 \times 1} \\
&= \sum_{k=0}^{\infty} (-4pqs^2)^k \frac{(-1/2) \times (-3/2) \times \cdots \times (3/2-k)(-1/2-(k-1))}{k!} \\
&= (1-4pqs^2)^{-1/2},
\end{aligned} \tag{3.4.8}$$

$|4pqs^2| < 1.$[3] □

Remark We note that, taking $s = 1$, by (1.6.1) we have

$$\begin{aligned}
H(1) &= \sum_{k=0}^{\infty} \mathbb{P}(S_k = 0 \mid S_0 = 0) \\
&= \sum_{k=0}^{\infty} \mathbb{E}[\mathbb{1}_{\{S_k=0\}} \mid S_0 = 0] \\
&= \mathbb{E}\left[\sum_{k=0}^{\infty} \mathbb{1}_{\{S_k=0\}} \,\middle|\, S_0 = 0 \right],
\end{aligned}$$

hence $H(1) = 1/\sqrt{1-4pq}$ represents the mean number of visits of the random walk $(S_n)_{n\in\mathbb{N}}$ to state 0.

Next, based on the convolution equation (3.4.3) of Lemma 3.1 we compute $G_{T_0^r}(s)$ in the next Lemma 3.3 by deriving and solving an Eq. (3.4.13) for $G_{T_0^r}(s)$. This method has some similarities with the z-transform method used in electrical engineering.

[3]We used the formula $(1+x)^\alpha = \sum_{k=0}^{\infty} \frac{x^k}{k!} \alpha(\alpha-1) \times \cdots \times (\alpha-(k-1))$, cf. Relation (A.8).

Lemma 3.3 *The probability generating function $G_{T_0^r}$ of the first return time T_0^r to ⓪ is given by*

$$\boxed{G_{T_0^r}(s) = 1 - \frac{1}{H(s)} = 1 - \sqrt{1 - 4pqs^2}, \qquad 4pqs^2 < 1.} \tag{3.4.9}$$

Proof We have, taking into account the relations $g(1) = \mathbb{P}(T_0^r = 1 \mid S_0 = 0) = 0$ and $h(0) = 0$,

$$\begin{aligned}
G_{T_0^r}(s)H(s) &= \left(\sum_{n=1}^{\infty} s^n g(n)\right)\left(\sum_{k=0}^{\infty} s^k h(k)\right) \\
&= \sum_{n=2}^{\infty}\sum_{k=0}^{\infty} s^{n+k} g(n)h(k) = \sum_{k=0}^{\infty}\sum_{n=2}^{\infty} s^{n+k} g(n)h(k) \\
&= \sum_{l=2}^{\infty} s^l \sum_{k=0}^{l-2} g(l-k)h(k) \\
&= \sum_{l=1}^{\infty} s^l h(l) && (3.4.10) \\
&= \sum_{l=1}^{\infty} s^l \mathbb{P}(S_l = 0 \mid S_0 = 0) && (3.4.11) \\
&= -1 + \sum_{l=0}^{\infty} s^l \mathbb{P}(S_l = 0 \mid S_0 = 0) = H(s) - 1, && (3.4.12)
\end{aligned}$$

where from line (3.4.10) to line (3.4.11) we have applied the change of variable $(k, n) \longmapsto (k, l)$ with $l = n + k$, and from line (3.4.11) to line (3.4.12) we have used the convolution equation (3.4.3) of Lemma 3.1. This shows that $G_{T_0^r}(s)$ satisfies the equation

$$\boxed{G_{T_0^r}(s)H(s) = H(s) - 1, \qquad 4pqs^2 < 1.} \tag{3.4.13}$$

Solving (3.4.13) yields the value of $G_{T_0^r}(s)$ for all s such that $4pqs^2 < 1$. □

See Exercise 3.4-(e) for another derivation of (3.4.9) based on first step analysis.[4]

---

[4]"Any good theorem should have several proofs, the more the better. For two reasons: usually, different proofs have different strengths and weaknesses, and they generalise in different directions - they are not just repetitions of each other. Some of them are good for this application, some are

We will apply our knowledge of $G_{T_0^r}(s)$ to the computation of the first return time distribution of T_0^r, the probability of return to (0) in finite time, and the mean return time to (0).

Probability of Return to Zero in Finite Time

The probability that the first return to (0) occurs within a finite time is

$$\begin{aligned}\mathbb{P}(T_0^r < \infty \mid S_0 = 0) &= \mathbb{E}[\mathbb{1}_{\{T_0^r<\infty\}} \mid S_0 = 0] = \mathbb{E}[1^{T_0^r}\mathbb{1}_{\{T_0^r<\infty\}} \mid S_0 = 0] \\ &= G_{T_0^r}(1) = 1 - \sqrt{1-4pq} \\ &= 1 - |2p-1| = 1 - |p-q| = \begin{cases} 2q, & p \geq 1/2, \\ 2p, & p \leq 1/2, \end{cases} \\ &= 2\min(p,q), \end{aligned} \tag{3.4.14}$$

hence

$$\boxed{\mathbb{P}(T_0^r = \infty \mid S_0 = 0) = |2p-1| = |p-q|.} \tag{3.4.15}$$

Note that in (2.2.13) above we have shown that the probability of hitting state (0) in finite time starting from any state (k) with $k \geq 1$ is given by

$$\mathbb{P}(T_0^r < \infty \mid S_0 = k) = \min\left(1, \left(\frac{q}{p}\right)^k\right), \qquad k \geq 1, \tag{3.4.16}$$

i.e.

$$\mathbb{P}(T_0^r = \infty \mid S_0 = k) = \max\left(0, 1 - \left(\frac{q}{p}\right)^k\right), \qquad k \geq 1,$$

cf. also Exercises 3.2-(c) and 3.4-(c).

In the non-symmetric case $p \neq q$, Relation (3.4.14) shows that

$$\mathbb{P}(T_0^r < \infty \mid S_0 = 0) < 1 \qquad \text{and} \qquad \mathbb{P}(T_0^r = \infty \mid S_0 = 0) > 0,$$

whereas in the symmetric case (or fair game) $p = q = 1/2$ we find that

$$\mathbb{P}(T_0^r < \infty \mid S_0 = 0) = 1 \qquad \text{and} \qquad \mathbb{P}(T_0^r = \infty \mid S_0 = 0) = 0,$$

i.e. the random walk returns to (0) with probability one.

See Exercise 3.4-(b) and 5.9-(a) for other derivations of (3.4.16).

---

good for that application. They all shed light on the area. If you cannot look at a problem from different directions, it is probably not very interesting; the more perspectives, the better !" - Sir Michael Atiyah.

Mean Return Time to Zero

(i) In the non-symmetric case $p \neq q$, by (3.4.15), the time T_0^r needed to return to state (0) is infinite with probability

$$\mathbb{P}(T_0^r = \infty \mid S_0 = 0) = |p - q| > 0,$$

hence the expected value[5]

$$\begin{aligned}\mathbb{E}[T_0^r \mid S_0 = 0] &= \infty \times \mathbb{P}(T_0^r = \infty \mid S_0 = 0) + \sum_{k=1}^{\infty} k\mathbb{P}(T_0^r = k \mid S_0 = 0) \\ &= \infty \end{aligned} \tag{3.4.17}$$

is infinite in that case.[6]

Note that starting from $S_0 = k \geq 1$, by (2.3.12) we have found that the mean hitting time of state (0) equals

$$\mathbb{E}[T_0^r \mid S_0 = k] = \begin{cases} \infty & \text{if } q \leq p, \\ \dfrac{k}{q-p} & \text{if } q > p. \end{cases} \tag{3.4.18}$$

In particular we have $\mathbb{P}(T_0^r < \infty \mid S_0 = k) = 1$ when $q > p$ and $k \geq 1$, which is consistent with (3.4.16). See Exercise 5.9-(b) for other derivations of (2.3.12)-(3.4.18) using the probability generating function $s \mapsto G_{T_0^r}(s)$.

Remark By (3.4.9), the truncated expectation $\mathbb{E}[T_0^r \mathbb{1}_{\{T_0^r < \infty\}} \mid S_0 = 0]$ satisfies

$$\begin{aligned}\mathbb{E}[T_0^r \mathbb{1}_{\{T_0^r < \infty\}} \mid S_0 = 0] &= \sum_{n=1}^{\infty} n\mathbb{P}(T_0^r = n \mid S_0 = 0) \\ &= G'_{T_0^r}(1) \\ &= \frac{4pqs}{\sqrt{1 - 4pqs^2}}\bigg|_{s=1} \\ &= \frac{4pq}{\sqrt{1 - 4pq}} \\ &= \frac{4pq}{|p - q|}, \end{aligned} \tag{3.4.19}$$

when $p \neq q$, which shows in particular from Lemma 1.4 that

---

[5]Note that the summation $\sum_{k=1}^{\infty} = \sum_{1 \leq k < \infty}$ actually *excludes* the value $k = \infty$.

[6]We use the convention $\infty \times 0 = 0$.

$$\begin{aligned}\mathbb{E}[T_0^r \mid T_0^r < \infty,\ S_0 = 0] &= \frac{1}{\mathbb{P}(T_0^r < \infty \mid S_0 = 0)} \mathbb{E}[T_0^r \mathbb{1}_{\{T_0^r < \infty\}} \mid S_0 = 0] \\ &= \frac{1}{\min(p, q)} \frac{2pq}{|p - q|} = 2\frac{\max(p, q)}{|p - q|}.\end{aligned}$$

(ii) In the symmetric case $p = q = 1/2$ we have $\mathbb{P}(T_0^r < \infty \mid S_0 = 0) = 1$ and

$$\mathbb{E}[T_0^r \mid S_0 = 0] = \mathbb{E}\left[T_0^r \mathbb{1}_{\{T_0^r < \infty\}} \mid S_0 = 0\right] = G'_{T_0^r}(1) = \infty \qquad (3.4.20)$$

as the slope of $s \mapsto G_{T_0^r}(s)$ in Fig. 3.4b) is infinite at $s = 1$, or by taking the limit as $p, q \to 1/2$ in (3.4.19) or (3.4.18).

When $p = q = 1/2$ the random walk returns to state (0) with probability *one* within a *finite* (random) time, while the average of this random time is *infinite*. This yields another example of a random variable T_0^r which is almost surely finite, while its expectation is infinite as in the St. Petersburg paradox.

This shows how even a fair game can be risky when the player's wealth is negative as it will take on average an infinite time to recover the losses.

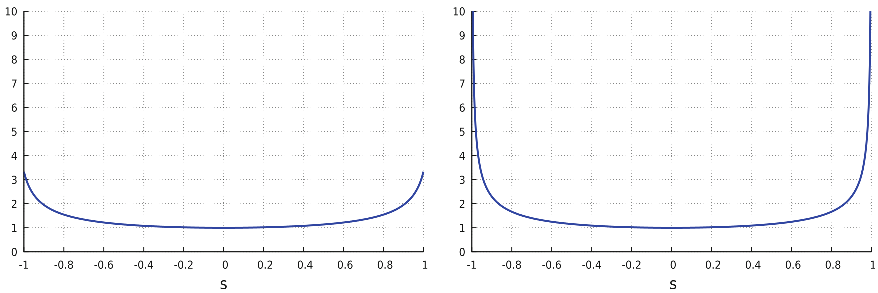
First Return Time Distribution

Proposition 3.4 can also be obtained from the path counting result of Exercise 3.8.

Proposition 3.4 *The probability distribution $\mathbb{P}(T_0^r = n \mid S_0 = 0)$ of the first return time T_0^r to (0) is given by*

$$\mathbb{P}(T_0^r = 2k \mid S_0 = 0) = \frac{1}{2k - 1}\binom{2k}{k}(pq)^k, \qquad k \in \mathbb{N}, \qquad (3.4.21)$$

with $\mathbb{P}(T_0^r = 2k + 1 \mid S_0 = 0) = 0$, $k \in \mathbb{N}$.



(a) Graph of $G_{T_0^r}(s)$ with $p = 0.35$. (b) Graph of $G_{T_0^r}(s)$ with $p = 0.5$.

Fig. 3.4 Probability generating functions of T_0^r for $p = 0.35$ and $p = 0.5$

Proof By applying a Taylor expansion to $s \longmapsto 1-(1-4pqs^2)^{1/2}$ in (3.4.9), we get

$$\begin{aligned} G_{T_0^r}(s) &= 1-(1-4pqs^2)^{1/2} \\ &= 1-\sum_{k=0}^{\infty}\frac{1}{k!}(-4pqs^2)^k\left(\frac{1}{2}-0\right)\left(\frac{1}{2}-1\right)\times\cdots\times\left(\frac{1}{2}-(k-1)\right) \\ &= \frac{1}{2}\sum_{k=1}^{\infty}s^{2k}\frac{(4pq)^k}{k!}\left(1-\frac{1}{2}\right)\times\cdots\times\left(k-1-\frac{1}{2}\right), \end{aligned} \tag{3.4.22}$$

where we used (13.9) for $\alpha = 1/2$. By identification of (3.4.22) with the expansion

$$G_{T_0^r}(s) = \sum_{n=0}^{\infty} s^n \mathbb{P}(T_0^r = n \mid S_0 = 0), \qquad -1 \le s \le 1,$$

we obtain

$$\begin{aligned} \mathbb{P}(T_0^r = 2k \mid S_0 = 0) = g(2k) & \\ &= \frac{(4pq)^k}{k!}\frac{1}{2}\left(1-\frac{1}{2}\right)\times\cdots\times\left(k-1-\frac{1}{2}\right) \\ &= \frac{(4pq)^k}{2k!}\prod_{m=1}^{k-1}\left(m-\frac{1}{2}\right) \\ &= \frac{1}{2k-1}\binom{2k}{k}(pq)^k, \qquad k \in \mathbb{N}, \end{aligned}$$

while $\mathbb{P}(T_0^r = 2k+1 \mid S_0 = 0) = g(2k+1) = 0, k \in \mathbb{N}$. This conclusion could also be obtained using (1.7.8) from the relation

$$\mathbb{P}(T_0^r = n \mid S_0 = 0) = \frac{1}{n!}\frac{\partial^n}{\partial s^n}G_{T_0^r}(s)_{|s=0}, \qquad n \in \mathbb{N}.$$

□

<u>Exercise</u>: Check that the formula (3.4.21) recovers (3.4.1) and (3.4.2) when $k = 0, 1, 2$.

Using the independence of increments of the random walk $(S_n)_{n\in\mathbb{N}}$ one can also show that the probability generating function of the first passage time

$$T_k = \inf\{n \ge 0 \ : \ S_n = k\}$$

to any level $k \ge 1$ is given by

$$G_{T_k}(s) = \left(\frac{1-\sqrt{1-4pqs^2}}{2qs}\right)^k, \qquad 4pqs^2 < 1, \quad q \le p, \tag{3.4.23}$$

from which the distribution of T_k can be computed given the series expansion of $G_{T_k}(s)$, cf. Exercise 3.4 below with $k = -i$.

The gambling process of Chap. 2 and the standard random walk $(S_n)_{n\in\mathbb{N}}$ will later be reconsidered as particular cases in the general framework of Markov chains of Chaps. 4 and 5.

Exercises

Exercise 3.1 We consider the simple random walk $(S_n)_{n\in\mathbb{N}}$ of Sect. 3.1 with independent increments and started at $S_0 = 0$, in which the probability of advance is p and the probability of retreat is $1 - p$.

(a) Enumerate all possible sample paths that conduct to $S_4 = 2$ starting from $S_0 = 0$.
(b) Show that

$$\mathbb{P}(S_4 = 2 \mid S_0 = 0) = \binom{4}{3} p^3(1-p) = \binom{4}{1} p^3(1-p).$$

(c) Show that we have

$$\mathbb{P}(S_n = k \mid S_0 = 0)$$
$$= \begin{cases} \binom{n}{(n+k)/2} p^{(n+k)/2}(1-p)^{(n-k)/2}, & n+k \text{ even and } |k| \le n, \\ \mathbb{P}(S_n = k \mid S_0 = 0) = 0, & n+k \text{ odd or } |k| > n. \end{cases} \tag{3.4.24}$$

(d) Show, by a direct argument using a "last step" analysis at time $n+1$ on random walks, that $p_{n,k} := \mathbb{P}(S_n = k \mid S_0 = 0)$ satisfies the difference equation

$$p_{n+1,k} = p p_{n,k-1} + q p_{n,k+1}, \tag{3.4.25}$$

under the boundary conditions $p_{0,0} = 1$ and $p_{0,k} = 0$, $k \neq 0$.
(e) Confirm that $p_{n,k} = \mathbb{P}(S_n = k \mid S_0 = 0)$ given by (3.4.24) satisfies the equation (3.4.25) and its boundary conditions.

Exercise 3.2 Consider a random walk $(S_n)_{n\in\mathbb{N}}$ on $\mathbb{Z}$ with independent increments and probabilities p, resp. $q = 1 - p$ of moving up by one step, resp. down by one step. Let

$$T_0 = \inf\{n \ge 0 \ : \ S_n = 0\}$$

denote the hitting time of state 0.

(a) Explain why for any $k \ge 1$ we have

$$\mathbb{E}[T_0 \mid S_0 = k] = k\, \mathbb{E}[T_0 \mid S_0 = 1],$$

and compute $\mathbb{E}[T_0 \mid S_0 = 1]$ using first step analysis when $q > p$. What can we conclude when $p \geq q$?

(b) Explain why, by the Markov property, we have

$$\mathbb{P}(T_0 < \infty \mid S_0 = k) = (\mathbb{P}(T_0 < \infty \mid S_0 = 1))^k, \qquad k \geq 1.$$

(c) Using first step analysis for random walks, show that $\alpha := \mathbb{P}(T_0 < \infty \mid S_0 = 1)$ satisfies the quadratic equation

$$p\alpha^2 - \alpha + q = p(\alpha - q/p)(\alpha - 1) = 0,$$

and give the values of $\mathbb{P}(T_0 < \infty \mid S_0 = 1)$ and $\mathbb{P}(T_0 = \infty \mid S_0 = 1)$ in the cases $p < q$ and $p \geq q$ respectively.

Exercise 3.3 Consider the random walk

$$S_n := X_1 + \cdots + X_n, \qquad n \geq 1,$$

with $S_0 = 0$, where $(X_k)_{k\geq 1}$ is a sequence of Bernoulli random variables with

$$\mathbb{P}(X_k = 1) = p \in (0, 1), \qquad \mathbb{P}(X_k = -1) = q \in (0, 1),$$

and $p + q = 1$. Recall that the probability generating function (PGF)

$$G_{T_0^r}(s) = \sum_{k=0}^{\infty} s^k \mathbb{P}(T_0^r = k), \qquad s \in [-1, 1], \tag{3.4.26}$$

of the first return time $T_0^r := \inf\{S_n = 0 \ : \ n \geq 1\}$ to state (0) is given by

$$G_{T_0^r}(s) = 1 - \sqrt{1 - 4pqs^2}, \qquad s \in [-1, 1]. \tag{3.4.27}$$

(a) Compute $\mathbb{P}(T_0^r = 0)$ and $\mathbb{P}(T_0^r < \infty)$ from $G_{T_0^r}$.
(b) By differentiation of (3.4.26) and (3.4.27), compute $\mathbb{P}(T_0^r = 1)$, $\mathbb{P}(T_0^r = 2)$, $\mathbb{P}(T_0^r = 3)$ and $\mathbb{P}(T_0^r = 4)$ using the PGF $G_{T_0^r}$.
(c) Compute $\mathbb{E}[T_0^r \mid T_0^r < \infty]$ using the PGF $G_{T_0^r}$.

Exercise 3.4 Consider a simple random walk $(X_n)_{n\geq 0}$ on $\mathbb{Z}$ with respective probabilities p and q of increment and decrement. Let

$$T_0 := \inf\{n \geq 0 \ : \ X_n = 0\}$$

denote the first hitting time of state (0), and consider the probability generating function

$$G_i(s) := \mathbb{E}\left[s^{T_0} \mid X_0 = i\right], \qquad -1 < s < 1, \quad i \in \mathbb{Z}.$$

(a) By a first step analysis argument, find the finite difference equation satisfied by $G_i(s)$, and its boundary condition(s) at $i = 0$ and $i = \pm\infty$.
(b) Find the value of $G_i(s)$ for all $i \in \mathbb{Z}$ and $s \in (0, 1)$, and recover the result of (3.4.23) on the probability generating function of the hitting time T_0 of $\textcircled{0}$ starting from state $\textcircled{i}$.
(c) Recover Relation (2.2.13)–(3.4.16) using $G_i(s)$.
(d) Recover Relation (2.3.12)–(3.4.18) by differentiation of $s \mapsto G_i(s)$.
(e) Recover the result of (3.4.9) on the probability generating function of the return time T_0^r to $\textcircled{0}$.

Exercise 3.5 Using the probability distribution (3.4.21) of T_0^r, recover the fact that $\mathbb{E}[T_0^r \mathbb{1}_{\{T_0^r<\infty\}} \mid S_0 = 0] = \infty$, when $p = q = 1/2$.

Exercise 3.6 Consider a sequence $(X_k)_{k\geq 1}$ of independent Bernoulli random variables with

$$\mathbb{P}(X_k = 1) = p, \quad \text{and} \quad \mathbb{P}(X_k = -1) = q, \qquad k \geq 1,$$

where $p + q = 1$, and let the process $(M_n)_{n\in\mathbb{N}}$ be defined by $M_0 := 0$ and

$$M_n := \sum_{k=1}^{n} 2^{k-1} X_k, \qquad n \geq 1.$$

(a) Compute $\mathbb{E}[M_n]$ for all $n \geq 0$.
(b) Consider the hitting time $\tau := \inf\{n \geq 1 \ : \ M_n = 1\}$ and the stopped process

$$M_{\min(n,\tau)} = M_n \mathbb{1}_{\{n<\tau\}} + \mathbb{1}_{\{\tau\leq n\}}, \qquad n \in \mathbb{N}.$$

Determine the possible values of $M_{\min(n,\tau)}$, and the probability distribution of $M_{\min(n,\tau)}$ at any time $n \geq 1$.
(c) Give an interpretation of the stopped process $(M_{\min(n,\tau)})_{n\in\mathbb{N}}$ in terms of strategy in a game started at $M_0 = 0$.
(d) Based on the result of part (b), compute $\mathbb{E}[M_{\min(n,\tau)}]$ for all $n \geq 1$.

Exercise 3.7 Winning streaks. Consider a sequence $(X_n)_{n\geq 1}$ of independent Bernoulli random variables with the distribution

$$\mathbb{P}(X_n = 1) = p, \qquad \mathbb{P}(X_n = 0) = q, \qquad n \geq 1,$$

with $q := 1 - p$. For some $m \geq 1$, let $T^{(m)}$ denote the time of the first appearance of m consecutive "1" in the sequence $(X_n)_{n\geq 1}$. For example, for $m = 4$ the following sequence

$$(\overset{\overset{1}{\downarrow}}{0}, \overset{\overset{2}{\downarrow}}{1}, \overset{\overset{3}{\downarrow}}{1}, \overset{\overset{4}{\downarrow}}{0}, \underbrace{\overset{\overset{5}{\downarrow}}{1}, \overset{\overset{6}{\downarrow}}{1}, \overset{\overset{7}{\downarrow}}{1}, \overset{\overset{8}{\downarrow}}{1}}_{4 \text{ times}}, 0, 1, 1, 0, \ldots)$$

yields $T^{(4)} = 8$.

(a) Compute $\mathbb{P}(T^{(m)} < m), \mathbb{P}(T^{(m)} = m), \mathbb{P}(T^{(m)} = m+1)$, and $\mathbb{P}(T^{(m)} = m+2)$.

(b) Show that the probability generating function

$$G_{T^{(m)}}(s) := \mathbb{E}\left[s^{T^{(m)}} \mathbb{1}_{\{T^{(m)}<\infty\}}\right], \qquad s \in (-1, 1),$$

satisfies

$$G_{T^{(m)}}(s) = p^m s^m + \sum_{k=0}^{m-1} p^k q s^{k+1} G_{T^{(m)}}(s), \qquad s \in (-1, 1). \tag{3.4.28}$$

Hint: Look successively at all possible starting patterns of the form

$$(1, \ldots, 1, \overset{\overset{k}{\downarrow}}{1}, 0, \ldots),$$

where $k = 0, 1, \ldots, m$, compute their respective probabilities, and apply a "k-step analysis" argument.

(c) From (3.4.28), compute the probability generating function $G_{T^{(m)}}$ of $T^{(m)}$ for all $s \in (-1, 1)$.

Hint: We have

$$\sum_{k=0}^{m-1} x^k = \frac{1-x^m}{1-x}, \qquad x \in (-1, 1).$$

(d) From the probability generating function $G_{T^{(m)}}(s)$, compute $\mathbb{E}[T^{(m)}]$ for all $m \geq 1$.

Hint: It can be simpler to differentiate inside (3.4.28) and to use the relation

$$(1-x)\sum_{k=0}^{m-1}(k+1)x^k + mx^m = \frac{1-x^m}{1-x}, \qquad x \in (-1, 1).$$

Exercise 3.8 Consider a random walk $(S_n)_{n\in\mathbb{N}}$ on $\mathbb{Z}$ with increments ± 1, started at $S_0 = 0$. Recall that the number of paths joining states 0 and $\boxed{2k}$ over $2m$ time steps is

$$\binom{2m}{m+k}. \tag{3.4.29}$$

(a) Compute the total number of paths joining $S_1 = 1$ to $S_{2n-1} = 1$.
Hint: Apply the formula (3.4.29).

(b) Compute the total number of paths joining $S_1 = 1$ to $S_{2n-1} = -1$.
Hint: Apply the formula (3.4.29).

(c) Show that to every one path joining $S_1 = 1$ to $S_{2n-1} = 1$ by *crossing or hitting* $\textcircled{0}$ we can associate one path joining $S_1 = 1$ to $S_{2n-1} = -1$, in a one-to-one correspondence.
Hint: Draw a sample path joining $S_1 = 1$ to $S_{2n-1} = 1$, and reflect it in such a way that the reflected path then joins $S_1 = 1$ to $S_{2n-1} = -1$.

(d) Compute the total number of paths joining $S_1 = 1$ to $S_{2n-1} = 1$ by *crossing or hitting* $\textcircled{0}$.
Hint: Combine the answers to part (b) and part (c).

(e) Compute the total number of paths joining $S_1 = 1$ to $S_{2n-1} = 1$ *without* crossing or hitting $\textcircled{0}$.
Hint: Combine the answers to part (a) and (d).

(f) Give the total number of paths joining $S_0 = 0$ to $S_{2n} = 0$ *without* crossing or hitting $\textcircled{0}$ between time 1 and time $2n - 1$.
Hint: Apply two times the answer to part (e). A drawing is recommended.

Exercise 3.9 Range process. Consider the random walk $(S_n)_{n\geq 0}$ defined by $S_0 = 0$ and

$$S_n := X_1 + \cdots + X_n, \qquad n \geq 1,$$

where $(X_k)_{k\geq 1}$ is an *i.i.d.*[7] family of $\{-1, +1\}$-valued random variables with distribution

$$\begin{cases} \mathbb{P}(X_k = +1) = p, \\ \mathbb{P}(X_k = -1) = q, \end{cases}$$

$k \geq 1$, where $p + q = 1$. We let R_n denote the *range* of $(S_0, S_1, \ldots, S_n)$, i.e. the (random) number of distinct values appearing in the sequence $(S_0, S_1, \ldots, S_n)$.

(a) Explain why

$$R_n = 1 + \left(\sup_{k=0,1,\ldots,n} S_k \right) - \left(\inf_{k=0,1,\ldots,n} S_k \right),$$

and give the value of R_0 and R_1.

(b) Show that for all $k \geq 1$, $R_k - R_{k-1}$ is a Bernoulli random variable, and that

$$\mathbb{P}(R_k - R_{k-1} = 1) = \mathbb{P}(S_k - S_0 \neq 0, S_k - S_1 \neq 0, \ldots, S_k - S_{k-1} \neq 0).$$

(c) Show that for all $k \geq 1$ we have

$$\mathbb{P}(R_k - R_{k-1} = 1) = \mathbb{P}(X_1 \neq 0, X_1 + X_2 \neq 0, \ldots, X_1 + \cdots + X_k \neq 0).$$

(d) Show why the telescoping identity $R_n = R_0 + \sum_{k=1}^{n} (R_k - R_{k-1})$ holds for all $n \in \mathbb{N}$.

(e) Show that $\mathbb{P}(T_0^r = \infty) = \lim_{k\to\infty} \mathbb{P}(T_0^r > k)$.

[7] Independent and identically distributed.

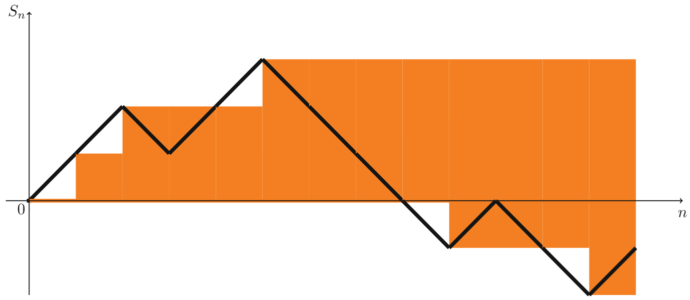


Fig. 3.5 Illustration of the range process

(f) From the results of Questions (c) and (d), show that

$$\mathbb{E}[R_n] = \sum_{k=0}^{n} \mathbb{P}(T_0^r > k), \qquad n \in \mathbb{N},$$

where $T_0^r = \inf\{n \geq 1 \ : \ S_n = 0\}$ is the time of first return to (0) of the random walk.

(g) From the results of Questions (e) and (f), show that

$$\mathbb{P}(T_0^r = \infty) = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}[R_n].$$

(h) Show that

$$\lim_{n \to \infty} \frac{1}{n} \mathbb{E}[R_n] = 0.$$

when $p = 1/2$, and that $\mathbb{E}[R_n] \simeq_{n \to \infty} n|p - q|$, when $p \neq 1/2$.[8]

In Fig. 3.5 the height at time n of the colored area coincides with $R_n - 1$.

[8]The meaning of $f(n) \simeq_{n \to \infty} g(n)$ is $\lim_{n \to \infty} f(n)/g(n) = 1$, provided that $g(n) \neq 0$, $n \geq 1$.

Chapter 4
Discrete-Time Markov Chains

In this chapter we start the general study of discrete-time Markov chains by focusing on the Markov property and on the role played by transition probability matrices. We also include a complete study of the time evolution of the two-state chain, which represents the simplest example of Markov chain.

4.1 Markov Property

We consider a discrete-time stochastic process $(Z_n)_{n\in\mathbb{N}}$ taking values in a discrete state space $\mathbb{S}$, typically $\mathbb{S} = \mathbb{Z}$.

The $\mathbb{S}$-valued process $(Z_n)_{n\in\mathbb{N}}$ is said to be *Markov*, or to have the *Markov property* if, for all $n \geq 1$, the probability distribution of Z_{n+1} is determined by the state Z_n of the process at time n, and does not depend on the past values of Z_k for $k = 0, 1, \ldots, n-1$.

In other words, for all $n \geq 1$ and all $i_0, i_1, \ldots, i_n, j \in \mathbb{S}$ we have

$$\mathbb{P}(Z_{n+1} = j \mid Z_n = i_n,\ Z_{n-1} = i_{n-1}, \ldots, Z_0 = i_0) = \mathbb{P}(Z_{n+1} = j \mid Z_n = i_n).$$

In particular we have

$$\mathbb{P}(Z_{n+1} = j \mid Z_n = i_n,\ Z_{n-1} = i_{n-1}) = \mathbb{P}(Z_{n+1} = j \mid Z_n = i_n),$$

and

$$\mathbb{P}(Z_2 = j \mid Z_1 = i_1, Z_0 = i_0) = \mathbb{P}(Z_2 = j \mid Z_1 = i_1).$$

Note that this feature is apparent in the statement of Lemma 2.2. In addition, we have the following facts.

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4_4

1. The first order transition probabilities can be used for the complete computation of the probability distribution of the process as

$$
\begin{aligned}
&\mathbb{P}(Z_n = i_n, Z_{n-1} = i_{n-1}, \ldots, Z_0 = i_0)\\
&\quad = \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1}, \ldots, Z_0 = i_0)\mathbb{P}(Z_{n-1} = i_{n-1}, \ldots, Z_0 = i_0)\\
&\quad = \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1})\mathbb{P}(Z_{n-1} = i_{n-1} \mid Z_{n-2} = i_{n-2}, \ldots, Z_0 = i_0)\\
&\qquad \times \mathbb{P}(Z_{n-2} = i_{n-2}, \ldots, Z_0 = i_0)\\
&\quad = \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1})\mathbb{P}(Z_{n-1} = i_{n-1} \mid Z_{n-2} = i_{n-2})\\
&\qquad \times \mathbb{P}(Z_{n-2} = i_{n-2} \mid Z_{n-3} = i_{n-3}, \ldots, Z_0 = i_0)\mathbb{P}(Z_{n-3} = i_{n-3}, \ldots, Z_0 = i_0)\\
&\quad = \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1})\mathbb{P}(Z_{n-1} = i_{n-1} \mid Z_{n-2} = i_{n-2})\\
&\qquad \times \mathbb{P}(Z_{n-2} = i_{n-2} \mid Z_{n-3} = i_{n-3})\mathbb{P}(Z_{n-3} = i_{n-3}, \ldots, Z_0 = i_0),
\end{aligned}
$$

   which shows, reasoning by induction, that

$$
\begin{aligned}
&\mathbb{P}(Z_n = i_n, Z_{n-1} = i_{n-1}, \ldots, Z_0 = i_0) \qquad (4.1.1)\\
&\qquad = \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1}) \cdots \mathbb{P}(Z_1 = i_1 \mid Z_0 = i_0)\mathbb{P}(Z_0 = i_0),
\end{aligned}
$$

   or

$$
\begin{aligned}
&\mathbb{P}(Z_n = i_n, Z_{n-1} = i_{n-1}, \ldots, Z_1 = i_1 \mid Z_0 = i_0) \qquad (4.1.2)\\
&\qquad = \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1}) \cdots \mathbb{P}(Z_1 = i_1 \mid Z_0 = i_0),
\end{aligned}
$$

   $i_0, i_1, \ldots, i_n \in \mathbb{S}$.
2. By the *law of total probability* (1.3.1) applied to the events $A_k = \{Z_2 = i_2$ and $Z_1 = k\}$, $k \in \mathbb{S}$, under the probability measure $\mathbb{P}(\cdot \mid Z_0 = i_0)$ we also have

$$
\begin{aligned}
\mathbb{P}(Z_2 = i_2 \mid Z_0 = i_0) &= \sum_{i_1 \in \mathbb{S}} \mathbb{P}(Z_2 = i_2 \text{ and } Z_1 = i_1 \mid Z_0 = i_0)\\
&= \sum_{i_1 \in \mathbb{S}} \mathbb{P}(Z_2 = i_2 \mid Z_1 = i_1)\mathbb{P}(Z_1 = i_1 \mid Z_0 = i_0),
\end{aligned}
$$

   $i_0, i_2 \in \mathbb{S}$, and

$$
\begin{aligned}
\mathbb{P}(Z_1 = i_1) &= \sum_{i_0 \in \mathbb{S}} \mathbb{P}(Z_1 = i_1, Z_0 = i_0)\\
&= \sum_{i_0 \in \mathbb{S}} \mathbb{P}(Z_1 = i_1 \mid Z_0 = i_0)\mathbb{P}(Z_0 = i_0), \qquad i_1 \in \mathbb{S}. \qquad (4.1.3)
\end{aligned}
$$

Example

The random walk

$$
S_n = X_1 + X_2 + \cdots + X_n, \qquad n \in \mathbb{N}, \qquad (4.1.4)
$$

considered in Chap. 3, where $(X_n)_{n\geq 1}$ is a sequence of independent $\mathbb{Z}$-valued random increments, is a discrete-time Markov chain with $\mathbb{S} = \mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

Indeed, the value of S_{n+1} depends only on S_n and on the value of the next increment X_{n+1}. In other words, for all $j, i_n, \ldots, i_1 \in \mathbb{Z}$ we have (note that $S_0 = 0$ here)

$$
\begin{aligned}
&\mathbb{P}(S_{n+1} = j \mid S_n = i_n,\ S_{n-1} = i_{n-1}, \ldots, S_1 = i_1) \qquad (4.1.5)\\
&= \frac{\mathbb{P}(S_{n+1} = j, S_n = i_n, S_{n-1} = i_{n-1}, \ldots, S_1 = i_1)}{\mathbb{P}(S_n = i_n,\ S_{n-1} = i_{n-1}, \ldots, S_1 = i_1)}\\
&= \frac{\mathbb{P}(S_{n+1} - S_n = j - i_n, S_n - S_{n-1} = i_n - i_{n-1}, \ldots, S_2 - S_1 = i_2 - i_1, S_1 = i_1)}{\mathbb{P}(S_n - S_{n-1} = i_n - i_{n-1}, \ldots, S_2 - S_1 = i_2 - i_1, S_1 = i_1)}\\
&= \frac{\mathbb{P}(X_{n+1} = j - i_n, X_n = i_n - i_{n-1}, \ldots, X_2 = i_2 - i_1, X_1 = i_1)}{\mathbb{P}(X_n = i_n - i_{n-1}, \ldots, X_2 = i_2 - i_1, X_1 = i_1)}\\
&= \frac{\mathbb{P}(X_{n+1} = j - i_n)\mathbb{P}(X_n = i_n - i_{n-1}, \ldots, X_2 = i_2 - i_1, X_1 = i_1)}{\mathbb{P}(X_n = i_n - i_{n-1}, \ldots, X_2 = i_2 - i_1, X_1 = i_1)}\\
&= \mathbb{P}(X_{n+1} = j - i_n)\\
&= \frac{\mathbb{P}(X_{n+1} = j - i_n)\mathbb{P}(X_n + \cdots + X_1 = i_n)}{\mathbb{P}(X_1 + \cdots + X_n = i_n)}\\
&= \frac{\mathbb{P}(X_{n+1} = j - i_n, X_n + \cdots + X_1 = i_n)}{\mathbb{P}(X_1 + \cdots + X_n = i_n)}\\
&= \frac{\mathbb{P}(X_{n+1} = j - i_n \text{ and } S_n = i_n)}{\mathbb{P}(S_n = i_n)} = \frac{\mathbb{P}(S_{n+1} = j \text{ and } S_n = i_n)}{\mathbb{P}(S_n = i_n)}\\
&= \mathbb{P}(S_{n+1} = j \mid S_n = i_n).
\end{aligned}
$$

In addition, the Markov chain $(S_n)_{n\in\mathbb{N}}$ is *time homogeneous* if the random sequence $(X_n)_{n\geq 1}$ is identically distributed.

In particular we have

$$\mathbb{P}(S_{n+1} = j \mid S_n = i) = \mathbb{P}(X_{n+1} = j - i),$$

hence the transition probability from state i to state j of a random walk with independent increments depends only on the difference $j - i$ and on the distribution of X_{n+1}.

More generally, all processes with independent increments are Markov processes. However, *not all Markov chains have independent increments*. In fact, the Markov chains of interest in this chapter do not have independent increments.

4.2 Transition Matrix

As seen above, the random evolution of a Markov chain $(Z_n)_{n\in\mathbb{N}}$ is determined by the data of

$$P_{i,j} := \mathbb{P}(Z_1 = j \mid Z_0 = i), \qquad i, j \in \mathbb{S}, \tag{4.2.1}$$

which coincides with the probability $\mathbb{P}(Z_{n+1} = j \mid Z_n = i)$ which is independent of $n \in \mathbb{N}$. In this case the Markov chain $(Z_n)_{n\in\mathbb{N}}$ is said to be *time homogeneous*. This data can be encoded into a matrix indexed by $\mathbb{S}^2 = \mathbb{S} \times \mathbb{S}$, called the *transition matrix* of the Markov chain:

$$[\ P_{i,j}\ ]_{i,j\in\mathbb{S}} = [\ \mathbb{P}(Z_1 = j \mid Z_0 = i)\ ]_{i,j\in\mathbb{S}}\ ,$$

also written on $\mathbb{S} := \mathbb{Z}$ as

$$P = [\ P_{i,j}\ ]_{i,j\in\mathbb{S}} = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & ⋰ \\ \cdots & P_{-2,-2} & P_{-2,-1} & P_{-2,0} & P_{-2,1} & P_{-2,2} & \cdots \\ \cdots & P_{-1,-2} & P_{-1,-1} & P_{-1,0} & P_{-1,1} & P_{-1,2} & \cdots \\ \cdots & P_{0,-2} & P_{0,-1} & P_{0,0} & P_{0,1} & P_{0,2} & \cdots \\ \cdots & P_{1,-2} & P_{1,-1} & P_{1,0} & P_{1,1} & P_{1,2} & \cdots \\ \cdots & P_{2,-2} & P_{2,-1} & P_{2,0} & P_{2,1} & P_{2,2} & \cdots \\ ⋰ & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The notion of transition matrix is related to that of (weighted) adjacency matrix in graph theory.

Note the inversion of the order of indices (i, j) between $\mathbb{P}(Z_{n+1} = j \mid Z_n = i)$ and $P_{i,j}$. In particular, the initial state i is a *row number* in the matrix, while the final state j corresponds to a *column number*.

By the *law of total probability* (1.3.1) applied to the probability measure $\mathbb{P}(\cdot \mid Z_0 = i)$ we have the relation

$$\sum_{j\in\mathbb{S}} \mathbb{P}(Z_1 = j \mid Z_0 = i) = \mathbb{P}(\cup_{j\in\mathbb{S}}\{Z_1 = j\} \mid Z_0 = i) = \mathbb{P}(\Omega) = 1, \qquad i \in \mathbb{N}, \tag{4.2.2}$$

i.e. the *rows* of the transition matrix satisfy the condition

$$\sum_{j\in\mathbb{S}} P_{i,j} = 1,$$

for every row index $i \in \mathbb{S}$.

Using the matrix notation $P = (P_{i,j})_{i,j\in\mathbb{S}}$, and Relation (4.1.1) we find

$$\mathbb{P}(Z_n = i_n, Z_{n-1} = i_{n-1}, \dots, Z_0 = i_0) = P_{i_{n-1},i_n} \cdots P_{i_0,i_1} \mathbb{P}(Z_0 = i_0),$$

$i_0, i_1, \dots, i_n \in \mathbb{S}$, and we rewrite (4.1.3) as

$$\mathbb{P}(Z_1 = i) = \sum_{j\in\mathbb{S}} \mathbb{P}(Z_1 = i \mid Z_0 = j)\mathbb{P}(Z_0 = j) = \sum_{j\in\mathbb{S}} P_{j,i}\mathbb{P}(Z_0 = j), \quad i \in \mathbb{S}. \tag{4.2.3}$$

A state $k \in \mathbb{S}$ is said to be *absorbing* if $P_{k,k} = 1$.

In the sequel we will often consider $\mathbb{S} = \mathbb{N} = \{0, 1, 2, \ldots\}$ and $\mathbb{N}$-valued Markov chains, in which case the *transition matrix* $\left[\ \mathbb{P}(Z_{n+1} = j \mid Z_n = i)\ \right]_{i,j\in\mathbb{N}}$ of the chain is written as

$$\left[\ P_{i,j}\ \right]_{i,j\in\mathbb{N}} = \begin{bmatrix} P_{0,0} & P_{0,1} & P_{0,2} & \cdots \\ P_{1,0} & P_{1,1} & P_{1,2} & \cdots \\ P_{2,0} & P_{2,1} & P_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

From (4.2.2) we have

$$\sum_{j=0}^{\infty} P_{i,j} = 1,$$

for all $i \in \mathbb{N}$.

In case the Markov chain $(Z_k)_{k\in\mathbb{N}}$ takes values in the finite state space $\mathbb{S} = \{0, 1, \ldots, N\}$ its $(N+1) \times (N+1)$ transition matrix will simply have the form

$$\left[\ P_{i,j}\ \right]_{0\le i,j\le N} = \begin{bmatrix} P_{0,0} & P_{0,1} & P_{0,2} & \cdots & P_{0,N} \\ P_{1,0} & P_{1,1} & P_{1,2} & \cdots & P_{1,N} \\ P_{2,0} & P_{2,1} & P_{2,2} & \cdots & P_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{N,0} & P_{N,1} & P_{N,2} & \cdots & P_{N,N} \end{bmatrix}.$$

Still on the finite state space $\mathbb{S} = \{0, 1, \ldots, N\}$, Relation (4.2.3) can be restated in the language of matrix and vector products using the shorthand notation:

$$\boxed{\eta = \pi P, \tag{4.2.4}}$$

where

$$\eta = [\mathbb{P}(Z_1 = 0), \ldots, \mathbb{P}(Z_1 = N)] = [\eta_0, \eta_1, \ldots, \eta_N] \in \mathbb{R}^{N+1}$$

is the row vector "distribution of Z_1",

$$\pi = [\mathbb{P}(Z_0 = 0), \ldots, \mathbb{P}(Z_0 = N)] = [\pi_0, \ldots, \pi_N] \in \mathbb{R}^{N+1}$$

is the row vector representing the probability distribution of Z_0, and

$$[\eta_0, \eta_1, \ldots, \eta_N] = [\pi_0, \ldots, \pi_N] \times \begin{bmatrix} P_{0,0} & P_{0,1} & P_{0,2} & \cdots & P_{0,N} \\ P_{1,0} & P_{1,1} & P_{1,2} & \cdots & P_{1,N} \\ P_{2,0} & P_{2,1} & P_{2,2} & \cdots & P_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{N,0} & P_{N,1} & P_{N,2} & \cdots & P_{N,N} \end{bmatrix}. \tag{4.2.5}$$

Invariant Vectors

A row vector π such that $\pi = \pi P$ is said to be *invariant* or *stationary* by the transition matrix P.

For example, in case the matrix P takes the form

$$P = \left[\, P_{i,j} \,\right]_{0 \le i,j \le N} = \begin{bmatrix} \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \end{bmatrix},$$

with all rows equal and $\pi_0 + \pi_1 + \cdots + \pi_N = 1$, then we have $\pi = \pi P$, i.e. π is an invariant (or stationary) distribution for P.

4.3 Examples of Markov Chains

The wide range of applications of Markov chains to engineering, physics and biology has already been mentioned in the introduction. Here we consider some more specific examples.

(i) *Random walk.*

The transition matrix $\left[\, P_{i,j} \,\right]_{i,j \in \mathbb{S}}$ of the unrestricted random walk (4.1.4) is given by

$$
\left[P_{i,j} \right]_{i,j\in\mathbb{S}} = \begin{array}{c} \\ \\ i-2 \\ i-1 \\ i \\ i+1 \\ i+2 \\ \\ \end{array}
\begin{array}{c}
\begin{array}{ccccccc} & & i-1 & i & i+1 & & \end{array} \\
\left[\begin{array}{ccccccc}
\ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \iddots \\
\cdots & 0 & p & 0 & 0 & 0 & \cdots \\
\cdots & q & 0 & p & 0 & 0 & \cdots \\
\cdots & 0 & q & 0 & p & 0 & \cdots \\
\cdots & 0 & 0 & q & 0 & p & \cdots \\
\cdots & 0 & 0 & 0 & q & 0 & \cdots \\
\iddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right]
\end{array}. \tag{4.3.1}
$$

(ii) *Gambling process.*

The transition matrix $\left[P_{i,j} \right]_{0\le i,j\le S}$ of the gambling process on $\{0, 1, \ldots, S\}$ with absorbing states 0 and S is given by

$$
P = \left[P_{i,j} \right]_{0\le i,j\le S} = \begin{bmatrix}
1 & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 & 0 \\
q & 0 & p & 0 & \cdots & \cdots & 0 & 0 & 0 & 0 \\
0 & q & 0 & p & \cdots & \cdots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & \cdots & q & 0 & p & 0 \\
0 & 0 & 0 & 0 & \cdots & \cdots & 0 & q & 0 & p \\
0 & 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 & 1
\end{bmatrix}.
$$

(iii) *Credit rating.*

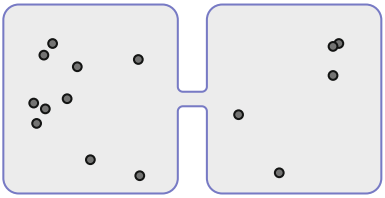
[transition probabilities are expressed in %].

Rating at the start of a year Rating at the end of the year

	AAA	AA	A	BBB	BB	B	CCC	D	N.R.	Total
AAA	90.34	5.62	0.39	0.08	0.03	0	0	0	3.5	100
AA	0.64	88.78	6.72	0.47	0.06	0.09	0.02	0.01	3.21	100
A	0.07	2.16	87.94	4.97	0.47	0.19	0.01	0.04	4.16	100
BBB	0.03	0.24	4.56	84.26	4.19	0.76	0.15	0.22	5.59	100
BB	0.03	0.06	0.4	6.09	76.09	6.82	0.96	0.98	8.58	100
B	0	0.09	0.29	0.41	5.11	74.62	3.43	5.3	10.76	100
CCC	0.13	0	0.26	0.77	1.66	8.93	53.19	21.94	13.14	100
D	0	0	0	0	1.0	3.1	9.29	51.29	37.32	100
N.R.	0	0	0	0	0	0.1	8.55	74.06	17.07	100

We note that higher ratings are more stable since the diagonal coefficients of the matrix go decreasing. On the other hand starting from the rating AA it is easier to be downgraded (probability 6.72%) than to be upgraded (probability 0.64%).

(iv) *Ehrenfest chain.*
Two volumes of air (left and right), containing a total of N balls, are connected by a pipe.



At each time step, one picks a ball at random and moves it to the other side. Let $Z_n \in \{0, 1, \ldots, N\}$ denote the number of balls in the left side at time n. The transition probabilities $\mathbb{P}(Z_{n+1} = j \mid Z_n = i)$, $0 \leq i, j \leq N$, are given by

$$\mathbb{P}(Z_{n+1} = k + 1 \mid Z_n = k) = \frac{N - k}{N}, \qquad k = 0, 1, \ldots, N - 1, \tag{4.3.2}$$

and

$$\mathbb{P}(Z_{n+1} = k - 1 \mid Z_n = k) = \frac{k}{N}, \qquad k = 1, 2, \ldots, N, \tag{4.3.3}$$

with

$$P = \begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 & 0 \\ 1/N & 0 & (N-1)/N & \cdots & \cdots & 0 & 0 \\ 0 & 2/N & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 3/N & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \ddots & 3/N & 0 & 0 \\ 0 & 0 & \cdots & \ddots & 0 & 2/N & 0 \\ 0 & 0 & \cdots & \cdots & (N-1)/N & 0 & 1/N \\ 0 & 0 & \cdots & \cdots & 0 & 1 & 0 \end{bmatrix},$$

cf. Exercises 6.7 and 7.3, Problem 7.23 on modified Ehrenfest chains, and Exercise 4.9 on the Bernoulli–Laplace chain.

(v) *Markov chains in music.*
By a statistical analysis of note transitions, every type of music can be encoded into a Markov chain. An example of such an analysis is presented in the next transition matrix.

	A	A♯	B	C	D	E	F	G	G♯
A	4/19	0	3/19	0	2/19	1/19	0	6/19	3/19
A♯	1	0	0	0	0	0	0	0	0
B	7/15	0	1/15	4/15	0	3/15	0	0	0
C	0	0	6/15	3/15	6/15	0	0	0	0
D	0	0	0	3/11	3/11	5/11	0	0	0
E	4/19	1/19	0	3/19	0	5/19	4/19	1/19	1/19
F	0	0	0	1/5	0	0	1/5	0	3/5
G	1/5	0	1/5	2/5	0	0	0	1/5	0
G♯	0	0	3/4	0	0	1/4	0	0	0

(vi) *Text generation.*
Markov chains can be used to generate sentences in a given language, based on a statistical analysis on the transition between words in a sample text. The state space of the Markov chain can be made of different word sequences.

Other applications of Markov chains include:

- Memory management in computer science,
- Logistics, supply chain management, and waiting queues,
- Modeling of insurance claims,
- Board games, e.g. Snakes and Ladders,
- Genetics, cf. the Wright–Fisher model.
- Random fields in imaging,
- Artificial intelligence, learning theory and machine learning.

Graph Representation

Whenever possible we will represent a Markov chain using a graph, as in the following example with transition matrix, see Fig. 4.1.

$$P = \begin{bmatrix} 0 & 0.2 & 0.8 & 0 & 0 \\ 0.4 & 0 & 0 & 0.6 & 0 \\ 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.4 & 0.6 \\ 0 & 0.4 & 0.6 & 0 & 0 \end{bmatrix}. \qquad (4.3.4)$$

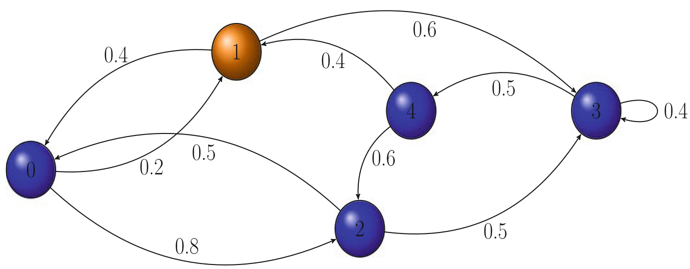


Fig. 4.1 Graph of a five-state Markov chain

4.4 Higher-Order Transition Probabilities

As noted above, the transition matrix P is a convenient way to record $\mathbb{P}(Z_{n+1} = j \mid Z_n = i)$, $i, j \in \mathbb{S}$, into an array of data.

However, it is *much more than that*, as already hinted at in Relation (4.2.4). Suppose for example that we are interested in the two-step transition probability

$$\mathbb{P}(Z_{n+2} = j \mid Z_n = i).$$

This probability does not appear in the transition matrix P, but it can be computed by first step analysis, applying the *law of total probability* (1.3.1) to the probability measure $\mathbb{P}(\cdot \mid Z_n = i)$ as follows.

(i) 2-step transitions. Denoting by $\mathbb{S}$ the state space of the process, we have

$$\begin{aligned}
\mathbb{P}(Z_{n+2} = j \mid Z_n = i) &= \sum_{l \in \mathbb{S}} \mathbb{P}(Z_{n+2} = j \text{ and } Z_{n+1} = l \mid Z_n = i) \\
&= \sum_{l \in \mathbb{S}} \frac{\mathbb{P}(Z_{n+2} = j,\ Z_{n+1} = l,\ Z_n = i)}{\mathbb{P}(Z_n = i)} \\
&= \sum_{l \in \mathbb{S}} \frac{\mathbb{P}(Z_{n+2} = j,\ Z_{n+1} = l,\ Z_n = i)}{\mathbb{P}(Z_{n+1} = l \text{ and } Z_n = i)} \frac{\mathbb{P}(Z_{n+1} = l \text{ and } Z_n = i)}{\mathbb{P}(Z_n = i)} \\
&= \sum_{l \in \mathbb{S}} \mathbb{P}(Z_{n+2} = j \mid Z_{n+1} = l \text{ and } Z_n = i)\mathbb{P}(Z_{n+1} = l \mid Z_n = i) \\
&= \sum_{l \in \mathbb{S}} \mathbb{P}(Z_{n+2} = j \mid Z_{n+1} = l)\mathbb{P}(Z_{n+1} = l \mid Z_n = i) \\
&= \sum_{l \in \mathbb{S}} P_{i,l} P_{l,j} \\
&= [P^2]_{i,j}, \qquad i, j \in \mathbb{S},
\end{aligned}$$

where we used (4.2.1). In other words, using matrix product notation, we find

$$\begin{aligned}
&(\mathbb{P}(Z_{n+2} = j \mid Z_n = i))_{0 \le i,j \le N} \\
&= \begin{bmatrix}
P_{0,0} & P_{0,1} & P_{0,2} & \cdots & P_{0,N} \\
P_{1,0} & P_{1,1} & P_{1,2} & \cdots & P_{1,N} \\
P_{2,0} & P_{2,1} & P_{2,2} & \cdots & P_{2,N} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
P_{N,0} & P_{N,1} & P_{N,2} & \cdots & P_{N,N}
\end{bmatrix} \times \begin{bmatrix}
P_{0,0} & P_{0,1} & P_{0,2} & \cdots & P_{0,N} \\
P_{1,0} & P_{1,1} & P_{1,2} & \cdots & P_{1,N} \\
P_{2,0} & P_{2,1} & P_{2,2} & \cdots & P_{2,N} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
P_{N,0} & P_{N,1} & P_{N,2} & \cdots & P_{N,N}
\end{bmatrix}.
\end{aligned}$$

(i) k-step transitions. More generally, for all $k \in \mathbb{N}$ we have the recursion

$$\begin{aligned}
\mathbb{P}(Z_{n+k+1} = j \mid Z_n = i) &= \sum_{l\in\mathbb{S}} \mathbb{P}(Z_{n+k+1} = j \text{ and } Z_{n+k} = l \mid Z_n = i) \\
&= \sum_{l\in\mathbb{S}} \frac{\mathbb{P}(Z_{n+k+1} = j,\ Z_{n+k} = l,\ Z_n = i)}{\mathbb{P}(Z_n = i)} \\
&= \sum_{l\in\mathbb{S}} \frac{\mathbb{P}(Z_{n+k+1} = j,\ Z_{n+k} = l,\ Z_n = i)}{\mathbb{P}(Z_{n+k} = l \text{ and } Z_n = i)} \frac{\mathbb{P}(Z_{n+k} = l \text{ and } Z_n = i)}{\mathbb{P}(Z_n = i)} \\
&= \sum_{l\in\mathbb{S}} \mathbb{P}(Z_{n+k+1} = j \mid Z_{n+k} = l \text{ and } Z_n = i)\mathbb{P}(Z_{n+k} = l \mid Z_n = i) \\
&= \sum_{l\in\mathbb{S}} \mathbb{P}(Z_{n+k+1} = j \mid Z_{n+k} = l)\mathbb{P}(Z_{n+k} = l \mid Z_n = i) \\
&= \sum_{l\in\mathbb{S}} \mathbb{P}(Z_{n+k} = l \mid Z_n = i) P_{l,j}.
\end{aligned}$$

We have just checked that the family of matrix

$$\left[\ \mathbb{P}(Z_{n+k} = j \mid Z_n = i)\ \right]_{i,j\in\mathbb{S}}, \qquad k \geq 1,$$

satisfies the same induction relation as the *matrix power* P^k, i.e.

$$[P^{k+1}]_{i,j} = \sum_{l\in\mathbb{S}} [P^k]_{i,l} P_{l,j},$$

hence by induction on $k \geq 0$ the equality

$$\boxed{\left[\ \mathbb{P}(Z_{n+k} = j \mid Z_n = i)\ \right]_{i,j\in\mathbb{S}} = \left[\ [P^k]_{i,j}\ \right]_{i,j\in\mathbb{S}} = P^k}$$

holds not only for $k = 0$ and $k = 1$, but also for all $k \in \mathbb{N}$.

⚠Note that in general we have $[P^k]_{i,j} \neq (P_{i,j})^k$, $i, j \in \mathbb{S}$.

The matrix product relation

$$P^{m+n} = P^m P^n = P^n P^m,$$

which reads

$$[P^{m+n}]_{i,j} = \sum_{l\in\mathbb{S}} [P^m]_{i,l} [P^n]_{l,j} = \sum_{l\in\mathbb{S}} [P^n]_{i,l} [P^m]_{l,j}, \qquad i, j \in \mathbb{S},$$

can now be interpreted as

$$\begin{aligned}\mathbb{P}(Z_{n+m} = j \mid Z_0 = i) &= \sum_{l \in \mathbb{S}} \mathbb{P}(Z_m = j \mid Z_0 = l)\mathbb{P}(Z_n = l \mid Z_0 = i)\\ &= \sum_{l \in \mathbb{S}} \mathbb{P}(Z_n = j \mid Z_0 = l)\mathbb{P}(Z_m = l \mid Z_0 = i),\end{aligned}$$

$i, j \in \mathbb{S}$, which is called the *Chapman-Kolmogorov* equation, cf. also the triple (1.2.2).

Example The gambling process $(Z_n)_{n\geq 0}$.

Taking $S = 4$ and $p = 40\%$, the transition matrix of the gambling process on $\mathbb{S} = \{0, 1, \ldots, 4\}$ of Chap. 2 reads

$$P = \left[\ P_{i,j}\ \right]_{0\leq i,j\leq 4} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.6 & 0 & 0.4 & 0 & 0 \\ 0 & 0.6 & 0 & 0.4 & 0 \\ 0 & 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4.4.1}$$

and we can check by hand that:

$$\begin{aligned} P^2 &= P \times P \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.6 & 0 & 0.4 & 0 & 0 \\ 0 & 0.6 & 0 & 0.4 & 0 \\ 0 & 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.6 & 0 & 0.4 & 0 & 0 \\ 0 & 0.6 & 0 & 0.4 & 0 \\ 0 & 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.6 & 0.24 & 0 & 0.16 & 0 \\ 0.36 & 0 & 0.48 & 0 & 0.16 \\ 0 & 0.36 & 0 & 0.24 & 0.4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Exercise: From the above matrix (4.4.1), check that

$$\begin{aligned} &\mathbb{P}(Z_2 = 4 \mid Z_0 = 2) = [P^2]_{2,4} = 0.16,\\ &\mathbb{P}(Z_2 = 1 \mid Z_0 = 2) = [P^2]_{2,1} = 0, \qquad \text{and}\\ &\mathbb{P}(Z_2 = 2 \mid Z_0 = 2) = [P^2]_{2,2} = 0.48. \end{aligned}$$

Example. The fifth order transitions of the chain with Markov matrix (4.3.4) can be computed from the fifth matrix power

$$P^5 = \begin{bmatrix} 0.14352 & 0.09600 & 0.25920 & 0.30160 & 0.19968 \\ 0.15840 & 0.10608 & 0.24192 & 0.30400 & 0.18960 \\ 0.17040 & 0.10920 & 0.23280 & 0.30880 & 0.17880 \\ 0.17664 & 0.11520 & 0.22800 & 0.30928 & 0.17088 \\ 0.14904 & 0.09600 & 0.25440 & 0.30520 & 0.19536 \end{bmatrix}.$$

Note that for large transition orders (for example 1000 time steps) we get

$$P^{1000} = \begin{bmatrix} 0.16273 & 0.10613 & 0.24056 & 0.30660 & 0.18396 \\ 0.16273 & 0.10613 & 0.24056 & 0.30660 & 0.18396 \\ 0.16273 & 0.10613 & 0.24056 & 0.30660 & 0.18396 \\ 0.16273 & 0.10613 & 0.24056 & 0.30660 & 0.18396 \\ 0.16273 & 0.10613 & 0.24056 & 0.30660 & 0.18396 \end{bmatrix},$$

which suggests a *convergence phenomenon* in large time for the Markov chain, see Chap. 7 for details.

Example For the simple random walk of Chap. 3, computing the probability to travel from 0 to $\boxed{2k}=\boxed{10}$ in $2n = 20$ time steps involves a summation over $\binom{20}{10+5} = \binom{2n}{n+k} = 15504$ paths, which can be evaluated by computing $\left[P^{20}\right]_{0,10}$, cf. also Fig. 3.1.

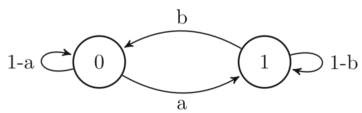
4.5 The Two-State Discrete-Time Markov Chain

The above discussion shows that there is some interest in computing the n-th order transition matrix P^n. Although this is generally difficult, this is actually possible when the number of states equals two, i.e. $\mathbb{S} = \{0, 1\}$.

To close this chapter we provide a complete study of the two-state Markov chain, whose transition matrix has the form

$$P = \begin{bmatrix} 1-a & a \\ & \\ b & 1-b \end{bmatrix}, \tag{4.5.1}$$

with $a \in [0, 1]$ and $b \in [0, 1]$.



We also have

$$\mathbb{P}(Z_{n+1} = 1 \mid Z_n = 0) = a, \qquad \mathbb{P}(Z_{n+1} = 0 \mid Z_n = 0) = 1 - a,$$

and

$$\mathbb{P}(Z_{n+1} = 0 \mid Z_n = 1) = b, \qquad \mathbb{P}(Z_{n+1} = 1 \mid Z_n = 1) = 1 - b.$$

The power

$$P^n = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}^n$$

of the transition matrix P is computed for all $n \geq 0$ in the next Proposition 4.1. We always exclude the case $a = b = 0$ since it corresponds to the trivial case where $P = \mathrm{I_d}$ is the identity matrix (constant chain).

Proposition 4.1 *We have*

$$P^n = \frac{1}{a+b} \begin{bmatrix} b + a(1-a-b)^n & a(1-(1-a-b)^n) \\ b(1-(1-a-b)^n) & a + b(1-a-b)^n \end{bmatrix}, \qquad n \in \mathbb{N}.$$

Proof This result will be proved by a *diagonalization* argument. The matrix P has two eigenvectors

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -a \\ b \end{bmatrix},$$

with respective eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 1 - a - b$.

Hence P can be written in the *diagonal form*

$$P = M \times D \times M^{-1}, \tag{4.5.2}$$

i.e.

$$P = \begin{bmatrix} 1 & -a \\ 1 & b \end{bmatrix} \times \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \times \begin{bmatrix} \dfrac{b}{a+b} & \dfrac{a}{a+b} \\ -\dfrac{1}{a+b} & \dfrac{1}{a+b} \end{bmatrix}.$$

As a consequence of (4.5.2), we have

$$\begin{aligned} P^n &= (M \times D \times M^{-1})^n = (M \times D \times M^{-1}) \cdots (M \times D \times M^{-1}) \\ &= M \times D \times \cdots \times D \times M^{-1} = M \times D^n \times M^{-1}, \end{aligned}$$

where

$$D^n = \begin{bmatrix} 1 & 0 \\ 0 & \lambda_2^n \end{bmatrix}, \qquad n \in \mathbb{N}.$$

hence

$$\begin{aligned} P^n &= \begin{bmatrix} 1 & -a \\ \\ 1 & b \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ \\ 0 & \lambda_2^n \end{bmatrix} \times \begin{bmatrix} \dfrac{b}{a+b} & \dfrac{a}{a+b} \\ \\ -\dfrac{1}{a+b} & \dfrac{1}{a+b} \end{bmatrix} \\ &= \frac{1}{a+b} \begin{bmatrix} b & a \\ \\ b & a \end{bmatrix} + \frac{\lambda_2^n}{a+b} \begin{bmatrix} a & -a \\ \\ -b & b \end{bmatrix} \\ &= \frac{1}{a+b} \begin{bmatrix} b + a\lambda_2^n & a(1-\lambda_2^n) \\ \\ b(1-\lambda_2^n) & a + b\lambda_2^n \end{bmatrix}. \end{aligned} \tag{4.5.3}$$

□

For an alternative proof of Proposition 4.1, see also Exercise 1.4.1 p. 5 of [Nor98] in which P^n is written as

$$P^n = \begin{bmatrix} 1 - a_n & -a_n \\ \\ b_n & 1 - b_n \end{bmatrix}$$

and the relation $P^{n+1} = P \times P^n$ is used to find induction relations for a_n and b_n, cf. the solution of Exercise 7.17 for a similar analysis.

From the result of Proposition 4.1 we may now compute the probabilities

$$\mathbb{P}(Z_n = 0 \mid Z_0 = 0) = \frac{b + a\lambda_2^n}{a+b}, \quad \mathbb{P}(Z_n = 1 \mid Z_0 = 0) = \frac{a(1-\lambda_2^n)}{a+b} \tag{4.5.4}$$

and

$$\mathbb{P}(Z_n = 0 \mid Z_0 = 1) = \frac{b(1-\lambda_2^n)}{a+b}, \quad \mathbb{P}(Z_n = 1 \mid Z_0 = 1) = \frac{a + b\lambda_2^n}{a+b}. \tag{4.5.5}$$

As an example, the value of $\mathbb{P}(Z_3 = 0 \mid Z_0 = 0)$ could also be computed using pathwise analysis as

$$\mathbb{P}(Z_3 = 0 \mid Z_0 = 0) = (1-a)^3 + ab(1-b) + 2(1-a)ab,$$

which coincides with (4.5.4), i.e.

$$\mathbb{P}(Z_3 = 0 \mid Z_0 = 0) = \frac{b + a(1-a-b)^3}{a+b},$$

for $n = 3$. Under the condition

$$-1 < \lambda_2 = 1 - a - b < 1,$$

which is equivalent to $(a, b) \neq (0, 0)$ and $(a, b) \neq (1, 1)$, we can let n go to infinity in (4.5.3) to derive the large time behavior, or limiting distribution, of the Markov chain:

$$\lim_{n\to\infty} P^n = \lim_{n\to\infty} \begin{bmatrix} \mathbb{P}(Z_n = 0 \mid Z_0 = 0) & \mathbb{P}(Z_n = 1 \mid Z_0 = 0) \\ \mathbb{P}(Z_n = 0 \mid Z_0 = 1) & \mathbb{P}(Z_n = 1 \mid Z_0 = 1) \end{bmatrix} = \frac{1}{a+b} \begin{bmatrix} b & a \\ b & a \end{bmatrix}.$$

Note that convergence will be faster when $a + b$ is closer to 1.

Hence we have

$$\lim_{n\to\infty} \mathbb{P}(Z_n = 1 \mid Z_0 = 0) = \lim_{n\to\infty} \mathbb{P}(Z_n = 1 \mid Z_0 = 1) = \frac{a}{a+b} \tag{4.5.6}$$

and

$$\lim_{n\to\infty} \mathbb{P}(Z_n = 0 \mid Z_0 = 0) = \lim_{n\to\infty} \mathbb{P}(Z_n = 0 \mid Z_0 = 1) = \frac{b}{a+b}. \tag{4.5.7}$$

Consequently,

$$\pi = [\pi_0, \pi_1] := \left[\frac{b}{a+b}, \frac{a}{a+b} \right] \tag{4.5.8}$$

is a *limiting distribution* as n goes to infinity, provided that $(a, b) \neq (1, 1)$. In other words, whatever the initial state Z_0, the probability of being at 1 after a "large" time becomes close to $a/(a+b)$, while the probability of being at 0 becomes close to $b/(a+b)$.

In case $a = b = 0$, we have

$$P = I_d = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

the chain is constant and it clearly admits its initial distribution as limiting distribution. In case $a = b = 1$, we have

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

and there is no limiting distribution as the chain switches indefinitely between state (0) and (1)

The notions of limiting and invariant (or stationary) distributions will be treated in Chap. 7 in the general framework of Markov chains, see for example Proposition 7.7.

Remarks

(i) The limiting distribution π in (4.5.8) is *invariant* (or stationary) by P in the sense that

$$\begin{aligned}\pi P &= \frac{1}{a+b}[b,a]\begin{bmatrix}1-a & a\\ b & 1-b\end{bmatrix} = \frac{1}{a+b}\begin{bmatrix}b(1-a)+ab\\ ab+a(1-b)\end{bmatrix}^T\\ &= \frac{1}{a+b}[b,a] = \pi,\end{aligned}$$

i.e. π is invariant (or stationary) with respect to P, and the invariance relation (4.2.4):

$$\pi = \pi P,$$

which means that $\mathbb{P}(Z_1 = k) = \pi_k$ if $\mathbb{P}(Z_0 = k) = \pi_k$, $k = 0, 1$. For example, the distribution $\pi = [1/2, 1/2]$ is clearly invariant (or stationary) for the swapping chain with $a = b = 1$ and transition matrix

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

while $\pi := [1/3, 2/3]$ will *not* be invariant (or stationary) for this chain. This is a two-state particular case of the circular chain of Example (7.2.4).

(ii) If $a + b = 1$, one sees that

$$P^n = \begin{bmatrix}b & a\\ b & a\end{bmatrix} = P$$

for all $n \in \mathbb{N}$ and we find

$$\mathbb{P}(Z_n = 1 \mid Z_k = 0) = \mathbb{P}(Z_n = 1 \mid Z_k = 1) = \mathbb{P}(Z_n = 1) = a$$

and

$$\mathbb{P}(Z_n = 0 \mid Z_k = 0) = \mathbb{P}(Z_n = 0 \mid Z_k = 1) = \mathbb{P}(Z_n = 0) = b$$

for all $k = 0, 1, \ldots, n-1$, regardless of the initial distribution $[\mathbb{P}(Z_0 = 0), \mathbb{P}(Z_0 = 1)]$. In this case, Z_n is independent of Z_k as we have

$$\mathbb{P}(Z_n = 1,\ Z_k = j) = \mathbb{P}(Z_n = i \mid Z_k = j)\mathbb{P}(Z_k = j) = \mathbb{P}(Z_n = i)\mathbb{P}(Z_k = j),$$

$i, j = 0, 1$, $0 \leq k < n$, and $(Z_n)_{n\in\mathbb{N}}$ is an *i.i.d* sequence of random variables with distribution $(1-a, a) = (b, a)$ over $\{0, 1\}$.

(iii) A given proportion $p = a/(a+b) \in (0, 1)$ of visits to state 1 in the long run can be reached by any $a \in (0, p]$ and $b \in (0, 1-p]$ satisfying $a = bp/(1-p)$. Smaller values of a and b will lead to increased *stickiness*. The case $(a, b) = (p, 1-p)$ satisfies $a + b = 1$ and corresponds to minimal stickiness, i.e. to the independence of the sequence $(Z_n)_{n\in\mathbb{N}}$.

(iv) When $a = b = 1$ in (4.5.1) the limit $\lim_{n\to\infty} P^n$ does not exist as we have

$$P^n = \begin{cases} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & n = 2k, \\ \\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, & n = 2k+1, \end{cases}$$

and the chain is indefinitely switching at each time step from one state to the other.

In Figs. 4.2 and 4.3 we consider a simulation of the two-state random walk with transition matrix

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{bmatrix},$$

i.e. $a = 0.2$ and $b = 0.4$. Figure 4.2 represents a sample path $(x_n)_{n=0,1,\ldots,100}$ of the chain, while Fig. 4.3 represents the sample average

$$y_n = \frac{1}{n+1}(x_0 + x_1 + \cdots + x_n), \qquad n = 0, 1, \ldots, 100,$$

which counts the proportion of values of the chain in the state 1. This proportion is found to converge to $a/(a+b) = 1/3$. This is actually a consequence of the Ergodic Theorem, cf. Theorem 7.12 in Chap. 7.

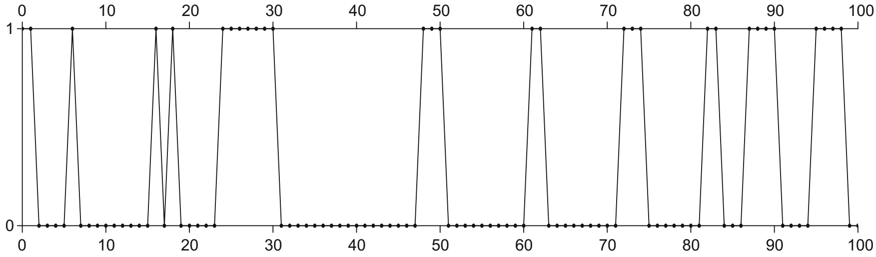


Fig. 4.2 Sample path of a two-state chain in continuous time with $a = 0.2$ and $b = 0.4$

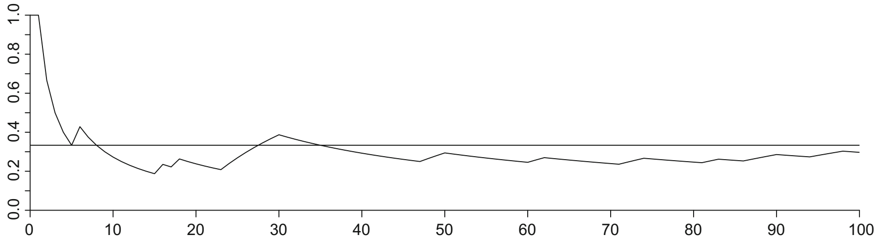


Fig. 4.3 The proportion of chain values at ① tends to $1/3 = a/(a + b)$

```
# Dimension of the transition matrix
d=2
# Parameter definition
a=0.2; b=0.4;
# Definition of the transition matrix
P=matrix(c(1−a,a,b,1−b),nrow=d,ncol=d,byrow=TRUE)
# Number of time steps
N=100
# Encoding of chain values
Z=array(N+1);
for(ll in seq(1,N)) {
Z[1]=sample(d,size=1,prob=P[2,])
# Random simulation of Z[j+1] given Z[j]
for (j in seq(1,N)) Z[j+1]=sample(d,size=1,prob=P[Z[j],])
Y=array(N+1);S=0;
# Computation of the average over the l first steps
for(l in seq(1,N+1)) { Z[l]=Z[l]−1; S=S+Z[l]; Y[l]=S/l; }
X=array(N+1); for(l in seq(1,N+1)) { X[l]=l−1; }
par(mfrow=c(2,1))
plot(X,Y,type="l",yaxt="n",xaxt="n",xlim=c(0,N),xlab="",
ylim=c(0,1),ylab="",xaxs="i",col="black",main="",bty="n")
segments(0,a/(a+b),N,a/(a+b))
axis(2,pos=0,at=c(0.0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0))
axis(1,pos=0,at=seq(0,N,10),outer=TRUE)
plot(X,Z,type="o",xlab="",ylab="",xlim=c(0,N),yaxt="n",
xaxt="n",xaxs="i",col="black",main="",pch=20,bty="n")
axis(1,pos=1,at=seq(0,N+1,10),outer=TRUE,padj=−4,tcl=0.5)
axis(1,pos=0,at=seq(0,N+1,10),outer=TRUE)
axis(2,las=2,at=0:1)
readline(prompt = "Pause. Press <Enter> to continue...")}
```

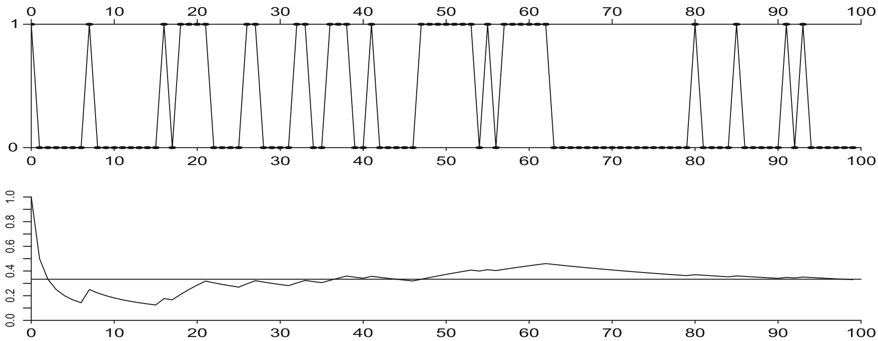


Fig. 4.4 Convergence graph for the two-state Markov chain with $a = 0.2$ and $b = 0.4$

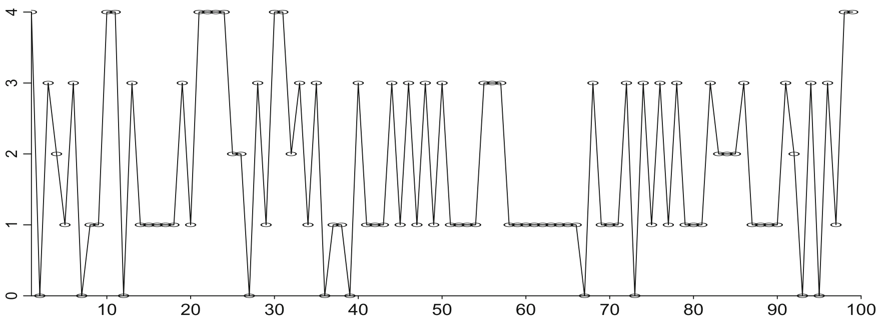


Fig. 4.5 Sample path of a five-state Markov chain

We close this chapter with two other sample paths of Markov chains in Figs. 4.4 and 4.5. In the next Fig. 4.4 we check again that the proportion of chain values in the state 1 converges to $1/3$ for a two-state Markov chain.

In Fig. 4.5 we draw a sample path of a five-state Markov chain.

Exercises

Exercise 4.1 Consider a *symmetric* random walk $(S_n)_{n\in\mathbb{N}}$ on $\mathbb{Z}$ with independent increments ± 1 chosen with equal probability $1/2$, started at $S_0 = 0$.

(a) Is the process $Z_n := 2S_n + 1$ a Markov chain?
(b) Is the process $Z_n := (S_n)^2$ a Markov chain?

Exercise 4.2 Consider the Markov chain $(Z_n)_{n\geq 0}$ with state space $\mathbb{S} = \{1, 2\}$ and transition matrix

$$P = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{bmatrix} 0.4 & 0.6 \\ 0.8 & 0.2 \end{bmatrix} \end{array}.$$

(a) Compute $\mathbb{P}(Z_7 = 1 \text{ and } Z_5 = 2 \mid Z_4 = 1 \text{ and } Z_3 = 2)$.
(b) Compute $\mathbb{E}[Z_2 \mid Z_1 = 1]$.

Exercise 4.3 Consider a transition probability matrix P of the form

$$P = \left[\ P_{i,j}\ \right]_{0 \le i,j \le N} = \begin{bmatrix} \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \end{bmatrix},$$

where $\pi = [\pi_0, \pi_1, \ldots, \pi_N] \in [0,1]^{N+1}$ is a vector such that $\pi_0 + \pi_1 + \cdots + \pi_N = 1$.

(a) Compute P^n for all $n \ge 2$.
(b) Show that the vector π is an invariant (or stationary) distribution for P.
(c) Show that if $\mathbb{P}(Z_0 = i) = \pi_i$, $i = 0, 1, \ldots, N$, then Z_n is independent of Z_k for all $0 \le k < n$, and $(Z_n)_{n\in\mathbb{N}}$ is an *i.i.d* sequence of random variables with distribution $\pi = [\pi_0, \pi_1, \ldots, \pi_N]$ over $\{0, 1, \ldots, N\}$.

Exercise 4.4 Consider a $\{0,1\}$-valued "hidden" two-state Markov chain $(X_n)_{n\in\mathbb{N}}$ with *transition probability matrix*

$$P = \begin{bmatrix} P_{0,0} & P_{0,1} \\ P_{1,0} & P_{1,1} \end{bmatrix} = \begin{bmatrix} \mathbb{P}(X_1 = 0 \mid X_0 = 0) & \mathbb{P}(X_1 = 1 \mid X_0 = 0) \\ \mathbb{P}(X_1 = 0 \mid X_0 = 1) & \mathbb{P}(X_1 = 1 \mid X_0 = 1) \end{bmatrix},$$

and initial distribution

$$\pi = [\pi_0, \pi_1] = [\mathbb{P}(X_0 = 0), \mathbb{P}(X_0 = 1)].$$

We observe a process $(O_k)_{k\in\mathbb{N}}$ whose state $O_k \in \{a, b\}$ at every time $k \in \mathbb{N}$ has a conditional distribution given $X_k \in \{0,1\}$ denoted by

$$M = \begin{bmatrix} m_{0,a} & m_{0,b} \\ m_{1,a} & m_{1,b} \end{bmatrix} = \begin{bmatrix} \mathbb{P}(O_k = a \mid X_k = 0) & \mathbb{P}(O_k = b \mid X_k = 0) \\ \mathbb{P}(O_k = a \mid X_k = 1) & \mathbb{P}(O_k = b \mid X_k = 1) \end{bmatrix},$$

called the *emission probability* matrix.

(a) Using elements of π, P and M, compute $\mathbb{P}(X_0 = 1, X_1 = 1)$ and the probability

$$\mathbb{P}((O_0, O_1) = (a, b) \text{ and } (X_0, X_1) = (1, 1))$$

of observing the sequence $(O_0, O_1) = (a, b)$ when $(X_0, X_1) = (1, 1)$.
Hint: By independence, the conditional probability of observing $(O_0, O_1) = (a, b)$ given that $(X_0, X_1) = (1, 1)$ splits as

$$\mathbb{P}((O_0, O_1) = (a, b) \mid (X_0, X_1) = (1, 1)) = \mathbb{P}(O_0 = a \mid X_0 = 1)\mathbb{P}(O_1 = b \mid X_1 = 1).$$

(b) Find the probability$\mathbb{P}((O_0, O_1) = (a, b))$ that the observed sequence is (a, b).
Hint: Use the law of total probability based on all possible values of (X_0, X_1).

(c) Compute the probabilities

$$\mathbb{P}(X_1 = 1 \mid (O_0, O_1) = (a, b)), \quad \text{and} \quad \mathbb{P}(X_1 = 0 \mid (O_0, O_1) = (a, b)).$$

Exercise 4.5 Consider a two-dimensional random walk $(S_n)_{n\in\mathbb{N}}$ started at $S_0 = (0, 0)$ on $\mathbb{Z}^2$, where, starting from a location $S_n = (i, j)$ the chain can move to any of the points $(i + 1, j + 1)$, $(i + 1, j - 1)$, $(i - 1, j + 1)$, $(i - 1, j - 1)$ with equal probability $1/4$.

(a) Suppose in addition that the random walk cannot visit any site more than once, as in a *snake game*. Is the resulting system a Markov chain? Justify your answer.

(b) Let $S_n = (X_n, Y_n)$ denote the coordinates of S_n at time n and let $Z_n := X_n^2 + Y_n^2$. Is $(Z_n)_{n\in\mathbb{N}}$ a Markov chain? Justify your answer.

Hint: Use the fact that a same value of Z_n may correspond to different locations of (X_n, Y_n) on the circle, for example $(X_n, Y_n) = (5, 0)$ and $(X_n, Y_n) = (4, 3)$ when $Z_n = 25$.

Questions (a) and (b) above are independent.

Exercise 4.6 The Elephant Random Walk $(S_n)_{n\in\mathbb{N}}$ [ST08] is a discrete-time $\mathbb{Z}$-valued random walk

$$S_n := X_1 + \cdots + X_n, \qquad n \in \mathbb{N},$$

whose increments $X_k = S_k - S_{k-1}$, $k \geq 1$, are recursively defined as follows:

- At time $n = 1$, X_1 is a Bernoulli $\{-1, +1\}$-valued random variable with

$$\mathbb{P}(X_1 = +1) = p \quad \text{and} \quad \mathbb{P}(X_1 = -1) = q = 1 - p \in (0, 1).$$

- At any subsequent time $n \geq 2$, one draws randomly an integer time index $k \in \{1, \ldots, n - 1\}$ with uniform probability, and lets $X_n := X_k$ with probability p, and $X_n := -X_k$ with probability $q := 1 - p$.

Does the Elephant Random Walk $(S_n)_{n\in\mathbb{N}}$ have the Markov property?

Exercise 4.7 Consider a Markov chain $(X_n)_{n\geq 0}$ with state space $\mathbb{S} = \{0, 1\}$ and transition matrix

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array}\!\!\begin{array}{c} \begin{array}{cc} 0 & \quad 1 \end{array} \\ \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} \end{array},$$

where $a, b > 0$, and define a new stochastic process $(Z_n)_{n\geq 1}$ by $Z_n = (X_{n-1}, X_n)$, $n \geq 1$. Argue that $(Z_n)_{n\geq 1}$ is a Markov chain and write down its transition matrix. Start by determining the state space of $(Z_n)_{n\geq 1}$.

Exercise 4.8 Given $p \in [0, 1)$, consider the Markov chain $(X_n)_{n\geq 0}$ on the state space $\{0, 1, 2\}$ having the transition matrix

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} p & q & 0 \\ 0 & p & q \\ 0 & 0 & 1 \end{bmatrix} \end{array},$$

with $q := 1 - p$.

(a) Give the probability distribution of the first hitting time

$$T_2 := \inf\big\{n \geq 0 \ : \ X_n = 2\big\}.$$

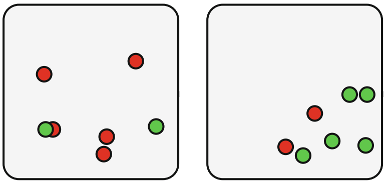
of state 2 starting from $X_0 = 0$.
Hint: The sum $Z = X_1 + \cdots + X_n$ of n independent geometric random variables on $\{1, 2, \ldots\}$ has the negative binomial distribution

$$\mathbb{P}(Z = k \mid X_0 = 1) = \binom{k-1}{k-d}(1-p)^d p^{k-d}, \qquad k \geq d.$$

(b) Compute the mean hitting time $\mathbb{E}[T_2 \mid X_0 = 0]$ of state 2 starting from $X_0 = 0$.
Hint: We have

$$\sum_{k=1}^{\infty} kp^{k-1} = \frac{1}{(1-p)^2} \quad \text{and} \quad \sum_{k=2}^{\infty} k(k-1)p^{k-2} = \frac{2}{(1-p)^3}, \qquad 0 \leq p < 1.$$

Exercise 4.9 Bernoulli–Laplace chain. Consider two boxes and a total of $2N$ balls made of N red balls and N green balls. At time 0, a number $k = X_0$ of red balls and a number $N - k$ of green balls are placed in the first box, while the remaining $N - k$ red balls and k green balls are placed in the second box.



At each unit of time, one ball is chosen randomly out of N in each box, and the two balls are interchanged. Write down the transition matrix of the Markov chain $(X_n)_{n\in\mathbb{N}}$ with state space $\{0, 1, 2, \ldots, N\}$, representing the number of red balls in the first box. Start for example from $N = 5$.

Exercise 4.10 (a) After winning k dollars, a gambler either receives $k + 1$ dollars with probability p, or has to *quit* the game and lose everything with probability $q = 1 - p$. Starting from *one* dollar, find a model for the time evolution of the wealth of the player using a Markov chain whose transition probability matrix P will be described explicitly along with its powers P^n of all orders $n \geq 1$.

(b) (Success runs Markov chain). We modify the model of Question (a) by allowing the gambler to start playing again and win with probability p after reaching state 0. Write down the corresponding transition probability matrix P, and compute P^n for all $n \geq 2$.

Exercise 4.11 Let $(X_k)_{k\in\mathbb{N}}$ be the Markov chain with transition matrix

$$P = \begin{bmatrix} 1/4 & 0 & 1/2 & 1/4 \\ 0 & 1/5 & 0 & 4/5 \\ 0 & 1 & 0 & 0 \\ 1/3 & 1/3 & 0 & 1/3 \end{bmatrix}.$$

A new process is defined by letting

$$Z_n := \begin{cases} 0 & \text{if } X_n = 0 \text{ or } X_n = 1, \\ X_n & \text{if } X_n = 2 \text{ or } X_n = 3, \end{cases}$$

i.e.

$$Z_n = X_n \mathbb{1}_{\{X_n \in \{2,3\}\}}, \qquad n \geq 0.$$

(a) Compute

$$\mathbb{P}(Z_{n+1} = 2 \mid Z_n = 0 \text{ and } Z_{n-1} = 2) \ \text{ and } \ \mathbb{P}(Z_{n+1} = 2 \mid Z_n = 0 \text{ and } Z_{n-1} = 3),$$

$n \geq 1$.

(b) Is $(Z_n)_{n\in\mathbb{N}}$ a Markov chain?

Exercise 4.12 [OSA+09]
Abeokuta, one of the major towns of the defunct Western Region of Nigeria, has recently seen an astronomic increase in vehicular activities. The intensity of vehicle traffic at the Lafenwa intersection which consists of Ayetoro, Old Bridge and Ita-Oshin routes, is modeled according to three states L/M/H = {Low / Moderate / High}.

(a) During year 2005, low intensity incoming traffic has been observed at Lafenwa intersection for $\eta_L = 50\%$ of the time, moderate traffic has been observed for $\eta_M = 40\%$ of the time, while high traffic has been observed during $\eta_H = 10\%$ of the time.

Given the correspondence table

incoming traffic	vehicles per hour
L (low intensity)	360
M (medium intensity)	505
H (high intensity)	640

compute the average incoming traffic per hour in year 2005.

(b) The analysis of incoming daily traffic volumes at Lafenwa intersection between years 2004 and 2005 shows that the probability of switching states within {L, M, H} is given by the Markov transition probability matrix

$$P = \begin{bmatrix} 2/3 & 1/6 & 1/6 \\ 1/3 & 1/2 & 1/6 \\ 1/6 & 2/3 & 1/6 \end{bmatrix}.$$

Based on the knowledge of P and $\eta = [\eta_L, \eta_M, \eta_H]$, give a projection of the respective proportions of traffic in the states L/M/H for year 2006.

(c) Based on the result of Question (b), give a projected estimate for the average incoming traffic per hour in year 2006.

(d) By solving the equation $\pi = \pi P$ for the invariant (or stationary) probability distribution $\pi = [\pi_L, \pi_M, \pi_H]$, give a long term projection of steady traffic at Lafenwa intersection. *Hint*: we have $\pi_L = 11/24$.

Chapter 5
First Step Analysis

Starting with this chapter we introduce the systematic use of the first step analysis technique, in a general framework that covers the examples of random walks already treated in Chaps. 2 and 3. The main applications of first step analysis are the computation of hitting probabilities, mean hitting and absorption times, mean first return times, and average number of returns to a given state.

5.1 Hitting Probabilities

Let us consider a Markov chain $(Z_n)_{n\in\mathbb{N}}$ with state space $\mathbb{S}$, and let $A \subset \mathbb{S}$ denote a subset of $\mathbb{S}$. We are interested in the first time T_A the chain hits the subset A, with

$$T_A = \inf\{n \geq 0 \ : \ Z_n \in A\}, \tag{5.1.1}$$

with $T_A = 0$ if $Z_0 \in A$ and

$$T_A = \infty \quad \text{if} \quad \{n \geq 0 \ : \ Z_n \in A\} = \emptyset,$$

i.e. if $Z_n \notin A$ for all $n \in \mathbb{N}$. Similarly to the gambling problem of Chap. 2, we would like to compute the probabilities

$$g_l(k) = \mathbb{P}(Z_{T_A} = l \mid Z_0 = k)$$

of hitting the set $A \subset \mathbb{S}$ through state $l \in A$ starting from $k \in \mathbb{S}$, where Z_{T_A} represents the location of the chain $(Z_n)_{n\in\mathbb{N}}$ at the hitting time T_A.

This computation can be achieved by first step analysis, using the *law of total probability* (1.3.1) for the probability measure $\mathbb{P}(\cdot \mid Z_0 = k)$ and the Markov property, as follows.

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4_5

For all $k \in \mathbb{S} \setminus A$ we have $T_A \geq 1$ given that $Z_0 = k$, hence we can write

$$\begin{aligned}
g_l(k) &= \mathbb{P}(Z_{T_A} = l \mid Z_0 = k) \\
&= \sum_{m \in \mathbb{S}} \mathbb{P}(Z_{T_A} = l \mid Z_1 = m \text{ and } Z_0 = k)\mathbb{P}(Z_1 = m \mid Z_0 = k) \\
&= \sum_{m \in \mathbb{S}} \mathbb{P}(Z_{T_A} = l \mid Z_1 = m)\mathbb{P}(Z_1 = m \mid Z_0 = k) \\
&= \sum_{m \in \mathbb{S}} P_{k,m}\mathbb{P}(Z_{T_A} = l \mid Z_1 = m) \\
&= \sum_{m \in \mathbb{S}} P_{k,m}\mathbb{P}(Z_{T_A} = l \mid Z_0 = m) \\
&= \sum_{m \in \mathbb{S}} P_{k,m} g_l(m), \qquad k \in \mathbb{S} \setminus A, \quad l \in A,
\end{aligned}$$

where the relation

$$\mathbb{P}(Z_{T_A} = l \mid Z_1 = m) = \mathbb{P}(Z_{T_A} = l \mid Z_0 = m)$$

follows from the fact that the probability of ruin does not depend on the initial time the counter is started, as in Lemma 2.2.

Hence we have

$$g_l(k) = \sum_{m \in \mathbb{S}} P_{k,m} g_l(m) = P_{k,l} + \sum_{m \in \mathbb{S} \setminus A} P_{k,m} g_l(m), \tag{5.1.2}$$

$k \in \mathbb{S}$, $l \in A$, under the boundary conditions

$$g_l(k) = \mathbb{P}(Z_{T_A} = l \mid Z_0 = k) = \mathbb{1}_{\{k=l\}} = \begin{cases} 1 \text{ if } k = l, \\ 0 \text{ if } k \neq l, \end{cases} \qquad k \in A, \ l \in S,$$

since $T_A = 0$ whenever one starts from $Z_0 \in A$. Equation (5.1.2) can be rewritten in matrix form as

$$g_l = P g_l, \qquad l \in A, \tag{5.1.3}$$

where g is a column vector, under the boundary condition

$$g_l(k) = \mathbb{P}(Z_{T_A} = l \mid Z_0 = k) = \mathbb{1}_{\{l\}}(k) = \begin{cases} 1, \ k = l, \\ 0, \ k \neq l, \end{cases}$$

for all $k \in A$ *and* $l \in \mathbb{S}$.

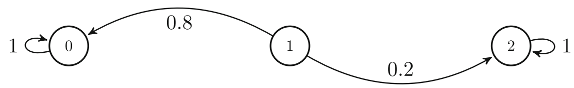
In addition, the hitting probabilities $g_l(k) = \mathbb{P}(Z_{T_A} = l \mid Z_0 = k)$ satisfy the condition

$$\begin{aligned} 1 &= \mathbb{P}(T_A = \infty \mid Z_0 = k) + \sum_{l \in A} \mathbb{P}(Z_{T_A} = l \mid Z_0 = k) \\ &= \mathbb{P}(T_A = \infty \mid Z_0 = k) + \sum_{l \in A} g_l(k), \end{aligned} \tag{5.1.4}$$

for all $k \in \mathbb{S}$.

Note that we may have $\mathbb{P}(T_A = \infty \mid Z_0 = k) > 0$, for example in the following chain with $A = \{0\}$ and $k = 1$ we have

$$\mathbb{P}(T_0 = \infty \mid Z_0 = 1) = 0.2.$$



In case the transition matrix P satisfies

$$P_{k,l} = \mathbb{1}_{\{k=l\}}$$

for all $k, l \in A$, the set A is said to be *absorbing*.

The next lemma will be used in Chap. 8 on branching Processes.

Lemma 5.1 *Assume that state* $j \in \mathbb{S}$ *is absorbing. Then for all* $i \in \mathbb{S}$ *we have*

$$\mathbb{P}(T_j < \infty \mid Z_0 = i) = \lim_{n \to \infty} \mathbb{P}(Z_n = j \mid Z_0 = i).$$

Proof We have

$$\{T_j < \infty\} = \bigcup_{n \geq 1} \{Z_n = j\},$$

because the finiteness of T_j means that Z_n becomes equal to j for some $n \in \mathbb{N}$. In addition, since $j \in \mathbb{S}$ is absorbing it holds that

$$\{Z_n = j\} \subset \{Z_{n+1} = j\}, \qquad n \in \mathbb{N},$$

hence given that $\{Z_0 = i\}$, by (1.2.3) we have

$$\begin{aligned} \alpha_1 &= \mathbb{P}(T_j < \infty \mid Z_0 = i) = \mathbb{P}\left(\bigcup_{n \geq 1} \{Z_n = j\} \,\middle|\, Z_0 = i \right) \\ &= \mathbb{P}\left(\lim_{n \to \infty} \{Z_n = j\} \,\middle|\, Z_0 = i \right) = \lim_{n \to \infty} \mathbb{P}(\{Z_n = j\} \mid Z_0 = i). \end{aligned} \tag{5.1.5}$$

□

Block Triangular Transition Matrices

Assume now that the state space is $\mathbb{S} = \{0, 1, \dots, N\}$ and the transition matrix P has the form

$$P = \begin{bmatrix} Q & R \\ & \\ 0 & I_d \end{bmatrix}, \tag{5.1.6}$$

where Q is a square $(r+1) \times (r+1)$ matrix, R is a $(r+1) \times (N-r)$ matrix, and

$$\mathrm{I_d} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

is the $(N-r) \times (N-r)$ identity matrix, in which case the states in $\{r+1, r+2, \dots, N\}$ are *absorbing*.

If the set $A := \{r+1, r+2, \dots, N\}$ is made of the absorbing states of the chain, we have the boundary conditions

$$g_l(m) = \mathbb{1}_{\{m=l\}}, \qquad l = 0, 1, \dots, N, \quad m = r+1, r+2, \dots, N, \tag{5.1.7}$$

hence the Eq. (5.1.2) can be rewritten as

$$\begin{aligned} g_l(k) &= \sum_{m=0}^{N} P_{k,m} g_l(m) \\ &= \sum_{m=0}^{r} P_{k,m} g_l(m) + \sum_{m=r+1}^{N} P_{k,m} g_l(m) \\ &= \sum_{m=0}^{r} P_{k,m} g_l(m) + P_{k,l} \\ &= \sum_{m=0}^{r} Q_{k,m} g_l(m) + R_{k,l}, \qquad k = 0, 1, \dots, r, \quad l = r+1, \dots, N, \end{aligned}$$

from (5.1.7) and since $P_{k,l} = R_{k,l}$, $k = 0, 1, \dots, r$, $l = r+1, \dots, N$. Hence we have

$$g_l(k) = \sum_{m=0}^{r} Q_{k,m} g_l(m) + R_{k,l}, \qquad k = 0, 1, \dots, r, \quad l = r+1, \dots, N.$$

Remark In the case of the two-state Markov chain with transition matrix (4.5.1) and $A = \{0\}$ we simply find $g_0(0) = 1$ and

$$g_0(1) = b + (1 - b) \times g_0(1),$$

hence $g_0(1) = 1$ if $b > 0$ and $g_0(1) = 0$ if $b = 0$.

Examples Consider a Markov chain on $\{0, 1, 2, 3\}$ with transition matrix of the form

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & b & c & d \\ \alpha & \beta & \gamma & \eta \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.1.8}$$

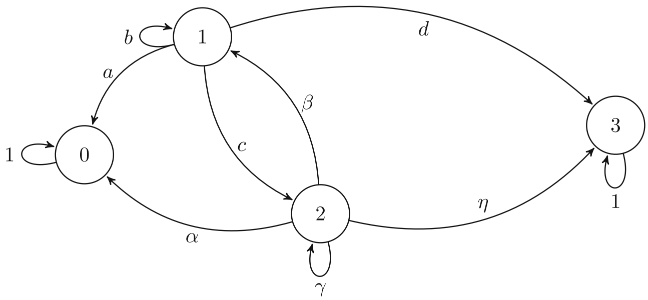
Let $A = \{0, 3\}$ denote the absorbing states of the chain, and let

$$T_{0,3} = \inf\{n \geq 0 \ : \ X_n = 0 \text{ or } X_n = 3\}$$

and compute the probabilities

$$g_0(k) = \mathbb{P}(X_{T_{0,3}} = 0 \mid X_0 = k)$$

of hitting state ⓪ first within $\{0, 3\}$ starting from $k = 0, 1, 2, 3$. The chain has the following graph:



(5.1.9)

Noting that ⓪ and ③ are absorbing states, and writing the relevant rows of the first step analysis matrix equation $g = Pg$, we have

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = a \times 1 + b g_0(1) + c g_0(2) + d \times 0 \\ g_0(2) = \alpha \times 1 + \beta g_0(1) + \gamma g_0(2) + \eta \times 0 \\ g_0(3) = 0, \end{cases}$$

i.e.

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = a + bg_0(1) + cg_0(2) \\ g_0(2) = \alpha + \beta g_0(1) + \gamma g_0(2) \\ g_0(3) = 0, \end{cases}$$

which has for solution

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = \dfrac{c\alpha + a(1-\gamma)}{(1-b)(1-\gamma) - c\beta} \\ g_0(2) = \dfrac{a\beta + \alpha(1-b)}{(1-b)(1-\gamma) - c\beta} \\ g_0(3) = 0. \end{cases} \tag{5.1.10}$$

We have $g_l(0) = g_l(3) = 0$ for $l = 1, 2$, and by a similar analysis, letting

$$g_3(k) := \mathbb{P}(X_{T_{0,3}} = 3 \mid X_0 = k), \qquad k = 0, 1, 2, 3,$$

we find

$$\begin{cases} g_3(0) = 0 \\ g_3(1) = \dfrac{c\eta + d(1-\gamma)}{(1-b)(1-\gamma) - c\beta} \\ g_3(2) = \dfrac{\beta d + \eta(1-b)}{(1-b)(1-\gamma) - c\beta} \\ g_3(3) = 1, \end{cases}$$

and we note that

$$g_0(1) + g_3(1) = \frac{c\alpha + a(1-\gamma)}{(1-b)(1-\gamma) - c\beta} + \frac{c\eta + d(1-\gamma)}{(1-b)(1-\gamma) - c\beta} = 1,$$

since $\alpha + \eta = 1 - \gamma - \beta$ and $a + d = 1 - b - c$, and similarly

$$g_0(2) + g_3(2) = \frac{a\beta + \alpha(1-b)}{(1-b)(1-\gamma) - c\beta} + \frac{\beta d + \eta(1-b)}{(1-b)(1-\gamma) - c\beta} = 1.$$

We also check that in case $a = d$ and $\alpha = \eta$ we have

$$g_0(1) = \frac{c\alpha + a(\beta + 2\alpha)}{(c + 2a)(\beta + 2\alpha) - c\beta} = \frac{c\alpha + a\beta + 2a\alpha}{2c\alpha + 2a\beta + 4a\alpha} = g_0(2) = \frac{1}{2},$$

and

$$g_0(1) = g_3(1) = g_0(2) = g_3(2) = \frac{1}{2}.$$

Note that, letting

$$T_0 := \inf\{n \geq 0 \ : \ X_n = 0\} \quad \text{and} \quad T_3 := \inf\{n \geq 0 \ : \ X_n = 0\},$$

we also have

$$g_0(k) = \mathbb{P}(X_{T_{0,3}} = 0 \mid X_0 = k) = \mathbb{P}(T_0 < \infty \mid X_0 = k)$$

and

$$g_3(k) = \mathbb{P}(X_{T_{0,3}} = 3 \mid X_0 = k) = \mathbb{P}(T_3 < \infty \mid X_0 = k)$$

$k = 0, 1, 2, 3$.

5.2 Mean Hitting and Absorption Times

We are now interested in the mean hitting time

$$h_A(k) := \mathbb{E}[T_A \mid Z_0 = k]$$

it takes for the chain to hit the set $A \subset \mathbb{S}$ starting from a state $k \in \mathbb{S}$. In case the set A is absorbing we refer to $h_A(k)$ as the *mean absorption time* into A starting from the state k.

Clearly, since $T_A = 0$ whenever $X_0 = k \in A$, we have

$$h_A(k) = 0, \qquad \text{for all } k \in A.$$

In addition, for all $k \in \mathbb{S} \setminus A$, by first step analysis using the *law of total expectation* (1.6.11) applied to the probability measure $\mathbb{P}(\cdot \mid Z_0 = l)$, the Markov property and Lemma 1.4 we have

$$\begin{aligned}
h_A(k) &= \mathbb{E}[T_A \mid Z_0 = k] \\
&= \sum_{l \in \mathbb{S}} \mathbb{E}\big[T_A \mathbb{1}_{\{Z_1 = l\}} \mid Z_0 = k\big] \\
&= \frac{1}{\mathbb{P}(Z_0 = k)} \sum_{l \in \mathbb{S}} \mathbb{E}\big[T_A \mathbb{1}_{\{Z_1 = l\}} \mathbb{1}_{\{Z_0 = k\}}\big] \\
&= \frac{1}{\mathbb{P}(Z_0 = k)} \sum_{l \in \mathbb{S}} \mathbb{E}\big[T_A \mathbb{1}_{\{Z_1 = l \text{ and } Z_0 = k\}}\big] \\
&= \sum_{l \in \mathbb{S}} \mathbb{E}\big[T_A \mid Z_1 = l \text{ and } Z_0 = k\big] \frac{\mathbb{P}(Z_1 = l \text{ and } Z_0 = k)}{\mathbb{P}(Z_0 = k)}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{l \in \mathbb{S}} \mathbb{E}[T_A \mid Z_1 = l \text{ and } Z_0 = k]\mathbb{P}(Z_1 = l \mid Z_0 = k) \\
&= \sum_{l \in \mathbb{S}} \mathbb{E}[1 + T_A \mid Z_0 = l]\mathbb{P}(Z_1 = l \mid Z_0 = k) \\
&= \sum_{l \in \mathbb{S}} (1 + \mathbb{E}[T_A \mid Z_0 = l])\mathbb{P}(Z_1 = l \mid Z_0 = k) \\
&= \sum_{l \in \mathbb{S}} \mathbb{P}(Z_1 = l \mid Z_0 = k) + \sum_{l \in \mathbb{S}} \mathbb{P}(Z_1 = l \mid Z_0 = k)\mathbb{E}[T_A \mid Z_0 = l] \\
&= 1 + \sum_{l \in \mathbb{S}} \mathbb{P}(Z_1 = l \mid Z_0 = k)\mathbb{E}[T_A \mid Z_0 = l] \\
&= 1 + \sum_{l \in \mathbb{S}} P_{k,l} h_A(l), \qquad k \in \mathbb{S} \setminus A,
\end{aligned}
$$

where the relation

$$\mathbb{E}[T_A \mid Z_1 = l,\ Z_0 = k] = 1 + \mathbb{E}[T_A \mid Z_0 = l]$$

can be justified as in the proof of Lemma 2.3.

Hence we have

$$h_A(k) = 1 + \sum_{l \in \mathbb{S}} P_{k,l} h_A(l) = 1 + \sum_{l \in \mathbb{S} \setminus A} P_{k,l} h_A(l), \quad k \in \mathbb{S} \setminus A, \tag{5.2.1}$$

under the boundary conditions

$$h_A(k) = \mathbb{E}[T_A \mid Z_0 = k] = 0, \qquad k \in A, \tag{5.2.2}$$

Condition (5.2.2) implies that (5.2.1) becomes

$$h_A(k) = 1 + \sum_{l \in \mathbb{S} \setminus A} P_{k,l} h_A(l), \qquad k \in \mathbb{S} \setminus A.$$

This equation can be rewritten in matrix form as

$$h_A = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + P h_A,$$

by considering only the rows with index $k \in \mathbb{S} \setminus A$, under the boundary conditions

$$h_A(k) = 0, \qquad k \in A.$$

Block Triangular Transition Matrices

When the transition matrix P has the form (5.1.6) and $A = \{r+1, r+2, \ldots, N\}$, Eq. (5.2.1) rewrites as

$$\begin{aligned} h_A(k) &= 1 + \sum_{l=0}^{N} P_{k,l} h_A(l) \\ &= 1 + \sum_{l=0}^{r} P_{k,l} h_A(l) + \sum_{l=r+1}^{N} P_{k,l} h_A(l) \\ &= 1 + \sum_{l=0}^{r} P_{k,l} h_A(l), \qquad 0 \le k \le r, \end{aligned}$$

since $h_A(l) = 0$, $l = r+1, r+2, \ldots, N$, i.e.

$$h_A(k) = 1 + \sum_{l=0}^{r} P_{k,l} h_A(l), \qquad 0 \le k \le r,$$

with $h_A(k) = 0$, $k = r+1, \ldots, n$.

Two-State Chain

In the case of the two-state Markov chain with transition matrix (4.5.1) with $A = \{0\}$ we simply find $h_{\{0\}}(0) = 0$ and

$$h_{\{0\}}(1) = b \times 1 + (1-b)(1 + h_{\{0\}}(1)) = 1 + (1-b) h_{\{0\}}(1), \tag{5.2.3}$$

with solution

$$h_{\{0\}}(1) = b \sum_{k=1}^{\infty} k(1-b)^k = \frac{1}{b},$$

and similarly we find

$$h_{\{1\}}(0) = a \sum_{k=1}^{\infty} k(1-a)^k = \frac{1}{a},$$

with $h_{\{0\}}(0) = h_{\{1\}}(1) = 0$, cf. also (5.3.3) below.

Utility Functionals

The above can be generalized to derive an equation for an expectation of the form

$$h_A(k) := \mathbb{E}\left[\sum_{i=0}^{T_A} f(X_i) \,\Big|\, X_0 = k\right], \qquad k = 0, 1, \ldots, N,$$

where $f(\cdot)$ is a given utility function, as follows:

$$\begin{aligned}
h_A(k) &= \mathbb{E}\left[\sum_{i=0}^{T_A} f(X_i) \,\Big|\, X_0 = k\right] \\
&= \sum_{m=0}^{r} P_{k,m}\left(f(k) + \mathbb{E}\left[\sum_{i=1}^{T_A} f(X_i) \,\Big|\, X_1 = m\right]\right) \\
&= \sum_{m=0}^{r} P_{k,m} f(k) + \sum_{m=0}^{r} P_{k,m}\mathbb{E}\left[\sum_{i=1}^{T_A} f(X_i) \,\Big|\, X_1 = m\right] \\
&= f(k)\sum_{m=0}^{r} P_{k,m} + \sum_{m=0}^{r} P_{k,m}\mathbb{E}\left[\sum_{i=0}^{T_A} f(X_i) \,\Big|\, X_0 = m\right] \\
&= f(k) + \sum_{m=0}^{r} P_{k,m} h_A(m), \qquad k \in A^c := \{0, 1, \ldots, r\},
\end{aligned}$$

with $A := \{r+1, \ldots, N\}$, hence

$$h_A(k) = f(k) + \sum_{m=0}^{r} P_{k,m} h_A(m), \qquad k \in A^c = \{0, 1, \ldots, r\},$$

with the boundary condition

$$h_A(k) = 0, \qquad k \in A = \{r+1, \ldots, n\},$$

see also Exercise 5.20.

Examples

- When $f = \mathbb{1}_{A^c} = \mathbb{1}_{\{0,1,\ldots,r\}}$ is the indicator function over the set A^c, i.e.

$$f(X_i) = \mathbb{1}_{A^c}(X_i) = \begin{cases} 1 & \text{if } X_i \notin A, \\ \\ 0 & \text{if } X_i \in A, \end{cases}$$

  the quantity $h_A(k)$ coincides with the mean hitting time of the set A starting from k. In particular, when $A = \{m\}$ this recovers the equation

$$h_{\{m\}}(k) = 1 + \sum_{\substack{l \in \mathbb{S} \\ l \neq m}} P_{k,l} h_{\{m\}}(l), \qquad k \in \mathbb{S} \setminus \{m\}, \tag{5.2.4}$$

with $h_{\{m\}}(m) = 1$.

- When f is the indicator function $f = \mathbb{1}_{\{l\}}$, i.e.

$$f(X_i) = \mathbb{1}_{\{l\}}(X_i) = \begin{cases} 1 & \text{if } X_i = l, \\ \\ 0 & \text{if } X_i \neq l, \end{cases}$$

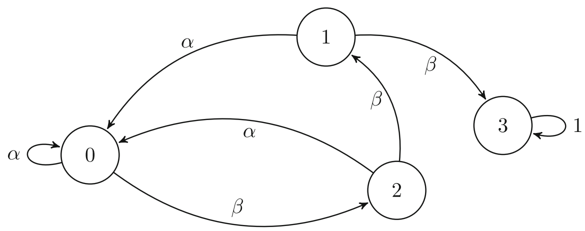
with $l \in A^c$, the quantity $h_A(k)$ will yield the mean number of visits to state ⓛ starting from ⓚ before hitting the set A.

- See Exercises 5.19, 5.20, and also Problem 5.22 for a complete solution in case $f(k) = k$ and $(X_k)_{k \geq 0}$ is the gambling process of Chap. 2.

Examples Consider the Markov chain whose transition probability matrix is given by

$$P = \left[P_{i,j} \right]_{0 \leq i,j \leq 3} = \begin{bmatrix} \alpha & 0 & \beta & 0 \\ \alpha & 0 & 0 & \beta \\ \alpha & \beta & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$. Taking $A := \{3\}$, determine the mean time it takes to reach state ③ starting from state ⓪. We observe that state ③ is absorbing:



Let

$$h_3(k) = \mathbb{E}[T_3 \mid X_0 = k]$$

denote the mean (hitting) time to reach state ③, after starting from state $k = 0, 1, 2, 3$. We get

$$\begin{cases} h_3(0) = \alpha(1 + h_3(0)) + \beta(1 + h_3(2)) = 1 + \alpha h_3(0) + \beta h_3(2) \\ h_3(1) = \alpha(1 + h_3(0)) + \beta(1 + h_3(3)) = 1 + \alpha h_3(0) \\ h_3(2) = \alpha(1 + h_3(0)) + \beta(1 + h_3(1)) = 1 + \alpha h_3(0) + \beta h_3(1) \\ h_3(3) = 0, \end{cases}$$

which, using the relation $\alpha = 1 - \beta$, yields

$$h_3(3) = 0, \qquad h_3(1) = \frac{1}{\beta^3}, \qquad h_3(2) = \frac{1+\beta}{\beta^3}, \qquad h_3(0) = \frac{1+\beta+\beta^2}{\beta^3}.$$

Since state 3 can only be reached from state 1 with probability β, it is natural that the hitting times go to infinity as β goes to zero. We also check that $h_3(3) < h_3(1) < h_3(2) < h_3(0)$, as can be expected from the above graph. In addition, $(h_3(1), h_3(2), h_3(0))$ converge to $(1, 2, 3)$ as β goes to 1, as can be expected.

5.3 First Return Times

Consider now the first *return* time T_j^r to state $j \in \mathbb{S}$, defined by

$$T_j^r := \inf\{n \geq 1 \ : \ X_n = j\},$$

with

$$T_j^r = \infty \quad \text{if } X_n \neq j \text{ for all } n \geq 1.$$

Note that in contrast with the definition (5.1.1) of the hitting time T_j, the infimum is taken here for $n \geq 1$ as it takes at least one step out of the initial state in order to *return* to state j. Nevertheless we have $T_j = T_j^r$ if the chain is started from a state i different from j.

Denote by

$$\mu_j(i) = \mathbb{E}[T_j^r \mid X_0 = i] \geq 1$$

the *mean return time* to state $j \in \mathbb{S}$ after starting from state $i \in \mathbb{S}$.

Mean return times can also be computed by first step analysis. We have

$$\begin{aligned} \mu_j(i) &= \mathbb{E}[T_j^r \mid X_0 = i] \\ &= 1 \times \mathbb{P}(X_1 = j \mid X_0 = i) \\ &\quad + \sum_{\substack{l \in \mathbb{S} \\ l \neq j}} \mathbb{P}(X_1 = l \mid X_0 = i)(1 + \mathbb{E}[T_j^r \mid X_0 = l]) \end{aligned}$$

$$
\begin{aligned}
&= P_{i,j} + \sum_{\substack{l \in \mathbb{S} \\ l \neq j}} P_{i,l}(1 + \mu_j(l)) \\
&= P_{i,j} + \sum_{\substack{l \in \mathbb{S} \\ l \neq j}} P_{i,l} + \sum_{\substack{l \in \mathbb{S} \\ l \neq j}} P_{i,l}\mu_j(l) \\
&= \sum_{l \in \mathbb{S}} P_{i,l} + \sum_{\substack{l \in \mathbb{S} \\ l \neq j}} P_{i,l}\mu_j(l) \\
&= 1 + \sum_{\substack{l \in \mathbb{S} \\ l \neq j}} P_{i,l}\mu_j(l),
\end{aligned}
$$

hence

$$\mu_j(i) = 1 + \sum_{\substack{l \in \mathbb{S} \\ l \neq j}} P_{i,l}\mu_j(l), \qquad i, j \in \mathbb{S}. \tag{5.3.1}$$

Hitting Times *Versus* Return Times

Note that the mean return time equation in (5.3.1) does not include any boundary condition, in contrast with the mean hitting time Eq. (5.2.4) in Sect. 5.2. In addition, the time T_i^r to return to state (i) is always at least one by construction, hence $\mu_i(i) \geq 1$ cannot vanish, while we always have $h_i(i) = 0$, $i \in \mathbb{S}$. On the other hand, by definition we have

$$h_i(j) = \mathbb{E}[T_i^r \mid X_0 = j] = \mathbb{E}[T_i \mid X_0 = j] = \mu_i(j),$$

for all $i \neq j$, and for $i = j$ the mean return time $\mu_j(j)$ can be computed from the hitting times $h_j(l)$, $l \neq j$, by first step analysis as

$$
\begin{aligned}
\mu_j(j) &= \sum_{l \in S} P_{j,l}(1 + h_j(l)) \\
&= P_{j,j} + \sum_{l \neq j} P_{j,l}(1 + h_j(l)) \\
&= \sum_{l \in S} P_{j,l} + \sum_{l \neq j} P_{j,l}h_j(l) \\
&= 1 + \sum_{l \neq j} P_{j,l}h_j(l), \qquad j \in \mathbb{S},
\end{aligned} \tag{5.3.2}
$$

which in agreement with (5.3.1) when $i = j$.

In practice we may prefer to compute first the hitting times $h_i(j) = 0$ under the boundary conditions $h_i(i) = 0$, and then to recover the return time $\mu_i(i)$ from (5.3.2), $i, j \in \mathbb{S}$.

Examples

(i) *Mean return times for the two-state Markov chain.*
The mean return time $\mu_0(i) = \mathbb{E}[T_0^r \mid X_0 = i]$ to state ⓪ starting from state ⓘ $\in \{0, 1\}$ satisfies

$$\begin{cases} \mu_0(0) = (1-a) \times 1 + a(1+\mu_0(1)) = 1 + a\mu_0(1) \\ \\ \mu_0(1) = b \times 1 + (1-b)(1+\mu_0(1)) = 1 + (1-b)\mu_0(1) \end{cases}$$

which yields

$$\mu_0(0) = 1 + \frac{a}{b} \quad \text{and} \quad \mu_0(1) = h_0(1) = \frac{1}{b}, \tag{5.3.3}$$

cf. also (5.2.3) above for the computation of $\mu_0(1) = h_0(1) = 1/b$ as a mean hitting time. In the two-state case, the distribution of T_0^r given $X_0 = 0$ is given by

$$f_{0,0}^{(n)} := \mathbb{P}(T_0^r = n \mid X_0 = 0) = \begin{cases} 0 & \text{if } n = 0, \\ \\ 1-a & \text{if } n = 1, \\ \\ ab(1-b)^{n-2} & \text{if } n \geq 2, \end{cases} \tag{5.3.4}$$

hence (5.3.3) can be directly recovered as[1]

$$\begin{aligned} \mu_0(0) &= \mathbb{E}[T_0^r \mid X_0 = 0] \\ &= \sum_{n=0}^{\infty} n\mathbb{P}(T_0^r = n \mid X_0 = 0) \\ &= \sum_{n=0}^{\infty} n f_{0,0}^{(n)} \\ &= 1 - a + ab\sum_{n=2}^{\infty} n(1-b)^{n-2} \\ &= 1 - a + ab\sum_{n=0}^{\infty} (n+2)(1-b)^{n} \\ &= 1 - a + ab(1-b)\sum_{n=0}^{\infty} n(1-b)^{n-1} + 2ab\sum_{n=0}^{\infty}(1-b)^n \\ &= \frac{a+b}{b} = 1 + \frac{a}{b}, \end{aligned} \tag{5.3.5}$$

[1] We are using the identities $\sum_{k=0}^{\infty} r^k = (1-r)^{-1}$ and $\sum_{k=1}^{\infty} kr^{k-1} = (1-r)^{-2}$, cf. (A.3) and (A.4).

where we used the identity (A.4).
Similarly we check that

$$\begin{cases} \mu_1(0) = 1 + (1-a)\mu_1(0) \\ \\ \mu_1(1) = 1 + b\mu_1(0), \end{cases}$$

which yields

$$\mu_1(0) = h_1(0) = \frac{1}{a} \quad \text{and} \quad \mu_1(1) = 1 + \frac{b}{a},$$

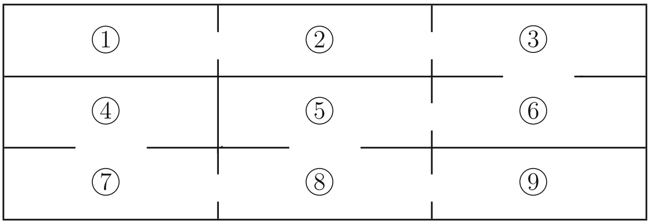
and can be directly recovered by

$$\mu_1(1) = 1 - b + ab \sum_{n=0}^{\infty} (n+2)(1-a)^n = \frac{a+b}{a} = 1 + \frac{b}{a}, \tag{5.3.6}$$

as in (5.3.3) and (5.3.5) above, by swapping a with b and state ⓪ with state ①.

(ii) *Maze problem.*
Mazes provide natural examples of Markovian systems as their users tend rely on their current positions and to forget past information. More generally, Markovian systems can be used as an approximation of a non-Markovian reality.

Consider a fish placed in an aquarium with 9 compartments:



(5.3.7)

The fish moves randomly: at each time step it changes compartments and if it finds $k \geq 1$ exit doors from one compartment, it will choose one of them with probability $1/k$, i.e. the transition matrix is

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Find the average time to come back to state ① starting from state ①.

Letting

$$T_l^r = \inf\{n \geq 1 \ : \ X_n = l\}$$

denote the first return time to state ①, and defining

$$\mu_1(k) := \mathbb{E}[T_1^r \mid X_0 = k]$$

the mean return time to state ① starting from ⓚ, we have

$$\begin{cases} \mu_1(1) = 1 + \mu_1(2) \\ \\ \mu_1(2) = \frac{1}{2}(1+0) + \frac{1}{2}(1+\mu_1(3)) = 1 + \frac{1}{2}\mu_1(3) \\ \\ \mu_1(3) = \frac{1}{2}(1+\mu_1(2)) + \frac{1}{2}(1+\mu_1(6)) = 1 + \frac{1}{2}\mu_1(2) + \frac{1}{2}\mu_1(6) \\ \\ \mu_1(4) = 1 + \mu_1(7) \\ \\ \mu_1(5) = \frac{1}{2}(1+\mu_1(8)) + \frac{1}{2}(1+\mu_1(6)) = 1 + \frac{1}{2}\mu_1(8) + \frac{1}{2}\mu_1(6) \\ \\ \mu_1(6) = \frac{1}{2}(1+\mu_1(3)) + \frac{1}{2}(1+\mu_1(5)) = 1 + \frac{1}{2}\mu_1(3) + \frac{1}{2}\mu_1(5) \\ \\ \mu_1(7) = 1 + \frac{1}{2}\mu_1(4) + \frac{1}{2}\mu_1(8) = \frac{1}{2}(1+\mu_1(4)) + \frac{1}{2}(1+\mu_1(8)) \\ \\ \mu_1(8) = \frac{1}{3}(1+\mu_1(7)) + \frac{1}{3}(1+\mu_1(5)) + \frac{1}{3}(1+\mu_1(9)) \\ \\ \qquad\quad = 1 + \frac{1}{3}\left(\mu_1(7) + \mu_1(5) + \mu_1(9)\right) \\ \\ \mu_1(9) = 1 + \mu_1(8), \end{cases}$$

or

$$\begin{cases} \mu_1(1) = 1 + \mu_1(2), \quad \mu_1(2) = 1 + \frac{1}{2}\mu_1(3), \quad \mu_1(3) = 2 + \frac{2}{3}\mu_1(6), \\ \\ \mu_1(4) = 1 + \mu_1(7), \quad \mu_1(5) = 1 + \frac{1}{2}\mu_1(8) + \frac{1}{2}\mu_1(6), \quad 0 = 30 + 3\mu_1(8) - 5\mu_1(6), \\ \\ \mu_1(7) = 3 + \mu_1(8), \quad 0 = 80 + 5\mu_1(6) - 5\mu_1(8), \quad \mu_1(9) = 1 + \mu_1(8), \end{cases}$$

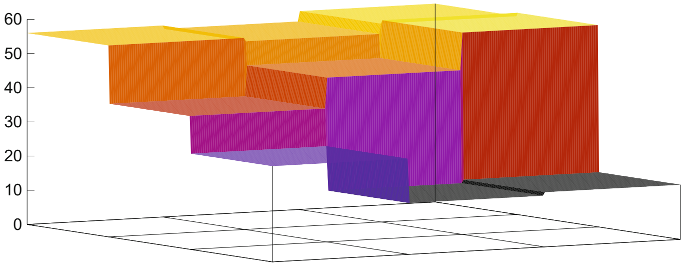


Fig. 5.1 Mean return times to state 0 on the maze (5.3.7)

which yields

$$\mu_1(1) = 16, \ \mu_1(2) = 15, \ \mu_1(3) = 28, \ \mu_1(4) = 59, \ \mu_1(5) = 48, \qquad (5.3.8)$$
$$\mu_1(6) = 39, \ \mu_1(7) = 58, \ \mu_1(8) = 55, \ \mu_1(9) = 56.$$

Consequently, it takes on average 16 steps to come back to 1 starting from 1, and 59 steps to reach 1 starting from 4. This data is illustrated in the following picture in which the numbers represent the average time it takes to return to 1 starting from a given state.

$\mu_1(1) = 16$	$\mu_1(2) = 15$	$\mu_1(3) = 28$
$\mu_1(4) = 59$	$\mu_1(5) = 48$	$\mu_1(6) = 39$
$\mu_1(7) = 58$	$\mu_1(8) = 55$	$\mu_1(9) = 56$

The next Fig. 5.1 represents the mean return times to state 0 according to the initial state on the maze (5.3.7).

5.4 Mean Number of Returns

Return Probabilities

In the sequel we let

$$p_{ij} = \mathbb{P}(T_j^r < \infty \mid X_0 = i) = \mathbb{P}(X_n = j \text{ for some } n \geq 1 \mid X_0 = i)$$

denote the probability of return to state (j) in finite time[2] starting from state (i). The probability p_{ii} of return to state (i) within a finite time starting from state (i) can be computed as follows:

$$\begin{aligned}
p_{ii} &= \mathbb{P}(X_n = i \text{ for some } n \geq 1 \mid X_0 = i) \\
&= \mathbb{P}\Big(\bigcup_{n\geq 1}\{X_n = i\}\Big| X_0 = i\Big) \\
&= \sum_{n=1}^{\infty} \mathbb{P}(X_n = i, X_{n-1} \neq i, \ldots, X_1 \neq i \mid X_0 = i) \\
&= \sum_{n=1}^{\infty} f_{i,i}^{(n)},
\end{aligned} \tag{5.4.1}$$

where

$$f_{i,j}^{(n)} := \mathbb{P}(T_j^r = n \mid X_0 = i) = \mathbb{P}(X_n = j, X_{n-1} \neq j, \ldots, X_1 \neq j \mid X_0 = i),$$

$i, j \in \mathbb{S}$, is the probability distribution of T_j^r given that $X_0 = i$, with

$$f_{i,i}^{(0)} = \mathbb{P}(T_i^r = 0 \mid X_0 = i) = 0.$$

Note that we have

$$f_{i,i}^{(1)} = \mathbb{P}(X_1 = i \mid X_0 = i) = P_{i,i}, \qquad i \in \mathbb{S}.$$

Convolution Equation

By conditioning on the first return time $k \geq 1$, the return time probability distribution $f_{i,i}^{(k)} = \mathbb{P}(T_i^r = k \mid X_0 = i)$ satisfies the convolution equation

$$\begin{aligned}
[P^n]_{i,i} &= \mathbb{P}(X_n = i \mid X_0 = i) \\
&= \sum_{k=1}^{n} \mathbb{P}(X_k = i, X_{k-1} \neq i, \ldots, X_1 \neq i \mid X_0 = i)\mathbb{P}(X_n = i \mid X_k = i) \\
&= \sum_{k=1}^{n} \mathbb{P}(X_k = i, X_{k-1} \neq i, \ldots, X_1 \neq i \mid X_0 = i)\mathbb{P}(X_{n-k} = i \mid X_0 = i) \\
&= \sum_{k=1}^{n} f_{i,i}^{(k)} [P^{n-k}]_{i,i},
\end{aligned}$$

which extends the convolution equation (3.4.7) from random walks to the more general setting of Markov chains.

[2] When $(i) \neq (j)$, p_{ij} is the probability of *visiting* state (j) in finite time after starting from state (i).

The return probabilities p_{ij} will be used below to compute the average number of returns to a given state, and the distribution $f_{i,j}^{(k)}$, $k \geq 1$, of T_j^r given that $X_0 = i$ will be useful in Sect. 6.4 on positive and null recurrence.

Number of Returns

Let

$$R_j := \sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=j\}} \tag{5.4.2}$$

denote the number of returns[3] to state j by the chain $(X_n)_{n\in\mathbb{N}}$. The next proposition shows that, given $\{X_0 = i\}$, R_j has a zero-modified geometric distribution with initial mass $1 - p_{ij}$.

The next proposition shows that, given $\{X_0 = i\}$, R_j has a zero-modified geometric distribution with initial mass $1 - p_{ij}$.

Proposition 5.1 *The probability distribution of the number of returns R_j to state j given that $\{X_0 = i\}$ is given by*

$$\mathbb{P}(R_j = m \mid X_0 = i) = \begin{cases} 1 - p_{ij}, & m = 0, \\ \\ p_{ij} \times (p_{jj})^{m-1} \times (1 - p_{jj}), & m \geq 1, \end{cases}$$

Proof When the chain never visits state j starting from $X_0 = i$ we have $R_j = 0$, and this happens with probability

$$\begin{aligned} \mathbb{P}(R_j = 0 \mid X_0 = i) &= \mathbb{P}(T_j^r = \infty \mid X_0 = i) \\ &= 1 - \mathbb{P}(T_j^r < \infty \mid X_0 = i) \\ &= 1 - p_{ij}. \end{aligned}$$

Next, when the chain $(X_n)_{n\in\mathbb{N}}$ makes a number $R_j = m \geq 1$ of visits to state j starting from state i, it makes a first visit to state j with probability $p_{i,j}$ and then makes $m - 1$ returns to state j, each with probability p_{jj}. After those m visits, it never returns to state j, and this event occurs with probability $1 - p_{jj}$. Hence, given that $\{X_0 = i\}$ we have

$$\mathbb{P}(R_j = m \mid X_0 = i) = \begin{cases} p_{ij} \times (p_{jj})^{m-1} \times (1 - p_{jj}), & m \geq 1, \\ \\ 1 - p_{ij}, & m = 0, \end{cases}$$

by the same argument as in (5.3.4) above.

In case $i = j$, R_i is simply the number of returns to state i starting from state i, and it has the geometric distribution

[3]Here, R_j is called a number of returns because the time counter is started at $n = 1$ and excludes the initial state.

$$\mathbb{P}(R_i = m \mid X_0 = i) = (1 - p_{ii})(p_{ii})^m, \qquad m \geq 0.$$

Proposition 5.2 *We have*

$$\mathbb{P}(R_j < \infty \mid X_0 = i) = \begin{cases} 1 - p_{ij}, & \textit{if } p_{jj} = 1, \\ \\ 1, & \textit{if } p_{jj} < 1. \end{cases}$$

Proof We note that

$$\begin{aligned} \mathbb{P}(R_j < \infty \mid X_0 = i) &= \mathbb{P}(R_j = 0 \mid X_0 = i) + \sum_{m=1}^{\infty} \mathbb{P}(R_j = m \mid X_0 = i) \\ &= 1 - p_{ij} + p_{ij}(1 - p_{jj}) \sum_{m=1}^{\infty} (p_{jj})^{m-1} \\ &= \begin{cases} 1 - p_{ij}, & \text{if } p_{jj} = 1, \\ \\ 1, & \text{if } p_{jj} < 1. \end{cases} \end{aligned}$$

□

We also have

$$\mathbb{P}(R_j = \infty \mid X_0 = i) = \begin{cases} p_{ij}, & \text{if } p_{jj} = 1, \\ \\ 0, & \text{if } p_{jj} < 1. \end{cases}$$

In particular if $p_{jj} = 1$, i.e. state j is recurrent, we have

$$\mathbb{P}(R_j = m \mid X_0 = i) = 0, \qquad m \geq 1,$$

and in this case,

$$\begin{cases} \mathbb{P}(R_j < \infty \mid X_0 = i) = \mathbb{P}(R_j = 0 \mid X_0 = i) = 1 - p_{ij}, \\ \\ \mathbb{P}(R_j = \infty \mid X_0 = i) = 1 - \mathbb{P}(R_j < \infty \mid X_0 = i) = p_{ij}. \end{cases}$$

On the other hand, when $i = j$, by (1.5.13) we find

$$
\begin{aligned}
\mathbb{P}(R_i < \infty \mid X_0 = i) &= \sum_{m=0}^{\infty} \mathbb{P}(R_i = m \mid X_0 = i) \\
&= (1 - p_{ii}) \sum_{m=0}^{\infty} (p_{ii})^m \\
&= \begin{cases} 0, & \text{if } p_{ii} = 1, \\ \\ 1, & \text{if } p_{ii} < 1, \end{cases}
\end{aligned} \tag{5.4.3}
$$

hence

$$
\mathbb{P}(R_i = \infty \mid X_0 = i) = \begin{cases} 1, & \text{if } p_{ii} = 1, \\ \\ 0, & \text{if } p_{ii} < 1, \end{cases} \tag{5.4.4}
$$

i.e. the number of returns to a recurrent state is infinite with probability one.

Mean Number of Returns

The notion of *mean number of returns* will be needed for the classification of states of Markov chains in Chap. 6. By (A.4), when $p_{jj} < 1$ we have $\mathbb{P}(R_j < \infty \mid X_0 = i) = 1$ and[4]

$$
\mathbb{E}[R_j \mid X_0 = i] = \sum_{m=0}^{\infty} m \mathbb{P}(R_j = m \mid X_0 = i) \tag{5.4.5}
$$

$$
= (1 - p_{jj}) p_{ij} \sum_{m=1}^{\infty} m (p_{jj})^{m-1}
$$

$$
= \frac{p_{ij}}{1 - p_{jj}}, \tag{5.4.6}
$$

hence

$$
\mathbb{E}[R_j \mid X_0 = i] < \infty \quad \text{if} \quad p_{jj} < 1.
$$

If $p_{j,j} = 1$ then $\mathbb{E}[R_j \mid X_0 = i] = \infty$ unless $p_{i,j} = 0$, in which case $\mathbb{P}(R_j = 0 \mid X_0 = i) = 1$ and $\mathbb{E}[R_j \mid X_0 = i] = 0$. In particular, when $i = j$ we find the next proposition.

Proposition 5.3 *The mean number of returns to state ⓘ is given by*

$$
\mathbb{E}[R_i \mid X_0 = i] = \frac{p_{ii}}{1 - p_{ii}},
$$

and it is finite, i.e. $\mathbb{E}[R_i \mid X_0 = i] < \infty$, *if and only if* $p_{ii} < 1$.

[4]We are using the identity $\sum_{k=1}^{\infty} k r^{k-1} = (1 - r)^{-2}$, cf. (A.4).

More generally, by (5.4.2) we can also write

$$\begin{aligned}\mathbb{E}[R_j \mid X_0 = i] &= \mathbb{E}\left[\sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=j\}} \mid X_0 = i\right] \\ &= \sum_{n=1}^{\infty} \mathbb{E}[\mathbb{1}_{\{X_n=j\}} \mid X_0 = i] \\ &= \sum_{n=1}^{\infty} \mathbb{P}(X_n = j \mid X_0 = i) \\ &= \sum_{n=1}^{\infty} \left[P^n\right]_{i,j}. \end{aligned} \tag{5.4.7}$$

The above quantity coincides with

$$\mathbb{E}[R_j \mid X_0 = i] = -\mathbb{1}_{\{i=j\}} + \left[(I_d - P)^{-1}\right]_{i,j},$$

where $(I_d - P)^{-1}$ is the matrix inverse of $I_d - P$, by analogy with (A.3). Finally, if m is the *only* absorbing state, we can also write

$$\mathbb{E}[T_m \mid X_0 = i] = 1 + \sum_{\substack{j\in\mathbb{S}\\ j\neq m}} \mathbb{E}[R_j \mid X_0 = i] = \sum_{\substack{j\in\mathbb{S}\\ j\neq m}} \left[(I_d - P)^{-1}\right]_{i,j}, \quad i \neq m.$$

See [AKS93] for an application of this formula to the Snakes and Ladders game.

Exercises

Exercise 5.1 Consider a Markov chain $(X_n)_{n\in\mathbb{N}}$ with state space $\mathbb{S} = \{0, 1, 2, 3\}$ and transition probabilities

$$\begin{aligned} &\mathbb{P}(X_1 = 0 \mid X_0 = 0) = 1, && \mathbb{P}(X_1 = 3 \mid X_0 = 3) = 1, \\ &\mathbb{P}(X_1 = 0 \mid X_0 = 1) = 1/2, && \mathbb{P}(X_1 = 2 \mid X_0 = 1) = 1/2, \\ &\mathbb{P}(X_1 = 1 \mid X_0 = 2) = 1/3, && \mathbb{P}(X_1 = 3 \mid X_0 = 2) = 2/3. \end{aligned}$$

(a) Draw the graph of the chain and write down its transition matrix.
(b) Compute $\alpha := \mathbb{P}(T_3 < \infty \mid X_0 = 1)$ and $\beta := \mathbb{P}(T_3 < \infty \mid X_0 = 2)$, where

$$T_3 := \inf\{n \geq 0 \ : \ X_n = 3\}.$$

(c) Letting

$$T_{0,3} := \inf\{n \geq 0 \ : \ X_n = 0 \text{ or } X_n = 3\},$$

compute $\mathbb{E}[T_{0,3} \mid X_0 = 1]$ and $\mathbb{E}[T_{0,3} \mid X_0 = 2]$.

Exercise 5.2 Consider a Markov chain $(X_n)_{n\geq 0}$ with state space $\mathbb{S} = \{0, 1\}$ and transition matrix

$$P = \begin{bmatrix} 0.4 & 0.6 \\ \\ 0.8 & 0.2 \end{bmatrix}.$$

Compute the mean duration between two visits to state ①.

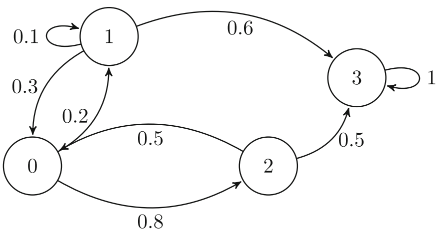
Exercise 5.3 Consider the Markov chain $(X_n)_{n\geq 0}$ on $\mathbb{S} = \{0, 1, 2\}$ whose transition probability matrix P is given by

$$P = [\quad 0 \quad 1 \quad 20 \quad 1 \quad 0 \quad 01 \quad 1/3 \quad 0 \quad 2/32 \quad 0 \quad 1 \quad 0].$$

(a) Draw a graph of the chain and find the probability $g_0(k)$ that the chain is absorbed into state ⓪ given that it started from states $k = 0, 1, 2$.

(b) Determine the mean time $h_0(k)$ it takes until the chain is absorbed into state ⓪, after starting from $k = 0, 1, 2$.

Exercise 5.4 Consider the Markov chain with the graph



(5.4.8)

and let

$$T_k^r := \inf\{n \geq 1 \;:\; X_n = k\}$$

denote the *return* time to state $k = 0, 1, 2, 3$.

(a) Find the probabilities

$$p_{k,2} := \mathbb{P}(T_2^r < \infty \mid X_0 = k),$$

for $k = 0, 1, 2, 3$.

(b) Find the probabilities

$$p_{k,1} := \mathbb{P}(T_1^r < \infty \mid X_0 = k), \qquad k = 0, 1, 2, 3.$$

Exercise 5.5 Consider a Markov chain $(X_n)_{n\in\mathbb{N}}$ with state space $\mathbb{S} = \{0, 1, 2, 3\}$ and transition probability matrix given by

$$[\ P_{i,j}\ ]_{0\le i,j\le 3} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.3 & 0 & 0.7 & 0 \\ 0 & 0.3 & 0 & 0.7 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(a) What are the absorbing states of the chain $(X_n)_{n\in\mathbb{N}}$?
(b) Given that the chain starts at ①, find its probability of absorption $g_0(k) = \mathbb{P}(T_0 < \infty \mid X_0 = k)$ into state ⓪ for $k = 0, 1, 2, 3$.
(c) Find the mean hitting times $h_1(k) = \mathbb{E}[T_1 \mid X_0 = k]$ of state ① starting from state ⓚ, for $k = 0, 1, 2, 3$.

Exercise 5.6 Consider a random walk with Markov transition matrix given by

$$[\ P_{i,j}\ ]_{0\le i,j\le 3} = \begin{bmatrix} 0.5 & 0 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0.5 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Compute the average time it takes to reach state ③ given that the chain is started at state ⓪.

Exercise 5.7 Consider the Markov chain $(X_n)_{n\ge 0}$ on $\{0, 1, 2, 3\}$ whose transition probability matrix P is given by

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(a) Draw the graph of this chain.
(b) Find the probability $g_0(k)$ that the chain is absorbed into state ⓪ given that it started from state $k = 0, 1, 2, 3$.
(c) Determine the mean time $h(k)$ it takes until the chain hits an absorbing state, after starting from $k = 0, 1, 2, 3$.

Exercise 5.8 Consider a discrete-time homogeneous Markov chain $(X_n)_{n\in\mathbb{N}}$ on a state space $\mathbb{S}$, and the first hitting time

$$T_A = \inf\{n \geq 0 \ : \ Z_n \in A\},$$

of a subset $A \subset$ of $\mathbb{S}$. Show that $(X_n)_{n\in\mathbb{N}}$ has the *strong Markov property* with respect to T_A, i.e. show that for all $n, m \geq 0$, $j \in \mathbb{S}$, and $(i_k)_{k\in\mathbb{N}} \subset \mathbb{S}$ we have

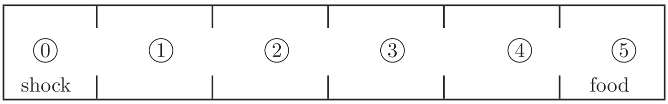
$$\mathbb{P}(X_{T_A+n} = j \mid X_{T_A} = i_0, \ldots, X_0 = i_{T_A} \text{ and } T_A < +\infty) = \mathbb{P}(X_n = j \mid X_0 = i_0).$$

Exercise 5.9 We consider the simple random walk $(S_n)_{n\in\mathbb{N}}$ of Chap. 3.

(a) Using first step analysis, recover the formula (3.4.16) for the probability $\mathbb{P}(T_0 < \infty \mid S_0 = k)$ of hitting state ⓪ in finite time starting from any state ⓚ ≥ 0 when $q < p$.
(b) Using first step analysis, recover the formula (3.4.18) giving the mean hitting time $\mathbb{E}[T_0 \mid S_0 = k]$ of state ⓪ from any state ⓚ ≥ 0 when $q > p$.

Exercise 5.10 A player tosses a fair six-sided die and records the number appearing on the uppermost face. The die is then tossed again and the second result is added to the first one. This procedure is repeated until the sum of all results becomes strictly greater than 10. Compute the probability that the game finishes with a cumulative sum equal to 13.

Exercise 5.11 A fish is put into the linear maze as shown, and its state at time n is denoted by $X_n \in \{0, 1, \ldots, 5\}$:



Starting from any state $k \in \{1, 2, 3, 4\}$, the fish moves to the right with probability p and to the left with probability q such that $p + q = 1$ and $p \in (0, 1)$. Consider the hitting times

$$T_0 = \inf\{n \geq 0 \ : \ X_n = 0\}, \quad \text{and} \quad T_5 = \inf\{n \geq 0 \ : \ X_n = 5\},$$

and $g(k) = \mathbb{P}(T_5 < T_0 \mid X_0 = k)$, $k = 0, 1, \ldots, 5$.

(a) Using first step analysis, write down the equation satisfied by $g(k)$, $k = 0, 1, \ldots, 5$, and give the values of $g(0)$ and $g(5)$.
(b) Assume that the fish is equally likely to move right or left at each step. Compute the probability that starting from state ⓚ it finds the food before getting shocked, for $k = 0, 1, \ldots, 5$.

Exercise 5.12 Starting from a state $m \geq 1$ at time k, the next state of a random device at time $k + 1$ is uniformly distributed among $\{0, 1, \ldots, m - 1\}$, with ⓪ as an absorbing state.

(a) Model the time evolution of this system using a Markov chain whose transition probability matrix will be given explicitly.
(b) Let $h_0(m)$ denote the mean time until the system reaches the state zero for the first time after starting from state m. Using first step analysis, write down the equation satisfied by $h_0(m)$, $m \geq 1$ and give the values of $h_0(0)$ and $h_0(1)$.
(c) Show that $h_0(m)$ is given by $h_0(m) = h_0(m-1) + \dfrac{1}{m}$, $m \geq 1$, and that

$$h_0(m) = \sum_{k=1}^{m} \frac{1}{k},$$

for all $m \in \mathbb{N}$.

Exercise 5.13 An individual is placed in a castle tower having three exits. Exit A leads to a tunnel that returns to the tower after three days of walk. Exit B leads to a tunnel that returns to the tower after one day of walk. Exit C leads to the outside. Since the inside of the tower is dark, each exit is chosen at random with probability $1/3$. The individual decides to remain outside after exiting the tower, and you may choose the number of steps it takes from Exit C to the outside of the tower, *e.g.* take it equal to 0 for simplicity.

(a) Show that this problem can be modeled using a Markov chain $(X_n)_{n\in\mathbb{N}}$ with four states. Draw the graph of the chain $(X_n)_{n\in\mathbb{N}}$.
(b) Write down the transition matrix of the chain $(X_n)_{n\in\mathbb{N}}$.
(c) Starting from inside the tower, find the average time it takes to exit the tower.

Exercise 5.14 A mouse is trapped in a maze. Initially it has to choose one of two directions. If it goes to the right, then it will wander around in the maze for three minutes and will then return to its initial position. If it goes to the left, then with probability $1/3$ it will depart the maze after two minutes of travelling, and with probability $2/3$ it will return to its initial position after five minutes of travelling. Assuming that the mouse is at all times equally likely to go to the left or to the right, what is the expected number of minutes that it will remain trapped in the maze?

Exercise 5.15 This exercise is a particular case of (5.1.8). Consider the Markov chain whose transition probability matrix P is given by

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.1 & 0.6 & 0.1 & 0.2 \\ 0.2 & 0.3 & 0.4 & 0.1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(a) Find the probability that the chain finishes at 0 given that it was started at state 1.
(b) Determine the mean time it takes until the chain reaches an absorbing state.

Exercise 5.16 Consider the Markov chain on $\{0, 1, 2\}$ with transition matrix

$$\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 3/4 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(a) Compute the probability $\mathbb{P}(T_2 < \infty \mid X_0 = 1)$ of *hitting* state ② in finite time starting from state ①, and the probability $\mathbb{P}(T_1^r < \infty \mid X_0 = 1)$ of *returning* to state ① in finite time.

(b) Compute the mean *return time* $\mu_1(1) = \mathbb{E}[T_1^r \mid X_0 = 1]$ to state ① and the mean *hitting time* $h_2(1) = \mathbb{E}[T_2 \mid X_0 = 1]$ of state ② starting from state ①.

Exercise 5.17 Taking $\mathbb{N} := \{0, 1, 2, \ldots\}$, consider the random walk $(Z_k)_{k\in\mathbb{N}} = (X_k, Y_k)_{k\in\mathbb{N}}$ on $\mathbb{N} \times \mathbb{N}$ with the transition probabilities

$$\begin{aligned}
&\mathbb{P}(X_{k+1} = x + 1,\ Y_{k+1} = y \mid X_k = x,\ Y_k = y) \\
&= \mathbb{P}(X_{k+1} = x,\ Y_{k+1} = y + 1 \mid X_k = x,\ Y_k = y) \\
&= \frac{1}{2},
\end{aligned}$$

$k \geq 0$, and let

$$A = \big\{(x, y) \in \mathbb{N} \times \mathbb{N} \ : \ x \geq 2,\ y \geq 2\big\}.$$

Let also

$$T_A := \inf\{n \geq 0 \ : \ X_n \geq 2 \text{ and } Y_n \geq 2\}$$

denote the hitting time of the set A by the random walk $(Z_k)_{k\in\mathbb{N}}$, and consider the mean hitting times

$$\mu_A(x, y) := \mathbb{E}[T_A \mid X_0 = x,\ Y_0 = y], \qquad x, y \in \mathbb{N}.$$

(a) Give the value of $\mu_A(x, y)$ when $x \geq 2$ and $y \geq 2$.

(b) Show that $\mu_A(x, y)$ solves the equation

$$\mu_A(x, y) = 1 + \frac{1}{2}\mu_A(x + 1, y) + \frac{1}{2}\mu_A(x, y + 1), \quad x, y \in \mathbb{N}. \tag{5.4.9}$$

(c) Show that $\mu_A(1, 2) = \mu_A(2, 1) = 2$ and $\mu_A(0, 2) = \mu_A(2, 0) = 4$.

(d) In each round of a game, a coin is thrown to two cans in such a way that each can has exactly 50% chance to receive the coin. Compute the mean time it takes until both cans contain at least \$2.

Exercise 5.18 Let $\mathbb{N} := \{0, 1, 2, \ldots\}$ and consider a random walk $(Z_k)_{k\in\mathbb{N}} = (X_k, Y_k)_{k\in\mathbb{N}}$ on $\mathbb{N} \times \mathbb{N}$ with the transition probabilities

$$\begin{aligned}&\mathbb{P}\big((X_{k+1}, Y_{k+1}) = (x+1, y) \mid (X_k, Y_k) = (x, y)\big)\\ &= \mathbb{P}\big((X_{k+1}, Y_{k+1}) = (x, y+1) \mid (X_k, Y_k) = (x, y)\big) = \frac{1}{2}, \quad (x, y) \in \mathbb{N} \times \mathbb{N},\end{aligned}$$

$k \geq 0$, and let

$$A := \big\{(x, y) \in \mathbb{N} \times \mathbb{N} \ : \ x \geq 3 \text{ or } y \geq 3\big\}.$$

Let also

$$T_A := \inf\big\{n \geq 0 \ : \ (X_n, Y_n) \in A\big\}$$

denote the first hitting time of the set A by the random walk $(Z_k)_{k\in\mathbb{N}} = (X_k, Y_k)_{k\in\mathbb{N}}$, and consider the mean hitting times

$$\mu_A(x, y) := \mathbb{E}\big[T_A \mid (X_0, Y_0) = (x, y)\big], \qquad (x, y) \in \mathbb{N} \times \mathbb{N}.$$

(a) Give the values of $\mu_A(x, y)$ when $(x, y) \in A$.
(b) By applying first step analysis, find an equation satisfied by $\mu_A(x, y)$ for $0 \leq x, y \leq 3$.
(c) Find the values of $\mu_A(x, y)$ for all $x, y \leq 3$ by solving the equation of part (b).
(d) Two players compete in a fair game in which only one of the two players will earn \$1 at each round. How many rounds does it take on average until the gain of one of the players reaches \$3, given that both of them started from zero?

Exercise 5.19 Let $(X_n)_{n\geq 0}$ be a Markov chain with state space $\mathbb{S}$ and transition probability matrix $(P_{ij})_{i,j\in\mathbb{S}}$. Our goal is to compute the expected value of the infinite discounted series

$$h(i) := \mathbb{E}\left[\sum_{n=0}^{\infty} \beta^n c(X_n) \,\middle|\, X_0 = i\right], \qquad i \in \mathbb{S},$$

where $\beta \in (0, 1)$ is the discount coefficient and $c(\cdot)$ is a utility function, starting from state (i). Show, by a first step analysis argument, that $h(i)$ satisfies the equation

$$h(i) = c(i) + \beta \sum_{j\in\mathbb{S}} P_{ij} h(j)$$

for every state $(i) \in \mathbb{S}$.

Exercise 5.20 Consider a *Markov Decision Process* (MDP) on a state space $\mathbb{S}$, with set of actions $\mathbb{A}$ and family $(P^a)_{a\in\mathbb{A}}$ of transition probability matrices

$$\begin{aligned}P : \mathbb{S} \times \mathbb{S} \times \mathbb{A} &\to [0, 1],\\ (k, l, a) &\longmapsto P^a_{k,l},\end{aligned}$$

and a *policy* $\pi : \mathbb{S} \to \mathbb{A}$ giving the action chosen at every given state in $\mathbb{S}$. By first step analysis, derive the *Bellman equation* for the optimal value function

$$V^*(k) = \max_{\pi} \mathbb{E}\left[\sum_{n=0}^{\infty} \gamma^n R(X_n) \,\Big|\, X_0 = k\right],$$

where $\gamma \in (0, 1)$ is a discount factor and $R : \mathbb{S} \to \mathbb{R}$ is a reward function.

Problem 5.21 Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain on $\{0, 1, \ldots, N\}$, $N \geq 1$, with transition matrix $P = \left[\ P_{i,j}\ \right]_{0\leq i,j\leq N}$.

(a) Consider the hitting times

$$T_0 = \inf\{n \geq 0 \ : \ X_n = 0\}, \qquad T_N = \inf\{n \geq 0 \ : \ X_n = N\},$$

and

$$g(k) = \mathbb{P}(T_0 < T_N \mid X_0 = k), \qquad k = 0, 1, \ldots, N.$$

What are the values of $g(0)$ and $g(N)$?

(b) Show, using first step analysis, that the function g satisfies the relation

$$g(k) = \sum_{l=0}^{N} P_{k,l}\, g(l), \qquad k = 1, \ldots, N-1. \tag{5.4.10}$$

(c) In this question and the following ones we consider the Wright-Fisher stochastic model in population genetics, in which the state X_n denotes the number of individuals in the population at time n, and

$$P_{k,l} = \mathbb{P}(X_{n+1} = l \mid X_n = k) = \binom{N}{l}\left(\frac{k}{N}\right)^l\left(1 - \frac{k}{N}\right)^{N-l},$$

$k, l = 0, 1, \ldots, N$. Write down the transition matrix P when $N = 3$.

(d) Show, from Question (b), that given that the solution to (5.4.10) is unique, we have

$$\mathbb{P}(T_0 < T_N \mid X_0 = k) = \frac{N-k}{N}, \qquad k = 0, 1, \ldots, N.$$

(e) Let

$$T_{0,N} = \inf\{n \geq 0 \ : \ X_n = 0 \text{ or } X_n = N\},$$

and

$$h(k) = \mathbb{E}[T_{0,N} \mid X_0 = k], \qquad k = 0, 1, \ldots, N.$$

What are the values of $h(0)$ and $h(N)$?

(f) Show, using first step analysis, Show, using first step analysis, that the function h satisfies the relation

$$h(k) = 1 + \sum_{l=0}^{N} P_{k,l} h(l), \qquad k = 1, 2, \ldots, N-1.$$

(g) Assuming that $N = 3$, compute

$$h(k) = \mathbb{E}[T_{0,3} \mid X_0 = k], \qquad k = 0, 1, 2, 3.$$

Problem 5.22 Consider a gambling process $(X_n)_{n\geq 0}$ on the state space $\{0, 1, \ldots, N\}$, with transition probabilities

$$\mathbb{P}(X_{n+1} = k+1 \mid X_n = k) = p, \qquad \mathbb{P}(X_{n+1} = k-1 \mid X_n = k) = q,$$

$k = 1, 2, \ldots, N-1$, with $p + q = 1$. Let

$$\tau := \inf\{n \geq 0 \ : \ X_n = 0 \text{ or } X_n = N\}$$

denote the time until the process hits either state 0 or state N, and consider the expectation

$$h(k) := \mathbb{E}\left[\sum_{i=0}^{\tau-1} X_i \,\middle|\, X_0 = k\right],$$

of the random sum $\sum_{0\leq i<\tau} X_i$ of all chain values visited *before* the process hits 0 or N after starting from $k = 0, 1, 2, \ldots, N$.

(a) Give the values of $h(0)$ and $h(N)$.[5]

(b) Show, by first step analysis, that $h(k)$ satisfies the equations

$$h(k) = k + ph(k+1) + qh(k-1), \qquad k = 1, 2, \ldots, N-1. \qquad (5.4.11)$$

From now on we take $p = q = 1/2$.

(c) Find a particular solution of Eq. (5.4.11).

(d) Knowing that the solution of the homogeneous equation

$$f(k) = \frac{1}{2} f(k+1) + \frac{1}{2} f(k-1), \qquad k = 1, 2, \ldots, N-1,$$

---

[5]We apply the convention $\sum_{i=0}^{-1} = \sum_{0\leq i<0} = 0$.

takes the form $f(k) = C_1 + C_2 k$, show that the expectation $h(k)$ solution of (5.4.11) is given by

$$h(k) = k\frac{N^2 - k^2}{3}, \qquad k = 0, 1, \ldots, N.$$

(e) Compute $h(1)$ when $N = 2$ and explain why this result makes pathwise sense.

(f) Suppose that you start a business with initial monthly income of \$4 K. Every month the income you receive from that business may *increase* or *decrease* by \$1 K with equal probabilities $(1/2, 1/2)$. You decide to stop that business as soon as your monthly income hits the levels 0 or \$70 K, whichever comes first.

 (i) Compute the expected duration of your business (in number of months).[6]
 (ii) Compute your expected accumulated wealth until the month before you stop your business.
 (iii) Compute your expected accumulated wealth under the assumption that your monthly income remains constant equal to \$4 K over the same mean duration as in Question (f-i) above.
 (iv) Any comment?

[6]Recall that $\mathbb{E}[T_{0,N} \mid X_0 = k] = k(N - k)$, $k = 0, 1, \ldots, N$.

Chapter 6
Classification of States

In this chapter we present the notions of communicating, transient and recurrent states, as well as the concept of irreducibility of a Markov chain. We also examine the notions of positive and null recurrence, periodicity, and aperiodicity of such chains. Those topics will be important when analysing the long-run behavior of Markov chains in the next chapter.

6.1 Communicating States

Definition 6.1 A state $\textcircled{j} \in \mathbb{S}$ is to be *accessible* from another state $\textcircled{i} \in \mathbb{S}$, and we write $\textcircled{i} \longmapsto \textcircled{j}$, if there exists a *finite* integer $n \geq 0$ such that

$$[P^n]_{i,j} = \mathbb{P}(X_n = j \mid X_0 = i) > 0.$$

In other words, it is possible to travel from $\textcircled{i}$ to $\textcircled{j}$ with non-zero probability in a certain (random) number of steps. We also say that state $\textcircled{i}$ leads to state $\textcircled{j}$, and when $i \neq j$ we have

$$\mathbb{P}(T_j^r < \infty \mid X_0 = i) \geq \mathbb{P}(T_j^r \leq n \mid X_0 = i) \geq \mathbb{P}(X_n = j \mid X_0 = i) > 0.$$

Remark 6.1 Since $P^0 = \mathrm{I_d}$ and $[P^0]_{i,j} = \mathbb{P}(X_0 = j \mid X_0 = i) = \mathbb{1}_{\{i = j\}}$ the definition of accessibility states implicitly that any state $\textcircled{i}$ is always accessible from itself (in zero time steps) even if $P_{i,i} = 0$.

In case $\textcircled{i} \longmapsto \textcircled{j}$ and $\textcircled{j} \longmapsto \textcircled{i}$ we say that $\textcircled{i}$ and $\textcircled{j}$ *communicate*[1] and we write $\textcircled{i} \longleftrightarrow \textcircled{j}$.

The binary relation "$\longleftrightarrow$" is a called an *equivalence relation* as it satisfies the following properties:

[1]In graph theory, one says that $\textcircled{i}$ and $\textcircled{i}$ are *strongly connected*.

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4_6

(a) *Reflexivity*:
For all $i \in \mathbb{S}$ we have $\textcircled{i} \longleftrightarrow \textcircled{i}$.

(b) *Symmetry*:
For all $i, j \in \mathbb{S}$ we have that $\textcircled{i} \longleftrightarrow \textcircled{j}$ is equivalent to $\textcircled{j} \longleftrightarrow \textcircled{i}$.

(c) *Transitivity*:
For all $i, j, k \in \mathbb{S}$ such that $\textcircled{i} \longleftrightarrow \textcircled{j}$ and $\textcircled{j} \longleftrightarrow \textcircled{k}$, we have $\textcircled{i} \longleftrightarrow \textcircled{k}$.

Proof It is clear that the relation $\longleftrightarrow$ is reflexive and symmetric. The proof of transitivity can be stated as follows. If $\textcircled{i} \longmapsto \textcircled{j}$ and $\textcircled{j} \longmapsto \textcircled{k}$, there exists $a \geq 1$ and $b \geq 1$ such that

$$[P^a]_{i,j} > 0, \qquad [P^b]_{j,k} > 0.$$

Next, by (4.1.2), for all $n \geq a + b$ we have

$$\begin{aligned}
&\mathbb{P}(X_n = k \mid X_0 = i) \\
&= \sum_{l,m\in\mathbb{S}}^{\infty} \mathbb{P}(X_n = k, \ X_{n-b} = l, \ X_a = m \mid X_0 = i) \\
&= \sum_{l,m\in\mathbb{S}}^{\infty} \mathbb{P}(X_n = k \mid X_{n-b} = l)\mathbb{P}(X_{n-b} = l \mid X_a = m)\mathbb{P}(X_a = m \mid X_0 = i) \\
&\geq \mathbb{P}(X_n = k \mid X_{n-b} = j)\mathbb{P}(X_{n-b} = j \mid X_a = j)\mathbb{P}(X_a = j \mid X_0 = i) \\
&= [P^a]_{i,j}[P^{n-a-b}]_{j,j}[P^b]_{j,k} \qquad (6.1.1) \\
&\geq 0.
\end{aligned}$$

The conclusion follows by taking $n = a + b$, in which case we have

$$\mathbb{P}(X_n = k \mid X_0 = i) \geq [P^a]_{i,j}[P^b]_{j,k} > 0.$$

□

The *equivalence relation* '$\longleftrightarrow$" induces a *partition* of $\mathbb{S}$ into disjoint classes $A_1, A_2, \ldots, A_m$ such that $\mathbb{S} = A_1 \cup \cdots \cup A_m$, and

(a) we have $\textcircled{i} \longleftrightarrow \textcircled{j}$ for all $i, j \in A_q$, and

(b) we have $\textcircled{i} \not\longleftrightarrow \textcircled{j}$ whenever $i \in A_p$ and $j \in A_q$ with $p \neq q$.

The sets $A_1, A_2, \ldots, A_m$ are called the *communicating classes* of the chain.

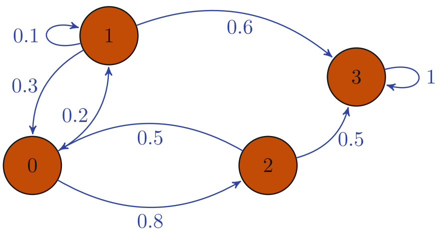
Definition 6.2 A Markov chain whose state space is made of a unique communicating class is said to be *irreducible*, otherwise the chain is said to be *reducible*.

The R package "markovchain" can be used to the irreducibility of a given chain.

```
install.packages("markovchain")
library (markovchain)
statesNames <- c("0", "1")
mcA <- new("markovchain",
transitionMatrix = matrix(c(0.7,0.3,0.1,0.9),
byrow = TRUE, nrow = 2,dimnames = list(statesNames, statesNames)))
is.irreducible(mcA)
```

Clearly, all states in $\mathbb{S}$ communicate when $(X_n)_{n\in\mathbb{N}}$ is irreducible. In case the ith column of a transition matrix P vanishes, i.e. $P_{k,i} = 0$, $i \in \mathbb{S}$, then state ⓘ cannot be reached from any other state and ⓘ becomes a communicating class on its own, as is the case of state ① in Exercise 4.10 for $n \geq 2$, or in Exercise 7.12. The same is true of absorbing states. However, having a returning loop with probability strictly lower than one is not sufficient to turn a given state into a communicating class on its own. Clearly, the existence of at least one absorbing state ⓘ with $P_{i,i} = 1$ makes a chain reducible.

Exercise: Find the communicating classes of the Markov chain with transition matrix (5.4.8) for the equivalence relation "$\longleftrightarrow$".



$$P = \begin{bmatrix} 0 & 0.2 & 0.8 & 0 \\ 0.3 & 0.1 & 0 & 0.6 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The above state space $\mathbb{S} = \{0, 1, 2, 3\}$ is partitioned into two communicating classes which are $\{0, 1, 2\}$ and $\{3\}$.

6.2 Recurrent States

Definition 6.3 A state ⓘ $\in \mathbb{S}$ is said to be *recurrent* if, starting from state ⓘ, the chain will return to state ⓘ within a finite (random) time, with probability 1, i.e.,

$$p_{i,i} := \mathbb{P}(T_i^r < \infty \mid X_0 = i) = \mathbb{P}(X_n = i \text{ for some } n \geq 1 \mid X_0 = i) = 1. \quad (6.2.1)$$

The next Proposition 6.4 uses the mean number of returns R_i to state ⓘ defined in (5.4.2).

Proposition 6.4 *For any state ⓘ $\in \mathbb{S}$, the following statements are equivalent:*

(i) the state ⓘ $\in \mathbb{S}$ is recurrent, i.e. $p_{i,i} = 1$,
(ii) the number of returns to ⓘ $\in \mathbb{S}$ is a.s.[2] infinite, i.e.

[2] almost surely.

$$\mathbb{P}(R_i = \infty \mid X_0 = i) = 1, \text{ i.e. } \mathbb{P}(R_i < \infty \mid X_0 = i) = 0, \tag{6.2.2}$$

(iii) the mean number of returns to ⓘ $\in \mathbb{S}$ *is* infinite, *i.e.*

$$\mathbb{E}[R_i \mid X_0 = i] = \infty, \tag{6.2.3}$$

(iv) we have

$$\sum_{n=1}^{\infty} f_{i,i}^{(n)} = 1, \tag{6.2.4}$$

where $f_{i,i}^{(n)} := \mathbb{P}(T_i^r = n \mid X_0 = i)$, $n \geq 1$, *is the distribution of* T_i^r.

Proof Part (i) follows by the definition (6.2.1) of recurrent states.

(ii) Relation (6.2.2) is equivalent to (6.2.1) by (5.4.3) and (5.4.4).
(iii) Relation (6.2.3) is equivalent to (6.2.1) by (5.4.5).
(iv) Relation (6.2.4) is equivalent to (6.2.1) by (5.4.1).

□

For example, state ⓪ is recurrent for the random walk of Chap. 3 when $p = q = 1/2$, while it is not recurrent if $p \neq q$ as by (3.4.14) we have

$$p_{0,0} = \mathbb{P}(T_0 < \infty) = 2\min(p, q). \tag{6.2.5}$$

As a consequence of (6.2.3), we have the following result.

Corollary 6.5 *A state* $i \in \mathbb{S}$ *is recurrent if and only if*

$$\sum_{n=1}^{\infty} [P^n]_{i,i} = \infty,$$

i.e. the above series diverges.

Proof For all $i, j \in \mathbb{S}$, by (5.4.7) we have

$$\begin{aligned} \mathbb{E}[R_j \mid X_0 = i] &= \mathbb{E}\left[\sum_{n=1}^{\infty} \mathbb{1}_{\{X_n = j\}} \,\Big|\, X_0 = i\right] = \sum_{n=1}^{\infty} \mathbb{E}[\mathbb{1}_{\{X_n = j\}} \mid X_0 = i] \\ &= \sum_{n=1}^{\infty} \mathbb{P}(X_n = j \mid X_0 = i) = \sum_{n=1}^{\infty} [P^n]_{i,j}, \end{aligned} \tag{6.2.6}$$

as in (5.4.7). To conclude we let $j = i$ and apply (6.2.3). □

Corollary 6.5 admits the following consequence, which shows that any state communicating with a recurrent state is itself recurrent. In other words, recurrence is a *class property*, as all states in a given communicating class are recurrent as soon as one of them is recurrent.

Corollary 6.6 *Let* ⓙ $\in \mathbb{S}$ *be a recurrent state. Then any state* ⓘ $\in \mathbb{S}$ *that communicates with state* ⓙ *is also recurrent.*

Proof By definition, since ⓘ $\longmapsto$ ⓙ and ⓙ $\longmapsto$ ⓘ, there exists $a \geq 1$ and $b \geq 1$ such that

$$[P^a]_{i,j} > 0 \qquad [P^b]_{j,i} > 0,$$

and from (6.1.1) applied with $k = i$ we find

$$\begin{aligned}
\sum_{n=a+b}^{\infty} [P^n]_{i,i} &= \sum_{n=a+b}^{\infty} \mathbb{P}(X_n = i \mid X_0 = i) \\
&\geq [P^a]_{i,j}[P^b]_{j,i} \sum_{n=a+b}^{\infty} P_{j,j}^{n-a-b} \\
&= [P^a]_{i,j}[P^b]_{j,i} \sum_{n=0}^{\infty} [P^n]_{j,j} \\
&= \infty,
\end{aligned}$$

which shows that state ⓘ is recurrent from Corollary 6.5 and the assumption that state ⓙ is recurrent. □

A communicating class $A \subset \mathbb{S}$ is therefore recurrent if any of its states is recurrent.

6.3 Transient States

A state ⓘ $\in \mathbb{S}$ is said to be *transient* when it is not recurrent, i.e., by (6.2.1),

$$p_{i,i} = \mathbb{P}(T_i^r < \infty \mid X_0 = i) = \mathbb{P}(X_n = i \text{ for some } n \geq 1 \mid X_0 = i) < 1, \quad (6.3.1)$$

or

$$\mathbb{P}(T_i^r = \infty \mid X_0 = i) > 0.$$

Proposition 6.7 *For any state* ⓘ $\in \mathbb{S}$*, the following statements are equivalent:*

(i) the state ⓘ $\in \mathbb{S}$ *is transient, i.e.* $p_{i,i} < 1$,
(ii) the number of returns to ⓘ $\in \mathbb{S}$ *is a.s.*[3] *finite, i.e.*

$$\mathbb{P}(R_i = \infty \mid X_0 = i) = 0, \ \textit{i.e.} \ \mathbb{P}(R_i < \infty \mid X_0 = i) = 1, \quad (6.3.2)$$

[3] almost surely.

(iii) the mean number of returns to ⓘ $\in \mathbb{S}$ *is* finite*, i.e.*

$$\mathbb{E}[R_i \mid X_0 = i] < \infty, \tag{6.3.3}$$

(iv) we have

$$\sum_{n=1}^{\infty} f_{i,i}^{(n)} < 1. \tag{6.3.4}$$

where $f_{i,i}^{(n)} := \mathbb{P}(T_i^r = n \mid X_0 = i)$*,* $n \geq 1$*, is the distribution of* T_i^r*.*

Proof This is a direct consequence of Proposition 6.4 and the definition (6.3.1) of transience. Regarding point (ii) and the Condition (6.3.2) we also note that the state ⓘ $\in \mathbb{S}$ is transient if and only if

$$\mathbb{P}(R_i = \infty \mid X_0 = i) < 1,$$

which, by (5.4.4) is equivalent to $\mathbb{P}(R_i = \infty \mid X_0 = i) = 0$. □

In other words, a state ⓘ $\in \mathbb{S}$ is *transient* if and only if

$$\mathbb{P}(R_i < \infty \mid X_0 = i) > 0,$$

which by (5.4.3) is equivalent to

$$\mathbb{P}(R_i < \infty \mid X_0 = i) = 1,$$

i.e. the number of returns to state $i \in \mathbb{S}$ is finite with a non-zero probability which is necessarily equal to one. As a consequence of Corollary 6.5 we also have the following result.

Corollary 6.8 *A state* $i \in \mathbb{S}$ *is transient if and only if*

$$\sum_{n=1}^{\infty} [P^n]_{i,i} < \infty,$$

i.e. the above series converges*.*

By Corollary 6.8 and the relation

$$\sum_{n=0}^{\infty} \left[P^n\right]_{i,j} = \left[(I_d - P)^{-1}\right]_{i,j}, \qquad i, j \in \mathbb{S},$$

we find that a chain with finite state space is transient if the matrix $I_d - P$ is invertible.

Clearly, any absorbing state is recurrent, and any state that leads to an absorbing state is transient.

In addition, if a state ⓘ $\in \mathbb{S}$ communicates with a transient state ⓙ then ⓘ is also transient (otherwise the state ⓙ would be recurrent by Corollary 6.6). In other words, transience is a *class property*, as all states in a given communicating class are transient as soon as one of them is transient.

Example

For the two-state Markov chain of Sect. 4.5, Relations (4.5.4) and (4.5.5) show that

$$\sum_{n=1}^{\infty}[P^n]_{0,0} = \sum_{n=1}^{\infty}\frac{b+a\lambda_2^n}{a+b} = \begin{cases} \infty, & \text{if } b>0, \\ \\ \displaystyle\sum_{n=1}^{\infty}(1-a)^n < \infty, & \text{if } b=0 \text{ and } a>0, \end{cases}$$

hence state ⓪ is transient if $b=0$ and $a>0$, and recurrent otherwise. Similarly we have

$$\sum_{n=1}^{\infty}[P^n]_{1,1} = \sum_{n=1}^{\infty}\frac{a+b\lambda_2^n}{a+b} = \begin{cases} \infty, & \text{if } a>0, \\ \\ \displaystyle\sum_{n=1}^{\infty}(1-b)^n < \infty, & \text{if } a=0 \text{ and } b>0, \end{cases}$$

hence state ① is transient if $a=0$ and $b>0$, and recurrent otherwise.

The above results can be recovered by a simple first step analysis for $g_i(j) = \mathbb{P}(T_i < \infty \mid X_0 = j)$, $i, j \in \{0, 1\}$, i.e.

$$\begin{cases} g_0(0) = ag_0(1) + 1 - a \\ \\ g_0(1) = b + (1-b)g_0(1) \\ \\ g_1(0) = (1-a)g_1(0) + a \\ \\ g_1(1) = bg_1(0) + 1 - b, \end{cases}$$

which shows that $g_0(0) = 1$ if $b > 0$ and $g_1(1) = 1$ if $a > 0$.
We close this section with the following result for Markov chains with finite state space.

Theorem 6.9 *Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain with* finite *state space $\mathbb{S}$. Then $(X_n)_{n\in\mathbb{N}}$ has at least one recurrent state.*

Proof Recall that from (5.4.5) we have

$$\mathbb{E}[R_j \mid X_0 = i] = p_{i,j}(1-p_{j,j})\sum_{n=1}^{\infty} n(p_{j,j})^{n-1} = \frac{p_{i,j}}{1-p_{j,j}},$$

for any states ⓘ, ⓙ $\in \mathbb{S}$. Assuming that the state ⓙ $\in \mathbb{S}$ is transient we have $p_{j,j} < 1$ by (6.3.1), hence

$$\mathbb{E}[R_j \mid X_0 = i] = \sum_{n=1}^{\infty}[P^n]_{i,j} < \infty,$$

by (6.3.3) which implies[4] that

$$\lim_{n\to\infty}[P^n]_{i,j} = 0$$

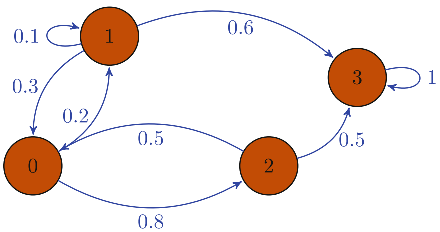
for all transient states $j \in \mathbb{S}$. In case all states in $\mathbb{S}$ were transient, since $\mathbb{S}$ is finite, by the *law of total probability* (4.2.2) we would have

$$0 = \sum_{j\in\mathbb{S}}\lim_{n\to\infty}[P^n]_{i,j} = \lim_{n\to\infty}\sum_{j\in\mathbb{S}}[P^n]_{i,j} = \lim_{n\to\infty} 1 = 1,$$

which is a contradiction. Hence not all states can be transient, and there exists at least one recurrent state. □

Exercises:

(i) Find which states are transient and recurrent in the chain (5.4.8).



$$P = \begin{bmatrix} 0 & 0.2 & 0.8 & 0 \\ 0.3 & 0.1 & 0 & 0.6 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

State ③ is clearly recurrent since we have $T_3^r = 1$ with probability one when $X_0 = 3$. State ② is transient because

$$1 - p_{2,2} = \mathbb{P}(T_2^r = \infty \mid X_0 = 2) = \frac{4}{7} \geq \mathbb{P}(X_1 = 3 \mid X_0 = 2) = 0.5 > 0, \tag{6.3.5}$$

and state ① is transient because

$$\mathbb{P}(T_1^r = \infty \mid X_0 = 1) = 0.8 \geq \mathbb{P}(X_1 = 3 \mid X_0 = 1) = 0.6, \tag{6.3.6}$$

see the Exercise 5.4 for the computations of

$$p_{1,1} = \mathbb{P}(T_2^r < \infty \mid X_0 = 2) = 0.8$$

---

[4]For any sequence $(a_n)_{n\geq 0}$ of *nonnegative* real numbers, $\sum_{n=0}^{\infty} a_n < \infty$ implies $\lim_{n\to\infty} a_n = 0$.

and

$$p_{2,2} = \mathbb{P}(T_2^r < \infty \mid X_0 = 2) = \frac{3}{7}.$$

By Corollary 6.6, the states ⓪ and ① are transient because they communicate with state ②.

(ii) Which are the recurrent states in the simple random walk $(S_n)_{n\in\mathbb{N}}$ of Chap. 3 on $\mathbb{S} = \mathbb{Z}$?

First, we note that this random walk is irreducible as all states communicate when $p \in (0, 1)$. The simple random walk $(S_n)_{n\in\mathbb{N}}$ on $\mathbb{S} = \mathbb{Z}$ has the transition matrix

$$P_{i,i+1} = p, \qquad P_{i,i-1} = q = 1 - p, \qquad i \in \mathbb{Z}.$$

We have

$$[P^n]_{i,i} = \mathbb{P}(S_n = i \mid S_0 = i) = \mathbb{P}(S_n = 0 \mid S_0 = 0),$$

with

$$\mathbb{P}(S_{2n} = 0) = \binom{2n}{n} p^n q^n \qquad \text{and} \qquad \mathbb{P}(S_{2n+1} = 0) = 0, \qquad n \in \mathbb{N}.$$

Hence

$$\begin{aligned}\sum_{n=0}^{\infty} [P^n]_{0,0} &= \sum_{n=0}^{\infty} \mathbb{P}(S_n = 0 \mid S_0 = 0) = \sum_{n=0}^{\infty} \mathbb{P}(S_{2n} = 0 \mid S_0 = 0)\\ &= \sum_{n=0}^{\infty} \binom{2n}{n} p^n q^n = H(1) = \frac{1}{\sqrt{1-4pq}},\end{aligned}$$

and

$$\mathbb{E}[R_0 \mid S_0 = 0] = \sum_{n=1}^{\infty} \mathbb{P}(S_n = 0 \mid S_0 = 0) = \frac{1}{\sqrt{1-4pq}} - 1,$$

where $H(s)$ is defined in (3.4.8).

Consequently, by Corollary 6.5, all states ⓘ $\in \mathbb{Z}$ are recurrent when $p = q = 1/2$, whereas by Corollary 6.8 they are all transient when $p \neq q$, cf. Corollary 6.6.

Alternatively we could reach the same conclusion by directly using (3.4.14) and (6.2.1) which state that

$$\mathbb{P}(T_i^r < \infty \mid X_0 = i) = 2\min(p, q).$$

6.4 Positive Versus Null Recurrence

The expected time of return (or mean recurrence time) to a state ⓘ $\in \mathbb{S}$ is given by

$$\begin{aligned}\mu_i(i) :&= \mathbb{E}[T_i^r \mid X_0 = i] \\ &= \sum_{n=1}^{\infty} n\mathbb{P}(T_i^r = n \mid X_0 = i) \\ &= \sum_{n=1}^{\infty} n f_{i,i}^{(n)}.\end{aligned}$$

Recall that when state ⓘ is recurrent we have $\mathbb{P}(T_i^r < \infty \mid X_0 = i) = 1$, i.e. the random return time T_i^r is almost surely *finite* starting from state ⓘ, nevertheless this yields no information on the finiteness of its expectation $\mu_i(i) = \mathbb{E}[T_i^r \mid X_0 = i]$, cf. the example (1.6.5).

Definition 6.10 A *recurrent* state $i \in \mathbb{S}$ is said to be:

(a) *positive recurrent* if the mean return time to ⓘ is *finite*, i.e.

$$\mu_i(i) = \mathbb{E}[T_i^r \mid X_0 = i] < \infty,$$

(b) *null recurrent* if the mean return time to ⓘ is *infinite*, i.e.

$$\mu_i(i) = \mathbb{E}[T_i^r \mid X_0 = i] = \infty.$$

Exercise: Which states are positive/null recurrent in the simple random walk $(S_n)_{n\in\mathbb{N}}$ of Chap. 3 on $\mathbb{S} = \mathbb{Z}$?

From (3.4.20) and (3.4.17) we know that $\mathbb{E}[T_i^r \mid S_0 = i] = \infty$ for all values of $p \in (0, 1)$, hence all states of the random walk on $\mathbb{Z}$ are null recurrent when $p = 1/2$, while all states are transient when $p \neq 1/2$ due to (3.4.14).

The following Theorem 6.11 shows in particular that a Markov chain with finite state space cannot have any null recurrent state, cf. e.g. Corollary 2.3 in [Kij97], and also Corollary 3.7 in [Asm03].

Theorem 6.11 *Assume that the state space $\mathbb{S}$ of a Markov chain $(X_n)_{n\in\mathbb{N}}$ is* finite. *Then all recurrent states in $\mathbb{S}$ are also positive recurrent.*

As a consequence of Definition 6.2, Corollary 6.6, and Theorems 6.9 and 6.11 we have the following corollary.

Corollary 6.12 *Let $(X_n)_{n\in\mathbb{N}}$ be an irreducible Markov chain with* finite *state space $\mathbb{S}$. Then* all *states of $(X_n)_{n\in\mathbb{N}}$ are positive recurrent.*

6.5 Periodicity and Aperiodicity

Given a state $i \in \mathbb{S}$, consider the sequence

$$\{n \geq 1 \ : \ [P^n]_{i,i} > 0\}$$

of integers which represent the possible travel times from state ⓘ to itself.

Definition 6.13 The *period* of the state $i \in \mathbb{S}$ is the greatest common divisor of the sequence

$$\{n \geq 1 \ : \ [P^n]_{i,i} > 0\}.$$

A state having period 1 is said to be *aperiodic*, which is the case in particular if $P_{i,i} > 0$, i.e. when a state admits a returning loop with nonzero probability.

In particular, any absorbing state is both aperiodic and recurrent. A *recurrent* state $i \in \mathbb{S}$ is said to be *ergodic* if it is both *positive recurrent* and *aperiodic*.

If $[P^n]_{i,i} = 0$ for all $n \geq 1$ then the set $\{n \geq 1 \ : \ [P^n]_{i,i} > 0\}$ is empty and by convention the period of state ⓘ is defined to be 0. In this case, state ⓘ is also transient.

Note also that if

$$\{n \geq 1 \ : \ [P^n]_{i,i} > 0\}$$

contains two distinct numbers that are relatively prime to each other (i.e. their greatest common divisor is 1) then state ⓘ aperiodic.

Proposition 6.14 shows that periodicity is a *class property*, as all states in a given communicating class have same periodicity.

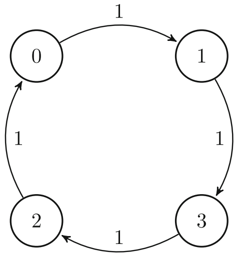
Proposition 6.14 *All states that belong to a same communicating class have the same period.*

Proof Assume that state ⓘ has period d_i, that ⓙ communicates with ⓘ, and let $n \in \{m \geq 1 \ : \ [P^m]_{j,j} > 0\}$. Since ⓘ and ⓙ communicate, there exists $k, l \geq 1$ such that $[P^k]_{i,j} > 0$ and $[P^l]_{j,i} > 0$, hence by (6.1.1) we have $[P^{k+l}]_{i,i} > 0$ hence $k + l$ is a multiple of d_i. Similarly by (6.1.1) we also have $[P^{n+k+l}]_{i,i} > 0$, hence $n + k + l$ and n are multiples of d_i, which implies $d_j \geq d_i$. Exchanging the roles of ⓘ and ⓙ we obtain similarly that $d_i \geq d_j$. □

A Markov chain is said to be *aperiodic* when *all* of its states are aperiodic. Note that any state that communicates with an aperiodic state becomes itself aperiodic. In particular, if a communicating class contains an aperiodic state then the whole class becomes aperiodic.

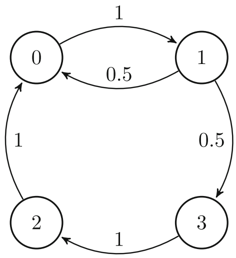
Examples

(i) The chain



clearly has periodicity equal to 4.

(ii) Consider the following chain:

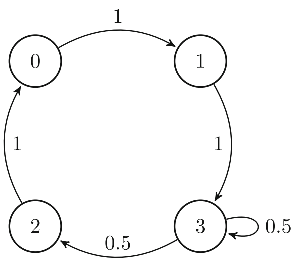


Here we have

$$\begin{aligned}
\{n \geq 1 \ : \ [P^n]_{0,0} > 0\} &= \{2, 4, 6, 8, 10, \ldots\}, \\
\{n \geq 1 \ : \ [P^n]_{1,1} > 0\} &= \{2, 4, 6, 8, 10, \ldots\}, \\
\{n \geq 1 \ : \ [P^n]_{2,2} > 0\} &= \{4, 6, 8, 10, 12, \ldots\}, \\
\{n \geq 1 \ : \ [P^n]_{3,3} > 0\} &= \{4, 6, 8, 10, 12, \ldots\},
\end{aligned}$$

hence *all* states have period 2, and this is also consequence of Proposition 6.14.

(iii) Consider the following chain:



(6.5.1)

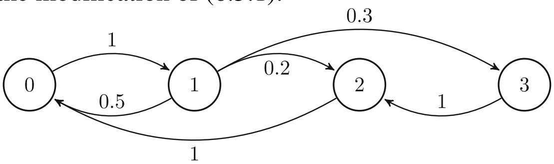
Here we have

$$\begin{aligned}
\{n \geq 1 \ : \ [P^n]_{0,0} > 0\} &= \{4, 5, 6, 7, \ldots\}, \\
\{n \geq 1 \ : \ [P^n]_{1,1} > 0\} &= \{4, 5, 6, 7, \ldots\},
\end{aligned}$$

$$\{n \geq 1 \ : \ [P^n]_{2,2} > 0\} = \{4, 5, 6, 7, \ldots\},$$
$$\{n \geq 1 \ : \ [P^n]_{3,3} > 0\} = \{4, 5, 6, 7, \ldots\},$$

hence *all* states have period 1, see also Proposition 6.14.

(iv) Next, consider the modification of (6.5.1):



Here the chain is aperiodic since we have

$$\{n \geq 1 \ : \ [P^n]_{0,0} > 0\} = \{2, 3, 4, 5, 6, 7, \ldots\},$$
$$\{n \geq 1 \ : \ [P^n]_{1,1} > 0\} = \{2, 3, 4, 5, 6, 7, \ldots\},$$
$$\{n \geq 1 \ : \ [P^n]_{2,2} > 0\} = \{3, 4, 5, 6, 7, 8, \ldots\},$$
$$\{n \geq 1 \ : \ [P^n]_{3,3} > 0\} = \{4, 6, 7, 8, 9, 10, \ldots\},$$

hence *all* states have period 1.

Exercises:

(i) What is the periodicity of the simple random walk $(S_n)_{n\in\mathbb{N}}$ of Chap. 3 on $\mathbb{S} = \mathbb{Z}$?

By (3.3.3) We have

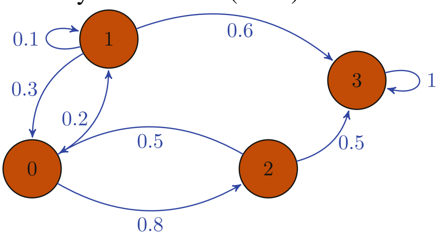
$$[P^{2n}]_{i,i} = \binom{2n}{n} p^n q^n > 0 \quad \text{and} \quad [P^{2n+1}]_{i,i} = 0, \qquad n \in \mathbb{N},$$

hence

$$\{n \geq 1 \ : \ [P^n]_{i,i} > 0\} = \{2, 4, 6, 8, \ldots\},$$

and the chain has period 2.

(ii) Find the periodicity of the chain (5.4.8).



$$\begin{bmatrix} 0 & 0.2 & 0.8 & 0 \\ 0.3 & 0.1 & 0 & 0.6 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

States ⓪, ①, ② and ③ have period 1, hence the chain is aperiodic.

(iii) The chain of Fig. 4.1 is aperiodic since it is irreducible and state ③ has a returning loop.

Exercises

Exercise 6.1 Consider a Markov chain $(X_n)_{n\geq 0}$ on the state space $\{0, 1, 2, 3\}$, with transition matrix

$$\begin{bmatrix} 1/3 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

(a) Draw the graph of this chain and find its communicating classes. Is this Markov chain reducible? Why?
(b) Find the periods of states ⓪, ①, ②, and ③.
(c) Compute $\mathbb{P}(T_0 < \infty \mid X_0 = 0)$, $\mathbb{P}(T_0 = \infty \mid X_0 = 0)$, and $\mathbb{P}(R_0 < \infty \mid X_0 = 0)$.
(d) Which state(s) is (are) absorbing, recurrent, and transient?

Exercise 6.2 Consider the Markov chain on $\{0, 1, 2\}$ with transition matrix

$$\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/4 & 3/4 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(a) Is the chain irreducible? Give its communicating classes.
(b) Which states are absorbing, transient, recurrent, positive recurrent?
(c) Find the period of every state.

Exercise 6.3 Consider a Markov chain $(X_n)_{n\geq 0}$ on the state space $\{0, 1, 2, 3, 4\}$, with transition matrix

$$\begin{bmatrix} 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

(a) Draw the graph of this chain.
(b) Find the periods of states ⓪, ①, ②, and ③.
(c) Which state(s) is (are) absorbing, recurrent, and transient?
(d) Is the Markov chain reducible? Why?

Exercise 6.4 Consider the Markov chain with transition matrix

$$[\, P_{i,j} \,]_{0\leq i,j\leq 5} = \begin{bmatrix} 1/2 & 0 & 1/4 & 0 & 0 & 1/4 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1/6 & 1/2 & 1/6 & 0 & 0 & 1/6 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

(a) Is the chain reducible? If yes, find its communicating classes.
(b) Determine the transient and recurrent states of the chain.
(c) Find the period of each state.

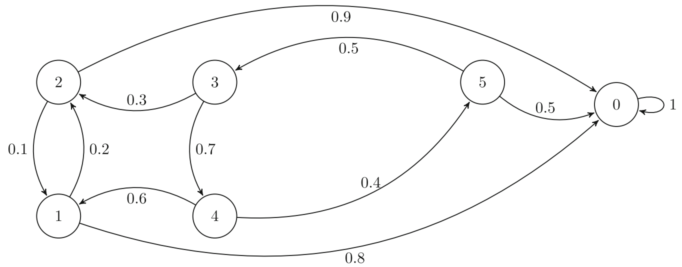
Exercise 6.5 Consider the Markov chain with transition matrix

$$\begin{bmatrix} 0.8 & 0 & 0.2 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0.3 & 0.4 & 0 & 0.3 \end{bmatrix}.$$

(a) Is the chain irreducible? If not, give its communicating classes.
(b) Find the period of each state. Which states are absorbing, transient, recurrent, positive recurrent?

Exercise 6.6 In the following chain, find:

(a) the communicating class(es),
(b) the transient state(s),
(c) the recurrent state(s),
(d) the positive recurrent state(s),
(e) the period of every state.



Exercise 6.7 Consider two boxes containing a total of N balls. At each unit of time one ball is chosen randomly among N and moved to the other box.

(a) Write down the transition matrix of the Markov chain $(X_n)_{n\in\mathbb{N}}$ with state space $\{0, 1, 2, \ldots, N\}$, representing the number of balls in the first box.
(b) Determine the periodicity, transience and recurrence of the Markov chain.

Exercise 6.8

(a) Is the Markov chain of Exercise 4.10-(a) recurrent? positive recurrent?
(b) Find the periodicity of every state.
(c) Same questions for the success runs Markov chain of Exercise 4.10-(b).

Problem 6.9 Let $\alpha > 0$ and consider the Markov chain with state space $\mathbb{N}$ and transition matrix given by

$$P_{i,i-1} = \frac{1}{\alpha + 1}, \quad P_{i,i+1} = \frac{\alpha}{\alpha + 1}, \qquad i \geq 1.$$

and a reflecting barrier at 0, such that $P_{0,1} = 1$. Compute the mean return times $\mathbb{E}[T_k^r \mid X_0 = k]$ for $k \in \mathbb{N}$, and show that the chain is positive recurrent if and only if $\alpha < 1$.

Chapter 7
Long-Run Behavior of Markov Chains

This chapter is concerned with the large time behavior of Markov chains, including the computation of their limiting and stationary distributions. Here the notions of recurrence, transience, and classification of states introduced in the previous chapter play a major role.

7.1 Limiting Distributions

Definition 7.1 A Markov chain $(X_n)_{n\in\mathbb{N}}$ is said to admit a *limiting probability distribution* if the following conditions are satisfied:

(i) the limits

$$\lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i) \tag{7.1.1}$$

exist for all $i, j \in \mathbb{S}$, and

(ii) they form a *probability distribution* on $\mathbb{S}$, i.e.

$$\sum_{j\in\mathbb{S}} \lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i) = 1, \tag{7.1.2}$$

for all $i \in \mathbb{S}$.

Note that Condition (7.1.2) is always satisfied if the limits (7.1.1) exist and the state space $\mathbb{S}$ is finite.

As remarked in (4.5.6) and (4.5.7) above, the two-state Markov chain has a limiting distribution given by

$$[\pi_0, \pi_1] = \left[\frac{b}{a+b}, \frac{a}{a+b}\right], \tag{7.1.3}$$

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4_7

provided that $(a, b) \neq (0, 0)$ and $(a, b) \neq (1, 1)$, while the corresponding mean return times are given from (5.3.3) by

$$(\mu_0(0), \mu_1(1)) = \left(1 + \frac{a}{b}, 1 + \frac{b}{a}\right),$$

i.e. the limiting probabilities are given by the inverses

$$[\pi_0, \pi_1] = \left[\frac{b}{a+b}, \frac{a}{a+b}\right] = \left[\frac{1}{\mu_0(0)}, \frac{1}{\mu_1(1)}\right] = \left[\frac{\mu_1(0)}{\mu_0(1) + \mu_1(0)}, \frac{\mu_0(1)}{\mu_0(1) + \mu_1(0)}\right].$$

This fact is not a simple coincidence, and it is actually a consequence of the following more general result, which shows that the longer it takes on average to return to a state, the smaller the probability is to find the chain in that state. Recall that a chain $(X_n)_{n\in\mathbb{N}}$ is said to be recurrent, resp. aperiodic, if all its states are recurrent, resp. aperiodic.

Theorem 7.2 (Theorem IV.4.1 in [KT81]) *Consider a Markov chain $(X_n)_{n\in\mathbb{N}}$ satisfying the following 3 conditions:*

(i) recurrence,
(ii) aperiodicity, and
(iii) irreducibility.

Then we have

$$\lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i) = \frac{1}{\mu_j(j)}, \qquad i, j \in \mathbb{S}, \tag{7.1.4}$$

independently of the initial state $i \in \mathbb{S}$, where

$$\mu_j(j) = \mathbb{E}[T_j^r \mid X_0 = j] \in [1, \infty]$$

is the mean return time to state $(j) \in \mathbb{S}$.

In Theorem 7.2, Condition (i), resp. Condition (ii), is satisfied from Proposition 6.14, resp. from Proposition 6.6, provided that at least one state is aperiodic, resp. recurrent, since the chain is irreducible.

The conditions stated in Theorem 7.2 are sufficient, but they are not all necessary. For example, a Markov chain may admit a limiting distribution when the recurrence and irreducibility Conditions (i) and (iii) above are not satisfied.

Note that the limiting probability (7.1.4) is independent of the initial state (i), and it vanishes whenever the state (i) is transient or null recurrent, cf. Proposition 7.4 below. In the case of the two-state Markov chain this result in consistent with (4.5.6), (4.5.7), and (7.1.3). However it does not apply to e.g. the simple random walk of Chap. 3 which is *not* recurrent when $p \neq q$ from (6.2.5), and has period 2.

For an aperiodic chain with finite state space, we can show that the limit $\lim_{n\to\infty} \mathbb{P}(X_n = i \mid X_0 = j)$ exists for all $i, j \in \mathbb{S}$ by breaking the chain into communicating classes, however it may depend on the initial state (j). This however does not apply to the random walk of Chap. 3 which is not aperiodic and has an infinite state space, although it can be turned into an aperiodic chain by allowing a draw as in Exercise 2.1.

The following sufficient condition is a consequence of Theorem IV.1.1 in [KT81].

Proposition 7.3 *Consider a Markov chain* $(X_n)_{n\in\mathbb{N}}$ *with* finite *state space* $\mathbb{S} = \{0, 1, \ldots, N\}$, *whose transition matrix* P *is* regular, *i.e. there exists* $n \geq 1$ *such that all entries of the power matrix* P^n *are non-zero. Then* $(X_n)_{n\in\mathbb{N}}$ *admits a limiting probability distribution* $\pi = (\pi_i)_{i=0,1,\ldots,N}$ *given by*

$$\pi_j = \lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i), \qquad 0 \leq i, j \leq N, \tag{7.1.5}$$

A chain with finite state space is regular if it is aperiodic and irreducible, cf. Proposition 1.7 of [LPW09].

We close this section with the following proposition, whose proof uses an argument similar to that of Theorem 6.9.

Proposition 7.4 *Let* $(X_n)_{n\in\mathbb{N}}$ *be a Markov chain with a transient state* $(j) \in \mathbb{S}$. *Then we have*

$$\lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i) = 0,$$

for all $(i) \in \mathbb{S}$.

Proof Since (j) is a transient state, the probability p_{jj} of return to (j) in finite time satisfies $p_{jj} < 1$ by definition, hence by Relation (5.4.5) p. 133, the expected number of returns to (j) starting from state (j) is finite[1]:

$$\begin{aligned}
\mathbb{E}[R_j \mid X_0 = i] &= \mathbb{E}\left[\sum_{n=1}^{\infty} \mathbb{1}_{\{X_n=j\}} \,\middle|\, X_0 = i\right] \\
&= \sum_{n=1}^{\infty} \mathbb{E}\left[\mathbb{1}_{\{X_n=j\}} \mid X_0 = i\right] \\
&= \sum_{n=1}^{\infty} \mathbb{P}(X_n = j \mid X_0 = i) \\
&= \frac{p_{ij}}{1 - p_{jj}} < \infty.
\end{aligned}$$

[1] The exchange of infinite sums and expectation is valid in particular for nonnegative series.

The convergence of the above series implies the convergence to 0 of its general term, i.e.

$$\lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i) = 0$$

for all ⓘ $\in \mathbb{S}$, which is the expected conclusion. □

7.2 Stationary Distributions

Definition 7.5 A *probability distribution* on $\mathbb{S}$ is any a family $\pi = (\pi_i)_{i\in\mathbb{S}}$ in $[0,1]$ such that

$$\sum_{i\in\mathbb{S}} \pi_i = 1.$$

Next, we state the definition of stationary distribution.

Definition 7.6 A *probability distribution* π on $\mathbb{S}$ is said to be *stationary* if, starting X_0 at time 0 with the distribution $(\pi_i)_{i\in\mathbb{S}}$, it turns out that the distribution of X_1 is still $(\pi_i)_{i\in\mathbb{S}}$ at time 1.
In other words, $(\pi_i)_{i\in\mathbb{S}}$ is stationary for the Markov chain with transition matrix P if, letting

$$\mathbb{P}(X_0 = i) := \pi_i, \qquad i \in \mathbb{S},$$

at time 0, implies

$$\mathbb{P}(X_1 = i) = \mathbb{P}(X_0 = i) = \pi_i, \quad i \in \mathbb{S},$$

at time 1. This also means that

$$\pi_j = \mathbb{P}(X_1 = j) = \sum_{i\in\mathbb{S}} \mathbb{P}(X_1 = j \mid X_0 = i)\mathbb{P}(X_0 = i) = \sum_{i\in\mathbb{S}} \pi_i P_{i,j}, \quad j \in \mathbb{S},$$

i.e. the distribution π is stationary if and only if the vector π is *invariant* (or stationary) by the matrix P, that means

$$\pi = \pi P. \tag{7.2.1}$$

Note that in contrast with (5.1.3), the multiplication by P in (7.2.1) is on the *right* hand side and *not* on the left. The relation (7.2.1) can be rewritten as the *balance condition*

$$\sum_{i\in\mathbb{S}} \pi_i P_{i,k} = \pi_k = \pi_k \sum_{j\in\mathbb{S}} P_{k,j} = \sum_{j\in\mathbb{S}} \pi_k P_{k,j}, \tag{7.2.2}$$

which can be illustrated as follows (Fig. 7.1):

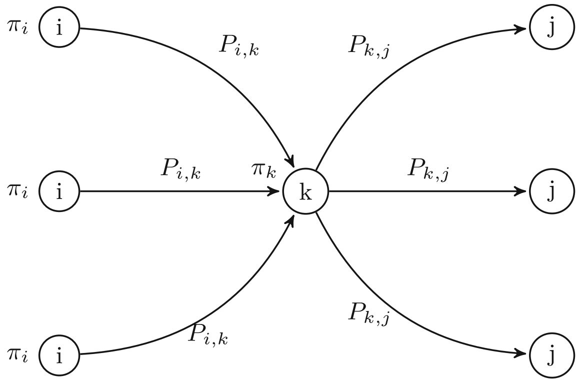


Fig. 7.1 Global balance condition (discrete time)

We also note that the stationarity and limiting properties of distributions are quite different concepts. If the chain is started in the stationary distribution then it will remain in that distribution at *any* subsequent time step (which is stronger than saying that the chain will reach that distribution after an *infinite* number of time steps). On the other hand, in order to reach the limiting distribution the chain can be started from any given initial distribution or even from any fixed given state, and it will converge to the limiting distribution if it exists. Nevertheless, the limiting and stationary distribution may coincide in some situations as in Theorem 7.8 below.

More generally, assuming that X_n has the invariant (or stationary) distribution π at time n, i.e. $\mathbb{P}(X_n = i) = \pi_i$, $i \in \mathbb{S}$, we have

$$\begin{aligned}\mathbb{P}(X_{n+1} = j) &= \sum_{i \in \mathbb{S}} \mathbb{P}(X_{n+1} = j \mid X_n = i)\mathbb{P}(X_n = i)\\ &= \sum_{i \in \mathbb{S}} P_{i,j}\mathbb{P}(X_n = i) = \sum_{i \in \mathbb{S}} P_{i,j}\pi_i\\ &= [\pi P]_j = \pi_j, \qquad j \in \mathbb{S},\end{aligned}$$

since the transition matrix of $(X_n)_{n\in\mathbb{N}}$ is time homogeneous, hence

$$\mathbb{P}(X_n = j) = \pi_j, \quad j \in \mathbb{S}, \quad \Longrightarrow \quad \mathbb{P}(X_{n+1} = j) = \pi_j, \quad j \in \mathbb{S}.$$

By induction on $n \geq 0$, this yields

$$\mathbb{P}(X_n = j) = \pi_j, \qquad j \in \mathbb{S}, \quad n \geq 1,$$

i.e. the chain $(X_n)_{n\in\mathbb{N}}$ remains in the same distribution π at all times $n \geq 1$, provided that it has been started with the stationary distribution π at time $n = 0$.

Proposition 7.7 *Assume that $\mathbb{S} = \{0, 1, \ldots, N\}$ is finite and that the limiting distribution* (7.1.5)

$$\pi_j^{(i)} := \lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i)$$

exists for all $i, j \in \mathbb{S}$, i.e. we have

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \pi_0^{(0)} & \pi_1^{(0)} & \cdots & \pi_N^{(0)} \\ \pi_0^{(1)} & \pi_1^{(1)} & \cdots & \pi_N^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_0^{(N)} & \pi_1^{(N)} & \cdots & \pi_N^{(N)} \end{bmatrix}.$$

Then for every $i = 0, 1, \ldots, N$, the vector $\pi^{(i)} := (\pi_j^{(i)})_{j\in\{0,1,\ldots,N\}}$ is a stationary distribution and we have

$$\pi^{(i)} = \pi^{(i)} P, \tag{7.2.3}$$

i.e. $\pi^{(i)}$ is invariant (or stationary) by P, $i = 0, 1, \ldots, N$.

Proof We have

$$\begin{aligned} \pi_j^{(i)} &:= \lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i) \\ &= \lim_{n\to\infty} \mathbb{P}(X_{n+1} = j \mid X_0 = i) \\ &= \lim_{n\to\infty} \sum_{l\in\mathbb{S}} \mathbb{P}(X_{n+1} = j \mid X_n = l)\mathbb{P}(X_n = l \mid X_0 = i) \\ &= \lim_{n\to\infty} \sum_{l\in\mathbb{S}} P_{l,j}\mathbb{P}(X_n = l \mid X_0 = i) \\ &= \sum_{l\in\mathbb{S}} P_{l,j} \lim_{n\to\infty} \mathbb{P}(X_n = l \mid X_0 = i) \\ &= \sum_{l\in\mathbb{S}} \pi_l^{(i)} P_{l,j}, \qquad i, j \in \mathbb{S}, \end{aligned}$$

where we exchanged limit and summation because the state space $\mathbb{S}$ is assumed to be finite, which shows that

$$\pi^{(i)} = \pi^{(i)} P,$$

i.e. (7.2.3) holds and $\pi^{(i)}$ is a stationary distribution, $i = 0, 1, \ldots, N$. □

Proposition 7.7 can be applied in particular when the limiting distribution $\pi_j := \lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i)$ does not depend on the initial state ⓘ, i.e.

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \pi_0 & \pi_1 & \cdots & \pi_N \\ \pi_0 & \pi_1 & \cdots & \pi_N \\ \vdots & \vdots & \ddots & \vdots \\ \pi_0 & \pi_1 & \cdots & \pi_N \end{bmatrix}.$$

For example, the limiting distribution (7.1.3) of the two-state Markov chain is also an invariant distribution, i.e. it satisfies (7.2.1). In particular we have the following result.

Theorem 7.8 (Theorem IV.4.2 in [KT81]) *Assume that the Markov chain $(X_n)_{n\in\mathbb{N}}$ satisfies the following* 3 *conditions:*

(i) positive recurrence,
(ii) aperiodicity, and
(iii) irreducibility.

Then the chain $(X_n)_{n\in\mathbb{N}}$ admits a limiting distribution

$$\pi_i := \lim_{n\to\infty} \mathbb{P}(X_n = i \mid X_0 = j) = \lim_{n\to\infty} [P^n]_{j,i} = \frac{1}{\mu_i(i)}, \qquad i, j \in \mathbb{S},$$

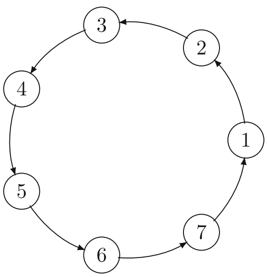
which also forms a stationary distribution $(\pi_i)_{i\in\mathbb{S}} = (1/\mu_i(i))_{i\in\mathbb{S}}$, uniquely determined by the equation

$$\pi = \pi P.$$

In Theorem 7.8 above, Condition (ii), is satisfied from Proposition 6.14, provided that at least one state is aperiodic, since the chain is irreducible.

See Exercise 7.21 for an application of Theorem 7.8 on an infinite state space.

In the following trivial example of a finite circular chain, Theorems 7.2 and 7.8 cannot be applied since the chain is not aperiodic, and it clearly does not admit a limiting distribution. However, Theorem 7.10 below applies and the chain admits a stationary distribution: one can easily check that $\mu_k(k) = n$ and $\pi_k = 1/n = 1/\mu_k(k)$, $k = 1, 2, \ldots, n$, with $n = 7$.



$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (7.2.4)$$

In view of Theorem 6.11, we have the following corollary of Theorem 7.8:

Corollary 7.9 *Consider an irreducible aperiodic Markov chain with finite state space. Then the limiting probabilities*

$$\pi_i := \lim_{n\to\infty} \mathbb{P}(X_n = i \mid X_0 = j) = \frac{1}{\mu_i(i)}, \qquad i, j \in \mathbb{S},$$

exist and form a stationary distribution which is uniquely determined by the equation

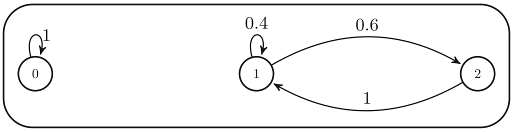
$$\pi = \pi P.$$

Corollary 7.9 can also be applied separately to derive a stationary distribution on each closed component of a reducible chain.

The convergence of the two-state chain to its stationary distribution has been illustrated in Fig. 4.4. Before proceeding further we make some comments on the assumptions of Theorems 7.2 and 7.8.

Remarks

- *Irreducibility.*
  The irreducibility assumption on the chain in Theorems 7.2 and 7.8 is truly required in general, as a reducible chain may have a limiting distribution that depends on the initial state as in the following trivial example on the state space $\{0, 1, 2\}$:

  in which the chain is aperiodic and positive recurrent, but not irreducible. Note that the sub-chain $\{1, 2\}$ admits $[\pi_1, \pi_2] = [1/1.6, 0.6/1.6]$ as stationary and limiting distribution, however any vector of the form $(1 - \alpha, \alpha\pi_1, \alpha\pi_2)$ is also a stationary distribution on $\mathbb{S} = \{0, 1, 2\}$ for any $\alpha \in [0, 1]$, showing the non uniqueness of the stationary distribution.
  More generally, in case the chain is not irreducible we can split it into subchains and consider the subproblems separately. For example, when the state space $\mathbb{S}$ is a finite set it admits at least one communicating class $A \subset \mathbb{S}$ that leads to no other class, and admits a stationary distribution π_A by Corollary 7.11 since it is irreducible, hence a chain with finite state space $\mathbb{S}$ admits at least one stationary distribution of the form $(0, 0, \ldots, 0, \pi_A)$.
  Similarly, the constant two-state Markov chain with transition matrix $P = I_d$ is reducible, it admits an infinity of stationary distributions, and a limiting distribution which is dependent on the initial state.
- *Aperiodicity.*
  The conclusions of Theorems 7.2, 7.8 and Corollary 7.9 ensure the existence of the limiting distribution by requiring the aperiodicity of the Markov chain. Indeed, the limiting distribution may not exist when the chain is not aperiodic. For example, the two-state Markov chain with transition matrix

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

is not aperiodic (both states have period 2) and it has no limiting distribution because[2]

$$\mathbb{P}(X_{2n} = 1 \mid X_0) = 1 \quad \text{and} \quad \mathbb{P}(X_{2n+1} = 1 \mid X_0) = 0, \qquad n \in \mathbb{N}.$$

The chain does have an invariant (or stationary) distribution π solution of $\pi = \pi P$, and given by

$$\pi = [\pi_0, \pi_1] = \left[\frac{1}{2}, \frac{1}{2}\right].$$

- *Positive recurrence.*
Theorems 7.8 and 7.10 below, and Corollary 7.9 do not apply to the unrestricted random walk $(S_n)_{n\in\mathbb{N}}$ of Chap. 3, because this chain is not positive recurrent, cf. Relations (3.4.20) and (3.4.17), and admits no stationary distribution.
If a stationary distribution $\pi = (\pi_i)_{i\in\mathbb{Z}}$ existed it would satisfy the equation $\pi = \pi P$ which, according to (4.3.1), would read

$$\pi_i = p\pi_{i-1} + q\pi_{i+1}, \qquad i \in \mathbb{Z},$$

i.e.

$$(p+q)\pi_i = p\pi_{i-1} + q\pi_{i+1}, \qquad i \in \mathbb{Z},$$

or

$$\pi_{i+1} - \pi_i = \frac{p}{q}(\pi_i - \pi_{i-1}), \qquad i \in \mathbb{Z}.$$

As in the direct solution method of p. 48, this implies

$$\pi_{i+1} - \pi_i = \left(\frac{p}{q}\right)^i (\pi_1 - \pi_0), \qquad i \in \mathbb{N},$$

so that by a telescoping summation argument we have

$$\begin{aligned}
\pi_k &= \pi_0 + \sum_{i=0}^{k-1}(\pi_{i+1} - \pi_i) \\
&= \pi_0 + (\pi_1 - \pi_0)\sum_{i=0}^{k-1}\left(\frac{p}{q}\right)^i \\
&= \pi_0 + (\pi_1 - \pi_0)\frac{1-(p/q)^k}{1-p/q}, \qquad k \in \mathbb{N},
\end{aligned}$$

[2]This two-state chain is a particular case of the circular chain (7.2.4) for $n = 2$.

which cannot satisfy the condition $\sum_{k\in\mathbb{Z}} \pi_k = 1$, with $p \neq q$. When $p = q = 1/2$ we similarly obtain

$$\pi_k = \pi_0 + \sum_{i=0}^{k-1} (\pi_{i+1} - \pi_i) = \pi_0 + k(\pi_1 - \pi_0), \qquad k \in \mathbb{Z},$$

and in this case as well, the sequence $(\pi_k)_{k\in\mathbb{N}}$ cannot satisfy the condition $\sum_{k\in\mathbb{Z}} \pi_k = 1$, and we conclude that the chain does not admit a stationary distribution. Hence the stationary distribution of a Markov chain may not exist at all.
In addition, it follows from (3.3.3) and the Stirling approximation formula that

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}(S_{2n} = 2k \mid S_0 = 0) &= \lim_{n\to\infty} \frac{(2n)!}{(n+k)!(n-k)!} p^{n+k} q^{n-k} \\
&\leq \lim_{n\to\infty} \frac{(2n)!}{2^{2n} n!^2} \\
&= \lim_{n\to\infty} \frac{1}{\sqrt{\pi n}} = 0, \qquad k \in \mathbb{N},
\end{aligned}$$

so that the limiting distribution does not exist as well. Here, Theorem 7.2 cannot be applied because the chain is not aperiodic (it has period 2), however aperiodicity and irreducibility are not sufficient in general when the state space is infinite, cf. e.g. the model of Exercise 2.1.

The following theorem gives sufficient conditions for the existence of a stationary distribution, without requiring aperiodicity or finiteness of the state space. As noted above, the limiting distribution may not exist in this case.

Theorem 7.10 ([BN96], Theorem 4.1) *Consider a Markov chain $(X_n)_{n\in\mathbb{N}}$ satisfying the following two conditions:*

(i) positive recurrence, and
(ii) irreducibility.

Then the probabilities

$$\pi_i = \frac{1}{\mu_i(i)}, \qquad i \in \mathbb{S},$$

form a stationary distribution which is uniquely determined by the equation $\pi = \pi P$.

Note that the conditions stated in Theorem 7.10 are sufficient, but they are not all necessary. For example, Condition (ii) is not necessary as the trivial constant chain, whose transition matrix $P = I_d$ is reducible, does admit a stationary distribution.

Note that the positive recurrence assumption in Theorem 7.2 is required in general on infinite state spaces. For example, the process in Exercise 7.21 is positive recurrent for $\alpha < 1$ only, whereas no stationary distribution exists when $\alpha \geq 1$.

As a consequence of Corollary 6.12 we have the following corollary of Theorem 7.10, which does not require aperiodicity for the stationary distribution to exist.

Corollary 7.11 ([BN96]) *Let* $(X_n)_{n\in\mathbb{N}}$ *be an irreducible Markov chain with* finite *state space* $\mathbb{S}$. *Then the probabilities*

$$\pi_k = \frac{1}{\mu_k(k)}, \qquad k \in \mathbb{S},$$

form a stationary distribution which is uniquely determined by the equation

$$\pi = \pi P.$$

According to Corollary 7.11, the limiting distribution and stationary distribution both exist (and coincide) when the chain is irreducible aperiodic with finite state space, and in this case we have $\pi_k > 0$ for all $k \in \mathbb{S}$ by Corollaries 6.12 and 7.11. When the chain is irreducible it is usually easier to compute the stationary distribution, which will give us the limiting distribution.

Under the assumptions of Theorem 7.8, if the stationary and limiting distributions both exist then they are equal and in this case we only need to compute one of them. However, in some situations only the stationary distribution might exist. According to Corollary 7.11 above the stationary distribution always exists when the chain is irreducible with finite state space, nevertheless the limiting distribution may not exist if the chain is not aperiodic, consider for example the two-state switching chain with $a = b = 1$.

Finding a Limiting Distribution

In summary:

- We usually attempt first to compute the stationary distribution whenever possible, and this also gives the limiting distribution when it exists. For this, we first check whether the chain is positive recurrent, aperiodic and irreducible, in which case the limiting distribution can be found by solving $\pi = \pi P$ according to Theorem 7.8.
- In case the above properties are not satisfied we need to compute the limiting distribution by taking the limit $\lim_{n\to\infty} P^n$ of the powers P^n of the transition matrix, if possible by decomposing the state space in communicating classes as in e.g. Exercise 7.11. This can turn out to be much more complicated and done only in special cases. If the chain has period $d \geq 2$ we may need to investigate the limits $\lim_{n\to\infty} P^{nd}$ instead, see e.g. Exercise 7.3 and (7.2.5)–(7.2.6) below.

To further summarize, we note that by Theorem 7.2 we have

(a) irreducible + recurrent + aperiodic $\Longrightarrow$ existence of a limiting distribution, by Theorem 7.8 we get

(b) irreducible + positive recurrent + aperiodic $\Longrightarrow$ existence of a limiting distribution which is also stationary, and by Theorem 7.10 we get
(c) irreducible + positive recurrent $\Longrightarrow$ existence of a stationary distribution.

In addition, the limiting or stationary distribution $\pi = (\pi_i)_{i\in\mathbb{S}}$ satisfies

$$\pi_i = \frac{1}{\mu_i(i)}, \qquad i \in \mathbb{S},$$

in all above cases (a), (b) and (c).

The Ergodic Theorem

The *Ergodic Theorem*, cf. e.g. Theorem 1.10.2 of [Nor98] states the following.

Theorem 7.12 *Assume that the chain $(X_n)_{n\in\mathbb{N}}$ is* irreducible. *Then the sample average time spent at to state ⓘ converges almost surely to $1/\mu_i(i)$, i.e.*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\{X_k=i\}} = \frac{1}{\mu_i(i)}, \qquad i \in \mathbb{S}.$$

In case $(X_n)_{n\in\mathbb{N}}$ is also positive recurrent, Theorem 7.10 shows that we also have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\{X_k=i\}} = \pi_i, \qquad i \in \mathbb{S},$$

where $(\pi_i)_{i\in\mathbb{S}}$ is the stationary distribution of $(X_n)_{n\in\mathbb{N}}$. We refer to Fig. 4.4 for an illustration of convergence in the setting of the Ergodic Theorem 7.12.

Example. Consider the maze random walk (5.3.7) with transition matrix

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

The equation $\pi = \pi P$ yields

$$\begin{cases} \pi_1 = \dfrac{1}{2}\pi_2 \\ \pi_2 = \pi_1 + \dfrac{1}{2}\pi_3 \\ \pi_3 = \dfrac{1}{2}\pi_2 + \dfrac{1}{2}\pi_6 \\ \pi_4 = \dfrac{1}{2}\pi_7 \\ \pi_5 = \dfrac{1}{2}\pi_6 + \dfrac{1}{3}\pi_8 \\ \pi_6 = \dfrac{1}{2}\pi_3 + \dfrac{1}{2}\pi_5 \\ \pi_7 = \pi_4 + \dfrac{1}{3}\pi_8 \\ \pi_8 = \dfrac{1}{2}\pi_5 + \dfrac{1}{2}\pi_7 + \pi_9 \\ \pi_9 = \dfrac{1}{3}\pi_8, \end{cases} \qquad \text{hence} \qquad \begin{cases} \pi_1 = \dfrac{1}{2}\pi_2 \\ \pi_2 = \pi_3 \\ \pi_3 = \pi_6 \\ \pi_4 = \dfrac{1}{2}\pi_3 \\ \dfrac{1}{2}\pi_7 = \dfrac{1}{3}\pi_8 \\ \pi_6 = \pi_5 \\ \pi_6 = \pi_7 \\ \pi_9 = \dfrac{1}{3}\pi_8, \end{cases}$$

and

$$\begin{aligned} 1 &= \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6 + \pi_7 + \pi_8 + \pi_9 \\ &= \pi_1 + 2\pi_1 + 2\pi_1 + \pi_1 + 2\pi_1 + 2\pi_1 + 2\pi_1 + 3\pi_1 + \pi_1 \\ &= 16\pi_1, \end{aligned}$$

hence

$$\begin{aligned} &\pi_1 = \frac{1}{16},\ \pi_2 = \frac{2}{16},\ \pi_3 = \frac{2}{16},\ \pi_4 = \frac{1}{16}, \\ &\pi_5 = \frac{2}{16},\ \pi_6 = \frac{2}{16},\ \pi_7 = \frac{2}{16},\ \pi_8 = \frac{3}{16},\ \pi_9 = \frac{1}{16}, \end{aligned}$$

cf. Figs. 7.3 and 7.2, and we check that since $\mu_1(1) = 16$ by (5.3.8), we indeed have

$$\pi_1 = \frac{1}{\mu_1(1)} = \frac{1}{16},$$

according to Corollary 7.11.

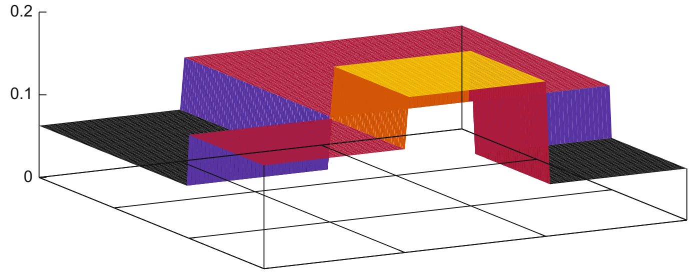


Fig. 7.2 Stationary distribution on the maze (5.3.7)

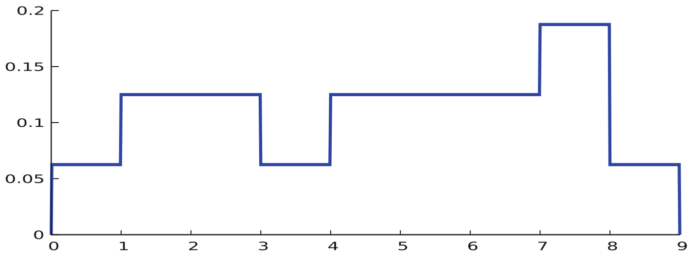
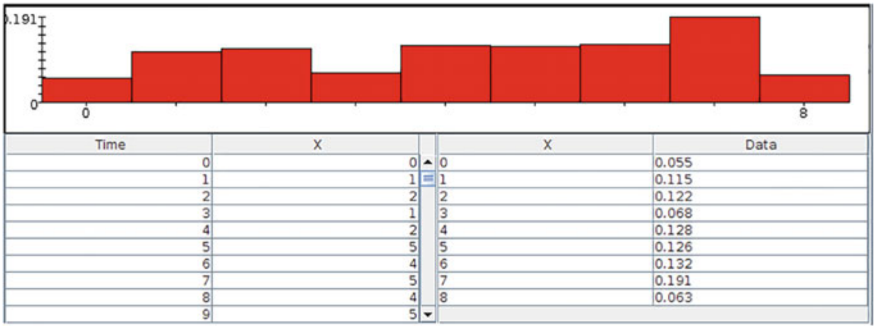


Fig. 7.3 Stationary distribution by state numbering



Time	X	X	Data
0	0	0	0.055
1	1	1	0.115
2	2	2	0.122
3	1	3	0.068
4	2	4	0.128
5	5	5	0.126
6	4	6	0.132
7	5	7	0.191
8	4	8	0.063
9	5		

Fig. 7.4 Simulated stationary distribution

The stationary probability distribution of Figs. 7.3 and 7.2 can be compared to the proportions of time spent at each state simulated in Fig. 7.4.

Note that this chain has period 2 and the matrix powers $(P^n)_{n\in\mathbb{N}}$ do not converge as n tends to infinity, i.e. it does not admit a limiting distribution. In fact, using the following Matlab/Octave commands:

```
P = [0,1,0,0,0,0,0,0,0;
1/2,0,1/2,0,0,0,0,0,0;
0,1/2,0,0,0,1/2,0,0,0;
0,0,0,0,0,0,1,0,0;
0,0,0,0,0,1/2,0,1/2,0;
0,0,1/2,0,1/2,0,0,0,0;
0,0,0,1/2,0,0,0,1/2,0;
0,0,0,0,1/3,0,1/3,0,1/3;
0,0,0,0,0,0,0,1,0]
mpower(P,1000)
mpower(P,1001)
```

we see that

$$\lim_{n\to\infty} P^{2n} = \begin{bmatrix} 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/8 \\ 0 & 1/4 & 0 & 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/8 \\ 0 & 1/4 & 0 & 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/8 \\ 0 & 1/4 & 0 & 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/8 \\ 0 & 1/4 & 0 & 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/8 \end{bmatrix}, \tag{7.2.5}$$

and

$$\lim_{n\to\infty} P^{2n+1} = \begin{bmatrix} 0 & 1/4 & 0 & 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/8 \\ 0 & 1/4 & 0 & 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/8 \\ 0 & 1/4 & 0 & 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/8 \\ 0 & 1/4 & 0 & 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/8 \\ 0 & 1/4 & 0 & 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \end{bmatrix}, \tag{7.2.6}$$

which shows that, although $(P^n)_{n\geq 1}$ admits two converging subsequences, $\lim_{n\to\infty} P^n$ does not exist, therefore the chain does not admit a limiting distribution.

7.3 Markov Chain Monte Carlo

The goal of the Markov Chain Monte Carlo (MCMC) method, or Metropolis algorithm, is to generate random samples according to a target distribution $\pi = (\pi_i)_{i \in \mathbb{S}}$ via a Markov chain that admits π as limiting and stationary distribution. It applies in particular in the setting of huge state spaces $\mathbb{S}$.

A Markov chain $(X_k)_{k \in \mathbb{N}}$ with transition matrix P on a state space $\mathbb{S}$ is said to satisfy the *detailed balance* (or reversibility) condition with respect to the probability distribution $\pi = (\pi_i)_{i \in S}$ if

$$\pi_i P_{i,j} = \pi_j P_{j,i}, \qquad i, j \in \mathbb{S}. \tag{7.3.1}$$

See Fig. 7.5. Note that the detailed balance condition (7.3.1) implies the global balance condition (7.2.2) as, by summation over $i \in \mathbb{S}$ in (7.3.1) we have

$$\sum_{i \in \mathbb{S}} \pi_i P_{i,j} = \sum_{i \in \mathbb{S}} \pi_j P_{j,i} = \pi_j \sum_{i \in \mathbb{S}} P_{j,i} = \pi_j, \quad j \in \mathbb{S},$$

which shows that $\pi P = \pi$, i.e. π is a stationary distribution for P, cf. e.g. Problem 7.23-(c).

If the transition matrix P satisfies the detailed balance condition with respect to π then the probability distribution of X_n will naturally converge to the stationary distribution π in the long run, e.g. under the hypotheses of Theorem 7.8, i.e. when the chain $(X_k)_{k \in \mathbb{N}}$ is positive recurrent, aperiodic, and irreducible.

In general, however, the detailed balance (or reversibility) condition (7.3.1) may not be satisfied by π and P. In this case one can construct a modified transition matrix $\tilde{P}$ that will satisfy the detailed balance condition with respect to π. This modified transition matrix $\tilde{P}$ is defined by

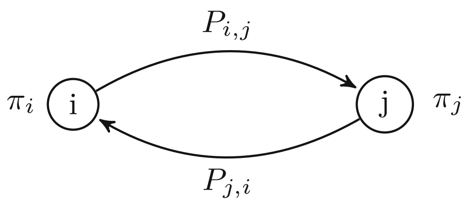


Fig. 7.5 Detailed balance condition (discrete time)

$$\tilde{P}_{i,j} := P_{i,j} \times \min\left(1, \frac{\pi_j P_{j,i}}{\pi_i P_{i,j}}\right) = \begin{cases} P_{j,i}\dfrac{\pi_j}{\pi_i} & \text{if } \pi_j P_{j,i} < \pi_i P_{i,j}, \\ \\ P_{i,j} & \text{if } \pi_j P_{j,i} \geq \pi_i P_{i,j}, \end{cases}$$

for $i \neq j$, and

$$\tilde{P}_{i,i} = 1 - \sum_{k \neq i \in \mathbb{S}} \tilde{P}_{i,k} = P_{i,i} + \sum_{i \neq k \in \mathbb{S}} P_{i,k}\left(1 - \min\left(1, \frac{\pi_j P_{j,i}}{\pi_i P_{i,j}}\right)\right), \quad i \in \mathbb{S}.$$

Clearly, we have $\tilde{P} = P$ when the detailed balance (or reversibility) condition (7.3.1) is satisfied. In the general case, we can check that for $i \neq j$ we have

$$\pi_i \tilde{P}_{i,j} = \left\{ \begin{array}{ll} P_{j,i}\pi_j = \pi_j \tilde{P}_{j,i} & \text{if } \pi_j P_{j,i} < \pi_i P_{i,j}, \\ \\ \pi_i P_{i,j} = \pi_j \tilde{P}_{j,i} & \text{if } \pi_j P_{j,i} \geq \pi_i P_{i,j}, \end{array} \right\} = \pi_j \tilde{P}_{j,i},$$

hence $\tilde{P}$ satisfies the detailed balance condition with respect to π (the condition is obviously satisfied when $i = j$). Therefore, the random simulation of $(\tilde{X}_n)_{n\in\mathbb{N}}$ according to the transition matrix $\tilde{P}$ will provide samples of the distribution π in the long run as n tends to infinity, provided that the chain $(\tilde{X}_n)_{n\in\mathbb{N}}$ is positive recurrent, aperiodic, and irreducible.

In Table 7.1 we summarize the definitions introduced in this chapter and in Chap. 6.

Table 7.1 Summary of Markov chain properties

Property	Definition
Absorbing (state)	$P_{i,i} = 1$
Recurrent (state)	$\mathbb{P}(T_i^r < \infty \mid X_0 = i) = 1$
Transient (state)	$\mathbb{P}(T_i^r < \infty \mid X_0 = i) < 1$
Positive recurrent (state)	Recurrent *and* $\mathbb{E}[T_i^r \mid X_0 = i] < \infty$
Null recurrent (state)	Recurrent *and* $\mathbb{E}[T_i^r \mid X_0 = i] = \infty$
Aperiodic (state or chain)	Period(s) = 1
Ergodic (state or chain)	Positive recurrent *and* aperiodic
Irreducible (chain)	*All* states communicate
Regular (chain)	All coefficients of P^n are > 0 for some $n \geq 1$
Stationary distribution π	Obtained from solving $\pi = \pi P$

Exercises

Exercise 7.1 We consider the Markov chain of Exercise 4.10-(a).

(a) Is the chain irreducible, aperiodic, recurrent, positive recurrent?
(b) Does it admit a stationary distribution?
(c) Does it admit a limiting distribution?

Exercise 7.2 We consider the success runs Markov chain of Exercise 4.10-(b).

(a) Is the success runs chain irreducible, aperiodic, recurrent, positive recurrent?
(b) Does it admit a stationary distribution?
(c) Does it admit a limiting distribution?

Exercise 7.3 We consider the Ehrenfest chain (4.3.2)–(4.3.3).

(a) Is the Ehrenfest chain irreducible, aperiodic, recurrent, positive recurrent?
(b) Does it admit a stationary distribution?
(c) Does it admit a limiting distribution?

Hint: Try a binomial distribution.

Exercise 7.4 Consider the Bernoulli–Laplace chain $(X_n)_{n\in\mathbb{N}}$ of Exercise 4.9 with state space $\{0, 1, 2, \ldots, N\}$ and transition matrix

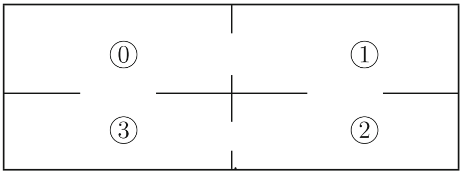
$$P = \begin{bmatrix}
0 & 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\
1/N^2 & 2(N-1)/N^2 & (N-1)^2/N^2 & 0 & \cdots & \cdots & 0 & 0 \\
0 & 2^2/N^2 & 4(N-2)/N^2 & (N-2)^2/N^2 & \cdots & \cdots & 0 & 0 \\
0 & 0 & 3^2/N^2 & 0 & \cdots & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & \vdots & 0 & 3^2/N^2 & 0 & 0 \\
0 & 0 & \cdots & 0 & (N-2)^2/N^2 & 4(N-2)/N^2 & 2^2/N^2 & 0 \\
0 & 0 & \cdots & 0 & 0 & (N-1)^2/N^2 & 2(N-1)/N^2 & 1/N^2 \\
0 & 0 & \cdots & 0 & 0 & 0 & 1 & 0
\end{bmatrix},$$

i.e.

$$P_{k,k-1} = \frac{k^2}{N^2}, \quad P_{k,k} = \frac{2k(N-k)^2}{N^2}, \quad P_{k,k+1} = \frac{(N-k)^2}{N^2}, \quad k = 1, 2, \ldots, N-1.$$

(a) Is the Bernoulli–Laplace chain irreducible, aperiodic, recurrent, positive recurrent?
(b) Does it admit a stationary distribution?
(c) Does it admit a limiting distribution?

Exercise 7.5 Consider a robot evolving in the following circular maze, moving from one room to the other according to a Markov chain with equal probabilities.



Let $X_n \in \{0, 1, 2, 3\}$ denote the state of the robot at time $n \in \mathbb{N}$.

(a) Write down the transition matrix P of the chain.
(b) By first step analysis, compute the mean *return* times $\mu_0(k)$ from state $k = 0, 1, 2, 3$ to state ⓪.[3]
(c) Guess an invariant (or stationary) probability distribution $[\pi_0, \pi_1, \pi_2, \pi_3]$ for the chain, and show that it does satisfy the condition $\pi = \pi P$.

Exercise 7.6 A signal processor is analysing a sequence of signals that can be either distorted or non-distorted. It turns out that on average, 1 out of 4 signals following a distorted signal are distorted, while 3 out of 4 signals are non-distorted following a non-distorted signal.

(a) Let $X_n \in \{D, N\}$ denote the state of the nth signal being analysed by the processor. Show that the process $(X_n)_{n\geq 1}$ can be modeled as a Markov chain and determine its transition matrix.
(b) Compute the stationary distribution of $(X_n)_{n\geq 1}$.
(c) In the long run, what fraction of analysed signals are distorted?
(d) Given that the last observed signal was distorted, how long does it take on average until the next non-distorted signal?
(e) Given that the last observed signal was non-distorted, how long does it take on average until the next distorted signal?

Exercise 7.7 Consider a Markov chain $(X_n)_{n\geq 0}$ on the state space $\{0, 1, 2, 3\}$ with transition probability matrix P given by

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.2 & 0 & 0.8 & 0 \\ 0.3 & 0 & 0.7 & 0 \\ 0.4 & 0.6 & 0 & 0 \end{bmatrix}.$$

(a) Draw the graph of this chain. Is the chain reducible?
(b) Find the recurrent, transient, and absorbing state(s) of this chain.
(c) Compute the fraction of time spent at state ⓪ in the long run.
(d) On the average, how long does it take to reach state ⓪ after starting from state ②?

[3] You may use the symmetry of the problem to simplify the calculations.

Exercise 7.8 Consider the transition probability matrix

$$P = [\ P_{i,j}\ ]_{0\leq i,j\leq 3} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.1 & 0.4 & 0.2 & 0.3 \\ 0.2 & 0.2 & 0.5 & 0.1 \\ 0.3 & 0.3 & 0.4 & 0 \end{bmatrix}.$$

(a) Compute the limiting distribution $[\pi_0, \pi_1, \pi_2, \pi_3]$ of this Markov chain.
(b) Compute the average time $\mu_0(1)$ it takes to the chain to travel from state ① to state ⓪.
Hint: The data of the first row of the matrix P should play no role in the computation of $\mu_0(k)$, $k = 0, 1, 2, 3$.
(c) Prove by direct computation that the relation $\pi_0 = 1/\mu_0(0)$ holds, where $\mu_0(0)$ represents the mean return time to state ⓪ for this chain.

Exercise 7.9 Consider the Markov chain with transition probability matrix

$$[\ P_{i,j}\ ]_{0\leq i,j\leq 3} = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/4 & 0 & 3/4 & 0 \\ 0 & 1/3 & 0 & 2/3 \\ 1/2 & 0 & 1/2 & 0 \end{bmatrix}.$$

(a) Show that the chain is periodic[4] and compute its period.
(b) Determine the stationary distribution of this chain.

Exercise 7.10 The lifetime of a given component of a machine is a discrete random variable T with distribution

$$\mathbb{P}(T = 1) = 0.1, \quad \mathbb{P}(T = 2) = 0.2, \quad \mathbb{P}(T = 3) = 0.3, \quad \mathbb{P}(T = 4) = 0.4.$$

The component is immediately replaced with a new component upon failure, and the machine starts functioning with a new component. Compute the long run probability of finding the machine about to fail at the next time step.

Exercise 7.11 Suppose that a Markov chain has the one-step transition probability matrix P on the state space $\{A, B, C, D, E\}$ given by

$$P = \begin{bmatrix} 0.6 & 0.4 & 0 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 & 0 \\ 0.2 & 0 & 0.4 & 0 & 0.4 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Find $\lim_{n\to\infty} \mathbb{P}(X_n = A \mid X_0 = C)$.

[4] A chain is periodic when all states have the same period.

Exercise 7.12 Consider a Markov chain $(X_n)_{n\geq 0}$ on the state space $\{0, 1, 2, 3, 4\}$ with transition probability matrix P given by

$$P = \begin{bmatrix} 1/3 & 2/3 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1/7 & 6/7 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

(a) Draw the graph of this chain.
(b) Identify the communicating class(es).
(c) Find the recurrent, transient, and absorbing state(s) of this chain.
(d) Find $\lim_{n\to\infty} \mathbb{P}(X_n = 0 \mid X_0 = 4)$.

Exercise 7.13 Three out of 4 trucks passing under a bridge are followed by a car, while only 1 out of every 5 cars passing under that same bridge is followed by a truck. Let $X_n \in \{C, T\}$ denote the nature of the nth vehicle passing under the bridge, $n \geq 1$.

(a) Show that the process $(X_n)_{n\geq 1}$ can be modeled as a Markov chain and write down its transition matrix.
(b) Compute the stationary distribution of $(X_n)_{n\geq 1}$.
(c) In the long run, what fraction of vehicles passing under the bridge are trucks?
(d) Given that the last vehicle seen was a truck, how long does it take on average until the next truck is seen under that same bridge?

Exercise 7.14 Consider a discrete-time Markov chain $(X_n)_{n\in\mathbb{N}}$ on $\mathbb{S} = \{1, 2, \ldots, N\}$, whose transition matrix $P = (P_{i,j})_{1\leq i,j\leq N}$ is assumed to be *symmetric*, i.e. $P_{i,j} = P_{j,i}$, $1 \leq i, j \leq N$,

(a) Find an invariant (or stationary) distribution of the chain.
Hint: The equation $\pi P = \pi$ admits an easy solution.
(b) Assume further that $P_{i,i} = 0$, $1 \leq i \leq N$, and that $P_{i,j} > 0$ for all $1 \leq i < j \leq N$. Find the period of every state.

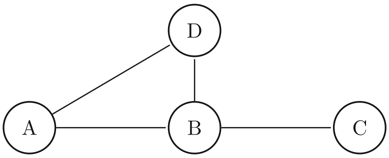
Exercise 7.15 Consider the Markov chain with transition matrix

$$\begin{bmatrix} q & p & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ q & 0 & 0 & p & 0 \\ q & 0 & 0 & 0 & p \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix},$$

where $p, q \in (0, 1)$ satisfy $p + q = 1$.

(a) Compute the stationary distribution $[\pi_0, \pi_1, \pi_2, \pi_3, \pi_4]$ of this chain.
(b) Compute the limiting distribution of the chain.

Exercise 7.16 Four players A, B, C, D are connected by the following network, and play by exchanging a token.



At each step of the game, the player who holds the token chooses another player he is connected to, and sends the token to that player.

(a) Assuming that the player choices are made at random and are equally distributed, model the states of the token as a Markov chain $(X_n)_{n\geq 1}$ on $\{A, B, C, D\}$ and give its transition matrix.
(b) Compute the stationary distribution $[\pi_A, \pi_B, \pi_C, \pi_D]$ of $(X_n)_{n\geq 1}$.
Hint: To simplify the resolution, start by arguing that we have $\pi_A = \pi_D$.
(c) Compute the mean return times $\mu_D(i)$, $i \in \{A, B, C, D\}$. On average, how long does player D have to wait to recover the token?
(d) In the long run, what is the probability that player D holds the token?

Exercise 7.17 Consider the Markov chain with transition probability matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & b & c \end{bmatrix},$$

with $a + b + c = 1$.

(a) Compute the power P^n for all $n \geq 2$.
(b) Does the chain admit a limiting distribution? If yes, compute this distribution.
(c) Does the chain admit a stationary distribution? Compute this distribution if it exists.

Exercise 7.18 Consider a game server that can become offline with probability p and can remain online with probability $q = 1 - p$ on any given day. Assume that the random time N it takes to fix the server has the geometric distribution

$$\mathbb{P}(N = k) = \beta(1 - \beta)^{k-1}, \qquad k \geq 1,$$

with parameter $\beta \in (0, 1)$. We let $X_n = 1$ when the server is online on day n, and $X_n = 0$ when it is offline.

(a) Show that the process $(X_n)_{n\in\mathbb{N}}$ can be modeled as a discrete-time Markov chain and write down its transition matrix.
(b) Compute the probability that the server is online in the long run, in terms of the parameters β and p.

Exercise 7.19 Let $(X_n)_{n\in\mathbb{N}}$ be an irreducible aperiodic Markov chain on the finite state space $\mathbb{S} = \{1, 2, \ldots, N\}$.

(a) Show that there *exists* a state $i \in \{1, 2, \ldots, N\}$ such that the mean return time $\mu_i(i)$ from state i to itself is lower or equal to N, i.e. $\mu_i(i) \leq N$.
(b) Show that there *exists* a state $i \in \{1, 2, \ldots, N\}$ such that the mean return time $\mu_i(i)$ from state i to itself is higher or equal to N, i.e. $\mu_i(i) \geq N$.

Exercise 7.20 Consider a Markov chain on the state space $\{1, 2, \ldots, N\}$. For any $i \in \{2, \ldots, N-1\}$, the chain has probability $p \in (0, 1)$ to switch from state i to state $\boxed{i+1}$, and probability $q = 1 - p$ to switch from i to $\boxed{i\text{-}1}$. When the chain reaches state i=1 it rebounds to state 2 with probability p or stays at state 1 with probability q. Similarly, after reaching state N it rebounds to state $\boxed{N\text{-}1}$ with probability q, or remains at N with probability p.

(a) Write down the transition probability matrix of this chain.
(b) Is the chain reducible?
(c) Determine the absorbing, transient, recurrent, and positive recurrent states of this chain.
(d) Compute the stationary distribution of this chain.
(e) Compute the limiting distribution of this chain.

Exercise 7.21 (*Problem 6.9 continued*). Let $\alpha > 0$ and consider the Markov chain with state space $\mathbb{N}$ and transition matrix given by

$$P_{i,i-1} = \frac{1}{\alpha+1}, \quad P_{i,i+1} = \frac{\alpha}{\alpha+1}, \qquad i \geq 1.$$

and a reflecting barrier at 0, such that $P_{0,1} = 1$.

(a) Show that when $\alpha < 1$ this chain admits a stationary distribution of the form

$$\pi_k = \alpha^{k-1}(1-\alpha^2)/2, \qquad k \geq 1,$$

where the value of π_0 has to be determined.
(b) Does the chain admit a stationary distribution when $\alpha \geq 1$?
(c) Show that the chain is positive recurrent when $\alpha < 1$.

Exercise 7.22 Consider two discrete-time stochastic processes $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ on a state space $\mathbb{S}$, such that

$$X_n = Y_n, \qquad n \geq \tau,$$

where τ is a random time called the *coupling time* of $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$.

(a) Show that for all $x \in \mathbb{S}$ and $n \in \mathbb{N}$ we have

$$\mathbb{P}(X_n = x) \leq \mathbb{P}(Y_n = x) + \mathbb{P}(\tau > n) \qquad x \in \mathbb{S}, \quad n \in \mathbb{N}.$$

Hint: Use the law of total probability as $\mathbb{P}(A) = \mathbb{P}(A \cap \{\tau \leq n\}) + \mathbb{P}(A \cap \{\tau > n\})$.

(b) Show that for all $n \in \mathbb{N}$ we have

$$\sup_{x \in \mathbb{S}} |\mathbb{P}(X_n = x) - \mathbb{P}(Y_n = x)| \leq \mathbb{P}(\tau > n), \qquad n \in \mathbb{N}.$$

Problem 7.23 Reversibility is a fundamental issue in physics as it is akin to the idea of "traveling backward in time". This problem studies the reversibility of Markov chains, and applies it to the computation of stationary and limiting distributions. Given $N \geq 1$ and $(X_k)_{k=0,1,\ldots,N}$ a Markov chain with transition matrix P on a state space S, we let

$$Y_k := X_{N-k}, \qquad k = 0, 1, \ldots, N,$$

denote the *time reversal* of $(X_k)_{k=0,1,\ldots,N}$.

(a) Assume that X_k has same distribution $\pi = (\pi_i)_{i \in S}$ for every $k = 0, 1, \ldots, N$, i.e.

$$\mathbb{P}(X_k = i) = \pi_i, \qquad i \in S, \quad k = 0, 1, \ldots, N.$$

Show that the process $(Y_k)_{k=0,1,\ldots,N}$ is a (forward) Markov chain (i.e. $(X_k)_{k=0,1,\ldots,N}$ has the backward Markov property), and find the transition matrix

$$[\mathbb{P}(Y_{n+1} = j \mid Y_n = i)]_{i,j}$$

in terms of P and π.
Hint: Use the basic definition of conditional probabilities to compute

$$[\mathbb{P}(Y_{n+1} = j \mid Y_n = i)]_{i,j},$$

and then show that $(Y_n)_{n=0,1,\ldots,N}$ has the Markov property

$$\mathbb{P}(Y_{n+1} = j \mid Y_n = i_n, \ldots, Y_0 = i_0) = \mathbb{P}(Y_{n+1} = j \mid Y_n = i_n).$$

(b) We say that $(X_k)_{k=0,1,\ldots,N}$ is *reversible* for π when $(X_k)_{k=0,1,\ldots,N}$ and $(Y_k)_{k=0,1,\ldots,N}$ have same transition probabilities.
Write down this reversibility condition in terms of P and π. From now on we refer to that condition as the *detailed balance condition*, which can be stated independently of N.
Hint: By "same transition probabilities" we mean

$$\mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(Y_{n+1} = j \mid Y_n = i).$$

(c) Show that if $(X_k)_{k \in \mathbb{N}}$ is *reversible* for π, then π is also a stationary distribution for $(X_k)_{k \in \mathbb{N}}$.
Hint: The fact that $\sum_i P_{j,i} = 1$ plays a role here.

(d) Show that if an irreducible positive recurrent aperiodic Markov chain is reversible for its stationary distribution π then we have

$$P_{k_1,k_2}P_{k_2,k_3}\cdots P_{k_{n-1},k_n}P_{k_n,k_1} = P_{k_1,k_n}P_{k_n,k_{n-1}}\cdots P_{k_3,k_2}P_{k_2,k_1} \tag{7.3.2}$$

for all sequences $\{k_1, k_2, \ldots, k_n\}$ of states and $n \geq 2$.
Hint: This is a standard algebraic manipulation.

(e) Show that conversely, if an irreducible positive recurrent aperiodic Markov chain satisfies Condition (7.3.2) for all sequences $\{k_1, k_2, \ldots, k_n\}$ of states, $n \geq 2$, then it is reversible for its stationary distribution π.
Hint: This question is more difficult and here you need to apply Theorem 7.8.

(f) From now on we assume that $S = \{0, 1, \ldots, M\}$ and that P is the transition matrix

$$P_{i,i+1} = \frac{1}{2} - \frac{i}{2M}, \qquad P_{i,i} = \frac{1}{2}, \qquad P_{i,i-1} = \frac{i}{2M}, \qquad 1 \leq i \leq M-1,$$

of the *modified* Ehrenfest chain, with $P_{0,0} = P_{0,1} = P_{M,M-1} = P_{M,M} = 1/2$. Find a probability distribution π for which the chain $(X_k)_{k\in\mathbb{N}}$ is reversible.
Hint: The reversibility condition will yield a relation that can be used to compute π by induction. Remember to make use of the condition $\sum_i \pi_i = 1$.

(g) Confirm that the distribution of Question (f) is invariant (or stationary) by checking explicitly that the equality

$$\pi = \pi P$$

does hold.

(h) Show, by quoting the relevant theorem, that π is also the limiting distribution of the *modified* Ehrenfest chain $(X_k)_{k\geq 0}$.

(i) Show, by the result of Question (h), that

$$\lim_{n\to\infty} \mathbb{E}[X_n \mid X_0 = i] = \frac{M}{2},$$

for all $i = 0, 1, \ldots, M$.

(j) Show, using first step analysis and induction on $n \geq 0$, that we have

$$\mathbb{E}\left[X_n - \frac{M}{2}\,\middle|\, X_0 = i\right] = \left(i - \frac{M}{2}\right)\left(1 - \frac{1}{M}\right)^n, \qquad n \geq 0,$$

for all $i = 0, 1, \ldots, M$, and that this relation can be used to recover the result of Question (i).
Hint: Letting

$$h_n(i) = \mathbb{E}\left[X_n - \frac{M}{2}\,\middle|\, X_0 = i\right], \qquad n \geq 0,$$

in order to prove the formula by induction one has to

(i) show that the formula holds for $h_0(i)$ when $n = 0$;
(ii) show that assuming that the formula holds for $h_n(i)$, it then holds for $h_{n+1}(i)$.

It can help to start by proving the formula for $h_1(i)$ when $n = 1$ by first step analysis.

Chapter 8
Branching Processes

Abstract Branching processes are used as a tool for modeling in genetics, biomolecular reproduction, population growth, genealogy, disease spread, photomultiplier cascades, nuclear fission, earthquake triggering, queueing models, viral phenomena, social networks, neuroscience, etc. This chapter mainly deals with the computation of probabilities of extinction and explosion in finite time for branching processes.

8.1 Construction and Examples

Consider a time-dependent population made of a number X_n of individuals at generation $n \geq 0$. In the branching process model, each of these X_n individuals may have a random number of descendants born at time $n + 1$.

For each $k = 1, 2, \ldots, X_n$ we let Y_k denote the number of descendants of individual n^o k. That means, we have $X_0 = 1$, $X_1 = Y_1$, and at time $n + 1$, the new population size X_{n+1} will be given by

$$X_{n+1} = Y_1 + \cdots + Y_{X_n} = \sum_{k=1}^{X_n} Y_k, \tag{8.1.1}$$

where the $(Y_k)_{k\geq 1}$ form a sequence of independent, identically distributed, nonnegative integer valued random variables which are assumed to be almost surely finite, i.e.

$$\mathbb{P}(Y_k < \infty) = \sum_{n=0}^{\infty} \mathbb{P}(Y_k = n) = 1.$$

Note that in this model the X_n individuals of generation n "die" at time $n + 1$ as they are not considered in the sum (8.1.1). In order to keep them at the next generation we would have to modify (8.1.1) into

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4_8

$$X_{n+1} = X_n + Y_1 + \cdots + Y_{X_n},$$

however we will *not* adopt this convention, and we will rely on (8.1.1) instead.

As a consequence of (8.1.1), the branching process $(X_n)_{n\in\mathbb{N}}$ is a Markov process with state space $\mathbb{S} = \mathbb{N}$ and transition matrix given by

$$P = \left[\ P_{i,j}\ \right]_{i,j\in\mathbb{N}} = \begin{bmatrix} 1 & 0 & 0 & \cdots \\ \mathbb{P}(Y_1 = 0) & \mathbb{P}(Y_1 = 1) & \mathbb{P}(Y_1 = 2) & \cdots \\ P_{2,0} & P_{2,1} & P_{2,2} & \cdots \\ P_{3,0} & P_{3,1} & P_{3,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{8.1.2}$$

Note that state ⓪ is absorbing since by construction we always have

$$P_{0,0} = \mathbb{P}(X_{n+1} = 0 \mid X_n = 0) = 1, \qquad n \in \mathbb{N}.$$

Figure 8.1 represents an example of branching process with $X_0 = 1$ and $Y_1 = 6$, hence

$$X_1 = Y_{X_0} = Y_1 = 6,$$

then successively

$$(Y_k)_{k=1,2,\ldots,X_1} = (Y_1, Y_2, Y_3, Y_4, Y_5, Y_6) = (0, 4, 1, 2, 2, 3)$$

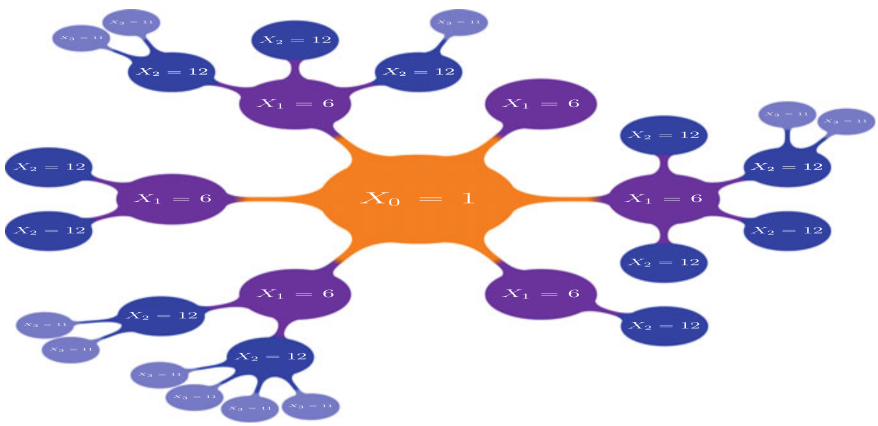


Fig. 8.1 Example of branching process

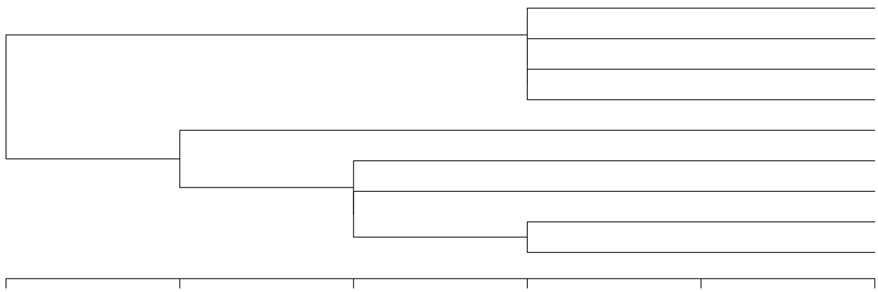


Fig. 8.2 Sample graph of a branching process

and

$$\begin{aligned} X_2 &= Y_1 + \cdots + Y_{X_1} \\ &= Y_1 + Y_2 + Y_3 + Y_4 + Y_5 + Y_6 \\ &= 0 + 4 + 1 + 2 + 2 + 3 \\ &= 12, \end{aligned}$$

then

$$\begin{aligned} (Y_k)_{k=1,2,\ldots,X_2} &= (Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8, Y_9, Y_{10}, Y_{11}, Y_{12}) \\ &= (0, 2, 0, 0, 0, 4, 2, 0, 0, 2, 0, 1), \end{aligned}$$

and

$$\begin{aligned} X_3 &= Y_1 + \cdots + Y_{X_2} \\ &= Y_1 + Y_2 + Y_3 + Y_4 + Y_5 + Y_6 + Y_7 + Y_8 + Y_9 + Y_{10} + Y_{11} + Y_{12} \\ &= 0 + 2 + 0 + 0 + 0 + 4 + 2 + 0 + 0 + 2 + 0 + 1 \\ &= 11. \end{aligned}$$

The next Fig. 8.2 presents another sample tree for the path of a branching process.

In Fig. 8.2 above the branching process starts from $X_0 = 2$, with $X_1 = 3$, $X_2 = 5$, $X_3 = 9$, $X_4 = 9$, $X_5 = 9$. However, in the sequel and except if otherwise specified, all branching processes will start from $X_0 = 1$.

See [SSB08] and [IM11] for results on the modeling of the offspring distribution of Y_1 based on social network and internet data and the use of power tail distributions. The use of power tail distributions leads to probability generating functions of polylogarithmic form.

8.2 Probability Generating Functions

Let now $G_1(s)$ denote the probability generating function of $X_1 = Y_1$, defined as

$$G_1(s) := \mathbb{E}[s^{X_1} \mid X_0 = 1] = \mathbb{E}[s^{Y_1}] = \sum_{k=0}^{\infty} s^k \mathbb{P}(Y_1 = k), \quad -1 \leq s \leq 1,$$

cf. (1.7.3), denote the probability generating function of the (almost surely finite) random variable $X_1 = Y_1$, with

$$\begin{cases} G_1(0) = \mathbb{P}(Y_1 = 0), \\ G_1(1) = \displaystyle\sum_{n=0}^{\infty} \mathbb{P}(Y_1 = n) = \mathbb{P}(Y_1 < \infty) = 1, \\ \mu := G_1'(1) = \displaystyle\sum_{k=0}^{\infty} k\mathbb{P}(Y_1 = k) = \mathbb{E}[X_1 \mid X_0 = 1] = \mathbb{E}[Y_1]. \quad (8.2.1a) \end{cases}$$

More generally, letting $G_n(s)$ denote the probability generating function of X_n, defined as

$$G_n(s) := \mathbb{E}[s^{X_n} \mid X_0 = 1] = \sum_{k=0}^{\infty} s^k \mathbb{P}(X_n = k \mid X_0 = 1), \quad -1 \leq s \leq 1,$$

$n \in \mathbb{N}$, we have

$$\begin{cases} G_0(s) = s, \qquad -1 \leq s \leq 1, \\ G_n(0) = \mathbb{P}(X_n = 0 \mid X_0 = 1), \qquad n \in \mathbb{N}, \quad (8.2.2a) \\ \mu_n := \mathbb{E}[X_n \mid X_0 = 1] = G_n'(1) = \displaystyle\sum_{k=0}^{\infty} k\mathbb{P}(X_n = k \mid X_0 = 1), \quad (8.2.2b) \end{cases}$$

cf. (1.7.5). When $X_0 = k$ we can view the branching tree as the union of k independent trees started from $X_0 = 1$ and we can write X_n as the sum of independent random variables

$$X_n = \sum_{l=1}^{k} X_n^{(l)}, \qquad n \in \mathbb{N},$$

where $X_n^{(l)}$ denotes the size of the tree $n^o l$ at time n, with $X_n^{(l)} = 1$, $l = 1, 2, \ldots, k$. In this case, we have

$$
\begin{aligned}
\mathbb{E}\left[s^{X_n} \mid X_0 = k\right] &= \mathbb{E}\left[s^{\sum_{l=1}^k X_n^{(k)}} \,\middle|\, X_0^{(1)} = 1, \ldots, X_0^{(k)} = 1\right] \\
&= \prod_{l=1}^{k} \mathbb{E}\left[s^{X_n^{(l)}} \mid X_0^{(l)} = 1\right] \\
&= \left(\mathbb{E}\left[s^{X_n} \mid X_0 = 1\right]\right)^k \\
&= (G_n(s))^k, \quad -1 \le s \le 1, \quad n \in \mathbb{N}.
\end{aligned}
$$

The next proposition provides an algorithm for the computation of the probability generating function G_n.

Proposition 8.1 *We have the recurrence relation*

$$
G_{n+1}(s) = G_n(G_1(s)) = G_1(G_n(s)), \quad -1 \le s \le 1, \qquad n \in \mathbb{N}. \tag{8.2.3}
$$

Proof By the identity (1.6.13) on random products we have

$$
\begin{aligned}
G_{n+1}(s) &= \mathbb{E}[s^{X_{n+1}} \mid X_0 = 1] \\
&= \mathbb{E}[s^{Y_1 + \cdots + Y_{X_n}} \mid X_0 = 1] \\
&= \mathbb{E}\left[\prod_{l=1}^{X_n} s^{Y_l} \,\middle|\, X_0 = 1\right] \\
&= \sum_{k=0}^{\infty} \mathbb{E}\left[\prod_{l=1}^{X_n} s^{Y_l} \,\middle|\, X_n = k\right] \mathbb{P}(X_n = k \mid X_0 = 1) \\
&= \sum_{k=0}^{\infty} \mathbb{E}\left[\prod_{l=1}^{k} s^{Y_l} \,\middle|\, X_n = k\right] \mathbb{P}(X_n = k \mid X_0 = 1) \\
&= \sum_{k=0}^{\infty} \mathbb{E}\left[\prod_{l=1}^{k} s^{Y_l}\right] \mathbb{P}(X_n = k \mid X_0 = 1) \\
&= \sum_{k=0}^{\infty} \left(\prod_{l=1}^{k} \mathbb{E}\left[s^{Y_l}\right]\right) \mathbb{P}(X_n = k \mid X_0 = 1) \\
&= \sum_{k=0}^{\infty} (\mathbb{E}[s^{Y_1}])^k \mathbb{P}(X_n = k \mid X_0 = 1) \\
&= G_n(\mathbb{E}[s^{Y_1}]) \\
&= G_n(G_1(s)), \qquad -1 \le s \le 1.
\end{aligned}
$$

□

Instead of (8.2.3) we may also write

$$
G_n(s) = G_1(G_1(\cdots(G_1(s), \cdots)), \qquad -1 \le s \le 1, \tag{8.2.4}
$$

and

$$G_n(s) = G_1(G_{n-1}(s)) = G_{n-1}(G_1(s)), \qquad -1 \le s \le 1.$$

Mean Population Size

In case the random variable Y_k is equal to a deterministic constant $\mu \in \mathbb{N}$, the population size at generation $n \ge 0$ will clearly be equal to μ^n. The next proposition shows that for branching processes, this property admits a natural extension to the random case.

Proposition 8.2 *The mean population size μ_n at generation $n \ge 0$ is given by*

$$\mu_n = \mathbb{E}[X_n \mid X_0 = 1] = (\mathbb{E}[X_1 \mid X_0 = 1])^n = \mu^n, \qquad n \ge 1, \tag{8.2.5}$$

where $\mu = \mathbb{E}[Y_1]$ is given by (8.2.1a).

Proof By (8.2.4), (8.2.2b) and the chain rule of derivation we have

$$\begin{aligned}
\mu_n &= G_n'(1) \\
&= \frac{d}{ds} G_1(G_{n-1}(s))_{|s=1} \\
&= G_{n-1}'(1) G_1'(G_{n-1}(1)) \\
&= G_{n-1}'(1) G_1'(1) \\
&= \mu \times \mu_{n-1},
\end{aligned}$$

hence $\mu_1 = \mu$, $\mu_2 = \mu \times \mu_1 = \mu^2$, $\mu_3 = \mu \times \mu_2 = \mu^3$, and by induction on $n \ge 1$ we obtain (8.2.5). □

Similarly we find

$$\mathbb{E}[X_n \mid X_0 = k] = k\, \mathbb{E}[X_n \mid X_0 = 1] = k\mu^n, \qquad n \ge 1,$$

hence starting from $X_0 = k \ge 1$, the average of X_n goes to infinity when $\mu > 1$. On the other hand, μ_n converges to 0 when $\mu < 1$.

Examples

(i) Supercritical case. When $\mu > 1$ the average population size $\mu_n = \mu^n$ grows to infinity as n tends to infinity, and we say that the branching process $(X_n)_{n\in\mathbb{N}}$ is *supercritical*.
This condition holds in particular when $\mathbb{P}(Y_1 \ge 1) = 1$ and Y_1 is not almost surely equal to 1, i.e. $\mathbb{P}(Y_1 = 1) < 1$. Indeed, under those conditions we have $\mathbb{P}(Y_1 \ge 2) > 0$ and

$$
\begin{aligned}
\mu = \mathbb{E}[Y_1] &= \sum_{n=1}^{\infty} n\mathbb{P}(Y_1 = n) \\
&\geq \mathbb{P}(Y_1 = 1) + 2\sum_{n=2}^{\infty} \mathbb{P}(Y_1 = n) \\
&= \mathbb{P}(Y_1 = 1) + \mathbb{P}(Y_1 \geq 2) \\
&> \mathbb{P}(Y_1 \geq 1) \\
&= 1,
\end{aligned}
$$

hence $\mu > 1$.

(ii) Critical case. When $\mu = 1$ we have $\mu_n = (\mu)^n = 1$ for all $n \in \mathbb{N}$, and we say that the branching process $(X_n)_{n\in\mathbb{N}}$ is *critical*.

(iii) Subcritical case. In case $\mu < 1$, the average population size $\mu_n = \mu^n$ tends to 0 as n tends to infinity and we say that the branching process $(X_n)_{n\in\mathbb{N}}$ is *subcritical*. In this case we necessarily have $\mathbb{P}(Y_1 = 0) > 0$ as

$$
\begin{aligned}
1 > \mu = \mathbb{E}[Y_1] &= \sum_{n=1}^{\infty} n\mathbb{P}(Y_1 = n) \\
&\geq \sum_{n=1}^{\infty} \mathbb{P}(Y_1 = n) \\
&= \mathbb{P}(Y_1 \geq 1) \\
&= 1 - \mathbb{P}(Y_0 = 0),
\end{aligned}
$$

although the converse is not true in general.

The variance $\sigma_n^2 = \mathrm{Var}[X_n \mid X_0 = 1]$ of X_n given that $X_0 = 1$ can be shown in a similar way to satisfy the recurrence relation

$$
\sigma_{n+1}^2 = \sigma^2\mu^n + \mu^2\sigma_n^2,
$$

cf. also Relation (1.7.6), where $\sigma^2 = \mathrm{Var}[Y_1]$, which shows by induction that

$$
\sigma_n^2 = \mathrm{Var}[X_n \mid X_0 = 1] = \begin{cases} n\sigma^2, & \mu = 1, \\ \sigma^2\mu^{n-1}\dfrac{1-\mu^n}{1-\mu} = \sigma^2\displaystyle\sum_{k=0}^{n-1}\mu^{n+k-1}, & \mu \neq 1, \end{cases}
$$

$n \geq 1$ cf. e.g. pages 180–181 of [KT81], and Exercise 8.3-(a) for an application. We also have

$$
\mathrm{Var}[X_n \mid X_0 = k] = k\mathrm{Var}[X_n \mid X_0 = 1] = k\sigma_n^2, \qquad k, n \in \mathbb{N},
$$

due to Relation (1.6.12) for the variance of a sum of independent random variables.

8.3 Extinction Probabilities

Here we are interested in the time to extinction[1]

$$T_0 := \inf\{n \geq 0 \ : \ X_n = 0\},$$

and in the extinction probability

$$\alpha_k := \mathbb{P}(T_0 < \infty \mid X_0 = k)$$

within a finite time, after starting from $X_0 = k$. Note that the word "extinction" can have negative as well as positive meaning, for example when the branching process is used to model the spread of an infection.

Proposition 8.3 *The probability distribution of T_0 can be expressed using the probability generating function G_n as*

$$\mathbb{P}(T_0 = n \mid X_0 = 1) = G_n(0) - G_{n-1}(0) = G_1(G_{n-1}(0)) - G_{n-1}(0), \quad n \geq 1,$$

with $\mathbb{P}(T_0 = 0 \mid X_0 = 1) = 0$.

Proof By the relation $\{X_{n-1} = 0\} \subset \{X_n = 0\}$, we have

$$\{T_0 = n\} = \{X_n = 0\} \cap \{X_{n-1} \geq 1\} = \{X_n = 0\} \setminus \{X_{n-1} = 0\}$$

and

$$\begin{aligned}
\mathbb{P}(T_0 = n \mid X_0 = 1) &= \mathbb{P}(\{X_n = 0\} \cap \{X_{n-1} \geq 1\} \mid X_0 = 1) \\
&= \mathbb{P}(\{X_n = 0\} \setminus \{X_{n-1} = 0\} \mid X_0 = 1) \\
&= \mathbb{P}(\{X_n = 0\}) - \mathbb{P}(\{X_{n-1} = 0\} \mid X_0 = 1) \\
&= G_n(0) - G_{n-1}(0) \\
&= G_1(G_{n-1}(0)) - G_{n-1}(0), \quad n \geq 1,
\end{aligned}$$

where we applied Proposition 8.1. □

First, we note that by the independence assumption, starting from $X_0 = k \geq 2$ independent individuals, we have

$$\alpha_k = \mathbb{P}(T_0 < \infty \mid X_0 = k) = (\mathbb{P}(T_0 < \infty \mid X_0 = 1))^k = (\alpha_1)^k, \quad k \geq 1. \quad (8.3.1)$$

Indeed, given k individuals at generation 0, each of them will start independently a new branch of offsprings, and in order to have extinction of the whole population, all of k branches should become extinct. Since the k branches behave independently, α_k

[1] We normally start from $X_0 \geq 1$.

is the product of the extinction probabilities for each branch, which yields $\alpha_k = (\alpha_1)^k$ since these extinction probabilities are all equal to α_1 and there are k of them.

The next proposition is a consequence of Lemma 5.1.

Proposition 8.4 *We have*

$$\alpha_1 = \lim_{n\to\infty} G_n(0).$$

Proof Since state ⓪ is absorbing, by Lemma 5.1 with $j = 0$ and $i = 1$ we find

$$\alpha_1 = \mathbb{P}(T_0 < \infty \mid X_0 = 1) = \lim_{n\to\infty} \mathbb{P}(X_n = 0 \mid X_0 = 1) = \lim_{n\to\infty} G_n(0).$$

□

The next proposition shows that the extinction probability α_1 can be computed as the solution of an equation.

Proposition 8.5 *The extinction probability*

$$\alpha_1 := \mathbb{P}(T_0 < \infty \mid X_0 = 1)$$

is a solution of the equation

$$\alpha = G_1(\alpha). \tag{8.3.2}$$

Proof By first step analysis we have

$$\begin{aligned}
\alpha_1 &= \mathbb{P}(T_0 < \infty \mid X_0 = 1) \\
&= \mathbb{P}(X_1 = 0 \mid X_0 = 1) + \sum_{k=1}^{\infty} \mathbb{P}(T_0 < \infty \mid X_1 = k)\mathbb{P}(X_1 = k \mid X_0 = 1) \\
&= \mathbb{P}(Y_1 = 0) + \sum_{k=1}^{\infty} \mathbb{P}(T_0 < \infty \mid X_0 = k)\mathbb{P}(Y_1 = k) \\
&= \sum_{k=0}^{\infty} (\alpha_1)^k \mathbb{P}(Y_1 = k) \\
&= G_1(\alpha_1),
\end{aligned}$$

hence the extinction probability α_1 solves (8.3.2). □

Note that from the above proof we find

$$\alpha_1 \geq \mathbb{P}(X_1 = 0 \mid X_0 = 1) = \mathbb{P}(Y_1 = 0), \tag{8.3.3}$$

which shows that the extinction probability is non-zero whenever $\mathbb{P}(Y_1 = 0) > 0$. On the other hand, any solution α of (8.3.2) also satisfies

$$\alpha = G_1(G_1(\alpha)), \quad \alpha = G_1(G_1(G_1(\alpha))),$$

and more generally

$$\alpha = G_n(\alpha), \qquad n \geq 1, \tag{8.3.4}$$

by Proposition 8.1. On the other hand the solution of (8.3.2) may not be unique, for example $\alpha = 1$ is always solution of (8.3.2) since $G_1(1) = 1$, and it may not be equal to the extinction probability. The next proposition clarifies this point.

Proposition 8.6 *The extinction probability*

$$\alpha_1 := \mathbb{P}(T_0 < \infty \mid X_0 = 1)$$

is the smallest solution of the equation $\alpha = G_1(\alpha)$.

Proof By Proposition 8.4 we have $\alpha_1 = \lim_{n\to\infty} G_n(0)$. Next, we note that the function $s \longmapsto G_1(s)$ is increasing because

$$G_1'(s) = \mathbb{E}[Y_1 s^{Y_1 - 1}] > 0, \qquad s \in [0, 1).$$

Hence $s \longmapsto G_n(s)$ is also increasing by Proposition 8.1, and for any solution $\alpha \geq 0$ of (8.3.2) we have, by (8.3.4),

$$0 \leq G_n(0) \leq G_n(\alpha) = \alpha, \qquad n \geq 1,$$

and taking the limit in this inequality as n goes to infinity we get

$$0 \leq \alpha_1 = \lim_{n\to\infty} G_n(0) \leq \alpha,$$

by (5.1.5) and Proposition 8.4, hence the extinction probability α_1 is always smaller than any solution α of (8.3.2). This fact can also be recovered from Proposition 8.4 and

$$\begin{aligned}
\alpha &= \lim_{n\to\infty} G_n(\alpha) \\
&= \lim_{n\to\infty} \left(G_n(0) + \sum_{k=1}^{\infty} \alpha^k \mathbb{P}(X_n = k \mid X_0 = 1) \right) \\
&\geq \lim_{n\to\infty} G_n(0) \\
&= \alpha_1.
\end{aligned}$$

Therefore α_1 is the smallest solution of (8.3.2). □

Since $G_1(0) = \mathbb{P}(Y_1 = 0)$ we have

$$\mathbb{P}(Y_1 = 0) = G_1(0) \leq G_1(\alpha_1) = \alpha_1,$$

which recovers (8.3.3).

On the other hand, if $\mathbb{P}(Y_1 \geq 1) = 1$ then we have $G_1(0) = 0$, which implies $\alpha_1 = 0$ by Proposition 8.6.

Note that from Lemma 5.1, Proposition 8.4, and (8.3.1), the transition matrix (8.1.2) satisfies

$$\lim_{n\to\infty} \big([P^n]_{i,0}\big)_{i\in\mathbb{N}} = \begin{bmatrix} 1 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 \\ \alpha_1 \\ (\alpha_1)^2 \\ (\alpha_1)^3 \\ \vdots \end{bmatrix}.$$

Examples

(i) Assume that Y_1 has a Bernoulli distribution with parameter $p \in (0, 1)$, i.e.

$$\mathbb{P}(Y_1 = 1) = p, \qquad \mathbb{P}(Y_1 = 0) = 1 - p.$$

Compute the extinction probability of the associated branching process.

In this case the branching process is actually a two-state Markov chain with transition matrix

$$P = \begin{bmatrix} 1 & 0 \\ 1-p & p \end{bmatrix},$$

and we have

$$G_n(0) = \mathbb{P}(X_n = 0 \mid X_0 = 1) = (1-p)\sum_{k=0}^{n-1} p^k = 1 - p^n, \tag{8.3.5}$$

where we used the geometric series (A.2), hence as in (5.15) the extinction probability α_1 is given by

$$\begin{aligned} \alpha_1 &= \mathbb{P}(T_0 < \infty \mid X_0 = 1) = \mathbb{P}\bigg(\bigcup_{n\geq 1}\{X_n = 0\}\,\bigg|\, X_0 = 1\bigg) \\ &= \lim_{n\to\infty} \mathbb{P}(X_n = 0 \mid X_0 = 1) \\ &= \lim_{n\to\infty} G_n(0) = 1, \end{aligned}$$

provided that $p = \mathbb{E}[Y_1] < 1$, otherwise we have $\alpha_1 = 0$ when $p = 1$. The value of α_1 can be recovered using the generating function

$$G_1(s) = \mathbb{E}[s^{Y_1}] = \sum_{k=0}^{\infty} s^k \mathbb{P}(Y_1 = k) = 1 - p + ps, \tag{8.3.6}$$

for which the unique solution of $G_1(\alpha) = \alpha$ is the extinction probability $\alpha_1 = 1$, as shown in the next Fig. 8.3.

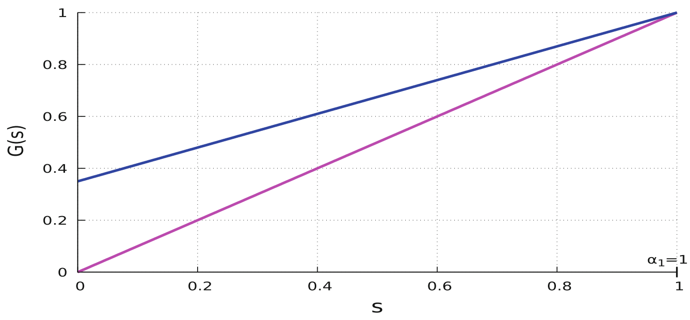


Fig. 8.3 Generating function of Y_1 with $p = 0.65$

From (8.2.4) and (8.3.6) we can also show by induction on $n \geq 1$ as in Exercise 8.2 that

$$G_n(s) = p^n s + (1-p)\sum_{k=0}^{n-1} p^k = 1 - p^n + p^n s,$$

which recovers (8.3.5) from (1.7.5) or (8.2.2a) as

$$\mathbb{P}(X_n = 0 \mid X_0 = 1) = G_n(0) = 1 - p^n.$$

We also have $\mathbb{E}[X_n] = p^n$, $n \geq 0$.

(ii) Same question as in (i) above for

$$\mathbb{P}(Y_1 = 2) = p, \qquad \mathbb{P}(Y_1 = 0) = q = 1 - p.$$

Here, we will directly use the probability generating function

$$\begin{aligned} G_1(s) = \mathbb{E}[s^{Y_1}] &= \sum_{k=0}^{\infty} s^k \mathbb{P}(Y_1 = k) \\ &= s^0 \mathbb{P}(Y_1 = 0) + s^2 \mathbb{P}(Y_1 = 2) = 1 - p + ps^2. \end{aligned}$$

We check that the solutions of

$$G_1(\alpha) = 1 - p + p\alpha^2 = \alpha,$$

i.e.[2]

$$p\alpha^2 - \alpha + q = p(\alpha - 1)(\alpha - q/p) = 0, \tag{8.3.7}$$

[2]Remark that (8.3.7) is identical to the characteristic Eq. (2.2.15).

with $q = 1 - p$, are given by

$$\left\{ \frac{1+\sqrt{1-4pq}}{2p}, \frac{1-\sqrt{1-4pq}}{2p} \right\} = \left\{ 1, \frac{q}{p} \right\}, \qquad p \in (0, 1]. \tag{8.3.8}$$

Hence the extinction probability is $\alpha_1 = 1$ if $q \geq p$, and it is equal to $\alpha_1 = q/p < 1$ if $q < p$, or equivalently if $\mathbb{E}[Y_1] > 1$, due to the relation $\mathbb{E}[Y_1] = 2p$.

(iii) Assume that Y_1 has the geometric distribution with parameter $p \in (0, 1)$, i.e.

$$\mathbb{P}(Y_1 = n) = (1-p)p^n, \qquad n \geq 0,$$

with $\mu = \mathbb{E}[Y_1] = p/q$. We have

$$G_1(s) = \mathbb{E}[s^{Y_1}] = \sum_{n=0}^{\infty} s^n \mathbb{P}(Y_1 = n) = (1-p) \sum_{n=0}^{\infty} p^n s^n = \frac{1-p}{1-ps}. \tag{8.3.9}$$

The equation $G_1(\alpha) = \alpha$ reads

$$\frac{1-p}{1-p\alpha} = \alpha,$$

i.e.

$$p\alpha^2 - \alpha + q = p(\alpha - 1)(\alpha - q/p) = 0,$$

which is identical to (2.2.15) and (8.3.7) with $q = 1 - p$, and has for solutions (8.3.8). Hence the finite time extinction probability is

$$\begin{aligned} \alpha_1 &= \mathbb{P}(T_0 < \infty \mid X_0 = 1) \\ &= \min\left(1, \frac{q}{p}\right) = \begin{cases} \dfrac{q}{p}, & p \geq 1/2, \text{ (super)critical case,} \\ \\ 1, & p \leq 1/2, \text{ (sub)critical case.} \end{cases} \end{aligned}$$

Note that we have $\alpha_1 < 1$ if and only if $\mathbb{E}[Y_1] > 1$, due to the equality $\mathbb{E}[Y_1] = p/q$. As can be seen from Figs. 8.4 and 8.5, the extinction probability α_1 is equal to 1 when $p \leq 1/2$, meaning that extinction within a finite time is certain in that case. Note that we also find

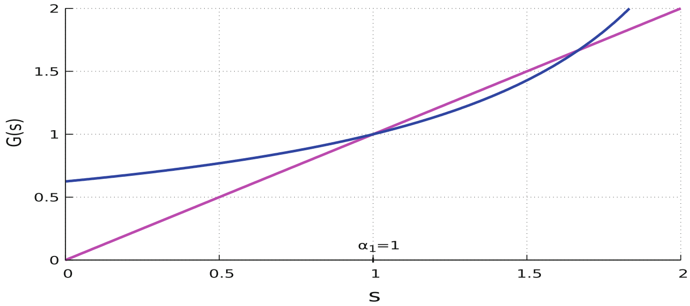


Fig. 8.4 Generating function of Y_1 with $p = 3/8 < 1/2$ and $\alpha_1 = 1$

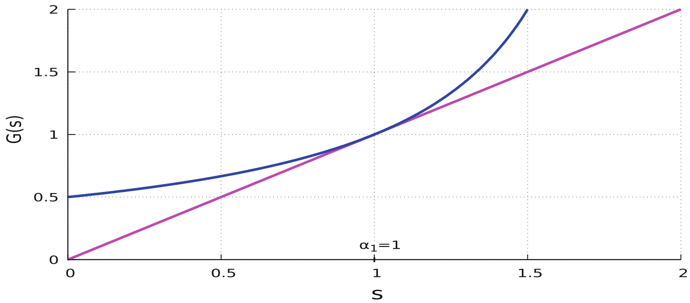


Fig. 8.5 Generating function of Y_1 with $p = 1/2$ and $\alpha_1 = 1$

$$\mathbb{P}(T_0 < \infty \mid X_0 = k) = \min\left(1, \left(\frac{q}{p}\right)^k\right)$$
$$= \begin{cases} \left(\dfrac{q}{p}\right)^k, & p \geq 1/2, \text{ (super)critical case,} \\ 1, & p \leq 1/2, \text{ (sub)critical case.} \end{cases}$$

which incidentally coincides with the finite time hitting probability found in (3.4.16) for the simple random walk started from $k \geq 1$.
Next in Fig. 8.5 is a graph of the generating function $s \longmapsto G_1(s)$ for $p = 1/2$. The graph of generating function in Fig. 8.6 corresponds to $p = 3/4$.
We also have $\mu_n = (\mathbb{E}[Y_1])^n = (p/q)^n$, $n \geq 1$.

(iv) Assume now that Y_1 is the sum of two independent geometric variables with parameter $1/2$, i.e. it has the *negative binomial distribution*

$$\mathbb{P}(Y_1 = n) = \binom{n+r-1}{r-1} q^r p^n = (n+1) q^r p^n = (n+1) q^2 p^n, \qquad n \geq 0,$$

with $r = 2$, cf. (1.5.12).

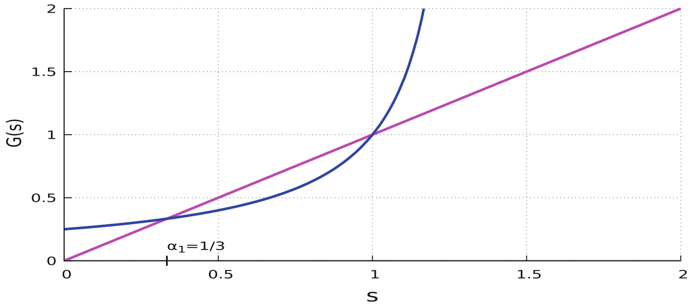


Fig. 8.6 Generating function of Y_1 with $p = 3/4 > 1/2$ and $\alpha_1 = q/p = 1/3$

In this case we have[3]

$$G_1(s) = \mathbb{E}[s^{Y_1}] = \sum_{n=0}^{\infty} s^n \mathbb{P}(Y_1 = n)$$
$$= q^2 \sum_{n=0}^{\infty} (n+1) p^n s^n = \left(\frac{1-p}{1-ps}\right)^2, \qquad -1 \le s \le 1.$$

When $p = 1/2$ we check that $G_1(\alpha) = \alpha$ reads

$$s^3 - 4s^2 + 4s - 1 = 0,$$

which is an equation of degree 3 in the unknown s. Now, since $\alpha = 1$ is solution of this equation we can factorise it as follows:

$$(s-1)(s^2 - 3s + 1) = 0,$$

and we check that the smallest nonnegative solution of this equation is given by

$$\alpha_1 = \frac{1}{2}(3 - \sqrt{5}) \simeq 0.382$$

which is the extinction probability, as illustrated in the next Fig. 8.7. Here we have $\mathbb{E}[Y_1] = 2$.

The next graph illustrates the extinction of a branching process in finite time when Y_1 has the geometric distribution with $p = 1/2$, in which case there is extinction within finite time with probability 1 (Fig. 8.8).

In Table 8.1 we summarize some questions and their associated solution methods introduced in this chapter and the previous ones.

---

[3]Here, Y_1 is the sum of two independent geometric random variables, and G_1 is the square of the generating function (8.3.9) of the geometric distribution.

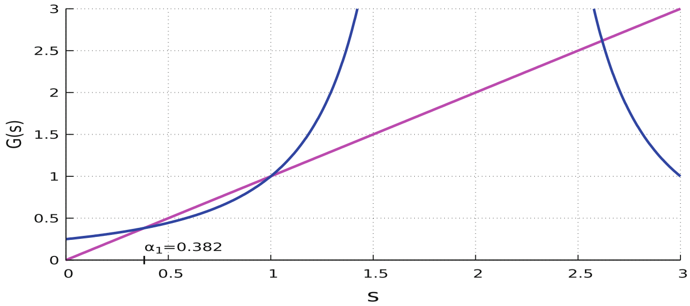


Fig. 8.7 Probability generating function of Y_1

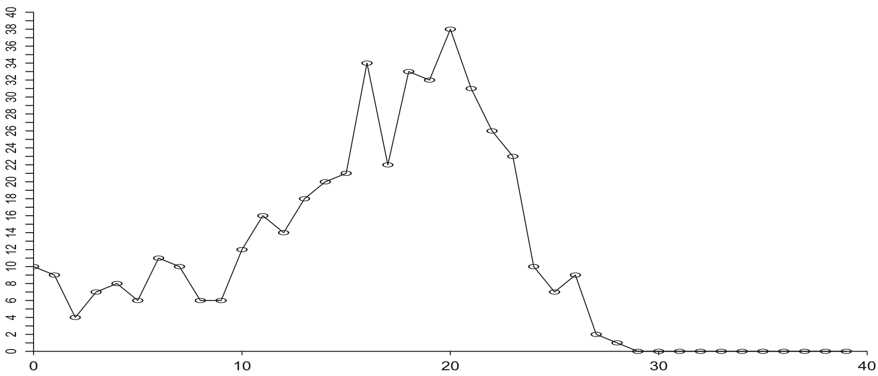


Fig. 8.8 Sample path of a branching process $(X_n)_{n\geq 0}$

Table 8.1 Summary of computing methods

How to compute	*Method*
The expected value $\mathbb{E}[X]$	Sum the values of X weighted by their probabilities
Uses of $G_X(s)$	$G_X(0) = \mathbb{P}(X = 0)$
	$G_X(1) = \mathbb{P}(X < \infty)$
	$G'_X(1) = \mathbb{E}[X]$
The hitting probabilities $g(k)$	Solve[a] $g = Pg$ for $g(k)$
The mean hitting times $h(k)$	Solve[a] $h = \mathbb{1} + Ph$ for $h(k)$
The stationary distribution π	Solve[b] $\pi = \pi P$ for π
The extinction probability α_1	Solve $G_1(\alpha) = \alpha$ for α and choose the smallest solution
$\lim_{n\to\infty}\left(\begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}\right)^n$	$\begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{bmatrix}$

[a]Be sure to write only the relevant rows of the system under the appropriate boundary conditions
[b] Remember that the values of $\pi(k)$have to add up to 1

Exercises

Exercise 8.1 A parent particle can be divided into 0, 1 or 2 particles with probabilities $1/5$, $3/5$, and $1/5$, respectively. It disappears after splitting. Starting with one particle, the ancestor, let us denote by X_n the size of the corresponding branching process at the nth generation.

(a) Find $P(X_2 > 0)$.
(b) Find $P(X_2 = 1)$.
(c) Find the probability that $X_1 = 2$ given that $X_2 = 1$.

Exercise 8.2 Each individual in a population has a random number Y of offsprings, with

$$\mathbb{P}(Y = 0) = 1/2, \qquad \mathbb{P}(Y = 1) = 1/2.$$

Let X_n denote the size of the population at time $n \in \mathbb{N}$, with $X_0 = 1$.

(a) Compute the generating function $G_1(s) = \mathbb{E}[s^Y]$ of Y for $s \in \mathbb{R}_+$.
(b) Let $G_n(s) = \mathbb{E}\left[s^{X_n}\right]$ denote the generating function of X_n. Show that

$$G_n(s) = 1 - \frac{1}{2^n} + \frac{s}{2^n}, \qquad s \in \mathbb{R}. \tag{8.3.10}$$

(c) Compute the probability $\mathbb{P}(X_n = 0 \mid X_0 = 1)$ that the population is extinct at time n.
(d) Compute the average size $\mathbb{E}[X_n \mid X_0 = 1]$ of the population at step n.
(e) Compute the extinction probability of the population starting from one individual at time 0.

Exercise 8.3 Each individual in a population has a random number ξ of offsprings, with distribution

$$\mathbb{P}(\xi = 0) = 0.2, \qquad \mathbb{P}(\xi = 1) = 0.5, \qquad \mathbb{P}(\xi = 2) = 0.3.$$

Let X_n denote the number of individuals in the population at the nth generation, with $X_0 = 1$.

(a) Compute the mean and variance of X_2.
(b) Give the probability distribution of the random variable X_2.
(c) Compute the probability that the population is extinct by the fourth generation.
(d) Compute the expected number of offsprings at the tenth generation.
(e) What is the probability of extinction of this population?

Exercise 8.4 Each individual in a population has a random number Y of offsprings, with

$$\mathbb{P}(Y = 0) = c, \qquad \mathbb{P}(Y = 1) = b, \qquad \mathbb{P}(Y = 2) = a,$$

where $a + b + c = 1$.

(a) Compute the generating function $G_1(s)$ of Y for $s \in [-1, 1]$.
(b) Compute the probability that the population is extinct at time 2, starting from 1 individual at time 0.
(c) Compute the probability that the population is extinct at time 2, starting from 2 individuals at time 0.
(d) Show that when $0 < c \leq a$ the probability of eventual extinction of the population, starting from 2 individuals at time 0, is $(c/a)^2$.
(e) What is this probability equal to when $0 < a < c$?

Exercise 8.5 Consider a branching process $(Z_n)_{n\geq 0}$ in which the offspring distribution at each generation is binomial with parameter $(2, p)$, i.e.

$$\mathbb{P}(Y = 0) = q^2, \quad \mathbb{P}(Y = 1) = 2pq, \quad \mathbb{P}(Y = 2) = p^2,$$

with $q := 1 - p$.

(a) Compute the probability generating function G_Y of Y.
(b) Compute the extinction probability of this process, starting from $Z_0 = 1$.
(c) Compute the probability that the population becomes extinct for the first time in the second generation ($n = 2$), starting from $Z_0 = 1$.
(d) Suppose that the initial population size Z_0 is a Poisson random variable with parameter $\lambda > 0$. Compute the extinction probability in this case.

Exercise 8.6 A cell culture is started with one red cell at time 0. After one minute the red cell dies and two new cells are born according to the following probability distribution:

Color configuration	Probability
2 red cells	1/4
1 red cell + 1 white cell	2/3
2 white cells	1/12

The above procedure is repeated minute after minute for any red cell present in the culture. On the other hand, the white cells can only live for one minute, and disappear after that without reproducing. We assume that the cells behave independently.

(a) What is the probability that no white cells have been generated until time n included?
(b) Compute the extinction probability of the whole cell culture.
(c) Same questions as above for the following probability distribution:

Color configuration	Probability
2 red cells	1/3
1 red cell + 1 white cell	1/2
2 white cells	1/6

Exercise 8.7 Using first step analysis, show that if $(X_n)_{n\geq 0}$ is a subcritical branching process, i.e. $\mu = \mathbb{E}[Y_1] < 1$, the time to extinction $T_0 := \inf\{n \geq 0 \ : \ X_n = 0\}$ satisfies $\mathbb{E}[T_0 \mid X_0 = 1] < \infty$.

Exercise 8.8 Consider a branching process $(X_n)_{n\geq 0}$ started at $X_0 = 1$, in which the numbers Y_k of descendants of individual $n^o k$ form an *i.i.d.* sequence with the negative binomial distribution

$$\mathbb{P}(Y_k = n) = (n+1)q^2 p^n, \qquad n \geq 0, \quad k \geq 1,$$

where $0 < q = 1 - p < 1$.

(a) Compute the probability generating function

$$G_1(s) := \mathbb{E}\left[s^{Y_k}\right] = \sum_{n=0}^{\infty} s^n \mathbb{P}(Y_1 = n)$$

of Y_k, $k \geq 1$.

b) Compute the extinction probability $\alpha_1 := \mathbb{P}(T_0 < \infty \mid X_0 = 1)$ of the branching process $(X_n)_{n\geq 0}$ in finite time.

Exercise 8.9 Families in a given society have children until the birth of the first girl, after which the family stops having children. Let X denote the number of male children of a given husband.

(a) Assuming that girls and boys are equally likely to be born, compute the probability distribution of X.
(b) Compute the probability generating function $G_X(s)$ of X.
(c) What is the probability that a given man has no male descendant (patrilineality) by the time of the third generation?
(d) Suppose now that one fourth of the married couples have no children at all while the others continue to have children until the first girl, and then cease childbearing. What is the probability that the wife's female line of descent (matrilineality) will cease to exist by the third generation?

Exercise 8.10 Consider a branching process $(Z_k)_{k\in\mathbb{N}}$ with $Z_0 = 1$ and offspring distribution given by

$$\mathbb{P}(Z_1 = 0) = \frac{1-p-q}{1-p} \quad \text{and} \quad \mathbb{P}(Z_1 = k) = qp^{k-1}, \qquad k = 1, 2, 3, \ldots,$$

where $0 \leq p < 1$ and $0 \leq q \leq 1 - p$.

(a) Find the probability generating function of Z_1.
(b) Compute $\mathbb{E}[Z_1]$.
(c) Find the value of q for which $\mathbb{E}[Z_1] = 1$, known as the *critical value*.
(d) Using the critical value of q, show by induction that determine the probability generating function of Z_k is given by

$$G_{Z_k}(s) = \frac{kp - (kp + p - 1)s}{1 - p + kp - kps}, \qquad -1 < s < 1,$$

for all $k \geq 1$.

Problem 8.11 Consider a branching process with *i.i.d.* offspring sequence $(Y_k)_{k\geq 1}$. The number of individuals in the population at generation $n + 1$ is given by the relation $X_{n+1} = Y_1 + \cdots + Y_{X_n}$, with $X_0 = 1$.

(a) Let

$$Z_n = \sum_{k=1}^{n} X_k,$$

denote the total number of individuals generated from time 1 to n. Compute $\mathbb{E}[Z_n]$ as a function of $\mu = \mathbb{E}[Y_1]$.

(b) Let $Z = \sum_{k=1}^{\infty} X_k$. denote the total number of individuals generated from time 1 to infinity. Compute $\mathbb{E}[Z]$ and show that it is finite when $\mu < 1$.

In the sequel we work under the condition $\mu < 1$.

(c) Let

$$H(s) = \mathbb{E}\left[s^Z\right], \qquad -1 \leq s \leq 1,$$

denote the generating function of Z.

Show, by first step analysis, that the relation

$$H(s) = G_1(sH(s)), \qquad 0 \leq s \leq 1,$$

holds, where $G_1(x)$ is the probability generating function of Y_1.

(d) In the sequel we assume that Y_1 has the geometric distribution $\mathbb{P}(Y_1 = k) = qp^k$, $k \in \mathbb{N}$, with $p \in (0, 1)$ and $q = 1 - p$. Compute $H(s)$ for $s \in [0, 1]$.
(e) Using the expression of the generating function $H(s)$ computed in Question (d), check that we have $H(0) = \lim_{s \searrow 0} H(s)$, where $H(0) = \mathbb{P}(Z = 0) = \mathbb{P}(Y_1 = 0) = G_1(0)$.
(f) Using the generating function $H(s)$ computed in Question (d), recover the value of $\mathbb{E}[Z]$ found in Question (b).

(g) Assume that each of the Z individuals earns an income U_k, $k = 1, 2, \ldots, Z$, where $(U_k)_{k\geq 1}$ is an *i.i.d.* sequence of random variables with finite expectation $\mathbb{E}[U]$ and distribution function $F(x) = \mathbb{P}(U \leq x)$.

Compute the expected value of the sum of gains of all the individuals in the population.

(h Compute the probability that none of the individuals earns an income higher than $x > 0$.

(i) Evaluate the results of Questions (g) and (h) when U_k has the exponential distribution with $F(x) = 1 - \mathrm{e}^{-x}$, $x \in \mathbb{R}_+$.

Chapter 9
Continuous-Time Markov Chains

In this chapter we start the study of *continuous-time* stochastic processes, which are families $(X_t)_{t\in\mathbb{R}_+}$ of random variables indexed by $\mathbb{R}_+$. Our aim is to make the transition from discrete to continuous-time Markov chains, the main difference between the two settings being the replacement of the transition matrix with the continuous-time *infinitesimal generator* of the process. We will start with the two fundamental examples of the Poisson and birth and death processes, followed by the construction of continuous-time Markov chains and their generators in more generality. From the point of view of simulations, the use of continuous-time Markov chains does not bring any special difficulty as any continuous-time simulation is actually based on discrete-time samples. From a theoretical point of view, however, the rigorous treatment of the continuous-time Markov property is much more demanding than its discrete-time counterpart, notably due to the use of the strong Markov property. Here we focus on the understanding of the continuous-time case by simple calculations, and we will refer to the literature for the use of the strong Markov property.

9.1 The Poisson Process

The *standard Poisson process* $(N_t)_{t\in\mathbb{R}_+}$ is a continuous-time *counting* process, i.e. $(N_t)_{t\in\mathbb{R}_+}$ has jumps of size $+1$ only, and its paths are constant (and right-continuous) in between jumps. The next Fig. 9.1 represents a sample path of a Poisson process.

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4_9

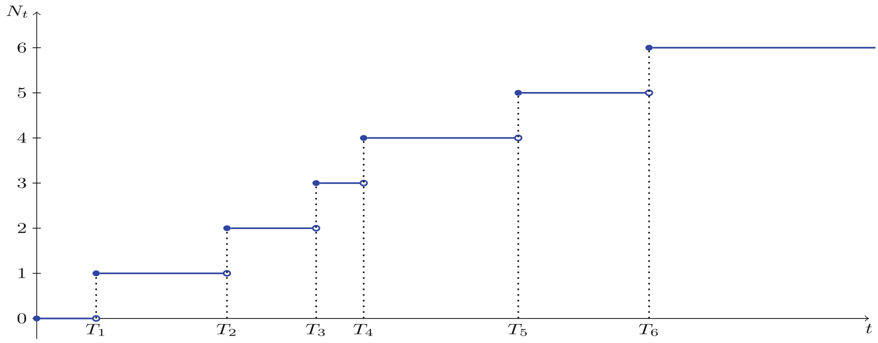


Fig. 9.1 Sample path of a Poisson process $(N_t)_{t\in\mathbb{R}_+}$

We denote by $(T_k)_{k\geq 1}$ the increasing sequence of jump times of $(N_t)_{t\in\mathbb{R}_+}$, which can be defined from the (right-continuous) Poisson process path $(N_t)_{t\in\mathbb{R}_+}$ by noting that T_k is the first hitting time of state k, i.e.

$$T_k = \inf\{t \in \mathbb{R}_+ \ : \ N_t = k\}, \qquad k \geq 1,$$

with

$$\lim_{k\to\infty} T_k = \infty.$$

The value N_t at time t of the Poisson process can be recovered from its jump times $(T_k)_{k\geq 1}$ as

$$N_t = \sum_{k=1}^{\infty} k \mathbb{1}_{[T_k,T_{k+1})}(t) = \sum_{k=1}^{\infty} \mathbb{1}_{[T_k,\infty)}(t), \qquad t \in \mathbb{R}_+,$$

where

$$\mathbb{1}_{[T_k,\infty)}(t) = \begin{cases} 1 \text{ if } t \geq T_k, \\ \\ 0 \text{ if } 0 \leq t < T_k, \end{cases}$$

and

$$\mathbb{1}_{[T_k,T_{k+1})}(t) = \begin{cases} 1 \text{ if } T_k \leq t < T_{k+1}, \qquad k \geq 0, \\ \\ 0 \text{ if } 0 \leq t < T_k \text{ or } t \geq T_{k+1}, \qquad k \geq 0. \end{cases}$$

with $T_0 = 0$.

In addition, $(N_t)_{t\in\mathbb{R}_+}$ is assumed to satisfy the following conditions:

(i) Independence of increments: for all $0 \leq t_0 < t_1 < \cdots < t_n$ and $n \geq 1$ the increments

$$N_{t_1} - N_{t_0}, \ldots, \ N_{t_n} - N_{t_{n-1}},$$

over the disjoint time intervals $[t_0, t_1)$, $[t_1, t_2)$, ..., $[t_{n-2}, t_{n-1})$, $[t_{n-1}, t_n]$ are mutually independent random variables.

(ii) Stationarity of increments: $N_{t+h} - N_{s+h}$ has the same distribution as $N_t - N_s$ for all $h > 0$ and $0 \leq s \leq t$.

The meaning of the above stationarity condition is that for all fixed $k \in \mathbb{N}$ we have

$$\mathbb{P}(N_{t+h} - N_{s+h} = k) = \mathbb{P}(N_t - N_s = k),$$

for all $h > 0$ and $0 \leq s \leq t$.

The stationarity of increments means that for all $k \in \mathbb{N}$, the probability $\mathbb{P}(N_{t+h} - N_{s+h} = k)$ does not depend on $h > 0$.

Based on the above assumption, a natural question arises:

what is the distribution of N_t at time t?

We already know that N_t takes values in $\mathbb{N}$ and therefore it has a discrete distribution for all $t \in \mathbb{R}_+$. It is a remarkable fact that the distribution of the increments of $(N_t)_{t\in\mathbb{R}_+}$, can be completely determined from the above conditions, as shown in the following theorem.

As seen in the next result, the random variable $N_t - N_s$ has the *Poisson distribution* with parameter $\lambda(t - s)$.

Theorem 9.1 *Assume that the counting process $(N_t)_{t\in\mathbb{R}_+}$ satisfies the independence and stationarity Conditions (i) and (ii) above. Then we have*

$$\mathbb{P}(N_t - N_s = k) = \mathrm{e}^{-\lambda(t-s)} \frac{(\lambda(t-s))^k}{k!}, \qquad k \in \mathbb{N}, \quad 0 \leq s \leq t,$$

for some constant $\lambda > 0$.

Theorem 9.1 shows in particular that

$$\mathbb{E}[N_t - N_s] = \lambda(t - s) \quad \text{and} \quad \mathrm{Var}[N_t - N_s] = \lambda(t - s),$$

$0 \leq s \leq t$, cf. Relations (14.4) and (14.5) in the solution of Exercise 1.3-(a).

The parameter $\lambda > 0$ is called the *intensity* of the process and it can be recovered given from $\mathbb{P}(N_h = 1) = \lambda h \mathrm{e}^{-\lambda h}$ as the limit

$$\lambda = \lim_{h \searrow 0} \frac{1}{h} \mathbb{P}(N_h = 1). \tag{9.1.1}$$

Proof of Theorem 9.1. We only quote the main steps of the proof and we refer to [BN96] for the complete argument. Using the independence and stationarity of increments, we show that the probability generating function

$$G_t(u) := \mathbb{E}[u^{N_t}], \qquad -1 \leq u \leq 1,$$

satisfies

$$G_t(u) := (G_1(u))^t, \qquad -1 \le u \le 1,$$

which implies that

$$G_t(u) := \mathrm{e}^{-tf(u)}, \qquad -1 \le u \le 1,$$

for some function $f(u)$ of u. Still relying on the independence and stationarity of increments, it can be shown that $f(u)$ takes the form

$$f(u) = \lambda \times (1-u), \qquad -1 \le u \le 1,$$

where $\lambda > 0$ is given by (9.1.1). $\square$

In particular, given that $N_0 = 0$, the random variable N_t has a Poisson distribution with parameter λt:

$$\mathbb{P}(N_t = k) = \frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}, \qquad t \in \mathbb{R}_+.$$

From (9.1.1) above we see that[1]

$$\begin{cases} \mathbb{P}(N_h = 0) = \mathrm{e}^{-\lambda h} = 1 - \lambda h + o(h), \qquad h \searrow 0, \\ \\ \mathbb{P}(N_h = 1) = h\lambda \mathrm{e}^{-\lambda h} \simeq \lambda h, \qquad h \searrow 0, \end{cases}$$

and more generally that

$$\begin{cases} \mathbb{P}(N_{t+h} - N_t = 0) = \mathrm{e}^{-\lambda h} = 1 - \lambda h + o(h), \qquad h \searrow 0, & (9.1.2a) \\ \\ \mathbb{P}(N_{t+h} - N_t = 1) = \lambda h \mathrm{e}^{-\lambda h} \simeq \lambda h, \qquad h \searrow 0, & (9.1.2b) \\ \\ \mathbb{P}(N_{t+h} - N_t = 2) = h^2 \dfrac{\lambda^2}{2} \mathrm{e}^{-\lambda h} \simeq h^2 \dfrac{\lambda^2}{2}, \qquad h \searrow 0, & (9.1.2c) \end{cases}$$

for all $t \in \mathbb{R}_+$. This means that within "short" time intervals $[kh, (k+1)h]$ of length $h = t/n > 0$, the increments $N_{(k+1)h} - N_{kh}$ can be approximated by independent Bernoulli random variables X_{kh} with parameter λh, whose sum

$$\sum_{k=0}^{n-1} X_{kh} \simeq \sum_{k=0}^{n-1} (N_{(k+1)h} - N_{kh}) = N_t - N_0 = N_t$$

[1] The notation $f(h) \simeq h^k$ means $\lim_{h\to 0} f(h)/h^k = 1$, and $f(h) = o(h)$ means $\lim_{h\to 0} f(h)/h = 0$.

converges in distribution as n goes to infinity to the Poisson random variable N_t with parameter λt. This remark can be used for the random simulation of Poisson process paths.

More generally, we have

$$\mathbb{P}(N_{t+h} - N_t = k) \simeq h^k \frac{\lambda^k}{k!}, \qquad h \searrow 0, \qquad t > 0.$$

In order to determine the distribution of the first jump time T_1 we note that we have the equivalence

$$\{T_1 > t\} \Longleftrightarrow \{N_t = 0\},$$

which implies

$$\mathbb{P}(T_1 > t) = \mathbb{P}(N_t = 0) = \mathrm{e}^{-\lambda t}, \qquad t \geq 0,$$

i.e. T_1 has an exponential distribution with parameter $\lambda > 0$.

In order to prove the next proposition we note that more generally, we have the equivalence

$$\{T_n > t\} = \{N_t < n\}, \qquad t \geq 0, \quad n \geq 1.$$

Indeed, stating that the nth jump time T_n is strictly larger that t is equivalent to saying that at most $n - 1$ jumps of the Poisson process have occurred over the interval $[0, t]$, i.e. $N_t \leq n - 1$. The next proposition shows that T_n has a gamma distribution with parameter (λ, n) for $n \geq 1$, also called the Erlang distribution in queueing theory.

Proposition 9.2 *The random variable T_n has the gamma probability density function*

$$x \longmapsto \lambda^n \mathrm{e}^{-\lambda x} \frac{x^{n-1}}{(n-1)!}$$

$x \in \mathbb{R}_+$, $n \geq 1$.

Proof For $n = 1$ we have

$$\mathbb{P}(T_1 > t) = \mathbb{P}(N_t = 0) = \mathrm{e}^{-\lambda t}, \qquad t \in \mathbb{R}_+,$$

and by induction on $n \geq 1$, assuming that

$$\mathbb{P}(T_{n-1} > t) = \lambda \int_t^\infty \mathrm{e}^{-\lambda s} \frac{(\lambda s)^{n-2}}{(n-2)!} ds,$$

at the rank $n-1$ with $n \geq 2$, we obtain

$$\begin{aligned}
\mathbb{P}(T_n > t) &= \mathbb{P}(T_n > t \geq T_{n-1}) + \mathbb{P}(T_{n-1} > t) \\
&= \mathbb{P}(N_t = n-1) + \mathbb{P}(T_{n-1} > t) \\
&= \mathrm{e}^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} + \lambda \int_t^\infty \mathrm{e}^{-\lambda s} \frac{(\lambda s)^{n-2}}{(n-2)!} ds \\
&= \lambda \int_t^\infty \mathrm{e}^{-\lambda s} \frac{(\lambda s)^{n-1}}{(n-1)!} ds, \qquad t \in \mathbb{R}_+,
\end{aligned}$$

which proves the desired relation at the rank n, where we applied an integration by parts on $\mathbb{R}_+$ to derive the last line. □

Let now

$$\tau_k = T_{k+1} - T_k, \qquad k \geq 1,$$

denote the time spent in state $k \in \mathbb{N}$, with $T_0 = 0$. In addition to Proposition 9.2 we could show the following proposition which is based on the *strong Markov property*, see e.g. Theorem 6.5.4 of [Nor98], (9.2.4) below and Exercise 5.8 in discrete time.

Proposition 9.3 *The random inter-jump times*

$$\tau_k := T_{k+1} - T_k$$

spent in state $k \in \mathbb{N}$ form a sequence of independent identically distributed random variables having the exponential distribution with parameter $\lambda > 0$, i.e.

$$\mathbb{P}(\tau_0 > t_0, \tau_1 > t_1, \ldots, \tau_n > t_n) = \mathrm{e}^{-\lambda(t_0+t_1+\cdots+t_n)}, \qquad t_0, t_1, \ldots, t_n \in \mathbb{R}_+.$$

Random samples of Poisson process jump times can be generated using the following *R* code.

```
lambda = 2.0
n = 10
for (k in 1:n){tauk <- rexp(n)/lambda; Ti <- cumsum(tauk)}
tauk
Ti
```

Similarly, random samples of Poisson process paths can be generated using the following code.

```
n<-100
x<-cumsum(rexp(50,rate=0.5))
y<-cumsum(c(0,rep(1,50)))
plot(stepfun(x,y),xlim = c(0,10),do.points = F,main="L=0.5")
```

In other words, we have

$$\mathbb{P}(\tau_0 > t_0, \tau_1 > t_1, \ldots, \tau_n > t_n) = \mathbb{P}(\tau_0 > t_0) \times \cdots \times \mathbb{P}(\tau_n > t_n) \qquad (9.1.3)$$
$$= \prod_{k=0}^{n} \mathrm{e}^{-\lambda t_k}$$
$$= \mathrm{e}^{-\lambda(t_0+\cdots+t_n)},$$

for all $t_0, t_1, \ldots, t_n \in \mathbb{R}_+$. In addition, from Proposition 9.2 the sum

$$T_k = \tau_0 + \tau_1 + \cdots + \tau_{k-1}, \qquad k \geq 1,$$

has a gamma distribution with parameter (λ, k), cf. also Exercise 9.12 for a proof in the particular case $k = 2$.

As the expectation of the exponentially distributed random variable τ_k with parameter $\lambda > 0$ is given by

$$\mathbb{E}[\tau_k] = \lambda \int_0^\infty x \mathrm{e}^{-\lambda x} dx = \frac{1}{\lambda},$$

we can check that the higher the intensity λ (i.e. the higher the probability of having a jump within a small interval), the smaller is the time spent in each state $k \in \mathbb{N}$ on average. Poisson random samples on arbitrary spaces will be considered in Chap. 11.

9.2 Continuous-Time Markov Chains

A $\mathbb{S}$-valued continuous-time stochastic process $(X_t)_{t\in\mathbb{R}_+}$ is said to be *Markov*, or to have the *Markov property* if, for all $t \in [s, \infty)$, the probability distribution of X_t given the past of the process up to time s is determined by the state X_s of the process at time s, and does not depend on the past values of X_u for $u < s$. In other words, for all

$$0 < s_1 < \cdots < s_{n-1} < s < t$$

we have

$$\mathbb{P}(X_t = j \mid X_s = i_n,\ X_{s_{n-1}} = i_{n-1}, \ldots, X_{s_1} = i_0) = \mathbb{P}(X_t = j \mid X_s = i_n). \qquad (9.2.1)$$

In particular we have

$$\mathbb{P}(X_t = j \mid X_s = i_n \text{ and } X_{s_{n-1}} = i_{n-1}) = \mathbb{P}(X_t = j \mid X_s = i_n).$$

Example

The Poisson process $(N_t)_{t\in\mathbb{R}_+}$ considered in Sect. 9.1 is a continuous-time Markov chain because it has independent increments by Condition (i) p. 212. The birth and death processes discussed below are also continuous-time Markov chains, although they may not have independent increments.

More generally, any continuous-time process $(X_t)_{t\in\mathbb{R}_+}$ with independent increments has the Markov property. Indeed, for all $j, i_n, \ldots, i_1 \in \mathbb{S}$ we have (note that $X_0 = 0$ here)

$$\begin{aligned}
&\mathbb{P}(X_t = j \mid X_s = i_n,\ X_{s_{n-1}} = i_{n-1}, \ldots, X_{s_1} = i_1)\\
&= \frac{\mathbb{P}(X_t = j, X_s = i_n, X_{s_{n-1}} = i_{n-1}, \ldots, X_{s_1} = i_1)}{\mathbb{P}(X_s = i_n,\ X_{s_{n-1}} = i_{n-1}, \ldots, X_{s_1} = i_1)}\\
&= \frac{\mathbb{P}(X_t - X_s = j - i_n, X_s = i_n, \ldots, X_{s_2} = i_2, X_{s_1} = i_1)}{\mathbb{P}(X_s = i_n, , \ldots, X_{s_2} = i_2, X_{s_1} = i_1)}\\
&= \frac{\mathbb{P}(X_t - X_s = j - i_n)\mathbb{P}(X_s = i_n, X_{s_{n-1}} = i_{n-1}, \ldots, X_{s_2} = i_2, X_{s_1} = i_1)}{\mathbb{P}(X_s = i_n, X_{s_{n-1}} = i_{n-1}, \ldots, X_{s_2} = i_2, X_{s_1} = i_1)}\\
&= \mathbb{P}(X_t - X_s = j - i_n) = \frac{\mathbb{P}(X_t - X_s = j - i_n)\mathbb{P}(X_s = i_n)}{\mathbb{P}(X_s = i_n)}\\
&= \frac{\mathbb{P}(X_t - X_s = j - i_n \text{ and } X_s = i_n)}{\mathbb{P}(X_s = i_n)}\\
&= \frac{\mathbb{P}(X_t = j \text{ and } X_s = i_n)}{\mathbb{P}(X_s = i_n)} = \mathbb{P}(X_t = j \mid X_s = i_n),
\end{aligned}$$

cf. (4.1.5) for the discrete-time version of this argument. Hence, continuous-time processes with *independent increments* are Markov chains. However, *not all continuous-time Markov chains have independent increments*, and in fact the continuous-time Markov chains of interest in this chapter will not have independent increments.

Birth Process

The pure birth process behaves similarly to the Poisson process, by making the parameter of every exponential inter-jump time dependent on the current state of the process.

In other words, a continuous-time Markov chain $(X^b_t)_{t\in\mathbb{R}_+}$ such that[2]

$$\begin{aligned}
\mathbb{P}(X^b_{t+h} = i + 1 \mid X^b_t = i) &= \mathbb{P}(X^b_{t+h} - X^b_t = 1 \mid X^b_t = i)\\
&\simeq \lambda_i h, \qquad h \searrow 0, \quad i \in \mathbb{S},
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{P}(X^b_{t+h} = X^b_t \mid X^b_t = i) &= \mathbb{P}(X^b_{t+h} - X^b_t = 0 \mid X^b_t = i)\\
&= 1 - \lambda_i h + o(h), \qquad h \searrow 0, \quad i \in \mathbb{S}, \qquad (9.2.2)
\end{aligned}$$

is called a *pure birth process* with (possibly) state-dependent birth rates $\lambda_i \geq 0, i \in \mathbb{S}$, see Fig. 9.2. Its inter-jump times $(\tau_k)_{k\geq 0}$ form a sequence of exponential independent random variables with state-dependent parameters.

[2]Recall that by definition $f(h) \simeq g(h)$, $h \to 0$, if and only if $\lim_{h\to 0} f(h)/g(h) = 1$.

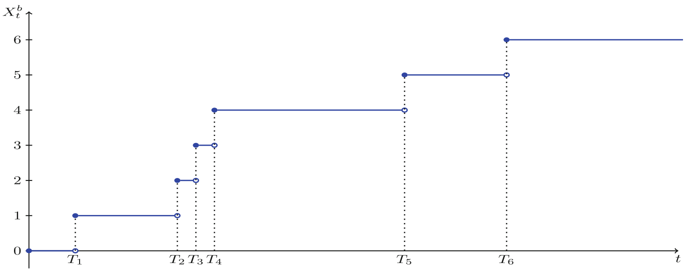


Fig. 9.2 Sample path of a birth process $(X_t^b)_{t\in\mathbb{R}_+}$

This process is stationary in time because the rates λ_i, $i \in \mathbb{N}$, are independent of time t. The Poisson process $(N_t)_{t\in\mathbb{R}_+}$ is a pure birth process with state-independent birth rates $\lambda_i = \lambda > 0$, $i \in \mathbb{N}$.

As a consequence of (9.2.2) we can recover the fact that the time $\tau_{i,i+1}$ spent in state i by the pure birth process $(X_t^b)_{t\in\mathbb{R}_+}$ started at state i at time 0 before it moves to state $\boxed{i+1}$ has an exponential distribution with parameter λ_i. Indeed we have, using the Markov property in continuous time,

$$\begin{aligned}
\mathbb{P}(\tau_{i,i+1} > t+h \mid \tau_{i,i+1} > t \text{ and } X_0^b = i) &= \frac{\mathbb{P}(\tau_{i,i+1} > t+h \mid X_0^b = i)}{\mathbb{P}(\tau_{i,i+1} > t \mid X_0^b = i)} \\
&= \frac{\mathbb{P}(X_{t+h}^b = i \mid X_0^b = i)}{\mathbb{P}(X_t^b = i \mid X_0^b = i)} \\
&= \frac{\mathbb{P}(X_{t+h}^b = i \text{ and } X_0^b = i)\mathbb{P}(X_0^b = i)}{\mathbb{P}(X_t^b = i \text{ and } X_0^b = i)\mathbb{P}(X_0^b = i)} \\
&= \frac{\mathbb{P}(X_{t+h}^b = i \text{ and } X_0^b = i)}{\mathbb{P}(X_t^b = i \text{ and } X_0^b = i)} \\
&= \frac{\mathbb{P}(X_{t+h}^b = i,\ X_t^b = i,\ X_0^b = i)}{\mathbb{P}(X_t^b = i \text{ and } X_0^b = i)} \\
&= \mathbb{P}(X_{t+h}^b = i \mid X_t^b = i \text{ and } X_0^b = i) \\
&= \mathbb{P}(X_{t+h}^b = i \mid X_t^b = i) \\
&= \mathbb{P}(X_h^b = i \mid X_0^b = i) \\
&= \mathbb{P}(\tau_{i,i+1} > h \mid X_0^b = i) \\
&= 1 - \lambda_i h + o(h), \qquad (9.2.3)
\end{aligned}$$

which is often referred to as the *memoryless property* of Markov processes. In other words, since the above ratio is independent of $t > 0$ we get

$$\mathbb{P}(\tau_{i,i+1} > t+h \mid \tau_{i,i+1} > t \text{ and } X_0^b = i) = \mathbb{P}(\tau_{i,i+1} > h \mid X_0^b = i),$$

which means that the distribution of the waiting time after time t does not depend on t, cf. (12.1.1) in Chap. 12 for a similar argument.

From (9.2.3) we have

$$\begin{aligned}
&\frac{\mathbb{P}(\tau_{i,i+1} > t + h \mid X_0^b = i) - \mathbb{P}(\tau_{i,i+1} > t \mid X_0^b = i)}{h\mathbb{P}(\tau_{i,i+1} > t \mid X_0^b = i)} \\
&= \frac{\mathbb{P}(\tau_{i,i+1} > t + h \mid X_0^b = i)}{h\mathbb{P}(\tau_{i,i+1} > t \mid X_0^b = i)} - 1 \\
&= \mathbb{P}(\tau_{i,i+1} > t + h \mid \tau_{i,i+1} > t \text{ and } X_0^b = i) - 1 \\
&\simeq -\lambda_i, \qquad h \to 0,
\end{aligned}$$

which can be read as the differential equation

$$\frac{d}{dt} \log \mathbb{P}(\tau_{i,i+1} > t \mid X_0^b = i) = -\lambda_i,$$

where "log" denotes the *natural logarithm* "ln", with solution

$$\mathbb{P}(\tau_{i,i+1} > t \mid X_0^b = i) = \mathrm{e}^{-\lambda_i t}, \qquad t \in \mathbb{R}_+, \tag{9.2.4}$$

i.e. $\tau_{i,i+1}$ is an exponentially distributed random variable with parameter λ_i, and the mean time spent at state i before switching to state $\boxed{i+1}$ is given by

$$\mathbb{E}[\tau_{i,i+1} \mid X_0^b = i] = \frac{1}{\lambda_i}, \qquad i \in \mathbb{S},$$

see (9.4.9) below for the general case of continuous-time Markov chains. More generally, and similarly to (9.1.3) it can also be shown as a consequence of the *strong Markov property* that the sequence $(\tau_{j,j+1})_{j \geq i}$ is made of independent random variables which are respectively exponentially distributed with parameters λ_j, $j \geq i$.

Letting $T_{i,j}^b = \tau_{i,i+1} + \cdots + \tau_{j-1,j}$ denote the hitting time of state j starting from state i by the birth process $(X_t^b)_{t \in \mathbb{R}_+}$, we have the representation

$$X_t^b = i + \sum_{i<j<\infty} \mathbb{1}_{[T_{i,j}^b, \infty)}(t), \qquad t \in \mathbb{R}_+.$$

Note that since the pure birth process has stationary increments, by Theorem 9.1 it can have independent increments only when the rates $\lambda_i = \lambda$ are state independent, i.e. when $(X_t^b)_{t \in \mathbb{R}_+}$ is a standard Poisson process with intensity $\lambda > 0$.

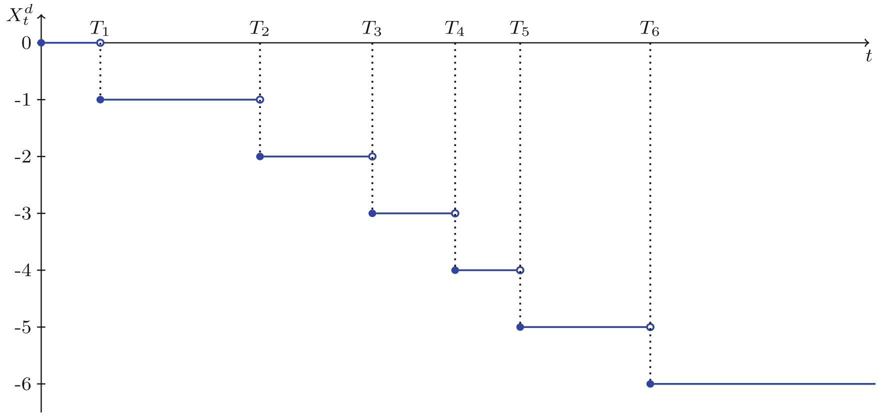


Fig. 9.3 Sample path of a death process $(X_t^d)_{t\in\mathbb{R}_+}$

Death Process

A continuous-time Markov chain $(X_t^d)_{t\in\mathbb{R}_+}$ such that

$$\begin{cases} \mathbb{P}(X_{t+h}^d - X_t^d = -1 \mid X_t^d = i) \simeq \mu_i h, & h \searrow 0, \quad i \in \mathbb{S}, \\ \mathbb{P}(X_{t+h}^d - X_t^d = 0 \mid X_t^d = i) = 1 - \mu_i h + o(h), & h \searrow 0, \quad i \in \mathbb{S}, \end{cases}$$

is called a *pure death process* with (possibly) state-dependent death rates $\mu_i \geq 0, i \in \mathbb{S}$. Its inter-jump times $(\tau_k)_{k\geq 0}$ form a sequence of exponential independent random variables with state-dependent parameters (Fig. 9.3).

In the case of a pure death process $(X_t^d)_{t\in\mathbb{R}_+}$ we denote by $\tau_{i,i-1}$ the time spent in state i by $(X_t^d)_{t\in\mathbb{R}_+}$ before it moves to state $\boxed{i\text{-}1}$. Similarly to the pure birth process, that the sequence $(\tau_{j,j-1})_{j\leq i}$ is made of independent random variables which are exponentially distributed with parameter μ_j, $j \leq i$, which

$$\mathbb{P}(\tau_{j,j-1} > t) = \mathrm{e}^{-\mu_j t}, \qquad t \in \mathbb{R}_+,$$

and

$$\mathbb{E}[\tau_{i,i-1}] = \frac{1}{\mu_i}, \qquad i \in \mathbb{S}.$$

Letting $T_{i,j}^d = \tau_{i,i-1} + \cdots + \tau_{j+1,j}$ denote the hitting time of state j starting from state i by the death process $(X_t^d)_{t\in\mathbb{R}_+}$ we have the representation

$$X_t^d = i - \sum_{-\infty<j<i} \mathbb{1}_{[T_{i,j}^d,\infty)}(t), \qquad t \in \mathbb{R}_+.$$

When $(N_t)_{t\in\mathbb{R}_+}$ is a Poisson process, the process $(-N_t)_{t\in\mathbb{R}_+}$ is a pure death process with state-independent death rates $\mu_n = \lambda > 0$, $n \in \mathbb{N}$.

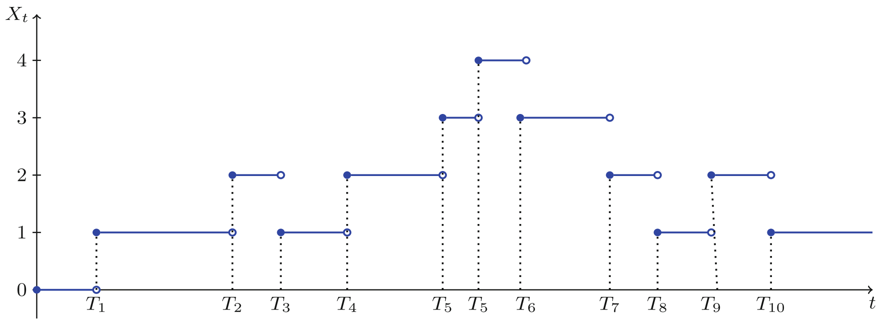


Fig. 9.4 Sample path of a birth and death process $(X_t)_{t\in\mathbb{R}_+}$

Birth and Death Process

A continuous-time Markov chain $(X_t)_{t\in\mathbb{R}_+}$ such that, for all $i \in \mathbb{S}$,

$$\begin{cases} \mathbb{P}(X_{t+h} - X_t = 1 \mid X_t = i) \simeq \lambda_i h, \quad h \searrow 0, & \text{(9.2.5a)} \\ \mathbb{P}(X_{t+h} - X_t = -1 \mid X_t = i) \simeq \mu_i h, \quad h \searrow 0, \text{ and} & \\ \mathbb{P}(X_{t+h} - X_t = 0 \mid X_t = i) = 1 - (\lambda_i + \mu_i)h + o(h), \quad h \searrow 0, & \text{(9.2.5b)} \end{cases}$$

is called a *birth and death process* with (possibly) state-dependent birth rates $\lambda_i \geq 0$ and death rates $\mu_i \geq 0$, $i \in \mathbb{S}$ (Fig. 9.4).

The birth and death process $(X_t)_{t\in\mathbb{R}_+}$ can be built as

$$X_t = X_t^b + X_t^d, \qquad t \in \mathbb{R}_+,$$

in which case the time τ_i spent in state (i) by $(X_t)_{t\in\mathbb{R}_+}$ satisfies the identity in distribution

$$\tau_i = \min(\tau_{i,i+1}, \tau_{i,i-1})$$

i.e. τ_i is an exponentially distributed random variable with parameter $\lambda_i + \mu_i$ and

$$\mathbb{E}[\tau_i] = \frac{1}{\lambda_i + \mu_i}.$$

Indeed, since $\tau_{i,i+1}$ and $\tau_{i,i-1}$ are two independent exponentially distributed random variables with parameters λ_i and μ_i, we have

$$\begin{aligned} \mathbb{P}(\min(\tau_{i,i+1} \text{ and } \tau_{i,i-1}) > t) &= \mathbb{P}(\tau_{i,i+1} > t \text{ and } \tau_{i,i-1} > t) \\ &= \mathbb{P}(\tau_{i,i+1} > t)\mathbb{P}(\tau_{i,i-1} > t) \\ &= \mathrm{e}^{-t(\lambda_i+\mu_i)}, \qquad t \in \mathbb{R}_+, \end{aligned}$$

hence $\tau_i = \min(\tau_{i,i+1}, \tau_{i,i-1})$ is an exponentially distributed random variable with parameter $\lambda_i + \mu_i$, cf. also (1.5.8) in Chap. 1.

9.3 Transition Semigroup

The transition semigroup of the continuous time Markov process $(X_t)_{t\in\mathbb{R}_+}$ is the family $(P(t))_{t\in\mathbb{R}_+}$ of matrices determined by

$$P_{i,j}(t) := \mathbb{P}(X_{t+s} = j \mid X_s = i), \qquad i, j \in \mathbb{S}, \quad s, t \in \mathbb{R}_+,$$

where we assume that the probability $\mathbb{P}(X_{t+s} = j \mid X_s = i)$ does not depend on $s \in \mathbb{R}_+$. In this case the Markov process $(X_t)_{t\in\mathbb{R}_+}$ is said to be *time homogeneous*.

Definition 9.4 A continuous-time Markov chain $(X_t)_{t\in\mathbb{R}_+}$ is *irreducible* if for all $t > 0$, $P(t)$ is the transition matrix of an irreducible discrete-time chain.

Note that we always have

$$P(0) = \mathrm{I_d}.$$

This data can be recorded as a time-dependent matrix indexed by $\mathbb{S}^2 = \mathbb{S} \times \mathbb{S}$, called the *transition semigroup* of the Markov process:

$$\left[\, P_{i,j}(t) \,\right]_{i,j\in\mathbb{S}} = \left[\, \mathbb{P}(X_{t+s} = j \mid X_s = i) \,\right]_{i,j\in\mathbb{S}},$$

also written as

$$\left[\, P_{i,j}(t) \,\right]_{i,j\in\mathbb{S}} = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & ⋰ \\ \cdots & P_{-2,-2}(t) & P_{-2,-1}(t) & P_{-2,0}(t) & P_{-2,1}(t) & P_{-2,2}(t) & \cdots \\ \cdots & P_{-1,-2}(t) & P_{-1,-1}(t) & P_{-1,0}(t) & P_{-1,1}(t) & P_{-1,2}(t) & \cdots \\ \cdots & P_{0,-2}(t) & P_{0,-1}(t) & P_{0,0}(t) & P_{0,1}(t) & P_{0,2}(t) & \cdots \\ \cdots & P_{1,-2}(t) & P_{1,-1}(t) & P_{1,0}(t) & P_{1,1}(t) & P_{1,2}(t) & \cdots \\ \cdots & P_{2,-2}(t) & P_{2,-1}(t) & P_{2,0}(t) & P_{2,1}(t) & P_{2,2}(t) & \cdots \\ ⋰ & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

As in the discrete-time case, note the inversion of the order of indices (i, j) between $\mathbb{P}(X_{t+s} = j \mid X_s = i)$ and $P_{i,j}(t)$. In particular, the initial state i correspond to a *row number* in the matrix $P(t)$, while the final state j corresponds to a *column number*.

Due to the relation

$$\sum_{j \in \mathbb{S}} \mathbb{P}(X_{t+s} = j \mid X_s = i) = 1, \qquad i \in \mathbb{S}, \tag{9.3.1}$$

all *rows* of the transition matrix semigroup $(P(t))_{t \in \mathbb{R}_+}$ satisfy the condition

$$\sum_{j \in \mathbb{S}}^{\infty} P_{i,j}(t) = 1,$$

for $i \in \mathbb{S}$. In the sequel we will only consider $\mathbb{N}$-valued Markov process, and in this case the *transition semigroup* $(P(t))_{t \in \mathbb{R}_+}$ of the Markov process is written as

$$P(t) = \left[\, P_{i,j}(t) \,\right]_{i,j \in \mathbb{N}} = \begin{bmatrix} P_{0,0}(t) & P_{0,1}(t) & P_{0,2}(t) & \cdots \\ P_{1,0}(t) & P_{1,1}(t) & P_{1,2}(t) & \cdots \\ P_{2,0}(t) & P_{2,1}(t) & P_{2,2}(t) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

From (9.3.1) we have

$$\sum_{j=0}^{\infty} P_{i,j}(t) = 1,$$

for all $i \in \mathbb{N}$ and $t \in \mathbb{R}_+$.

<u>Exercise</u>: Write down the transition semigroup $\left[\, P_{i,j}(t) \,\right]_{i,j \in \mathbb{N}}$ of the Poisson process $(N_t)_{t \in \mathbb{R}_+}$.

We can show that

$$\left[\, P_{i,j}(t) \,\right]_{i,j \in \mathbb{N}} = \begin{bmatrix} e^{-\lambda t} & \lambda t e^{-\lambda t} & \dfrac{\lambda^2 t^2}{2} e^{-\lambda t} & \cdots \\ 0 & e^{-\lambda t} & \lambda t e^{-\lambda t} & \cdots \\ 0 & 0 & e^{-\lambda t} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Indeed we have

$$\begin{aligned}
P_{i,j}(t) &= \mathbb{P}(N_{s+t} = j \mid N_s = i) = \frac{\mathbb{P}(N_{s+t} = j \text{ and } N_s = i)}{\mathbb{P}(N_s = i)} \\
&= \frac{\mathbb{P}(N_{s+t} - N_s = j - i \text{ and } N_s = i)}{\mathbb{P}(N_s = i)} = \frac{\mathbb{P}(N_{s+t} - N_s = j - i)\mathbb{P}(N_s = i)}{\mathbb{P}(N_s = i)} \\
&= \mathbb{P}(N_{s+t} - N_s = j - i) = \begin{cases} \mathrm{e}^{-\lambda t} \dfrac{(\lambda t)^{j-i}}{(j-i)!} & \text{if } j \geq i, \\ \\ 0 & \text{if } j < i. \end{cases}
\end{aligned}$$

In case the Markov process $(X_t)_{t\in\mathbb{R}_+}$ takes values in the finite state space $\{0, 1, \ldots, N\}$ its transition semigroup will simply have the form

$$P(t) = \left[\, P_{i,j}(t) \,\right]_{0\leq i,j\leq N} = \begin{bmatrix} P_{0,0}(t) & P_{0,1}(t) & P_{0,2}(t) & \cdots & P_{0,N}(t) \\ P_{1,0}(t) & P_{1,1}(t) & P_{1,2}(t) & \cdots & P_{1,N}(t) \\ P_{2,0}(t) & P_{2,1}(t) & P_{2,2}(t) & \cdots & P_{2,N}(t) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{N,0}(t) & P_{N,1}(t) & P_{N,2}(t) & \cdots & P_{N,N}(t) \end{bmatrix}.$$

As noted above, the semigroup matrix $P(t)$ is a convenient way to record the values of $\mathbb{P}(X_{t+s} = j \mid X_s = i)$ in a table.

Proposition 9.5 *The family $(P(t))_{t\in\mathbb{R}_+}$ satisfies the relation*

$$\boxed{P(s+t) = P(s)P(t) = P(t)P(s), \qquad (9.3.2)}$$

which is called the semigroup property.

Proof Using the Markov property and denoting by $\mathbb{S}$ the state space of the process, by standard arguments based on the *law of total probability* (1.3.1) for the probability measure $\mathbb{P}(\cdot \mid X_0 = i)$ and the Markov property (9.2.1), we have

$$\begin{aligned}
P_{i,j}(t+s) &= \mathbb{P}(X_{t+s} = j \mid X_0 = i) \\
&= \sum_{l\in\mathbb{S}} \mathbb{P}(X_{t+s} = j \text{ and } X_s = l \mid X_0 = i) = \sum_{l\in\mathbb{S}} \frac{\mathbb{P}(X_{t+s} = j,\ X_s = l,\ X_0 = i)}{\mathbb{P}(X_0 = i)} \\
&= \sum_{l\in\mathbb{S}} \frac{\mathbb{P}(X_{t+s} = j,\ X_s = l,\ X_0 = i)}{\mathbb{P}(X_s = l \text{ and } X_0 = i)} \frac{\mathbb{P}(X_s = l \text{ and } X_0 = i)}{\mathbb{P}(X_0 = i)} \\
&= \sum_{l\in\mathbb{S}} \mathbb{P}(X_{t+s} = j \mid X_s = l \text{ and } X_0 = i)\mathbb{P}(X_s = l \mid X_0 = i) \\
&= \sum_{l\in\mathbb{S}} \mathbb{P}(X_{t+s} = j \mid X_s = l)\mathbb{P}(X_s = l \mid X_0 = i) = \sum_{l\in\mathbb{S}} P_{i,l}(s)P_{l,j}(t) \\
&= [P(s)P(t)]_{i,j},
\end{aligned}$$

$i, j \in \mathbb{S}$, $s, t \in \mathbb{R}_+$. We have shown the relation

$$P_{i,j}(s+t) = \sum_{l\in\mathbb{S}} P_{i,l}(s)P_{l,j}(t),$$

which leads to (9.3.2). □

From (9.3.2) property one can check in particular that the matrices $P(s)$ and $P(t)$ *commute*, i.e. we have

$$P(s)P(t) = P(t)P(s), \qquad s, t \in \mathbb{R}_+.$$

Example

For the transition semigroup $(P(t))_{t\in\mathbb{R}_+}$ of the Poisson process we can check by hand computation that

$$P(s)P(t) = \begin{bmatrix} e^{-\lambda s} & \lambda s e^{-\lambda s} & \frac{\lambda^2}{2}s^2 e^{-\lambda s} & \cdots \\ 0 & e^{-\lambda s} & \lambda s e^{-\lambda s} & \cdots \\ 0 & 0 & e^{-\lambda s} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \times \begin{bmatrix} e^{-\lambda t} & \lambda t e^{-\lambda t} & \frac{\lambda^2}{2}t^2 e^{-\lambda t} & \cdots \\ 0 & e^{-\lambda t} & \lambda t e^{-\lambda t} & \cdots \\ 0 & 0 & e^{-\lambda t} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$= \begin{bmatrix} e^{-\lambda(s+t)} & \lambda(s+t)e^{-\lambda(s+t)} & \frac{\lambda^2}{2}(s+t)^2 e^{-\lambda(s+t)} & \cdots \\ 0 & e^{-\lambda(s+t)} & \lambda(s+t)\, e^{-\lambda(s+t)} & \cdots \\ 0 & 0 & e^{-\lambda(s+t)} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$= P(s+t).$$

The above identity can be recovered by the following calculation, for all $0 \leq i \leq j$, which amounts to saying that the sum of two independent Poisson random variables with parameters s and t has a Poisson distribution with parameter $s+t$, cf. (14.7) in the solution of Exercise 1.6-(a). We have $P_{i,j}(s) = 0$, $i > j$, and $P_{l,j}(t) = 0$, $l > j$, hence

$$\begin{aligned}
[P(s)P(t)]_{i,j} &= \sum_{l=0}^{\infty} P_{i,l}(s)P_{l,j}(t) = \sum_{l=i}^{j} P_{i,l}(s)P_{l,j}(t) \\
&= \mathrm{e}^{-\lambda s-\lambda t} \sum_{l=i}^{j} \frac{(\lambda s)^{l-i}}{(l-i)!} \frac{(\lambda t)^{j-l}}{(j-l)!} = \mathrm{e}^{-\lambda s-\lambda t} \frac{1}{(j-i)!} \sum_{l=i}^{j} \binom{j-i}{l-i} (\lambda s)^{l-i} (\lambda t)^{j-l} \\
&= \mathrm{e}^{-\lambda s-\lambda t} \frac{1}{(j-i)!} \sum_{l=0}^{j-i} \binom{j-i}{l} (\lambda s)^{l} (\lambda t)^{j-i-l} = \mathrm{e}^{-\lambda(s+t)} \frac{1}{(j-i)!} (\lambda s + \lambda t)^{j-i} \\
&= P_{i,j}(s+t), \qquad s, t \in \mathbb{R}_+.
\end{aligned}$$

9.4 Infinitesimal Generator

The infinitesimal generator of a continuous-time Markov process allows us to encode all properties of the process $(X_t)_{t\in\mathbb{R}_+}$ in a single matrix.

By differentiating the semigroup relation (9.3.2) with respect to t we get, by componentwise differentiation and assuming a finite state space $\mathbb{S}$,

$$\begin{aligned}
P'(t) &= \lim_{h \searrow 0} \frac{P(t+h) - P(t)}{h} = \lim_{h \searrow 0} \frac{P(t)P(h) - P(t)}{h} \\
&= P(t) \lim_{h \searrow 0} \frac{P(h) - P(0)}{h} = P(t)Q,
\end{aligned}$$

where

$$Q := P'(0) = \lim_{h \searrow 0} \frac{P(h) - P(0)}{h}$$

is called the *infinitesimal generator* of $(X_t)_{t\in\mathbb{R}_+}$.

When $\mathbb{S} = \{0, 1, \ldots, N\}$ we will denote by $\lambda_{i,j}$, $i, j \in \mathbb{S}$, the entries of the infinitesimal generator matrix $Q = (\lambda_{i,j})_{i,j\in\mathbb{S}}$, i.e.

$$Q = \frac{dP(t)}{dt}_{|t=0} = [\ \lambda_{i,j}\ ]_{0 \le i,j \le N} = \begin{bmatrix} \lambda_{0,0} & \lambda_{0,1} & \lambda_{0,2} & \cdots & \lambda_{0,N} \\ \lambda_{1,0} & \lambda_{1,1} & \lambda_{1,2} & \cdots & \lambda_{1,N} \\ \lambda_{2,0} & \lambda_{2,1} & \lambda_{2,2} & \cdots & \lambda_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{N,0} & \lambda_{N,1} & \lambda_{N,2} & \cdots & \lambda_{N,N} \end{bmatrix}. \tag{9.4.1}$$

Denoting $Q = [\lambda_{i,j}]_{i,j\in\mathbb{S}}$, for all $i \in \mathbb{S}$ we have

$$\sum_{j\in\mathbb{S}} \lambda_{i,j} = \sum_{j\in\mathbb{S}} P'_{i,j}(0) = \frac{d}{dt} \sum_{j\in\mathbb{S}} P_{i,j}(t)_{|t=0} = \frac{d}{dt} \mathbb{1}_{|t=0} = 0,$$

hence the *rows* of the infinitesimal generator matrix $Q = [\lambda_{i,j}]_{i,j\in\mathbb{S}}$ always add up to 0, i.e.

$$\sum_{l\in\mathbb{S}} \lambda_{i,l} = \lambda_{i,i} + \sum_{l\neq i} \lambda_{i,l} = 0,$$

or

$$\lambda_{i,i} = -\sum_{l\neq i} \lambda_{i,l}. \tag{9.4.2}$$

Note that a state (i) such that $\lambda_{i,j} = 0$ for all $j \in \mathbb{S}$ is absorbing.

$$P'(t) = P(t)Q, \qquad t > 0, \tag{9.4.3}$$

is called the *forward Kolmogorov equation*, cf. (1.2.2). In a similar way we can show that

$$P'(t) = QP(t), \qquad t > 0, \tag{9.4.4}$$

which is called the *backward Kolmogorov equation*.

The forward and backward Kolmogorov equations (9.4.3)–(9.4.4) can be solved either using the matrix exponential e^{tQ} defined as

$$\exp(tQ) := \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n = \mathrm{I_d} + \sum_{n=1}^{\infty} \frac{t^n}{n!} Q^n, \tag{9.4.5}$$

or by viewing the Kolmogorov equations (9.4.3)–(9.4.4) component by component as systems of differential equations.

In (9.4.5) above, $Q^0 = \mathrm{I_d}$ is the identity matrix, written as

$$\mathrm{I_d} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

when the state space is $\mathbb{S} = \{0, 1, \ldots, N\}$. Using matrix exponentials, the solution of (9.4.3) is given by

$$\boxed{P(t) = P(0)\exp(tQ) = \exp(tQ), \quad t \in \mathbb{R}_+.}$$

We will often use the first order approximation in $h \to 0$ of

$$P(h) = \exp(hQ) = \mathrm{I_d} + \sum_{n=1}^{\infty} \frac{h^n}{n!} Q^n = \mathrm{I_d} + hQ + \frac{h^2}{2!}Q^2 + \frac{h^3}{3!}Q^3 + \frac{h^4}{4!}Q^4 + \cdots,$$

given by

$$P(h) = \mathrm{I_d} + hQ + o(h), \qquad h \searrow 0, \tag{9.4.6}$$

where $o(h)$ is a function such that $\lim_{h\to 0} o(h)/h = 0$, i.e.

$$P(h) = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} + h \begin{bmatrix} \lambda_{0,0} & \lambda_{0,1} & \lambda_{0,2} & \cdots & \lambda_{0,N} \\ \lambda_{1,0} & \lambda_{1,1} & \lambda_{1,2} & \cdots & \lambda_{1,N} \\ \lambda_{2,0} & \lambda_{2,1} & \lambda_{2,2} & \cdots & \lambda_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{N,0} & \lambda_{N,1} & \lambda_{N,2} & \cdots & \lambda_{N,N} \end{bmatrix} + o(h), \quad h \searrow 0.$$

Relation (9.4.6) yields the transition probabilities over a small time interval of length $h > 0$, as:

$$\mathbb{P}(X_{t+h} = j \mid X_t = i) = P_{i,j}(h) = \begin{cases} \lambda_{i,j}h + o(h), & i \neq j, \quad h \searrow 0, \\ 1 + \lambda_{i,i}h + o(h), & i = j, \quad h \searrow 0, \end{cases}$$

and by (9.4.2) we also have

$$\mathbb{P}(X_{t+h} = j \mid X_t = i) = P_{i,j}(h) = \begin{cases} \lambda_{i,j} h + o(h), & i \neq j, \quad h \searrow 0, \\ \\ 1 - h \sum_{l \neq i} \lambda_{i,l} + o(h), & i = j, \quad h \searrow 0. \end{cases} \tag{9.4.7}$$

For example, in the case of a two-state continuous-time Markov chain we have

$$Q = \begin{bmatrix} -\alpha & \alpha \\ \\ \beta & -\beta \end{bmatrix},$$

with $\alpha, \beta \geq 0$, and

$$\begin{aligned} P(h) &= \mathrm{I_d} + hQ + o(h) \\ &= \begin{bmatrix} 1 & 0 \\ \\ 0 & 1 \end{bmatrix} + h \begin{bmatrix} -\alpha & \alpha \\ \\ \beta & -\beta \end{bmatrix} + o(h) \\ &= \begin{bmatrix} 1 - \alpha h & \alpha h \\ \\ \beta h & 1 - \beta h \end{bmatrix} + o(h), \end{aligned} \tag{9.4.8}$$

as $h \searrow 0$. In this case, $P(h)$ above has the same form as the transition matrix (4.5.1) of a *discrete-time* Markov chain with "small" time step $h > 0$ and "small" transition probabilities $h\alpha$ and $h\beta$, namely $h\alpha$ is the probability of switching from state 0 to state 1, and $h\beta$ is the probability of switching from state 1 to state 0 within a short period of time $h > 0$.

We note that since

$$\mathbb{P}(X_{t+h} = j \mid X_t = i) \simeq \lambda_{i,j} h, \qquad h \searrow 0, \quad i \neq j,$$

and

$$\mathbb{P}(X_{t+h} \neq j \mid X_t = i) = 1 - \lambda_{i,j} h + o(h), \qquad h \searrow 0, \quad i \neq j,$$

the transition of the process $(X_t)_{t \in \mathbb{R}_+}$ from state i to state j behaves identically to that of a Poisson process with intensity $\lambda_{i,j}$, cf. (9.1.2a)–(9.1.2b) above. Similarly to the Poisson, birth and death processes, the relation

$$\mathbb{P}(X_{t+h} = j \mid X_t = i) = \lambda_{i,j} h + o(h), \qquad h \searrow 0, \quad i \neq j,$$

shows that the time $\tau_{i,j}$ spent in state i "before moving to state $j \neq i$", i.e. given the first jump is to state j, is an *exponentially distributed random variable* with parameter $\lambda_{i,j}$, i.e.

$$\mathbb{P}(\tau_{i,j} > t) = \mathrm{e}^{-\lambda_{i,j} t}, \qquad t \in \mathbb{R}_+, \tag{9.4.9}$$

and we have

$$\mathbb{E}[\tau_{i,j}] = \lambda_{i,j} \int_0^\infty t \mathrm{e}^{-t\lambda_{i,j}} dt = \frac{1}{\lambda_{i,j}}, \qquad i \neq j.$$

When $i = j$ we have

$$\mathbb{P}(X_{t+h} \neq i \mid X_t = i) \simeq h \sum_{l \neq i} \lambda_{i,l} = -\lambda_{i,i} h, \qquad h \searrow 0,$$

and

$$\mathbb{P}(X_{t+h} = i \mid X_t = i) = 1 - h \sum_{l \neq i} \lambda_{i,l} + o(h) = 1 + \lambda_{i,i} h + o(h), \qquad h \searrow 0,$$

hence, by the same Poisson process analogy, the time τ_i spent in state (i) before the next transition to a different state is an exponentially distributed random variable with parameter $\sum_{j \neq i} \lambda_{i,j}$, i.e.

$$\mathbb{P}(\tau_i > t) = \exp\left(-t \sum_{j \neq i} \lambda_{i,j}\right) = \mathrm{e}^{t\lambda_{i,i}}, \qquad t \in \mathbb{R}_+.$$

In other words, we can also write the time τ_i spent in state (i) as

$$\tau_i = \min_{\substack{j \in \mathbb{S} \\ j \neq i}} \tau_{i,j},$$

and this recovers the fact that τ_i is an exponential random variable with parameter $\sum_{j \neq i} \lambda_{i,j}$, since

$$\begin{aligned}
\mathbb{P}(\tau_i > t) &= \mathbb{P}\left(\min_{\substack{j \in \mathbb{S} \\ j \neq i}} \tau_{i,j} > t\right) \\
&= \prod_{\substack{j \in \mathbb{S} \\ j \neq i}} \mathbb{P}(\tau_{i,j} > t) \\
&= \exp\left(-t \sum_{j \neq i} \lambda_{i,j}\right) = \mathrm{e}^{t\lambda_{i,i}}, \qquad t \in \mathbb{R}_+.
\end{aligned}$$

cf. (1.5.8) in Chap. 1. In addition we have

$$\mathbb{E}[\tau_i] = \sum_{l\neq i} \lambda_{i,l} \int_0^\infty t \exp\left(-t \sum_{l\neq i} \lambda_{i,l}\right) dt = \frac{1}{\sum_{l\neq i} \lambda_{i,l}} = -\frac{1}{\lambda_{i,i}},$$

and the times $(\tau_k)_{k\in\mathbb{S}}$ spent in each state $k \in \mathbb{S}$ form a sequence of independent random variables.

Examples

(i) *Two-state* continuous-time Markov chain.
For the two-state continuous-time Markov chain with generator

$$Q = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix},$$

the mean time spent at state 0 is $1/\alpha$, whereas the mean time spent at state 1 is $1/\beta$. We will come back to this example in more detail in the following Sect. 9.5.

(ii) *Poisson process*.
The generator of the Poisson process is given by $\lambda_{i,j} = \mathbb{1}_{\{j=i+1\}}\lambda$, $i \neq j$, i.e.

$$Q = [\ \lambda_{i,j}\ ]_{i,j\in\mathbb{N}} = \begin{bmatrix} -\lambda & \lambda & 0 & \cdots \\ 0 & -\lambda & \lambda & \cdots \\ 0 & 0 & -\lambda & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

From the relation $P(h) = \mathrm{I_d} + hQ + o(h)$ we recover the infinitesimal transition probabilities of the Poisson process as

$$\mathbb{P}(N_{t+h} - N_t = 1) = \mathbb{P}(N_{t+h} = i + 1 \mid N_t = i) \simeq \lambda h,$$

$h \searrow 0$, $i \in \mathbb{N}$, and

$$\mathbb{P}(N_{t+h} - N_t = 0) = \mathbb{P}(N_{t+h} = i \mid N_t = i) = 1 - \lambda h + o(h),$$

$h \searrow 0$, $i \in \mathbb{N}$.

(iii) *Pure birth process.*
The generator of the pure birth process on $\mathbb{N} = \{0, 1, 2, \ldots\}$ is

$$Q = [\ \lambda_{i,j}\ ]_{i,j\in\mathbb{N}} = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots \\ \\ 0 & -\lambda_1 & \lambda_1 & \cdots \\ \\ 0 & 0 & -\lambda_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

in which the rate λ_i is (possibly) state-dependent. From the relation

$$P(h) = \mathrm{I_d} + hQ + o(h), \qquad h \searrow 0,$$

i.e.

$$P(h) = \begin{bmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} + h \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots \\ \\ 0 & -\lambda_1 & \lambda_1 & \cdots \\ \\ 0 & 0 & -\lambda_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} + o(h), \quad h \searrow 0,$$

we recover the infinitesimal transition probabilities of the pure birth process as

$$\mathbb{P}(X_{t+h} - X_t = 1 \mid X_t = i) = \mathbb{P}(X_{t+h} = i + 1 \mid X_t = i) \simeq \lambda_i h,$$

$h \searrow 0$, $i \in \mathbb{N}$, and

$$\mathbb{P}(X_{t+h} - X_t = 0 \mid X_t = i) = \mathbb{P}(X_{t+h} = i \mid X_t = i) = 1 - \lambda_i h + o(h),$$

$h \searrow 0$, $i \in \mathbb{N}$.

(iv) *Pure death process.*
The generator of the pure death process on $-\mathbb{N} = \{\ldots, -2, -1, 0\}$ is

$$Q = [\ \lambda_{i,j}\ ]_{i,j\leq 0} = \begin{bmatrix} \cdots & 0 & \mu_0 & -\mu_0 \\ \\ \cdots & \mu_1 & -\mu_1 & 0 \\ \\ \cdots & -\mu_2 & 0 & 0 \\ \iddots & \vdots & \vdots & \vdots \end{bmatrix}.$$

From the relation

$$P(h) = \begin{bmatrix} \cdots & 0 & 0 & 1 \\ \cdots & 0 & 1 & 0 \\ \cdots & 1 & 0 & 0 \\ \iddots & \vdots & \vdots & \vdots \end{bmatrix} + h \begin{bmatrix} \cdots & 0 & \mu_0 & -\mu_0 \\ \cdots & \mu_1 & -\mu_1 & 0 \\ \cdots & -\mu_2 & 0 & 0 \\ \iddots & \vdots & \vdots & \vdots \end{bmatrix} + o(h), \quad h \searrow 0,$$

we recover the infinitesimal transition probabilities

$$\mathbb{P}(X_{t+h} - X_t = -1 \mid X_t = i) = \mathbb{P}(X_{t+h} = i - 1 \mid X_t = i) \simeq \mu_i h, \quad h \searrow 0,$$

$i \in \mathbb{S}$, and

$$\mathbb{P}(X_{t+h} = i \mid X_t = i) = \mathbb{P}(X_{t+h} - X_t = 0 \mid X_t = i) = 1 - \mu_i h + o(h), \quad h \searrow 0,$$

$i \in \mathbb{S}$, of the pure death process.

(v) *Birth and death process* on $\{0, 1, \ldots, N\}$.
By (9.2.5a)–(9.2.5b) and (9.4.7), the generator of the birth and death process on $\{0, 1, \ldots, N\}$ is

$$[\, \lambda_{i,j} \,]_{0 \le i,j \le N} = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots\cdots & \cdots & 0 & 0 \\ \mu_1 & -\lambda_1 - \mu_1 & \lambda_1 & 0 & \cdots\cdots & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots \ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots \ \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots \ \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \ \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots \ 0 & \mu_{N-1} & -\lambda_{N-1} - \mu_{N-1} & \lambda_{N-1} \\ 0 & 0 & 0 & 0 & \cdots \ 0 & 0 & \mu_N & -\mu_N \end{bmatrix},$$

with $\mu_0 = \lambda_N = 0$.

From the Relation (9.4.6) we have

$$P(h) = \mathrm{I_d} + hQ + o(h), \qquad h \searrow 0,$$

and we recover the infinitesimal transition probabilities

$$\mathbb{P}(X_{t+h} - X_t = 1 \mid X_t = i) \simeq \lambda_i h, \qquad h \searrow 0, \quad i = 0, 1, \ldots, N,$$

and

$$\mathbb{P}(X_{t+h} - X_t = -1 \mid X_t = i) \simeq \mu_i h, \qquad h \searrow 0, \quad i = 0, 1, \ldots, N,$$

and

$$\mathbb{P}(X_{t+h} - X_t = 0 \mid X_t = i) = 1 - (\lambda_i + \mu_i)h + o(h), \qquad h \searrow 0, \quad i = 0, 1, \ldots, N,$$

of the birth and death process on $\{0, 1, \ldots, N\}$, with $\mu_0 = \lambda_N = 0$.
Recall that the time τ_i spent in state (i) is an exponentially distributed random variable with parameter $\lambda_i + \mu_i$ and we have

$$P_{i,i}(t) \geq \mathbb{P}(\tau_i > t) = \mathrm{e}^{-t(\lambda_i+\mu_i)}, \qquad t \in \mathbb{R}_+,$$

and

$$\mathbb{E}[\tau_i] = \frac{1}{\lambda_i + \mu_i}.$$

In the case of a pure birth process we find

$$P_{i,i}(t) = \mathbb{P}(\tau_i > t) = \mathrm{e}^{-t\lambda_i}, \qquad t \in \mathbb{R}_+,$$

and similarly for a pure death process. This allows us in particular to compute the diagonal entries in the matrix exponential $P(t) = \exp(tQ)$, $t \in \mathbb{R}_+$.
When $\mathbb{S} = \{0, 1, \ldots, N\}$ with $\lambda_i = \lambda$ and $\mu_i = \mu$, $i = 1, 2, \ldots, N-1$, and $\lambda_0 = \mu_N = 0$, the above birth and death process becomes a continuous-time analog of the discrete-time gambling process.

9.5 The Two-State Continuous-Time Markov Chain

In this section we consider a continuous-time Markov process with state space $\mathbb{S} = \{0, 1\}$, in the same way as in Sect. 4.5 which covered the two-state Markov chain in discrete time.

In this case the infinitesimal generator Q of $(X_t)_{t\in\mathbb{R}_+}$ has the form

$$Q = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}, \tag{9.5.1}$$

with $\alpha, \beta \geq 0$. The forward Kolmogorov equation (9.4.3) reads

$$P'(t) = P(t) \times \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}, \qquad t > 0, \tag{9.5.2}$$

i.e.

$$\begin{bmatrix} P'_{0,0}(t) & P'_{0,1}(t) \\ P'_{1,0}(t) & P'_{1,1}(t) \end{bmatrix} = \begin{bmatrix} P_{0,0}(t) & P_{0,1}(t) \\ P_{1,0}(t) & P_{1,1}(t) \end{bmatrix} \times \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}, \qquad t > 0,$$

or

$$\begin{cases} P'_{0,0}(t) = -\alpha P_{0,0}(t) + \beta P_{0,1}(t), & P'_{0,1}(t) = \alpha P_{0,0}(t) - \beta P_{0,1}(t), \\ \\ P'_{1,0}(t) = -\alpha P_{1,0}(t) + \beta P_{1,1}(t), & P'_{1,1}(t) = \alpha P_{1,0}(t) - \beta P_{1,1}(t), \end{cases}$$

$t > 0$, which is a system of four differential equations with initial condition

$$P(0) = \begin{bmatrix} P_{0,0}(0) & P_{0,1}(0) \\ \\ P_{1,0}(0) & P_{1,1}(0) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \\ 0 & 1 \end{bmatrix} = \mathrm{I_d}.$$

The solution of the forward Kolmogorov equation (9.5.2) is given by the matrix exponential

$$P(t) = P(0)\exp(tQ) = \exp(tQ) = \exp\left(t \begin{bmatrix} -\alpha & \alpha \\ \\ \beta & -\beta \end{bmatrix} \right),$$

which is computed in the next Proposition 9.6.

Proposition 9.6 *The solution $P(t)$ of the forward Kolmogorov equation* (9.5.2) *is given by*

$$P(t) = \begin{bmatrix} P_{0,0}(t) & P_{0,1}(t) \\ \\ P_{1,0}(t) & P_{1,1}(t) \end{bmatrix} \tag{9.5.3}$$
$$= \begin{bmatrix} \dfrac{\beta}{\alpha+\beta} + \dfrac{\alpha}{\alpha+\beta}\mathrm{e}^{-t(\alpha+\beta)} & \dfrac{\alpha}{\alpha+\beta} - \dfrac{\alpha}{\alpha+\beta}\mathrm{e}^{-t(\alpha+\beta)} \\ \\ \dfrac{\beta}{\alpha+\beta} - \dfrac{\beta}{\alpha+\beta}\mathrm{e}^{-t(\alpha+\beta)} & \dfrac{\alpha}{\alpha+\beta} + \dfrac{\beta}{\alpha+\beta}\mathrm{e}^{-t(\alpha+\beta)} \end{bmatrix},$$

$t \in \mathbb{R}_+$.

Proof We will compute the matrix exponential e^{tQ} by the diagonalization technique. The matrix Q has two eigenvectors

$$\begin{bmatrix} 1 \\ \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -\alpha \\ \\ \beta \end{bmatrix},$$

with respective eigenvalues $\lambda_1 = 0$ and $\lambda_2 = -\alpha - \beta$. Hence it can be put in the *diagonal form*

$$Q = M \times D \times M^{-1}$$

as follows:

$$Q = \begin{bmatrix} 1 & -\alpha \\ 1 & \beta \end{bmatrix} \times \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \times \begin{bmatrix} \dfrac{\beta}{\alpha+\beta} & \dfrac{\alpha}{\alpha+\beta} \\ -\dfrac{1}{\alpha+\beta} & \dfrac{1}{\alpha+\beta} \end{bmatrix}.$$

Consequently we have

$$\begin{aligned}
P(t) &= \exp(tQ) = \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n = \sum_{n=0}^{\infty} \frac{t^n}{n!} (M \times D \times M^{-1})^n \\
&= \sum_{n=0}^{\infty} \frac{t^n}{n!} M \times D^n \times M^{-1} = M \times \left(\sum_{n=0}^{\infty} \frac{t^n}{n!} D^n \right) \times M^{-1} \\
&= M \times \exp(tD) \times M^{-1} \\
&= \begin{bmatrix} 1 & -\alpha \\ 1 & \beta \end{bmatrix} \times \begin{bmatrix} e^{t\lambda_1} & 0 \\ 0 & e^{t\lambda_2} \end{bmatrix} \times \begin{bmatrix} \dfrac{\beta}{\alpha+\beta} & \dfrac{\alpha}{\alpha+\beta} \\ -\dfrac{1}{\alpha+\beta} & \dfrac{1}{\alpha+\beta} \end{bmatrix} \\
&= \begin{bmatrix} 1 & -\alpha \\ 1 & \beta \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ 0 & e^{-t(\alpha+\beta)} \end{bmatrix} \times \begin{bmatrix} \dfrac{\beta}{\alpha+\beta} & \dfrac{\alpha}{\alpha+\beta} \\ -\dfrac{1}{\alpha+\beta} & \dfrac{1}{\alpha+\beta} \end{bmatrix} \\
&= \frac{1}{\alpha+\beta} \begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix} + \frac{e^{-t(\alpha+\beta)}}{\alpha+\beta} \begin{bmatrix} \alpha & -\alpha \\ -\beta & \beta \end{bmatrix} \\
&= \begin{bmatrix} \dfrac{\beta}{\alpha+\beta} + \dfrac{\alpha}{\alpha+\beta} e^{-t(\alpha+\beta)} & \dfrac{\alpha}{\alpha+\beta} - \dfrac{\alpha}{\alpha+\beta} e^{-t(\alpha+\beta)} \\ \dfrac{\beta}{\alpha+\beta} - \dfrac{\beta}{\alpha+\beta} e^{-t(\alpha+\beta)} & \dfrac{\alpha}{\alpha+\beta} + \dfrac{\beta}{\alpha+\beta} e^{-t(\alpha+\beta)} \end{bmatrix},
\end{aligned}$$

$t > 0$. □

From Proposition 9.6 we obtain the probabilities

$$
\begin{cases}
\mathbb{P}(X_h = 0 \mid X_0 = 0) = \dfrac{\beta + \alpha \mathrm{e}^{-(\alpha+\beta)h}}{\alpha+\beta}, \\[2ex]
\mathbb{P}(X_h = 1 \mid X_0 = 0) = \dfrac{\alpha}{\alpha+\beta}(1 - \mathrm{e}^{-(\alpha+\beta)h}), \\[2ex]
\mathbb{P}(X_h = 0 \mid X_0 = 1) = \dfrac{\beta}{\alpha+\beta}(1 - e^{-h(\alpha+\beta)h}), \\[2ex]
\mathbb{P}(X_h = 1 \mid X_0 = 1) = \dfrac{\alpha + \beta \mathrm{e}^{-(\alpha+\beta)h}}{\alpha+\beta}, \qquad h \in \mathbb{R}_+.
\end{cases}
$$

In other words, (9.5.3) can be rewritten as

$$
P(h) = \begin{bmatrix} 1 - \dfrac{\alpha}{\alpha+\beta}(1 - \mathrm{e}^{-(\alpha+\beta)h}) & \dfrac{\alpha}{\alpha+\beta}(1 - \mathrm{e}^{-(\alpha+\beta)h}) \\[3ex] \dfrac{\beta}{\alpha+\beta}(1 - \mathrm{e}^{-(\alpha+\beta)h}) & 1 - \dfrac{\beta}{\alpha+\beta}(1 - \mathrm{e}^{-(\alpha+\beta)h}) \end{bmatrix}, \qquad h > 0, \tag{9.5.4}
$$

hence, since

$$
1 - \mathrm{e}^{-(\alpha+\beta)h} \simeq h(\alpha+\beta), \qquad h \searrow 0,
$$

the expression (9.5.4) above recovers (9.4.8) as $h \searrow 0$, i.e. we have

$$
P(h) = \begin{bmatrix} 1 - h\alpha & h\alpha \\[2ex] h\beta & 1 - h\beta \end{bmatrix} + o(h) = \mathrm{I_d} + hQ + o(h), \qquad h \searrow 0,
$$

which recovers (9.5.1).

From these expressions we can determine the large time behavior of the continuous-time Markov chain by taking limits as t goes to infinity:

$$
\lim_{t\to\infty} P(t) = \lim_{t\to\infty} \begin{bmatrix} \mathbb{P}(X_t = 0 \mid X_0 = 0) & \mathbb{P}(X_t = 1 \mid X_0 = 0) \\[2ex] \mathbb{P}(X_t = 0 \mid X_0 = 1) & \mathbb{P}(X_t = 1 \mid X_0 = 1) \end{bmatrix} = \frac{1}{\alpha+\beta} \begin{bmatrix} \beta & \alpha \\[2ex] \beta & \alpha \end{bmatrix},
$$

whenever $\alpha > 0$ or $\beta > 0$, whereas if $\alpha = \beta = 0$ we simply have

$$
P(t) = \begin{bmatrix} 1 & 0 \\[2ex] 0 & 1 \end{bmatrix} = \mathrm{I_d}, \qquad t \in \mathbb{R}_+,
$$

and the chain is constant. Note that in continuous time the limiting distribution of the two-state chain always exists (unlike in the discrete-time case), and convergence will be faster when $\alpha + \beta$ is larger. Hence we have

$$\lim_{t\to\infty} \mathbb{P}(X_t = 1 \mid X_0 = 0) = \lim_{t\to\infty} \mathbb{P}(X_t = 1 \mid X_0 = 1) = \frac{\alpha}{\alpha+\beta}$$

and

$$\lim_{t\to\infty} \mathbb{P}(X_t = 0 \mid X_0 = 0) = \lim_{t\to\infty} \mathbb{P}(X_t = 0 \mid X_0 = 1) = \frac{\beta}{\alpha+\beta}$$

and

$$\pi = [\pi_0, \pi_1] = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right) = \left(\frac{1/\alpha}{1/\alpha+1/\beta}, \frac{1/\beta}{1/\alpha+1/\beta}\right)$$

appears as a *limiting distribution* as t goes to infinity, provided that $(\alpha, \beta) \neq (0, 0)$. This means that whatever the starting point X_0, the probability of being at (1) after a "large" time is close to $\alpha/(\alpha+\beta)$, while the probability of being at (0) is close to $\beta/(\alpha+\beta)$.

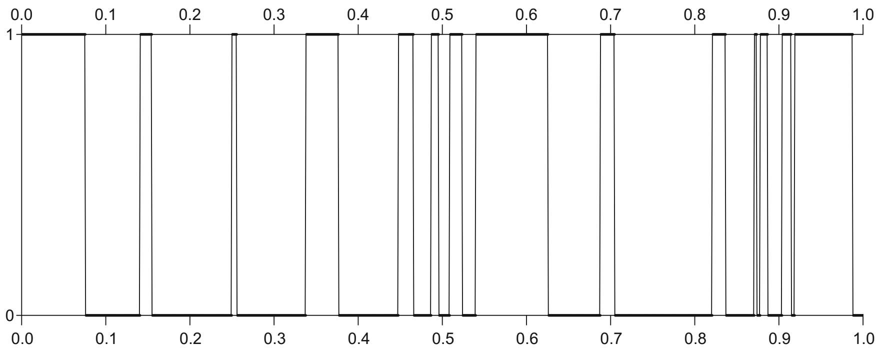


Fig. 9.5 Sample path of a continuous-time two-state chain with $\alpha = 20$ and $\beta = 40$

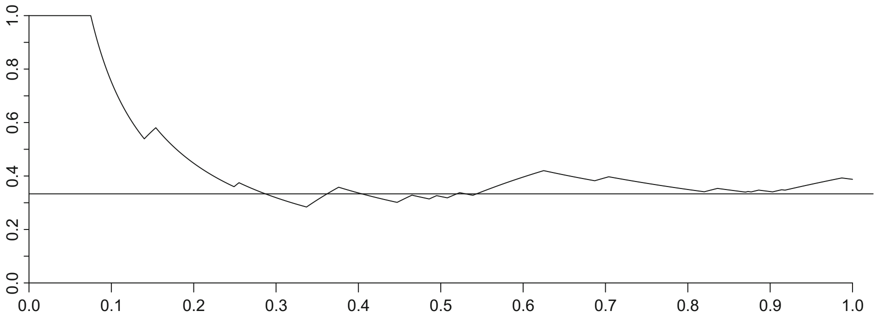


Fig. 9.6 The proportion of process values in the state 1 converges to $1/3 = \alpha/(\alpha+\beta)$

Next we consider a simulation of the two-state continuous Markov chain with infinitesimal generator

$$Q = \begin{bmatrix} -20 & 20 \\ \\ 40 & -40 \end{bmatrix},$$

i.e. $\alpha = 20$ and $\beta = 40$. Figure 9.5 represents a sample path $(x_t)_{t\in\mathbb{R}_+}$ of the continuous-time chain, while Fig. 9.6 represents the sample average

$$y_t = \frac{1}{t}\int_0^t x_s ds, \qquad t \in [0, 1],$$

which counts the proportion of values of the chain in the state 1. This proportion is found to converge to $\alpha/(\alpha+\beta) = 1/3$, as a consequence of the Ergodic Theorem in continuous time, see (9.6.4) below.

9.6 Limiting and Stationary Distributions

A probability distribution $\pi = (\pi_i)_{i\in\mathbb{S}}$ is said to be *stationary* for $P(t)$ if it satisfies the equation

$$\pi P(t) = \pi, \qquad t \in \mathbb{R}_+.$$

In the next proposition we show that the notion of stationary distribution admits a simpler characterization.

Proposition 9.7 *The probability distribution $\pi = (\pi_i)_{i\in\mathbb{S}}$ is stationary if and only if it satisfies the equation*

$$\pi Q = 0.$$

Proof Assuming that $\pi Q = 0$, we have

$$\begin{aligned} \pi P(t) &= \pi \exp(tQ) = \pi \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n \\ &= \pi \left(\mathrm{I_d} + \sum_{n=1}^{\infty} \frac{t^n}{n!} Q^n \right) \\ &= \pi + \sum_{n=1}^{\infty} \frac{t^n}{n!} \pi Q^n \\ &= \pi, \end{aligned}$$

since $\pi Q^n = \pi Q Q^{n-1} = 0$, $n \geq 1$. Conversely, the relation $\pi = \pi P(t)$ shows, by differentiation at $t = 0$, that

$$0 = \pi P'(0) = \pi Q.$$

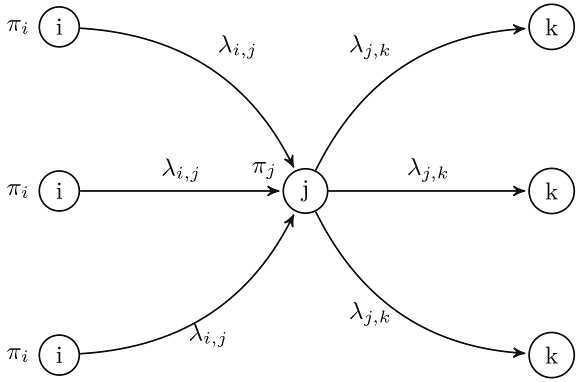


Fig. 9.7 Global balance condition (continuous time)

□

When $\mathbb{S} = \{0, 1, \ldots, N\}$ and the generator Q has the form (9.4.1) the relation $\pi Q = 0$ reads

$$\pi_0 \lambda_{0,j} + \pi_1 \lambda_{1,j} + \cdots + \pi_N \lambda_{N,j} = 0, \qquad j = 0, 1, \ldots, N,$$

i.e.,

$$\sum_{\substack{i=0 \\ i \neq j}}^{N} \pi_i \lambda_{j,i} = -\pi_j \lambda_{j,j}, \qquad j = 0, 1, \ldots, N,$$

hence from (9.4.2) we find the balance condition

$$\sum_{\substack{i=0 \\ i \neq j}}^{N} \pi_i \lambda_{i,j} = \sum_{\substack{k=0 \\ k \neq j}}^{N} \pi_j \lambda_{j,k},$$

which can be interpreted by stating the equality between incoming and outgoing "flows" into and from state j are equal for all $j = 0, 1, \ldots, N$ (Fig. 9.7).

Next is the continuous-time analog of Proposition 7.7 in Sect. 7.2.

Proposition 9.8 *Consider a continuous-time Markov chain $(X_t)_{t \in \mathbb{R}_+}$ on a* finite state space, *which admits a limiting distribution given by*

$$\pi_j := \lim_{t \to \infty} \mathbb{P}(X_t = j \mid X_0 = i) = \lim_{t \to \infty} P_{i,j}(t), \qquad j \in \mathbb{S}, \tag{9.6.1}$$

independent of the initial state $i \in \mathbb{S}$. Then we have

$$\pi Q = 0, \tag{9.6.2}$$

i.e. π is a stationary distribution for the chain $(X_t)_{t \in \mathbb{R}_+}$.

Proof Taking $\mathbb{S} = \{0, 1, \ldots, N\}$ we note that by (9.6.1) since the limiting distribution is independent of the initial state it satisfies

$$\lim_{t\to\infty} P(t) = \begin{bmatrix} \lim\limits_{t\to\infty} P_{0,0}(t) & \cdots & \lim\limits_{t\to\infty} P_{0,N}(t) \\ \vdots & \ddots & \vdots \\ \lim\limits_{t\to\infty} P_{N,0}(t) & \cdots & \lim\limits_{t\to\infty} P_{N,N}(t) \end{bmatrix}$$

$$= \begin{bmatrix} \pi_0 & \pi_1 & \cdots & \pi_N \\ \vdots & \vdots & \ddots & \vdots \\ \pi_0 & \pi_1 & \cdots & \pi_N \end{bmatrix} = \begin{bmatrix} \pi \\ \vdots \\ \pi \end{bmatrix},$$

where π is the row vector

$$\pi = [\pi_0, \pi_1, \ldots, \pi_n].$$

By the forward Kolmogorov equation (9.4.3) and (9.6.1) we find that the limit of $P'(t)$ exists as $t \to \infty$ since

$$\lim_{t\to\infty} P'(t) = \lim_{t\to\infty} P(t)Q = \begin{bmatrix} \pi \\ \vdots \\ \pi \end{bmatrix} Q.$$

On the other hand, since $P'(t)$ converges as $t \to \infty$ we should have

$$\lim_{t\to\infty} P'(t) = 0,$$

for the integral

$$P(t) = P(0) + \int_0^t P'(s)ds \tag{9.6.3}$$

to converge as $t \to \infty$. This shows that

$$\begin{bmatrix} \pi Q \\ \vdots \\ \pi Q \end{bmatrix} = \begin{bmatrix} \pi \\ \vdots \\ \pi \end{bmatrix} Q = 0$$

by (9.4.3), hence we have $\pi Q = 0$, or $\sum\limits_{i\in\mathbb{S}} \pi_i \lambda_{i,j} = 0$, $j \in \mathbb{S}$. □

Equation (9.6.2) is actually equivalent to

$$\pi = \pi(\mathrm{I_d} + hQ), \qquad h > 0,$$

which yields the stationary distribution of a discrete-time Markov chain with transition matrix $P(h) = \mathrm{I_d} + hQ + o(h)$ on "small" discrete intervals of length $h \searrow 0$.

Proposition 9.8 admits the following partial converse. More generally it can be shown, cf. Corollary 6.2 of [Lal], that an irreducible continuous-time Markov chain admits its stationary distribution π as limiting distribution, similarly to the discrete-time Theorem 7.8, cf. also Proposition 1.1 p. 9 of [Sig] in the positive recurrent case.

Proposition 9.9 *Consider an irreducible, continuous-time Markov chain* $(X_t)_{t\in\mathbb{R}_+}$ *on a* finite state space*, with stationary distribution* π*, i.e.*

$$\pi Q = 0,$$

and assume that the matrix Q *is diagonalizable. Then* $(X_t)_{t\in\mathbb{R}_+}$ *admits* π *as limiting distribution, i.e.*

$$\lim_{t\to\infty} P_{i,j}(t) = \lim_{t\to\infty} \mathbb{P}(X_t = j \mid X_0 = i) = \pi_j, \quad j \in \mathbb{S},$$

which is independent of the initial state $i \in \mathbb{S}$.

Proof By e.g. Theorem 2.1 in Chap. 10 of [KT81], since the chain is irreducible, $\lambda_1 = 0$ is an eigenvalue of Q with multiplicity one and eigenvector $u^{(1)} = (1, 1, \ldots, 1)$. In addition, the remaining eigenvectors $u^{(2)}, \ldots, u^{(n)} \in \mathbb{R}^n$ with eigenvalues $\lambda_2, \ldots, \lambda_n$ are orthogonal to the invariant (or stationary) distribution $[\pi_1, \pi_2, \ldots, \pi_n]$ of $(X_t)_{t\in\mathbb{R}_+}$ as we have $\lambda_k \langle u^{(k)}, \pi\rangle_{\mathbb{R}^n} = \langle Q u^{(k)}, \pi\rangle_{\mathbb{R}^n} = \langle u^{(k)}, Q^T\pi\rangle_{\mathbb{R}^n} = \pi^T Q u^{(k)} = 0$, $k = 2, \ldots, n$. Hence by diagonalization we have $Q = M^{-1}DM$ where the matrices M and M^{-1} take the form

$$M = \begin{bmatrix} \pi_1 & \cdots & \pi_n \\ M_{2,1} & \cdots & M_{2,n} \\ \vdots & \ddots & \vdots \\ M_{n,1} & \cdots & M_{n,n} \end{bmatrix} \quad \text{and} \quad M^{-1} = \begin{bmatrix} 1 & u_1^{(2)} & \cdots & u_1^{(n)} \\ 1 & u_2^{(2)} & \cdots & u_2^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & u_n^{(2)} & \cdots & u_n^{(n)} \end{bmatrix},$$

and D is the diagonal matrix $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. This allows us to compute the transition probabilities of $(X_t)_{t\in\mathbb{R}_+}$ as

$$P_{i,j}(t) = \mathbb{P}(X_t = j \mid X_0 = i) = [\exp(tQ)]_{i,j} = [M^{-1}\exp(tD)M]_{i,j}$$

where $\exp(tD)$ is the diagonal matrix

$$\exp(tD) = \mathrm{diag}(1, \mathrm{e}^{t\lambda_2}, \ldots, \mathrm{e}^{-t\lambda_n}),$$

and the eigenvalues $\lambda_2, \ldots, \lambda_n$ have to be strictly negative, hence we have

$$
\begin{aligned}
\lim_{t\to\infty} \; [\mathbb{P}(X_t = j \mid X_0 = i)]_{1\le i,j\le n} &= \lim_{t\to\infty} [M^{-1}\exp(tD)M]_{1\le i,j\le n} \\
&= \begin{bmatrix} 1 & u_1^{(2)} & \cdots & u_1^{(n)} \\ 1 & u_2^{(2)} & \cdots & u_2^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & u_n^{(2)} & \cdots & u_n^{(n)} \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \pi_1 & \cdots & \pi_n \\ M_{2,1} & \cdots & M_{2,n} \\ \vdots & \ddots & \vdots \\ M_{n,1} & \cdots & M_{n,n} \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \pi_1 & \cdots & \pi_n \\ M_{2,1} & \cdots & M_{2,n} \\ \vdots & \ddots & \vdots \\ M_{n,1} & \cdots & M_{n,n} \end{bmatrix} \\
&= \begin{bmatrix} \pi_1 & \cdots & \pi_n \\ \pi_1 & \cdots & \pi_n \\ \vdots & \ddots & \vdots \\ \pi_1 & \cdots & \pi_n \end{bmatrix}.
\end{aligned}
$$

□

The discrete-time Ergodic Theorem 7.12 also admits a continuous-time version with a similar proof, stating that if the chain $(X_t)_{t\in\mathbb{R}_+}$ is *irreducible* then the sample average of the number of visits to state (i) converges almost surely to π_i, i.e.,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \mathbb{1}_{\{X_t=i\}} dt = \pi_i, \qquad i \in \mathbb{S}. \tag{9.6.4}$$

Examples

(i) *Two-state Markov chain.*
Consider the two-state Markov chain with infinitesimal generator

$$Q = \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix},$$

the limiting distribution solves $\pi Q = 0$, i.e.

$$\begin{cases} 0 = -\alpha\pi_0 + \beta\pi_1 \\ 0 = \alpha\pi_0 - \beta\pi_1, \end{cases}$$

with $\pi_0 + \pi_1 = 1$, i.e.

$$\pi = [\pi_0, \pi_1] = \left[\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right]. \tag{9.6.5}$$

(ii) *Birth and death process* on $\mathbb{N}$.
Next, consider the birth and death process on $\mathbb{N}$ with infinitesimal generator

$$Q = [\,\lambda_{i,j}\,]_{i,j\in\mathbb{N}} = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 & \cdots \\ \mu_1 & -\lambda_1-\mu_1 & \lambda_1 & 0 & 0 & \cdots \\ 0 & \mu_2 & -\lambda_2-\mu_2 & \lambda_2 & 0 & \cdots \\ 0 & 0 & \mu_3 & -\lambda_3-\mu_3 & \lambda_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix},$$

the stationary distribution solves $\pi Q = 0$, i.e.

$$\begin{cases} 0 = -\lambda_0\pi_0 + \mu_1\pi_1 \\ 0 = \lambda_0\pi_0 - (\lambda_1+\mu_1)\pi_1 + \mu_2\pi_2 \\ 0 = \lambda_1\pi_1 - (\lambda_2+\mu_2)\pi_2 + \mu_3\pi_3 \\ \quad\vdots \\ 0 = \lambda_{j-1}\pi_{j-1} - (\lambda_j+\mu_j)\pi_j + \mu_{j+1}\pi_{j+1}, \\ \quad\vdots \end{cases}$$

i.e.

$$\begin{cases} \pi_1 = \dfrac{\lambda_0}{\mu_1}\pi_0 \\ \pi_2 = -\dfrac{\lambda_0}{\mu_2}\pi_0 + \dfrac{\lambda_1+\mu_1}{\mu_2}\pi_1 = -\dfrac{\lambda_0}{\mu_2}\pi_0 + \dfrac{\lambda_1+\mu_1}{\mu_2}\dfrac{\lambda_0}{\mu_1}\pi_0 = \dfrac{\lambda_1}{\mu_2}\dfrac{\lambda_0}{\mu_1}\pi_0 \\ \pi_3 = -\dfrac{\lambda_1}{\mu_3}\pi_1 + \dfrac{\lambda_2+\mu_2}{\mu_3}\pi_2 = -\dfrac{\lambda_1}{\mu_3}\dfrac{\lambda_0}{\mu_1}\pi_0 + \dfrac{\lambda_2+\mu_2}{\mu_3}\dfrac{\lambda_1}{\mu_2}\dfrac{\lambda_0}{\mu_1}\pi_0 = \dfrac{\lambda_2}{\mu_3}\dfrac{\lambda_1}{\mu_2}\dfrac{\lambda_0}{\mu_1}\pi_0 \\ \quad\vdots \\ \pi_{j+1} = \dfrac{\lambda_j\cdots\lambda_0}{\mu_{j+1}\cdots\mu_1}\pi_0. \\ \quad\vdots \end{cases}$$

Using the convention

$$\lambda_{j-1}\cdots\lambda_0 = \prod_{l=0}^{j-1}\lambda_l = 1 \quad \text{and} \quad \mu_j\cdots\mu_1 = \prod_{l=1}^{j}\mu_l = 1$$

in the case $j = 0$, we have

$$1 = \sum_{j=0}^{\infty} \pi_j = \pi_0 + \pi_0 \sum_{j=0}^{\infty} \frac{\lambda_j \cdots \lambda_0}{\mu_{j+1} \cdots \mu_1} = \pi_0 + \pi_0 \sum_{j=1}^{\infty} \frac{\lambda_{j-1} \cdots \lambda_0}{\mu_j \cdots \mu_1}$$
$$= \pi_0 \sum_{i=0}^{\infty} \frac{\lambda_0 \cdots \lambda_{i-1}}{\mu_1 \cdots \mu_i},$$

hence

$$\pi_0 = \frac{1}{\displaystyle\sum_{i=0}^{\infty} \frac{\lambda_0 \cdots \lambda_{i-1}}{\mu_1 \cdots \mu_i}},$$

and

$$\pi_j = \frac{\lambda_0 \cdots \lambda_{j-1}}{\displaystyle\mu_1 \cdots \mu_j \sum_{i=0}^{\infty} \frac{\lambda_0 \cdots \lambda_{i-1}}{\mu_1 \cdots \mu_i}}, \qquad j \in \mathbb{N}.$$

When $\lambda_i = \lambda$, $i \in \mathbb{N}$, and $\mu_i = \mu$, $i \geq 1$, this gives

$$\pi_j = \frac{\lambda^j}{\displaystyle\mu^j \sum_{i=0}^{\infty} (\lambda/\mu)^i} = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^j, \qquad j \in \mathbb{N}.$$

provided that $\lambda < \mu$, hence in this case the stationary distribution is the geometric distribution with parameter λ/μ, otherwise the stationary distribution does not exist.

(iii) *Birth and death process* on $\mathbb{S} = \{0, 1, \ldots, N\}$.
The birth and death process on $\mathbb{S} = \{0, 1, \ldots, N\}$ has the infinitesimal generator

$$Q = [\,\lambda_{i,j}\,]_{0 \le i,j \le N} = \begin{bmatrix}
-\lambda_0 & \lambda_0 & 0 & \cdots & \cdots & 0 & 0 & 0 & 0 \\
\mu_1 & -\lambda_1 - \mu_1 & \lambda_1 & \cdots & \cdots & 0 & 0 & 0 & 0 \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots & 0 & 0 & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\
0 & 0 & 0 & \cdots & \cdots & 0 & \mu_{N-1} & -\lambda_{N-1} - \mu_{N-1} & \lambda_{N-1} \\
0 & 0 & 0 & \cdots & \cdots & 0 & 0 & \mu_N & -\mu_N
\end{bmatrix},$$

we can apply (1.6.6) with $\lambda_j = 0$, $j \geq N$, and $\mu_j = 0$, $j \geq N + 1$, which yields

$$\pi_j = \frac{\lambda_{j-1} \cdots \lambda_0}{\displaystyle\mu_j \cdots \mu_1 \sum_{i=0}^{N} \frac{\lambda_{i-1} \cdots \lambda_0}{\mu_i \cdots \mu_1}}, \qquad j \in \{0, 1, \ldots, N\},$$

and coincides with (9.6.5) when $N = 1$.
When $\lambda_i = \lambda$, $i \in \mathbb{N}$, and $\mu_i = \mu$, $i \geq 1$, this gives

$$\pi_j = \frac{1 - \lambda/\mu}{1 - (\lambda/\mu)^{N+1}} \left(\frac{\lambda}{\mu}\right)^j, \qquad j = 0, 1, \ldots, N,$$

which is a truncated geometric distribution since $\pi_j = 0$ for all $j > N$ and any $\lambda, \mu > 0$.

9.7 The Discrete-Time Embedded Chain

Consider the sequence $(T_n)_{n\in\mathbb{N}}$ the sequence of jump times of the continuous-time Markov process $(X_t)_{t\in\mathbb{R}_+}$, defined recursively by $T_0 = 0$, then

$$T_1 = \inf\{t > 0 \ : \ X_t \neq X_0\},$$

and

$$T_{n+1} = \inf\{t > T_n \ : \ X_t \neq X_{T_n}\}, \qquad n \in \mathbb{N}.$$

The *embedded chain* of $(X_t)_{t\in\mathbb{R}_+}$ is the *discrete-time* Markov chain $(Z_n)_{n\in\mathbb{N}}$ defined by $Z_0 = X_0$ and

$$Z_n := X_{T_n}, \qquad n \geq 1.$$

The next Fig. 9.8 shows the graph of the embedded chain of a birth and death process.

The results of Chaps. 2–8 can now be applied to the discrete-time embedded chain. The next Fig. 9.9 represents the discrete-time embedded chain associated to the path of Fig. 9.8, in which we have $Z_0 = 0$, $Z_1 = 1$, $Z_2 = 2$, $Z_3 = 3$, $Z_4 = 4$, $Z_5 = 3$, $Z_6 = 4$, $Z_7 = 3, \ldots$

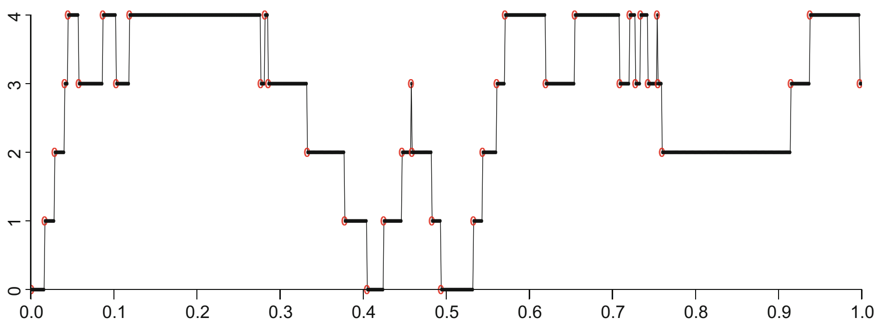


Fig. 9.8 Birth and death process $(X_t)_{t\in\mathbb{R}_+}$ with its embedded chain $(Z_n)_{n\in\mathbb{N}}$

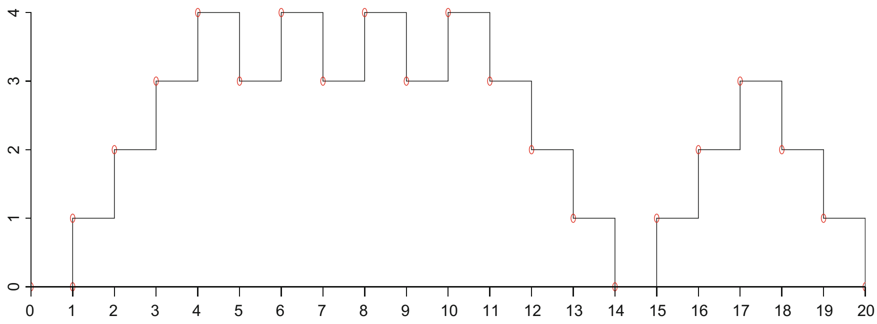


Fig. 9.9 Discrete-time embedded chain $(Z_n)_{n\in\mathbb{N}}$ based on the path of Fig. 9.8

For example, if $\lambda_0 > 0$ and $\mu_0 > 0$, the embedded chain of the two-state continuous-time Markov chain has the transition matrix

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{9.7.1}$$

which switches permanently between the states 0 and 1.

In case one of the states $\{0, 1\}$ is absorbing the transition matrix becomes

$$P = \begin{cases} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \ \lambda_0 = 0, \ \mu_1 > 0, \\ \\ \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \ \lambda_0 > 0, \ \mu_1 = 0, \\ \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \lambda_0 = 0, \ \mu_1 = 0. \end{cases}$$

Birth and Death Embedded Chain

More generally, consider now the birth and death process with infinitesimal generator

$$Q = [\,\lambda_{i,j}\,]_{0\le i,j\le N} = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots & \cdots & 0 & 0 & 0 & 0 \\ \mu_1 & -\lambda_1-\mu_1 & \lambda_1 & \cdots & \cdots & 0 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots & \vdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & \mu_{N-1} & -\lambda_{N-1}-\mu_{N-1} & \lambda_{N-1} \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & \mu_N & -\mu_N \end{bmatrix}.$$

Given that a transition occurs from state (i) in a "short" time interval $[t, t+h]$, the probability that the chain switches to state $\boxed{i+1}$ is given by

$$\begin{aligned}\mathbb{P}(X_{t+h} = i+1 \mid \ X_t = i \text{ and } X_{t+h} - X_t \neq 0) &= \frac{\mathbb{P}(X_{t+h} = i+1 \text{ and } X_t = i)}{\mathbb{P}(X_{t+h} - X_t \neq 0 \text{ and } X_t = i)} \\ &= \frac{\mathbb{P}(X_{t+h} = i+1 \text{ and } X_t = i)}{\mathbb{P}(X_{t+h} - X_t \neq 0 \mid X_t = i)\mathbb{P}(X_t = i)} \\ &= \frac{\mathbb{P}(X_{t+h} = i+1 \mid X_t = i)}{\mathbb{P}(X_{t+h} - X_t \neq 0 \mid X_t = i)} \\ &= \frac{\mathbb{P}(X_{t+h} - X_t = 1 \mid X_t = i)}{\mathbb{P}(X_{t+h} - X_t \neq 0 \mid X_t = i)} \\ &\simeq \frac{\lambda_i h}{\lambda_i h + \mu_i h} \\ &= \frac{\lambda_i}{\lambda_i + \mu_i}, \qquad h \searrow 0, \quad i \in \mathbb{S},\end{aligned}$$

where we applied (9.4.7), hence the transition matrix of the embedded chain satisfies

$$P_{i,i+1} = \lim_{h \searrow 0} \mathbb{P}(X_{t+h} = i+1 \mid X_t = i \text{ and } X_{t+h} - X_t \neq 0) = \frac{\lambda_i}{\lambda_i + \mu_i}, \quad i \in \mathbb{S}. \tag{9.7.2}$$

This result can also be obtained from (1.5.9) which states that

$$\mathbb{P}(\tau_{i,i+1} < \tau_{i,i-1}) = \frac{\lambda_i}{\lambda_i + \mu_i}. \tag{9.7.3}$$

Similarly the probability that a given transition occurs from (i) to $\boxed{i\text{-}1}$ is

$$\mathbb{P}(X_{t+h} = i-1 \mid X_t = i \text{ and } X_{t+h} - X_t \neq 0) = \frac{\mu_i}{\lambda_i + \mu_i}, \qquad h \searrow 0, \quad i \in \mathbb{S},$$

which can also be obtained from (1.5.9) which states that

$$\mathbb{P}(\tau_{i,i-1} < \tau_{i,i+1}) = \frac{\mu_i}{\lambda_i + \mu_i}.$$

Hence we have

$$P_{i,i-1} = \lim_{h \searrow 0} \mathbb{P}(X_{t+h} = i-1 \mid X_t = i \text{ and } X_{t+h} - X_t \neq 0) = \frac{\mu_i}{\lambda_i + \mu_i}, \quad i \in \mathbb{S},$$

and the embedded chain $(Z_n)_{n \in \mathbb{N}}$ has the transition matrix

$$P = [\ P_{i,j}\ ]_{i,j\in\mathbb{S}}$$

$$= \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \dfrac{\mu_1}{\lambda_1+\mu_1} & 0 & \dfrac{\lambda_1}{\lambda_1+\mu_1} & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots & \vdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \dfrac{\mu_{N-1}}{\lambda_{N-1}+\mu_{N-1}} & 0 & \dfrac{\lambda_{N-1}}{\lambda_{N-1}+\mu_{N-1}} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \end{bmatrix},$$

provided that $\lambda_0 > 0$ and $\mu_N > 0$. When $N = 1$, this coincides with (9.7.1). In case $\lambda_0 = \mu_N = 0$, states 0 and N are both absorbing since the birth rate starting from 0 and the death rate starting from N are both 0, hence the transition matrix of the embedded chain can be written as

$$P = [\ P_{i,j}\ ]_{0\le i,j\le N}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \dfrac{\mu_1}{\lambda_1+\mu_1} & 0 & \dfrac{\lambda_1}{\lambda_1+\mu_1} & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots & \vdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \dfrac{\mu_{N-1}}{\lambda_{N-1}+\mu_{N-1}} & 0 & \dfrac{\lambda_{N-1}}{\lambda_{N-1}+\mu_{N-1}} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \end{bmatrix},$$

which is the transition matrix of a gambling type process on $\{0, 1, \ldots, N\}$. When $N = 1$ this yields $P = \mathrm{I_d}$, which is consistent with the fact that a two-state Markov chain with two absorbing states is constant.

For example, for a continuous-time chain with infinitesimal generator

$$Q = [\ \lambda_{i,j}\ ]_{0\le i,j\le 4} = \begin{bmatrix} -10 & 10 & 0 & 0 & 0 \\ 10 & -20 & 10 & 0 & 0 \\ 0 & 10 & -30 & 20 & 0 \\ 0 & 0 & 10 & -40 & 30 \\ 0 & 0 & 0 & 20 & -20 \end{bmatrix},$$

the transition matrix of the embedded chain is

$$P = [\ P_{i,j}\ ]_{0\le i,j\le 4} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/3 & 0 & 2/3 & 0 \\ 0 & 0 & 1/4 & 0 & 3/4 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

In case the states (0) and (4) are absorbing, i.e.

$$Q = [\ \lambda_{i,j}\ ]_{0\le i,j\le 4} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 10 & -20 & 10 & 0 & 0 \\ 0 & 10 & -30 & 20 & 0 \\ 0 & 0 & 10 & -40 & 30 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

the transition matrix of the embedded chain becomes

$$P = [\ P_{i,j}\ ]_{0\le i,j\le 4} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/3 & 0 & 2/3 & 0 \\ 0 & 0 & 1/4 & 0 & 3/4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

More generally, by (9.4.7) and (9.4.2) we could also show that the embedded chain of a continuous-time Markov chain with generator Q of the form (9.4.1) has the transition matrix

$$P = [\ P_{i,j}\ ]_{i,j\in\mathbb{S}}$$

$$= \begin{bmatrix} 0 & -\frac{\lambda_{0,1}}{\lambda_{0,0}} & -\frac{\lambda_{0,2}}{\lambda_{0,0}} & \cdots\cdots & -\frac{\lambda_{0,N-1}}{\lambda_{0,0}} & -\frac{\lambda_{0,N}}{\lambda_{0,0}} \\ -\frac{\lambda_{1,0}}{\lambda_{1,1}} & 0 & -\frac{\lambda_{1,2}}{\lambda_{1,1}} & \cdots\cdots & -\frac{\lambda_{1,N-1}}{\lambda_{1,1}} & -\frac{\lambda_{1,N}}{\lambda_{1,1}} \\ \vdots & \ddots & \vdots & \vdots\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots\ \vdots & \ddots & \vdots \\ -\frac{\lambda_{N-1,0}}{\lambda_{N-1,N-1}} & -\frac{\lambda_{N-1,1}}{\lambda_{N-1,N-1}} & \cdots & \cdots\cdots & 0 & -\frac{\lambda_{N-1,N}}{\lambda_{N-1,N-1}} \\ -\frac{\lambda_{N,0}}{\lambda_{N,N}} & -\frac{\lambda_{N,1}}{\lambda_{N,N}} & \cdots & \cdots\cdots & -\frac{\lambda_{N,N-1}}{\lambda_{N,N}} & 0 \end{bmatrix},$$

provided that $\lambda_{i,i} > 0$, $i = 0, 1, \ldots, N$.

9.8 Mean Absorption Time and Probabilities

Absorption Probabilities

The absorption probabilities of the continuous-time process $(X_t)_{t\in\mathbb{R}_+}$ can be computed based on the behaviour of the embedded chain $(Z_n)_{n\in\mathbb{N}}$. In fact the continuous waiting time between two jumps has no influence on the absorption probabilities. Here we consider only the simple example of birth and death processes, which can be easily generalized to more complex situations.

The basic idea is to perform a first step analysis on the underlying discrete-time embedded chain. Assume that state $\textcircled{0}$ is absorbing, i.e. $\lambda_0 = 0$, and let

$$T_0 = \inf\{t \in \mathbb{R}_+ \ : \ X_t = 0\}$$

denote the absorption time of the chain into state $\textcircled{0}$. Let now

$$g_0(k) = \mathbb{P}(T_0 < \infty \mid X_0 = k), \qquad k = 0, 1, \ldots, N,$$

denote the probability of absorption in $\textcircled{0}$ starting from state $k \in \{0, 1, \ldots, N\}$. We have the boundary condition $g_0(0) = 1$, and by first step analysis on the chain $(Z_n)_{n\geq 1}$ we get

$$g_0(k) = \frac{\lambda_k}{\lambda_k + \mu_k} g_0(k+1) + \frac{\mu_k}{\lambda_k + \mu_k} g_0(k-1), \qquad k = 1, 2, \ldots, N-1.$$

When the rates $\lambda_k = \lambda$ and $\mu_k = \mu$ are independent of $k \in \{1, 2, \ldots, N-1\}$, this equation becomes

$$g_0(k) = p g_0(k+1) + q g_0(k-1), \qquad k = 1, 2, \ldots, N-1,$$

which is precisely Eq. (2.2.6) for the gambling process with

$$p = \frac{\lambda}{\lambda + \mu} \quad \text{and} \quad q = \frac{\mu}{\lambda + \mu}.$$

When $\lambda_0 = \mu_N = 0$ we have the boundary conditions

$$g_0(0) = 1 \quad \text{and} \quad g_0(N) = 0$$

since the state $\textcircled{N}$ becomes absorbing, and the solution becomes

$$g_0(k) = \frac{(\mu/\lambda)^k - (\mu/\lambda)^N}{1 - (\mu/\lambda)^N}, \qquad k = 0, 1, \ldots, N,$$

when $\lambda \neq \mu$, according to (2.2.11). When $\lambda \neq \mu$, Relation (2.2.12) shows that

$$g_0(k) = \frac{N-k}{N} = 1 - \frac{k}{N}, \qquad k = 0, 1, \ldots, N.$$

Mean Absorption Time

We may still use the embedded chain $(Z_n)_{n\in\mathbb{N}}$ to compute the mean absorption time, using the mean inter-jump times. Here, unlike in the case of absorption probabilities, the random time spent by the continuous-time process $(X_t)_{t\in\mathbb{R}_+}$ should be taken into account in the calculation. We consider a birth and death process on $\{0, 1, \ldots, N\}$ with absorbing states 0 and N.

Recall that the the mean time spent at state i, given that the next transition is from i to $\boxed{i+1}$, is equal to

$$\mathbb{E}[\tau_{i,i+1}] = \frac{1}{\lambda_i}, \qquad i = 1, 2, \ldots, N-1,$$

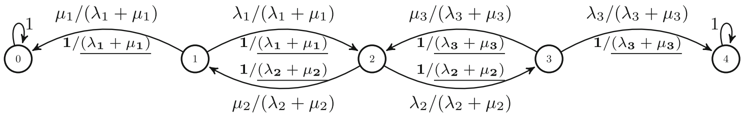
and the mean time spent at state i, given that the next transition is from i to $\boxed{i\text{-}1}$, is equal to

$$\mathbb{E}[\tau_{i,i-1}] = \frac{1}{\mu_i}, \qquad i = 1, 2, \ldots, N-1.$$

We associate a *weighted* graph to the Markov chain $(Z_n)_{n\in\mathbb{N}}$ that includes the average

$$\mathbb{E}[\tau_i] = \frac{1}{\lambda_i + \mu_i}$$

of the time $\tau_i = \min(\tau_{i,i-1}, \tau_{i,i+1})$ spent in state i before the next transition, $i = 1, 2, \ldots, N-1$. In the next graph, which is drawn for $N = 4$, the weights are *underlined*:



with $\lambda_0 = \mu_4 = 0$.

Proposition 9.10 *The mean absorption times*

$$h_{0,N}(i) = \mathbb{E}\left[T_{0,N} \mid X_0 = i\right], \qquad i = 0, 1, \ldots, N,$$

into states $\{0, N\}$ *starting from state* $i \in \{0, 1, \ldots, N\}$ *satisfy the boundary conditions* $h_{0,N}(0) = h_{0,N}(N) = 0$ *the first step analysis equation*

$$h_{0,N}(i) = \frac{1}{\lambda_i + \mu_i} + \frac{\mu_i}{\lambda_i + \mu_i} h_{0,N}(i-1) + \frac{\lambda_i}{\lambda_i + \mu_i} h_{0,N}(i+1),$$

$i = 1, 2, \ldots, N-1.$

Proof By first step analysis on the discrete-time embedded chain $(Z_n)_{n\geq 1}$ with transition matrix

$$P = [\ P_{i,j}\ ]_{i,j\in\mathbb{S}} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \frac{\mu_1}{\lambda_1+\mu_1} & 0 & \frac{\lambda_1}{\lambda_1+\mu_1} & 0 & \cdots & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots & \vdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \frac{\lambda_{N-1}}{\lambda_{N-1}+\mu_{N-1}} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix},$$

we have

$$\begin{aligned} h_{0,N}(i) &= \frac{\mu_i}{\lambda_i+\mu_i}\big(\mathbb{E}[\tau_i] + h_{0,N}(i-1)\big) + \frac{\lambda_i}{\lambda_i+\mu_i}\big(\mathbb{E}[\tau_i] + h_{0,N}(i+1)\big) \\ &= \frac{\mu_i}{\lambda_i+\mu_i}\left(\frac{1}{\lambda_i+\mu_i} + h_{0,N}(i-1)\right) + \frac{\lambda_i}{\lambda_i+\mu_i}\left(\frac{1}{\lambda_i+\mu_i} + h_{0,N}(i+1)\right), \end{aligned} \tag{9.8.1}$$

$i = 1, 2, \ldots, N-1$. □

Note that by conditioning on the independent exponential random variables $\tau_{i,i-1}$ and $\tau_{i,i+1}$ we can also show that

$$\mathbb{E}[\tau_i \mid \tau_{i,i-1} < \tau_{i,i+1}] = \mathbb{E}[\tau_i \mid \tau_{i,i+1} < \tau_{i,i-1}] = \mathbb{E}[\tau_i] = \frac{1}{\lambda_i+\mu_i},$$

$i = 1, 2, \ldots, N-1$, cf. (1.7.9) in Exercise 1.4-(a), hence (9.8.1) can be rewritten as

$$\begin{aligned} h_{0,N}(i) = \frac{\mu_i}{\lambda_i+\mu_i}&\big(\mathbb{E}[\tau_i \mid \tau_{i,i-1} < \tau_{i,i+1}] + h_{0,N}(i-1)\big) \\ &+ \frac{\lambda_i}{\lambda_i+\mu_i}\big(\mathbb{E}[\tau_i \mid \tau_{i,i+1} < \tau_{i,i-1}] + h_{0,N}(i+1)\big). \end{aligned}$$

When the rates $\lambda_i = \lambda$ and $\mu_i = \mu$ are independent of $i \in \{1, 2, \ldots, N-1\}$, this equation becomes

$$h_{0,N}(i) = \frac{1}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}h_{0,N}(i+1) + \frac{\mu}{\lambda+\mu}h_{0,N}(i-1),$$

$i = 1, 2, \ldots, N-1$, which is a modification of Eq. (2.3.6), by replacing the discrete time step by the average time $1/(\lambda+\mu)$ spent at any state. Rewriting the equation as

$$h_{0,N}(i) = \frac{1}{\lambda + \mu} + p h_{0,N}(i+1) + q h_{0,N}(i-1),$$

$i = 1, 2, \ldots, N-1$, or

$$(\lambda + \mu) h_{0,N}(i) = 1 + p(\lambda + \mu) h_{0,N}(i+1) + q(\lambda + \mu) h_{0,N}(i-1),$$

$i = 1, 2, \ldots, N-1$, with

$$p = \frac{\lambda}{\lambda + \mu} \quad \text{and} \quad q = \frac{\mu}{\lambda + \mu},$$

we find from (2.3.11) that, with $r = q/p = \mu/\lambda$,

$$(\lambda + \mu) h_{0,N}(k) = \frac{1}{q - p}\left(k - N\frac{1 - r^k}{1 - r^N}\right),$$

i.e.

$$\boxed{h_{0,N}(k) = \frac{1}{\mu - \lambda}\left(k - N\frac{1 - (\mu/\lambda)^k}{1 - (\mu/\lambda)^N}\right), \qquad k = 0, 1, \ldots, N,} \quad (9.8.2)$$

when $\lambda \neq \mu$. In the limit $\lambda \to \mu$ we find by (2.3.17) that

$$h_{0,N}(k) = \frac{1}{2\mu} k(N - k), \qquad k = 0, 1, \ldots, N.$$

This solution is similar to that of the gambling problem with draw Exercise 2.1 as we multiply the solution of the gambling problem in the fair case by the average time $1/(2\mu)$ spent in any state in $\{1, 2, \ldots, N-1\}$.

The mean absorption time for the embedded chain $(Z_n)_{n \in \mathbb{N}}$ can be recovered by dividing (9.8.2) by the mean time $\mathbb{E}[\tau_i] = 1/(\lambda + \mu)$ between two jumps, as

$$\frac{\lambda + \mu}{\mu - \lambda}\left(k - N\frac{1 - (\mu/\lambda)^k}{1 - (\mu/\lambda)^N}\right), \qquad k = 0, 1, \ldots, N, \tag{9.8.3}$$

which coincides with (2.3.11) in the non-symmetric case with $p = \lambda/(\lambda + \mu)$ and $p = \mu/(\lambda + \mu)$, and recovers (2.3.17), i.e.

$$k(N - k), \qquad k = 0, 1, \ldots, N,$$

in the symmetric case $\lambda = \mu$.

In Table 9.1 we gather some frequent questions and their corresponding solution methods.

Table 9.1 Summary of computing methods

How to compute	Method
The infinitesimal generator $Q = (\lambda_{i,j})_{i,j\in\mathbb{S}}$	$Q = \dfrac{dP(t)}{dt}_{\mid t=0} = P'(0)$
The semigroup $(P(t))_{t\in\mathbb{R}_+}$	$P(t) = \exp(tQ)$, $t \in \mathbb{R}_+$, $P(h) = \mathrm{I_d} + hQ + o(h)$, $h \searrow 0$
The stationary distribution π	Solve* $\pi Q = 0$ for π
The probability distribution of the time $\tau_{i,j}$ spent in $i \to j$	Exponential distribution $(\lambda_{i,j})$
The probability distribution of the time τ_i spent at state i	Exponential distribution $\left(\sum_{l\neq i} \lambda_{i,l}\right)$
$\lim_{t\to\infty} \exp\left(t \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix}\right)$	$\begin{bmatrix} \dfrac{\beta}{\alpha+\beta} & \dfrac{\alpha}{\alpha+\beta} \\ \dfrac{\beta}{\alpha+\beta} & \dfrac{\alpha}{\alpha+\beta} \end{bmatrix}$
Hitting probabilities	Solve† $g = Pg$ for the embedded chain
Mean hitting times	Use the embedded chain with weighted links using mean inter-jump times

*Remember that the values of $\pi(k)$ have to add up to 1

†Be sure to write only the relevant rows of the system under the appropriate boundary conditions

Exercises

Exercise 9.1 A workshop has five machines and one repairman. Each machine functions until it fails at an exponentially distributed random time with rate $\mu = 0.20$ per hour. On the other hand, it takes a exponentially distributed random time with (rate) $\lambda = 0.50$ per hour to repair a given machine. We assume that the machines behave independently of one another, and that

(i) up to five machines can operate at any given time,
(ii) at most one can be under repair at any time.

Compute the proportion of time the repairman is idle in the long run.

Exercise 9.2 Two types of consultations occur at a database according to two independent Poisson processes: "read" consultations arrive at the rate (or intensity) λ_R and "write" consultations arrive at the rate (or intensity) λ_W.

(a) What is the probability that the time interval between two consecutive "read" consultations is larger than $t > 0$?
(b) What is the probability that during the time interval $[0, t]$, at most three "write" consultations arrive?
(c) What is the probability that the next arriving consultation is a "read" consultation?

(d) Determine the distribution of the number of arrived "read" consultations during $[0, t]$, given that in this interval a total number of n consultations occurred.

Exercise 9.3 Consider two machines, operating simultaneously and independently, where both machines have an exponentially distributed time to failure with mean $1/\mu$. There is a single repair facility, and the repair times are exponentially distributed with rate λ.

(a) In the long run, what is the probability that no machines are operating when $\lambda = \mu = 1$?
(b) We now assume that at most one machine can operate at any time. Namely, while one machine is working, the other one may be either under repair or already fixed and waiting to take over. How does this modify your answer to question (a)?

Exercise 9.4 Passengers arrive at a cable car station according to a Poisson process with intensity $\lambda > 0$. Each car contains at most 4 passengers, and a cable car arrives immediately and leaves with 4 passengers as soon as there are at least 4 people in the queue. We let $(X_t)_{t\in\mathbb{R}_+}$ denote the number of passengers in the waiting queue at time $t \geq 0$.

(a) Explain why $(X_t)_{t\in\mathbb{R}_+}$ is a continuous-time Markov chain with state space $\mathbb{S} = \{0, 1, 2, 3\}$, and give its matrix infinitesimal generator Q.
(b) Compute the limiting distribution $\pi = [\pi_0, \pi_1, \pi_2, \pi_3]$ of $(X_t)_{t\in\mathbb{R}_+}$.
(c) Compute the mean time between two departures of cable cars.

Exercise 9.5 [MT15] We consider a stock whose prices can only belong to the following five ticks:

$$\$10.01;\quad \$10.02;\quad \$10.03;\quad \$10.04;\quad \$10.05,$$

numbered $k = 1, 2, 3, 4, 5$.
At time t, the order book for this stock contains exactly $N_t^{(k)}$ sell orders at the price tick n$^o k$, $k = 1, 2, 3, 4, 5$, where $(N_t^{(k)})_{t\in\mathbb{R}_+}$ are independent Poisson processes with same intensity $\lambda > 0$. In addition,

- any sell order can be cancelled after an exponentially distributed random time with parameter $\mu > 0$,
- buy market orders are submitted according to another Poisson process with intensity $\theta > 0$, and are filled instantly at the lowest order price present in the book.

Order cancellations can occur as a result of various trading algorithms such as, e.g., "spoofing", "layering", or "front running".

(a) Show that the *total* number of sell orders L_t in the order book at time t forms a continuous-time Markov chain, and write down its infinitesimal generator Q.
(b) It is estimated that 95% percent of high-frequency trader orders are later cancelled. What relation does this imply between μ and λ?

Exercise 9.6 The size of a fracture in a rock formation is modeled by a continuous-time pure birth process with parameters

$$\lambda_k = (1+k)^\rho, \qquad k \geq 1,$$

i.e. the growth rate of the fracture is a power of $1+k$, where k is the current fracture length. Show that when $\rho > 1$, the mean time for the fracture length to grow to infinity is finite. Conclude that the time to failure of the rock formation is almost-surely finite.[3]

Exercise 9.7 Customers arrive at a processing station according to a Poisson process with rate $\lambda = 0.1$, i.e. on average one customer per ten minutes. Processing of customer queries starts as soon as the third customer enters the queue.

(a) Compute the expected time until the start of the customer service.
(b) Compute the probability that no customer service occurs within the first hour.

Exercise 9.8 Suppose that customers arrive at a facility according to a Poisson process having rate $\lambda = 3$. Let N_t be the number of customers that have arrived up to time t and let T_n be the arrival time of the nth customer, $n = 1, 2, \ldots$ Determine the following (conditional) probabilities and (conditional) expectations, where $0 < t_1 < t_2 < t_3 < t_4$.

(a) $\mathbb{P}(N_{t_3} = 5 \mid N_{t_1} = 1)$.
(b) $\mathbb{E}[N_{t_1} N_{t_4} (N_{t_3} - N_{t_2})]$.
(c) $\mathbb{E}[N_{t_2} \mid T_2 > t_1]$.

Exercise 9.9 Let $(X_t)_{t\in\mathbb{R}_+}$ be a birth and death process on $\{0, 1, 2\}$ with birth and death parameters $\lambda_0 = 2\alpha$, $\lambda_1 = \alpha$, $\lambda_2 = 0$, and $\mu_0 = 0$, $\mu_1 = \beta$, $\mu_2 = 2\beta$. Determine the stationary distribution of $(X_t)_{t\in\mathbb{R}_+}$.

Exercise 9.10 Let $(X_t)_{t\in\mathbb{R}_+}$ be a birth and death process on $0, 1, \ldots, N$ with birth and death parameters $\lambda_n = \alpha(N-n)$ and $\mu_n = \beta n$, respectively. Determine the stationary distribution of $(X_t)_{t\in\mathbb{R}_+}$.

Exercise 9.11 Consider a pure birth process with birth rates $\lambda_0 = 1, \lambda_1 = 3, \lambda_2 = 2, \lambda_3 = 5$. Compute $P_{0,n}(t)$ for $n = 0, 1, 2, 3$.

Exercise 9.12 Consider a pure birth process $(X_t)_{t\in\mathbb{R}_+}$ started at $X_0 = 0$, and let T_k denote the time until the kth birth. Show that

$$\mathbb{P}(T_1 > t \text{ and } T_2 > t+s) = P_{0,0}(t)(P_{0,0}(s) + P_{0,1}(s)).$$

Determine the joint probability density function of (T_1, T_2), and then the joint density of $(\tau_0, \tau_1) := (T_1, T_2 - T_1)$.

---

[3]Recall that a finite-valued random variable may have an infinite mean.

Exercise 9.13 Cars pass a certain street location with identical speeds, according to a Poisson process with rate $\lambda > 0$. A woman at that location needs T units of time to cross the street, i.e. she waits until it appears that no car will cross that location within the next T time units.

(a) Find the probability that her waiting time is 0.
(b) Find her expected waiting time.
(c) Find the total average time it takes to cross the street.
(d) Assume that, due to other factors, the crossing time in the absence of cars is an independent exponentially distributed random variable with parameter $\mu > 0$. Find the total average time it takes to cross the street in this case.

Exercise 9.14 A machine is maintained at random times, such that the inter-service times $(\tau_k)_{k\geq 0}$ are *i.i.d.* with exponential distribution of parameter $\mu > 0$. The machine breaks down if it has not received maintenance for more than T units of time. After breaking down it is automatically repaired.

(a) Compute the probability that the machine breaks down before its first maintenance after it is started.
(b) Find the expected time until the machine breaks down.
(c) Assuming that the repair time is exponentially distributed with parameter $\lambda > 0$, find the proportion of time the machine is working.

Exercise 9.15 A system consists of two machines and two repairmen. Each machine can work until failure at an exponentially distributed random time with parameter 0.2. A failed machine can be repaired only by one repairman, within an exponentially distributed random time with parameter 0.25. We model the number X_t of working machines at time $t \in \mathbb{R}_+$ as a continuous-time Markov process.

(a) Complete the missing entries in the matrix

$$Q = \begin{bmatrix} \square & 0.5 & 0 \\ 0.2 & \square & \square \\ 0 & \square & -0.4 \end{bmatrix}$$

of its generator.
(b) Calculate the long-run probability distribution $[\pi_0, \pi_1, \pi_2]$ of X_t.
(c) Compute the average number of working machines in the long run.
(d) Given that a working machine can produce 100 units every hour, how many units can the system produce per hour in the long run?
(e) Assume now that in case a single machine is under failure then both repairmen can work on it, therefore dividing the mean repair time by a factor 2. Complete the missing entries in the matrix

$$Q = \begin{bmatrix} -0.5 & \square & \square \\ \square & -0.7 & \square \\ \square & \square & -0.4 \end{bmatrix}$$

of the modified generator and calculate the long run probability distribution $[\pi_0, \pi_1, \pi_2]$ for X_t.

Exercise 9.16 Let $X_1(t)$ and $X_2(t)$ be two independent two-state Markov chains on $\{0, 1\}$ and having the same infinitesimal matrix

$$\begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}.$$

Argue that $Z(t) := X_1(t) + X_2(t)$ is a Markov chain on the state space $\mathbb{S} = \{0, 1, 2\}$ and determine the transition semigroup $P(t)$ of $Z(t)$.

Exercise 9.17 Consider a two-state discrete-time Markov chain $(\xi_n)_{n\geq 0}$ on $\{0, 1\}$ with transition matrix

$$\begin{bmatrix} 0 & 1 \\ 1-\alpha & \alpha \end{bmatrix}. \tag{9.8.4}$$

Let $(N_t)_{t\in\mathbb{R}_+}$ be a Poisson process with parameter $\lambda > 0$, and let the

$$X_t = \xi_{N_t}, \qquad t \in \mathbb{R}_+,$$

i.e. $(X_t)_{t\in\mathbb{R}_+}$ is a two-state birth and death process.

(a) Compute the mean return time $\mathbb{E}[T_0^r \mid X_0 = 0]$ of X_t to state 0, where T_0^r is defined as

$$T_0^r = \inf\{t > T_1 \ : \ X_t = 0\}$$

and

$$T_1 = \inf\{t > 0 \ : \ X_t = 1\}$$

is the first hitting time of state 1. Note that the return time

(b) Compute the mean return time $\mathbb{E}[T_1^r \mid X_0 = 1]$ of X_t to state 1, where T_1^r is defined as

$$T_1^r = \inf\{t > T_0 \ : \ X_t = 1\}$$

and

$$T_0 = \inf\{t > 0 \ : \ X_t = 0\}$$

is the first hitting time of state 0. The return time T_1^r to 1 starting from 1 is evaluated by switching first to state 0 before returning to state 1.

(c) Show that $(X_t)_{t\in\mathbb{R}_+}$ is a two-state birth and death process and determine its generator matrix Q in terms of α and λ.

Problem 9.18 Let $(N_t^1)_{t\in\mathbb{R}_+}$ and $(N_t^2)_{t\in\mathbb{R}_+}$ be two independent Poisson processes with intensities $\lambda_1 > 0$ and $\lambda_2 > 0$.

(a) Show that $(N_t^1 + N_t^2)_{t\in\mathbb{R}_+}$ is a Poisson process and find its intensity.
(b) Consider the difference

$$M_t = N_t^1 - N_t^2, \qquad t \in \mathbb{R}_+,$$

and that $(M_t)_{t\in\mathbb{R}_+}$ has stationary independent increments.
(c) Find the distribution of $M_t - M_s$, $0 < s < t$.
(d) Compute

$$\lim_{t\to\infty} \mathbb{P}(|M_t| \leq c)$$

for any $c > 0$.
(e) Suppose that N_t^1 denotes the number of clients arriving at a taxi station during the time interval $[0, t]$, and that N_t^2 denotes the number of taxis arriving at that same station during the same time interval $[0, t]$.
How do you interpret the value of M_t depending on its sign?
How do you interpret the result of Question (d)?

Problem 9.19 We consider a birth and death process $(X_t)_{t\in\mathbb{R}_+}$ on $\{0, 1, \ldots, N\}$ with transition semigroup $(P(t))_{t\in\mathbb{R}}$ and birth and death rates

$$\lambda_n = (N-n)\lambda, \qquad \mu_n = n\mu, \qquad n = 0, 1, \ldots, N.$$

This process is used for the modeling of membrane channels in neuroscience.

(a) Write down the infinitesimal generator Q of $(X_t)_{t\in\mathbb{R}_+}$.
(b) From the forward Kolmogorov equation $P'(t) = P(t)Q$, show that for all $n = 0, 1, \ldots, N$ we have

$$\begin{cases} P'_{n,0}(t) = -\lambda_0 P_{n,0}(t) + \mu_1 P_{n,1}(t), \\ \\ P'_{n,k}(t) = \lambda_{k-1} P_{n,k-1}(t) - (\lambda_k + \mu_k) P_{n,k}(t) + \mu_{k+1} P_{n,k+1}(t), \\ \\ P'_{n,N}(t) = \lambda_{N-1} P_{n,N-1}(t) - \mu_N P_{n,N}(t), \end{cases}$$

$k = 1, 2, \ldots, N-1$.
(c) Let

$$G_k(s,t) = \mathbb{E}\left[s^{X_t} \mid X_0 = k\right] = \sum_{n=0}^{N} s^n \mathbb{P}(X_t = n \mid X_0 = k) = \sum_{n=0}^{N} s^n P_{k,n}(t)$$

denote the generating function of X_t given that $X_0 = k \in \{0, 1, \ldots, N\}$. From the result of Question (b), show that $G_k(s,t)$ satisfies the *partial differential equation*

$$\frac{\partial G_k}{\partial t}(s,t) = \lambda N(s-1)G_k(s,t) + (\mu + (\lambda-\mu)s - \lambda s^2)\frac{\partial G_k}{\partial s}(s,t), \quad (9.8.5)$$

with $G_k(s,0) = s^k$, $k = 0, 1, \ldots, N$.

(d) Verify that the solution of (9.8.5) is given by

$$G_k(s,t) = \frac{1}{(\lambda+\mu)^N}(\mu + \lambda s + \mu(s-1)\mathrm{e}^{-(\lambda+\mu)t})^k(\mu + \lambda s - \lambda(s-1)\mathrm{e}^{-(\lambda+\mu)t})^{N-k},$$

$k = 0, 1, \ldots, N$.

(e) Show that

$$\begin{aligned}\mathbb{E}[X_t \mid X_0 = k] = {} & \frac{k}{(\lambda+\mu)^N}(\lambda + \mu\mathrm{e}^{-(\lambda+\mu)t})(\mu+\lambda)^{k-1}(\mu+\lambda)^{N-k} \\ & + \frac{N-k}{(\lambda+\mu)^N}(\mu+\lambda)^k(\lambda - \lambda\mathrm{e}^{-(\lambda+\mu)t})(\mu+\lambda)^{N-k-1}.\end{aligned}$$

(f) Compute

$$\lim_{t\to\infty} \mathbb{E}[X_t \mid X_0 = k]$$

and show that it does not depend on $k \in \{0, 1, \ldots, N\}$.

Chapter 10
Discrete-Time Martingales

As mentioned in the introduction, stochastic processes can be classified into two main families, namely *Markov processes* on the one hand, and *martingales* on the other hand. Markov processes have been our main focus of attention so far, and in this chapter we turn to the notion of martingale. In particular we will give a precise mathematical meaning to the description of martingales stated in the introduction, which says that when $(X_n)_{n\in\mathbb{N}}$ is a martingale, the best possible estimate at time $n \in \mathbb{N}$ of the future value X_m at time $m > n$ is X_n itself. The main application of martingales will be to recover in an elegant way the previous results on gambling processes of Chap. 2. Before that, let us state many recent applications of stochastic modeling are relying on the notion of martingale. In financial mathematics for example, the notion of martingale is used to characterize the fairness and equilibrium of a market model.

10.1 Filtrations and Conditional Expectations

Before dealing with martingales we need to introduce the important notion of *filtration* generated by a discrete-time stochastic process $(X_n)_{n\in\mathbb{N}}$. The *filtration* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ generated by a stochastic process $(X_n)_{n\in\mathbb{N}}$ taking its values in a state space $\mathbb{S}$, is the family of σ-algebras

$$\mathcal{F}_n := \sigma(X_0, X_1, \ldots, X_n), \qquad n \geq 0,$$

which denote the collections of events generated by $X_0, X_1, \ldots, X_n$. Examples of such events include

$$\{X_0 \leq a_0,\ X_1 \leq a_1,\ \ldots,\ X_n \leq a_n\}$$

for $a_0, a_1, \ldots, a_n$ a given fixed sequence of real numbers. Note that we have the inclusion $\mathcal{F}_n \subset \mathcal{F}_{n+1}$, $n \in \mathbb{N}$.

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4_10

One refers to $\mathcal{F}_n$ as the *information* generated by $(X_k)_{k\in\mathbb{N}}$ up to time n, and to $(\mathcal{F}_n)_{n\in\mathbb{N}}$ as the *information flow* generated by $(X_n)_{n\in\mathbb{N}}$. We say that a random variable is $\mathcal{F}_n$-measurable whenever F can be written as a function $F = f(X_0, X_1, \ldots, X_n)$ of $(X_0, X_1, \ldots, X_n)$.

Consider for example the simple random walk

$$S_n := X_1 + X_2 + \cdots + X_n, \qquad n \in \mathbb{N},$$

where $(X_k)_{k\geq 1}$ is a sequence of independent, identically distributed $\{-1, 1\}$ valued random variables. The filtration (or information flow) $(\mathcal{F}_n)_{n\in\mathbb{N}}$ generated by $(S_n)_{n\in\mathbb{N}}$ is given by $\mathcal{F}_0 = \big\{\emptyset, \Omega\big\}$,

$$\mathcal{F}_1 = \big\{\emptyset, \{X_1 = 1\}, \{X_1 = -1\}, \Omega\big\},$$

and

$$\mathcal{F}_2 = \sigma\Big(\big\{\emptyset, \{X_1 = 1, X_2 = 1\}, \{X_1 = 1, X_2 = -1\}, \{X_1 = -1, X_2 = 1\}, \\ \{X_1 = -1, X_2 = -1\}, \Omega\big\}\Big).$$

The notation $\mathcal{F}_n$ is useful to represent a quantity of information available at time n, and various sub σ-algebras of $\mathcal{F}_n$ can be defined such as e.g.

$$\mathcal{G}_2 := \big\{\emptyset, \{X_1 = 1, X_2 = -1\} \cup \{X_1 = -1, X_2 = 1\}, \\ \{X_1 = 1, X_2 = 1\} \cup \{X_1 = -1, X_2 = -1\}, \Omega\big\},$$

which contains less information than $\mathcal{F}_2$, as it only tells whether the increments X_1, X_2 have same signs.

We now review the definition of *conditional expectation*, cf. also Sect. 1.6. Given F a random variable with finite mean the conditional expectation $\mathbb{E}[F \mid \mathcal{F}_n]$ refers to

$$\mathbb{E}[F \mid X_0, X_1, \ldots, X_n] = \mathbb{E}[F \mid X_0 = k_0, \ldots, X_n = k_n]_{k_0 = X_0, \ldots, k_n = X_n},$$

given that $X_0, X_1, \ldots, X_n$ are respectively equal to $k_0, k_1, \ldots, k_n \in \mathbb{S}$.

The conditional expectation $\mathbb{E}[F \mid \mathcal{F}_n]$ is itself a random variable that depends only on the values of $X_0, X_1, \ldots, X_n$, i.e. on the history of the process up to time $n \in \mathbb{N}$. It can also be interpreted as the best possible estimate of F in mean square sense, given the values of $X_0, X_1, \ldots, X_n$, cf. (1.6.17).

A stochastic process $(Z_n)_{n\in\mathbb{N}}$ is said to be $\mathcal{F}_n$*-adapted* if the value of Z_n depends on no more than the information available up to time n in $\mathcal{F}_n$, that means, the value of Z_n is some function of $X_0, X_1, \ldots, X_n$, $n \in \mathbb{N}$.

In particular, any $\mathcal{F}_n$*-adapted* process $(Z_n)_{n\in\mathbb{N}}$ satisfies

$$\mathbb{E}[Z_n \mid \mathcal{F}_n] = Z_n, \qquad n \in \mathbb{N}.$$

10.2 Martingales - Definition and Properties

We now turn to the definition of *martingale*.

Definition 10.1 An integrable,[1] discrete-time stochastic process $(Z_n)_{n\in\mathbb{N}}$ is a *martingale* with respect to $(\mathcal{F}_n)_{n\in\mathbb{N}}$ if $(Z_n)_{n\in\mathbb{N}}$ is $\mathcal{F}_n$-adapted and satisfies the property

$$\mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] = Z_n, \qquad n \in \mathbb{N}. \tag{10.2.1}$$

The process $(Z_n)_{n\in\mathbb{N}}$ is a martingale with respect to $(\mathcal{F}_n)_{n\in\mathbb{N}}$ if, given the information $\mathcal{F}_n$ known up to time n, the best possible estimate of Z_{n+1} is simply Z_n.
Exercise. Using the *tower property* of conditional expectations, show that Definition (10.2.1) can be equivalently stated by saying that

$$\mathbb{E}[M_n \mid \mathcal{F}_k] = M_k, \qquad 0 \le k < n.$$

A particular property of martingales is that their expectation is constant over time.

Proposition 10.2 *Let* $(Z_n)_{n\in\mathbb{N}}$ *be a martingale. We have*

$$\mathbb{E}[Z_n] = \mathbb{E}[Z_0], \qquad n \in \mathbb{N}.$$

Proof From the tower property (1.6.10) we have:

$$\mathbb{E}[Z_{n+1}] = \mathbb{E}[\mathbb{E}[Z_{n+1} \mid \mathcal{F}_n]] = \mathbb{E}[Z_n], \qquad n \in \mathbb{N},$$

hence by induction on $n \in \mathbb{N}$ we have

$$\mathbb{E}[Z_{n+1}] = \mathbb{E}[Z_n] = \mathbb{E}[Z_{n-1}] = \cdots = \mathbb{E}[Z_1] = \mathbb{E}[Z_0], \qquad n \in \mathbb{N}.$$

□

Examples of Martingales

1. Any centered[2] integrable process $(S_n)_{n\in\mathbb{N}}$ with independent increments is a martingale with respect to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ generated by $(S_n)_{n\in\mathbb{N}}$.

   Indeed, in this case we have

$$\begin{aligned}\mathbb{E}[S_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}[S_n \mid \mathcal{F}_n] + \mathbb{E}\left[S_{n+1} - S_n \mid \mathcal{F}_n\right] \\ &= \mathbb{E}[S_n \mid \mathcal{F}_n] + \mathbb{E}[S_{n+1} - S_n] \\ &= \mathbb{E}[S_n \mid \mathcal{F}_n] = S_n, \qquad n \in \mathbb{N}.\end{aligned}$$

[1] Integrable means $\mathbb{E}[|Z_n|] < \infty$ for all $n \in \mathbb{N}$.

[2] A random variable X_n is said to be *centered* if $\mathbb{E}[X_n] = 0$.

In addition to being a martingale, a process $(X_n)_{n\in\mathbb{N}}$ with centered independent increments is also a Markov process, cf. Sect. 4.1. However, not all martingales have the Markov property, and not all Markov processes are martingales. In addition, there are martingales and Markov processes which do not have independent increments.

2. Given $F \in L^2(\Omega)$ a square-integrable random variable and $(\mathcal{F}_n)_{n\in\mathbb{N}}$ a filtration, the process $(X_n)_{n\in\mathbb{N}}$ defined by $X_n := \mathbb{E}[F \mid \mathcal{F}_n]$ is an $(\mathcal{F}_n)_{n\in\mathbb{N}}$-martingale under the probability measure $\mathbb{P}$, as follows from the tower property:

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[\mathbb{E}[F \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n] = \mathbb{E}[F \mid \mathcal{F}_n] = X_n, \quad n \in \mathbb{N}, \tag{10.2.2}$$

from the tower property (1.6.10).

10.3 Stopping Times

Next, we turn to the definition of *stopping time*. If an event occurs at a (random) stopping time, it should be possible, at any time $n \in \mathbb{N}$, to determine whether the event has already occured, based on the information available at time n. This idea is formalized in the next definition.

Definition 10.3 A *stopping time* is a random variable $\tau : \Omega \longrightarrow \mathbb{N}$ such that

$$\{\tau > n\} \in \mathcal{F}_n, \qquad n \in \mathbb{N}. \tag{10.3.1}$$

The meaning of Relation (10.3.1) is that the knowledge of $\{\tau > n\}$ depends only on the information present in $\mathcal{F}_n$ up to time n, i.e. on the knowledge of $X_0, X_1, \ldots, X_n$.

Note that condition (10.3.1) is equivalent to the condition

$$\{\tau \leq n\} \in \mathcal{F}_n, \qquad n \in \mathbb{N},$$

since $\mathcal{F}_n$ is *stable by complement* and $\{\tau \leq n\} = \{\tau > n\}^c$.

Not every $\mathbb{N}$-valued random variable is a stopping time, however, hitting times provide natural examples of stopping times.

Proposition 10.4 *The first hitting time*

$$T_x := \inf\{k \geq 0 \ : \ X_k = x\}$$

of $x \in \mathbb{S}$ is a stopping time.

Proof We have

$$\begin{aligned}\{T_x > n\} &= \{X_0 \neq x, \ X_1 \neq x, \ \ldots, X_n \neq x\} \\ &= \{X_0 \neq x\} \cap \{X_1 \neq x\} \cap \cdots \cap \{X_n \neq x\} \in \mathcal{F}_n, \quad n \in \mathbb{N},\end{aligned}$$

since

$$\{X_0 \neq x\} \in \mathcal{F}_0 \subset \mathcal{F}_n, \ \{X_1 \neq x\} \in \mathcal{F}_1 \subset \mathcal{F}_n, \ \ldots, \ \{X_n \neq x\} \in \mathcal{F}_n, \quad n \in \mathbb{N}.$$

□

On the other hand, the first time

$$T = \inf \left\{ k \geq 0 \ : \ X_k = \max_{l=0,1,\ldots,N} X_l \right\}$$

the process $(X_k)_{k\in\mathbb{N}}$ reaches its maximum over $\{0, 1, \ldots, N\}$ is not a stopping time. Indeed, it is not possible to decide whether $\{T \leq n\}$, i.e. the maximum has been reached before time n, based on the information available at time n.
Exercise: Show that the minimum $\tau \wedge \nu = \min(\tau, \nu)$ of two stopping times is a stopping time.

Definition 10.5 Given $(Z_n)_{n\in\mathbb{N}}$ a stochastic process and $\tau : \Omega \longrightarrow \mathbb{N}$ a stopping time, the *stopped process*

$$(Z_{\tau\wedge n})_{n\in\mathbb{N}} = (Z_{\min(\tau,n)})_{n\in\mathbb{N}}$$

is defined as

$$Z_{\tau\wedge n} = Z_{\min(\tau,n)} = \begin{cases} Z_n & \text{if } n < \tau, \\ \\ Z_\tau & \text{if } n \geq \tau, \end{cases}$$

Using indicator functions we may also write

$$Z_{\tau\wedge n} = Z_n \mathbb{1}_{\{n<\tau\}} + Z_\tau \mathbb{1}_{\{n\geq\tau\}}, \qquad n \in \mathbb{N}.$$

The following Fig. 10.1 is an illustration of the path of a stopped process.

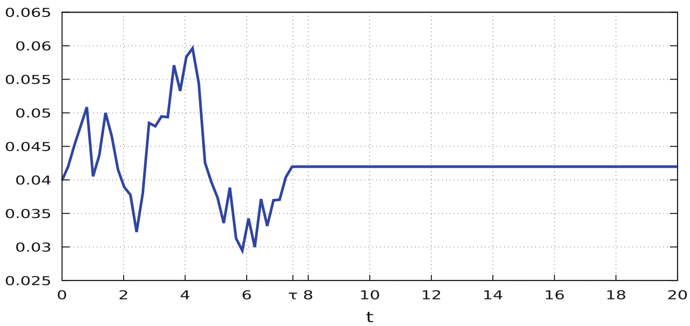


Fig. 10.1 Stopped process

The following Theorem 10.6 is called the *stopping time theorem*, it is due to J.L. Doob (1910-2004).

Theorem 10.6 *Assume that* $(M_n)_{n\in\mathbb{N}}$ *is a martingale with respect to* $(\mathcal{F}_n)_{n\in\mathbb{N}}$. *Then the* stopped process $(M_{\tau\wedge n})_{n\in\mathbb{N}}$ *is also a martingale with respect to* $(\mathcal{F}_n)_{n\in\mathbb{N}}$.

Proof Writing

$$M_n = M_0 + \sum_{l=1}^{n}(M_l - M_{l-1}) = M_0 + \sum_{l=1}^{\infty} \mathbb{1}_{\{l\le n\}}(M_l - M_{l-1}),$$

we find

$$M_{\tau\wedge n} = M_0 + \sum_{l=1}^{\tau\wedge n}(M_l - M_{l-1}) = M_0 + \sum_{l=1}^{n} \mathbb{1}_{\{l\le \tau\}}(M_l - M_{l-1}),$$

and for $k \le n$ we find

$$\begin{aligned}
\mathbb{E}\big[M_{\tau\wedge n} \mid \mathcal{F}_k\big] &= M_0 + \sum_{l=1}^{n} \mathbb{E}\big[\mathbb{1}_{\{l\le\tau\}}(M_l - M_{l-1}) \mid \mathcal{F}_k\big] \\
&= M_0 + \sum_{l=1}^{k} \mathbb{E}\big[\mathbb{1}_{\{l\le\tau\}}(M_l - M_{l-1}) \mid \mathcal{F}_k\big] + \sum_{l=k+1}^{n} \mathbb{E}\big[\mathbb{1}_{\{l\le\tau\}}(M_l - M_{l-1}) \mid \mathcal{F}_k\big] \\
&= M_0 + \sum_{l=1}^{k}(M_l - M_{l-1})\mathbb{E}\big[\mathbb{1}_{\{l\le\tau\}} \mid \mathcal{F}_k\big] \\
&\quad + \sum_{l=k+1}^{n} \mathbb{E}\big[\mathbb{E}\big[(M_l - M_{l-1})\mathbb{1}_{\{l-1<\tau\}} \mid \mathcal{F}_{l-1}\big] \mid \mathcal{F}_k\big] \\
&= M_0 + \sum_{l=1}^{k}(M_l - M_{l-1})\mathbb{1}_{\{l\le\tau\}} \\
&\quad + \sum_{l=k+1}^{n} \mathbb{E}\big[\mathbb{1}_{\{l-1<\tau\}} \underbrace{\mathbb{E}\big[(M_l - M_{l-1}) \mid \mathcal{F}_{l-1}\big]}_{=0} \mid \mathcal{F}_k\big] \\
&= M_0 + \sum_{l=1}^{\tau\wedge k}(M_l - M_{l-1}) \\
&= M_{\tau\wedge k},
\end{aligned}$$

$k = 0, 1, \ldots, n$, where we used the tower property and the fact that

$$\{\tau \ge l\} = \{\tau > l-1\} \in \mathcal{F}_{l-1} \subset \mathcal{F}_l \subset \mathcal{F}_k, \qquad 1 \le l \le k.$$

□

By Theorem 10.6 we know that the stopped process $(M_{\tau\wedge n})_{n\in\mathbb{N}}$ is a martingale, hence its expectation is constant by Proposition 10.2. As a consequence, if τ is a stopping time bounded by a constant $N > 0$, i.e. $\tau \le N$, we have

$$\mathbb{E}[M_\tau] = \mathbb{E}[M_{\tau\wedge N}] = \mathbb{E}[M_{\tau\wedge 0}] = \mathbb{E}[M_0]. \tag{10.3.2}$$

As a consequence of (10.3.2), if $(M_n)_{n\in\mathbb{N}}$ is a martingale and $\tau \le N$ and $\nu \le N$ are two *bounded* stopping times bounded by a constant $N > 0$, we have

$$\mathbb{E}[M_\tau] = \mathbb{E}[M_\nu] = \mathbb{E}[M_0]. \tag{10.3.3}$$

In case τ is only *a.s.* finite, i.e. $\mathbb{P}(\tau < \infty) = 1$, we may also write

$$\mathbb{E}[M_\tau] = \mathbb{E}\left[\lim_{n\to\infty} M_{\tau\wedge n}\right] = \lim_{n\to\infty} \mathbb{E}[M_{\tau\wedge n}] = \mathbb{E}[M_0],$$

provided that the limit and expectation signs can be exchanged, however this may not be always the case. In some situations the exchange of limit and expectation signs can be difficult to justify, nevertheless the exchange is possible when the stopped process $(M_{\tau\wedge n})_{n\in\mathbb{N}}$ is bounded in absolute value, i.e. $|M_{\tau\wedge n}| \le K$ *a.s.*, $n \in \mathbb{N}$, for some constant $K > 0$.

Analog statements can be proved for *sub*martingales, cf. Exercise 10.2 for this notion.

10.4 Ruin Probabilities

In the sequel we will show that, as an application of the stopping time theorem, the ruin probabilities computed for simple random walks in Chap. 2 can be recovered in a simple and elegant way.

Consider the standard random walk (or gambling process) $(S_n)_{n\in\mathbb{N}}$ on $\{0, 1, \ldots, B\}$ with independent $\{-1, 1\}$-valued increments with

$$\mathbb{P}(S_{n+1} - S_n = +1) = p \quad \text{and} \quad \mathbb{P}(S_{n+1} - S_n = -1) = q, \quad n \in \mathbb{N},$$

as introduced in Sect. 2.1. Let

$$T_{0,B} : \Omega \longrightarrow \mathbb{N}$$

be the first hitting time of the boundary $\{0, B\}$, defined by

$$\tau := T_{0,B} := \inf\{n \ge 0 \ : \ S_n = B \text{ or } S_n = 0\}.$$

One checks easily that the event $\{\tau > n\}$ depends only on the history of $(S_k)_{k\in\mathbb{N}}$ up to time n since for $k \in \{1, 2, \dots, B-1\}$ we have

$$\{\tau > n\} = \{0 < S_0 < B\} \cap \{0 < S_1 < B\} \cap \cdots \cap \{0 < S_n < B\},$$

hence τ is a stopping time.

We will recover the ruin probabilities

$$\mathbb{P}(S_\tau = 0 \mid S_0 = k), \qquad k = 0, 1, \dots, B,$$

computed in Chap. 2 in three steps, first in the unbiased case $p = q = 1/2$ (note that the hitting time τ can be shown to be *a.s.* finite, cf. e.g. the identity (2.2.29)).

Step 1. The process $(S_n)_{n\in\mathbb{N}}$ is a martingale.

We note that the process $(S_n)_{n\in\mathbb{N}}$ has independent increments, and in the unbiased case $p = q = 1/2$ those increments are centered:

$$\mathbb{E}[S_{n+1} - S_n] = 1 \times p + (-1) \times q = 1 \times \frac{1}{2} + (-1) \times \frac{1}{2} = 0, \tag{10.4.1}$$

hence $(S_n)_{n\in\mathbb{N}}$ is a *martingale* by Point 1 p. 265.

Step 2. The stopped process $(S_{\tau\wedge n})_{n\in\mathbb{N}}$ is also a martingale, as a consequence of Theorem 10.6.

Step 3. Since the stopped process $(S_{\tau\wedge n})_{n\in\mathbb{N}}$ is a martingale by Theorem 10.6, we find that its expectation $\mathbb{E}[S_{\tau\wedge n} \mid S_0 = k]$ is constant in $n \in \mathbb{N}$ by Proposition 10.2, which gives

$$k = \mathbb{E}[S_0 \mid S_0 = k] = \mathbb{E}[S_{\tau\wedge n} \mid S_0 = k], \qquad k = 0, 1, \dots, B.$$

Letting n go to infinity we get

$$\begin{aligned}\mathbb{E}[S_\tau \mid S_0 = k] &= \mathbb{E}\left[\lim_{n\to\infty} S_{\tau\wedge n} \mid S_0 = k\right] \\ &= \lim_{n\to\infty} \mathbb{E}[S_{\tau\wedge n} \mid S_0 = k] = k,\end{aligned}$$

where the exchange between limit and expectation is justified by the boundedness $|S_{\tau\wedge n}| \le B$ *a.s.*, $n \in \mathbb{N}$. Hence we have

$$\begin{cases} 0 \times \mathbb{P}(S_\tau = 0 \mid S_0 = k) + B \times \mathbb{P}(S_\tau = B \mid S_0 = k) = \mathbb{E}[S_\tau \mid S_0 = k] = k \\ \\ \mathbb{P}(S_\tau = 0 \mid S_0 = k) + \mathbb{P}(S_\tau = B \mid S_0 = k) = 1, \end{cases}$$

which shows that

$$\mathbb{P}(S_\tau = B \mid S_0 = k) = \frac{k}{B} \quad \text{and} \quad \mathbb{P}(S_\tau = 0 \mid S_0 = k) = 1 - \frac{k}{B},$$

$k = 0, 1, \ldots, B$, which recovers (2.2.21) without use of boundary conditions, and with short calculations. Namely, the solution has been obtained in a simple way without solving any finite difference equation, demonstrating the power of the martingale approach.

Next, let us turn to the biased case where $p \neq q$. In this case the process $(S_n)_{n\in\mathbb{N}}$ is no longer a martingale, and in order to use Theorem 10.6 we need to construct a martingale of a different type. Here we note that the process

$$M_n := \left(\frac{q}{p}\right)^{S_n}, \qquad n \in \mathbb{N},$$

is a martingale with respect to $(\mathcal{F}_n)_{n\in\mathbb{N}}$.

Step 1. The process $(M_n)_{n\in\mathbb{N}}$ is a martingale.

Indeed, we have

$$\begin{aligned}
\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}\left[\left(\frac{q}{p}\right)^{S_{n+1}} \Big| \mathcal{F}_n\right] = \mathbb{E}\left[\left(\frac{q}{p}\right)^{S_{n+1}-S_n} \left(\frac{q}{p}\right)^{S_n} \Big| \mathcal{F}_n\right] \\
&= \left(\frac{q}{p}\right)^{S_n} \mathbb{E}\left[\left(\frac{q}{p}\right)^{S_{n+1}-S_n} \Big| \mathcal{F}_n\right] \\
&= \left(\frac{q}{p}\right)^{S_n} \mathbb{E}\left[\left(\frac{q}{p}\right)^{S_{n+1}-S_n}\right] \\
&= \left(\frac{q}{p}\right)^{S_n} \left(\frac{q}{p}\mathbb{P}(S_{n+1} - S_n = 1) + \left(\frac{q}{p}\right)^{-1} \mathbb{P}(S_{n+1} - S_n = -1)\right) \\
&= \left(\frac{q}{p}\right)^{S_n} \left(p\frac{q}{p} + q\left(\frac{q}{p}\right)^{-1}\right) \\
&= \left(\frac{q}{p}\right)^{S_n} (q + p) = \left(\frac{q}{p}\right)^{S_n} = M_n,
\end{aligned}$$

$n \in \mathbb{N}$. In particular, the expectation of $(M_n)_{n\in\mathbb{N}}$ is constant over time by Proposition 10.2 since it is a martingale, i.e. we have

$$\left(\frac{q}{p}\right)^k = \mathbb{E}[M_0 \mid S_0 = k] = \mathbb{E}[M_n \mid S_0 = k], \qquad k = 0, 1, \ldots, B, \quad n \in \mathbb{N}.$$

Step 2. The stopped process $(M_{\tau\wedge n})_{n\in\mathbb{N}}$ is also a martingale, as a consequence of Theorem 10.6.

Step 3. Since the stopped process $(M_{\tau\wedge n})_{n\in\mathbb{N}}$ remains a martingale by Theorem 10.6, its expected value $\mathbb{E}[M_{\tau\wedge n} \mid S_0 = k]$ is constant in $n \in \mathbb{N}$ by Proposition 10.2. This gives

$$\left(\frac{q}{p}\right)^k = \mathbb{E}[M_0 \mid S_0 = k] = \mathbb{E}[M_{\tau\wedge n} \mid S_0 = k].$$

Next, letting n go to infinity we find

$$\begin{aligned}\left(\frac{q}{p}\right)^k &= \mathbb{E}[M_0 \mid S_0 = k] = \lim_{n\to\infty} \mathbb{E}[M_{\tau\wedge n} \mid S_0 = k] \\ &= \mathbb{E}\left[\lim_{n\to\infty} M_{\tau\wedge n} \mid S_0 = k\right] \\ &= \mathbb{E}[M_\tau \mid S_0 = k],\end{aligned}$$

hence

$$\begin{aligned}\left(\frac{q}{p}\right)^k &= \mathbb{E}[M_\tau \mid S_0 = k] \\ &= \left(\frac{q}{p}\right)^B \mathbb{P}\left(M_\tau = \left(\frac{q}{p}\right)^B \mid S_0 = k\right) + \left(\frac{q}{p}\right)^0 \mathbb{P}\left(M_\tau = \left(\frac{q}{p}\right)^0 \mid S_0 = k\right) \\ &= \left(\frac{q}{p}\right)^B \mathbb{P}\left(M_\tau = \left(\frac{q}{p}\right)^B \mid S_0 = k\right) + \mathbb{P}(M_\tau = 1 \mid S_0 = k).\end{aligned}$$

Solving the system of equations

$$\begin{cases} \left(\frac{q}{p}\right)^k = \left(\frac{q}{p}\right)^B \mathbb{P}\left(M_\tau = \left(\frac{q}{p}\right)^B \mid S_0 = k\right) + \mathbb{P}(M_\tau = 1 \mid S_0 = k) \\ \mathbb{P}\left(M_\tau = \left(\frac{q}{p}\right)^B \Big| S_0 = k\right) + \mathbb{P}(M_\tau = 1 \mid S_0 = k) = 1, \end{cases}$$

gives

$$\begin{aligned}\mathbb{P}(S_\tau = B \mid S_0 = k) &= \mathbb{P}\left(M_\tau = \left(\frac{q}{p}\right)^B \Big| S_0 = k\right) \qquad (10.4.2) \\ &= \frac{(q/p)^k - 1}{(q/p)^B - 1}, \qquad k = 0, 1, \ldots, B,\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}(S_\tau = 0 \mid S_0 = k) &= \mathbb{P}(M_\tau = 1 \mid S_0 = k)\\ &= 1 - \frac{(q/p)^k - 1}{(q/p)^B - 1},\\ &= \frac{(q/p)^B - (q/p)^k}{(q/p)^B - 1},\end{aligned}$$

$k = 0, 1, \ldots, B$, which recovers (2.2.11).

10.5 Mean Game Duration

In this section we show that the mean game durations $\mathbb{E}[\tau \mid S_0 = k]$ computed in Sect. 2.3 can also be recovered as a second application of the stopping time theorem.

In the case of a fair game $p = q = 1/2$ the martingale method can be used by noting that $(S_n^2 - n)_{n\in\mathbb{N}}$ is also a martingale.

Step 1. The process $(S_n^2 - n)_{n\in\mathbb{N}}$ is a martingale.

We have

$$\begin{aligned}&\mathbb{E}[S_{n+1}^2 - (n+1) \mid \mathcal{F}_n] = \mathbb{E}[(S_n + S_{n+1} - S_n)^2 - (n+1) \mid \mathcal{F}_n]\\ &= \mathbb{E}[S_n^2 + (S_{n+1} - S_n)^2 + 2S_n(S_{n+1} - S_n) - (n+1) \mid \mathcal{F}_n]\\ &= \mathbb{E}[S_n^2 - n - 1 \mid \mathcal{F}_n] + \mathbb{E}[(S_{n+1} - S_n)^2 \mid \mathcal{F}_n] + 2\mathbb{E}[S_n(S_{n+1} - S_n) \mid \mathcal{F}_n]\\ &= S_n^2 - n - 1 + \mathbb{E}[(S_{n+1} - S_n)^2 \mid \mathcal{F}_n] + 2S_n\mathbb{E}[S_{n+1} - S_n \mid \mathcal{F}_n]\\ &= S_n^2 - n - 1 + \mathbb{E}[(S_{n+1} - S_n)^2] + 2S_n\mathbb{E}[S_{n+1} - S_n]\\ &= S_n^2 - n - 1 + \mathbb{E}[(S_{n+1} - S_n)^2]\\ &= S_n^2 - n, \qquad n \in \mathbb{N}.\end{aligned}$$

Step 2. The stopped process $(S_{\tau\wedge n}^2 - \tau \wedge n)_{n\in\mathbb{N}}$ is also a martingale, as a consequence of Theorem 10.6.

Step 3. Since the stopped process $(S_{\tau\wedge n}^2 - \tau \wedge n)_{n\in\mathbb{N}}$ is also a martingale, its expectation $\mathbb{E}[S_{\tau\wedge n}^2 - \tau \wedge n \mid S_0 = k]$ is constant in $n \in \mathbb{N}$ by Proposition 10.2, hence we have

$$k^2 = \mathbb{E}\left[S_0^2 - 0 \mid S_0 = k\right] = \mathbb{E}\left[S_{\tau\wedge n}^2 - \tau \wedge n \mid S_0 = k\right],$$

and after taking the limit as n goes to infinity,

$$\begin{aligned}
k^2 &= \lim_{n\to\infty} \mathbb{E}\left[S_{\tau\wedge n}^2 - \tau\wedge n \mid S_0 = k\right] \\
&= \mathbb{E}\left[\lim_{n\to\infty} S_{\tau\wedge n}^2 - \lim_{n\to\infty} \tau\wedge n \mid S_0 = k\right] \\
&= \mathbb{E}[S_\tau^2 - \tau \mid S_0 = k],
\end{aligned}$$

since $S_{\tau\wedge n}^2 \in [0, B^2]$ for all $n \in \mathbb{N}$ and $n \longmapsto \tau \wedge n$ is nondecreasing, and this gives[3, 4]

$$\begin{aligned}
k^2 &= \mathbb{E}[S_\tau^2 - \tau \mid S_0 = k] \\
&= \mathbb{E}[S_\tau^2 \mid S_0 = k] - \mathbb{E}[\tau \mid S_0 = k] \\
&= B^2 \mathbb{P}(S_\tau = B \mid S_0 = k) + 0^2 \times \mathbb{P}(S_\tau = 0 \mid S_0 = k) - \mathbb{E}[\tau \mid S_0 = k],
\end{aligned}$$

i.e.

$$\begin{aligned}
\mathbb{E}[\tau \mid S_0 = k] &= B^2 \mathbb{P}(S_\tau = B \mid S_0 = k) - k^2 \\
&= B^2 \frac{k}{B} - k^2 \\
&= k(B - k),
\end{aligned}$$

$k = 0, 1, \ldots, B$, which recovers (2.3.17).

Finally we show how to recover the value of the mean game duration, i.e. the mean hitting time of the boundaries $\{0, B\}$ in the non-symmetric case $p \neq q$.

Step 1. The process $S_n - (p - q)n$ is a martingale.

In this case we note that although $(S_n)_{n\in\mathbb{N}}$ does not have centered increments and is not a martingale, the compensated process

$$S_n - (p - q)n, \qquad n \in \mathbb{N},$$

is a martingale because, in addition to being independent, its increments are centered random variables:

$$\mathbb{E}[S_n - S_{n-1} - (p - q)] = \mathbb{E}[S_n - S_{n-1}] - (p - q) = 0,$$

by (10.4.1).

Step 2. The stopped process $(S_{\tau\wedge n} - (p - q)(\tau \wedge n))_{n\in\mathbb{N}}$ is also a martingale, as a consequence of Theorem 10.6.

[3] By application of the *dominated convergence* theorem.

[4] By application of the *monotone convergence* theorem.

Step 3. The expectation $\mathbb{E}[S_{\tau\wedge n} - (p-q)(\tau \wedge n) \mid S_0 = k]$ is constant in $n \in \mathbb{N}$.

Step 4. Since the stopped process $(S_{\tau\wedge n} - (p-q)(\tau \wedge n))_{n\in\mathbb{N}}$ is a martingale, we have

$$k = \mathbb{E}[S_0 - 0 \mid S_0 = k] = \mathbb{E}[S_{\tau\wedge n} - (p-q)(\tau \wedge n) \mid S_0 = k],$$

and after taking the limit as n goes to infinity,

$$\begin{aligned} k &= \lim_{n\to\infty} \mathbb{E}[S_{\tau\wedge n} - (p-q)(\tau \wedge n) \mid S_0 = k] \\ &= \mathbb{E}\left[\lim_{n\to\infty} S_{\tau\wedge n} - (p-q) \lim_{n\to\infty} \tau \wedge n \mid S_0 = k\right] \\ &= \mathbb{E}[S_\tau - (p-q)\tau \mid S_0 = k], \end{aligned}$$

which gives

$$\begin{aligned} k &= \mathbb{E}[S_\tau - (p-q)\tau \mid S_0 = k] \\ &= \mathbb{E}[S_\tau \mid S_0 = k] - (p-q)\mathbb{E}[\tau \mid S_0 = k] \\ &= B \times \mathbb{P}(S_\tau = B \mid S_0 = k) + 0 \times \mathbb{P}(S_\tau = 0 \mid S_0 = k) - (p-q)\mathbb{E}[\tau \mid S_0 = k], \end{aligned}$$

i.e.

$$\begin{aligned} (p-q)\mathbb{E}[\tau \mid S_0 = k] &= B \times \mathbb{P}(S_\tau = B \mid S_0 = k) - k \\ &= B\frac{(q/p)^k - 1}{(q/p)^B - 1} - k, \end{aligned}$$

from (10.4.2), hence

$$\mathbb{E}[\tau \mid S_0 = k] = \frac{1}{p-q}\left(B\frac{(q/p)^k - 1}{(q/p)^B - 1} - k\right), \qquad k = 0, 1, \ldots, B,$$

which recovers (2.3.11).

In Table 10.1 we summarize the family of martingales used to treat the above problems.

Table 10.1 List of martingales

Probabilities		
Problem	Unbiased	Biased
Ruin probability	S_n	$\left(\frac{q}{p}\right)^{S_n}$
Mean game duration	$S_n^2 - n$	$S_n - (p-q)n$

Exercises

Exercise 10.1 Consider a sequence $(X_n)_{n\geq 1}$ of independent Bernoulli random variables with

$$\mathbb{P}(X_n = 1) = \mathbb{P}(X_n = -1) = 1/2, \qquad n \geq 1,$$

and the process $(M_n)_{n\in\mathbb{N}}$ defined by $M_0 := 0$ and

$$M_n := \sum_{k=1}^{n} 2^{k-1} X_k, \qquad n \geq 1.$$

See (Fig. 10.2), Note that when $X_1 = X_2 = \cdots = X_{n-1} = -1$ and $X_n = 1$, we have

$$M_n = -\sum_{k=1}^{n-1} 2^{k-1} + 2^{n-1} = -\frac{1 - 2^{n-1}}{1-2} + 2^{n-1} = 1, \qquad n \geq 1.$$

(a) Show that the process $(M_n)_{n\in\mathbb{N}}$ is a martingale.

(b) Is the random time

$$\tau := \inf\{n \geq 1 \ : \ M_n = 1\}$$

a stopping time?

(c) Consider the stopped process

$$M_{\tau\wedge n} := M_n \mathbb{1}_{\{n<\tau\}} + \mathbb{1}_{\{\tau\leq n\}} = \begin{cases} M_n = 1 - 2^n & \text{if } n < \tau, \\ M_\tau = 1 & \text{if } n \geq \tau, \end{cases}$$

$n \in \mathbb{N}$, See (Fig. 10.3). Give an interpretation of $(M_{n\wedge\tau})_{n\in\mathbb{N}}$ in terms of betting strategy for a gambler starting a game at $M_0 = 0$.

(d) Determine the two possible values of $M_{\tau\wedge n}$ and the probability distribution of $M_{\tau\wedge n}$ at any time $n \geq 1$.

(e) Show, using the result of Question (d), that we have

$$\mathbb{E}[M_{\tau\wedge n}] = 0, \qquad n \in \mathbb{N}.$$

(f) Show that the result of Question (e) can be recovered using the stopping time theorem.

Exercise 10.2 Let $(M_n)_{n\in\mathbb{N}}$ be a discrete-time *sub*martingale with respect to a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$, with $\mathcal{F}_0 = \{\emptyset, \Omega\}$, i.e. we have

$$M_n \leq \mathbb{E}[M_{n+1} \mid \mathcal{F}_n], \qquad n \geq 0.$$

Fig. 10.2 Possible paths of the process $(M_n)_{n\in\mathbb{N}}$

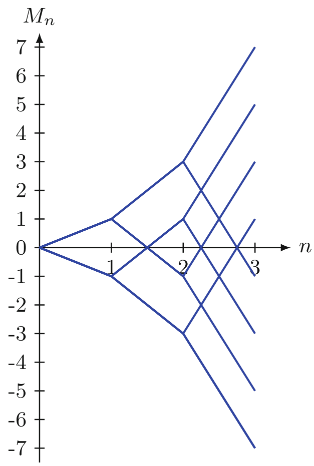
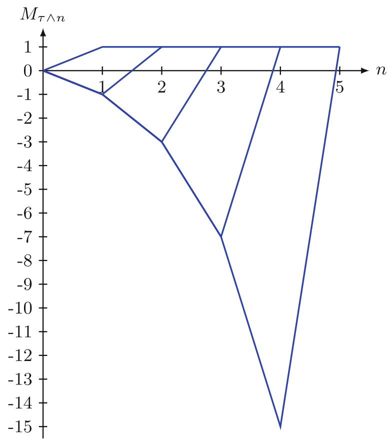


Fig. 10.3 Possible paths of the stopped process $(M_{\tau\wedge n})_{n\in\mathbb{N}}$



(a) Show that we have $\mathbb{E}[M_{n+1}] \geq \mathbb{E}[M_n]$, $n \geq 0$, i.e. a *sub*martingale has an *increasing* expectation.

(b) Show that independent increment processes whose increments have nonnegative expectation are examples of *sub*martingales.

(c) (Doob-Meyer decomposition) Show that there exists two processes $(N_n)_{n\in\mathbb{N}}$ and $(A_n)_{n\in\mathbb{N}}$ such that

 (i) $(N_n)_{n\in\mathbb{N}}$ is a martingale with respect to $(\mathcal{F}_n)_{n\in\mathbb{N}}$,

 (ii) $(A_n)_{n\in\mathbb{N}}$ is *non-decreasing*, i.e. $A_n \leq A_{n+1}$, *a.s.*, $n \in \mathbb{N}$,

 (iii) $(A_n)_{n\in\mathbb{N}}$ is predictable in the sense that A_n is $\mathcal{F}_{n-1}$-measurable, $n \in \mathbb{N}$, and

 (iv) $M_n = N_n + A_n$, $n \in \mathbb{N}$.

Hint: Let $A_0 = 0$ and

$$A_{n+1} := A_n + \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n], \quad n \geq 0,$$

and define $(N_n)_{n\in\mathbb{N}}$ in such a way that it satisfies the four required properties.

(d) Show that for all bounded stopping times σ and τ such that $\sigma \leq \tau$ *a.s.*, we have

$$\mathbb{E}[M_\sigma] \leq \mathbb{E}[M_\tau].$$

Hint: Use the Doob stopping time Theorem 10.6 for martingales and (10.3.3).

Exercise 10.3 Consider an asset price $(S_n)_{n=0,1,\ldots,N}$ which is a martingale under the risk-neutral measure $\mathbb{P}^*$, with respect to the filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$. Given the (convex) function $\phi(x) := (x - K)^+$, show that the price of an Asian option with payoff

$$\phi\left(\frac{S_1 + S_2 + \cdots + S_N}{N}\right)$$

is upper bounded by the price of the European *call* option with maturity N, i.e. show that

$$\mathbb{E}^*\left[\phi\left(\frac{S_1 + S_2 + \cdots + S_N}{N}\right)\right] \leq \mathbb{E}^*[\phi(S_N)].$$

Hint: Use in the following order:

(i) the convexity inequality

$$\phi\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right) \leq \frac{\phi(x_1) + \phi(x_2) + \cdots + \phi(x_n)}{n},$$

(ii) the martingale property of $(S_k)_{k\in\mathbb{N}}$,
(iii) the conditional Jensen inequality $\phi(\mathbb{E}[F \mid \mathcal{G}]) \leq \mathbb{E}[\phi(F) \mid \mathcal{G}]$,
(iv) the tower property of conditional expectations.

Exercise 10.4 A process $(M_n)_{n\in\mathbb{N}}$ is a *sub*martingale if it satisfies

$$M_k \leq \mathbb{E}[M_n \mid \mathcal{F}_k], \qquad k = 0, 1, \ldots, n.$$

(a) Show that the expectation $\mathbb{E}[M_n]$ of a *sub*martingale increases with time $n \in \mathbb{N}$.
(b) Consider the random walk given by $S_0 := 0$ and

$$S_n := \sum_{k=1}^{n} X_k = X_1 + X_2 + \cdots + X_n, \qquad n \geq 1,$$

where $(X_n)_{n\geq 1}$ is an i.i.d. Bernoulli sequence of $\{0, 1\}$-valued random variables with $\mathbb{P}(X_n = 1) = p$, $n \geq 1$. Under which condition on $\alpha \in \mathbb{R}$ is the process $(S_n - \alpha n)_{n\in\mathbb{N}}$ a *sub*martingale?

Exercise 10.5 Recall that a process $(M_n)_{n\in\mathbb{N}}$ is a *sub*martingale if it satisfies

$$M_k \leq \mathbb{E}[M_n \mid \mathcal{F}_k], \qquad k = 0, 1, \ldots, n.$$

(a) Show that any convex function $(\phi(M_n))_{n\in\mathbb{N}}$ of a martingale $(M_n)_{n\in\mathbb{N}}$ is itself a *sub*martingale. *Hint*: Use Jensen's inequality.
(b) Show that any convex nondecreasing function $\phi(M_n)$ of a *sub*martingale $(M_n)_{n\in\mathbb{N}}$ remains a *sub*martingale.

Problem 10.6 (a) Consider $(M_n)_{n\in\mathbb{N}}$ a *nonnegative martingale*. For any $x > 0$, let

$$\tau_x := \inf\{n \geq 0 \ : \ M_n \geq x\}.$$

Show that the random time τ_x is a stopping time.
(b) Show that for all $n \geq 0$ we have

$$\mathbb{P}\left(\max_{k=0,1,\ldots,n} M_k \geq x\right) \leq \frac{\mathbb{E}[M_n]}{x}. \tag{10.5.1}$$

Hint: Use the Markov inequality and the Doob stopping time Theorem 10.6 for the stopping time τ_x.
(c) Show that (10.5.1) remains valid when $(M_n)_{n\in\mathbb{N}}$ is a nonnegative *sub*martingale. *Hint*: Use the Doob stopping time theorem for *sub*martingales as in Exercise 10.2-(d).
(d) Show that for any $n \geq 0$ we have

$$\mathbb{P}\left(\max_{k=0,1,\ldots,n} M_k \geq x\right) \leq \frac{\mathbb{E}[(M_n)^2]}{x^2}, \qquad x > 0.$$

(e) Show that more generally we have

$$\mathbb{P}\left(\max_{k=0,1,\ldots,n} M_k \geq x\right) \leq \frac{\mathbb{E}[|M_n|^p]}{x^p}, \qquad x > 0,$$

for all $n \geq 0$ and $p \geq 1$.
(f) Given $(Y_n)_{n\geq 1}$ a sequence of centered independent random variables with same mean $\mathbb{E}[Y_n] = 0$ and variance $\sigma^2 = \mathrm{Var}[Y_n]$, $n \geq 1$, consider the random walk $S_n = Y_1 + Y_2 + \cdots + Y_n$, $n \geq 1$, with $S_0 = 0$.
Show that for all $n \geq 0$ we have

$$\mathbb{P}\left(\max_{k=0,1,\ldots,n} |S_k| \geq x\right) \leq \frac{n\sigma^2}{x^2}, \qquad x > 0.$$

(g) Show that for any (not necessarily nonnegative) *sub*martingale we have

$$\mathbb{P}\left(\max_{k=0,1,\ldots,n} M_k \geq x\right) \leq \frac{\mathbb{E}[M_n^+]}{x}, \qquad x > 0,$$

where $z^+ = \max(z, 0)$, $z \in \mathbb{R}$.

(h) A process $(M_n)_{n\in\mathbb{N}}$ is a *super*martingale[5] if it satisfies

$$\mathbb{E}[M_n \mid \mathcal{F}_k] \leq M_k, \qquad k = 0, 1, \ldots, n.$$

Show that for any *nonnegative super*martingale we have

$$\mathbb{P}\left(\max_{k=0,1,\ldots,n} M_k \geq x\right) \leq \frac{\mathbb{E}[M_0]}{x}, \qquad x > 0.$$

(i) Show that for any *nonnegative submartingale* $(M_n)_{n\in\mathbb{N}}$ and any convex nondecreasing nonnegative function ϕ we have

$$\mathbb{P}\left(\max_{k=0,1,\ldots,n} \phi(M_k) \geq x\right) \leq \frac{\mathbb{E}[\phi(M_n)]}{x}, \qquad x > 0.$$

Hint: Consider the stopping time

$$\tau_x^\phi := \inf\{n \geq 0 \; : \; M_n \geq x\}.$$

(j) Give an example of a nonnegative *super*martingale which is *not* a martingale.

[5] "This obviously inappropriate nomenclature was chosen under the malign influence of the noise level of radio's SUPERman program, a favorite supper-time program of Doob's son during the writing of [Doo53]", cf. [Doo84], historical notes, p. 808.

Chapter 11
Spatial Poisson Processes

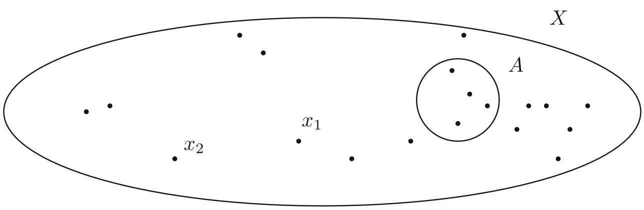
Spatial Poisson process are typically used to model the random scattering of configuration points within a plane or a three-dimensional space X. In case $X = \mathbb{R}_+$ is the real half line, these random points can be identified with the jump times $(T_k)_{k\geq 1}$ of the standard Poisson process $(N_t)_{t\in\mathbb{R}_+}$ introduced in Sect. 9.1. However, in contrast with the previous chapter, no time ordering is a priori imposed here on the index set X.

11.1 Spatial Poisson (1781–1840) Processes

In this section we present the construction of spatial Poisson processes on a space of configurations of $X \subset \mathbb{R}^d$, $d \geq 1$. The set

$$\Omega^X := \left\{\omega := (x_i)_{i=1}^N \subset X,\ N \in \mathbb{N} \cup \{\infty\}\right\},$$

is called the space of configurations on $X \subset \mathbb{R}^d$. The next figure illustrates a given configuration $\omega \in \Omega^X$.



N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4_11

Given A a (measurable) subset of X, we let

$$\omega(A) = \#\{x \in \omega \ : \ x \in A\} = \sum_{x \in \omega} \mathbb{1}_A(x)$$

denote the number of configuration points in ω that are contained in the set A.

Given $\rho : X \longrightarrow \mathbb{R}_+$ a nonnegative function, the Poisson probability measure $\mathbb{P}_\sigma^X$ with intensity $\sigma(dx) = \rho(x)dx$ on X is the only probability measure on Ω^X satisfying

(i) For any (measurable) subset A of X such that

$$\sigma(A) = \int_A \rho(x)dx = \int_{\mathbb{R}^d} \mathbb{1}_A(x)\rho(x)dx < \infty,$$

the number $\omega(A)$ of configuration points contained in A is a Poisson random variable with intensity $\sigma(A)$, i.e.

$$\mathbb{P}_\sigma^X(\omega \in \Omega^X \ : \ \omega(A) = n) = \mathrm{e}^{-\sigma(A)} \frac{(\sigma(A))^n}{n!}, \qquad n \in \mathbb{N}.$$

(ii) In addition, if $A_1, A_2, \ldots, A_n$ are disjoint subsets of X with $\sigma(A_k) < \infty$, $k = 1, 2, \ldots, n$, the $\mathbb{N}^n$-valued random vector

$$\omega \longmapsto (\omega(A_1), \ldots, \omega(A_n)), \qquad \omega \in \Omega^X,$$

is made of independent random variables for all $n \geq 1$.

In the remaining of this chapter we will assume that $\sigma(X) < \infty$ for simplicity. The Poisson measure $\mathbb{P}_\sigma^X$ can also be defined as

$$\mathbb{E}_{\mathbb{P}_\sigma^X}[F] = \mathrm{e}^{-\sigma(X)} \sum_{n=0}^{\infty} \frac{1}{n!} \int_{X^n} f_n(x_1, x_2, \ldots, x_n)\sigma(dx_1) \cdots \sigma(dx_n) \qquad (11.1.1)$$

for F written as

$$F(\omega) = \sum_{n=0}^{\infty} \mathbb{1}_{\{\omega(X)=n\}} f_n(x_1, x_2, \ldots, x_n)$$

where f_n is a symmetric integrable function of $\omega = \{x_1, x_2, \ldots, x_n\}$ when $\omega(X) = n$, $n \geq 1$, cf. e.g. Proposition 6.1.3 and Sect. 6.1 in [Pri09].

By applying the above to

$$F(\omega) = \mathbb{1}_{\{\omega(X)=n\}} \mathbb{1}_{A^n}(x_1, x_2, \ldots, x_n),$$

we find that the conditional distribution of $\omega = \{x_1, x_2, \ldots, x_n\}$ given that $\omega(X) = n$ is given by the formula

$$\begin{aligned}
\mathbb{P}_\sigma^X(\{x_1, \ldots, x_n\} \subset A \mid \omega(X) = n) &= \frac{\mathbb{P}_\sigma^X(\{x_1, \ldots, x_n\} \subset A \text{ and } \omega(X) = n)}{\mathbb{P}_\sigma^X(\omega(X) = n)} \\
&= \frac{1}{\mathbb{P}_\sigma^X(\omega(X) = n)} \mathbb{E}_{\mathbb{P}_\sigma^X}[\mathbb{1}_{\{\omega(X)=n\}} \mathbb{1}_{A^n}(x_1, x_2, \ldots, x_n)] \\
&= \left(\frac{\sigma(A)}{\sigma(X)}\right)^n. \qquad (11.1.2)
\end{aligned}$$

In many applications the intensity function $\rho(x)$ will be constant, i.e. $\rho(x) = \lambda > 0$, $x \in X$, where $\lambda > 0$ is called the intensity parameter, and

$$\sigma(A) = \lambda \int_A dx = \lambda \int_X \mathbb{1}_A(x) dx$$

represents the surface or volume of A in $\mathbb{R}^d$. In this case, (11.1.2) can be used to show that the random points $\{x_1, \ldots, x_n\}$ are uniformly distributed on A^n given that $\{\omega(A) = n\}$.

11.2 Poisson Stochastic Integrals

In the next proposition we consider the Poisson stochastic integral defined as

$$\int_X f(x)\omega(dx) := \sum_{x \in \omega} f(x),$$

for f an integrable function on X, and we compute its first and second order moments and cumulants via its characteristic function.

Proposition 11.1 *Let f be an integrable function on X. We have*

$$\mathbb{E}_{\mathbb{P}_\sigma^X}\left[\exp\left(i \int_X f(x)\omega(dx)\right)\right] = \exp\left(\int_X (e^{if(x)} - 1)\sigma(dx)\right).$$

Proof We assume that $\sigma(X) < \infty$. By (11.1.1) we have

$$\begin{aligned}
&\mathbb{E}_{\mathbb{P}_\sigma^X}\left[\exp\left(i \int_X f(x)\omega(dx)\right)\right] \\
&\qquad = e^{-\sigma(X)} \sum_{n=0}^{\infty} \frac{1}{n!} \int_X \cdots \int_X e^{i(f(x_1)+\cdots+f(x_n))} \sigma(dx_1) \cdots \sigma(dx_n).
\end{aligned}$$

$$= e^{-\sigma(X)} \sum_{n=0}^{\infty} \frac{1}{n!} \left(\int_X e^{if(x)} \sigma(dx) \right)^n$$
$$= \exp\left(\int_X (e^{if(x)} - 1)\sigma(dx) \right). \qquad \square$$

The characteristic function allows us to compute the expectation of $\int_X f(x)\omega(dx)$, as

$$\begin{aligned} \mathbb{E}_{\mathbb{P}_\sigma^X}\left[\int_X f(x)\omega(dx) \right] &= -i \frac{d}{d\varepsilon} \mathbb{E}_{\mathbb{P}_\sigma^X}\left[\exp\left(i\varepsilon \int_X f(x)\omega(dx) \right) \right]_{|\varepsilon=0} \\ &= -i \frac{d}{d\varepsilon} \exp\left(\int_X (e^{i\varepsilon f(x)} - 1)\sigma(dx) \right)_{|\varepsilon=0} \\ &= \int_X f(x)\sigma(dx), \end{aligned}$$

for f an integrable function on X. As a consequence, the *compensated* Poisson stochastic integral

$$\int_X f(x)\omega(dx) - \int_X f(x)\sigma(dx)$$

is a *centered* random variable, i.e. we have

$$\mathbb{E}_{\mathbb{P}_\sigma^X}\left[\int_X f(x)\omega(dx) - \int_X f(x)\sigma(dx) \right] = 0.$$

The variance can be similarly computed as

$$\mathbb{E}_{\mathbb{P}_\sigma^X}\left[\left(\int_X f(x)(\omega(dx) - \sigma(dx)) \right)^2 \right] = \int_X |f(x)|^2 \sigma(dx),$$

for all f in the space $L^2(X, \sigma)$ of functions which are square-integrable on X with respect to $\sigma(dx)$.

More generally, the logarithmic generating function

$$\log \mathbb{E}_{\mathbb{P}_\sigma^X}\left[\exp\left(\int_X f(x)\omega(dx) \right) \right] = \int_X (e^{f(x)} - 1)\sigma(dx) = \sum_{n=1}^{\infty} \frac{1}{n!} \int_X f^n(x)\sigma(dx),$$

shows that the cumulants of $\int_X f(x)\omega(dx)$ are given by

$$\kappa_n = \int_X f^n(x)\sigma(dx), \qquad n \geq 1. \tag{11.2.1}$$

11.3 Transformations of Poisson Measures

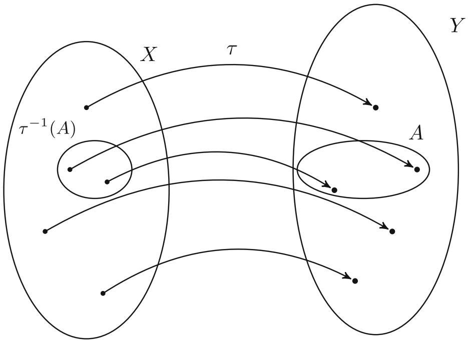
Consider a mapping $\tau : (X, \sigma) \longrightarrow (Y, \mu)$, and let

$$\tau_* : \Omega^X \longrightarrow \Omega^Y$$

be the transformed configuration defined by

$$\tau_*(\omega) := \{\tau(x) \ : \ x \in \omega\}, \qquad \omega \in \Omega^X,$$

as illustrated in the following figure.



Proposition 11.2 *The random configuration*

$$\begin{aligned} \Omega^X \ : \longrightarrow \ & \Omega^Y \\ \omega \longmapsto \ & \tau_*(\omega) \end{aligned}$$

has the Poisson distribution $\mathbb{P}_\mu^Y$ with intensity μ on Y, where μ is defined by

$$\mu(A) := \int_X \mathbb{1}_A(\tau(x))\sigma(dx) = \int_X \mathbb{1}_{\tau^{-1}(A)}(x)\sigma(dx) = \sigma(\tau^{-1}(A)),$$

for A a (measurable) subset of X.

Proof We have

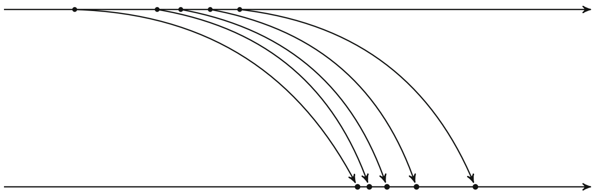
$$\begin{aligned} \mathbb{P}_\sigma^X(\tau_*\omega(A) = n) &= \mathbb{P}_\sigma^X(\omega(\tau^{-1}(A)) = n) \\ &= \mathrm{e}^{-\sigma(\tau^{-1}(A))} \frac{(\sigma(\tau^{-1}(A)))^n}{n!} \\ &= \mathrm{e}^{-\mu(A)} \frac{(\mu(A))^n}{n!}. \end{aligned}$$

More generally we can check that for all families $A_1, A_2, \ldots, A_n$ of disjoint subsets of X and $k_1, k_2, \ldots, k_n \in \mathbb{N}$, we have

$$\begin{aligned}
&\mathbb{P}^X_\sigma(\{\omega \in \Omega^X : \tau_*\omega(A_1) = k_1, \ldots, \tau_*\omega(A_n) = k_n\}) \\
&\quad = \prod_{i=1}^n \mathbb{P}^X_\sigma(\{\tau_*\omega(A_i) = k_i\}) \\
&\quad = \prod_{i=1}^n \mathbb{P}^X_\sigma(\{\omega(\tau^{-1}(A_i)) = k_i\}) \\
&\quad = \exp\left(-\sum_{i=1}^n \sigma(\tau^{-1}(A_i))\right) \prod_{i=1}^n \frac{(\sigma(\tau^{-1}(A_i)))^{k_i}}{k_i!} \\
&\quad = \exp\left(-\sum_{i=1}^n \mu(A_i)\right) \prod_{i=1}^n \frac{(\mu(A_i))^{k_i}}{k_i!} \\
&\quad = \prod_{i=1}^n \mathbb{P}^Y_\mu(\{\omega(A_i) = k_i\}) \\
&\quad = \mathbb{P}^Y_\mu(\{\omega(A_1) = k_1, \ldots, \omega(A_n) = k_n\}).
\end{aligned}$$

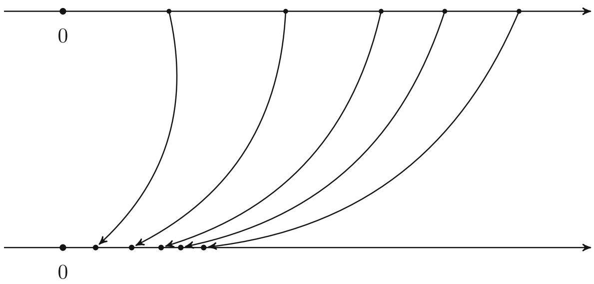
□

The next figure illustrates the transport of measure in the case of Gaussian intensities on $X = \mathbb{R}$.



For example in the case of a flat intensity $\rho(x) = \lambda$ on $X = \mathbb{R}_+$ the intensity becomes doubled under the mapping $\tau(x) = x/2$, since

$$\begin{aligned}
\mathbb{P}^X_\sigma(\tau_*\omega([0,t]) = n) &= \mathbb{P}^X_\sigma(\omega(\tau^{-1}([0,t])) = n) \\
&= \mathrm{e}^{-\sigma(\tau^{-1}([0,t]))} \frac{(\sigma(\tau^{-1}([0,t])))^n}{n!} \\
&= \mathrm{e}^{-2\lambda t}(2\lambda t)^n/n!.
\end{aligned}$$



Exercises

Exercise 11.1 Consider a standard Poisson process $(N_t)_{t\in\mathbb{R}_+}$ on $\mathbb{R}_+$ with intensity $\lambda = 2$ and jump times $(T_k)_{k\geq 1}$. Compute

$$\mathbb{E}[T_1 + T_2 + T_3 \mid N_2 = 2].$$

Exercise 11.2 Consider a spatial Poisson process on $\mathbb{R}^2$ with intensity $\lambda = 0.5$ per square meter. What is the probability that there are 10 events within a circle of radius 3 m.

Exercise 11.3 Some living organisms are distributed in space according to a Poisson process of intensity $\theta = 0.6$ units per mm^3. Compute the probability that more than two living organisms are found within a 10 mm^3 volume.

Exercise 11.4 Defects are present over a piece of fabric according to a Poisson process with intensity of one defect per piece of fabric. Both halves of the piece is checked separately. What is the probability that both inspections record at least one defect?

Exercise 11.5 Let $\lambda > 0$ and suppose that N points are independently and uniformly distributed over the interval $[0, N]$. Determine the probability distribution for the number of points in the interval $[0, \lambda]$ as $N \to \infty$.

Exercise 11.6 Suppose that $X(A)$ is a spatial Poisson process of discrete items scattered on the plane $\mathbb{R}^2$ with intensity $\lambda = 0.5$ per square meter. We let

$$D((x, y), r) = \{(u, v) \in \mathbb{R}^2 \ : \ (x - u)^2 + (y - v)^2 \leq r^2\}$$

denote the disc with radius r centered at (x, y) in $\mathbb{R}^2$. No evaluation of numerical expressions is required in this exercise.

(a) What is the probability that 10 items are found within the disk $D((0, 0), 3)$ with radius 3 meters centered at the origin?

(b) What is the probability that 5 items are found within the disk $D((0, 0), 3)$ *and* 3 items are found within the disk $D((x, y), 3)$ with $(x, y) = (7, 0)$?
(c) What is the probability that 8 items are found anywhere within $D((0, 0), 3) \bigcup D((x, y), 3)$ with $(x, y) = (7, 0)$?
(d) Given that 5 items are found within the disk $D((0, 0), 1)$, what is the probability that 3 of them are located within the disk $D((1/2, 0), 1/2)$ centered at $(1/2, 0)$ with radius $1/2$?

Chapter 12
Reliability Theory

This chapter consists in a short review of survival probabilities based on failure rate and reliability functions, in connection with Poisson processes having a time-dependent intensity.

12.1 Survival Probabilities

Let $\tau : \Omega \longrightarrow \mathbb{R}_+$ denote (random) the lifetime of an entity, and let $\mathbb{P}(\tau \geq t)$ denote its probability of surviving at least t years, $t > 0$. The probability of surviving up to a (deterministic) time T, given that the entity has already survived up to time t, is given by

$$\mathbb{P}(\tau > T \mid \tau > t) = \frac{\mathbb{P}(\tau > T \text{ and } \tau > t)}{\mathbb{P}(\tau > t)}$$
$$= \frac{\mathbb{P}(\tau > T)}{\mathbb{P}(\tau > t)}, \qquad 0 \leq t \leq T.$$

Let now

$$\lambda(t) := \lim_{h \searrow 0} \frac{\mathbb{P}(\tau < t + h \mid \tau > t)}{h}, \qquad t \in \mathbb{R}_+,$$

denote the *failure rate function* associated to τ. Letting $A = \{\tau < t + h\}$ and $B = \{\tau > t\}$ we note that $(\Omega \setminus A) \subset B$, hence $A \cap B = B \setminus A^c$, and

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4_12

$$\begin{aligned}
\lambda(t) &= \lim_{h \searrow 0} \frac{\mathbb{P}(\tau < t + h \mid \tau > t)}{h} \\
&= \frac{1}{\mathbb{P}(\tau > t)} \lim_{h \searrow 0} \frac{\mathbb{P}(\tau < t + h \text{ and } \tau > t)}{h} \\
&= \frac{1}{\mathbb{P}(\tau > t)} \lim_{h \searrow 0} \frac{\mathbb{P}(\tau > t) - \mathbb{P}(\tau > t + h)}{h} \\
&= -\frac{d}{dt} \log \mathbb{P}(\tau > t) \\
&= -\frac{1}{\mathbb{P}(\tau > t)} \frac{d}{dt} \mathbb{P}(\tau > t) \\
&= -\frac{1}{R(t)} \frac{d}{dt} R(t),
\end{aligned} \tag{12.1.1}$$

where the *reliability function* $R(t)$ is defined by

$$R(t) := \mathbb{P}(\tau > t), \qquad t \in \mathbb{R}_+.$$

This yields

$$R'(t) = -\lambda(t) R(t),$$

with $R(0) = \mathbb{P}(\tau > 0) = 1$, which has for solution

$$R(t) = \mathbb{P}(\tau > t) = R(0) \exp\left(-\int_0^t \lambda(u) du\right) = \exp\left(-\int_0^t \lambda(u) du\right), \tag{12.1.2}$$

$t \in \mathbb{R}_+$. Hence we have

$$\mathbb{P}(\tau > T \mid \tau > t) = \frac{R(T)}{R(t)} = \exp\left(-\int_t^T \lambda(u) du\right), \qquad t \in [0, T]. \tag{12.1.3}$$

In case the failure rate function $\lambda(t) = c$ is constant we recover the memoryless property of the exponential distribution with parameter $c > 0$, cf. (9.2.3).

Relation (12.1.2) can be recovered informally as

$$\mathbb{P}(\tau > T) = \prod_{0<t<T} \mathbb{P}(\tau > t + dt \mid \tau > t) = \prod_{0<t<T} \exp\left(-\lambda(t) dt\right),$$

which yields

$$\mathbb{P}(\tau > t) = \exp\left(-\int_0^t \lambda(s) ds\right), \qquad t \in \mathbb{R}_+,$$

in the limit.

12.2 Poisson Process with Time-Dependent Intensity

Recall that the random variable τ has the exponential distribution with parameter $\lambda > 0$ if

$$\mathbb{P}(\tau > t) = \mathrm{e}^{-\lambda t}, \qquad t \geq 0,$$

cf. (1.5.3). Given $(\tau_n)_{n\geq 0}$ a sequence of *i.i.d.* exponentially distributed random variables, letting

$$T_n = \tau_0 + \cdots + \tau_{n-1}, \qquad n \geq 1,$$

and

$$N_t = \sum_{n\geq 1} \mathbb{1}_{[T_n,\infty)}(t), \qquad t \in \mathbb{R}_+,$$

defines the standard Poisson process with intensity $\lambda > 0$ of Sect. 9.1 and we have

$$\mathbb{P}(N_t - N_s = k) = \mathrm{e}^{-\lambda(t-s)} \frac{(\lambda(t-s))^k}{k!}, \qquad k \geq 0.$$

The intensity of the Poisson process can in fact made time-dependent. For example under the time change

$$X_t = N_{\int_0^t \lambda(s)ds}$$

where $(\lambda(u))_{u\in\mathbb{R}_+}$ is a deterministic function of time, we have

$$\mathbb{P}(X_t - X_s = k) = \frac{\left(\int_s^t \lambda(u)du\right)^k}{k!} \exp\left(-\int_s^t \lambda(u)du\right), \qquad k \geq 0.$$

In this case we have

$$\mathbb{P}(X_{t+h} - X_t = 0) = \mathrm{e}^{-\lambda(t)h} + o(h) = 1 - \lambda(t)h + o(h), \qquad h \searrow 0, \quad (12.2.1)$$

and

$$\mathbb{P}(X_{t+h} - X_t = 1) = 1 - \mathrm{e}^{-\lambda(t)h} + o(h) \simeq \lambda(t)h, \qquad h \searrow 0, \quad (12.2.2)$$

which can also viewed as a pure birth process with time-dependent intensity. Letting τ_0 denote the first jump time of $(X_t)_{t\in\mathbb{R}_+}$, we have

$$R(t) = \mathbb{P}(\tau_0 > t) = \mathbb{P}(X_t = 0) = \exp\left(-\int_0^t \lambda(u)du\right), \qquad t \geq 0, \quad (12.2.3)$$

hence by (12.1.3) we find

$$\begin{aligned}\mathbb{P}(X_{t+h} = 0 \mid X_t = 0) &= \mathbb{P}(\tau_0 > t + h \mid \tau_0 > t)\\ &= \mathrm{e}^{-\lambda(t)h} + o(h) = 1 - \lambda(t)h + o(h), \qquad h \searrow 0,\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}(X_{t+h} \geq 1 \mid X_t = 0) &= \mathbb{P}(\tau_0 < t + h \mid \tau_0 > t) = 1 - \mathbb{P}(\tau_0 > t + h \mid \tau_0 > t)\\ &= 1 - \mathrm{e}^{-\lambda h} \simeq \lambda(t)h + o(h), \qquad h \searrow 0,\end{aligned}$$

which coincide respectively with $\mathbb{P}(X_{t+h} - X_t = 0)$ and $\mathbb{P}(X_{t+h} - X_t = 1)$ in (12.2.1) and (12.2.2) above, as $(X_t)_{t\in\mathbb{R}_+}$ has independent increments.

Cox Processes

The intensity process $\lambda(s)$ can also be made random. In this case, $(X_t)_{t\in\mathbb{R}_+}$ is called a *Cox process* and it may not have independent increments. For example, assume that $(\lambda_u)_{u\in\mathbb{R}_+}$ is a two-state Markov chain on $\{0, \lambda\}$, with transitions

$$\mathbb{P}(\lambda_{t+h} = \lambda \mid \lambda_t = 0) = \alpha h, \qquad h \searrow 0,$$

and

$$\mathbb{P}(\lambda_{t+h} = 0 \mid \lambda_t = \lambda) = \beta h, \qquad h \searrow 0.$$

In this case the probability distribution of N_t can be explicitly computed, cf. Chap. VI-7 in [KT81].

Renewal Processes

A *renewal process* is a counting process $(N_t)_{t\in\mathbb{R}_+}$ given by

$$N_t = \sum_{k\geq 1} k\mathbb{1}_{[T_k, T_{k+1})}(t) = \sum_{k\geq 1} \mathbb{1}_{[T_k,\infty)}(t), \qquad t \in \mathbb{R}_+,$$

in which $\tau_k = T_{k+1} - T_k$, $k \in \mathbb{N}$, is a sequence of independent identically distributed random variables. In particular, Poisson processes are renewal processes.

12.3 Mean Time to Failure

The mean time to failure is given, from (12.1.1), by

$$\begin{aligned}\mathbb{E}[\tau] &= \int_0^\infty t\frac{d}{dt}\mathbb{P}(\tau < t)dt = -\int_0^\infty t\frac{d}{dt}\mathbb{P}(\tau > t)dt\\ &= -\int_0^\infty tR'(t)dt = \int_0^\infty R(t)dt, \qquad (12.3.1)\end{aligned}$$

provided that $\lim_{t\searrow 0} tR(t) = 0$. For example when τ has the distribution function (12.2.3) we get

$$\mathbb{E}[\tau] = \int_0^\infty R(t)dt = \int_0^\infty \exp\left(-\int_0^t \lambda(u)du\right)dt.$$

In case the function $\lambda(t) = \lambda > 0$ is constant we recover the mean value

$$\mathbb{E}[\tau] = \int_0^\infty e^{-\lambda t}dt = \frac{1}{\lambda}$$

of the exponential distribution with parameter $\lambda > 0$.

Exercise

Exercise 12.1 Assume that the random time τ has the *Weibull* distribution with probability density

$$f_\beta(x) = \beta \mathbb{1}_{[0,\infty)} x^{\beta-1} e^{-t^\beta}, \qquad x \in \mathbb{R},$$

where $\beta > 0$ is a called the shape parameter.

(a) Compute the distribution function F_β of the Weibull distribution.
(b) Compute the reliability function $R(t) = \mathbb{P}(\tau > t)$.
(c) Compute the failure rate function $\lambda(t)$.
(d) Compute the mean time to failure.

Appendix A
Some Useful Identities

Here we present a summary of algebraic identities that are used in this text.

Indicator functions

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ \\ 0 & \text{if } x \notin A. \end{cases} \qquad \mathbb{1}_{[a,b]}(x) = \begin{cases} 1 & \text{if } a \le x \le b, \\ \\ 0 & \text{otherwise.} \end{cases}$$

Binomial coefficients

$$\binom{n}{k} := \frac{n!}{(n-k)!k!}, \qquad k = 0, 1, \ldots, n.$$

Exponential series

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \qquad x \in \mathbb{R}. \tag{A.1}$$

Geometric sum

$$\sum_{k=0}^{n} r^k = \frac{1 - r^{n+1}}{1 - r}, \qquad r \neq 1. \tag{A.2}$$

Geometric series

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1 - r}, \qquad -1 < r < 1. \tag{A.3}$$

Differentiation of geometric series

$$\sum_{k=1}^{\infty} k r^{k-1} = \frac{\partial}{\partial r} \sum_{k=0}^{\infty} r^k = \frac{\partial}{\partial r} \frac{1}{1 - r} = \frac{1}{(1 - r)^2}, \qquad -1 < r < 1. \tag{A.4}$$

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4

Binomial identities

$$\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} = (a+b)^n.$$

$$\sum_{k=0}^{n}\binom{n}{k} = 2^n. \tag{A.5}$$

$$\sum_{k=1}^{n}k\binom{n}{k} = n2^{n-1}.$$

$$\begin{aligned}\sum_{k=0}^{n}k\binom{n}{k}a^k b^{n-k} &= \sum_{k=1}^{n}\frac{n!}{(n-k)!(k-1)!}a^k b^{n-k}\\ &= \sum_{k=0}^{n-1}\frac{n!}{(n-1-k)!k!}a^{k+1}b^{n-1-k}\\ &= n\sum_{k=0}^{n-1}\binom{n-1}{k}a^{k+1}b^{n-1-k}\\ &= na(a+b)^{n-1}, \qquad n \geq 1,\end{aligned}$$

$$\begin{aligned}\sum_{k=0}^{n}k\binom{n}{k}a^k b^{n-k} &= a\frac{\partial}{\partial a}\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}\\ &= a\frac{\partial}{\partial a}(a+b)^n\\ &= na(a+b)^{n-1}, \qquad n \geq 1.\end{aligned}$$

Sums of integers

$$\sum_{k=1}^{n}k = \frac{n(n+1)}{2}. \tag{A.6}$$

$$\sum_{k=1}^{n}k^2 = \frac{n(n+1)(2n+1)}{6}. \tag{A.7}$$

Taylor expansion

$$(1+x)^{\alpha} = \sum_{k=0}^{\infty}\frac{x^k}{k!}\alpha(\alpha-1)\times\cdots\times(\alpha-(k-1)). \tag{A.8}$$

Differential equation

The solution of $f'(t) = cf(t)$ is given by $f(t) = f(0)\mathrm{e}^{ct}, \quad t \in \mathbb{R}_+.$ (A.9)

Appendix B
Solutions to Selected Exercises and Problems

Chapter 1 - Probability Background

Exercise 1.2 We write

$$Z = \sum_{k=1}^{N} X_k$$

where $\mathbb{P}(N = n) = 1/6$, $n = 1, 2, \ldots, 6$, and X_k is a Bernoulli random variable with parameter $1/2$, $k = 1, 2, \ldots, 6$.

(a) We have

$$\mathbb{E}[Z] = \mathbb{E}[\mathbb{E}[Z \mid N]] = \sum_{n=1}^{6} \mathbb{E}[Z \mid N = n]\mathbb{P}(N = n) = \frac{7}{4}, \tag{B.1}$$

where we applied (A.6). Concerning the variance, we have

$$\mathbb{E}[Z^2] = \mathbb{E}[\mathbb{E}[Z^2 \mid N]] = \frac{14}{3}, \tag{B.2}$$

where we used (A.7) and (B.1), hence

$$\mathrm{Var}[Z] = \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2 = \frac{14}{3} - \frac{49}{16} = \frac{77}{48}. \tag{B.3}$$

(b) We find

$$\mathbb{P}(Z = l) = \frac{1}{6} \sum_{n=\max(1,l)}^{6} \binom{n}{l} \left(\frac{1}{2}\right)^n, \qquad l = 0, 1, \ldots, 6.$$

(c) We have

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4

$$\mathbb{E}[Z] = \sum_{l=0}^{6} l\mathbb{P}(Z = l) = \frac{1}{6}\sum_{l=0}^{6} l \sum_{n=l}^{6} \left(\frac{1}{2}\right)^n \binom{n}{l} = \frac{7}{4},$$

which recovers (A.7), and where we applied (A.5). We also have

$$\mathbb{E}[Z^2] = \sum_{l=0}^{6} l^2\mathbb{P}(Z = l) = \frac{14}{3},$$

which recovers (B.3).

Exercise 1.3

(a) We assume that the sequence of Bernoulli trials is represented by a family $(X_k)_{k\geq 1}$ of independent Bernoulli random variables with distribution $\mathbb{P}(X_k = 1) = p$, $\mathbb{P}(X_k = 0) = 1 - p$, $k \geq 1$. We have

$$Z = X_1 + X_2 + \cdots + X_N = \sum_{k=1}^{N} X_k,$$

and, since $\mathbb{E}[X_k] = p$,

$$\mathbb{E}[Z] = \sum_{n=0}^{\infty} \left(\sum_{k=1}^{n} \mathbb{E}[X_k]\right) \mathbb{P}(N = n) = p\sum_{n=0}^{\infty} n\mathbb{P}(N = n) = p\mathbb{E}[N].$$

Next, the expectation of the Poisson random variable N with parameter $\lambda > 0$ is given as in (1.6.4) by

$$\mathbb{E}[N] = \sum_{n=0}^{\infty} n\mathbb{P}(N = n) = \mathrm{e}^{-\lambda}\sum_{n=0}^{\infty} n\frac{\lambda^n}{n!} = \lambda\mathrm{e}^{-\lambda}\sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = \lambda\mathrm{e}^{-\lambda}\mathrm{e}^{\lambda} = \lambda, \tag{B.4}$$

where we used the exponential series (A.1), hence $\mathbb{E}[Z] = p\lambda$. Concerning the variance we have, since $\mathbb{E}[X_k^2] = p$ we find

$$\mathbb{E}[Z^2] = \mathbb{E}\left[\left(\sum_{k=1}^{N} X_k\right)^2\right] = p(1 - p)\mathbb{E}[N] + p^2\mathbb{E}[N^2].$$

Next, we have

$$\mathbb{E}[N^2] = \sum_{n=0}^{\infty} n^2\mathbb{P}(N = n) = \mathrm{e}^{-\lambda}\sum_{n=0}^{\infty} n^2\frac{\lambda^n}{n!}\lambda + \lambda^2,$$

hence

$$\mathrm{Var}[N] = \mathbb{E}[N^2] - (\mathbb{E}[N])^2 = \lambda, \tag{B.5}$$

and $\mathrm{Var}[Z] = \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2 = p\lambda$.

(b) For $l \in \mathbb{N}$, using (1.3.1) with $B = \{Z = l\}$ and the fact that $\sum_{k=1}^{n} X_k$ has a binomial distribution with parameter (n, p), we have

$$\mathbb{P}(Z = l) = \sum_{n=0}^{\infty} \mathbb{P}\left(\sum_{k=1}^{N} X_k = l \,\Big|\, N = n\right) \mathbb{P}(N = n) = \frac{(\lambda p)^l}{l!} \mathrm{e}^{-p\lambda},$$

hence Z has a Poisson distribution with parameter $p\lambda$.

(c) From Question (b), Z is a Poisson random variable with parameter $p\lambda$, hence from (B.4) and (B.5) we have $\mathbb{E}[Z] = \mathrm{Var}[Z] = p\lambda$.

Exercise 1.5 Since U is uniformly distributed given L over the interval $[0, L]$, we have

$$f_{U|L=y}(x) = \frac{1}{y} \mathbb{1}_{[0,y]}(x), \qquad x \in \mathbb{R}, \quad y > 0,$$

hence by the definition (1.5.7) of the conditional density $f_{U|L=y}(x)$ we have

$$f_{(U,L)}(x, y) = f_{U|L=y}(x) f_L(y) = \mathbb{1}_{[0,y]}(x)\mathbb{1}_{[0,\infty)}(y)\mathrm{e}^{-y}. \tag{B.6}$$

Next, we from (B.6) we get

$$f_{(U,L-U)}(x, z) = f_{(U,L)}(x, x + z) = \mathbb{1}_{[0,\infty)}(x)\mathbb{1}_{[0,\infty)}(z)\mathrm{e}^{-x-z}.$$

Exercise 1.6

(a) Assuming that X and Y are independent Poisson random variables with parameters λ and μ, we have

$$\mathbb{P}(X + Y = n) = \sum_{k=0}^{n} \mathbb{P}(X = k \text{ and } X + Y = n) = \mathrm{e}^{-\lambda-\mu} \frac{(\lambda + \mu)^n}{n!}, \tag{B.7}$$

hence $X + Y$ has a Poisson distribution with parameter $\lambda + \mu$.

(b) We have

$$\mathbb{P}(X = k \mid X + Y = n) = \frac{\mathbb{P}(X = k \text{ and } X + Y = n)}{\mathbb{P}(X + Y = n)} = \binom{n}{k} \left(\frac{\lambda}{\lambda + \mu}\right)^k \left(\frac{\mu}{\lambda + \mu}\right)^{n-k}, \tag{B.8}$$

hence, given $X + Y = n$, the random variable X has a binomial distribution with parameters n and $\lambda/(\lambda + \mu)$.

(c) In this case, using the exponential probability density $f_\Lambda(x) = \theta \mathbb{1}_{[0,\infty)}(x)\mathrm{e}^{-\theta x}$, $x \in \mathbb{R}$, we find

$$\mathbb{P}(X = k) = \left(1 - \frac{1}{\theta + 1}\right)\left(\frac{1}{\theta + 1}\right)^k.$$

Therefore $X + Y$ has a negative binomial distribution with parameter $(r, p) = (2, 1/(\theta + 1))$, cf. (1.5.12), and we have

$$\mathbb{P}(X = k \mid X + Y = n) = \frac{\mathbb{P}(X = k)\mathbb{P}(Y = n - k)}{\mathbb{P}(X + Y = n)} = \frac{1}{n + 1}, \quad k = 0, 1, \ldots, n,$$

which shows that the distribution of X given $X + Y = n$ is the discrete uniform distribution on $\{0, 1, \ldots, n\}$.

(d) In case X and Y have the same parameter, i.e. $\lambda = \mu$, we have

$$\mathbb{P}(X = k \mid X + Y = n) = \binom{n}{k}\frac{1}{2^n}, \quad k = 0, 1, \ldots, n,$$

which becomes independent of λ. Hence, when λ is represented by a random variable Λ with probability density $x \longmapsto f_\Lambda(x)$ on $\mathbb{R}_+$, from (B.8) we get $\mathbb{P}(X = k \mid X + Y = n) = 2^{-n}\binom{n}{k}$, $k = 0, 1, \ldots, n$.

Exercise 1.7 Let C_1 denote the color of the first drawn pen, and let C_2 denote the color of the second drawn pen. We have $\mathbb{P}(C_1 = R) = \mathbb{P}(C_1 = G) = 1/2$ and $\mathbb{P}(C_2 = R \text{ and } C_1 = R) = 2/3$, $\mathbb{P}(C_2 = R \text{ and } C_1 = G) = 1/3$. On the other hand, we have

$$\begin{aligned}
&\mathbb{P}(C_2 = R) = \mathbb{P}(C_2 = R \text{ and } C_1 = R) + \mathbb{P}(C_2 = R \text{ and } C_1 = G)\\
&= \mathbb{P}(C_2 = R \mid C_1 = R)\mathbb{P}(C_1 = R) + \mathbb{P}(C_2 = R \mid C_1 = G)\mathbb{P}(C_1 = G)\\
&= \frac{2}{3} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{2} = \frac{1}{2},
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{P}(C_2 = G) = \mathbb{P}(C_2 = G \text{ and } C_1 = R) + \mathbb{P}(C_2 = G \text{ and } C_1 = G)\\
&= \mathbb{P}(C_2 = G \mid C_1 = R)\mathbb{P}(C_1 = R) + \mathbb{P}(C_2 = G \mid C_1 = G)\mathbb{P}(C_1 = G) = \frac{1}{2}.
\end{aligned}$$

Finally, the probability we wish to compute is

$$\mathbb{P}(C_1 = R \mid C_2 = R) = \frac{\mathbb{P}(C_1 = R \text{ and } C_2 = R)}{\mathbb{P}(C_2 = R)} = \mathbb{P}(C_2 = R \mid C_1 = R)\frac{\mathbb{P}(C_1 = R)}{\mathbb{P}(C_2 = R)} = \frac{2}{3}.$$

Exercise 1.8

(a) The probability that the system operates is

$$\mathbb{P}(X \geq 2) = \binom{3}{2}p^2(1 - p) + p^3 = 3p^2 - 2p^3,$$

where X is a binomial random variable with parameter $(3, p)$.

(b) The probability that the system operates is

$$\int_0^1 \mathbb{P}(X \geq 2 \mid p = x) d\mathbb{P}(p \leq x) = \int_0^1 \mathbb{P}(X \geq 2 \mid p = x) dx = \frac{1}{2},$$

similarly to Exercise 1.6-(b).

Chapter 2 - Gambling Problems

Exercise 2.1

(a) By first step analysis we have

$$f(k) = (1 - 2r) f(k) + r f(k+1) + r f(k-1),$$

which yields the equation

$$f(k) = \frac{1}{2} f(k+1) + \frac{1}{2} f(k-1), \qquad 1 \leq k \leq S - 1, \tag{B.9}$$

with the boundary conditions $f(0) = 1$ and $f(S) = 0$, which is identical to Equation (2.2.18).

We refer to this equation as the *homogeneous equation*.

(b) According to the result of (2.2.19) in Sect. 2.2 we know that the general solution of (B.9) has the form

$$f(k) = C_1 + C_2 k, \qquad k = 0, 1, \ldots, S$$

and after taking into account the boundary conditions we find $f(k) = \dfrac{S-k}{S}$, $k = 0, 1, \ldots, S$.

(c) By first step analysis we have

$$\begin{aligned} h(k) &= (1 - 2r)(1 + h(k)) + r(1 + h(k+1)) + r(1 + h(k-1)) \\ &= 1 + (1 - 2r)h(k) + r h(k+1) + r h(k-1), \end{aligned}$$

hence the equation

$$h(k) = \frac{1}{2r} + \frac{1}{2} h(k+1) + \frac{1}{2} h(k-1), \qquad 1 \leq k \leq S - 1,$$

with the boundary conditions $h(0) = 0$ and $h(S) = 0$, which is identical to (2.3.6) by changing $h(k)$ into $2r \times h(k)$ in (2.3.6).

(d) After trying a solution of the form $h(k) = Ck^2$ we find

$$Ck^2 = \frac{1}{2r} + \frac{1}{2} C(k+1)^2 + \frac{1}{2} C(k-1)^2,$$

hence C should be equal to $C = -1/(2r)$, hence $k \longmapsto -k^2/(2r)$ is a particular solution.

(e) Given the hint, the general solution has the form $h(k) = C_1 + C_2 k - k^2/2r$, $k = 0, 1, \ldots, S$, which gives

$$h(k) = \frac{k(S-k)}{2r}, \qquad k = 0, 1, \ldots, S, \tag{B.10}$$

after taking into account the boundary conditions.

(f) Starting from any state $k \in \{1, 2, \ldots, S-1\}$, the mean duration goes to infinity when r goes to zero.

Problem 2.7

(a) We have

$$g(k) = pg(k+1) + qg(k-1), \qquad k = 1, 2, \ldots, S, \tag{B.11}$$

with

$$g(0) = pg(1) + qg(0) \tag{B.12}$$

for $k = 0$, and the boundary condition $g(S) = 1$.

(b) We observe that the constant function $g(k) = C$ is solution of both (B.11) and (B.12) and the boundary condition $g(S) = 1$ yields $C = 1$, hence $g(k) = \mathbb{P}(W \mid X_0 = k) = 1$ for all $k = 0, 1, \ldots, S$.

(c) We have

$$g(k) = 1 + pg(k+1) + qg(k-1), \qquad k = 1, 2, \ldots, S-1, \tag{B.13}$$

with $g(0) = 1 + pg(1) + qg(0)$ for $k = 0$, and the boundary condition $g(S) = 0$.

(d) *Case* $p \neq q$. The solution of the homogeneous equation

$$g(k) = pg(k+1) + qg(k-1), \qquad k = 1, 2, \ldots, S-1,$$

has the form $g(k) = C_1 + C_2(q/p)^k$, $k = 1, 2, \ldots, S-1$, and we can check that $k \mapsto k/(p-q)$ is a particular solution. Hence the general solution of (B.13) has the form

$$g(k) = \frac{k}{q-p} + C_1 + C_2(q/p)^k, \qquad k = 0, 1, \ldots, S,$$

with

$$\begin{cases} 0 = g(S) = \dfrac{S}{q-p} + C_1 + C_2(q/p)^S, \\ \\ pg(0) = p(C_1 + C_2) = 1 + pg(1) = 1 + p\left(\dfrac{1}{q-p} + C_1 + C_2\dfrac{q}{p}\right), \end{cases}$$

which yields

$$g(k) = \mathbb{E}[T_S \mid X_0 = k] = \frac{S-k}{p-q} + \frac{q}{(p-q)^2}((q/p)^S - (q/p)^k),$$

$k = 0, 1, \ldots, S$.

Case $p = q = 1/2$. The solution of the homogeneous equation is given by

$$g(k) = C_1 + C_2 k, \qquad k = 1, 2, \ldots, S-1,$$

and the general solution to (B.13) has the form $g(k) = -k^2 + C_1 + C_2 k$, $k = 1, 2, \ldots, S$, with

$$\begin{cases} 0 = g(S) = -S^2 + C_1 + C_2 S, \\ \\ \dfrac{g(0)}{2} = \dfrac{C_1}{2} = 1 + \dfrac{g(1)}{2} = 1 + \dfrac{-1 + C_1 + C_2}{2}, \end{cases}$$

hence

$$g(k) = \mathbb{E}[T_S \mid X_0 = k] = (S + k + 1)(S - k), \quad k = 0, 1, \ldots, S.$$

(e) When $p \neq q$ we have

$$p_k := \mathbb{P}(T_S < T_0 \mid X_0 = k) = \frac{1 - (q/p)^k}{1 - (q/p)^S}, \qquad k = 0, 1, \ldots, S,$$

and when $p = q = 1/2$ we find $p_k = k/S$, $k = 0, 1, \ldots, S$.

(f) The equality holds because, given that we start from state $\boxed{k+1}$ at time 1, whether $T_S < T_0$ or $T_S > T_0$ does not depend on the past of the process before time 1. In addition it does not matter whether we start from state $\boxed{k+1}$ at time 1 or at time 0.

(g) We have

$$\begin{aligned} \mathbb{P}(X_1 = k+1 \mid X_0 = k \text{ and } T_S < T_0) &= \frac{\mathbb{P}(X_1 = k+1,\ X_0 = k,\ T_S < T_0)}{\mathbb{P}(X_0 = k \text{ and } T_S < T_0)} \\ &= p\frac{\mathbb{P}(T_S < T_0 \mid X_0 = k+1)}{\mathbb{P}(T_S < T_0 \mid X_0 = k)} = p\frac{p_{k+1}}{p_k}, \end{aligned}$$

$k = 0, 1, \ldots, S-1$. By the result of Question (e), when $p \neq q$ we find

$$\mathbb{P}(X_1 = k+1 \mid X_0 = k \text{ and } T_S < T_0) = p \frac{1-(q/p)^{k+1}}{1-(q/p)^k},$$

$k = 1, 2, \ldots, S-1$, and in case $p = q = 1/2$ we get

$$\mathbb{P}(X_1 = k+1 \mid X_0 = k \text{ and } T_S < T_0) = \frac{k+1}{2k},$$

$k = 1, 2, \ldots, S-1$. Note that this probability is higher than $p = 1/2$.

(h) Similarly, we have

$$\begin{aligned}
&\mathbb{P}(X_1 = k-1 \mid X_0 = k \text{ and } T_0 < T_S) \\
&= \frac{\mathbb{P}(X_1 = k-1,\ X_0 = k \text{ and } T_0 < T_S)}{\mathbb{P}(X_0 = k \text{ and } T_0 < T_S)} \\
&= q \frac{\mathbb{P}(T_0 < T_S \mid X_0 = k-1)}{\mathbb{P}(T_0 < T_S \mid X_0 = k)} = q \frac{1-p_{k-1}}{1-p_k},
\end{aligned}$$

$k = 1, 2, \ldots, S-1$. When $p \neq q$ this yields

$$\mathbb{P}(X_1 = k-1 \mid X_0 = k \text{ and } T_0 < T_S) = q \frac{(q/p)^{k-1} - (q/p)^S}{(q/p)^k - (q/p)^S},$$

$k = 1, 2, \ldots, S-1$, and when $p = q = 1/2$ we find

$$\mathbb{P}(X_1 = k-1 \mid X_0 = k \text{ and } T_0 < T_S) = \frac{S+1-k}{2(S-k)},$$

$k = 1, 2, \ldots, S-1$. Note that this probability is higher than $q = 1/2$.

(i) We find

$$h(k) = 1 + p \frac{p_{k+1}}{p_k} h(k+1) + \left(1 - p \frac{p_{k+1}}{p_k}\right) h(k-1), \tag{B.14}$$

$k = 1, 2, \ldots, S-1$, or, due to the first step equation $p_k = p p_{k+1} + q p_{k-1}$,

$$p_k h(k) = p_k + p p_{k+1} h(k+1) + q p_{k-1} h(k-1), \quad k = 1, 2, \ldots, S-1,$$

with the boundary condition $h(S) = 0$. When $p = q = 1/2$ we have $p_k = k/S$ by Question (e), hence (B.14) becomes

$$h(k) = 1 + \frac{k+1}{2k} h(k+1) + \frac{k-1}{2k} h(k-1), \quad k = 1, 2, \ldots, S-1.$$

(j) We have to solve

$$kh(k) = k + \frac{1}{2}(k+1)h(k+1) + \frac{1}{2}(k-1)h(k-1), \qquad k = 1, 2, \ldots, S-1,$$

with the boundary condition $h(S) = 0$. Letting $g(k) = kh(k)$ we check that $g(k)$ satisfies

$$g(k) = k + \frac{1}{2}g(k+1) + \frac{1}{2}g(k-1), \quad k = 1, 2, \ldots, S-1, \tag{B.15}$$

with the boundary conditions $g(0) = 0$ and $g(S) = 0$. We check that $g(k) = Ck^3$ is a particular solution when $C = -1/3$, hence the solution of (B.15) has the form $g(k) = -k^3/3 + C_1 + C_2 k$, by the homogeneous solution given in Sect. 2.3, where C_1 and C_2 are determined by the boundary conditions $0 = g(0) = C_1$ and

$$0 = g(S) = -\frac{1}{3}S^3 + C_1 + C_2 S,$$

i.e. $C_1 = 0$ and $C_2 = S^2/3$. Consequently, we have $g(k) = k(S^2 - k^2)/3$, $k = 0, 1, \ldots, S$, hence we have

$$h(k) = \mathbb{E}[T_S \mid X_0 = k,\ T_S < T_0] = \frac{S^2 - k^2}{3}, \qquad k = 1, 2, \ldots, S.$$

Chapter 3 - Random Walks

Exercise 3.1

(a) We find $\binom{4}{3} = \binom{4}{1} = 4$ paths, as follows.

(b) In each of the $\binom{4}{3} = \binom{4}{1} = \frac{4!}{3!} = 4$ paths there are 3 steps up (with probability p) and 1 step down (with probability $q = 1 - p$), hence the result.

(c) We consider two cases depending on the parity of n and k.

  (i) In case n and k are even, or written as $n = 2n'$ and $k = 2k'$, (3.3.3) shows that

$$\mathbb{P}(S_n = k \mid S_0 = 0) = \mathbb{P}(S_{2n'} = 2k' \mid S_0 = 0) = \binom{n}{(n+k)/2} p^{(n+k)/2} q^{(n-k)/2},$$

  $-n \leq k \leq n$.

  (ii) In case n and k are odd, or written as $n = 2n' + 1$ and $k = 2k' + 1$, (3.3.4) shows that

$$\mathbb{P}(S_n = k \mid S_0 = 0) = \binom{n}{(n+k)/2} p^{(n+k)/2} q^{(n-k)/2}, \qquad -n \leq k \leq n.$$

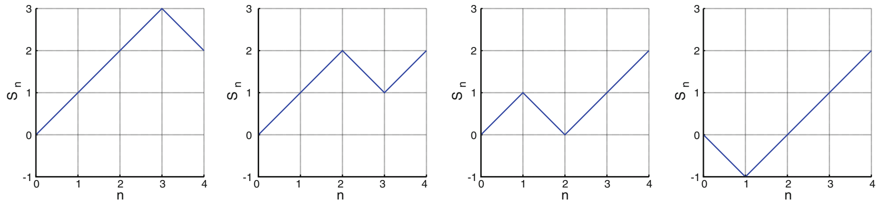


Fig. B.1 Four paths leading from 0 to 2 in four steps

(d) By a first step analysis started at state 0 we have, letting $p_{n,k} := \mathbb{P}(S_n = k)$, $p_{n+1,k} = pp_{n,k-1} + qp_{n,k+1}$, for all $n \in \mathbb{N}$ and $k \in \mathbb{Z}$.

(e) We consider two cases depending on the parity of $n + 1 + k$.

  (i) If $n + 1 + k$ is odd the equation is clearly satisfied as both the right hand side and left hand side of (3.4.25) are equal to 0.

  (ii) If $n + 1 + k$ is even we have $pp_{n,k-1} + qp_{n,k+1} p_{n+1,k}$, which shows that $p_{n,k}$ satisfies Equation (3.4.25). In addition we clearly have

$$p_{0,0} = \mathbb{P}(S_0 = 0) = 1 \quad \text{and} \quad p_{0,k} = \mathbb{P}(S_0 = k) = 0, \qquad k \neq 0.$$

Exercise 3.9

(a) Since the increment X_k takes its values in $\{-1, 1\}$, the set of distinct values in $\{S_0, S_1, \ldots, S_n\}$ is the integer interval

$$\left[\inf_{k=0,1,\ldots,n} S_k, \sup_{k=0,1,\ldots,n} S_k \right],$$

which has

$$R_n = 1 + \left(\sup_{k=0,1,\ldots,n} S_k \right) - \left(\inf_{k=0,1,\ldots,n} S_k \right)$$

elements. In addition we have $R_0 = 1$ and $R_1 = 2$.

(b) At each time step $k \geq 1$ the range can only either increase by one unit or remain constant, hence $R_k - R_{k-1} \in \{0, 1\}$ is a Bernoulli random variable. In addition we have the identity

$$\{R_k - R_{k-1} = 1\} = \{S_k \neq S_0, \ S_k \neq S_1, \ldots, \ S_k \neq S_{k-1}\},$$

hence, applying the probability $\mathbb{P}$ to both sides, we get

$$\mathbb{P}(R_k - R_{k-1} = 1) = \mathbb{P}(S_k - S_0 \neq 0, \ S_k - S_1 \neq 0, \ldots, \ S_k - S_{k-1} \neq 0).$$

(c) By the change of index

$$(X_1, X_2, \ldots, X_{k-1}, X_k) \longmapsto (X_k, X_{k-1}, \ldots, X_2, X_1)$$

under which $X_1 + X_2 + \cdots + X_l$ becomes $X_k + \cdots + X_{k-l+1}$, $l = 1, 2, \ldots, k$, we have

$$\mathbb{P}(R_k - R_{k-1} = 1) = \mathbb{P}(X_1 \neq 0,\ X_1 + X_2 \neq 0, \ldots,\ X_1 + \cdots + X_k \neq 0),$$

for all $k \geq 1$, since the sequence $(X_k)_{k\geq 1}$ is made of independent and identically distributed random variables.

(d) We have the telescoping sum

$$R_n = R_0 + \sum_{k=1}^{n}(R_k - R_{k-1}), \qquad n \in \mathbb{N}.$$

(e) By (1.2.4) we have

$$\mathbb{P}(T_0 = \infty) = \mathbb{P}\left(\bigcap_{k\geq 1}\{T_0 > k\}\right) = \lim_{k\to\infty}\mathbb{P}(T_0 > k),$$

since $\{T_0 > k+1\} \Longrightarrow \{T_0 > k\}$, $k \geq 1$, i.e. $(\{T_0 > k\})_{k\geq 1}$ is a decreasing sequence of events.

(f) Noting that $R_k - R_{k-1} \in \{0, 1\}$ is a Bernoulli random variable with $\mathbb{E}[R_k - R_{k-1}] = \mathbb{P}(R_k - R_{k-1} = 1)$, we find

$$\mathbb{E}[R_n] = \sum_{k=0}^{n}\mathbb{P}(T_0 > k).$$

(g) Let $\varepsilon > 0$. Since by Question (e) we have $\mathbb{P}(T_0 = \infty) = \lim_{k\to\infty}\mathbb{P}(T_0 > k)$, there exists $N \geq 1$ such that

$$|\mathbb{P}(T_0 = \infty) - \mathbb{P}(T_0 > k)| < \varepsilon, \qquad k \geq N.$$

Hence for $n \geq N$ we have

$$\left|\mathbb{P}(T_0 = \infty) - \frac{1}{n}\sum_{k=1}^{n}\mathbb{P}(T_0 > k)\right| \leq \frac{N}{n} + \varepsilon.$$

Then, choosing $N_0 \geq 1$ such that $(N+1)/n \leq \varepsilon$ for $n \geq N_0$, we get

$$\left|\mathbb{P}(T_0 = \infty) - \frac{1}{n}\mathbb{E}[R_n]\right| \leq \frac{1}{n} + \left|\mathbb{P}(T_0 = \infty) - \frac{1}{n}\sum_{k=1}^{n}\mathbb{P}(T_0 > k)\right| \leq 2\varepsilon,$$

$n \geq N_0$, which concludes the proof.

(h) From Relation (3.4.15) in Sect. 3.4 we have $\mathbb{P}(T_0 = +\infty) = |p - q|$, hence by the result of Question (g) we get $\lim_{n\to\infty} \mathbb{E}[R_n]/n = |p - q|$, when $p \neq q$, and $\lim_{n\to\infty} \mathbb{E}[R_n]/n = 0$ when $p = q = 1/2$.

Chapter 4 - Discrete-Time Markov Chains

Exercise 4.10

(a) Let S_n denote the wealth of the player at time $n \in \mathbb{N}$. The process $(S_n)_{n\in\mathbb{N}}$ is a Markov chain whose transition matrix is given by

$$P = \left[P_{i,j} \right]_{i,j\in\mathbb{N}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ q & 0 & p & 0 & 0 & 0 & \cdots \\ q & 0 & 0 & p & 0 & 0 & \cdots \\ q & 0 & 0 & 0 & p & 0 & \cdots \\ q & 0 & 0 & 0 & 0 & p & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

After n time steps we have $\mathbb{P}(S_n = n + 1 \mid S_0 = 1) = p^n$, $n \geq 1$, and P^n is given by

$$P^n = \left[[P^n]_{i,j} \right]_{i,j\in\mathbb{N}} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & p^n & 0 & 0 & 0 & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & 0 & p^n & 0 & 0 & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & 0 & 0 & p^n & 0 & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & 0 & 0 & 0 & p^n & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & p^n & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{B.16}$$

in which the n columns n^o 2 to $n + 1$ are identically 0.

(b) In this case the transition matrix P becomes

$$P = \left[P_{i,j} \right]_{i,j\in\mathbb{N}} = \begin{bmatrix} q & p & 0 & 0 & 0 & 0 & \cdots \\ q & 0 & p & 0 & 0 & 0 & \cdots \\ q & 0 & 0 & p & 0 & 0 & \cdots \\ q & 0 & 0 & 0 & p & 0 & \cdots \\ q & 0 & 0 & 0 & 0 & p & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

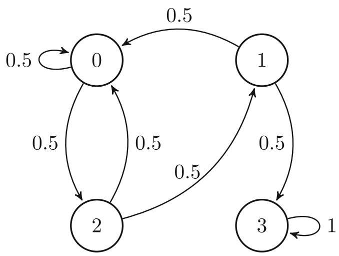
and by induction on $n \geq 2$ we find

$$P^n = \begin{bmatrix} q & qp & qp^2 & \dots & qp^{n-1} & p^n & 0 & 0 & 0 & \cdots \\ q & qp & qp^2 & \dots & qp^{n-1} & 0 & p^n & 0 & 0 & \cdots \\ q & qp & qp^2 & \dots & qp^{n-1} & 0 & 0 & p^n & 0 & \cdots \\ q & qp & qp^2 & \dots & qp^{n-1} & 0 & 0 & 0 & p^n & \cdots \\ q & qp & qp^2 & \dots & qp^{n-1} & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{B.17}$$

Chapter 5 - First Step Analysis

Exercise 5.5 This exercise is a particular case of the Example of Sect. 5.1, by taking $a := 0.3$, $b := 0$, $c := 0.7$, $d := 0$, $\alpha := 0$, $\beta := 0.3$, $\gamma := 0$, $\eta := 0.7$.

Exercise 5.6 We observe that state 3 is absorbing:



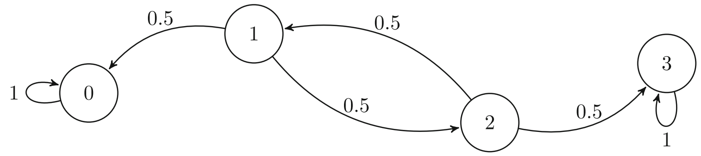
Let $h_3(k) := \mathbb{E}[T_3 \mid X_0 = k]$ denote the mean (hitting) time needed to reach state 3 after starting from state $k = 0, 1, 2, 3$. We have

$$\begin{cases} h_3(0) = 1 + \frac{1}{2}h_3(0) + \frac{1}{2}h_3(2) \\ \\ h_3(1) = 1 + \frac{1}{2}h_3(0) \\ \\ h_3(2) = 1 + \frac{1}{2}h_3(0) + \frac{1}{2}h_3(1) \\ \\ h_3(3) = 0, \end{cases}$$

which yields $h_3(3) = 0$, $h_3(1) = 8$, $h_3(2) = 12$, $h_3(0) = 14$. We check that $h_3(3) < h_3(1) < h_3(2) < h_3(0)$, as can be expected from the graph.

Exercise 5.7

(a) The chain has the following graph:



Note that this process is in fact a fair gambling process on the state space $\{0, 1, 2, 3\}$.

(b) Since the states ⓪ and ③ are absorbing, by first step analysis we find

$$g_0(0) = 1, \quad g_0(1) = \frac{2}{3}, \quad g_0(2) = \frac{1}{3}, \quad g_0(3) = 0.$$

(c) By first step analysis we find

$$h_{0,3}(0) = 0, \quad h_{0,3}(1) = 2, \quad h_{0,3}(2) = 2, \quad h_{0,3}(3) = 0.$$

Exercise 5.9

(a) Letting $f(k) := \mathbb{P}(T_0 < \infty \mid S_0 = k)$ we have the boundary condition $f(0) = 1$ and by first step analysis we find that $f(k)$ satisfies

$$f(k) = pf(k+1) + qf(k-1), \qquad k \geq 1,$$

which is (2.2.6), and has the general solution

$$f(k) = C_1 + C_2 r^k, \qquad k \in \mathbb{N}, \tag{B.18}$$

where $r = q/p$, by (2.2.16).

(i) In case $q \geq p$, $f(k)$ would tend to (positive or negative) infinity if $C_2 \neq 0$, hence we should have $C_2 = 0$, and $C_1 = f(0) = 1$, showing that $f(k) = 1$ for all $k \in \mathbb{N}$.

(ii) In case $q < p$, the probability of hitting ⓪ in finite time starting from ⓚ becomes 0 in the limit as k tends to infinity, i.e. we have

$$\lim_{k\to\infty} f(k) = \lim_{k\to\infty} \mathbb{P}(T_0 < \infty \mid S_0 = k) = 0, \qquad k \in \mathbb{N},$$

which shows that $C_1 = 0$.

On the other hand, the condition $f(0) = 1$ yields $C_2 = 1$, hence we find $f(k) = (q/p)^k$ for all $k \geq 0$.

(b) Letting $h(k) := \mathbb{E}[T_0 \mid S_0 = k]$ we have the boundary condition $h(0) = 0$ and by first step analysis we find that $h(k)$ satisfies

$$h(k) = 1 + ph(k+1) + qh(k-1), \qquad k \geq 1,$$

which is (2.3.6) and has a general solution of the form

$$h(k) = C_1 + C_2 r^k + \frac{1}{q-p} k, \qquad k \in \mathbb{N}, \tag{B.19}$$

by (2.3.9). Next, we note that by the Markov property we should have the decomposition

$$\begin{aligned} h(k+1) &= \mathbb{E}[T_0 \mid S_0 = k+1] \\ &= \mathbb{E}[T_0 \mid S_0 = 1] + \mathbb{E}[T_0 \mid S_0 = k] \\ &= h(1) + h(k), \qquad k \in \mathbb{N}, \end{aligned}$$

i.e. the mean time to go down from $k+1$ to 0 should be the sum of the mean time needed to go down one step plus the mean time needed to go down k steps. This shows that

$$h(k) = h(0) + kh(1) = kh(1),$$

hence by (B.19) we have $C_1 = C_2 = 0$, $h(1) = 1/(q-p)$, and

$$h(k) = \frac{k}{q-p}, \qquad k \in \mathbb{N}.$$

Exercise 5.10 First, we take a look at the complexity of the problem. Starting from 0 there are multiple ways to reach state 13 without reaching 11 or 12. For example:

$$13 = 3+4+1+5, \text{ or } 13 = 1+6+3+3, \text{ or } 13 = 1+1+2+1+3+1+4, \text{ etc.}$$

Clearly it would be difficult to enumerate all such possibilities, for this reason we use the framework of Markov chains. We denote by X_n the cumulative sum of dice outcomes after n rounds, and choose to model it as a Markov chain with n as a time index. We can represent X_n as

$$X_n = \sum_{k=1}^{n} \xi_k, \qquad n \geq 0,$$

where $(\xi_k)_{k \geq 1}$ is a family of independent random variables uniformly distributed over $\{1, 2, 3, 4, 5, 6\}$. The process $(X_n)_{n \geq 0}$ is a Markov chain since given the history of $(X_k)_{k=0,1,\ldots,n}$ up to time n, the value

$$X_{n+1} = X_n + \xi_{n+1}$$

depends only on X_n and on ξ_{n+1} which is independent of $X_0, X_1, \ldots, X_n$. The process $(X_n)_{n\geq 0}$ is actually a random walk with independent increments $\xi_1, \xi_2, \ldots$. The chain $(X_n)_{n\geq 0}$ has the transition matrix

$$
\left[\, P_{i,j} \,\right]_{i,j\in\mathbb{N}} = \begin{bmatrix}
0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}.
$$

Letting $A := \{11, 12, 13, 14, 15, 16\}$, we are looking at the probability

$$g_0 := \mathbb{P}(X_{T_A} = 13 \mid X_0 = 0)$$

of hitting the set A through and the set 13 after starting from state 0. More generally, letting

$$g_k := \mathbb{P}(X_{T_A} = 13 \mid X_0 = 0)$$

denote the probability of hitting the set A through the set 13 after starting from state k, we have $g_k = 0$ for all $k \geq 14$. By first step analysis we find the system of equations

$$g(k) = \frac{1}{6}\sum_{i=1}^{6} g_{k+i}, \qquad k \in \mathbb{N},$$

with solution

$$g_7 = \frac{7^3}{6^4}, \quad g_8 = \frac{7^2}{6^3}, \quad g_9 = \frac{7}{6^2}, \quad g_{10} = \frac{1}{6},$$

and

$$g_0 = \frac{7^{10} - 7^6 \times 6^4 - 4 \times 7^3 \times 6^6}{6^{11}} \simeq 0.181892636.$$

Exercise 5.11

(a) The transition matrix is given by

$$\begin{bmatrix} \times & \times & \times & \times & \times & \times \\ q & 0 & p & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & 0 \\ 0 & 0 & q & 0 & p & 0 \\ 0 & 0 & 0 & q & 0 & p \\ \times & \times & \times & \times & \times & \times \end{bmatrix}.$$

The information contained in the first and last lines of the matrix is not needed here because they have no influence on the result. We have $g(0) = 0$, $g(5) = 1$, and

$$g(k) = q \times g(k-1) + p \times g(k+1), \qquad 1 \le k \le 4. \tag{B.20}$$

(b) When $p = q = 1/2$ the probability that starting from state k the fish finds the food before getting shocked is obtained by solving Equation (B.20) rewritten as

$$g(k) = \frac{1}{2} \times g(k-1) + \frac{1}{2} \times g(k+1), \qquad 1 \le k \le 4.$$

Trying a solution of the form $g(k) = C_1 + kC_2$ under the boundary conditions $g(0) = 0$ and $g(5) = 1$, shows that $C_1 = 0$ and $C_2 = 1/5$, which yields $g(k) = k/5$, $k = 0, 1, \ldots, 5$.

Exercise 5.12

(a) The transition matrix is given by

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \cdots \\ 1 & 0 & 0 & 0 & 0 & 0 \cdots \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 \cdots \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \cdots \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 \cdots \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 & 0 \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \ddots \end{bmatrix}.$$

(b) We have

$$h_0(m) = \sum_{k=0}^{m-1} \frac{1}{m}(1 + h_0(k)) = 1 + \frac{1}{m} \sum_{k=0}^{m-1} h_0(k), \qquad m \ge 1,$$

and $h_0(0) = 0$, $h_0(1) = 1$.

(c) We have

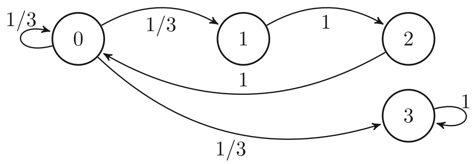
$$h_0(m) = 1 + \frac{1}{m} \sum_{k=0}^{m-1} h_0(k) = h_0(m-1) + \frac{1}{m}, \qquad m \ge 1,$$

hence

$$h_0(m) = h_0(m-1) + \frac{1}{m} = \sum_{k=1}^{m} \frac{1}{k}, \qquad m \geq 1.$$

Exercise 5.13

(a) Assuming that it takes one day per state transition, the graph of the chain can be drawn as



where state (0) represents the tower, states (1) and (2) represents the tunnel, and state (3) represents the outside.

(b) We have

$$P = \begin{bmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(c) By first step analysis we find

$$h_3(0) = \frac{1}{3}(1 + h_3(0)) + \frac{1}{3}(3 + h_3(0)) + \frac{1}{3},$$

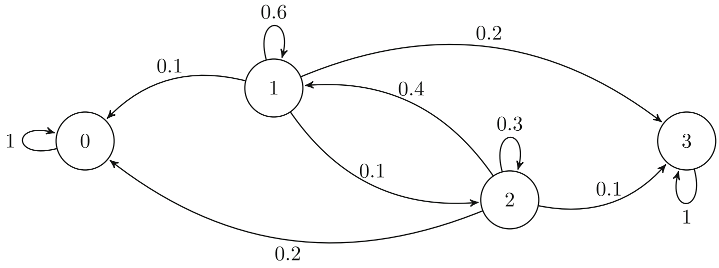
i.e. $h_3(0) = 5$, i.e. 4 times steps on average to reach the exit, plus one time step from the exit to the outside.

Exercise 5.14 The average time t spent inside the maze can be quickly computed by the following first step analysis using weighted links:

$$t = \frac{1}{2} \times (t+3) + \frac{1}{6} \times 2 + \frac{2}{6} \times (t+5),$$

which yields $t = 21$. We refer to Exercise 5.13 and its solution for a more detailed analyis of a similar problem.

Exercise 5.15 The chain has the following graph:



(a) Let us compute

$$g_0(k) = \mathbb{P}(T_0 < \infty \mid X_0 = k) = \mathbb{P}(X_{T_{\{0,3\}}} = 0 \mid X_0 = k), \qquad k = 0, 1, 2, 3.$$

Since states (0) and (3) are absorbing, by first step analysis we have

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = 0.1 \times g_0(0) + 0.6 \times g_0(1) + 0.1 \times g_0(2) + 0.2 \times g_0(3) \\ g_0(2) = 0.2 \times g_0(0) + 0.3 \times g_0(1) + 0.4 \times g_0(2) + 0.1 \times g_0(3) \\ g_0(3) = 0, \end{cases}$$

i.e.

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = 0.1 + 0.6 \times g_0(1) + 0.1 \times g_0(2) \\ g_0(2) = 0.2 + 0.3 \times g_0(1) + 0.4 \times g_0(2) \\ g_0(3) = 0, \end{cases}$$

which has for solution $g_0(0) = 1$, $g_0(1) = 8/21$, $g_0(2) = 11/21$, $g_0(3) = 0$, cf. also (5.1.10).

(b) Let

$$h_{0,3}(k) = \mathbb{E}[T_{\{0,3\}} \mid X_0 = k]$$

denote the mean time to reach the set $A = \{0, 3\}$ starting from $k = 0, 1, 2, 3$. By first step analysis, we have

$$\begin{cases} h_{0,3}(0) = 0 \\ h_{0,3}(1) = 0.1 \times 1 + 0.6 \times (1 + h_{0,3}(1)) + 0.1 \times (1 + h_{0,3}(2)) + 0.2 \times (1 + h_{0,3}(3)) \\ h_{0,3}(2) = 0.2 \times 1 + 0.4 \times (1 + h_{0,3}(1)) + 0.3 \times (1 + h_{0,3}(2)) + 0.1 \times (1 + h_{0,3}(3)) \\ h_{0,3}(3) = 0, \end{cases}$$

i.e.

$$\begin{cases} h_{0,3}(0) = 0 \\ h_{0,3}(1) = 1 + 0.6 \times h_{0,3}(1) + 0.1 \times h_{0,3}(2) \\ h_{0,3}(2) = 1 + 0.4 \times h_{0,3}(1) + 0.3 \times h_{0,3}(2) \\ h_{0,3}(3) = 0, \end{cases}$$

which has for solution $h_{0,3}(0) = 0$, $h_{0,3}(1) = 10/3$, $h_{0,3}(2) = 10/3$, $h_{0,3}(3) = 0$.

Note that the relation $h_{0,3}(1) = h_{0,3}(2)$ can be guessed from the symmetry of the problem.

Exercise 5.19 We have

$$\begin{aligned} h(k) &= \mathbb{E}\left[\sum_{i=0}^{\infty} \beta^i c(X_i) \,\Big|\, X_0 = k\right] \\ &= \mathbb{E}\left[c(X_0) \,\Big|\, X_0 = k\right] + \mathbb{E}\left[\sum_{i=1}^{\infty} \beta^i c(X_i) \,\Big|\, X_0 = k\right] \\ &= c(k) + \sum_{j \in S} P_{k,j} \mathbb{E}\left[\sum_{i=1}^{\infty} \beta^i c(X_i) \,\Big|\, X_1 = j\right] \\ &= c(k) + \beta \sum_{j \in S} P_{k,j} \mathbb{E}\left[\sum_{i=0}^{\infty} \beta^i c(X_i) \,\Big|\, X_0 = j\right] \\ &= c(k) + \beta \sum_{j \in S} P_{k,j} h(j), \qquad k \in S. \end{aligned}$$

However, this type of equation may be difficult to solve in general. We refer to Problem 5.22 for a particular case with explicit solution.

Problem 5.21

(a) The boundary conditions $g(0)$ and $g(N)$ are given by $g(0) = 1$ and $g(N) = 0$.
(b) We have

$$g(k) = \mathbb{P}(T_0 < T_N \mid X_0 = k) = \sum_{l=0}^{N} \mathbb{P}(T_0 < T_N \mid X_1 = l)\mathbb{P}(X_1 = l \mid X_0 = k)$$
$$= \sum_{l=0}^{N} g(l) P_{k,l}, \qquad k = 0, 1, \ldots, N.$$

(c) We find

$$[\ P_{i,j}\ ]_{0 \le i,j \le 3} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ (2/3)^3 & (2/3)^2 & 2/3^2 & 1/3^3 \\ 1/3^3 & 2/3^2 & (2/3)^2 & (2/3)^3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(d) Letting $g(k) = 1 - k/N$, we check that $g(k)$ satisfies the boundary conditions $g(0) = 1$ and $g(N) = 0$, and in addition we have

$$\begin{aligned} \sum_{l=0}^{N} g(l) P_{k,l} &= \sum_{l=0}^{N} \frac{N!}{(N-l)!l!} \left(\frac{k}{N}\right)^l \left(1 - \frac{k}{N}\right)^{N-l} \frac{N-l}{N} \\ &= \left(1 - \frac{k}{N}\right)\left(\frac{k}{N} + 1 - \frac{k}{N}\right)^{N-1} \\ &= \frac{N-k}{N} = g(k), \qquad k = 0, 1, \ldots, N, \end{aligned}$$

which allows us to conclude by uniqueness of the solution given two boundary conditions, cf. Exercise 5.9 for cases of non-uniqueness under a single boundary condition.

(e) The boundary conditions $h(0)$ and $h(N)$ are given by $h(0) = 0$ and $h(N) = 0$ since the states $\textcircled{0}$ and $\textcircled{N}$ are absorbing.

(f) We have

$$\begin{aligned} h(k) &= \mathbb{E}[T_{0,N} \mid X_0 = k] \\ &= \sum_{l=0}^{N} (1 + \mathbb{E}[T_{0,N} \mid X_1 = l])\mathbb{P}(X_1 = l \mid X_0 = k) \\ &= \sum_{l=0}^{N} (1 + \mathbb{E}[T_{0,N} \mid X_1 = l]) P_{k,l} = \sum_{l=0}^{N} P_{k,l} + \sum_{l=0}^{N} \mathbb{E}[T_{0,N} \mid X_1 = l] P_{k,l} \\ &= 1 + \sum_{l=1}^{N-1} \mathbb{E}[T_{0,N} \mid X_1 = l] P_{k,l} = 1 + \sum_{l=1}^{N-1} h(l) P_{k,l}, \end{aligned} \tag{B.21}$$

$k = 1, 2, \ldots, N - 1.$

(g) In this case, the Equation (B.21) reads

$$\begin{cases} h(0) = 0, \\ h(1) = 1 + \frac{4}{9}h(1) + \frac{2}{9}h(2) \\ h(2) = 1 + \frac{2}{9}h(1) + \frac{4}{9}h(2) \\ h(3) = 0, \end{cases}$$

which yields $h(0) = 0$, $h(1) = 3$, $h(2) = 3$, $h(3) = 0$.

Problem 5.22

(a) Since we consider the time until we hit either 0 or N, we have $h(N) = 0$ as well as $h(N) = 0$.

(b) We have

$$\begin{aligned} h(k) &= \mathbb{E}\left[\sum_{i=0}^{\tau-1} X_i \,\Big|\, X_0 = k\right] \\ &= \mathbb{E}\left[X_0 \,\Big|\, X_0 = k\right] + \mathbb{E}\left[\sum_{i=1}^{\tau-1} X_i \,\Big|\, X_0 = k\right] \\ &= k + p\mathbb{E}\left[\sum_{i=0}^{\tau-2} X_{i+1} \,\Big|\, X_1 = k+1\right] + q\mathbb{E}\left[\sum_{i=0}^{\tau-2} X_{i+1} \,\Big|\, X_1 = k-1\right] \qquad \text{(B.22)} \\ &= k + p\mathbb{E}\left[\sum_{i=0}^{\tau-1} X_i \,\Big|\, X_0 = k+1\right] + q\mathbb{E}\left[\sum_{i=0}^{\tau-1} X_i \,\Big|\, X_0 = k-1\right] \qquad \text{(B.23)} \\ &= k + ph(k+1) + qh(k-1), \qquad 1 \le k \le N-1, \end{aligned}$$

where we used the fact that $\tau - 1$ in (B.22) becomes τ in (B.23).

From now on we take $p = q = 1/2$.

(c) We check successively that $h(k) = C$, $h(k) = Ck$, $h(k) = Ck^2$ cannot be solution and that $h(k) = Ck^3$ is solution provided that $C = -1/3$.

(d) The general solution has the form $h(k) = -k^3/3 + C_1 + C_2 k$, and the boundary conditions show that

$$\begin{cases} 0 = h(0) = C_1, \\ 0 = h(N) = -\frac{N^3}{3} + C_1 + C_2 N, \end{cases}$$

hence $C_1 = 0$, $C_2 = N^2/3$, and

$$h(k) = -\frac{k^3}{3} + N^2\frac{k}{3} = \frac{k}{3}(N^2 - k^2) = k(N-k)\frac{N+k}{3}, \quad k = 0, 1, \ldots, N. \tag{B.24}$$

(e) When $N = 2$ we find $h(1) = 1$ since starting from $k = 1$ we can only move to state 0 or state $N = 2$ which ends the game with a cumulative sum equal to 1 in both cases.

(f) (i) We find an average of

$$\mathbb{E}[T_{0,N} \mid X_0 = 4] = 4(70-4) = 4(70-4) = 264 \text{ months } = 22 \text{ years}.$$

(ii) By (B.24) we find

$$h(4) = \frac{4}{3}(70^2 - 4^2) = \$6512\text{K} = \$6.512\text{M}.$$

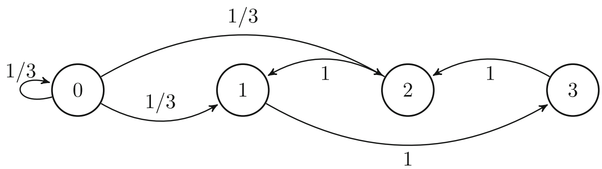
(iii) In that case we find $\$4\text{K} \times 264 = \$1056\text{K} = \$1.056\text{M}$.

(iv) It appears that starting a (potentially risky) business is more profitable on the average than keeping the same fixed initial income over an equivalent (average) period of time.

Chapter 6 - Classification of States

Exercise 6.1

(a) The graph of the chain is



This Markov chain is reducible because its state space can be partitioned into two communicating classes as $\mathbb{S} = \{0\} \cup \{1, 2, 3\}$.

(b) State 0 has period 1 and states 1, 2, 3 have period 3.

(c) We have

$$p_{0,0} = \mathbb{P}(T_0 < \infty \mid X_0 = 0) = \mathbb{P}(T_0 = 1 \mid X_0 = 0) = \frac{1}{3},$$

and

$$\mathbb{P}(T_0 = \infty \mid X_0 = 0) = 1 - \mathbb{P}(T_0 < \infty \mid X_0 = 0) = \frac{2}{3}.$$

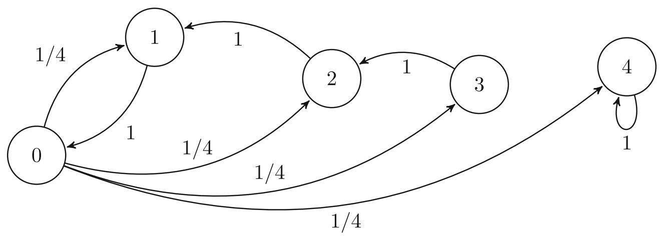
We also have

$$\mathbb{P}(R_0 < \infty \mid X_0 = 0) = \mathbb{P}(T_0 = \infty \mid X_0 = 0) \sum_{n=1}^{\infty} (\mathbb{P}(T_0 < \infty \mid X_0 = 0))^n$$
$$= \frac{2}{3} \sum_{n=1}^{\infty} \left(\frac{1}{3}\right)^n = 1.$$

(d) There are no absorbing states, state 0 is transient, and states 1, 2, 3 are recurrent by Corollary 6.6. State 0 is transient since $\mathbb{P}(R_0 < \infty \mid X_0 = 0) = 1$, as expected according to (5.4.3).

Exercise 6.3

(a) The chain has the following graph



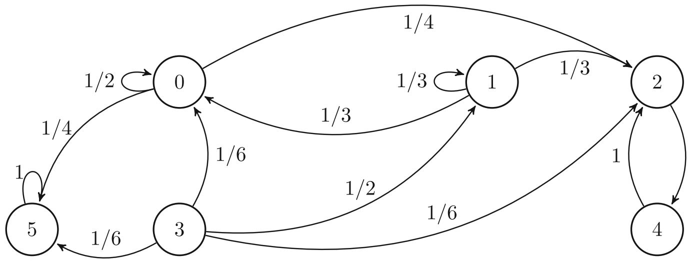
(b) All states 0, 1, 2 and 3 have period 1, which can be obtained as the greastest common divisor (GCD) of {2, 3} for states 0, 1, 2 and {4, 6, 7} for state 3. The chain is aperiodic.

(c) State 4 is absorbing (and therefore recurrent), state 0 is transient because
$$\mathbb{P}(T_0^r = \infty | X_0 = 0) \geq \frac{1}{4} > 0,$$
and the remaining states 1, 2, 3 are also transient because they communicate with the transient state 0, cf. Corollary 6.6. By pathwise or first step analysis we can actually check that

$$\mathbb{P}(T_0^r = \infty | X_0 = 0)$$
$$= \frac{1}{4}\big(\mathbb{P}(T_0^r = \infty | X_0 = 1) + \mathbb{P}(T_0^r = \infty | X_0 = 2) + \mathbb{P}(T_0^r = \infty | X_0 = 3)\big) = \frac{3}{4}.$$

(d) The Markov chain is reducible because its state space $\mathbb{S} = \{0, 1, 2, 3, 4\}$ can be partitioned into two communicating classes {0, 1, 2, 3} and {4}.

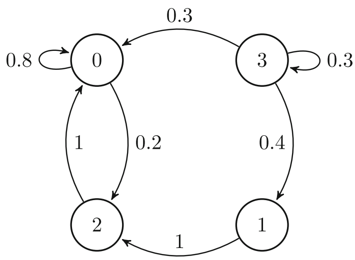
Exercise 6.4 The graph of the chain is



(a) The chain is reducible and its communicating classes are {0}, {1}, {3}, {5}, and {2, 4}.
(b) States 0, 1, 3 are transient and states 2, 4, 5 are recurrent.
(c) State 3 has period 0, states 2 and 4 have period 2, and states 0, 1, 5 are aperiodic.

Exercise 6.5

(a) The graph of the chain is



The chain is reducible, with communicating classes {0, 2}, {1}, {3}.
(b) States 0, 2, 3 have period 1 and state 1 has period 0. States 1 and 3 are transient, states 0 and 2 are recurrent by Theorem 6.9 and Corollary 6.6, and they are also positive recurrent since the state space is finite. There are no absorbing states.

Chapter 7 - Long-Run Behavior of Markov Chains

Exercise 7.5

(a) We have

$$\begin{bmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \end{bmatrix}.$$

(b) By first step analysis and symmetry of the maze we have $\mu_0(1) = \mu_0(3)$, hence

$$\mu_0(0) = 4, \quad \mu_0(1) = 3, \quad \mu_0(2) = 4, \quad \mu_0(3) = 3.$$

The symmetry of the problem shows that we have, $\mu_0(1) = \mu_0(3)$, which greatly simplifies the calculations.

(c) Clearly, the probability distribution $(\pi_0, \pi_1, \pi_2, \pi_3) = (1/4, 1/4, 1/4, 1/4)$ is invariant and satisfies the condition $\pi = \pi P$, see also Exercise 7.14.

Exercise 7.6

(a) Clearly, the transition from the current state to the next state depends only on the current state on the chain, hence the process is Markov. The transition matrix of the chain on the state space $\mathbb{S} = (D, N)$ is

$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = \begin{bmatrix} 1/4 & 3/4 \\ 1/4 & 3/4 \end{bmatrix}.$$

(b) The stationary distribution $\pi = (\pi_D, \pi_N)$ is solution of $\pi = \pi P$ under the condition $\pi_D + \pi_N = 1$, which yields $\pi_D = b/(a+b) = 1/4$ and $\pi_N = a/(a+b) = 3/4$.

(c) In the long run, by the Ergodic Theorem 7.12 we find that the fraction of distorted signals is $\pi_D = 1/4 = 25\%$.

(d) The average time $h_N(D) = \mu_N(D)$ to reach state N starting from state D satisfies

$$\mu_N(D) = (1-a)(1 + \mu_N(D)) + a \tag{B.25}$$

hence $\mu_N(D) = 1/a = 4/3$.

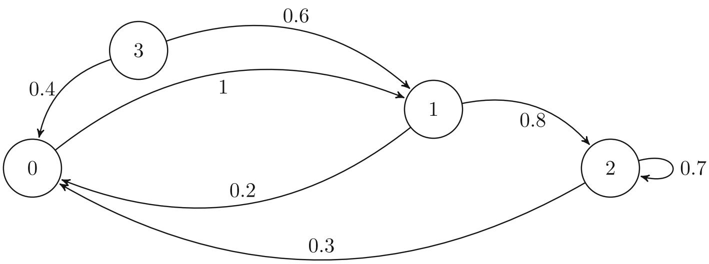
(e) The average time $\mu_D(N)$ to reach state D starting from state N satisfies

$$\mu_D(N) = (1-b)(1 + \mu_D(N)) + b \tag{B.26}$$

hence $\mu_N(D) = 1/b = 4$.

Exercise 7.7

(a) The chain has the following graph:



The chain is reducible and its communicating classes are $\{0, 1, 2\}$ and $\{3\}$.

(b) State 3 is transient because $\mathbb{P}(T_3 = \infty \mid X_0 = 3) = 0.4 + 0.6 = 1$, cf. (6.3.1), and states 0, 1, 2 are recurrent by Theorem 6.9 and Corollary 6.6.

(c) It suffices to consider the subchain on {0, 1, 2} with transition matrix

$$\tilde{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0.2 & 0 & 0.8 \\ 0.3 & 0 & 0.7 \end{bmatrix},$$

and to solve $\pi = \pi \tilde{P}$, i.e. which yields $\pi_1 = \pi_0$ and $0.3\pi_2 = 0.8\pi_1 = 0.8\pi_0$, with $1 = \pi_0 + \pi_1 + \pi_2 = 2\pi_0 + 8\pi_0/3$, i.e. $\pi_0 = 3/14$, $\pi_1 = 3/14$, $\pi_2 = 4/7$, and the fraction of time spent at state 0 in the long run is $3/14 \simeq 0.214$ as the limiting and stationary distributions coincide.

(d) Letting $h_0(k)$ denote the mean hitting time of state 0 starting from state k, we have

$$\begin{cases} h_0(0) = 0 \\ h_0(1) = 0.2(1 + h_0(0)) + 0.8(1 + h_0(2)) \\ h_0(2) = 0.3(1 + h_0(0)) + 0.7(1 + h_0(2)) \\ h_0(3) = 0.4(1 + h_0(0)) + 0.6(1 + h_0(1)), \end{cases}$$

hence $h_0(0) = 0$, $h_0(1) = 11/3$, $h_0(2) = 10/3$, $h_0(3) = 16/5$, and the mean time to reach state 0 starting from state 2 is found to be equal to $h_0(2) = 10/3$, which can also be recovered by pathwise analysis and the geometric series

$$h_0(2) = 0.3 \sum_{k=1}^{\infty} k(0.7)^{k-1} = \frac{0.3}{(1 - 0.7)^2} = \frac{10}{3}.$$

Note that the value of $h_0(2)$ could also be computed by restriction to the subchain {0, 1, 2}, by solving

$$\begin{cases} h_0(0) = 0 \\ h_0(1) = 0.2(1 + h_0(0)) + 0.8(1 + h_0(2)) \\ h_0(2) = 0.3(1 + h_0(0)) + 0.7(1 + h_0(2)). \end{cases}$$

Exercise 7.8

(a) First, we note that the chain has finite state space and it is irreducible, positive recurrent and aperiodic, hence by Theorem 7.8 its limiting distribution coincides with its stationary distribution which is the unique solution of $\pi = \pi P$. After calculations, this equation can be solved as

$$\pi_0 = c \times 161, \qquad \pi_1 = c \times 460, \qquad \pi_2 = c \times 320, \qquad \pi_3 = c \times 170.$$

The condition

$$1 = \pi_0 + \pi_1 + \pi_2 + \pi_3 = c \times 161 + c \times 460 + c \times 320 + c \times 170$$

shows that

$$\pi_0 = \frac{161}{1111}, \qquad \pi_1 = \frac{460}{1111}, \qquad \pi_2 = \frac{320}{1111}, \qquad \pi_3 = \frac{170}{1111}. \tag{B.27}$$

(b) We choose to solve this problem using mean return times since $\mu_0(i) = h_0(i)$, $i = 1, 2, 3$, however it could also be solved using mean hitting times $h_i(j)$. We have

$$\begin{cases} \mu_0(0) = 1 + \mu_0(1), \\ \\ \mu_0(1) = 0.1 + 0.4(1 + \mu_0(1)) + 0.2(1 + \mu_0(2)) + 0.3(1 + \mu_0(3)) \\ \qquad\quad = 1 + 0.4\mu_0(1) + 0.2\mu_0(2) + 0.3\mu_0(3), \\ \\ \mu_0(2) = 0.2 + 0.2(1 + \mu_0(1)) + 0.5(1 + \mu_0(2)) + 0.1(1 + \mu_0(3)) \\ \qquad\quad = 1 + 0.2\mu_0(1) + 0.5\mu_0(2) + 0.1\mu_0(3), \\ \\ \mu_0(3) = 1 + 0.3\mu_0(1) + 0.4\mu_0(2), \end{cases}$$

hence

$$\mu_0(1) = \frac{950}{161}, \qquad \mu_0(2) = \frac{860}{161}, \qquad \mu_0(3) = \frac{790}{161}.$$

Note that the data of the first row in the transition matrix is not needed in order to compute the mean return times.

(c) We find

$$\mu_0(0) = 1 + \mu_0(1) = 1 + \frac{950}{161} = \frac{161 + 950}{161} = \frac{1111}{161},$$

hence the relation $\pi_0 = 1/\mu_0(0)$ is satisfied from (B.27).

Exercise 7.9

(a) All states of this chain have period 2.
(b) The chain is irreducible and it has a finite state space, hence it is positive recurrent from Theorem 6.11. By Proposition 6.14, all states have period 2 hence the chain is not aperiodic, and for this reason Theorem 7.2 and Theorem 7.8 cannot be used and the chain actually has no limiting distribution. Nevertheless, Theorem 7.10 applies and shows that the equation $\pi = \pi P$ characterizes the stationary distribution.

Exercise 7.10 We choose to model the problem on the state space $\{1, 2, 3, 4\}$, meaning that the replacement of a component is immediate upon failure. Let X_n denote the remaining active time of the component at time n. Given that at time n there remains $X_n = k \geq 2$ units of time until failure, we know with certainty that at the next time step $n + 1$ there will remain $X_{n-1} = k - 1 \geq 1$ units of time until failure. Hence at any time $n \geq 1$ we have

$$X_n = 4 \quad \Longrightarrow X_{n+1} = 3 \quad \Longrightarrow X_{n+2} = 2 \quad \Longrightarrow X_{n+3} = 1,$$

whereas when $X_n = 1$ the component will become inactive at the next time step and will be immediately replaced by a new component of random lifetime $T \in \{1, 2, 3\}$. Hence we have

$$\mathbb{P}(X_{n+1} = k \mid X_n = 1) = \mathbb{P}(T = k), \qquad k = 1, 2, 3, 4,$$

and the process $(X_n)_{n\in\mathbb{N}}$ is a Markov chain on $\mathbb{S} = \{1, 2, 3, 4\}$, with transition matrix

$$P = \begin{bmatrix} \mathbb{P}(Y=1) & \mathbb{P}(Y=2) & \mathbb{P}(Y=3) & \mathbb{P}(Y=4) \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

We now look for the limit $\lim_{n\to\infty} \mathbb{P}(X_n = 1)$. Since the chain is irreducible, aperiodic (all states are checked to have period one) and its state space is finite, we know by Theorem 7.8 that $\pi_1 = \lim_{n\to\infty} \mathbb{P}(X_n = 1)$, where $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$ is the stationary distribution π uniquely determined from the equation $\pi = \pi P$, as follows:

$$\begin{cases} \pi_1 = 0.1\pi_1 + \pi_2 \\ \\ \pi_2 = 0.2\pi_1 + \pi_3 \\ \\ \pi_3 = 0.3\pi_1 + \pi_4 \\ \\ \pi_4 = 0.4\pi_1. \end{cases}$$

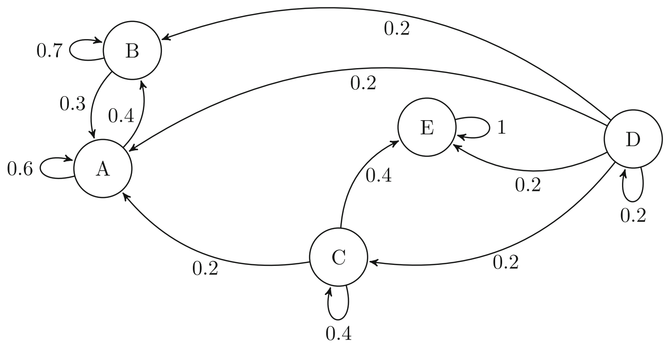
hence $\pi_2 = 0.9\pi_1$, $\pi_3 = 0.7\pi_1$, $\pi_4 = 0.4\pi_1$, under the condition $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$, i.e. $\pi_1 + 0.9\pi_1 + 0.7\pi_1 + 0.4\pi_1 = 1$, which yields $\pi_1 = 1/3$, $\pi_2 = 9/30$, $\pi_3 = 7/30$, $\pi_4 = 4/30$. This result can be confirmed by computing the limit of matrix powers $(P^n)_{n\in\mathbb{N}}$ as n tends to infinity using the following Matlab/Octave commands:

```
P = [0.1,0.2,0.3,0.4;
1,0,0,0;
0,1,0,0;
0,0,1,0;]
mpower(P,1000)
```

showing that

$$\lim_{n\to\infty} P^n = \begin{bmatrix} 0.33333 & 0.30000 & 0.23333 & 0.13333 \\ 0.33333 & 0.30000 & 0.23333 & 0.13333 \\ 0.33333 & 0.30000 & 0.23333 & 0.13333 \\ 0.33333 & 0.30000 & 0.23333 & 0.13333 \end{bmatrix}.$$

Exercise 7.11 The graph of the chain is as follows:



We note that the chain is reducible, and that its state space $\mathbb{S}$ can be partitioned into 4 communicating classes:

$$\mathbb{S} = \{A, B\} \cup \{C\} \cup \{D\} \cup \{E\},$$

where A, B are recurrent, E is absorbing, and C, D are transient.

Starting from state C, one can only return to C or end up in one of the absorbing classes $\{A, B\}$ or $\{E\}$. Let us denote by

$$T_{\{A,B\}} = \inf\{n \geq 0 \ : \ X_n \in \{A, B\}\}$$

the hitting time of $\{A, B\}$. We start by computing $\mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C)$. By first step analysis we find that this probability satisfies

$$\mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C) = 0.2 + 0.4 \times \mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C) + 0.4 \times 0,$$

hence $\mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C) = 1/3$, On the other hand, $\{A, B\}$ is a closed two-state chain with transition matrix

$$\begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix},$$

hence, starting from any state within $\{A, B\}$, the long run probability of being in A is given by

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}(X_n = A \mid X_0 \in \{A, B\}) \\
&\quad = \lim_{n\to\infty} \mathbb{P}(X_n = A \mid X_0 = A) \frac{\mathbb{P}(X_0 = A)}{\mathbb{P}(X_0 \in \{A, B\})} \\
&\qquad + \lim_{n\to\infty} \mathbb{P}(X_n = A \mid X_0 = B) \frac{\mathbb{P}(X_0 = B)}{\mathbb{P}(X_0 \in \{A, B\})} \\
&\quad = \frac{b}{a+b} \left(\frac{\mathbb{P}(X_0 = A)}{\mathbb{P}(X_0 \in \{A, B\})} + \frac{\mathbb{P}(X_0 = B)}{\mathbb{P}(X_0 \in \{A, B\})} \right)
\end{aligned}$$

$$= \frac{0.3}{0.3 + 0.4} = \frac{3}{7}.$$

Since

$$\{X_n = A\} \subset \{T_{A,B} \leq n\} \subset \{T_{A,B} < \infty\}, \qquad n \in \mathbb{N},$$

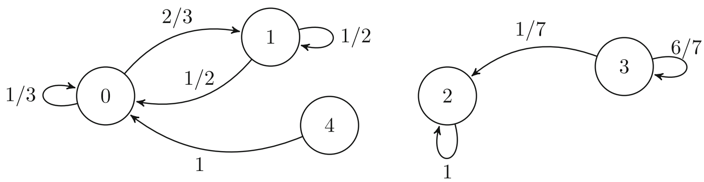
we conclude that

$$\begin{aligned}
\alpha &:= \lim_{n\to\infty} \mathbb{P}(X_n = A \mid X_0 = C) \\
&= \lim_{n\to\infty} \mathbb{P}\big(T_{A,B} < \infty \text{ and } X_n = A \mid X_0 = C\big) \\
&= \lim_{n\to\infty} \mathbb{P}\big(T_{A,B} < \infty \text{ and } X_{n+T_{A,B}} = A \mid X_0 = C\big) \\
&= \mathbb{P}\big(T_{A,B} < \infty \mid X_0 = C\big) \lim_{n\to\infty} \mathbb{P}\big(X_{n+T_{A,B}} = A \mid T_{A,B} < \infty \text{ and } X_0 = C\big) \\
&= \mathbb{P}\big(T_{A,B} < \infty \mid X_0 = C\big) \\
&\qquad \times \lim_{n\to\infty} \mathbb{P}\big(X_{n+T_{A,B}} = A \mid T_{A,B} < \infty,\ X_{T_{A,B}} \in \{A, B\},\ X_0 = C\big) \\
&= \mathbb{P}\big(T_{\{A,B\}} < \infty \mid X_0 = C\big) \lim_{n\to\infty} \mathbb{P}(X_n = A \mid X_0 \in \{A, B\}) \\
&= \frac{1}{3} \times \frac{3}{7} = \frac{1}{7},
\end{aligned}$$

where we used the strong Markov property, cf. Exercise 5.8.

Exercise 7.12

(a) The chain has the following graph:



(b) The communicating classes are {0, 1}, {2}, {3}, and {4}.
(c) States 3 and 4 are transient, states 0 and 1 are recurrent, and state 2 is absorbing (hence it is recurrent).
(d) By (4.5.7) we have

$$\lim_{n\to\infty} \mathbb{P}(X_n = 0 \mid X_0 = 4) = \lim_{n\to\infty} \mathbb{P}(X_n = 0 \mid X_0 = 0) = \frac{1/2}{2/3 + 1/2} = \frac{3}{7},$$

cf. also the Table 8.1.

Exercise 7.13

(a) The transition matrix P of the chain on the state space $\mathbb{S} = (C, T)$ is given by

$$\begin{bmatrix} 4/5 & 1/5 \\ 3/4 & 1/4 \end{bmatrix}.$$

(b) The stationary distribution $\pi = (\pi_C, \pi_T)$ is solution of $\pi = \pi P$ under the condition $\pi_C + \pi_T = 1$, which yields $\pi_C = 15/19$ and $\pi_T = 4/19$.

(c) In the long run, applying the Ergodic Theorem 7.12 we find that 4 out of 19 vehicles are trucks.

(d) Let $\mu_T(C)$ and $\mu_T(T)$ denote the mean return times to state (T) starting from (C) and (T), respecticely. By first step analysis we have

$$\begin{cases} \mu_T(C) = 1 + \dfrac{4}{5}\mu_T(C) \\ \\ \mu_T(T) = 1 + \dfrac{3}{4}\mu_T(C) \end{cases}$$

which has for solution $\mu_T(C) = 5$ and $\mu_T(T) = 19/4$.

Exercise 7.15

(a) We solve the system of equations

$$\begin{cases} \pi_0 = q(\pi_0 + \pi_1 + \pi_2 + \pi_3) + \pi_4 = q + p\pi_4 \\ \\ \pi_1 = p\pi_0 \\ \\ \pi_2 = p\pi_1 = p^2\pi_0 \\ \\ \pi_3 = p\pi_2 = p^3\pi_0 \\ \\ \pi_4 = p\pi_3 = p^4\pi_0, \end{cases}$$

which yields $1 = \pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 = \pi_0(1 + p + p^2 + p^3 + p^4)$, and

$$\begin{cases} \pi_0 = \dfrac{1}{1 + p + p^2 + p^3 + p^4} \\ \pi_1 = \dfrac{p}{1 + p + p^2 + p^3 + p^4} \\ \pi_2 = \dfrac{p^2}{1 + p + p^2 + p^3 + p^4} \\ \pi_3 = \dfrac{p^3}{1 + p + p^2 + p^3 + p^4} \\ \pi_4 = \dfrac{p^4}{1 + p + p^2 + p^3 + p^4}. \end{cases}$$

(b) Since the chain is irreducible and aperiodic with finite state space, its limiting distribution coincides with its stationary distribution.

Exercise 7.16

(a) The transition matrix P is given by

$$P = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/3 & 0 & 1/3 & 1/3 \\ 0 & 1 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix}.$$

(b) The chain is aperiodic, irreducible, and has finite state space hence we can apply Theorem 7.8 or Theorem 7.10. The equation $\pi P = \pi$ reads

$$\pi P = [\pi_A, \pi_B, \pi_C, \pi_D] \times \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/3 & 0 & 1/3 & 1/3 \\ 0 & 1 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3}\pi_B + \frac{1}{2}\pi_D \\ \frac{1}{2}\pi_A + \pi_C + \frac{1}{2}\pi_D \\ \frac{1}{3}\pi_B \\ \frac{1}{2}\pi_A + \frac{1}{3}\pi_B \end{bmatrix}$$

$$= [\pi_A, \pi_B, \pi_C, \pi_D],$$

i.e. $\pi_A = \pi_D = 2\pi_C$ and $\pi_B = 3\pi_C$, which, under the condition $\pi_A + \pi_B + \pi_C + \pi_D = 1$, gives $\pi_A = 1/4$, $\pi_B = 3/8$, $\pi_C = 1/8$, $\pi_D = 1/4$.

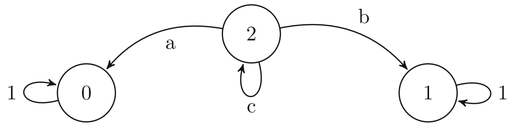
(c) We solve the system

$$\begin{cases} \mu_D(A) = \frac{1}{2} + \frac{1}{2}(1 + \mu_D(B)) = 1 + \frac{1}{2}\mu_D(B) \\ \mu_D(B) = \frac{1}{3} + \frac{1}{3}(1 + \mu_D(A)) + \frac{1}{3}(1 + \mu_D(C)) = 1 + \frac{1}{3}(\mu_D(A) + \mu_D(C)) \\ \mu_D(C) = 1 + \mu_D(B) \\ \mu_D(D) = \frac{1}{2}(1 + \mu_D(A)) + \frac{1}{2}(1 + \mu_D(B)) = 1 + \frac{1}{2}(\mu_D(A) + \mu_D(B)), \end{cases}$$

which has for solution $\mu_D(A) = 8/3$, $\mu_D(B) = 10/3$, $\mu_D(C) = 13/3$, $\mu_D(D) = 4$. On average, player D has to wait $\mu_D(D) = 4$ time units before recovering the token.

(d) This probability is $\pi_D = 0.25$, and we check that the relation $\mu(D) = 1/\pi_D = 4$ is satisfied.

Exercise 7.17 Clearly we may assume that $c < 1$, as the case $c = 1$ corresponds to the identity matrix, or to constant a chain. On the other hand, we cannot directly apply Theorem 7.8 since the chain is reducible. The chain has the following graph:



(a) By observation of

$$P^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a(1+c) & b(1+c) & c^2 \end{bmatrix}$$

and

$$P^3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a(1+c+c^2) & b(1+c+c^2) & c^3 \end{bmatrix}$$

we infer that P^n takes the general form

$$P^n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_n & b_n & c_n \end{bmatrix},$$

where a_n, b_n and c_n are coefficients to be determined by the following induction argument. Writing down the relation $P^{n+1} = P \times P^n$ as

$$P^{n+1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_{n+1} & b_{n+1} & c_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & b & c \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_n & b_n & c_n \end{bmatrix}$$

shows that we have the recurrence relations

$$\begin{cases} a_{n+1} = a + ca_n, \\ \\ b_{n+1} = b + cb_n, \\ \\ c_{n+1} = c \times c_n, \end{cases} \quad \text{which yield} \quad \begin{cases} a_n = a + ac + \cdots + ac^{n-1} = a\dfrac{1-c^n}{1-c}, \\ \\ b_n = b + bc + \cdots + bc^{n-1} = b\dfrac{1-c^n}{1-c}, \\ \\ c_{n+1} = c^n, \end{cases}$$

hence

$$P^n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a\dfrac{1-c^n}{1-c} & b\dfrac{1-c^n}{1-c} & c^n \end{bmatrix}.$$

(b) From the structure of P^n it follows that the chain admits a limiting distribution

$$\lim_{n\to\infty} P^n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \dfrac{a}{1-c} & \dfrac{b}{1-c} & 0 \end{bmatrix}.$$

which is dependent of the initial state, provided that $c < 1$. The limiting probabilities

$$\lim_{n\to\infty} \mathbb{P}(X_n = 0 \mid X_0 = 2) = \frac{a}{1-c},$$

resp.

$$\lim_{n\to\infty} \mathbb{P}(X_n = 1 \mid X_0 = 2) = \frac{b}{1-c},$$

correspond to the probability of moving to state 0, resp. the probability of moving to state 1, given one does not return to state 2.

In addition we have $\mathbb{P}(T_2^r = \infty \mid X_0 = 2) = a + b > 0$, hence state 2 is transient and the chain is not recurrent.

(c) By solving the equation $\pi = \pi P$ we find that the chain admits an infinity of stationary distributions of the form $(\pi_0, \pi_1, 0)$ with $\pi_0 + \pi_1 = 1$ when $c < 1$. We also note that here, all limiting distributions obtained in Question (b) are also stationary distributions on every row.

Exercise 7.18

(a) The process $(X_n)_{n\in\mathbb{N}}$ is a two-state Markov chain on $\{0, 1\}$ with transition matrix

$$\begin{bmatrix} \alpha & \beta \\ p & q \end{bmatrix}$$

and $\alpha = 1 - \beta$. The entries on the second line are easily obtained. Concerning the first line we note that $\mathbb{P}(N = 1) = \beta$ is the probability of switching from 0 to 1 in one time step, while the equality $\mathbb{P}(N = 2) = \beta(1 - \beta)$ shows that the probability of remaining at 0 for one time step is $1 - \beta$.

(b) This probability is given from the stationary distribution (π_0, π_1) as $\pi_1 = \beta/(p + \beta)$.

Exercise 7.20

(a) The $N \times N$ transition matrix of the chain is

$$P = \begin{bmatrix} q & p & 0 & \cdots & \cdots & 0 & 0 & 0 \\ q & 0 & p & \cdots & \cdots & 0 & 0 & 0 \\ 0 & q & 0 & \ddots & \ddots & 0 & 0 & 0 \\ 0 & 0 & q & \ddots & \ddots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & \ddots & 0 & p & 0 \\ 0 & 0 & 0 & \cdots & \cdots & q & 0 & p \\ 0 & 0 & 0 & \cdots & \cdots & 0 & q & p \end{bmatrix}.$$

(b) The chain is irreducible if $p \in (0, 1)$, and reducible if $p = 0$ or $p = 1$.

(c) If $p \in (0, 1)$ there are no absorbing states and all states are positive recurrent. If $p = 0$, state 1 is absorbing and all other states are transient. If $p = 1$, state N is absorbing and all other states are transient.

(d) The equation $\pi = \pi P$ yields

$$\pi_2 = \frac{p}{q}\pi_1 \quad \text{and} \quad \pi_N = \frac{p}{q}\pi_{N-1}, \qquad k = 1, 2, \ldots, N-1,$$

and

$$p(\pi_k - \pi_{k-1}) = q(\pi_{k+1} - \pi_k) \qquad k = 2, 3, \ldots, N-1.$$

We check that π_k given by

$$\pi_k = \frac{p^{k-1}}{q^{k-1}}\pi_1, \qquad k = 1, 2, \ldots, N,$$

satisfies the above conditions. The normalization condition

$$1 = \sum_{k=1}^{N} \pi_k = \pi_1 \sum_{k=1}^{N} \frac{p^{k-1}}{q^{k-1}} = \pi_1 \sum_{k=0}^{N-1} \left(\frac{p}{q}\right)^k = \pi_1 \frac{1 - (p/q)^N}{1 - p/q}$$

shows that

$$\pi_k = \frac{1 - p/q}{1 - (p/q)^N} \frac{p^{k-1}}{q^{k-1}}, \qquad k = 1, 2, \ldots, N,$$

provided that $0 < p \neq q < 1$. When $p = q = 1/2$ we find that the uniform distribution

$$\pi_k = \frac{1}{N}, \qquad k = 1, 2, \ldots, N,$$

is stationary. When $p = 0$ the stationary distribution is $\mathbb{1}_{\{0\}} = [1, 0, \ldots, 0, 0]$, and when $p = 1$ it is $\mathbb{1}_{\{N\}} = [0, 0, \ldots, 0, 1]$.

(e) The chain has finite state space and when $p \in (0, 1)$ it is irreducible and aperiodic, hence its limiting distribution coincides with its stationary distribution.

Problem 7.23

(a) We have

$$\mathbb{P}(Y_{n+1} = j \mid Y_n = i) = \frac{\mathbb{P}(Y_{n+1} = j \text{ and } Y_n = i)}{\mathbb{P}(Y_n = i)}$$
$$= \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i)}\mathbb{P}(X_{N-n} = i \mid X_{N-n-1} = j) = \frac{\pi_j}{\pi_i} P_{j,i}.$$

On the other hand, we have

$$\mathbb{P}(Y_{n+1} = j \mid Y_n = i_n, \ldots, Y_0 = i_0) = \frac{\mathbb{P}(Y_{n+1} = j, Y_n = i_n, \ldots, Y_0 = i_0)}{\mathbb{P}(Y_n = i_n, \ldots, Y_0 = i_0)}$$
$$= \frac{\mathbb{P}(X_{N-n-1} = j, X_{N-n} = i_n, \ldots, X_N = i_0)}{\mathbb{P}(X_{N-n} = i_n, \ldots, X_N = i_0)}$$
$$= \mathbb{P}(X_{N-n-1} = j \text{ and } X_{N-n} = i_n)$$
$$\times \frac{\mathbb{P}(X_{N-n+1} = i_{n-1}, \ldots, X_N = i_0 \mid X_{N-n-1} = j, X_{N-n} = i_n)}{\mathbb{P}(X_{N-n+1} = i_{n-1}, \ldots, X_N = i_0)}$$
$$= \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i_n)}\mathbb{P}(X_{N-n} = i_n \mid X_{N-n-1} = j) = \frac{\pi_j}{\pi_{i_n}} P_{j,i_n},$$

and this shows that

$$\mathbb{P}(Y_{n+1} = j \mid Y_n = i_n, \ldots, Y_0 = i_0) = \mathbb{P}(Y_{n+1} = j \mid Y_n = i_n) = \frac{\pi_j}{\pi_{i_n}} P_{j,i_n},$$

i.e. the time-reversed process $(Y_n)_{n=0,1,\ldots,N}$ has the Markov property.

(b) We find

$$P_{i,j} = \frac{\pi_i}{\pi_j} P_{j,i}, \tag{B.28}$$

i.e.

$$\pi_i P_{i,j} = \pi_j P_{j,i},$$

which is the *detailed balance condition* with respect to the probability distribution $\pi = (\pi_i)_{i \in S}$.

(c) We have

$$\pi_j = \sum_i \pi_j P_{j,i} = \sum_i \pi_i P_{i,j} = [\pi P]_j.$$

(d) According to the detailed balance condition (B.28) we have

$$P_{k_1,k_2} P_{k_2,k_3} \cdots P_{k_n,k_1} = P_{k_n,k_1} \prod_{i=1}^{n-1} P_{k_i,k_{i+1}} = P_{k_n,k_1} \prod_{i=1}^{n-1} \frac{\pi_{k_{i+1}}}{\pi_{k_i}} P_{k_{i+1},k_i}$$

$$= \frac{\pi_{k_n}}{\pi_{k_1}} P_{k_n,k_1} \prod_{i=1}^{n-1} P_{k_{i+1},k_i} = P_{k_1,k_n} \prod_{i=1}^{n-1} P_{k_{i+1},k_i},$$

holds for all sequences $\{k_1, k_2, \ldots, k_n\}$ of states and $n \geq 2$.

(e) If the Markov chain satisfies

$$P_{k_1,k_2} P_{k_2,k_3} \cdots P_{k_{n-1},k_n} P_{k_n,k_1} = P_{k_1,k_n} P_{k_n,k_{n-1}} \cdots P_{k_3,k_2} P_{k_2,k_1}$$

then by summation over the indexes $k_2, k_3, \ldots, k_{n-1}$, using the matrix power relation

$$[P^{n-1}]_{i,j} = \sum_{k_2,\ldots,k_{n-1}} P_{i,k_2} P_{k_2,k_3} \cdots P_{k_{n-1},j},$$

we get

$$[P^{n-1}]_{k_1,k_n} P_{k_n,k_1} = P_{k_1,k_n} [P^{n-1}]_{k_n,k_1}.$$

On the other hand, by taking the limit as n goes to infinity Theorem 7.8 shows that

$$\lim_{n\to\infty} [P^{n-1}]_{k_n,k_1} = \lim_{n\to\infty} [P^n]_{k_n,k_1} = \pi_{k_1}$$

since the limiting and stationary distributions coincide, and we get

$$\pi_{k_n} P_{k_n,k_1} = P_{k_1,k_n} \pi_{k_1},$$

which is the detailed balance condition.

(f) The detailed balance condition reads

$$\pi_i P_{i,i+1} = \pi_i \left(\frac{1}{2} - \frac{i}{2M} \right) = \pi_{i+1} P_{i+1,i} = \pi_{i+1} \frac{i+1}{2M},$$

hence

$$\frac{\pi_{i+1}}{\pi_i} = \frac{1 - i/M}{(i+1)/M} = \frac{M-i}{i+1},$$

which shows that

$$\pi_i = \frac{(M-i+1)}{i} \frac{(M-i+2)}{(i-1)} \cdots \frac{(M-1)}{2} \frac{M}{1} \pi_0 = \frac{M!}{i!(M-i)!} \pi_0 = \pi_0 \binom{M}{i},$$

$i = 0, 1, \ldots, M$, where the constant $\pi_0 > 0$ is given by

$$1 = \sum_i \pi_i = \pi_0 \sum_{i=0}^{M} \binom{M}{i} = \pi_0 2^M,$$

hence $\pi_0 = 2^{-M}$ and

$$\pi_i = \frac{1}{2^M}\binom{M}{i}, \qquad i = 0, 1, \ldots, M.$$

(g) We have

$$\begin{aligned}
[\pi P]_i &= P_{i+1,i}\pi_{i+1} + P_{i,i}\pi_i + P_{i-1,i}\pi_{i-1} \\
&= \frac{1}{2^M}\frac{i+1}{2M}\binom{M}{i+1} + \frac{1}{2^M}\left(\frac{1}{2} - \frac{i-1}{2M}\right)\binom{M}{i-1} + \frac{1}{2^M} \times \frac{1}{2}\binom{M}{i} \\
&= \frac{1}{2^M}\binom{M}{i},
\end{aligned}$$

which is also known as *Pascal's triangle*.

(h) The chain is positive recurrent, irreducible and aperiodic, therefore by Theorem 7.8 it admits a limiting distribution equal to π.

(i) We have

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}[X_n \mid X_0 = i] &= \lim_{n\to\infty} \sum_{j=0}^{M} j\mathbb{P}(X_n = j \mid X_0 = i) \\
&= \sum_{j=0}^{M} j \lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i) = \sum_{j=0}^{M} j\pi_j \\
&= \frac{1}{2^M}\sum_{j=0}^{M} j\binom{M}{j} = \frac{M}{2^M}\sum_{j=0}^{M-1}\frac{(M-1)!}{j!(M-1-j)!} = \frac{M}{2},
\end{aligned}$$

independently of $i = 0, 1, \ldots, M$.

(j) Clearly, the relation

$$\mathbb{E}\left[X_0 - \frac{M}{2}\,\Big|\, X_0 = i\right] = i - \frac{M}{2}$$

holds when $n = 0$. Next, assuming that the relation holds at the rank $n \geq 0$ we have

$$\begin{aligned}
h(i) &= \mathbb{E}\left[X_{n+1} - \frac{M}{2}\,\Big|\, X_0 = i\right] \\
&= P_{i,i+1}\mathbb{E}\left[X_{n+1} - \frac{M}{2}\,\Big|\, X_1 = i+1\right] + P_{i,i}\mathbb{E}\left[X_{n+1} - \frac{M}{2}\,\Big|\, X_1 = i\right] \\
&\quad + P_{i,i-1}\mathbb{E}\left[X_{n+1} - \frac{M}{2}\,\Big|\, X_1 = i-1\right]
\end{aligned}$$

$$= \left(\frac{1}{2} - \frac{i}{2M}\right) \mathbb{E}\left[X_{n+1} - \frac{M}{2} \,\Big|\, X_1 = i+1\right] + \frac{1}{2}\mathbb{E}\left[X_{n+1} - \frac{M}{2} \,\Big|\, X_1 = i\right]$$
$$+ \frac{i}{2M}\mathbb{E}\left[X_{n+1} - \frac{M}{2} \,\Big|\, X_1 = i-1\right]$$
$$= \left(\frac{1}{2} - \frac{i}{2M}\right)\left(i+1-\frac{M}{2}\right)\left(1-\frac{1}{M}\right)^n + \frac{1}{2}\left(i-\frac{M}{2}\right)\left(1-\frac{1}{M}\right)^n$$
$$+ \frac{i}{2M}\left(i-1-\frac{M}{2}\right)\left(1-\frac{1}{M}\right)^n = \left(i-\frac{M}{2}\right)\left(1-\frac{1}{M}\right)^{n+1}, \qquad n \geq 0,$$

for all $i = 0, 1, \ldots, M$.

Taking the limit as n goes to infinity we get

$$\lim_{n\to\infty} \mathbb{E}\left[X_n - \frac{M}{2} \,\Big|\, X_0 = i\right] = \lim_{n\to\infty}\left(i-\frac{M}{2}\right)\left(1-\frac{1}{M}\right)^n = 0,$$

hence $\lim_{n\to\infty} \mathbb{E}[X_n] = M/2$, for all $i = 0, 1, \ldots, M$, which recovers the result of Question (i).

Chapter 8 - Branching Processes

Exercise 8.2

(a) We have

$$G_1(s) = \mathbb{E}[s^Y] = s^0\mathbb{P}(Y=0) + s^1\mathbb{P}(Y=1) = \frac{1}{2} + \frac{1}{2}s, \qquad s \in \mathbb{R}.$$

(b) We prove this statement by induction. Clearly it holds at the order 1. Next, assuming that (8.3.10) holds at the order $n \geq 1$ we get

$$G_{n+1}(s) = G_1(G_n(s)) = G_1\left(1 - \frac{1}{2^n} + \frac{s}{2^n}\right)$$
$$= \frac{1}{2} + \frac{1}{2}\left(1 - \frac{1}{2^n} + \frac{s}{2^n}\right) = 1 - \frac{1}{2^{n+1}} + \frac{s}{2^{n+1}}.$$

(c) We have

$$\mathbb{P}(X_n = 0 \mid X_0 = 1) = G_n(0) = 1 - \frac{1}{2^n}.$$

(d) We have

$$\mathbb{E}[X_n \mid X_0 = 1] = G_n'(s)_{|s=1} = (\mathbb{E}[Y_1])^n = \frac{1}{2^n}.$$

(e) The extinction probability α is solution of $G_1(\alpha) = \alpha$, i.e. $\alpha = 1/2 + \alpha/2$, with unique solution $\alpha = 1$.

Exercise 8.3

(a) We have $G_1(s) = 0.2 + 0.5s + 0.3s^2$ and

$$\mathbb{E}[X_1] = \mathbb{E}[\xi] = G_1'(1) = 0.5 + 2 \times 0.3 = 1.1,$$

hence

$$\mathbb{E}[X_2] = (G_1'(1))^2 = (\mathbb{E}[\xi])^2 = (1.1)^2,$$

by Proposition 8.2. On the other hand, we have

$$\begin{aligned} G_2(s) &= G_1(G_1(s)) \\ &= G_1(0.2 + 0.5s + 0.3s^2) \\ &= 0.312 + 0.31s + 0.261s^2 + 0.09s^3 + 0.027s^4, \end{aligned}$$

with

$$G_2'(s) = 0.31 + 0.522s + 0.27s^2 + 0.108s^3$$

and

$$G_2''(s) = 0.522 + 0.54s + 0.324s^2,$$

hence

$$G_2'(1) = (G_1'(1))^2 = (1.1)^2 = 1.21 \quad \text{and} \quad G_2''(1) = 1.386,$$

and

$$\mathbb{E}[X_2^2] = G_2''(1) + G_2'(1) = 1.386 + 1.21 = 2.596.$$

By (1.7.6) this yields

$$\mathrm{Var}[X_2] = 2.596 - (1.21)^2.$$

(b) We have $G_2(s) = 0.312 + 0.31s + 0.261s^2 + 0.09s^3 + 0.027s^4$, hence

$$\mathbb{P}(X_2 = 0) = 0.312, \quad \mathbb{P}(X_2 = 1) = 0.31, \quad \mathbb{P}(X_2 = 2) = 0.261,$$

and

$$\mathbb{P}(X_2 = 3) = 0.09, \quad \mathbb{P}(X_2 = 4) = 0.027.$$

(c) We have $\mathbb{P}(X_4 = 0) = G_4(0) = G_2(G_2(0)) \simeq 0.44314$.
(d) We have $\mathbb{E}[X_{10}] = (\mathbb{E}[X_1])^{10} = (G_1'(1))^{10} = (1.1)^{10} = 2.59$, since the mean population size grows by 10% at each time step.

(e) The extinction probability α solves the equation

$$\alpha = G_1(\alpha) = 0.2 + 0.5\alpha + 0.3\alpha^2,$$

i.e. $0.3\alpha^2 - 0.5\alpha + 0.2 = 0.3(\alpha - 1)(\alpha - 2/3) = 0$, hence $\alpha = 2/3$.

Exercise 8.4

(a) We have $G_1(s) = \mathbb{P}(Y = 0) + s\mathbb{P}(Y = 1) + s^2\mathbb{P}(Y = 2) = as^2 + bs + c$, $s \in \mathbb{R}$.

(b) Letting X_n denote the number of individuals in the population at generation $n \geq 0$, we have

$$\mathbb{P}(X_2 = 0 \mid X_0 = 1) = G_1(G_1(0)) = G_1(c) = ac^2 + bc + c.$$

This probability can actually be recovered by pathwise analysis, by noting that in order to reach $\{X_2 = 0\}$ we should have either

(i) $Y_1 = 0$ with probability c, or
(ii) $Y_1 = 1$ with probability b and then $Y_1 = 0$ with probability c, or
(iii) $Y_1 = 2$ with probability a and then $Y_1 = 0$ (two times) with probability c,

which yields $\mathbb{P}(X_2 = 0 \mid X_0 = 1) = c + bc + ac^2$.

(c) We have

$$\mathbb{P}(X_2 = 0 \mid X_0 = 2) = (\mathbb{P}(X_2 = 0 \mid X_0 = 1))^2 = (ac^2 + bc + c)^2,$$

as in (8.3.1).

(d) The extinction probability α_1 given that $X_0 = 1$ is solution of $G_1(\alpha) = \alpha$, i.e.

$$a\alpha^2 + b\alpha + c = \alpha,$$

or

$$0 = a\alpha^2 - (a + c)\alpha + c = (\alpha - 1)(a\alpha - c)$$

from the condition $a + b + c = 1$. The extinction probability α_1 is known to be the smallest solution of $G_1(\alpha) = \alpha$, hence it is $\alpha_1 = c/a$ when $0 < c \leq a$. The extinction probability α_2 given that $X_0 = 2$ is $\alpha_2 = (\alpha_1)^2$.

(e) When $0 \leq a \leq c$ we have $\alpha_1 = 1$.

Exercise 8.6

(a) When only red cells are generated, their number at time $n - 1$ is s^{n-1}, hence the probability that only red cells are generated up to time n is

$$\frac{1}{4} \times \left(\frac{1}{4}\right)^2 \times \cdots \times \left(\frac{1}{4}\right)^{2^{n-1}} = \prod_{k=0}^{n-1} \left(\frac{1}{4}\right)^{2^k} \left(\frac{1}{4}\right)^{2^n - 1}, \qquad n \geq 0.$$

(b) Since white cells cannot reproduce, the extinction of the culture is equivalent to the extinction of the red cells, and this question can be solved as in the framework of Exercise 8.3. The probability distribution of the number Y of red cells produced from one red cell is

$$\mathbb{P}(Y=0)=\frac{1}{12},\qquad \mathbb{P}(Y=1)=\frac{2}{3},\qquad \mathbb{P}(Y=2)=\frac{1}{4},$$

which has the generating function

$$\begin{aligned}G_1(s)&=\mathbb{P}(Y=0)+s\mathbb{P}(Y=1)+s^2\mathbb{P}(Y=2)\\&=\frac{1}{12}+\frac{2s}{3}+\frac{s^2}{4}=\frac{1}{12}(1+8s+3s^2),\end{aligned}$$

hence the equation $G_1(\alpha)=\alpha$ reads

$$3\alpha^2-4\alpha+1=3(\alpha-1)(\alpha-1/3)=0,$$

which has $\alpha=1/3$ for smallest solution. Consequently, the extinction probability of the culture is equal to $1/3$.

(c) The probability that only red cells are generated from time 0 to time n is

$$\frac{1}{3}\times\left(\frac{1}{3}\right)^2\times\cdots\times\left(\frac{1}{3}\right)^{2^{n-1}}=\prod_{k=0}^{n-1}\left(\frac{1}{3}\right)^{2^k}=\left(\frac{1}{3}\right)^{2^n-1},$$

$n\geq 0$. The probability distribution

$$\mathbf{P}(Y=0)=\frac{1}{6},\qquad \mathbf{P}(Y=1)=\frac{2}{2},\qquad \mathbf{P}(Y=2)=\frac{1}{3},$$

of the number Y of red cells has the generating function

$$\begin{aligned}G_1(s)&=\mathbf{P}(Y=0)+s\mathbf{P}(Y=1)+s^2\mathbf{P}(Y=2)\\&=\frac{1}{6}+\frac{s}{2}+\frac{s^2}{3}=\frac{1}{12}(2+6s+4s^2),\end{aligned}$$

hence the equation $G_1(\alpha)=\alpha$ reads $1+3\alpha+2\alpha^2=6\alpha$, or

$$2\alpha^2-3\alpha+1=2(\alpha-1)(\alpha-1/2)=0,$$

which has $\alpha=1/2$ for smallest solution. Consequently, the extinction probability of the culture is equal to $1/2$.

Exercise 8.9

(a) We have $\mathbb{P}(X=k)=(1/2)^{k+1}$, $k\in\mathbb{N}$.

(b) The probability generating function of X is given by

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{\infty} s^k \mathbb{P}(X = k) = \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{s}{2}\right)^k = \frac{1}{2-s},$$

$-1 < s \leq 1$.

(c) The probability we are looking for is

$$\mathbb{P}(X_3 = 0 \mid X_0 = 0) = G_X(G_X(G_X(0))) = \frac{1}{2 - \frac{1}{2-1/2}} = \frac{3}{4}.$$

(d) Since giving birth to a girl is equivalent to having at least one child, and this happens to each couple with probability $1/4$, the probability we are looking for is equal to

$$\frac{1}{4} + \frac{3}{4} \times \frac{1}{4} + \left(\frac{3}{4}\right)^2 \times \frac{1}{4} = \frac{1}{4} \times \frac{1-(3/4)^3}{1-3/4} = 1 - (3/4)^3 = \frac{37}{64} = 0.578125.$$

It can also be recovered from

$$G_Z^{(3)}(s) = G_Z(G_Z(G_Z(s))) = \frac{37}{64} + \frac{27s}{64}$$

at $s = 0$, where G_Z is the probability generating function $G_Z(s) = 1/4 + 3s/4$.

Problem 8.11

(a) We have

$$\mathbb{E}[Z_n] = \sum_{k=1}^{n} \mathbb{E}[X_k] = \sum_{k=1}^{n} \mu^k = \mu \sum_{k=0}^{n-1} \mu^k = \mu \frac{1-\mu^n}{1-\mu}, \qquad n \in \mathbb{N}.$$

(b) We have

$$\mathbb{E}[Z] = \mathbb{E}\left[\sum_{k=1}^{\infty} X_k\right] = \sum_{k=1}^{\infty} \mathbb{E}[X_k] = \mu \sum_{k=0}^{\infty} \mu^k = \frac{\mu}{1-\mu}, \qquad n \in \mathbb{N},$$

provided that $\mu < 1$.

(c) We have

$$H(s) = \mathbb{E}\big[s^Z \mid X_0 = 1\big] = \sum_{k=0}^{\infty} \big(s\mathbb{E}\big[s^Z \mid X_0 = 1\big]\big)^k \mathbb{P}(Y_1 = k) = G_1(sH(s)).$$

(d) We have

$$H(s) = G_1(sH(s)) = \frac{1-p}{1-psH(s)},$$

hence

$$psH^2(s) - H(s) + q = 0,$$

and

$$H(s) = \frac{1 \pm \sqrt{1-4pqs}}{2ps} = \frac{1-\sqrt{1-4pqs}}{2ps},$$

where we have chosen the minus sign since the plus sign leads to $H(0) = +\infty$ whereas we should have $H(0) = \mathbb{P}(Z=0) \leq 1$. In addition we have $\mu = p/q < 1$ hence $p < 1/2 < q$ and the minus sign gives

$$H(1) = \frac{1-\sqrt{1-4pq}}{2p} = \frac{1-|q-p|}{2p} = 1.$$

(e) We have

$$\lim_{s\searrow 0^+} H(s) = \lim_{s\searrow 0^+} \frac{1-(1-2pqs)}{2ps} = q = \mathbb{P}(Z=0) = \mathbb{P}(Y_1=0) = H(0).$$

(f) We have

$$H'(s) = \frac{pq}{ps\sqrt{1-4pqs}} - \frac{1-\sqrt{1-4pqs}}{2ps^2},$$

and

$$\begin{aligned} H'(1) &= \frac{pq}{p\sqrt{1-4pq}} - \frac{1-\sqrt{1-4pq}}{2p} = \frac{pq}{p(q-p)} - \frac{1-(q-p)}{2p} \\ &= \frac{q}{q-p} - 1 = \frac{p}{q-p} = \frac{\mu}{1-\mu}, \end{aligned}$$

with $\mu = p/q$ for $p < 1/2$, which shows that

$$\mathbb{E}[Z] = \frac{\mu}{1-\mu}$$

and recovers the result of Question (b).

(g) We have

$$\mathbb{E}\left[\sum_{k=1}^{Z} U_k\right] = \sum_{n=0}^{\infty} \mathbb{E}\left[\sum_{k=1}^{Z} U_k \mid Z=n\right] \mathbb{P}(Z=n) \mathbb{E}[U_1]\mathbb{E}[Z] = \mathbb{E}[U_1]\frac{\mu}{1-\mu}.$$

(h) We have

$$\begin{aligned}\mathbb{P}(U_k < x,\ k = 1, 2, \ldots, Z) &= \sum_{n=0}^{\infty} \mathbb{P}(U_k < x,\ k = 1, 2, \ldots, n)\mathbb{P}(Z = n) \\ &= \sum_{n=0}^{\infty} (F(x))^n \mathbb{P}(Z = n) = H(F(x)).\end{aligned}$$

(i) We have

$$\mathbb{E}\left[\sum_{k=1}^{Z} U_k\right] = \mathbb{E}[U_1]\frac{\mu}{1-\mu} = \frac{\mu}{1-\mu} = \frac{p}{q-p}.$$

We find

$$\begin{aligned}\mathbb{P}(U_k < x,\ k = 1, 2, \ldots, Z) &= H(F(x)) = H(1 - e^{-x}) \\ &= \frac{1 - \sqrt{1 - 4pq(1 - e^{-x})}}{2p(1 - e^{-x})}.\end{aligned}$$

Chapter 9 - Continuous-Time Markov Chains

Exercise 9.1 We model the number of operating machines as a birth and death process $(X_t)_{t\in\mathbb{R}_+}$ on the state space $\{0, 1, 2, 3, 4, 5\}$. A new machine can only be added at the rate λ since the repairman can fix only one machine at a time. In order to determine the failure rate starting from state $k \in \{0, 1, 2, 3, 4, 5\}$, let us assume that the number of working machines at time t is $X_t = k$. It is known that the lifetime τ_i of machine $i \in \{0, \ldots, k\}$ is an exponentially distributed random variable with parameter $\mu > 0$. On the other hand, we know that the first machine to fail will do so at time $\min(\tau_1, \tau_2, \ldots, \tau_k)$, and we have

$$\begin{aligned}\mathbb{P}(\min(\tau_1, \tau_2, \ldots, \tau_k) > t) &= \mathbb{P}(\tau_1 > t, \tau_2 > t, \ldots, \tau_k > t) \\ &= \mathbb{P}(\tau_1 > t)\mathbb{P}(\tau_2 > t)\cdots\mathbb{P}(\tau_k > t) = (\mathrm{e}^{-\mu t})^k = \mathrm{e}^{-k\mu t},\end{aligned}$$

$t \in \mathbb{R}_+$, hence the time until the first machine failure is exponentially distributed with parameter $k\mu$, i.e. the birth rate μ_k of $(X_t)_{t\in\mathbb{R}_+}$ is $\mu_k = k\mu$, $k = 1, 2, 3, 4, 5$.

Consequently, the infinitesimal generator Q of $(X_t)_{t\in\mathbb{R}_+}$ is given by

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & 0 & 0 \\ \mu & -\mu-\lambda & \lambda & 0 & 0 & 0 \\ 0 & 2\mu & -2\mu-\lambda & \lambda & 0 & 0 \\ 0 & 0 & 3\mu & -3\mu-\lambda & \lambda & 0 \\ 0 & 0 & 0 & 4\mu & -4\mu-\lambda & \lambda \\ 0 & 0 & 0 & 0 & 5\mu & -5\mu \end{bmatrix},$$

with $\lambda = 0.5$ and $\mu = 0.2$. We look for a stationary distribution of the form

$$\pi = (\pi_0, \pi_1, \pi_2, \pi_3, \pi_4, \pi_5)$$

by solving $\pi Q = 0$, which yields

$$\pi_1 = \frac{\lambda}{\mu}\pi_0, \quad \pi_2 = \frac{\lambda}{2\mu}\pi_1, \quad \pi_3 = \frac{\lambda}{3\mu}\pi_2, \quad \pi_4 = \frac{\lambda}{4\mu}\pi_3, \quad \pi_5 = \frac{\lambda}{5\mu}\pi_4,$$

i.e.

$$\pi_1 = \frac{\lambda}{\mu}\pi_0, \quad \pi_2 = \frac{\lambda^2}{2\mu^2}\pi_0, \quad \pi_3 = \frac{\lambda^3}{3!\mu^3}\pi_0, \quad \pi_4 = \frac{\lambda^4}{4!\mu^4}\pi_0, \quad \pi_5 = \frac{\lambda^5}{5!\mu^5}\pi_0,$$

which is a truncated Poisson distribution with

$$\pi_0 + \frac{\lambda}{\mu}\pi_0 + \frac{\lambda^2}{2\mu^2}\pi_0 + \frac{\lambda^3}{3!\mu^3}\pi_0 + \frac{\lambda^4}{4!\mu^4}\pi_0 + \frac{\lambda^5}{5!\mu^5}\pi_0 = 1,$$

hence

$$\begin{aligned}\pi_0 &= \frac{1}{1 + \dfrac{\lambda}{\mu} + \dfrac{\lambda^2}{2\mu^2} + \dfrac{\lambda^3}{3!\mu^3} + \dfrac{\lambda^4}{4!\mu^4} + \dfrac{\lambda^5}{5!\mu^5}} \\ &= \frac{\mu^5}{\mu^5 + \lambda\mu^4 + \lambda^2\mu^3/2 + \lambda^3\mu^2/3! + \lambda^4\mu/4! + \lambda^5/5!}.\end{aligned}$$

Finally, since π_5 is the probability that all 5 machines are operating, the fraction of time the repairman is idle in the long run is

$$\pi_5 = \frac{\lambda^5}{120\mu^5 + 120\lambda\mu^4 + 60\lambda^2\mu^3 + 20\lambda^3\mu^2 + 5\lambda^4\mu + \lambda^5}.$$

Note that of at most two machines can be under repair, the infinitesimal generator Q of $(X_t)_{t\in\mathbb{R}_+}$ will become

$$Q = \begin{bmatrix} -2\lambda & 2\lambda & 0 & 0 & 0 & 0 \\ \mu & -\mu - 2\lambda & 2\lambda & 0 & 0 & 0 \\ 0 & 2\mu & -2\mu - 2\lambda & 2\lambda & 0 & 0 \\ 0 & 0 & 3\mu & -3\mu - 2\lambda & 2\lambda & 0 \\ 0 & 0 & 0 & 4\mu & -4\mu - \lambda & \lambda \\ 0 & 0 & 0 & 0 & 5\mu & -5\mu \end{bmatrix}.$$

Exercise 9.2

(a) Since the time τ_k^R spent between two Poisson arrivals n^o k and $k+1$ is an exponentially distributed random variable with parameter λ_R, the probability we are looking for is given by

$$\mathbb{P}(\tau_k^R > t) = e^{-\lambda_R t},$$

where N_t^R denotes a Poisson process with intensity λ_R.

(b) This probability is given by

$$\begin{aligned}\mathbb{P}(N_t^W \leq 3) &= \mathbb{P}(N_t^W = 0) + \mathbb{P}(N_t^W = 1) + \mathbb{P}(N_t^W = 2) + \mathbb{P}(N_t^W = 3)\\ &= e^{-\lambda_W t} + \lambda_W t e^{-\lambda_W t} + e^{-\lambda_W t}\frac{\lambda_W^2 t^2}{2} + e^{-\lambda_W t}\frac{\lambda_W^3 t^3}{6},\end{aligned}$$

where N_t^W denotes a Poisson process with intensity λ_W.

(c) This probability is given by the ratio $\mathbb{P}(\tau^R < \tau^W) = \lambda_R/(\lambda_W + \lambda_R)$ of arrival rates, as follows from the probability computation (1.5.9), where τ^R and τ^W are independent exponential random variables with parameters λ_R and λ_W, representing the time until the next "read", resp. "write" consultation.

(d) This distribution is given by $\mathbb{P}(N_t^R = k \mid N_t^R + Y = n)$ where N_t^R, N_t^W are independent Poisson random variables with parameters $\lambda_R t$ and $\lambda_W t$ respectively. We have

$$\begin{aligned}\mathbb{P}(N_t^R = k \mid N_t^R + N_t^W = n) &= \frac{\mathbb{P}(N_t^R = k \text{ and } N_t^R + N_t^W = n)}{\mathbb{P}(N_t^R + N_t^W = n)}\\ &= \binom{n}{k}\left(\frac{\lambda_R}{\lambda_R + \lambda_W}\right)^k \left(\frac{\lambda_W}{\lambda_R + \lambda_W}\right)^{n-k}, \quad k = 0, 1, \ldots, n,\end{aligned}$$

cf. (B.8) in the Exercise 1.6.

Exercise 9.3

(a) The number X_t of machines operating at time t is a birth and death process on $\{0, 1, 2\}$ with infinitesimal generator

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 \\ \mu & -(\lambda + \mu) & \lambda \\ 0 & 2\mu & -2\mu \end{bmatrix}.$$

The stationary distribution $\pi = (\pi_0, \pi_1, \pi_2)$ is solution of $\pi Q = 0$ under the condition $\pi_0 + \pi_1 + \pi_2 = 1$, which yields

$$(\pi_0, \pi_1, \pi_2) = \left(\frac{2\mu^2}{2\mu^2 + 2\lambda\mu + \lambda^2}, \frac{2\lambda\mu}{2\mu^2 + 2\lambda\mu + \lambda^2}, \frac{\lambda^2}{2\mu^2 + 2\lambda\mu + \lambda^2}\right),$$

i.e. the probability that no machine is operating is $\pi_0 = 2/5$ when $\lambda = \mu = 1$.

(b) The number X_t of machines operating at time t is now a birth and death process on $\{0, 1\}$. The time spent in state 0 is exponentially distributed with average $1/\lambda$. When the chain is in state 1, one machine is working while the other one may still be under repair, and the mean time $\mathbb{E}[T_0 \mid X_0 = 1]$ spent in state 1 before switching to state 0 has to be computed using first step analysis on the discrete-time embedded chain. We have

$$\begin{aligned}
&\mathbb{E}[T_0 \mid X_0 = 1] \\
&= \mathbb{P}(X_\lambda < X_\mu) \times \left(\frac{1}{\mu} + \mathbb{E}[T_0 \mid X_0 = 1] \right) + \mathbb{P}(X_\mu < X_\lambda) \times \frac{1}{\mu} \\
&= \frac{1}{\mu} + \mathbb{P}(X_\lambda < X_\mu) \times \mathbb{E}[T_0 \mid X_0 = 1] \\
&= \frac{1}{\mu} + \frac{\lambda}{\lambda + \mu} \mathbb{E}[T_0 \mid X_0 = 1],
\end{aligned}$$

where, by (9.7.2) and (9.7.3) or (1.5.9), $\mathbb{P}(X_\lambda < X_\mu) = \lambda/(\lambda + \mu)$ is the probability that an exponential random variable X_λ with parameter $\lambda > 0$ is smaller than another independent exponential random variable X_μ with parameter $\mu > 0$. In other words, $\mathbb{P}(X_\lambda < X_\mu)$ is the probability that the repair of the idle machine finishes before the working machine fails. This yields

$$\mathbb{E}[T_0 \mid X_0 = 1] = \frac{\lambda + \mu}{\mu^2},$$

hence the corresponding rate is $\mu^2/(\lambda + \mu)$ and the infinitesimal generator of the chain becomes

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \dfrac{1}{\mathbb{E}[T_0 \mid X_0 = 1]} & -\dfrac{1}{\mathbb{E}[T_0 \mid X_0 = 1]} \end{bmatrix} = \begin{bmatrix} -\lambda & \lambda \\ \dfrac{\mu^2}{\lambda + \mu} & -\dfrac{\mu^2}{\lambda + \mu} \end{bmatrix}.$$

The stationary distribution $\pi = (\pi_0, \pi_1)$ is solution of $\pi Q = 0$, i.e.

$$\begin{cases} 0 = -\lambda \pi_0 + \pi_1 \dfrac{\mu^2}{\lambda + \mu} \\ \\ 0 = \lambda \pi_0 - \pi_1 \dfrac{\mu^2}{\lambda + \mu} \end{cases}$$

under the condition $\pi_0 + \pi_1 = 1$, which yields

$$(\pi_0, \pi_1) = \left(\frac{\mu^2}{\mu^2 + \lambda\mu + \lambda^2}, \frac{\lambda\mu + \lambda^2}{\mu^2 + \lambda\mu + \lambda^2} \right),$$

i.e. the probability that no machine is operating when $\lambda = \mu = 1$ is $\pi_0 = 1/3$.

Exercise 9.6 The size of the crack is viewed as a continuous-time birth process taking values in $\{1, 2, 3, \ldots\}$ with state-dependent rate $\lambda_k = (1 + k)^\rho$, $k \geq 1$. Let us denote by τ_k the time spent at state $k \in \mathbb{N}$ between two increases, which is an exponentially distributed random variable with parameter λ_k. The time it takes for the crack length to grow to infinity is $\sum_{k=1}^{\infty} \tau_k$. It is known that $\sum_{k=1}^{\infty} \tau_k < \infty$ almost surely if the expectation $\mathbb{E}\left[\sum_{k=1}^{\infty} \tau_k\right]$ is finite, and in this situation the crack grows to infinity within a finite time. We have

$$\mathbb{E}\left[\sum_{k=1}^{\infty} \tau_k\right] = \sum_{k=1}^{\infty} \mathbb{E}[\tau_k] = \sum_{k=1}^{\infty} \frac{1}{\lambda_k} = \sum_{k=1}^{\infty} \frac{1}{(1+k)^\rho}.$$

By comparison with the integral of the function $x \longmapsto 1/(1+x)^\rho$ we get

$$\mathbb{E}\left[\sum_{k=1}^{\infty} \tau_k\right] = \sum_{k=1}^{\infty} \frac{1}{(1+k)^\rho} \le \int_0^{\infty} \frac{1}{(1+x)^\rho} dx = \frac{1}{\rho - 1} < \infty,$$

provided that $\rho > 1$. We conclude that the time for the crack to grow to infinite length is (almost surely) finite when $\rho > 1$. Similarly, we have

$$\mathbb{E}\left[\sum_{k=1}^{\infty} \tau_k\right] = \sum_{k=1}^{\infty} \frac{1}{(1+k)^\rho} \ge \int_1^{\infty} \frac{1}{(1+x)^\rho} dx = \infty,$$

hence the mean time for the crack to grow to infinite length is infinite when $\rho \le 1$.

Exercise 9.7

(a) This time is the expected value of the third jump time T_3, i.e.

$$\mathbb{E}[T_3] = \mathbb{E}[\tau_0] + \mathbb{E}[\tau_1] + \mathbb{E}[\tau_2] = \frac{3}{\lambda} = 30 \text{ minutes}.$$

(b) This probability is

$$\begin{aligned}\mathbb{P}(N_{60} < 3) &= \mathbb{P}(N_{60} = 0) + \mathbb{P}(N_{60} = 1) + \mathbb{P}(N_{60} = 2)\\ &= \mathrm{e}^{-60\lambda}(1 + 60\lambda + (60\lambda)^2/2)\\ &= 25\mathrm{e}^{-6} \simeq 0.062.\end{aligned}$$

Exercise 9.8

(a) By the independence of increments of the Poisson process $(N_t)_{t\in\mathbb{R}_+}$ we find

$$\begin{aligned}\mathbb{P}(N_{t_3} = 5 \mid N_{t_1} = 1) &= \frac{\mathbb{P}(N_{t_3} = 5 \text{ and } N_{t_1} = 1)}{\mathbb{P}(N_{t_1} = 1)}\\ &= \frac{\mathbb{P}(N_{t_3} - N_{t_1} = 4 \text{ and } N_{t_1} = 1)}{\mathbb{P}(N_{t_1} = 1)} = \frac{\mathbb{P}(N_{t_3} - N_{t_1} = 4)\mathbb{P}(N_{t_1} = 1)}{\mathbb{P}(N_{t_1} = 1)}\end{aligned}$$

$$= \mathbb{P}(N_{t_3} - N_{t_1} = 4) = \frac{(\lambda(t_3 - t_1))^4}{4!} e^{-\lambda(t_3 - t_1)}.$$

(b) We expand N_{t_4} into the telescoping sum

$$N_{t_4} = (N_{t_4} - N_{t_3}) + (N_{t_3} - N_{t_2}) + (N_{t_2} - N_{t_1}) + (N_{t_1} - N_0)$$

of independent increments on disjoint time intervals, to obtain

$$\begin{aligned}
&\mathbb{E}[N_{t_1} N_{t_4}(N_{t_3} - N_{t_2})] \\
&= \lambda^3 t_1 (t_4 - t_3)(t_3 - t_2) + \lambda^2 t_1 (t_3 - t_2)(1 + \lambda(t_3 - t_2)) \\
&\quad + \lambda^3 t_1 (t_2 - t_1)(t_3 - t_2) + \lambda^2 t_1 (1 + \lambda t_1)(t_3 - t_2).
\end{aligned}$$

(c) We have $\{T_2 > t\} = \{N_t \leq 1\}$, $t \in \mathbb{R}_+$, hence

$$\mathbb{E}[N_{t_2} \mid T_2 > t_1] = \mathbb{E}[N_{t_2} \mid N_{t_1} \leq 1] = \frac{1}{\mathbb{P}(N_{t_1} \leq 1)} \mathbb{E}[N_{t_2} \mathbb{1}_{\{N_{t_1} \leq 1\}}],$$

by (1.6.6). Now, using the independence of increments between $N_{t_2} - N_{t_1}$ and N_{t_1}, we have

$$\begin{aligned}
\mathbb{E}[N_{t_2} \mathbb{1}_{\{N_{t_1} \leq 1\}}] &= \mathbb{E}[(N_{t_2} - N_{t_1}) \mathbb{1}_{\{N_{t_1} \leq 1\}}] + \mathbb{E}[N_{t_1} \mathbb{1}_{\{N_{t_1} \leq 1\}}] \\
&= \mathbb{E}[N_{t_2} - N_{t_1}] \mathbb{P}(N_{t_1} \leq 1) + \mathbb{E}[N_{t_1} \mathbb{1}_{\{N_{t_1} \leq 1\}}],
\end{aligned}$$

hence

$$\begin{aligned}
&\mathbb{E}[N_{t_2} \mid T_2 > t_1] = \mathbb{E}[N_{t_2} \mid N_{t_1} \leq 1] = \frac{1}{\mathbb{P}(N_{t_1} \leq 1)} \mathbb{E}[N_{t_2} \mathbb{1}_{\{N_{t_1} \leq 1\}}] \\
&= \mathbb{E}[N_{t_2} - N_{t_1}] + \frac{\mathbb{P}(N_{t_1} = 1)}{\mathbb{P}(N_{t_1} \leq 1)} = \lambda(t_2 - t_1) + \frac{\lambda t_1 e^{-\lambda t_1}}{e^{-\lambda t_1} + \lambda t_1 e^{-\lambda t_1}}.
\end{aligned}$$

Exercise 9.10 This is an extension of Exercise 9.9. The generator of the process is given by

$$Q = \begin{bmatrix}
-\alpha N & \alpha N & 0 & \cdots & 0 & 0 & 0 \\
\beta & -\alpha(N-1) - \beta & \alpha(N-1) & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \beta(N-1) & -\alpha - \beta(N-1) & \alpha \\
0 & 0 & 0 & \cdots & 0 & \beta N & -\beta N
\end{bmatrix}.$$

Writing the equation $\pi Q = 0$ shows that we have $-\alpha N \pi_0 + \beta \pi_1 = 0$,

$$\beta(k+1)\pi_{k+1} - (\alpha(N-k) + \beta k)\pi_k + \alpha(N-(k-1))\pi_{k-1} = 0, \quad k = 1, 2, \ldots, N-1,$$

and $\alpha \pi_{N-1} - \beta N \pi_N = 0$, from which we deduce the recurrence relation

$$\pi_{k+1} = \frac{\alpha}{\beta}\frac{N-k}{k+1}\pi_k, \qquad k = 0, 1, \ldots, N-1,$$

and by induction on $k = 1, 2, \ldots, N$ we find

$$\begin{aligned}
\pi_k &= \left(\frac{\alpha}{\beta}\right)^k \frac{N(N-1)\cdots(N-k+1)}{k!}\pi_0 \\
&= \left(\frac{\alpha}{\beta}\right)^k \frac{N!}{(N-k)!k!}\pi_0 = \left(\frac{\alpha}{\beta}\right)^k \binom{N}{k}\pi_0,
\end{aligned}$$

$k = 0, 1, \ldots, N$. The condition $\pi_0 + \pi_1 + \cdots + \pi_N = 1$ shows that $\pi_0 = (1 + \alpha/\beta)^{-N}$ and we have

$$\pi_k = \left(\frac{\alpha}{\alpha+\beta}\right)^k \left(\frac{\beta}{\alpha+\beta}\right)^{N-k} \frac{N!}{(N-k)!k!}, \qquad k = 0, 1, \ldots, N,$$

hence the stationary distribution π is a binomial distribution with parameter $(N, p) = (N, \alpha/(\alpha+\beta))$.

Exercise 9.11 The generator Q of this pure birth process is given by

$$Q = [\,\lambda_{i,j}\,]_{i,j\in\mathbb{N}} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & \cdots \\ 0 & -3 & 3 & 0 & 0 & \cdots \\ 0 & 0 & -2 & 2 & 0 & \cdots \\ 0 & 0 & 0 & -5 & 5 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix},$$

hence the forward Kolmogorov equation $P'(t) = P(t)Q$ yields

$$\begin{cases}
P'_{0,0}(t) = -P_{0,0}(t), \\[2ex]
P'_{0,1}(t) = P_{0,0}(t) - 3P_{0,1}(t), \\[2ex]
P'_{0,2}(t) = 3P_{0,1}(t) - 2P_{0,2}(t), \\[2ex]
P'_{0,3}(t) = 2P_{0,2}(t) - 5P_{0,3}(t).
\end{cases}$$

The first equation is solved by (A.9) as

$$P_{0,0}(t) = P_{0,0}(0)\mathrm{e}^{-t} = \mathrm{e}^{-t}, \qquad t \in \mathbb{R}_+,$$

and this solution can be easily recovered from

$$P_{0,0}(t) = \mathbb{P}(X_t = 0 \mid X_0 = 0) = \mathbb{P}(\tau_0 > t) = \mathrm{e}^{-t}, \qquad t \in \mathbb{R}_+.$$

The second equation becomes

$$P'_{0,1}(t) = \mathrm{e}^{-t} - 3P_{0,1}(t), \tag{B.29}$$

and has the solution

$$P_{0,1}(t) = \frac{1}{2}(\mathrm{e}^{-t} - \mathrm{e}^{-3t}), \qquad t \in \mathbb{R}_+.$$

The remaining equations can be solved similarly by searching for a suitable particular solution. For

$$P'_{0,2}(t) = \frac{3}{2}(\mathrm{e}^{-t} - \mathrm{e}^{-3t}) - 2P_{0,2}(t),$$

we find

$$P_{0,2}(t) = \frac{3}{2}\mathrm{e}^{-3t}(1 - \mathrm{e}^{t})^2, \qquad t \in \mathbb{R}_+, \tag{B.30}$$

and for

$$P'_{0,3}(t) = 3\mathrm{e}^{-3t}(1 - \mathrm{e}^{t})^2 - 5P_{0,3}(t),$$

we find

$$P_{0,3}(t) = \frac{1}{4}\mathrm{e}^{-5t}(\mathrm{e}^{t} - 1)^3(1 + 3e^{t}), \qquad t \in \mathbb{R}_+. \tag{B.31}$$

Exercise 9.12 Noting that the two events

$$\{T_1 > t,\ T_2 > t + s\} = \{X_t = 0,\ 0 \leq X_{t+s} \leq 1\}$$

coincides for all $s, t \in \mathbb{R}_+$, we find that

$$\begin{aligned}
\mathbb{P}(T_1 > t \text{ and } T_2 > t + s) &= \mathbb{P}(X_t = 0 \text{ and } X_{t+s} \in \{0, 1\} \mid X_0 = 0) \\
&= \mathbb{P}(X_{t+s} \in \{0, 1\} \mid X_t = 0)\mathbb{P}(X_t = 0 \mid X_0 = 0) \\
&= \mathbb{P}(X_s \in \{0, 1\} \mid X_0 = 0)\mathbb{P}(X_t = 0 \mid X_0 = 0) \\
&= (\mathbb{P}(X_s = 0 \mid X_0 = 0) + \mathbb{P}(X_s = 1 \mid X_0 = 0))\mathbb{P}(X_t = 0 \mid X_0 = 0) \\
&= P_{0,0}(t)(P_{0,0}(s) + P_{0,1}(s)).
\end{aligned}$$

Next, we note that we have

$$P_{0,0}(t) = \mathrm{e}^{-\lambda_0 t}, \qquad t \in \mathbb{R}_+,$$

and

$$P_{0,1}(t) = \frac{\lambda_0}{\lambda_1 - \lambda_0}\left(\mathrm{e}^{-\lambda_0 t} - \mathrm{e}^{-\lambda_1 t}\right), \qquad t \in \mathbb{R}_+,$$

hence

$$\begin{aligned}\mathbb{P}(T_1 > t \text{ and } T_2 > t+s) &= \mathrm{e}^{-\lambda_0(t+s)} + \frac{\lambda_0}{\lambda_1-\lambda_0}\left(\mathrm{e}^{-\lambda_0(t+s)} - \mathrm{e}^{-\lambda_0 t-\lambda_1 s}\right)\\ &= \frac{\lambda_1}{\lambda_1-\lambda_0}\mathrm{e}^{-\lambda_0(t+s)} - \frac{\lambda_0}{\lambda_1-\lambda_0}\mathrm{e}^{-\lambda_0 t-\lambda_1 s}.\end{aligned}$$

Then, since

$$\mathbb{P}(T_1 > x \text{ and } T_2 > y) = \int_x^\infty \int_y^\infty f_{(T_1,T_2)}(u,v)dudv,$$

by (1.5.4) we get

$$\begin{aligned}f_{(T_1,T_2)}(x,y) &= \frac{\partial^2}{\partial y\partial x}\mathbb{P}(T_1 > x \text{ and } T_2 > y)\\ &= \frac{\partial^2}{\partial y\partial x}\left(\frac{\lambda_1}{\lambda_1-\lambda_0}\mathrm{e}^{-\lambda_0 y} - \frac{\lambda_0}{\lambda_1-\lambda_0}\mathrm{e}^{-\lambda_0 x-\lambda_1(y-x)}\right)\\ &= -\lambda_0\frac{\partial}{\partial y}\mathrm{e}^{-\lambda_0 x-\lambda_1(y-x)} = \lambda_0\lambda_1\mathrm{e}^{-\lambda_0 x-\lambda_1(y-x)},\end{aligned}$$

provided that $y \geq x \geq 0$. When $x > y \geq 0$ we have

$$f_{(T_1,T_2)}(x,y) = 0.$$

The density of (τ_0, τ_1) is given under the change of variable $T_0 = \tau_0$, $T_1 = \tau_0 + \tau_1$, by

$$f_{(\tau_0,\tau_1)}(s,t) = f_{(T_1,T_2)}(s, s+t) = \lambda_0\lambda_1\mathrm{e}^{-\lambda_0 s-\lambda_1 t}, \quad s,t \in \mathbb{R}_+,$$

which shows that τ_0, τ_1 are two independent exponentially distributed random variables with parameters λ_0 and λ_1, respectively.

Exercise 9.13 Let $(N_t)_{t\in\mathbb{R}_+}$ denote a Poisson process with intensity $\lambda > 0$.

(a) This probability is equal to

$$\mathbb{P}(N_T = 0) = \mathbb{P}(\tau_0 > T) = \mathrm{e}^{-\lambda T}.$$

(b) Let t denote the expected time we are looking for. When the woman attempts to cross the street, she can do so immediately with probability $\mathbb{P}(N_T = 0) = \mathbb{P}(\tau_0 > T)$, in which case the waiting time is 0. Otherwise, with probability $1 - \mathbb{P}(N_T = 0)$, she has to wait on average (using Lemma 1.4 and (1.6.14))

$$\begin{aligned}\mathbb{E}[\tau_0 \mid \tau_0 < T] &= \frac{1}{\mathbb{P}(\tau_0 < T)}\mathbb{E}[\tau_0 \mathbb{1}_{\{\tau_0 < T\}}] \\ &= \frac{\lambda}{1-\mathrm{e}^{-\lambda T}}\int_0^T x\mathrm{e}^{-\lambda x}dx \\ &= \frac{1-(1+\lambda T)\mathrm{e}^{-\lambda T}}{\lambda(1-\mathrm{e}^{-\lambda T})}\end{aligned}$$

for the first car to pass, after which the process is reinitialized and the average waiting time is again t. Hence by first step analysis in continuous time we find the equation

$$\begin{aligned}t &= 0 \times \mathbb{P}(N_T = 0) + (\mathbb{E}[\tau_0 \mid \tau_0 < T] + t) \times \mathbb{P}(\tau_0 \le T) \\ &= \frac{1-(1+\lambda T)\mathrm{e}^{-\lambda T}}{\lambda} + t(1-\mathrm{e}^{-\lambda T})\end{aligned} \tag{B.32}$$

with unknown t, and solution $t = (\mathrm{e}^{\lambda T} - 1 - \lambda T)/\lambda$.

(c) Denoting by t the mean time until she finishes crossing the street we have, by first step analysis in continuous time,

$$\begin{aligned}t &= \mathbb{E}\left[T\mathbb{1}_{\{T<\tau_0\}} + (\tau_0 + t)\mathbb{1}_{\{T \ge \tau_0\}}\right] \\ &= \lambda T\int_T^\infty \mathrm{e}^{-\lambda s}ds + \lambda\int_0^T (s+t)\mathrm{e}^{-\lambda s}ds \\ &= T\mathbb{P}(\tau_0 > T) + \lambda\int_0^T s\mathrm{e}^{-\lambda s}ds + t\mathbb{P}(\tau_0 < T) \\ &= T\mathrm{e}^{-\lambda T} + \frac{1-\mathrm{e}^{-\lambda T}(1+\lambda T)}{\lambda} + t(1-\mathrm{e}^{-\lambda T}),\end{aligned}$$

which yields $t = (\mathrm{e}^{\lambda T} - 1)/\lambda$.

(d) In this case, T becomes an independent exponentially distributed random variable with parameter $\mu > 0$, hence we can write

$$t = \mathbb{E}\left[\frac{e^{\lambda T}-1}{\lambda}\right] = \frac{1}{\mu - \lambda}$$

if $\mu > \lambda$, with $t = +\infty$ if $\mu \le \lambda$.

Exercise 9.15

(a) The generator Q of $(X_t)_{t\in\mathbb{R}_+}$ is given by

$$Q = \begin{bmatrix} -0.5 & 0.5 & 0 \\ 0.2 & -0.45 & 0.25 \\ 0 & 0.4 & -0.4 \end{bmatrix}.$$

(b) Solving for $\pi Q = 0$ we have

$$\pi Q = [\pi_0, \pi_1, \pi_2] \times \begin{bmatrix} -0.5 & 0.5 & 0 \\ 0.2 & -0.45 & 0.25 \\ 0 & 0.4 & -0.4 \end{bmatrix}$$
$$= \begin{bmatrix} -0.5 \times \pi_0 + 0.2 \times \pi_1 \\ 0.5 \times \pi_0 - 0.45 \times \pi_1 + 0.4 \times \pi_2 \\ 0.25 \times \pi_1 - 0.4 \times \pi_2 \end{bmatrix}^T = [0, 0, 0],$$

i.e. $\pi_0 = 0.4 \times \pi_1 = 0.64 \times \pi_2$ under the condition $\pi_0 + \pi_1 + \pi_2 = 1$, which gives $\pi_0 = 16/81$, $\pi_1 = 40/81$, $\pi_2 = 25/81$.

(c) In the long run the average is $0 \times \pi_0 + 1 \times \pi_1 + 2 \times \pi_2 = 40/81 + 50/81 = 90/81$.

(d) We find $100 \times 90/81 = 1000/9$.

(e) We have
$$Q = \begin{bmatrix} -0.5 & 0.5 & 0 \\ 0.2 & -0.7 & 0.5 \\ 0 & 0.4 & -0.4 \end{bmatrix},$$

and solving $\pi Q = 0$ shows that

$$[\pi_0, \pi_1, \pi_2] = \left[\frac{0.32}{2.12}, \frac{0.8}{2.12}, \frac{1}{2.12}\right] = [0.15094, 0.37736, 0.47170]\,.$$

Exercise 9.16 Both chains $(X_1(t))_{t\in\mathbb{R}}$ and $(X_2(t))_{t\in\mathbb{R}}$ have the same infinitesimal generator
$$Q = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}.$$

The infinitesimal generator of $Z(t) := X_1(t) + X_2(t)$ is given by

$$\begin{bmatrix} -2\lambda & 2\lambda & 0 \\ \mu & -\lambda - \mu & \lambda \\ 0 & 2\mu & -2\mu \end{bmatrix},$$

as the birth rate λ is doubled when both chains are in state 0, and the death rate μ is also doubled when both chains are in state 1. Recall that by Proposition 9.6, the semi-groups of $X_1(t)$ and $X_2(t)$ are given by

$$\begin{bmatrix} \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0) & \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0) \\ \\ \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 1) & \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 1) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-t(\lambda+\mu)} & \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu} e^{-t(\lambda+\mu)} \\ \\ \frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu} e^{-t(\lambda+\mu)} & \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} e^{-t(\lambda+\mu)} \end{bmatrix}.$$

As for the transition semi-group of $Z(t)$, we have

$$P_{0,0}(t) = \mathbb{P}(Z(t) = 0 \mid Z(0) = 0) = \left(\frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-t(\lambda+\mu)} \right)^2 .$$

For $P_{0,1}(t)$ we have

$$\begin{aligned} P_{0,1}(t) &= \mathbb{P}(Z(t) = 1 \mid Z(0) = 0) \\ &= \mathbb{P}(X_1(t) = 0 \text{ and } X_2(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 0) \\ &\quad + \mathbb{P}(X_1(t) = 1 \text{ and } X_2(t) = 0 \mid X_1(0) = 0 \text{ and } X_2(0) = 0) \\ &= \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0)\mathbb{P}(X_2(t) = 0 \mid X_2(0) = 0) \\ &\quad + \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0)\mathbb{P}(X_2(t) = 1 \mid X_2(0) = 0) \\ &= 2\mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0)\mathbb{P}(X_2(t) = 0 \mid X_2(0) = 0). \end{aligned}$$

Starting from $Z(0) = 1$ and ending at $Z(t) = 1$ we have two possibilities $(0, 1)$ or $(1, 0)$ for the terminal condition. As for the initial condition $Z(0) = 1$ the two possibilities $(0, 1)$ and $(1, 0)$ count for one only since they both give $Z(0) = 1$. Thus, in order to compute $P_{1,1}(t)$ we can choose to assign the value 0 to $X_1(0)$ and the value 1 to $X_2(0)$ without influencing the final result, as the other choice would lead to the same probability value. Hence for $P_{1,1}(t)$ we have

$$\begin{aligned} P_{1,1}(t) &= \mathbb{P}(Z(t) = 1 \mid Z(0) = 1) \\ &= \mathbb{P}(X_1(t) = 0 \text{ and } X_2(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1) \\ &\quad + \mathbb{P}(X_1(t) = 1 \text{ and } X_2(t) = 0 \mid X_1(0) = 0 \text{ and } X_2(0) = 1) \\ &= \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0)\mathbb{P}(X_2(t) = 1 \mid X_2(0) = 1) \\ &\quad + \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0)\mathbb{P}(X_2(t) = 0 \mid X_2(0) = 1). \end{aligned}$$

Concerning $P_{1,0}(t)$ we have

$$\begin{aligned} P_{1,0}(t) &= \mathbb{P}(Z(t) = 0 \mid Z(0) = 1) \\ &= \mathbb{P}(X_1(t) = 0 \text{ and } X_2(t) = 0 \mid X_1(0) = 0 \text{ and } X_2(0) = 1). \end{aligned}$$

On the other hand, we have

$$\begin{aligned} P_{1,2}(t) &= \mathbb{P}(Z(t) = 2 \mid Z(0) = 1) \\ &= \mathbb{P}(X_1(t) = 1 \text{ and } X_2(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1) \end{aligned}$$

$$= \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0)\mathbb{P}(X_2(t) = 1 \mid X_2(0) = 1).$$

Exercise 9.17 Starting from state 0, the process $X_t = \xi_{N_t}$ stays at state 0 during an exponentially distributed Poisson interjump time with parameter λ, after which N_t increases by one unit. In this case, $\xi_{N_t} = 0$ becomes $\xi_{N_t+1} = 1$ with probability 1, from the transition matrix (9.8.4), hence the birth rate of X_t from state 0 to state 1 is λ. Then, starting from state 1, the process X_t stays at 1 during an exponentially distributed time with parameter λ. The difference is that when N_t increases by one unit, $\xi_{N_t} = 1$ may move to $\xi_{N_t+1} = 0$ with probability $1 - \alpha$, or remain at $\xi_{N_t+1} = 1$ with probability α. In fact, due to the Markov property, X_t will remain at 1 during an exponentially distributed time whose expectation may be higher than $1/\lambda$ when $\alpha > 0$. We will compute the expectation of this random time.

(a) We have

$$\begin{aligned}\mathbb{E}\big[T_0^r \mid X_0 = 1\big] &= \alpha\left(\frac{1}{\lambda} + \mathbb{E}\big[T_0^r \mid X_0 = 1\big]\right) + (1-\alpha) \times \left(\frac{1}{\lambda} + 0\right) \\ &= \frac{1}{\lambda} + \alpha\mathbb{E}\big[T_0^r \mid X_0 = 1\big],\end{aligned}$$

hence

$$\mathbb{E}\big[T_0^r \mid X_0 = 1\big] = \frac{1}{\lambda(1-\alpha)} \tag{B.33}$$

and

$$\mathbb{E}\big[T_0^r \mid X_0 = 0\big] = \frac{1}{\lambda} + 1 \times \mathbb{E}\big[T_0^r \mid X_0 = 1\big] = \frac{2-\alpha}{\lambda(1-\alpha)}.$$

Note that (B.33) can also be recovered from (5.3.3) by letting $b = 1 - \alpha$ and multiplying by the average Poisson interjump time $1/\lambda$.

(b) We have

$$\begin{aligned}\mathbb{E}\big[T_1^r \mid X_0 = 1\big] &= \frac{1}{\lambda} + \alpha\mathbb{E}\big[T_1^r \mid X_0 = 1\big] + (1-\alpha)\mathbb{E}\big[T_1^r \mid X_0 = 0\big] \\ &= \frac{2-\alpha}{\lambda} + \alpha\mathbb{E}\big[T_1^r \mid X_0 = 1\big],\end{aligned}$$

since $\mathbb{E}\big[T_1^r \mid X_0 = 0\big] = 1/\lambda$, hence

$$\mathbb{E}\big[T_1^r \mid X_0 = 1\big] = \frac{2-\alpha}{\lambda(1-\alpha)}.$$

(c) This continuous-time first step analysis argument is similar to the one used in the solution of Exercise 9.3. Since

$$\mathbb{E}\big[T_0^r \mid X_0 = 1\big] = \frac{1}{\lambda(1-\alpha)},$$

it takes an exponential random time with parameter $\lambda(1-\alpha)$ for the process $(X_t)_{t\in\mathbb{R}_+}$ to switch from state 1 to state 0. Hence the death rate is

$$\frac{1}{\mathbb{E}\big[T_0^r \mid X_0 = 1\big]} = \lambda(1-\alpha),$$

and the infinitesimal generator Q of X_t is

$$\begin{bmatrix} -\lambda & \lambda \\ \dfrac{1}{\mathbb{E}\big[T_0^r \mid X_0 = 1\big]} & -\dfrac{1}{\mathbb{E}\big[T_0^r \mid X_0 = 1\big]} \end{bmatrix} = \begin{bmatrix} -\lambda & \lambda \\ (1-\alpha)\lambda & -(1-\alpha)\lambda \end{bmatrix}.$$

Problem 9.18

(a) We need to show the following properties.

(i) The process $(N_t^1 + N_t^2)_{t\in\mathbb{R}_+}$ is a counting process. Clearly, the jump heights are positive integers and they can only be equal to one since the probability that N_t^1 and N_t^2 jumps simultaneously is 0.

(ii) The process $(N_t^1 + N_t^2)_{t\in\mathbb{R}_+}$ has independent increments.

Letting $0 < t_1 < t_2 < \cdots < t_n$, the family

$$\begin{aligned}&\big(N_{t_n}^1 + N_{t_n}^2 - (N_{t_{n-1}}^1 + N_{t_{n-1}}^2), \ldots, N_{t_2}^1 + N_{t_2}^2 - (N_{t_1}^1 + N_{t_1}^2)\big)\\ &\quad = \big(N_{t_n}^1 - N_{t_{n-1}}^1 + N_{t_n}^2 - N_{t_{n-1}}^2, \ldots, N_{t_2}^1 - N_{t_1}^1 + N_{t_2}^2 - N_{t_1}^2\big)\end{aligned}$$

is a family of independent random variables. In order to see this we note that $N_{t_n}^1 - N_{t_{n-1}}^1$ is independent of

$$N_{t_{n-1}}^1 - N_{t_{n-2}}^1, \ldots, N_{t_2}^1 - N_{t_1}^1,$$

and of

$$N_{t_n}^2 - N_{t_{n-1}}^2, \ldots, N_{t_2}^2 - N_{t_1}^2,$$

hence it is also independent of

$$N_{t_{n-1}}^1 - N_{t_{n-2}}^1 + N_{t_{n-1}}^2 - N_{t_{n-2}}^2, \ldots, N_{t_2}^1 - N_{t_1}^1 + N_{t_2}^2 - N_{t_1}^2.$$

Similarly it follows that $N_{t_n}^2 - N_{t_{n-1}}^2$ is independent of

$$N_{t_{n-1}}^1 - N_{t_{n-2}}^1 + N_{t_{n-1}}^2 - N_{t_{n-2}}^2, \ldots, N_{t_2}^1 - N_{t_1}^1 + N_{t_2}^2 - N_{t_1}^2,$$

hence $N_{t_n}^1 + N_{t_n}^2 - (N_{t_{n-1}}^1 + N_{t_{n-1}}^2)$ is independent of

$$N^1_{t_{n-1}} - N^1_{t_{n-2}} + N^2_{t_{n-1}} - N^2_{t_{n-2}}, \ldots, N^1_{t_2} - N^1_{t_1} + N^2_{t_2} - N^2_{t_1}.$$

This shows the required mutual independence by induction on $n \geq 1$.

(iii) The process $(N^1_t + N^2_t)_{t\in\mathbb{R}_+}$ has stationary increments. We note that the distributions of the random variables $N^1_{t+h} - N^1_{s+h}$ and $N^2_{t+h} - N^2_{s+h}$ do not depend on $h \in \mathbb{R}_+$, hence by the law of total probability we check that

$$\begin{aligned}
&\mathbb{P}(N^1_{t+h} + N^2_{t+h} - (N^1_{s+h} + N^2_{s+h}) = n)\\
&= \sum_{k=0}^{n} \mathbb{P}(N^1_{t+h} - N^1_{s+h} = k)\mathbb{P}(N^2_{t+h} - N^2_{s+h} = n-k)
\end{aligned}$$

is independent of $h \in \mathbb{R}_+$.

The intensity of $N^1_t + N^2_t$ is $\lambda_1 + \lambda_2$.

(b) (i) The proof of independence of increments is similar to that of Question (a).

(ii) Concerning the stationarity of increments we have

$$\begin{aligned}
\mathbb{P}(M_{t+h} - M_t = n) &= \mathbb{P}(N^1_{t+h} - N^2_{t+h} - (N^1_{s+h} - N^2_{s+h}) = n)\\
&= \mathbb{P}(N^1_{t+h} - N^1_{s+h} - (N^2_{t+h} - N^2_{s+h}) = n)\\
&= \sum_{k=0}^{\infty} \mathbb{P}(N^1_{t+h} - N^1_{s+h} = n+k)\mathbb{P}(N^2_{t+h} - N^2_{s+h} = k)
\end{aligned}$$

which is independent of $h \in \mathbb{R}_+$ since the distributions of $N^1_{t+h} - N^1_{s+h}$ and $N^2_{t+h} - N^2_{s+h}$ are independent of $h \in \mathbb{R}_+$.

(c) For $n \in \mathbb{N}$ we have

$$\begin{aligned}
\mathbb{P}(M_t = n) &= \mathbb{P}(N^1_t - N^2_t = n)\\
&= \sum_{k=\max(0,-n)}^{\infty} \mathbb{P}(N^1_t = n+k)\mathbb{P}(N^2_t = k)\\
&= \mathrm{e}^{-(\lambda_1+\lambda_2)t} \sum_{k=\max(0,-n)}^{\infty} \frac{\lambda_1^{n+k}\lambda_2^k t^{n+2k}}{k!(n+k)!}\\
&= \left(\frac{\lambda_1}{\lambda_2}\right)^{n/2} \mathrm{e}^{-(\lambda_1+\lambda_2)t} \sum_{k=\max(0,-n)}^{\infty} \frac{(t\sqrt{\lambda_1\lambda_2})^{n+2k}}{(n+k)!k!}\\
&= \left(\frac{\lambda_1}{\lambda_2}\right)^{n/2} \mathrm{e}^{-(\lambda_1+\lambda_2)t} I_{|n|}(2t\sqrt{\lambda_1\lambda_2}),
\end{aligned}$$

where

$$I_n(x) = \sum_{k=0}^{\infty} \frac{(x/2)^{n+2k}}{k!(n+k)!}, \qquad x > 0,$$

is the modified Bessel function with parameter $n \geq 0$. When $n \leq 0$, by exchanging λ_1 and λ_2 we get

$$\mathbb{P}(M_t = n) = \mathbb{P}(-M_t = -n) = \left(\frac{\lambda_1}{\lambda_2}\right)^{n/2} \mathrm{e}^{-(\lambda_1+\lambda_2)t} I_{-n}(2t\sqrt{\lambda_1\lambda_2}),$$

hence in the general case we have

$$\mathbb{P}(M_t = n) = \left(\frac{\lambda_1}{\lambda_2}\right)^{n/2} \mathrm{e}^{-(\lambda_1+\lambda_2)t} I_{|n|}(2t\sqrt{\lambda_1\lambda_2}), \qquad n \in \mathbb{Z},$$

which is known as the *Skellam distribution*.

(d) From the bound[1] $I_{|n|}(y) < C_n y^{|n|}\mathrm{e}^y$, $y > 1$, we get

$$\begin{aligned}\mathbb{P}(M_t = n) &\leq C_n \mathrm{e}^{-(\lambda_1+\lambda_2)t} \left(\frac{\lambda_1}{\lambda_2}\right)^{n/2} (2t\sqrt{\lambda_1\lambda_2})^{|n|} \mathrm{e}^{2t\sqrt{\lambda_1\lambda_2}} \\ &= C_n \mathrm{e}^{-t(\sqrt{\lambda_1}-\sqrt{\lambda_2})^2} \left(\frac{\lambda_1}{\lambda_2}\right)^{n/2} (2t\sqrt{\lambda_1\lambda_2})^{|n|},\end{aligned}$$

which tends to 0 as t goes to infinity when $\lambda_1 \neq \lambda_2$. Hence we have[2]

$$\lim_{t\to\infty} \mathbb{P}(|M_t| < c) = \sum_{-c<k<c} \lim_{t\to\infty} \mathbb{P}(|M_t| = k) = 0, \qquad c > 0. \tag{B.34}$$

(e) When $M_t \geq 0$, M_t represents the number of waiting customers. When $M_t \leq 0$, $-M_t$ represents the number of waiting drivers. Relation (B.34) shows that for any fixed $c > 0$, the probability of having either more than c waiting customers or more than c waiting drivers is high in the long run.

Problem 9.19

(a) We have

$$Q = \begin{bmatrix} -N\lambda & N\lambda & 0 & \cdots & \cdots & 0 & 0 & 0 \\ \mu & -\mu-(N-1)\lambda & (N-1)\lambda & \cdots & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & (N-1)\mu & -(N-1)\mu-\lambda & \lambda \\ 0 & 0 & 0 & \cdots & \cdots & 0 & N\mu & -N\mu \end{bmatrix}.$$

(b) The system of equations follows by writing the matrix multiplication $P'(t) = P(t)Q$ term by term.

(c) We apply the result of Question (b) to

[1] See e.g. Theorem 2.1 of [Laf91] for a proof of this inequality.

[2] Treating the case $\lambda_1 = \lambda_2$ is more complicated and not required.

$$\frac{\partial G_k}{\partial t}(s,t) = \sum_{n=0}^{N} s^n P'_{k,n}(t),$$

and use the expression

$$\frac{\partial G_k}{\partial s}(s,t) = \sum_{n=1}^{N} n s^{n-1} P'_{k,n}(t).$$

(d) We have

$$\begin{aligned}
&\lambda N(s-1)G_k(s,t) + (\mu + (\lambda-\mu)s - \lambda s^2)\frac{\partial G_k}{\partial s}(s,t) - \frac{\partial G_k}{\partial t}(s,t) \\
&= -(s-1)(\lambda+\mu)(N-k)\Big((s-1)\mu e^{-(\lambda+\mu)t} + \lambda s + \mu\Big)^k \lambda\Big(-(s-1)\lambda e^{-(\lambda+\mu)t} + \lambda s + \mu\Big)^{N-k-1} e^{-(\lambda+\mu)t} \\
&+(s-1)(\lambda+\mu)k\mu\Big((s-1)\mu e^{-(\lambda+\mu)t} + \lambda s + \mu\Big)^{k-1}\Big(-(s-1)\lambda e^{-(\lambda+\mu)t} + \lambda s + \mu\Big)^{N-k} e^{-(\lambda+\mu)t} \\
&+(s-1)\Big((s-1)\mu e^{-(\lambda+\mu)t} + \lambda s + \mu\Big)^k N\lambda\Big(-(s-1)\lambda e^{-(\lambda+\mu)t} + \lambda s + \mu\Big)^{N-k} \\
&+\Big(\lambda s^2 - (\lambda-\mu)s - \mu\Big)(N-k)\Big(\lambda e^{-(\lambda+\mu)t} - \lambda\Big)\Big((s-1)\mu e^{-(\lambda+\mu)t} + \lambda s + \mu\Big)^k \Big(-(s-1)\lambda e^{-(\lambda+\mu)t} + \lambda s + \mu\Big)^{N-k-1} \\
&-\Big(\mu e^{-(\lambda+\mu)t} + \lambda\Big)k\Big((s-1)\mu e^{-(\lambda+\mu)t} + \lambda s + \mu\Big)^{k-1}\Big(-(s-1)\lambda e^{-(\lambda+\mu)t} + \lambda s + \mu\Big)^{N-k}\Big(\lambda s^2 - (\lambda-\mu)s - \mu\Big) \\
&= 0.
\end{aligned}$$

(e) This expression follows from the relation

$$\mathbb{E}[X_t \mid X_0 = k] = \frac{\partial G_k}{\partial s}(s,t)_{|s=1}$$

and the result of Question (d).

(f) We have

$$\lim_{t\to\infty} \mathbb{E}[X_t \mid X_0 = k] = k\frac{\lambda(\lambda+\mu)^{k-1}}{(\lambda+\mu)^N}(\mu+\lambda)^{N-k} + (N-k)\frac{(\mu+\lambda)\lambda^{k-1}}{(\lambda+\mu)^N}\lambda^{N-k} = \frac{N\lambda}{\lambda+\mu}.$$

Chapter 10 - Discrete-Time Martingales

Exercise 10.2

(a) From the tower property of conditional expectations we have:

$$\mathbb{E}[M_{n+1}] = \mathbb{E}[\mathbb{E}[M_{n+1} \mid \mathcal{F}_n]] \geq \mathbb{E}[M_n], \qquad n \geq 0.$$

(b) If $(Z_n)_{n\in\mathbb{N}}$ is a process with independent increments have negative expectation, we have

$$\begin{aligned}
\mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}[Z_n \mid \mathcal{F}_n] + \mathbb{E}\left[Z_{n+1} - Z_n \mid \mathcal{F}_n\right] \\
&= \mathbb{E}[Z_n \mid \mathcal{F}_n] + \mathbb{E}[Z_{n+1} - Z_n] \\
&\leq \mathbb{E}[Z_n \mid \mathcal{F}_n] = Z_n, \qquad n \geq 0.
\end{aligned}$$

(c) We let $A_0 := 0$, $A_{n+1} := A_n + \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n]$, $n \geq 0$, and

$$N_n := M_n - A_n, \qquad n \in \mathbb{N}. \tag{B.35}$$

(i) For all $n \in \mathbb{N}$, we have

$$\begin{aligned}
\mathbb{E}[N_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}[M_{n+1} - A_{n+1} \mid \mathcal{F}_n] \\
&= \mathbb{E}[M_{n+1} - A_n - \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] \mid \mathcal{F}_n] \\
&= \mathbb{E}[M_{n+1} - A_n \mid \mathcal{F}_n] - \mathbb{E}[\mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] \mid \mathcal{F}_n] \\
&= \mathbb{E}[M_{n+1} - A_n \mid \mathcal{F}_n] - \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] \\
&= -\mathbb{E}[A_n \mid \mathcal{F}_n] + \mathbb{E}[M_n \mid \mathcal{F}_n] = M_n - A_n = N_n,
\end{aligned}$$

hence $(N_n)_{n\in\mathbb{N}}$ is a martingale with respect to $(\mathcal{F}_n)_{n\in\mathbb{N}}$.

(ii) For all $n \in \mathbb{N}$, we have

$$\begin{aligned}
A_{n+1} - A_n &= \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] \\
&= \mathbb{E}[M_{n+1} \mid \mathcal{F}_n] - \mathbb{E}[M_n \mid \mathcal{F}_n] \\
&= \mathbb{E}[M_{n+1} \mid \mathcal{F}_n] - M_n \geq 0,
\end{aligned}$$

since $(M_n)_{n\in\mathbb{N}}$ is a submartingale.

(iii) By induction we have $A_{n+1} = A_n + \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n]$, $n \in \mathbb{N}$, which is $\mathcal{F}_n$-measurable if A_n is $\mathcal{F}_{n-1}$-measurable, $n \geq 1$.

(iv) This property is obtained by construction in (B.35).

(d) For all bounded stopping times σ and τ such that $\sigma \leq \tau$ *a.s.*, we have $\mathbb{E}[N_\sigma] = \mathbb{E}[N_\tau]$ by (10.3.3), hence

$$\begin{aligned}
\mathbb{E}[M_\sigma] &= \mathbb{E}[N_\sigma] + \mathbb{E}[A_\sigma] \\
&= \mathbb{E}[N_\tau] + \mathbb{E}[A_\sigma] \\
&\leq \mathbb{E}[N_\tau] + \mathbb{E}[A_\tau] \\
&= \mathbb{E}[M_\tau],
\end{aligned}$$

by (10.3.3), since $(M_n)_{n\in\mathbb{N}}$ is a martingale and $(A_n)_{n\in\mathbb{N}}$ is nondecreasing.

Chapter 11 - Spatial Poisson Processes

Exercise 11.2 The probability that there are 10 events within a circle of radius 3 meters is

$$\mathrm{e}^{-9\pi\lambda}\frac{(9\pi\lambda)^{10}}{10!} = \mathrm{e}^{-9\pi/2}\frac{(9\pi/2)^{10}}{10!} \simeq 0.0637.$$

Exercise 11.3 The probability that more than two living organisms are in this measured volume is

$$\begin{aligned}\mathbb{P}(N \geq 3) &= 1 - \mathbb{P}(N \leq 2) = 1 - \mathrm{e}^{-10\theta}\left(1 + 10\theta + \frac{(10\theta)^2}{2}\right) \\ &= 1 - \mathrm{e}^{-6}\left(1 + 6 + \frac{6^2}{2}\right) = 1 - 25\mathrm{e}^{-6} \simeq 0.938.\end{aligned}$$

Exercise 11.4 Let X_A, resp. X_B, the number of defects found by the first, resp. second, inspection. We know that X_A and X_B are independent Poisson random variables with intensities 0.5, hence the probability that both inspections yield defects is

$$\begin{aligned}\mathbb{P}(X_A \geq 1 \text{ and } X_B \geq 1) &= \mathbb{P}(X_A \geq 1)\mathbb{P}(X_B \geq 1) \\ &= (1 - \mathbb{P}(X_A = 0))(1 - \mathbb{P}(X_B = 0)) \\ &= (1 - \mathrm{e}^{-0.5})^2 \simeq 0.2212.\end{aligned}$$

Exercise 11.5 The number X_N of points in the interval $[0, \lambda]$ has a binomial distribution with parameter $(N, \lambda/N)$, i.e.

$$\mathbb{P}(X_N = k) = \binom{N}{k}\left(\frac{\lambda}{N}\right)^k\left(1 - \frac{\lambda}{N}\right)^{N-k}, \qquad k = 0, 1, \ldots, N,$$

and we find

$$\lim_{N\to\infty} \mathbb{P}(X_N = k) = \frac{\lambda^k}{k!}\lim_{N\to\infty}\left(\left(1 - \frac{\lambda}{N}\right)^{N-k}\prod_{i=0}^{k-1}\frac{N-i}{N}\right) = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!},$$

which is the Poisson distribution with parameter $\lambda > 0$.

Exercise 11.6

(a) Based on the area $\pi r^2 = 9\pi$, this probability is given by $\mathrm{e}^{-9\pi/2}(9\pi/2)^{10}/10!$.

(b) This probability is

$$\mathrm{e}^{-9\pi/2}\frac{(9\pi/2)^5}{5!} \times \mathrm{e}^{-9\pi/2}\frac{(9\pi/2)^3}{3!}.$$

(c) This probability is $\mathrm{e}^{-9\pi}(9\pi)^8/8!$.

(d) Since the location of points are uniformly distributed by (11.1.2), the probability that a point in the disk $D((0, 0), 1)$ is located in the subdisk $D((1/2, 0), 1/2)$ is given by the ratio $\pi/4/\pi = 1/4$ of their surfaces. Hence, given that 5 items are found in $D((0, 0), 1)$, the number of points located within $D((1/2, 0), 1/2)$ has a binomial distribution with parameter $(5, 1/4)$, cf. (B.8) in the solution of Exercise 1.6 and Exercise 9.2-(d), and we find the probability

$$\binom{5}{3}(1/4)^3(3/4)^2 = \frac{45}{512} \simeq 0.08789.$$

Chapter 12 - Reliability and Renewal Processes

Exercise 12.1

(a) We have

$$F_\beta(t) = \mathbb{P}(\tau < t) = \int_0^t f_\beta(x)dx = \beta \int_0^t x^{\beta-1} \mathrm{e}^{-x^\beta} dx = -\left[\mathrm{e}^{-x^\beta}\right]_0^t = 1 - \mathrm{e}^{-t^\beta},$$

$t \in \mathbb{R}_+$.

(b) We have

$$R(t) = \mathbb{P}(\tau > t) = 1 - F_\beta(t) = \mathrm{e}^{-t^\beta}, \qquad t \in \mathbb{R}_+.$$

(c) We have

$$\lambda(t) = -\frac{d}{dt} \log R(t) = \beta t^{\beta-1} \qquad t \in \mathbb{R}_+.$$

(d) By (12.3.1) we have

$$\mathbb{E}[\tau] = \int_0^\infty R(t)dt = \int_0^\infty \mathrm{e}^{-t^\beta} dt.$$

In particular this yields $\mathbb{E}[\tau] = \sqrt{\pi}/2$ when $\beta = 2$.

References

[AKS93] S.C. Althoen, L. King, K. Schilling, How long is a game of Snakes and Ladders? Math. Gazette **77**(478), 71–76 (1993)
[Asm03] S. Asmussen, Applied probability and queues, in *Applications of Mathematics*, vol. 51, 2nd edn. Stochastic Modelling and Applied Probability (Springer, New York)
[BN96] D. Bosq, H.T. Nguyen, *A Course in Stochastic Processes: Stochastic Models and Statistical Inference*. Mathematical and Statistical Methods (Kluwer, 1996)
[Çin75] E. Çinlar, *Introduction to Stochastic Processes* (Prentice-Hall Inc., Englewood Cliffs, 1975)
[Çın11] E. Çınlar, *Probability and Stochastics*, vol. 261, Graduate Texts in Mathematics (Springer, New York, 2011)
[Dev03] J.L. Devore, *Probability and Statistics for Engineering and the Sciences*, 6th edn. (Duxbury Press, 2003)
[Doo53] J.L. Doob, *Stochastic Processes* (Wiley, New York, 1953)
[Doo84] J.L. Doob, *Classical Potential Theory and Its Probabilistic Counterpart* (Springer, Berlin, 1984)
[Dur99] R. Durrett, *Essentials of Stochastic Processes* (Springer, 1999)
[GS01] G.R. Grimmett, D.R. Stirzaker, *Probability and Random Processes* (Oxford University Press, Oxford, 2001)
[IM11] J.L. Iribarren, E. Moro, Branching dynamics of viral information spreading. Phys. Rev. E **84**, 046116, 13 (2011)
[JP00] J. Jacod, P. Protter, *Probability Essentials* (Springer, Berlin, 2000)
[JS01] P.W. Jones, P. Smith, *Stochastic Processes: An Introduction*. Arnold Texts in Statistics (Hodder Arnold, 2001)
[Kal02] O. Kallenberg, *Foundations of Modern Probability*, 2nd edn. Probability and its Applications (Springer, New York, 2002)
[Kij97] M. Kijima, *Markov Processes for Stochastic Modeling*. Stochastic Modeling Series (Chapman & Hall, London, 1997)
[KT81] S. Karlin, H.M. Taylor, *A Second Course in Stochastic Processes* (Academic Press Inc., New York, 1981)
[Laf91] A. Laforgia, Bounds for modified Bessel functions. J. Comput. Appl. Math. **34**(3), 263–267 (1991)
[Lal] S.P. Lalley, Continuous time Markov chains, http://galton.uchicago.edu/lalley/Courses/313/ContinuousTime.pdf. Accessed 10 Dec 2015
[LPW09] D.A. Levin, Y. Peres, E.L. Wilmer, *Markov Chains and Mixing Times* (American Mathematical Society, Providence, RI, 2009). With a chapter by J.G. Propp, D.B. Wilson

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4

[Med10] J. Medhi, *Stochastic Processes*, 3rd edn. (New Age Science Limited, Tunbridge Wells, UK, 2010)

[MT15] I. Muni Toke, The order book as a queueing system: average depth and influence of the size of limit orders. Quant. Finan. **15**(5), 795–808 (2015)

[Nor98] J.R. Norris, Markov Chains, in *Cambridge Series in Statistical and Probabilistic Mathematics*, vol. 2 (Cambridge University Press, Cambridge, 1998)

[OSA+09] T. Olaleye, F.A. Sowunmi, S. Abiola, M.O. Salako, I.O. Eleyoowo, A Markov chain approach to the dynamics of vehicular traffic characteristics in Abeokuta metropolis. Res. J. Appl. Sci. Eng. Technol. **1**(3), 160–166 (2009)

[Pit99] J. Pitman, *Probability* (Springer, 1999)

[Pri09] N. Privault, *Stochastic Analysis in Discrete and Continuous Settings with Normal Martingales*, vol. 1982. Lecture Notes in Mathematics (Springer, Berlin, 2009)

[Pri13] N. Privault, *Understanding Markov Chains*. Springer Undergraduate Mathematics Series (Springer, 2013)

[Ros96] S.M. Ross, *Stochastic Processes*, 2nd edn. Wiley Series in Probability and Statistics: Probability and Statistics (Wiley, New York, 1996)

[Sig] K. Sigman, Continuous-time Markov chains, http://www.columbia.edu/~ks20/stochastic-I/stochastic-I-CTMC.pdf. Accessed 12 Sept 2014

[SSB08] S.H. Sellke, N.B. Shroff, S. Bagchi, Modeling and automated containment of worms. IEEE Trans. Depend. Secur. Comput. **5**(2), 71–86 (2008)

[ST08] G.M. Schütz, S. Trimper, Elephants can always remember: Exact long-range memory effects in a non-Markovian random walk. Phys. Rev. E **3** (2008)

[Ste01] J.M. Steele, *Stochastic Calculus and Financial Applications*, vol. 45, Applications of Mathematics (Springer, New York, 2001)

Subject Index

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4

Author Index

N. Privault, *Understanding Markov Chains*, Springer Undergraduate Mathematics Series, https://doi.org/10.1007/978-981-13-0659-4