

Contents

In	roduction 12	
1	Probability Background	21
	1.1 Probability Spaces and Events	21
	1.2 Probability Measures	26
	1.3 Conditional Probabilities and Independence	28
	1.4 Random Variables	30
	1.5 Probability Distributions	33
	1.6 Expectation of Random Variables	42
	1.7 Moment and Probability Generating Functions	55
	Exercises	61
2	Gambling Problems	65
	2.1 Constrained Random Walk	65
	2.2 Ruin Probabilities	68
	2.3 Mean Game Duration	85
	Exercises	96
3	Random Walks	106
-	3.1 Unrestricted Random Walk	106
	3.2 Mean and Variance	107
	3.3 Distribution	108
	3.4 First Return to Zero	110
	Exercises	123
4	Discrete-Time Markov Chains	141
-	4.1 Markov Property	- <u>141</u> 0

4.2	Transition matrix	144
4.3	Examples of Markov Chains	149
4.4	Higher-Order Transition Probabilities	155
4.5	The Two-State Discrete-Time Markov Chain	160
Exe	rcises	170

5	First Step Analysis	187
	5.1 Hitting Probabilities	187
	5.2 Mean Hitting and Absorption Times	195
	5.3 First Return Times	203
	5.4 Mean Number of Returns	211
	Exercises	217

6	Classification	of States
---	----------------	-----------

Classification of States	
6.1 Communicating States	242
6.2 Recurrent States	240
6.3 Transient States	248
6.4 Positive vs Null Recurrence	25
6.5 Periodicity and Aperiodicity	250
Exercises	260

Long-Run Behavior of Markov Chains	
7.1 Limiting Distributions	265
7.2 Stationary Distributions	268
7.3 Markov Chain Monte Carlo	284
Exercises	288

8	Branching Processes	310
	8.1 Construction and Examples	310
	8.2 Probability Generating Functions	314
	8.3 Extinction Probabilities	318
	Exercises	332

9	Continuous-Time Markov Chains	338
	9.1 The Poisson Process	. Y 33 8

9.2	Continuous-Time Markov Chains	3 45
9.3	Transition Semigroup	354
9.4	Infinitesimal Generator	359
9.5	The Two-State Continuous-Time Markov Chain	369
9.6	Limiting and Stationary Distributions	376
9.7	The Discrete-Time Embedded Chain	385
9.8	Mean Absorption Time and Probabilities	391
Exe	ercises	398
10 Dis	crete-Time Martingales	406
10.1	Filtrations and Conditional Expectations	406
10.2	2 Martingales - Definition and Properties	408
10.3	3 Stopping Times	410
10.4	4 Ruin Probabilities	415
10.5	5 Mean Game Duration	419
Exe	ercises	423
11 Spa	ntial Poisson Processes	431
11.1	Spatial Poisson (1781-1840) Processes	431
11.2	2 Poisson Stochastic Integrals	433
11.3	3 Transformations of Poisson Measures	435
11.4	4 Moments of Poisson Stochastic Integrals	438
11.5	5 Deviation Inequalities	442
Exe	ercises	445
12 Reli	iability Theory	448
12.1	l Survival Probabilities	448
12.2	2 Poisson Process with Time-Dependent Intensity	450
12.3	B Mean Time to Failure	452
Exe	ercises	453
Some	Useful Identities	455
Solutio	ons to the Exercises	458
Cha	$to ter 1 \ldots $	458

Chapter 2	473
Chapter 3	493
Chapter 4	525
Chapter 5	552
Chapter 6	610
Chapter 7	618
Chapter 8	675
Chapter 9	692
Chapter 10	726
Chapter 11	734
Chapter 12	739
Index	740
Author index	747
Bibliography	749

List of Figures

0.1	Sample trajectories of Brownian motion	15
0.2	Sample trajectory of 3 -dimensional Brownian motion	15
0.3	Twenty paths of a drifted Brownian motion, $d = 1$	16
0.4	Sample trajectories of a Poisson process	16
0.5	Sample trajectories of a compound Poisson process	17
0.6	Sample trajectories of a gamma process	17
0.7	Sample trajectories of a stable process	18
0.8	Sample trajectories of a Cauchy process	18
0.9	Sample trajectories of a variance gamma process	19
0.10	Sample trajectories of an inverse Gaussian process	19
0.11	Sample trajectories of a negative inverse Gaussian process	20
1.1	Probability computed as a volume integral	38
1.1 2.1	Probability computed as a volume integral \ldots	38 67
 1.1 2.1 2.2 	Probability computed as a volume integral \ldots	38 67 75
 1.1 2.1 2.2 2.3 	Probability computed as a volume integral $\dots \dots \dots$	38 67 75 83
 1.1 2.1 2.2 2.3 2.4 	Probability computed as a volume integral $\dots \dots \dots$	38 67 75 83 84
 1.1 2.1 2.2 2.3 2.4 2.5 	Probability computed as a volume integral \dots	38 67 75 83 84 85
 1.1 2.1 2.2 2.3 2.4 2.5 2.6 	Probability computed as a volume integral $\dots \dots \dots$	38 67 75 83 84 85 93
$1.1 \\ 2.1 \\ 2.2 \\ 2.3 \\ 2.4 \\ 2.5 \\ 2.6 \\ 2.7 \\$	Probability computed as a volume integral	38 67 75 83 84 85 93 95
$1.1 \\ 2.1 \\ 2.2 \\ 2.3 \\ 2.4 \\ 2.5 \\ 2.6 \\ 2.7 \\ 3.1$	Probability computed as a volume integral \dots Sample path of a gambling process $(X_n)_{n\in\mathbb{N}}$ \dots Ruin probability $f_{20}(k)$ function of $X_0 = k \in [0, 20]$ \dots Ruin probability $f_{20}(k)$ function of $X_0 = k \in [0, 20]$ \dots Ruin probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$ \dots Ruin probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$ \dots Ruin probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$ \dots Ruin probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$ \dots Ruin probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$ \dots Ruin probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$ \dots Ruin probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$ \dots Ruin probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$ \dots Ruin probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$ \dots Ruin probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$ \dots Ruin probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$ \dots Ruin Probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$ \dots Ruin Probability as a function of $p \in [0, 1]$ for $S = 20$ and $k = 10$ \dots Ruin Probability Appendix Appe	38 67 75 83 84 85 93 95 109

3.3	Last return to state 0 at time $k = 10$	193
3.4	Probability generating functions of T_0^r	121
3.5	Illustration of the range process [*]	133
3.6	Random walk with $cookies^*$	137
4.1	Mozart: variations KV 265 for Piano*	153
4.2	Graph of a five-state Markov chain*	155
4.3	Sample path of a two-state chain in continuous time	167
4.4	Proportion of chain values at state 1 converges to $a/(a+b)$	167
4.5	Convergence graph for the two-state Markov chain [*]	169
4.6	Sample path of a five-state Markov chain [*]	170
4.7	Lafenwa intersection	177
5.1	Mean return times to state 0 on the maze (5.3.7) $\ldots \ldots \ldots$	211
7.1	Global balance condition (discrete time)	270
7.2	Stationary distribution on the maze $(5.3.7)$	282
7.3	Stationary distribution by state numbering	282
7.4	Simulated stationary distribution	283
7.5	Detailed balance condition (discrete time)	285
7.6	Simulation of the voter model with $N=199^*$	305
7.7	Simulation of the voter model with $N=3^*$	309
8.1	Example of branching process	312
8.2	Sample graph of a branching process	313
8.3	Generating function of Y_1 with $p = 0.65$	324
8.4	Generating function of Y_1 with $p = 3/8 < 1/2$ and $\alpha_1 = 1$.	326
8.5	Generating function of Y_1 with $p = 1/2$ and $\alpha_1 = 1$	327
8.6	Generating function of Y_1 with $p = 3/4 > 1/2$ and $\alpha_1 = q/p = 1/3$.	327
8.7	Probability generating function of Y_1	329
8.8	Simulation of influenza spread [*] \ldots	4 v329 h
	· 1/11/0/12	

8.9	Sample path of a branching process $(X_n)_{n \ge 0}^*$	33 0
9.1	Sample path of a Poisson process $(N_t)_{t \in \mathbb{R}_+}$	339
9.2	Sample path of a birth process $(X_t^b)_{t\in\mathbb{R}_+}$	348
9.3	Sample path of a death process $(X_t^d)_{t \in \mathbb{R}_+}$	351
9.4	Sample path of a birth and death process $(X_t)_{t\in\mathbb{R}_+}$	353
9.5	Sample path of a continuous-time two-state chain	375
9.6	Proportion of process values in state 1 converges to $1/3 = \alpha/(\alpha + \beta)$	375
9.7	Global balance condition (continuous time) \ldots	377
9.8	Birth and death process with its embedded chain	385
9.9	Discrete-time embedded chain based on the path of Figure 9.8	386
10.1	Updated weather forceast	410
10.1 10.2	Stopped process	410
10.2	$ \begin{array}{c} \text{Stopped process} & \ldots & $	410
10.0	Possible paths of the stopped process $(M,)$	425
10.1	$1 \text{ ossible paths of the stopped process } (1 \tau \wedge n) n \in \mathbb{N} \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots $	120
S.1	Ruin probability $f_8(k)$ as a function of $X_0 = k \in [0,9]$	488
S.2	Mean game duration $h_{20}(k)$ as a function of $X_0 = k \in [0, 20]$	490
S.3	Standard deviation of $ au$ as a function of $X_0 = k \in [0, 20]$	493
S.4	Four paths leading from 0 to 2 in four time steps $\ldots \ldots \ldots$	494
S.5	Last return to state 0 at time $k = 10$	509
S.6	$Plot of estimate \$hmm\$emissionProbs[1,] \ldots \ldots$	547
S.7	$Plot of estimate \$hmm\$emissionProbs[2,] \dots \dots$	547
S.8	Plot of $\eta \mapsto (M_{0,\eta}/M_{0,"})((M_{1,"}-M_{1,\eta})/M_{1,"})^2$	549
S.9	Frequency analysis of alphabet letters	550
S.10	$Plot of estimate \$hmm\$emissionProbs[3,] \dots \dots$	551
S.11	$Plot of estimate \$hmm\$emissionProbs[1,] \dots \dots$	552
S.12	Graph of 928 paths leading from 0 to state 13^*	569
S.13	Graph of mean times to reach the set A with $N = 10^*$	581
S.14	Graph of mean times to reach the set A with $N = 10^*$	Y 583 0

S.15 Optimal value function with $p = 0$	ÌÔC
S.16 Optimal value function with $0 $	507
S.17 Optimal value function with $p = 1/2$	508
S.18 Optimal value function with $1/2 $	509
S.19 Stationary distribution as a function of $\varepsilon \in [0,1]$	59
S.20 Mean return times as functions of $\varepsilon \in [0,1]$	60
S.21 Graphs of distance to stationarity $d(n)$ and its upper bound $(1 - \theta)^n$	666
S.22 Probability of a majority of "+" as a function of $p \in [0, 1]$	573
S.23 Supremum deviation probability $\ldots \ldots \ldots$	'33
S.24 Martingale supremum as a function of time [*] \ldots	'35

^{*}Animated figures (work in Acrobat reader).

List of Tables

7.1	Summary of Markov chain properties	287
8.1	Summary of computing methods	331
9.1	Summary of computing methods	397
10.1	List of martingales	422
14.1	Number of paths starting from $k = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$.	569
14.2	Number of paths taking k rounds, $k = 3, 4, 5, 6, 7, 8, 9, 10, 11$.	569
14.3	Values of $\mu_A(x,y)$ with $N=2$	581
14.4	Values of $\mu_A(x,y)$ with $N=3$	583

Clicking on an exercise number inside the solution section will send to the original problem text inside the file. Conversely, clicking on the problem number sends the reader to the corresponding solution, however this feature should not be misused. This text contains 144 exercises and 26 problems with solutions, and 98 figures.



Introduction

A stochastic^{*} 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 an *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 that 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 time scale 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 time scale 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. $T = [0, 1], T = \mathbb{R}_+$). 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

^{*}From the Greek " $\sigma \tau \delta \chi o \varsigma$ " (stokhos), meaning "guess", or "conjecture".

processes would actually require additional background in *measure theory*, which is outside of the scope of this text. Measure theory $\frac{1}{13}$ 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 (1856-1922) Processes.

Roughly speaking, a stochastic 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 Chapter 1 the probabilistic tools required in the analysis of Markov chains, we consider simple gambling problems in Chapter 2, due to their practical usefulness and to the fact that they only require a minimal theoretical background. Next, in Chapter 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 Chapter 4 we introduce the general framework of Markov chains in discrete time, which includes the gambling process and the simple random walk of Chapters 2 and 3 as particular cases. In the subsequent Chapters 5, 6 and 7, Markov chains are considered from the point of view of first step analysis, which is introduced in detail in Chapter 5. The classification of states is reviewed in Chapter 6, with application to the long-run behavior of Markov chains in Chapter 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.q. for population dynamics or the control of disease outbreak, and they are considered in Chapter 8. Then in Chapter 9, we deal with Markov chains in continuous time, including Poisson and birth and death processes. Martingales are considered in Chapter 10, where they are used to recover in a simple and elegant way the main results of Chapter 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 Chapter 11. The end of that chapter also includes some recent results on moments and deviation inequalities for Poisson stochastic integrals, whose proofs should be accessible to advanced undergraduates. Reliability theory is an important engineering application of Markov chains, and it is reviewed in Chapter 12. All stochastic processes considered in this text have a discrete state space and discontinuous trajectories. Brownian motion, which is the first main example of a stochastic process having continuous trajectories in continuous time, remains outside the scope of this book.

We close this introduction with examples of simulations for the random paths of some stochastic processes with independent increments, that belong to the class of *Lévy processes* which are processes with independent and stationary increments. The graphs shown below are

for illustration purposes only. Indeed, apart from the Poisson process of Figure 0.4, most of these processes are however not within $\frac{1}{2}$ scope of this text, and they are presented as an incentive to delve deeper into the field of stochastic processes.

Examples

i) Standard Brownian motion, d = 1.



Figure 0.1: Sample trajectories of Brownian motion.

ii) Standard Brownian motion, d = 3.



Figure 0.2: Sample trajectory of 3-dimensional Brownian motion.



Figure 0.3: Twenty paths of a drifted Brownian motion, d = 1.

iv) The Poisson process, d = 1.



Figure 0.4: Sample trajectories of a Poisson process.

v) Compound Poisson process, d = 1.



Figure 0.5: Sample trajectories of a compound Poisson process.

vi) Gamma process, d = 1.



Figure 0.6: Sample trajectories of a gamma process.

vii) <u>Stable process</u>, d = 1.



Figure 0.7: Sample trajectories of a stable process.

viii) Cauchy process, d = 1.



Figure 0.8: Sample trajectories of a Cauchy process.

ix) Variance gamma process, d = 1.



Figure 0.9: Sample trajectories of a variance gamma process.

x) Inverse Gaussian process, d = 1.



Figure 0.10: Sample trajectories of an inverse Gaussian process.



Figure 0.11: Sample trajectories of a negative inverse Gaussian process.

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	21
1.2	Probability Measures	26
1.3	Conditional Probabilities and Independence	2 8
1.4	Random Variables	3 0
1.5	Probability Distributions	33
1.6	Expectation of Random Variables	42
1.7	Moment and Probability Generating Functions	55
Exe	rcises	61

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 distinguish 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: $\Omega = \{1, 2, 3, \dots, 52\}$.
- 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.

- 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: $\Omega = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}.$

♦ 23

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)
$$\boldsymbol{\emptyset} \in \boldsymbol{\mathcal{G}}$$
,

(ii) For all countable sequences $(A_n)_{n \ge 1}$ such that $A_n \in \mathcal{G}$, $n \ge 1$, we have $\bigcup_{n \ge 1} A_n \in \mathcal{G}$,

(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

$$\bigcap_{n \ge 1} A_n = \left(\bigcup_{n \ge 1} A_n^c\right)^c \in \mathcal{G}, \tag{1.1.1}$$
MH3512 AY19-20

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

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 $\Omega_{\text{He}}^{\circ}$ knowledge of whether the outcome is equal to 6 or not.

iv) Take

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

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

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

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

$$(1.1.2)$$

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

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

$$4 = \binom{n}{1} \text{ events of cardinality } 1,$$

$$6 = \binom{n}{2} \text{ events of cardinality } 2,$$

$$4 = \binom{n}{3} \text{ events of cardinality } 3,$$

$$1 = \binom{n}{4} \text{ event of cardinality } 4,$$
with $\underline{n = 4}$, for a total of

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 distinguish between the <u>outcome</u> $\omega \in \Omega$ and its associated <u>event</u> $\{\omega\} \in \mathcal{F}$, 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)
$$\mathbb{P}\left(\bigcup_{n \ge 1} A_n\right) = \sum_{n \ge 1} \mathbb{P}(A_n)$$
, whenever $A_k \cap A_l = \emptyset$, $k \ne l$.

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*

$$\mathbb{P}(A^c) = \mathbb{P}(\Omega \setminus A) = \mathbb{P}(\Omega) - \mathbb{P}(A) = 1 - \mathbb{P}(A).$$
 MH3512 AY19-20

When \boldsymbol{A} and \boldsymbol{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

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

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

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

and

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

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

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

In addition, we have the following convergence properties.

1. Let $(A_n)_{n \in \mathbb{N}}$ be a non-decreasing sequence of events, *i.e.* $A_n \subset A_{n+1}$, $n \in \mathbb{N}$. Then we have

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}}A_n\right) = \lim_{n\to\infty}\mathbb{P}(A_n).$$
(1.2.3)

MH3512 AY19-20

2. Let $(A_n)_{n \in \mathbb{N}}$ be a non-increasing sequence of events, i.e. $A_{n+1} \subset A_n$, $n \in \mathbb{N}$. Then we have

$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n).$$
(1.2.4)

 $\bigcirc 28$

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

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

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 $\frac{\mathbb{P}(A \cap W)}{\mathbb{P}(W)}$ of women who like apples, and - the relative proportion $\frac{\mathbb{P}(A \cap M)}{\mathbb{P}(M)}$ 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 <u>conditional probabilities</u>, 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

$$\mathbb{P}(A \mid Y) = rac{\mathbb{P}(Y \cap A)}{\mathbb{P}(Y)}$$
 MH3512 AY19-20

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

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

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:

$$egin{array}{rl} \mathbb{P}\left(B\cap igcup_{n\geqslant 1}A_n
ight) &=& \sum\limits_{n\geqslant 1}\mathbb{P}(B\cap A_n) \ &=& \sum\limits_{n\geqslant 1}\mathbb{P}(B\mid A_n)\mathbb{P}(A_n) \ &=& \sum\limits_{n\geqslant 1}\mathbb{P}(A_n\mid B)\mathbb{P}(B), \end{array}$$

for any family of disjoint events $(A_n)_{n \ge 1}$ with $A_i \cap A_j = \emptyset$, $i \ne j$, and $\mathbb{P}(B) > 0$, $n \ge 1$. 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)
$$\mathbb{P}\left(\bigcup_{n\geq 1} A_n \middle| B\right) = \sum_{n\geq 1} \mathbb{P}(A_n \mid B)$$
, whenever $A_k \cap A_l = \emptyset$, $k \neq l$.

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

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

(530)

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

$$\mathbb{P}\left(A \Big| igcup_{n \geqslant 1} B_n
ight)
eq \sum_{n \geqslant 1} \mathbb{P}(A \mid B_n),$$

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

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

Independent events

Two events A and B such that $\mathbb{P}(A), \mathbb{P}(B) > 0$ 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

$$egin{array}{lll} X &: & \Omega \longrightarrow \mathbb{R} \ & & & & & \ & & \omega \longmapsto X(\omega) \end{array} \end{array}$$
 MH3512 AY19-2

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

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

a random variable and A a (measurable)^{*} 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

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

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

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

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

$$\mathbb{1}_A \; : \; \Omega \longrightarrow \{0,1\}$$

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

$$\omega \mapsto \mathbb{1}_A(\omega)$$
 $\circlearrowright 32$

defined by

$$\mathbb{1}_A(\omega) = \left\{egin{array}{cc} 1 & ext{if } \omega \in A, \ 0 & ext{if } \omega \notin A, \end{array}
ight.$$

with the property

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

since

$$egin{aligned} &\omega\in A\cap B \iff \{\omega\in A ext{ and } \omega\in B\} \ &\iff \{\mathbbm{1}_A(\omega)=1 ext{ and } \mathbbm{1}_B(\omega)=1\} \ &\iff \mathbbm{1}_A(\omega)\mathbbm{1}_B(\omega)=1. \end{aligned}$$

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

$$\mathbb{1}_{\{k\}}(l) = egin{cases} 1 & ext{if } k = l, \ 0 & ext{if } k
eq l. \end{cases}$$

Given \boldsymbol{X} a random variable, we also let

$$\mathbbm{1}_{\{X=n\}} = \left\{egin{array}{ll} 1 & ext{if } X=n, \ 0 & ext{if } X
eq n, \end{array}
ight.$$
 MH3512 AY19-20

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

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

or

$$\{\mathbb{P}(X\leqslant a)\;:\;a\in\mathbb{R}\}, \hspace{1em} ext{or}\hspace{1em}\{\mathbb{P}(X\geqslant a)\;:\;a\in\mathbb{R}\},$$

see e.g. Corollary 3.8 in [Cin11].

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 \boldsymbol{A} and \boldsymbol{B} of \mathbb{R} .

Distributions admitting a density

We say that the distribution of X admits a probability <u>density</u> 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

$$\mathbb{P}(a \leqslant X \leqslant b) = \int_a^b f_X(x) dx.$$
 MH3512 AY19-20

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

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

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

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

and this is not a contradiction.

In particular, Remark 1.2 shows that

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

for $a \leq b$. 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

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

and

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

as

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

MH3512 AY19-20

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

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

ii) The *Gaussian* distribution.

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

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

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

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

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

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

We also have

$$\mathbb{P}(X>t)=\mathrm{e}^{-\lambda t}, \qquad t\in\mathbb{R}_+.$$
 MH3512(1,593)

iv) The gamma distribution.

The probability density function of the gamma distribution is given by

$$f(x):=rac{a^\lambda}{\Gamma(\lambda)}\mathbbm{1}_{[0,\infty)}(x)x^{\lambda-1}\mathrm{e}^{-ax}=\left\{egin{array}{c} rac{a^\lambda}{\Gamma(\lambda)}x^{\lambda-1}\mathrm{e}^{-ax}, & x\geqslant 0\ 0, & x<0, \end{array}
ight.$$

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

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

is the gamma function.

v) The *Cauchy* distribution.

The probability density function of the Cauchy distribution is given by

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

vi) The *lognormal* distribution.
The probability density function of the lognormal distribution is given by

$$f(x):=\mathbbm{1}_{[0,\infty)}(x)rac{1}{x\sigma\sqrt{2\pi}}\mathrm{e}^{-rac{(\mu-\log x)^2}{2\sigma^2}}=\left\{egin{array}{c} rac{1}{x\sigma\sqrt{2\pi}}\mathrm{e}^{-rac{(\mu-\log x)^2}{2\sigma^2}}, & x\geqslant 0\ 0, & x<0. \end{array}
ight.$$

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

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

is satisfied.

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

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

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

when

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

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

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

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

MH3512 AY19-20



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

and

$$(x,y)\longmapsto \mathbb{P}(X\geqslant x ext{ and } Y\geqslant y) = \int_x^\infty \int_y^\infty f_{(X,Y)}(s,t) ds dt,$$

as

$$f_{(X,Y)}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{(X,Y)}(x,y)$$
(1.5.4)
$$= \frac{\partial^2}{\partial x \partial y} \int_{-\infty}^x \int_{-\infty}^y f_{(X,Y)}(s,t) ds dt$$
(1.5.5)
$$= \frac{\partial^2}{\partial x \partial y} \int_x^\infty \int_y^\infty f_{(X,Y)}(s,t) ds dt,$$

38

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

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 <u>marginal densities</u> of (X, Y) and are given by

$$f_X(x) = \int_{-\infty}^\infty f_{(X,Y)}(x,y) dy, \qquad x \in \mathbb{R}, ext{MH3512} (1.5.6) ext{MH3512} (1.5.6)$$

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

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

$$\mathbb{P}(\min(X_1, \dots, X_n) > t) = \mathbb{P}(X_1 > t, \dots, X_n > t)
= \mathbb{P}(X_1 > t) \cdots \mathbb{P}(X_n > t)
= e^{-(\lambda_1 + \dots + \lambda_n)t}, \quad t \in \mathbb{R}_+,$$
(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

$$f_{(X_1,X_2)}(x,y)=f_{X_1}(x)f_{X_2}(y)=\lambda_1\lambda_2\mathrm{e}^{-\lambda_1x-\lambda_2y},\qquad x,y\geqslant 0,$$

we can write

$$egin{array}{rll} \mathbb{P}(X_1 < X_2) &=& \mathbb{P}(X_1 \leqslant X_2) \ &=& \int_0^\infty \int_0^y f_{(X_1,X_2)}(x,y) dx dy \ &=& \lambda_1 \lambda_2 \int_0^\infty \int_0^y \mathrm{e}^{-\lambda_1 x - \lambda_2 y} dx dy \end{array}$$

(39

MH3512 AY19-20

$$= rac{\lambda_1}{\lambda_1+\lambda_2},$$



and we note that

$$\mathbb{P}(X_1=X_2)=\lambda_1\lambda_2\int_{\{(x,y)\in\mathbb{R}^2_+\,:\,x=y\}}\mathrm{e}^{-\lambda_1x-\lambda_2y}dxdy=0.$$

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

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

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

MH3512 AY19-20

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

iii) The geometric distribution.

In this case, we have

$$\mathbb{P}(X=k) = (1-p)p^k, \quad k \in \mathbb{N},$$
(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,^{*}

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

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

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

where $p \in (0, 1)$ and $r \ge 1$ are parameters. Note that the sum of $r \ge 1$ 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

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

^{*}The notation "inf" stands for "infimum", meaning the smallest $n \ge 0$ such that $X_n = 0$, if such an n exists.

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

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

and we have

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

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

$$1=\mathbb{P}(X\in\mathbb{R})=\mathbb{P}(X\in\mathbb{N})=\sum\limits_{k\geqslant 0}\mathbb{P}(X=k)=0,$$

which is a contradiction.

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

$$\mathbb{P}(X = n \mid Y = k) = rac{\mathbb{P}(X = n ext{ and } Y = k)}{\mathbb{P}(Y = k)}, \quad n \in \mathbb{N},$$

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

1.6 Expectation of Random Variables

The <u>expectation</u>, or <u>expected value</u>, 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 $\Phi_{ay}^{3,3}$ 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

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

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

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

$$\mathbb{E}[\phi(X)] = \sum\limits_{k \geqslant 0} \phi(k) \mathbb{P}(X = k),$$

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

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

MH3512 AY19-20

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

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

provided that

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

Examples

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

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

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)</pre>

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

$$\mathbb{P}(X<\infty)=\sum\limits_{k\geqslant 0}\mathbb{P}(X=k),$$

and

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

and in general

$$\mathbb{E}[X] = +\infty imes \mathbb{P}(X = \infty) + \sum_{k \geqslant 0} k \mathbb{P}(X = k)$$

MH3512 AY19-20

ॳ

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 \boldsymbol{X} has the geometric distribution

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

with parameter p = 1/2, and

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

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

$$\mathbb{P}(\phi(X)<\infty)=\mathbb{P}(X<\infty)=\sum\limits_{k\geqslant 0}rac{1}{2^{k+1}}=1,$$

and

$$\mathbb{E}[\phi(X)] = \sum\limits_{k \geqslant 0} \phi(k) \mathbb{P}(X=k) = \sum\limits_{k \geqslant 0} rac{2^k}{2^{k+1}} = \sum\limits_{k \geqslant 0} rac{1}{2} = +\infty,$$

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

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

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

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

^{*}This is the St. Petersburg paradox.

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

$$\mathbb{E}\left[\mathrm{e}^{\lambda X}
ight] = \mu \int_{0}^{\infty} \mathrm{e}^{\lambda x} \mathrm{e}^{-\mu x} dx = \left\{egin{array}{cc} rac{1}{\mu-\lambda} < \infty & ext{if } \mu > \lambda, \ +\infty, & ext{if } \mu \leqslant \lambda. \end{array}
ight.$$

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

$$\mathbb{E}[X \mid A] = \sum_{k \geqslant 0} k \mathbb{P}(X = k \mid A) = \sum_{k \geqslant 0} k rac{\mathbb{P}(X = k ext{ and } A)}{\mathbb{P}(A)}$$

Lemma 1.4

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

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

$$\begin{split} \mathbb{E}[X \mid A] &= \frac{1}{\mathbb{P}(A)} \sum_{k \geqslant 0} k \mathbb{P}(X = k \mid A) \\ &= \frac{1}{\mathbb{P}(A)} \sum_{k \geqslant 0} k \mathbb{P}(\{X = k\} \cap A) = \frac{1}{\mathbb{P}(A)} \sum_{k \geqslant 0} k \mathbb{E}\left[\mathbb{1}_{\{X = k\} \cap A}\right] \\ &= \frac{1}{\mathbb{P}(A)} \sum_{k \geqslant 0} k \mathbb{E}\left[\mathbb{1}_{\{X = k\}} \mathbb{1}_{A}\right] = \frac{1}{\mathbb{P}(A)} \mathbb{E}\left[\mathbb{1}_{A} \sum_{k \geqslant 0} k \mathbb{1}_{\{X = k\}}\right] \\ &= \frac{1}{\mathbb{P}(A)} \mathbb{E}\left[\mathbb{1}_{A} X\right], \end{split}$$

where we used the relation

$$X = \sum\limits_{k \geqslant 0} k \mathbb{1}_{\{X = k\}}$$

which holds since \boldsymbol{X} takes only integer values.

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

$$\{X>0\}=\{1,3,5,7\}\subset\Omega,$$
 MH3512 AY19-20

047

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

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

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])</pre>

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

$$\mathbb{E}[X] = (1-p)\sum\limits_{k \geqslant 1} kp^k = rac{p}{1-p} = 3,$$

and

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

MH3512 AY19-20

 $\mathbf{348}$

ර 49

$$= \frac{p(1-p^{10}-10(1-p)p^9)}{(1-p)(1-p^{10})}$$

\$\approx 2.4032603455.

If the random variable $X: \Omega \longrightarrow \mathbb{N}$ is independent of the event A^* 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

$$\mathbb{E}[X \mid A] = \mathbb{E}[X]. \tag{1.6.7}$$

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

$$egin{array}{lll} \mathbbm{1}_A: \ \Omega \longrightarrow \{0,1\} \ \omega \longmapsto \mathbbm{1}_A:= \left\{egin{array}{lll} 1 & ext{if } \omega \in A, \ 0 & ext{if } \omega \notin A, \end{array}
ight.$$

shows that, in particular,

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

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

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

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

Proposition 1.5

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

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

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]], \qquad (1.6.8)$$

which is sometimes referred to as the tower property.

Proof

We have

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

where we used the marginal distribution

$$\mathbb{P}(X=n) = \sum_{k \ge 0} \mathbb{P}(X=n ext{ and } Y=k), \quad n \in \mathbb{N},$$

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

Taking

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

with $A_k := \{Y = k\}, k \in \mathbb{N}$, from (1.6.8) we also get the *law of total expectation*

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

$$= \sum_{k \ge 0} \mathbb{E}[X \mid Y = k] \mathbb{P}(Y = k)$$
(1.6.9)
MH3512 AY19-20

$$= \sum_{k \ge 0} \mathbb{E}[X \mid A_k] \mathbb{P}(A_k).$$

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

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

and we have

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 <u>variance</u> of a random variable \boldsymbol{X} is defined in general by

$$\operatorname{Var}[X] := \mathbb{E}[X^2] - (\mathbb{E}[X])^2,$$
 MH3512 AY19-20

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

$$\operatorname{Var}\left[\sum_{k=1}^{n} X_{k}\right] = \mathbb{E}\left[\left(\sum_{k=1}^{n} X_{k}\right)^{2}\right] - \left(\mathbb{E}\left[\sum_{k=1}^{n} X_{k}\right]\right)^{2}$$

$$= \mathbb{E}\left[\sum_{k=1}^{n} X_{k} \sum_{l=1}^{n} X_{l}\right] - \mathbb{E}\left[\sum_{k=1}^{n} X_{k}\right] \mathbb{E}\left[\sum_{l=1}^{n} X_{l}\right]$$

$$= \mathbb{E}\left[\sum_{k=1}^{n} \sum_{l=1}^{n} X_{k} X_{l}\right] - \sum_{k=1}^{n} \sum_{l=1}^{n} \mathbb{E}[X_{k}] \mathbb{E}[X_{l}]$$

$$= \sum_{k=1}^{n} \mathbb{E}[X_{k}^{2}] + \sum_{1 \leq k \neq l \leq n} \mathbb{E}[X_{k} X_{l}] - \sum_{k=1}^{n} (\mathbb{E}[X_{k}])^{2} - \sum_{1 \leq k \neq l \leq n} \mathbb{E}[X_{k}] \mathbb{E}[X_{l}]$$

$$= \sum_{k=1}^{n} (\mathbb{E}[X_{k}^{2}] - (\mathbb{E}[X_{k}])^{2})$$

$$= \sum_{k=1}^{n} \operatorname{Var}[X_{k}].$$
(1.6.10)

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* (1.6.8) or from the *law of total expectation* (1.6.9) as

$$\mathbb{E}\left[\sum_{k=1}^{Y} X_{k}\right] = \mathbb{E}\left[\mathbb{E}\left[\sum_{k=1}^{Y} X_{k} \middle| Y\right]\right]$$
$$= \sum_{n \geq 0} \mathbb{E}\left[\sum_{k=1}^{Y} X_{k} \middle| Y = n\right] \mathbb{P}(Y = n)$$
$$= \sum_{n \geq 0} \mathbb{E}\left[\sum_{k=1}^{n} X_{k} \middle| Y = n\right] \mathbb{P}(Y = n),$$
MH3512 AY19-20

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

$$\mathbb{E}\left[\sum\limits_{k=1}^{Y}X_{k}
ight] \;=\; \sum\limits_{n \geqslant 0}\mathbb{E}\left[\sum\limits_{k=1}^{n}X_{k}
ight]\mathbb{P}(Y=n) \ =\; \sum\limits_{n \geqslant 0}\mathbb{P}(Y=n)\sum\limits_{k=1}^{n}\mathbb{E}\left[X_{k}
ight].$$

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

$$\mathbb{E}\left[\prod_{k=1}^{Y} X_{k}\right] = \sum_{n \geq 0} \mathbb{E}\left[\prod_{k=1}^{n} X_{k}\right] \mathbb{P}(Y=n)$$

$$= \sum_{n \geq 0} \mathbb{P}(Y=n) \prod_{k=1}^{n} \mathbb{E}[X_{k}],$$
(1.6.11)

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

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,

$$\mathbb{E}[\phi(X)] = \int_{-\infty}^{\infty} \phi(x) f_X(x) dx, \qquad (1.6.12)$$

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

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

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

$$\mathbb{E}[\phi(X)] = rac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \phi(x) \mathrm{e}^{-(x-\mu)^2/(2\sigma^2)} dx.$$
 (1.6.13) (1.6.1

<u>Exercise</u>: 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

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

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

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

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

$$\mathbb{E}[\boldsymbol{X}] = \mathbb{E}[\mathbb{E}[\boldsymbol{X} \mid \boldsymbol{Y}]] \tag{1.6.14}$$

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

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

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

MH3512 AY19-20

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

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

1.7 Moment and Probability Generating Functions

Moment generating functions

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

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

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

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

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

$$\Phi_X(t) = \mathbb{E}\left[\mathrm{e}^{tX}
ight] = \sum\limits_{n \geqslant 0} rac{t^n}{n!} \, \mathbb{E}[X^n],$$

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

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 $-\infty_{H}t_{9_{12}}+\infty_{12}$

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

$$egin{array}{lll} G_X: [-1,1] \longrightarrow \mathbb{R} \ s \longmapsto G_X(s) \end{array}$$

defined by

$$G_X(s) := \mathbb{E}\left[s^X \mathbb{1}_{\{X < \infty\}}\right] = \sum_{n \ge 0} s^n \mathbb{P}(X = n), \qquad -1 \le s \le 1.$$
(1.7.1)

Note that the series summation in (1.7.1) is over the *finite* integers, which explains the presence of the truncating indicator $\mathbb{1}_{\{X < \infty\}}$ inside the expectation in (1.7.1). If the random variable $X : \Omega \longrightarrow \mathbb{N}$ is almost surely finite, *i.e.* $\mathbb{P}(X < \infty) = 1$, we simply have

$$G_X(s) = \mathbb{E}\left[s^X
ight] = \sum\limits_{n \geqslant 0} s^n \mathbb{P}(X=n), \qquad -1 \leqslant s \leqslant 1,$$

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

Examples

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

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

- --

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

Using the exponential series (A.1) we have

$$G_X(s) = \mathrm{e}^{\lambda(s-1)} = \mathrm{e}^{-\lambda} \sum_{n \geqslant 0} s^n rac{\lambda^n}{n!}, \qquad -1 \leqslant s \leqslant 1,$$
 $(1.7.2)$

hence by identification with (1.7.1) we find

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

i.e. \boldsymbol{X} has the Poisson distribution with parameter $\boldsymbol{\lambda}$.

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

$$G_X(s) = \sum\limits_{n \geqslant 0} s^n \mathbb{P}(X=n) = (1-p) \sum\limits_{n \geqslant 0} s^n p^n = rac{1-p}{1-ps}, \qquad -1 < s < 1,$$

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

We note that from (1.7.1) 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

$$G_X(1) = \sum\limits_{n \geqslant 0} \mathbb{P}(X=n) = \mathbb{P}(X < \infty) = \mathbb{E}\left[\mathbbm{1}_{\{X < \infty\}}
ight],$$

hence

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

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

$$G_X(0) = \mathbb{P}(X = 0).$$
 (1.7.3)

MH3512 AY19-20

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

$$G_X'(s) = \sum\limits_{n \geqslant 1} n s^{n-1} \mathbb{P}(X=n), \qquad -1 < s < 1,$$

hence we have*

$$G'_X(1^-) = \mathbb{E}[X] = \sum\limits_{k \geqslant 0} k \mathbb{P}(X=k),$$

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

iv) By computing the second derivative

$$egin{aligned} G_X''(s) &=& \sum\limits_{k \geqslant 2} k(k-1) s^{k-2} \mathbb{P}(X=k) \ &=& \sum\limits_{k \geqslant 0} k(k-1) s^{k-2} \mathbb{P}(X=k) \ &=& \sum\limits_{k \geqslant 0} k^2 s^{k-2} \mathbb{P}(X=k) - \sum\limits_{k \geqslant 0} k s^{k-2} \mathbb{P}(X=k), \qquad -1 < s < 1, \end{aligned}$$

we similarly find

$$egin{aligned} G_X''(1^-) &=& \sum\limits_{k \geqslant 0} k(k-1) \mathbb{P}(X=k) \ &=& \sum\limits_{k \geqslant 0} k^2 \mathbb{P}(X=k) - \sum\limits_{k \geqslant 0} k \mathbb{P}(X=k) \ &=& \mathbb{E}[X^2] - \mathbb{E}[X] \ &=& \mathbb{E}[X(X-1)], \end{aligned}$$

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

^{*}Here $G'_X(1^-)$ denotes the derivative on the left at the point s = 1.

More generally, using the k - th derivative of G_X we can compute the *factorial moment*

$$G_X^{(k)}(1^-) = \mathbb{E}[X(X-1)\cdots(X-k+1)], \quad k \ge 1,$$
 (1.7.4)

provided that $\mathbb{E}[|X^k|] < \infty$. In particular, we have

$$\operatorname{Var}[X] = G_X''(1^-) + G_X'(1^-)(1 - G_X'(1^-)), \qquad (1.7.5)$$

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

$$G_{X+Y}(s) = \mathbb{E}\left[s^{X+Y}\right] = \mathbb{E}\left[s^X s^Y\right] = \mathbb{E}\left[s^X\right] \mathbb{E}\left[s^Y\right] = G_X(s)G_Y(s), \qquad (1.7.6)$$

 $-1\leqslant s\leqslant 1$.

vi) The probability generating function can also be used from (1.7.1) to recover the distribution of the discrete random variable X as

$$\mathbb{P}(X=n) = \frac{1}{n!} \frac{\partial^n}{\partial s^n} G_X(s)_{|s=0}, \qquad n \in \mathbb{N},$$
(1.7.7)

extending (1.7.3) to all $n \ge 0$.

○ 59

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

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

From the generating function we also recover the mean

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

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

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

by (1.7.5).

Moments and cumulants

The <u>cumulants</u> of a random variable X are the sequence $(\kappa_n^X)_{n \in \mathbb{N}}$ of numbers defined by the logarithmic generating function

$$\log ig(\mathbb{E} \left[\mathrm{e}^{tX}
ight] ig) = \sum\limits_{n \geqslant 1} \kappa_n^X rac{t^n}{n!}, \qquad t \in \mathbb{R}$$

The n-th moment of a random variable X can be written in terms of its cumulants as

$$\mathbb{E}[X^{n}] = \sum_{k=0}^{n} \frac{1}{k!} \sum_{d_{1}+\dots+d_{k}=n} \frac{n!}{d_{1}!\dots d_{k}!} \kappa_{d_{1}}^{X} \dots \kappa_{d_{k}}^{X}$$

$$= \sum_{k=1}^{n} \sum_{B_{1},\dots,B_{k}} \kappa_{|B_{1}|}^{X} \dots \kappa_{|B_{k}|}^{X},$$
(1.7.8)

where the sum runs over the partitions B_1, \ldots, B_a of $\{1, \ldots, n\}$ with cardinality $|B_i|$ (Faà di Bruno formula). This also shows that

$$\mathbb{E}[X^n] = A_n(\kappa_1^X,\kappa_2^X,\ldots,\kappa_n^X),$$
 MH3512 AY19-20

where

$$A_n(x_1,\ldots,x_n)=n!\sum_{\substack{r_1+2r_2+\cdots+nr_n=n\r_1,\ldots,r_n\geqslant 0}}\prod_{l=1}^n \left(rac{1}{r_l!}\left(rac{x_l}{l!}
ight)^{r_l}
ight)$$

is the Bell polynomial of degree n. These relations follow from the Faà di Bruno formula, cf. *e.g.* § 2.4 and Relation (2.4.4) page 27 of [Luk70]. Indeed we have

$$\begin{split} \sum_{n \geqslant 0} \frac{t^n}{n!} \mathbb{E}[X^n] &= \mathbb{E}\left[e^{tX}\right] \\ &= \exp\left(\log\left(\mathbb{E}\left[e^{tX}\right]\right)\right) = \sum_{k \ge 0} \frac{1}{k!} \left(\log\left(\mathbb{E}\left[e^{tX}\right]\right)\right)^k \\ &= \sum_{k \ge 0} \frac{1}{k!} \left(\sum_{n \ge 1} \kappa_n^X \frac{t^n}{n!}\right)^k \\ &= \sum_{k \ge 0} \frac{1}{k!} \sum_{n \ge 1} t^n \sum_{\substack{d_1 + \dots + d_k = n \\ d_1 \ge 1, \dots, d_k \ge 1}} \frac{\kappa_{d_1}^X}{d_1!} \cdots \frac{\kappa_{d_1}^X}{d_k!} \\ &= \sum_{n \ge 0} \frac{t^n}{n!} \sum_{k = 0}^n \frac{1}{k!} \sum_{\substack{d_1 + \dots + d_k = n \\ d_1 \ge 1, \dots, d_k \ge 1}} \frac{n!}{d_1! \cdots d_k!} \frac{\kappa_{d_1}^X}{d_1!} \cdots \frac{\kappa_{d_k}^X}{d_k!}, \end{split}$$

which shows (1.7.8).

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

- 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 \boldsymbol{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

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

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

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

consider U a uniform random variable taking values in the interval [0, L] and let V = L - U.

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.

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.

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 Chapter 5.

2.1 Constrained Random Walk	. 65
2.2 Ruin Probabilities	. 68
2.3 Mean Game Duration	. 85
Exercises	. 96

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 Section 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 Amay 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 $n \in \mathbb{N}$.

MH3512 AY19-20

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 , $n \ge 0$, belongs to $\{1, 2, \ldots, S-1\}$ at the time step n, at the next step n + 1 we will have

$$X_{n+1} = \left\{egin{array}{cc} X_n+1 & ext{if Player A wins round } n+1, \ X_n-1 & ext{if Player B wins round } n+1. \end{array}
ight.$$

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.*

$$(X_n = 0) \Longrightarrow (X_{n+1} = 0)$$
 and $(X_n = S) \Longrightarrow (X_{n+1} = S)$,

or

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

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 Figure 2.1.

^{*}Daily entry levy: Singaporeans and permanent residents may have to start with $X_0 = -\$100$.



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

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 $\mathbb{P}(T = n)$, $n \ge 0$.

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

() 67

k = 1, 2, ..., S - 1, and in this case the chain is said to be <u>time homogeneous</u> 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

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

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

$$\mathbf{R}_{\mathbf{A}} = \text{``Player } \mathbf{A} \text{ loses all his capital at some time''} = \bigcup_{n \in \mathbb{N}} \{ \mathbf{X}_n = \mathbf{0} \},$$
(2.2.1)

and in computing the conditional probability

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

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

$$f_1(0) = \mathbb{P}(R_A \mid X_0 = 0) = 1,$$

$$f_1(1) = \mathbb{P}(R_A \mid X_0 = 1) = 0.$$
(2.2.3)

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

In this case we find

$$f_{2}(0) = \mathbb{P}(R_{A} \mid X_{0} = 0) = 1,$$

$$f_{2}(1) = \mathbb{P}(R_{A} \mid X_{0} = 1) = q,$$

$$f_{2}(2) = \mathbb{P}(R_{A} \mid X_{0} = 2) = 0.$$

(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

$$\begin{cases} f_{3}(0) = \mathbb{P}(R_{A} \mid X_{0} = 0) = 1, \\ f_{3}(1) = \mathbb{P}(R_{A} \mid X_{0} = 1) = q \sum_{n \geq 0} (pq)^{n} = \frac{q}{1 - pq}, \\ f_{3}(2) = \mathbb{P}(R_{A} \mid X_{0} = 2) = q^{2} \sum_{n \geq 0} (pq)^{n} = \frac{q^{2}}{1 - pq}, \\ f_{3}(3) = \mathbb{P}(R_{A} \mid X_{0} = 3) = 0. \end{cases}$$

$$(2.2.5)$$

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 $S \ge 4$, and increasingly difficult As_0S_0

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 Chapters 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

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

where we used Lemma 2.2 below on the last step.

In the case S = 3, Lemma 2.1 shows that

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

whose solution can be checked to be given by (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^{*}

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

subject to the boundary conditions

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

and

$$f_S(S) = \mathbb{P}(R_A \mid X_0 = S) = 0,$$
 (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 Equation (2.2.6) and the boundary conditions (2.2.7) and (2.2.8).

^{*}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_{3.512} AY_{19-20}

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.

\mathbf{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

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

$$= \mathbb{P}\left(\bigcup_{n \ge 0} \{X_n = 0\} \mid X_1 = k \pm 1, \ X_0 = k\right)$$

$$= \frac{\mathbb{P}\left(\left(\bigcup_{n \ge 0} \{X_n = 0\}\right) \cap \{X_1 = k \pm 1\} \cap \{X_0 = k\}\right)}{\mathbb{P}\left(\{X_1 = k \pm 1\} \cap \{X_0 = k\}\right)}$$
$$= \frac{\mathbb{P}\left(\bigcup_{n\geq 0} \left(\{X_{n}=0\} \cap \{X_{1}=k\pm 1\} \cap \{X_{0}=k\}\right)\right)}{\mathbb{P}\left(\{X_{1}=k\pm 1\} \cap \{X_{0}=k\}\right)}$$

$$= \frac{\mathbb{P}\left(\bigcup_{n\geq 2} \left(\{X_{n}=0\} \cap \{X_{1}=k\pm 1\} \cap \{X_{0}=k\}\right)\right)}{\mathbb{P}\left(\{X_{1}=k\pm 1\} \cap \{X_{0}=k\}\right)}$$

$$= \frac{\mathbb{P}\left(\left(\bigcup_{n\geq 2} \{X_{n}=0\}\right) \cap \{X_{1}=k\pm 1\} \cap \{X_{0}=k\}\right)}{\mathbb{P}\left(\{X_{1}=k\pm 1\} \cap \{X_{0}=k\}\right)}$$

$$= \frac{\mathbb{P}\left(\left(\bigcup_{n\geq 2} \{X_{n}=0\}\right) \cap \{X_{1}=k\pm 1\} \cap \{X_{1}-\tilde{X}_{0}=\pm 1\}\right)}{\mathbb{P}\left(\{X_{1}=k\pm 1\} \cap \{X_{1}-\tilde{X}_{0}=\pm 1\}\right)}$$

$$= \frac{\mathbb{P}\left(\left(\bigcup_{n\geq 2} \{X_{n}=0\}\right) \cap \{X_{1}=k\pm 1\}\right) \mathbb{P}\left(\{X_{1}-\tilde{X}_{0}=\pm 1\}\right)}{\mathbb{P}\left(\{X_{1}=k\pm 1\}\right) \mathbb{P}\left(\{X_{1}-\tilde{X}_{0}=\pm 1\}\right)}$$

$$= \mathbb{P}\left(\bigcup_{n\geq 2} \{X_{n}=0\} \mid \{X_{1}=k\pm 1\}\right)$$

$$= \mathbb{P}\left(\bigcup_{n\geq 1} \{X_{n}=0\} \mid \{X_{0}=k\pm 1\}\right)$$

$$= \mathbb{P}\left(\mathbb{Q}_{n\geq 1} \{X_{n}=0\} \mid \{X_{0}=k\pm 1\}\right)$$

$$= \mathbb{P}\left(\mathbb{Q}_{n\geq 1} \{X_{n}=0\} \mid \{X_{0}=k\pm 1\}\right)$$

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 (2.2.6) is given^{*} by

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

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

$$f_S(k) = \mathbb{P}(R_A \mid X_0 = k) = \frac{S-k}{S} = 1 - \frac{k}{S},$$
 (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 $k \ge 0$,

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

which represents the probability of hitting the origin starting from state (k), 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 Figure 2.2 the ruin probability (2.2.11) is plotted as a function of k for p = 0.45 and q = 0.55.

^{*}The techniques used to solve (2.2.6) can be found in MH1301 Discrete Mathematics.



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

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).

Standard solution method

We decide to look for a solution of (2.2.6) of the form *

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

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

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

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

^{*}Where did we get this idea? From intuition, experience, or empirically by multiple trials and errors.

$$\{a_1,a_2\} = \left(rac{1+\sqrt{1-4pq}}{2p},rac{1-\sqrt{1-4pq}}{2p}
ight) = \left(1,rac{q}{p}
ight) = (1,r),$$

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

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

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.*^{*} $r \neq 1$, and

$$f(k) = C_1 a_1^k = C_1$$
 and $f(k) = C_2 r^k$

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

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

where r = q/p and C_1 , C_2 are two constants to be determined from the boundary conditions, see also the command RSolve[f[k]=pf[k-p)f[k-1],f[k],k].

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

$$(f_S(0) = 1 = C_1 + C_2,$$

 $(f_S(S) = 0 = C_1 + C_2 r^S,$ (2.2.17)

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

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

^{*}From the Latin "id est" meaning "that is".

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

$$f_S(k) = C_1 + C_2 r^k = rac{r^k - r^S}{1 - r^S} = rac{(q/p)^k - (q/p)^S}{1 - (q/p)^S}, \ \ k = 0, 1, \dots, S$$

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

In this case, Equation (2.2.6) rewrites as

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

$$f_S(k) = f(k) + g(k) = C_1 + C_2 k,$$
 (2.2.19)

see also the command RSolve[f[k]=(1/2)f[k+1]+(1/2)f[k-1],f[k],k]. From (2.2.7), (2.2.8) and (2.2.19) we have

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

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

$$C_1 = 1$$
 and $C_2 = -1/S$, MH3512 AY19-20

which yields the quite intuitive solution

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

Direct solution method

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

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

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

$$p(f_S(k+1) - f_S(k)) - q(f_S(k) - f_S(k-1)) = 0, \qquad (2.2.22)$$

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

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

hence for k = 1 we have

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

and for k = 2 we find

$$f_S(3)-f_S(2)=rac{q}{p}(f_S(2)-f_S(1))=\left(rac{q}{p}
ight)^2(f_S(1)-f_S(0)).$$

Following by induction on $k \ge 2$, we can show that

$$f_S(k+1) - f_S(k) = \left(rac{q}{p}
ight)^k (f_S(1) - f_S(0)),$$

578

(2.2.23) MH3512 AY19-20 $k = 0, 1, \dots, S - 1$. 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

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

 $n = 1, 2, \ldots, S - 1$. 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

$$f_{S}(n) = f_{S}(0) + (f_{S}(1) - f_{S}(0)) \sum_{k=0}^{n-1} r^{k}$$

= $f_{S}(0) + \frac{1 - r^{n}}{1 - r} (f_{S}(1) - f_{S}(0)),$ (2.2.25)

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

Conditions (2.2.7) and (2.2.8) show that

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

hence

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

and combining this relation with (2.2.25) yields

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

MH3512 AY19-20

 $n = 0, 1, \dots, S$, 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

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

and (2.2.24) reads

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

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

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

and

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

 $n = 0, 1, \ldots, S$, 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 a continuous Laplace equation as

$$egin{array}{rcl} rac{\partial^2 f}{\partial x^2}(x) &\simeq & rac{\partial f}{\partial x}(x+1/2) - rac{\partial f}{\partial x}(x-1/2) \ &\simeq & (f(x+1)-f(x)-(f(x)-f(x-1))) \ &= & 0, \qquad x \in \mathbb{R}, \end{array}$$

which admits a solution of the form

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

S 80

MH3512 AY19-20

(2.2.26)

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

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

where

 $R_B:=$ "Player B loses all his capital at some time" $= igcup_{n\in\mathbb{N}} \{X_n=S\}.$

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

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

see also Exercise 2.3 below.

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

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

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

By (2.2.11) and (2.2.27) we can check that*

 $\mathbb{P}(R_A\cup R_B\mid X_0=k) \;=\; \mathbb{P}(R_A\mid X_0=k)+\mathbb{P}(R_B\mid X_0=k)$

^{*} Exercise: check by hand computation that the equality to 1 holds as stated.

$$= \frac{(q/p)^{k} - (q/p)^{S}}{1 - (q/p)^{S}} + \frac{1 - (q/p)^{k}}{1 - (q/p)^{S}}$$

$$= 1, \quad k = 0, 1, \dots, S, \qquad (2.2.29)$$

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

$$\mathbb{P}(R^c_A\cap R^c_B\mid X_0=k)=0, \qquad k=0,1,\ldots,S.$$

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

$$A_n := igcap_{k=1}^n \{X_{2k-1} = 1 ext{ and } X_{2k} = 2\}, \qquad n \geqslant 1,$$

the sequence $(A_n)_{n \ge 1}$ is non-increasing, hence

$$\mathbb{P}\left(\bigcap_{n \ge 1} \left\{ X_{2n-1} = 1 \text{ and } X_{2n} = 2 \right\} \middle| X_0 = 2 \right)$$

$$= \mathbb{P}\left(\bigcap_{n \ge 1} \bigcap_{k=1}^n \left\{ X_{2k-1} = 1 \text{ and } X_{2k} = 2 \right\} \middle| X_0 = 2 \right)$$

$$= \lim_{n \to \infty} \mathbb{P}\left(\bigcap_{k=1}^n \left\{ X_{2k-1} = 1 \text{ and } X_{2k} = 2 \right\} \middle| X_0 = 2 \right)$$

$$= p \lim_{n \to \infty} (pq)^n = 0,$$

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

MH3512 AY19-20

The expected terminal wealth of Player A starting from state (\mathbf{k}) is given by



Figure 2.3: Expected terminal wealth of Player A starting from state k = 1.

When the number S of states becomes large, (2.2.27) also shows that for all $k \ge 0$ we have

$$\lim_{S o\infty} \mathbb{P}(R_B \mid X_0 = k) = egin{cases} 0 & ext{if } p \leqslant q, \ 1 - \left(rac{q}{p}
ight)^k & ext{if } p > q, \end{cases}$$

which represents the complement of the probability (2.2.13) of hitting the origin starting from state (k), and is the probability that the process $(X_n)_{n \in \mathbb{N}}$ "escapes to infinity". MH3512 AY19-20





Figure 2.4: Ruin probability as a function of $p \in [0, 1]$ 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.* 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

^{*}In this game, the payout is 2 and the payout percentage is 2p.

2.3 Mean Game Duration

Let now

$$T_{0,S} = \inf\{n \geqslant 0 \; : \; X_n = 0 ext{ or } X_n = S\}$$

denote the time^{*} until any of the states \bigcirc or \bigcirc is 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 $n \ge 0$ such that $X_n = 0$ or $X_n = S$, see Figure 2.5.



Figure 2.5: Sample paths of a gambling process $(X_n)_{n \in \mathbb{N}}$.

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

Note that by (2.2.29) we have

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

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 $X_0 = k \in \{0, 1, \dots, S\}$. Clearly, we have the boundary conditions

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

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

i) <u>**S**</u> = 2.

We have

$$T_{0,2} = \left\{egin{array}{ll} 0 & ext{if } X_0 = 0, \ 1 & ext{if } X_0 = 1, \ 0 & ext{if } X_0 = 2, \end{array}
ight.$$

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) 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 $X_0 = 4$, 87

$$\left\{egin{array}{ll} \mathbb{P}(T_{0,3}=2k\mid X_0=1)=p^2(pq)^{k-1}, & k\geqslant 1, \ \mathbb{P}(T_{0,3}=2k+1\mid X_0=1)=q(pq)^k, & k\geqslant 0, \end{array}
ight.$$

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 $X_0 = 2$,

$$iggl\{ egin{array}{ll} \mathbb{P}(T_{0,3}=2k \mid X_0=2) = q^2(pq)^{k-1}, & k \geqslant 1, \ \mathbb{P}(T_{0,3}=2k+1 \mid X_0=2) = p(pq)^k, & k \geqslant 0, \end{array}$$

whereas $T_{0,3} = 0$ whenever $X_0 = 0$ or $X_0 = 3$.

As a consequence, we can directly compute

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

(2.3.2)MH3512 AY19-20 where we applied (A.4), and by exchanging p and q we get

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

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

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

Lemma 2.4 For all k = 1, 2, ..., 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, ..., 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

$$\begin{split} \mathbb{E}[T_{0,S} \mid X_0 = k] &= \frac{1}{\mathbb{P}(X_0 = k)} \mathbb{E}\left[T_{0,S} \mathbb{1}_{\{X_0 = k\}}\right] \\ &= \frac{1}{\mathbb{P}(X_0 = k)} \left(\mathbb{E}\left[T_{0,S} \mathbb{1}_{\{X_1 = k+1, X_0 = k\}}\right] + \mathbb{E}\left[T_{0,S} \mathbb{1}_{\{X_1 = k-1, X_0 = k\}}\right] \right) \\ &= \frac{\mathbb{P}(X_1 = k+1 \text{ and } X_0 = k)}{\mathbb{P}(X_0 = k)} \mathbb{E}[T_{0,S} \mid X_1 = k+1, X_0 = k] \\ &+ \frac{\mathbb{P}(X_1 = k-1 \text{ and } X_0 = k)}{\mathbb{P}(X_0 = k)} \mathbb{E}[T_{0,S} \mid X_1 = k-1, X_0 = k] \end{split}$$

MH3512 AY19-20

 $\mathbf{38}$

$$= \mathbb{P}(X_{1} = k + 1 \mid X_{0} = k) \mathbb{E}[T_{0,S} \mid X_{1} = k + 1, X_{0} = k] + \mathbb{P}(X_{1} = k - 1 \mid X_{0} = k) \mathbb{E}[T_{0,S} \mid X_{1} = k - 1, X_{0} = k] = p \mathbb{E}[T_{0,S} \mid X_{1} = k + 1, X_{0} = k] + q \mathbb{E}[T_{0,S} \mid X_{1} = k - 1, X_{0} = k]$$

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

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

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

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

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

$$(2.3.4)$$

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

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

whose solution can be checked to be given by (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

$$h_S(k) = p(1+h_S(k+1)) + q(1+h_S(k-1)) = 1 + ph_S(k+1) + qh_S(k-1), ext{MH3512 AY19-20}$$

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

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

$$(2.3.6)$$

$$(2.3.6)$$

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

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

or as the *difference equation*

$$p(h(k+1) - h(k)) - q(h(k) - h(k-1)) = -1, \qquad k = 1, 2, \dots, S - 1,$$
(2.3.7)

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

The equation

$$p(f(k+1) - f(k)) - q(f(k) - f(k-1)) = 0, \qquad k = 1, 2, \dots, S - 1,$$
(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 <u>homogeneous solution</u> of (2.3.8) is of the form $C_1 + C_2 r^k$. Next, searching for a <u>particular solution</u> of (2.3.7) of the form $k \mapsto Ck$ shows that C has to be equal to C = 1/(q-p). Therefore, when $r = q/p \neq 1$, the general solution of

^{*}The techniques used to solve (2.3.6) can be found in MH1301 Discrete Mathematics.

(2.3.7) has the form

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

ॳ91

see also the command RSolve[f[k]=1+pf[k+1]+(1-p)f[k-1],f[k],k]. From the boundary conditions (2.3.1a) and (2.3.1b) and from (2.3.9) we have

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

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

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

hence from (2.3.9) we get

$$h_{S}(k) = \mathbb{E}[T_{0,S} \mid X_{0} = k] = \frac{1}{q-p} \left(k - S \frac{1 - (q/p)^{k}}{1 - (q/p)^{S}} \right)$$

$$= \frac{1}{q-p} \left(k - S \mathbb{P}(R_{B} \mid X_{0} = k) \right), \quad k = 0, 1, 2, \dots, S,$$
 (2.3.11)

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. In particular, we note that

$$\mathbb{E}[T_{0,L} \mid S_0=k] < +\infty, \qquad k=0,1,2,\dots,L$$

When p = 1, *i.e.* r = 0, we can check easily that *e.g.* $h_S(k) = S - k$, $k = 0, 1, 2, \dots, S$. On the other hand, when the number S_0

of states becomes large, we find that for all $k \ge 1$,

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

with $h_{\infty}(0) = 0$, cf. also the symmetric case treated in (2.3.17) below when p = q = 1/2. In particular, for $k \ge 1$ we have

$$\mathbb{P}(T_0 < \infty \mid X_0 = k) < 1 \text{ and } \mathbb{E}[T_0 = \infty \mid X_0 = k] = \infty, \text{ for } p > q,$$

 $\mathbb{P}(T_0 < \infty \mid X_0 = k) = 1 \text{ and } \mathbb{E}[T_0 = \infty \mid X_0 = k] = \infty, \text{ for } p = q = \frac{1}{2},$
 $\mathbb{P}(T_0 < \infty \mid X_0 = k) = 1 \text{ and } \mathbb{E}[T_0 = \infty \mid X_0 = k] < \infty, \text{ for } p < q.$

We note in particular that the mean game duration $\mathbb{E}[T_0 = \infty \mid X_0 = k]$ is infinite in the fair game case p = q = 1/2. When $r = q/p \leq 1$, this yields an example of a random variable T_0 which is (almost surely^{*}) finite, while its expectation is infinite.[†] This situation is similar to that of the St. Petersburg paradox as in (1.6.5). Similarly, one can find sequences $(X_n)_{n \in \mathbb{N}}$ of random variables such that $X_n \to 0$ with probability one, while $\mathbb{E}[X_n] \to \infty$ as n tends to infinity.

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

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

and

$$\mathbb{E}[T_{0,3} \mid X_0 = 2] = \frac{1}{q-p} \left(2 - 3 \frac{1 - (q/p)^2}{1 - (q/p)^3} \right) = \frac{1+q}{1-pq},$$
(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), see for example *here*. In Figure 2.6 below the mean game duration (2.3.11) is plotted as a function of k for p = 0.45.

^{*&}quot;Almost surely" means "with probability 1".

 $^{^\}dagger \mathrm{Recall}$ that an infinite set of finite data values may have an infinite average.

[‡]This point is left as exercise.



Figure 2.6: Mean game duration $h_{20}(k)$ as a function of $X_0 = k \in [0, 20]$ 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).

Since r = 1 we see that $k \mapsto Ck$ can no longer be a particular solution of (2.3.7). However we can search for a particular solution of the form $k \mapsto Ck^2$, 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

$$h_S(k) = C_1 + C_2 k - k^2, \qquad k = 0, 1, 2, \dots, S,$$
(2.3.15)

see also the command RSolve[f[k]=1+(1/2)f[k+1]+(1/2)f[k-1],f[k],k]. From the boundary conditions (2.3.1a) and (2.3.1b) and from (2.3.15) we have

$$h_S(0) = 0 = C_1,$$
 $(2.3.16)$
 $h_S(S) = 0 = C_1 + C_2 S - S^2,$
MH3512 AY19-20

♦93

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

$$C_1 = 0$$
 and $C_2 = S$,

hence from (2.3.15) we get

$$h_S(k) = \mathbb{E}[T_{0,S} \mid X_0 = k] = k(S-k), \quad k = 0, 1, 2, \dots, S,$$

(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. In particular, we note that

$$\mathbb{E}[T_{0,S}\mid S_0=k]<+\infty, \qquad k=0,1,2,\ldots,S.$$

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 Poisson equation

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

which has for solution

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

Note that (2.3.17) can also be recovered from (2.3.11) by letting p go to 1/2. In the next Figure 2.7 the expected game duration (2.3.11) is plotted as a function of p for S = 20 and k = 10.



 $\bigcirc 95$

Figure 2.7: 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 2 we have noticed an interesting connection between analysis and probability. That is, a probabilistic quantity such as $k \mapsto \mathbb{P}(R_A \mid X_0 = k)$ or $k \mapsto \mathbb{E}[T_{0,S} \mid X_0 = k]$ 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 AV19-20

Exercises

Exercise 2.1 We consider a gambling problem with the possibility of a draw,^{*} *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 \boldsymbol{A} , and let

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

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, ..., 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 (a) and (b) and probabilities $p \neq q$ of moving by ± 1 , the probability of hitting state (a) before hitting state (b) after starting

^{*}Also called "lazy random walk"

from state $Z_0 = k \in \{a, a+1, \dots, b-1, b\}$ is given by

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

597

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 (S) 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 $m \in \{1, 2, ..., S 1\}$ after starting from $X_0 = k \in \{m + 1, ..., S\}$.
- b) Using Relation (2.3.18), give the probability of coming back in finite time to the given state $m \in \{1, 2, ..., S-1\}$ after starting from $X_0 = k \in \{0, 1, ..., m-1\}$.
- c) Using first step analysis, give the probability of coming back to state (m) 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 (a) before hitting state (b) after starting from state $Z_0 = k$ is given by

$$\frac{b-k}{b-a}, \qquad k=a, a+1, \dots, b-1, b.$$
 (2.3.19)

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

$$au_k:=\inf\{n\geqslant 0 \ : X_n=k\}$$

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

$$p_k := \mathbb{P}(au_{k+1} < au_0 \mid X_0 = k), \qquad k = 0, 1, \dots, S-1,$$
 MH3512 AY19-20

denote the probability of hitting state (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

$$p_k = p + q p_{k-1} p_k, \qquad k = 1, 2, \dots, S - 1,$$
(2.3.20)

i.e.

$$p_k = rac{p}{1 - q p_{k-1}}, \qquad k = 1, 2, \dots, S-1.$$

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

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

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

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

$$= \begin{bmatrix} 0 \\ q \\ p \\ 0 \\ 1 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

MH3512 AY19-20

♦98

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

In order to be ruined, Player A has to visit state \bigcirc <u>twice</u>. 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

$$\left\{egin{array}{l} \mathbb{P}(X_{n+1}=k+1\mid X_n=k)=p, \quad k=0,1,\ldots,S-1, \ \mathbb{P}(X_{n+1}=k-1\mid X_n=k)=q, \quad k=1,2,\ldots,S-1, \end{array}
ight.$$

with

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

MH3512 AY19-20

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

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

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 $(0, 1, \ldots, S)$.

b) Obtain $\mathbb{P}(W \mid X_0 = k)$ for all $k = 0, 1, \dots, S$ 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

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

denote the first hitting time of S by the process $(X_n)_{n \ge 0}$. 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 $k = 0, 1, \dots, S$ 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 associated *homogeneous equation* for k = 1, 2, ..., S - 1 using the results of Section 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

$$T_0=\inf\{n\geqslant 0\;:\;X_n=0\}$$

denote the first hitting time of 0 by the process $(X_n)_{n \ge 0}$. Using the results of Section 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 Chapter 2, $\{T_S < T_0\}$ denotes the event "Player A wins the game".

f) Explain why the equality

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

holds for $k \in \{0, 1, \dots, S-1\}$ (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

$$\mathbb{P}(X_1 = k+1 \mid X_0 = k \text{ and } T_S < T_0) = p \frac{\mathbb{P}(T_S < T_0 \mid X_0 = k+1)}{\mathbb{P}(T_S < T_0 \mid X_0 = k)} = p \frac{p_{k+1}}{p_k},$$
(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_{20}^2

No particular difficulty here, the proof should be a straightforward application of the definition of conditional probabilities.

h) Compute the probability

 $\mathbb{P}(X_1 = k - 1 \mid X_0 = k \text{ and } T_0 < T_S), \quad k = 1, 2, \dots, S,$

of a downward step given that state (1) 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 $\underline{p = 1/2}$ 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 Section 2.3.

Problem 2.8 Let $S \ge 1$. We consider a discrete-time process $(X_n)_{n\ge 0}$ that models the wealth of a gambler within $\{0, 1, \ldots, S\}$, with the transition probabilities

$$\mathbb{P}(X_{n+1}=k+2 \mid X_n=k) = p, \ \ \mathbb{P}(X_{n+1}=k-1 \mid X_n=k) = 2p,$$

and

$$\mathbb{P}(X_{n+1}=k\mid X_n=k)=r, \ \ k\in\mathbb{Z},$$

for all $n \in \mathbb{N} = \{0, 1, 2, \ldots\}$, where p > 0, $r \ge 0$, and 3p + r = 1. We let

 $au:=\inf\{n\geqslant 0\;:\;X_n\leqslant 0 ext{ or }X_n\geqslant S\}.$ MH3512 AY19-20

a) Consider the probability

$$g(k):=\mathbb{P}(X_{ au} \geqslant S \mid X_0=k)$$

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, ..., 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

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

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, ..., S 1, and its boundary conditions.
- g) Find a particular solution of the equation of Question (e).
- h) Solve the equation of Question (c).

Hint: the general solution of the equation is the sum of a particular solution and a solution of the associated *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 ([AH12]). Consider a gambling process $(X_n)_{n \ge 0}$ on the state space $\mathbb{S} = \{0, 1, \dots, S\}$, with transition probabilities

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

 $k = 1, 2, \dots, S - 1$, with p + q = 1. Let

$$au := \inf\{n \geqslant 0 \; : \; X_n = 0 ext{ or } X_n = S\}$$
 MH3512 AY19-20

denote the time until the process hits either state 0 or state S, and consider the second moment

$$h(k):=\mathbb{E}ig[au^2\mid X_0=kig],$$

of au after starting from $k = 0, 1, 2, \dots, S$.

- 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 $\mathbb{E}[\tau \mid X_0 = k-1]$, $k = 1, 2, \dots, S-1$.
- c) From now on we take p = q = 1/2. Recall that in this case we have

$$\mathbb{E}[au \mid X_0=k]=(S-k)k, \qquad k=0,1,\dots,S.$$

Show that the function h(k) satisfies the finite difference equation

$$h(k) = -1 + 2(S-k)k + \frac{1}{2}h(k+1) + \frac{1}{2}h(k-1), \qquad k = 1, 2, \dots, S-1.$$
(2.3.24)

d) Knowing that

$$k\longmapsto rac{2}{3}k^2-rac{2S}{3}k^3+rac{k^4}{3}$$

is a particular solution of the equation (2.3.24) of Question (c), and that the solution of the homogeneous equation

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

takes the form

$$f(k) = C_1 + C_2 k_2$$

compute the value of the expectation h(k) solution of (2.3.24) for all $k = 0, 1, \ldots, S$.

e) Compute the variance

$$v(k) = \mathbb{E}\left[au^2 \mid X_0 = k
ight] - \left(\mathbb{E}[au \mid X_0 = k]
ight)^2$$
 MH3512 AY19-20

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.

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	106
3.2	Mean and Variance	107
3.3	Distribution	108
3.4	First Return to Zero	110
Exe	Exercises	

3.1 Unrestricted Random Walk

The simple unrestricted random walk $(S_n)_{n \ge 0}$, also called the *Bernoulli* random walk, is defined by $S_0 = 0$ and

$$S_n = \sum_{k=1}^n X_k = X_1 + \dots + X_n, \qquad n \ge 1,$$

where the random walk *increments* $(X_k)_{k \ge 1}$ 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

$$\left\{egin{array}{l} \mathbb{P}(X_k=+1)=p,\ \mathbb{P}(X_k=-1)=q, \qquad k\geqslant 1, \end{array}
ight.$$

with p + q = 1.

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

$$egin{aligned} & \mathrm{Var}[X_n] \;=\; \mathbb{E}[X_n^2] - (\mathbb{E}[X_n])^2 \ & = \; 1 imes q + 1 imes p - (2p-1)^2 \ & = \; 4p(1-p) = 4pq. \end{aligned}$$

As a consequence, we find that

$$\mathbb{E}[S_n \mid S_0 = 0] = \mathbb{E}\left[\sum\limits_{k=1}^n X_k
ight] = \sum\limits_{k=1}^n \mathbb{E}\left[X_k
ight] = n(2p-1) = n(p-q),$$

and the variance can be computed by (1.6.10) as

$$\operatorname{Var}[S_n \mid S_0 = 0] = \operatorname{Var}\left[\sum_{k=1}^n X_k
ight] = \sum_{k=1}^n \operatorname{Var}[X_k] = 4npq,$$

where we used (1.6.10).

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

$$\mathbb{P}(S_{2n} = 2k + 1 \mid S_0 = 0) = 0, \qquad k \in \mathbb{Z}, \ n \in \mathbb{N},$$

$$\mathbb{P}(S_{2n+1} = 2k \mid S_0 = 0) = 0, \qquad k \in \mathbb{Z}, \ n \in \mathbb{N},$$

$$(3.3.1)$$

and

$$\mathbb{P}(S_n = k \mid S_0 = 0) = 0, \quad \text{for } k < -n \text{ or } k > n,$$
(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 (2k) in 2n steps. We find that this number of paths is

$$egin{pmatrix} 2n \ n+k \end{pmatrix} = egin{pmatrix} 2n \ n-k \end{pmatrix}$$

which represents the number of ways to arrange n + k upwards steps (or n - k downwards steps) within 2n time steps.

Hence we have
***** 109

$$\mathbb{P}(S_{2n} = 2k \mid S_0 = 0) = \binom{2n}{n+k} p^{n+k} q^{n-k}, \quad -n \leq k \leq n,$$
(3.3.3)

in addition to (3.3.1) and (3.3.2). In Figure 3.1 we enumerate the $120 = \binom{10}{7} = \binom{10}{3}$ possible paths corresponding to n = 5 and k = 2.

Figure 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

$$\mathbb{P}(S_{2n+1} = 2k+1 \mid S_0 = 0) = \binom{2n+1}{n+k+1} p^{n+k+1} q^{n-k}, \quad -n \leq k \leq n,$$
(3.3.4)

^{*}Animated figure (works in Acrobat Reader).

i.e. $(2n + 1 + S_{2n+1})/2$ is a binomial random variable with parameter (2n + 1, p), and

$$\mathbb{P}\left(rac{2n+1+S_{2n+1}}{2}=k \ \Big| \ S_0=0
ight) \ = \ \mathbb{P}\left(S_{2n+1}=2k-2n-1 \ | \ S_0=0
ight) \ = \ inom{(2n+1)}{k}p^kq^{2n+1-k},$$

 $k=0,1,\ldots,2n+1.$

ii) Show that $n + S_{2n}/2$ is a binomial^{*} random variable with parameter (2n, p), *i.e.*, show that

$$\mathbb{P}\left(n+rac{S_{2n}}{2}=k \bigm| S_0=0
ight) \;=\; \mathbb{P}\left(S_{2n}=2k-2n \mid S_0=0
ight) \ =\; {\binom{2n}{k}}p^kq^{2n-k}, \qquad k=0,1,\dots,2n$$

3.4 First Return to Zero

Let

$$T_0^r:=\inf\{n\geqslant 1\;:\;S_n=0\}$$

denote the time of first return to (0) of the random walk started at (0), with the convention $\inf \emptyset = +\infty$.[†] 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 Figure 3.2.

^{*}Note that S_{2n} is always an even number after we start from $S_0 = 0$.

[†]Recall that the notation "inf" stands for "infimum", meaning the smallest $n \ge 0$ such that $S_n = 0$, with $T_0^r = +\infty$ if no such $n \ge 0$ exists.



Figure 3.2: Sample path of the random walk $(S_n)_{n \in \mathbb{N}}$.

We are interested in computing the distribution

$$g(n)=\mathbb{P}(T_0^r=n\mid S_0=0), \qquad n\geqslant 1,$$

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

$$\mathbb{P}(T_0^r = 1 \mid S_0 = 0) = 0, \quad \mathbb{P}(T_0^r = 2 \mid S_0 = 0) = 2pq, \tag{3.4.1}$$

and

$$\mathbb{P}(T_0^r = 4 \mid S_0 = 0) = 2p^2 q^2, \tag{3.4.2}$$

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 $n \ge 3$.

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 3.4 below.

Lemma 3.1

(Convolution equation). The function

$$g: \{1,2,3,\ldots\} \longrightarrow [0,1] \ n \longmapsto g(n)$$

defined by

$$g(n):=\mathbb{P}(T_0^r=n\mid S_0=0), \qquad n\geqslant 1,$$

satisfies the convolution equation

$$h(n) = \sum_{k=0}^{n-2} g(n-k)h(k), \qquad n \ge 1,$$
(3.4.3)

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

$$h(2n) = \binom{2n}{n} p^n q^n, \quad \text{and} \quad h(2n+1) = 0, \qquad n \in \mathbb{N}.$$
(3.4.4)

Proof

We first partition the event $\{S_n = 0\}$ into

$$\{S_n=0\}=igcup_{k=0}^{n-2}ig\{S_k=0,\;S_{k+1}
eq 0,\ldots,S_{n-1}
eq 0,\;S_n=0ig\},\qquad n\geqslant 1,$$

according to all possible times k = 0, 1, ..., n - 2 of *last* return to state (0) before time n, with $\{S_1 = 0\} = \emptyset$ since we are starting from $S_0 = 0$, see Figure 3.3.





Then we have

 $h(n):=\mathbb{P}(S_n=0\mid S_0=0)$

$$= \sum_{k=0}^{n-2} \mathbb{P}(S_k = 0, \ S_{k+1} \neq 0, \dots, S_{n-1} \neq 0, \ S_n = 0 \mid S_0 = 0)$$

$$= \sum_{k=0}^{n-2} \mathbb{P}(S_{k+1} \neq 0, \dots, S_{n-1} \neq 0, \ S_n = 0 \mid S_k = 0, \ S_0 = 0) \mathbb{P}(S_k = 0 \mid S_0 = 0)$$

$$= \sum_{k=0}^{n-2} \mathbb{P}(S_{k+1} \neq 0, \dots, S_{n-1} \neq 0, \ S_n = 0 \mid S_k = 0) \mathbb{P}(S_k = 0 \mid S_0 = 0)$$

$$= \sum_{k=0}^{n-2} \mathbb{P}(S_1 \neq 0, \dots, S_{n-k-1} \neq 0, \ S_{n-k} = 0 \mid S_0 = 0) \mathbb{P}(S_k = 0 \mid S_0 = 0)$$

$$= \sum_{k=0}^{n-2} \mathbb{P}(T_0^r = n - k \mid S_0 = 0) \mathbb{P}(S_k = 0 \mid S_0 = 0)$$

$$= \sum_{k=0}^{n-2} \mathbb{P}(T_0^r = n - k \mid S_0 = 0) \mathbb{P}(S_k = 0 \mid S_0 = 0)$$

$$= \sum_{k=0}^{n-2} h(k)g(n-k), \quad n \ge 1,$$

$$(3.4.7)$$

$$where from (3.4.5) to (3.4.6) we applied a shift of k steps in time, from time k + 1 to time 1.$$

We now need to solve the convolution equation (3.4.3) for $g(n) = \mathbb{P}(T_0^r = n \mid S_0 = 0)$, $n \ge 1$, knowing that g(1) = 0. For this we will derive a simple equation for the probability generating function

$$egin{array}{lll} G_{T_0^r} & \colon & [-1,1] \longrightarrow \mathbb{R} \ & s \longmapsto G_{T_0^r}(s) \end{array}$$

of the random variable T_0^r , defined by

$$G_{T_0^r}(s) := \mathbb{E}\left[s^{T_0^r}\mathbb{1}_{\{T_0^r < \infty\}} \left| S_0 = 0
ight] = \sum_{n \geqslant 0} s^n \mathbb{P}(T_0^r = n \mid S_0 = 0) = \sum_{n \geqslant 0} s^n g(n),$$

 $-1 \leq s \leq 1$, cf. (1.7.1).

Recall that the knowledge of $G_{T_0^r}(s)$ provides certain information on the distribution of T_0^r , such as the probability

$$\mathbb{P}(T_0^r < \infty \mid S_0 = 0) = \mathbb{E}\left[\mathbb{1}_{\{T_0^r < \infty\}} \left| S_0 = 0
ight] = G_{T_0^r}(1)$$

and the expectation

$$\mathbb{E}\left[T_{0}^{r}\mathbb{1}_{\{T_{0}^{r}<\infty\}}ig|S_{0}=0
ight]=\sum\limits_{n\geqslant1}n\mathbb{P}(T_{0}^{r}=n\mid S_{0}=0)=G_{T_{0}^{r}}^{\prime}(1^{-}).$$

In Lemma 3.4 below we will compute $G_{T_0^r}(s)$ for all $s \in [-1, 1]$. First, let the function

$$egin{array}{lll} H & \colon \mathbb{R} \longrightarrow \mathbb{R} \ s \longmapsto H(s) \end{array}$$

be defined by

$$H(s):=\sum\limits_{k\geqslant 0}h(k)s^k=\sum\limits_{k\geqslant 0}s^k\mathbb{P}(S_k=0\mid S_0=0),\qquad -1\leqslant s\leqslant 1.$$

In the following lemma we show that the function H(s) can be computed in closed form.

Proposition 3.2

We have

$$H(s) = (1 - 4pqs^2)^{-1/2}, \qquad |s| < rac{1}{2\sqrt{pq}}$$

\mathbf{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

$$H(s) = \sum_{k \ge 0} s^{k} \mathbb{P}(S_{k} = 0 \mid S_{0} = 0)$$

$$= \sum_{k \ge 0} s^{2k} \mathbb{P}(S_{2k} = 0 \mid S_{0} = 0) = \sum_{k \ge 0} s^{2k} \binom{2k}{k} p^{k} q^{k}$$

$$= \sum_{k \ge 0} (pq)^{k} s^{2k} \frac{(2k)(2k-1)(2k-2)(2k-3) \times \dots \times 4 \times 3 \times 2 \times 1)}{(k(k-1) \times \dots \times 2 \times 1)^{2}}$$
(3.4.8)

MH3512 AY19-20

() 115

$$= \sum_{k \ge 0} (4pq)^k s^{2k} \frac{k(k-1/2)(k-2/2)(k-3/2) \times \dots \times (4/2) \times (3/2) \times (2/2) \times (1/2)}{(k(k-1) \times \dots \times 2 \times 1)^2}$$

$$= \sum_{k \ge 0} (4pq)^k s^{2k} \frac{(k-1/2)(k-3/2) \times \dots \times (3/2) \times (1/2)}{k(k-1) \times \dots \times 2 \times 1}$$

$$= \sum_{k \ge 0} (-1)^k (4pq)^k s^{2k} \frac{(-1/2 - (k-1))(3/2 - k) \times \dots \times (-3/2) \times (-1/2)}{k(k-1) \times \dots \times 2 \times 1}$$

$$= \sum_{k \ge 0} (-4pqs^2)^k \frac{(-1/2) \times (-3/2) \times \dots \times (3/2 - k)(-1/2 - (k-1))}{k!}$$

$$= (1 - 4pqs^2)^{-1/2},$$

 $|4pqs^2| < 1$,^{*a*} see for example *here*.

^{*a*}We used the formula
$$(1+x)^{\alpha} = \sum_{k \ge 0} \frac{x^k}{k!} \alpha(\alpha-1) \times \cdots \times (\alpha-(k-1))$$
, cf. Relation (A.8).

Remark 3.3

We note that, taking s = 1, by (1.6.1) we have

$$egin{array}{rcl} H(1) &=& \sum\limits_{k \geqslant 0} \mathbb{P}(S_k = 0 \mid S_0 = 0) \ &=& \sum\limits_{k \geqslant 0} \mathbb{E}[\mathbbm{1}_{\{S_k = 0\}} \mid S_0 = 0] \ &=& \mathbb{E}\left[\sum\limits_{k \geqslant 0} \mathbbm{1}_{\{S_k = 0\}} \mid S_0 = 0
ight], \end{array}$$

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 3.4 by deriving and solving an Equation (3.4.13) for $G_{T_0^r}(s)$. This method has some similarities with the *z*-transform method used in electrical engineering.

Lemma 3.4

The probability generating function $G_{T_0^r}$ of the first return time T_0^r to \bigcirc is given by

$$G_{T_0^r}(s) = 1 - \frac{1}{H(s)} = 1 - \sqrt{1 - 4pqs^2}, \quad 4pqs^2 < 1.$$
 (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,

$$G_{T_0^r}(s)H(s) = \left(\sum_{n \ge 1} s^n g(n)\right) \left(\sum_{k \ge 0} s^k h(k)\right)$$

= $\sum_{n \ge 2} \sum_{k \ge 0} s^{n+k} g(n)h(k) = \sum_{k \ge 0} \sum_{n \ge 2} s^{n+k} g(n)h(k)$ (3.4.10)

$$= \sum_{l \ge 2} s^l \sum_{k=0}^{l-2} g(l-k)h(k)$$
(3.4.11)

$$= \sum_{l \ge 1} s^l h(l) \tag{3.4.12}$$

$$egin{array}{ll} &=& \sum\limits_{l \geqslant 1} s^l \mathbb{P}(S_l=0 \mid S_0=0) \ &=& -1 + \sum\limits_{l \geqslant 0} s^l \mathbb{P}(S_l=0 \mid S_0=0) = H(s) - 1, \end{array}$$

where from line (3.4.10) to line (3.4.11) we have applied the change of variable $(k, n) \mapsto (k, l)$ 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

0117

$$G_{T_0^r}(s)H(s) = H(s) - 1, \qquad 4pqs^2 < 1.$$

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.*

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

$$\begin{split} \mathbb{P}(T_0^r < \infty \mid S_0 = 0) &= \mathbb{E}\left[\mathbbm{1}_{\{T_0^r < \infty\}} \mid S_0 = 0\right] = \mathbb{E}\left[\mathbbm{1}^{T_0^r} \mathbbm{1}_{\{T_0^r < \infty\}} \mid S_0 = 0\right] \\ &= G_{T_0^r}(1) = 1 - \sqrt{1 - 4pq} \\ &= 1 - |2p - 1| = 1 - |p - q| = \begin{cases} 2q, & p \ge 1/2, \\ 2p, & p \le 1/2, \\ &= 2\min(p, q), \end{cases} \end{split}$$
(3.4.14)

(3.4.13)

hence

$$\mathbb{P}(T_0^r = \infty \mid S_0 = 0) = |2p - 1| = |p - q|.$$
(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 $k \ge 1$

^{*&}quot;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 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. MH3512 AY19-20

is given by

$$\mathbb{P}(T_0^r < \infty \mid S_0 = k) = \min\left(1, \left(\frac{q}{p}\right)^k\right), \qquad k \ge 1,$$
(3.4.16)

i.e.

$$\mathbb{P}(T_0^r=\infty\mid S_0=k)=\max\left(0,1-\left(rac{q}{p}
ight)^k
ight),\qquad k\geqslant 1,$$

cf. also Exercise 3.2-(c) and Exercise 3.4-(c).

In the non-symmetric case $p \neq q$, Relation (3.4.14) shows that

$$\mathbb{P}(T_0^r < \infty \mid S_0 = 0) < 1$$
 and $\mathbb{P}(T_0^r = \infty \mid S_0 = 0) > 0$,

whereas in the symmetric case (or fair game) p=q=1/2 we find that

$$\mathbb{P}(T_0^r < \infty \mid S_0 = 0) = 1$$
 and $\mathbb{P}(T_0^r = \infty \mid S_0 = 0) = 0$,

i.e. the random walk returns to \bigcirc with probability one.

See Exercises 3.4-(b) and 5.10-(a) for other derivations of (3.4.16).

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 *

$$egin{array}{lll} \mathbb{E}[T_0^r \mid S_0 = 0] &= & \infty imes \mathbb{P}(T_0^r = \infty \mid S_0 = 0) + \sum\limits_{k \geqslant 1} k \mathbb{P}(T_0^r = k \mid S_0 = 0) \ &= & \infty \end{array}$$

*Note that the summation $\sum_{k \ge 1} = \sum_{1 \le k < \infty}$ actually *excludes* the value $k = \infty$.

MH3512 AY19-20 $\,$

(3.4.17)

(119

is infinite in that case.*

(120

Note that starting from $\underline{S_0 = k \ge 1}$, by (2.3.12) we have found that the mean hitting time of state \bigcirc equals

$$\mathbb{E}[T_0^r \mid S_0 = k] = \begin{cases} \infty & \text{if } q \leq p, \\ \frac{k}{q-p} & \text{if } q > p. \end{cases}$$
(3.4.18)

In particular we have $\mathbb{P}(T_0^r < \infty \mid S_0 = k) = 1$ when q > p and $k \ge 1$, which is consistent with (3.4.16). See Exercise 5.10-(b) for other derivations of (2.3.12)-(3.4.18) using the probability generating function $s \mapsto G_{T_0^r}(s)$.

Remark 3.5

By (3.4.9), the truncated expectation $\mathbb{E}[T_0^r \mathbb{1}_{\{T_0^r < \infty\}} \mid S_0 = 0]$ satisfies

$$\mathbb{E}[T_0^r \mathbb{1}_{\{T_0^r < \infty\}} \mid S_0 = 0] = \sum_{n \ge 1} n \mathbb{P}(T_0^r = n \mid S_0 = 0)$$

$$= G'_{T_0^r}(1^-)$$

$$= \frac{4pqs}{\sqrt{1 - 4pqs^2}}|_{s=1}$$

$$= \frac{4pq}{\sqrt{1 - 4pq}}$$

$$= \frac{4pq}{|p-q|}, \qquad (3.4.19)$$

when $p \neq q$, see for example *here* and in Figure 3.4-a), which shows in particular from Lemma 1.4 that

$$\mathbb{E}[T_0^r \mid T_0^r < \infty, \; S_0 = 0] \; = \; rac{1}{\mathbb{P}(T_0^r < \infty \mid S_0 = 0)} \, \mathbb{E}[T_0^r \mathbbm{1}_{\{T_0^r < \infty\}} \mid S_0 = 0]$$

^{*}We use the convention $\infty \times 0 = 0$.



Figure 3.4: Probability generating functions of T_0^r for p = 0.35 and p = 0.5.

ii) In the symmetric case p=q=1/2 we have $\mathbb{P}(T_0^r<\infty\mid S_0=0)=1$ and

$$\mathbb{E}[T_0^r \mid S_0 = 0] = \mathbb{E}\left[T_0^r \mathbb{1}_{\{T_0^r < \infty\}} \middle| S_0 = 0\right] = G'_{T_0^r}(1^-) = \infty$$
(3.4.20)

as the slope of $s \mapsto G_{T_0^r}(s)$ in Figure 3.4-b) 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. $_{MH3512 \text{ AY19-20}}$

First return time distribution

Proposition 3.6 can also be obtained from the path counting result of Exercise 3.8.

Proposition 3.6

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

$$\mathbb{P}(T_0^r = 2k \mid S_0 = 0) = \frac{1}{2k - 1} \binom{2k}{k} (pq)^k, \quad k \in \mathbb{N},$$
(3.4.21)

with $\mathbb{P}(T_0^r = 2k + 1 \mid S_0 = 0) = 0, k \in \mathbb{N}$.

Proof

By applying a Taylor expansion to $s \mapsto 1 - (1 - 4pqs^2)^{1/2}$ in (3.4.9), we get

$$\begin{aligned} G_{T_0^r}(s) &= 1 - (1 - 4pqs^2)^{1/2} \\ &= 1 - \sum_{k \ge 0} \frac{1}{k!} (-4pqs^2)^k \left(\frac{1}{2} - 0\right) \left(\frac{1}{2} - 1\right) \times \dots \times \left(\frac{1}{2} - (k - 1)\right) \\ &= \frac{1}{2} \sum_{k \ge 1} s^{2k} \frac{(4pq)^k}{k!} \left(1 - \frac{1}{2}\right) \times \dots \times \left(k - 1 - \frac{1}{2}\right), \end{aligned}$$
(3.4.22)

where we used (A.8) for $\alpha = 1/2$. By identification of (3.4.22) with the expansion

$$G_{T_0^r}(s) = \sum\limits_{n \geqslant 0} s^n \mathbb{P}(T_0^r = n \mid S_0 = 0), \qquad -1 \leqslant s \leqslant 1,$$

we obtain

$$egin{aligned} \mathbb{P}(T_0^r = 2k \mid S_0 = 0) &= g(2k) \ &= rac{(4pq)^k}{k!} rac{1}{2} \left(1 - rac{1}{2}
ight) imes \cdots imes \left(k - 1 - rac{1}{2}
ight) \end{aligned}$$

MH3512 AY19-20

$$egin{array}{lll} &=& rac{(4pq)^k}{2k!} \prod\limits_{m=1}^{k-1} \left(m-rac{1}{2}
ight) \ &=& rac{1}{2k-1} {2k \choose k} (pq)^k, \qquad k\in\mathbb{N}, \end{array}$$

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 (1.7.7) from the relation

$$\mathbb{P}(T_0^r=n\mid S_0=0)=rac{1}{n!}rac{\partial^n}{\partial s^n}G_{T_0^r}(s)_{ert s=0},\qquad n\in\mathbb{N}$$

Exercise: 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

$$T_k = \inf\{n \geqslant 0 \; : \; S_n = k\}$$

to any level $k \ge 1$ is given by

$$G_{T_k}(s) = \left(\frac{1 - \sqrt{1 - 4pqs^2}}{2qs}\right)^k, \qquad 4pqs^2 < 1, \quad q \leq p, \tag{3.4.23}$$

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 Chapter 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 Chapters 4 and 5.

Exercises

Exercise 3.1 We consider the simple random walk $(S_n)_{n \in \mathbb{N}}$ of Section 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

$$\mathbb{P}(S_4=2 \mid S_0=0) = {4 \choose 3} p^3 (1-p) = {4 \choose 1} p^3 (1-p)$$

c) Show that we have

$$\mathbb{P}(S_n = k \mid S_0 = 0)$$

$$= \begin{cases} \binom{n}{(n+k)/2} p^{(n+k)/2} (1-p)^{(n-k)/2}, & n+k \text{ even and } |k| \leq n, \\ \mathbb{P}(S_n = k \mid S_0 = 0) = 0, & n+k \text{ odd or } |k| > n. \end{cases}$$
(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

$$p_{n+1,k} = pp_{n,k-1} + qp_{n,k+1}, (3.4.25)$$

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

$$T_0=\inf\{n\geqslant 0 \;:\; S_n=0\}$$

denote the hitting time of state $(\mathbf{0})$.

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 $p \ge q$?

b) Explain why, by the Markov property, we have

$$\mathbb{P}(T_0<\infty\mid S_0=k)=\left(\mathbb{P}(T_0<\infty\mid S_0=1)
ight)^k, \qquad k\geqslant 1.$$

MH3512 AY19-20

c) Using first step analysis for random walks, show that $\alpha := \mathbb{P}(T_0 < \infty \mid S_0 = 1)$ satisfies the quadratic equation

$$plpha^2-lpha+q=p(lpha-q/p)(lpha-1)=0,$$

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 $p \ge q$ respectively.

Exercise 3.3 Consider the random walk

$$S_n := X_1 + \dots + X_n, \qquad n \ge 1,$$

with $\underline{S_0} = 0$, where $(X_k)_{k \ge 1}$ 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)

$$G_{T_0^r}(s) = \sum_{k \ge 0} s^k \mathbb{P}(T_0^r = k), \qquad s \in [-1, 1],$$
(3.4.26)

of the first return time $T_0^r := \inf\{S_n = 0 : n \ge 1\}$ to state (0) is given by

$$G_{T_0^r}(s) = 1 - \sqrt{1 - 4pqs^2}, \quad s \in [-1, 1].$$
 (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 $(X_n)_{n \ge 0}$ on \mathbb{Z} with respective probabilities p and q of increment and decrement. Let

$$T_0:=\inf\{n \geqslant 0 \; : \; X_n=0\}$$
 MH3512 AY19-2

denote the first hitting time of state (0), and consider the probability generating function

$$G_i(s) := \mathbb{E} ig[s^{T_0} \mid X_0 = i ig], \qquad -1 < s < 1, \quad i \in \mathbb{Z}.$$

- 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 (1) starting from state (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 (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 $(X_k)_{k \ge 1}$ of independent Bernoulli random variables with

$$\mathbb{P}(X_k=1)=p, \quad ext{and} \quad \mathbb{P}(X_k=-1)=q, \quad k \geqslant 1,$$

where p + q = 1, and let the process $(M_n)_{n \in \mathbb{N}}$ be defined by $M_0 := 0$ and

$$M_n:=\sum\limits_{k=1}^n 2^{k-1}X_k,\qquad n\geqslant 1.$$

- a) Compute $\mathbb{E}[M_n]$ for all $n \ge 0$.
- b) Consider the hitting time $au := \inf\{n \ge 1 \; : \; M_n = 1\}$ and the stopped process

$$M_{\min(n, au)} = M_n \mathbb{1}_{\{n < au\}} + \mathbb{1}_{\{ au \leqslant n\}}, \qquad n \in \mathbb{N}.$$

Determine the possible values of $M_{\min(n,\tau)}$, and the probability distribution of $M_{\min(n,\tau)}$ at any time $n \ge 1$.

MH3512 AY19-20

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 $n \ge 1$.

Exercise 3.7 Winning streaks. Consider a sequence $(X_n)_{n \ge 1}$ of independent Bernoulli random variables with the distribution

$$\mathbb{P}(X_n=1)=p, \qquad \mathbb{P}(X_n=0)=q, \qquad n \geqslant 1,$$

with q := 1 - p. For some $m \ge 1$, let $T^{(m)}$ denote the time of the first appearance of m consecutive "1" in the sequence $(X_n)_{n\ge 1}$. For example, for m = 4 the following sequence

$$(\overset{1}{0},\overset{2}{1},\overset{3}{1},\overset{4}{0},\overset{5}{\underbrace{\downarrow}},\overset{6}{\underbrace{\downarrow}},\overset{7}{\underbrace{\downarrow}},\overset{8}{\underbrace{\downarrow}},0,1,1,0,\ldots)$$

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

$$G_{T^{(m)}}(s) := \mathbb{E} \left[s^{T^{(m)}} \mathbb{1}_{\{T^{(m)} < \infty\}}
ight], \qquad s \in (-1,1),$$

satisfies

$$G_{T^{(m)}}(s) = p^m s^m + \sum_{k=0}^{m-1} p^k q s^{k+1} G_{T^{(m)}}(s), \qquad s \in (-1,1).$$
(3.4.28)

Hint: Look successively at all possible starting patterns of the form

$$(1,\ldots,1,\overset{\scriptscriptstyle{k}}{1},0,\ldots),$$

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 $s \in (-1, 1)$.

MH3512 AY19-20

 $\circ 127$

Hint: We have

$$\sum_{k=0}^{m-1} x^k = rac{1-x^m}{1-x}, \qquad x \in (-1,1).$$

d) From the probability generating function $G_{T^{(m)}}(s)$, compute $\mathbb{E}\left[T^{(m)}\right]$ for all $m \ge 1$.

Hint: It can be simpler to differentiate inside (3.4.28) and to use the relation

$$(1-x)\sum_{k=0}^{m-1}(k+1)x^k+mx^m=rac{1-x^m}{1-x},\qquad x\in(-1,1).$$

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 (1) and (2k) over 2m time steps is

$$\binom{2m}{m+k}.$$
(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 (1) 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 (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 (1). Hint: Combine the answers to part (a) and (d).

MH3512 AY19-20 $\,$

○128

f) Give the total number of paths joining $S_0 = 0$ to $S_{2n} = 0$ without crossing or hitting (1) between time 1 and time 2n - 1. $\bigcirc 129$ *Hint:* Apply two times the answer to part (e). A drawing is recommended.

Problem 3.9 Time spent above zero by a random walk. Consider a symmetric random walk $(S_n)_{n \in \mathbb{N}}$ started at $S_0 = 0$. We let

$$T_{2n}^+ := 2\sum\limits_{r=1}^n \mathbb{1}_{\{S_{2r-1} \geqslant 1\}}$$

denote an even estimate of the time spent strictly above the level 0 by the random walk between time 0 and time 2n. We also let

$$T_0:=\inf\{n\geqslant 1\ :\ S_n=0\}$$

denote an even estimate of the time of first return of $(S_n)_{n \in \mathbb{N}}$ to (0).

- a) Compute $\mathbb{P}(S_{2n} = 2k)$ for $k = 0, 1, \ldots, n$.
- b) Show the *convolution equation*

$$\mathbb{P}(S_{2n}=0)=\sum\limits_{r=1}^{n}\mathbb{P}(T_{0}=2r)\mathbb{P}(S_{2n-2r}=0), \hspace{0.5cm} n\geqslant 1.$$

c) By partitioning the event $\{T_{2n}^+ = 2k\}$ according to all possible times $2r = 2, 4, \ldots, 2n$ of *first* return to state (1) until time 2n, show the convolution equation

$$egin{aligned} \mathbb{P}(T_{2n}^+ = 2k) &= \sum\limits_{r=1}^n \mathbb{P}ig(T_0 = 2r, T_{2n}^+ = 2kig) \ &= rac{1}{2} \sum\limits_{r=1}^k \mathbb{P}ig(T_0 = 2rig) \mathbb{P}ig(T_{2n-2r}^+ = 2k-2rig) \ &+ rac{1}{2} \sum\limits_{r=1}^{n-k} \mathbb{P}ig(T_0 = 2rig) \mathbb{P}ig(T_{2n-2r}^+ = 2kig), \quad n \geqslant 1 \end{aligned}$$

d) Show that

$$\mathbb{P}(T_{2n}^+=2k)=\mathbb{P}(S_{2k}=0)\mathbb{P}(S_{2n-2k}=0), \qquad 0\leqslant k\leqslant n_2$$

MH3512 AY19-20

solves the convolution equation of Question (c).

e) Using the Stirling approximation $n! \simeq (n/e)^n \sqrt{2\pi n}$ as n tends to ∞ , compute the limit

$$\lim_{n o\infty} \mathbb{P}ig(T^+_{2n}/(2n)\leqslant xig) = \lim_{n o\infty} \sum_{0\leqslant k\leqslant nx} \mathbb{P}ig(T^+_{2n}/(2n)=k/nig),$$

and find the limiting distribution of $T_{2n}^+/(2n)$ as n tends to infinity.

Problem 3.10 Range process. Consider the random walk $(S_n)_{n \ge 0}$ defined by $S_0 = 0$ and

$$S_n := X_1 + \dots + X_n, \qquad n \ge 1,$$

where $(X_k)_{k \ge 1}$ is an *i.i.d.* * family of $\{-1, +1\}$ -valued random variables with distribution

$$ig egin{array}{ll} \mathbb{P}(X_k=+1)=p, \ \mathbb{P}(X_k=-1)=q, \end{array}$$

 $k \ge 1$, where p + q = 1. We let R_n denote the <u>range</u> 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

$$R_n = 1 + \left(\sup_{k=0,1,...,n} S_k
ight) - \left(\inf_{k=0,1,...,n} S_k
ight),$$

and give the value of R_0 and R_1 .

b) Show that for all $k \ge 1, R_k - R_{k-1}$ is a Bernoulli random variable, and that

$$\mathbb{P}(R_k - R_{k-1} = 1) = \mathbb{P}(S_k - S_0 \neq 0, S_k - S_1 \neq 0, \dots, S_k - S_{k-1} \neq 0).$$

^{*}independent and identically distributed.

c) Show that for all $k \ge 1$ we have

$$\mathbb{P}(R_k - R_{k-1} = 1) = \mathbb{P}(X_1 \neq 0, X_1 + X_2 \neq 0, \dots, X_1 + \dots + X_k \neq 0).$$

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 $\mathbb{P}(T_0^r = \infty) = \lim_{k \to \infty} \mathbb{P}(T_0^r > k).$
- 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 $T_0^r = \inf\{n \ge 1 : S_n = 0\}$ is the time of first return to (0) of the random walk.

g) From the results of Questions (e) and (f), show that

$$\mathbb{P}(T_0^r=\infty)=\lim_{n o\infty}rac{1}{n}\,\mathbb{E}[R_n].$$

h) Show that

when
$$p = 1/2$$
, and that $\mathbb{E}[R_n] \simeq_{n \to \infty} n|p-q|$, when $p \neq 1/2$.*

Hints and comments on Problem 3.10.

- a) No mathematical computation is needed here, a credible explanation (in words) is sufficient. It may be of interest to also compute $\mathbb{E}[\mathbf{R}_2]$.
- b) Show first that the two events

$$\{R_k - R_{k-1} = 1\}$$
 and $\{S_k - S_0 \neq 0, S_k - S_1 \neq 0, \dots, S_k - S_{k-1} \neq 0\}$

MH3512 AY19-20

^{*}The meaning of $f(n) \simeq_{n \to \infty} g(n)$ is $\lim_{n \to \infty} f(n)/g(n) = 1$, provided that $g(n) \neq 0, n \ge 1$.

are the same.

c) Here the events

 $\{R_k - R_{k-1} = 1\}$ and $\{X_1 \neq 0, X_1 + X_2 \neq 0, \dots, X_1 + \dots + X_k \neq 0\}$

are not the same, however we can show the equality between the probabilities.

- d) Telescoping identities can be useful in many situations, not restricted to probability or stochastic processes.
- e) Use point 2. on page 27, after identifying what the events A_k are.
- f) The basic identity $\mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$ can be used.
- g) A mathematically rigorous proof is asked here. The following definition of limit may be used:

A real number a is said to be the limit of the sequence $(x_n)_{n \ge 1}$, written $a = \lim_{n \to \infty} x_n$, if and only if for every real number $\varepsilon > 0$, there exists a natural number N such that for every n > N we have $|x_n - a| < \varepsilon$.

Alternatively, we may use the notion of Cesàro mean and state and apply the relevant theorem.

h) Use the formula (3.4.15) giving $\mathbb{P}(T_0^r = \infty)$.

Figure 3.5: Illustration of the range process.*

In Figure 3.5 the height at time n of the colored area coincides with $R_n - 1$.

Problem 3.11 Recurrent random walk in d dimensions. Let $\{e_1, e_2, \ldots, e_d\}$ denote the canonical basis of \mathbb{R}^d , *i.e.*

$$e_k=(0,\ldots, {\displaystyle 1\atop {\stackrel{\uparrow}{k}}}, 0,\ldots, 0), \qquad k=1,2,\ldots, d.$$

Consider the symmetric \mathbb{Z}^d -valued random walk

$$S_n = X_1 + \dots + X_n, \qquad n \in \mathbb{N},$$

started at $S_0 = \vec{0} = (0, 0, \dots, 0)$, where $(X_n)_{n \ge 1}$ is a sequence of independent uniformly distributed random variables

$$X_n \in \{e_1,e_2,\ldots,e_d,-e_1,-e_2,\ldots,-e_d\}, \qquad n \geqslant 1,$$

^{*}Animated figure (works in Acrobat Reader).

with distribution

$$\mathbb{P}(X_n=e_k)=\mathbb{P}(X_n=-e_k)=rac{1}{2d},\qquad k=1,2,\ldots,d.$$

Let

$$T^r_{ec 0}:=\inf\{n \geqslant 1 \; : \; S_n=ec 0\}$$

denote the first return time of $(S_n)_{n \in \mathbb{N}}$ to $\vec{0} = (0, 0, \dots, 0)$. The random walk is said to be <u>recurrent</u> if $\mathbb{P}(T_{\vec{0}}^r < \infty) = 1$.

a) Show that the probability distribution $\mathbb{P}(T^r_{\vec{0}} = n), n \ge 1$, satisfies the equation

$$\mathbb{P}ig|(S_n=ec{0}ig)=\sum\limits_{k=2}^n \mathbb{P}ig|(T^r_{ec{0}}=k)\mathbb{P}ig|(S_{n-k}=ec{0}ig), \qquad n \geqslant 1.$$

b) Show that

$$\sum_{n=2}^{m} \mathbb{P}(T_{\vec{0}}^r = n) \ge 1 - \frac{1}{\sum_{n=0}^{m} \mathbb{P}(S_n = \vec{0})}, \qquad m \ge 1.$$
(3.4.30)

Hint: Start by showing that

$$\sum_{n=1}^m \mathbb{P}(S_n = \vec{0}) = \sum_{k=2}^m \mathbb{P}(T_{\vec{0}}^r = k) \sum_{l=0}^{m-k} \mathbb{P}(S_l = \vec{0}).$$

c) Show that under the condition

$$\sum_{n \ge 0} \mathbb{P}(S_n = \vec{0}) = \infty,$$

we have $\mathbb{P}(T^r_{\vec{0}} < \infty) = 1$.

d) Show that

$$\sum_{n=2}^{m} \mathbb{P}(T_{\vec{0}}^{r} = n) \leqslant \frac{\sum_{n=2}^{2m} \mathbb{P}(S_{n} = \vec{0})}{\sum_{n=0}^{m} \mathbb{P}(S_{n} = \vec{0})}, \qquad m \ge 1.$$
(3.4.31)

Hint: Start by showing that

$$\sum_{n=1}^{2m} \mathbb{P}(S_n = \vec{0}) = \sum_{n=2}^{2m} \mathbb{P}(T_{\vec{0}}^r = n) \sum_{l=0}^{2m-n} \mathbb{P}(S_l = \vec{0}).$$

() 134

MH3512 AY19-20 $\,$

e) Show that under the condition

$$\sum\limits_{n \geqslant 0} \mathbb{P}ig(S_n = ec{0}ig) < \infty,$$

we have $\mathbb{P}(T^r_{\vec{0}} = \infty) > 0.$

f) When $\underline{d=1}$, compute $\mathbb{P}(S_{2n} = 0)$, $n \ge 1$, and show that the one-dimensional random walk is recurrent, *i.e.* we have $\mathbb{P}(T_0^r < \infty) = 1$.

Hint: Use Stirling's approximation $n! \simeq (n/e)^n \sqrt{2\pi n}$ as n tends to ∞ .

g) When $\underline{d=2}$, show that we have

$$\mathbb{P}(S_{2n} = \vec{0}) = \frac{(2n)!}{4^{2n}(n!)^2} \sum_{k=0}^n \binom{n}{k}^2, \qquad n \ge 1,$$
(3.4.32)

and that the two-dimensional random walk is recurrent, *i.e.* we have $\mathbb{P}(T_{\vec{0}}^r < \infty) = 1$.

Hint: Use the combinatorial identity

$$\binom{2n}{n} = \sum\limits_{k=0}^n \binom{n}{k}^2$$

and Stirling's approximation.

h) Given $i_1, i_2, \ldots, i_d \in \mathbb{N}$, count all paths starting from $\vec{0}$ and returning to $\vec{0}$ via i_k "forward" steps in the direction e_k and i_k "backward" steps in the direction $-e_k$, $k = 1, 2, \ldots, d$. Deduce an expression for $\mathbb{P}(S_{2n} = \vec{0})$ that generalizes (3.4.32) to all $d \ge 2$.

Hint: Use multinomial coefficients.

i) Based on the Euclidean division $n = a_n d + b_n$ where $b_n \in \{0, 1, \dots, d-1\}$, show that we have

$$\sum_{n \geqslant 1} \mathbb{P}ig(S_{2n} = ec{0}ig) \leqslant \sum_{n \geqslant 1} rac{(2n)!}{2^{2n} d^n n! (a_n!)^d a_n^{b_n}}$$
MH3512 AY19-20

Hint: Use the bound

$$i_1!i_2!\cdots i_d! \geqslant (a_n!)^d(a_n+1)^{b_n}$$

which is valid for any $n = i_1 + \cdots + i_d$, and the identity

$$d^n = \sum\limits_{\substack{i_1+\dots+i_d=n\ i_1,i_2,\dots,i_d\geqslant 0}} rac{n!}{i_1!\cdots i_d!}.$$

j) Applying Stirling's approximation, to n!, (2n)! and $a_n!$, show that there exists a constant C > 0 such that for all n sufficiently large we have

$$rac{(2n)!}{2^{2n}d^nn!(a_n!)^d} \leqslant rac{C}{n^{d/2}}$$

Hint: We have $\lim_{m\to\infty} (1 + x/m)^m = e^x$ for all $x \in \mathbb{R}$.

k) Is the random walk recurrent when $d \ge 3$?

Problem 3.12 ([AR05], [BW03]). Random walks in a cookie environment, also called excited random walks, can be used to model the behavior of primitive organisms.

In the absence of cookies the random walk is symmetric, with probabilities 1/2 of going up and down, and it can *rebound* to (1) with probability 1/2 after hitting state (0).

When the random walk encounters a cookie, its behavior becomes modified and it restarts with probabilities p and $\underline{q} = 1 - p$ of moving up, resp. down, where $p \in [0, 1]$. Every encountered cookie is eaten by the organism, and when the random walk reaches an empty spot it restarts with equal probabilities 1/2 of moving up or down. In this case the organism wanders without a preferred direction. The random walk is *attracted* by the cookies when p > 1/2, and *repulsed* when p < 1/2.

Figure 3.6: Random walk with cookies.*

- a) Does the cookie random walk $(S_n)_{n \in \mathbb{N}}$ have the Markov property? Explain your answer.
- b) Suppose that at time $n \ge 1$ the random walk has just eaten a cookie at state $x \ge 1$, after eating all cookies at states $1, 2, \ldots, x 1$. Show that in this case, the probability of reaching x + 1 before reaching 0 is given by

$$p+(1-p)rac{x-1}{x+1}=1-rac{2q}{x+1},\qquad x\geqslant 1.$$

Hint: Use first step analysis together with formula (2.2.12) page 74, see also page 81.

c) For any $x \ge 1$, let τ_x denote the first hitting time

$$au_x := \inf\{n \geqslant 1 \ : \ S_n = x\}.$$
 MH3512 AY19-20

Give the value of $\mathbb{P}(\tau_1 < \tau_0 \mid S_0 = 0)$, and show that for all $x \ge 1$ we have^{*}

$$\mathbb{P}(au_x < au_0 \mid S_0 = 0) = rac{1}{2} \exp\left(\sum\limits_{l=2}^x \log\left(1 - rac{2q}{l}
ight)
ight), \qquad x \geqslant 1,$$

where "log" denotes the *natural logarithm* "ln".

d) Show the bound $\mathbb{P}(\tau_x < \tau_0 \mid S_0 = 0) \leqslant (x/2)^{-2q}$ for $x \ge 2$.

Hint: Use comparison arguments between integrals and series.

e) Show that $\mathbb{P}(\tau_0 < \infty \mid S_0 = 0) = 1$, provided that p < 1.[†]

Hint: Consider $\mathbb{P}(\tau_0 \leq \tau_x \mid S_0 = 0)$ and use Relation (1.2.3) page 27.

- f) Starting from state (0), compute the mean time needed by the random walk to reach state (1). *Hint:* Apply first step analysis, or locate and use the relevant result in Problem 2.7.
- g) Suppose that a cookie has just been eaten at state $x \ge 1$, after eating all cookies at states $1, 2, \ldots, x 1$. Show that the mean time to reach the next cookie at state x + 1 is 1 + q(4x + 2).

Hint: Locate and apply the relevant results in Problem 2.7 together with first step analysis.

h) Show that the mean time to reach state (x) starting from (0) is given by

$$\mathbb{E}[au_x \mid S_0 = 0] = 1 - 2q + x + 2qx^2, \qquad x \geqslant 1.$$

Hint: Perform a summation on the results of Questions (f) and (g).

i) Suppose that a cookie has just been eaten at state $x \ge 1$, after eating all cookies at states $1, 2, \ldots, x - 1$. Show that, given one

^{*}We use the convention $\sum_{k=2} a_k = 0$ for any sequence (a_k) .

[†]In this case the cookie random walk is said to be *recurrent*.

does not hit $(\mathbf{0})$, the probabilities of moving up to $(\mathbf{x} + \mathbf{1})$, resp. down to $(\mathbf{x} - \mathbf{1})$, are

$$rac{p}{1-2q/(x+1)} \quad ext{and} \quad \quad rac{q(x-1)/(x+1)}{1-2q/(x+1)}, \qquad x \geqslant 1.$$

Hint: Proceed similarly to Question (g) of Problem 2.7.

j) Suppose that a cookie has just been eaten at state $x \ge 1$, after eating all cookies at states $1, 2, \ldots, x - 1$. Show that the mean time to reach the next cookie at state x + 1 given one does not hit (1) is

$$1+rac{q-2q/(x+1)}{1-2q/(x+1)}rac{4x}{3},\qquad x\geqslant 1.$$

Hint: Locate and use the relevant result in Problem 2.7.

Problem 3.13 ([AR05]). A random walk $(S_n)_{n \in \mathbb{N}}$ with cookies on $\{1, 2, 3, \ldots\}$ is symmetric in the absence of cookies, and restarts with probabilities p and q = 1 - p of moving up, resp. down, when it encounters a cookie, where $p \in [0, 1)$. The random walk starts at state (0), which is empty of cookie.

For any $x \ge 1$, let τ_x denote the first hitting time

$$au_x := \inf\{n \geqslant 1 \; : \; S_n = x\}, \qquad x \geqslant 1.$$

Recall that the probability of eating at least \boldsymbol{x} cookies before returning to the origin $(\mathbf{0})$ is given by

$$\mathbb{P}(\tau_x < \tau_0 \mid S_0 = 0) = \frac{1}{2} \prod_{l=2}^x \left(1 - \frac{2q}{l} \right), \qquad x \ge 1,$$
(3.4.33)

and that the random walk is recurrent, *i.e.* it returns to the origin 0 in finite time whenever p < 1, that means we have $\mathbb{P}(\tau_0 < \infty \mid S_0 = 0) = 1$.

a) Let X denote the number of cookies eaten by the random walk before returning to the origin 0. Show that

$$\mathbb{P}(X=0)=1/2,\qquad \mathbb{P}(X=1)=q/2,$$

and, using (3.4.33), that the distribution of satisfies

$$\mathbb{P}(X=x) = \frac{q}{x+1} \prod_{l=2}^{x} \left(1 - \frac{2q}{l}\right), \qquad x \ge 2.$$
(3.4.34)

b) Show from (3.4.34) that the average number $\mathbb{E}[X]$ of cookies eaten before returning to the origin (0) is finite, *i.e.* $\mathbb{E}[X] < \infty$, if and only if q > 1/2.

Hint: There exists constants $c_q, C_q > 0$ such that

$$rac{c_q}{x^{2q}}\leqslant \prod\limits_{l=2}^x \left(1-rac{2q}{l}
ight)\leqslant rac{C_q}{x^{2q}}, \qquad x\geqslant 2.$$

(140

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	141
4.2	Transition matrix	144
4.3	Examples of Markov Chains	149
4.4	Higher-Order Transition Probabilities	155
4.5	The Two-State Discrete-Time Markov Chain	160
Exe	$\operatorname{Exercises} \ldots \ldots$	

4.1 Markov Property

We consider a discrete-time stochastic process $(Z_n)_{n \in \mathbb{N}}$ taking values in a discrete state space S, typically $S = \mathbb{Z}$.

The S-valued process $(Z_n)_{n \in \mathbb{N}}$ is said to be *Markov*, or to have the *Markov property* if, for all $n \ge 1$, 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 $n \ge 1$ and all $i_0, i_1, \ldots, i_n, j \in \mathbb{S}$ we have

$$\mathbb{P}(Z_{n+1}=j\mid Z_n=i_n,\ Z_{n-1}=i_{n-1},\ldots,Z_0=i_0)=\mathbb{P}(Z_{n+1}=j\mid Z_n=i_n).$$
 MH3512 AY19-20

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.

1. The first order transition probabilities can be used for the complete computation of the probability distribution of the process as

$$\begin{split} \mathbb{P}(Z_n = i_n, Z_{n-1} = i_{n-1}, \dots, Z_0 = i_0) \\ &= \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1}, \dots, Z_0 = i_0) \mathbb{P}(Z_{n-1} = i_{n-1}, \dots, Z_0 = i_0) \\ &= \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1}) \mathbb{P}(Z_{n-1} = i_{n-1} \mid Z_{n-2} = i_{n-2}, \dots, Z_0 = i_0) \\ &\times \mathbb{P}(Z_{n-2} = i_{n-2}, \dots, Z_0 = i_0) \\ &= \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1}) \mathbb{P}(Z_{n-1} = i_{n-1} \mid Z_{n-2} = i_{n-2}) \\ &\times \mathbb{P}(Z_{n-2} = i_{n-2} \mid Z_{n-3} = i_{n-3}, \dots, Z_0 = i_0) \mathbb{P}(Z_{n-3} = i_{n-3}, \dots, Z_0 = i_0) \\ &= \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1}) \mathbb{P}(Z_{n-1} = i_{n-1} \mid Z_{n-2} = i_{n-2}) \\ &\times \mathbb{P}(Z_{n-2} = i_{n-2} \mid Z_{n-3} = i_{n-3}) \mathbb{P}(Z_{n-3} = i_{n-3}, \dots, Z_0 = i_0), \end{split}$$

which shows, reasoning by induction, that

$$\mathbb{P}(Z_n = i_n, Z_{n-1} = i_{n-1}, \dots, Z_0 = i_0)
= \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1}) \cdots \mathbb{P}(Z_1 = i_1 \mid Z_0 = i_0) \mathbb{P}(Z_0 = i_0),$$
(4.1.1)

or

$$\mathbb{P}(Z_n = i_n, Z_{n-1} = i_{n-1}, \dots, Z_1 = i_1 \mid Z_0 = i_0)
= \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1}) \cdots \mathbb{P}(Z_1 = i_1 \mid Z_0 = i_0),$$
(4.1.2)

 $i_0, i_1, \ldots, i_n \in \mathbb{S}$.

2. By the *law of total probability* (1.3.1) applied to the events $A_k = \{Z_2 = i_2 \text{ and } Z_1 = k\}, k \in \mathbb{S}$, under the probability measure $\mathbb{P}(\cdot \mid Z_0 = i_0)$ we also have

$$\begin{split} \mathbb{P}(Z_2 = i_2 \mid Z_0 = i_0) &= \sum_{i_1 \in \mathbb{S}} \mathbb{P}(Z_2 = i_2 \text{ and } Z_1 = i_1 \mid Z_0 = i_0) \\ &= \sum_{i_1 \in \mathbb{S}} \mathbb{P}(Z_2 = i_2 \mid Z_1 = i_1) \mathbb{P}(Z_1 = i_1 \mid Z_0 = i_0), \end{split}$$

 $i_0, i_2 \in \mathbb{S}$, and

$$\mathbb{P}(Z_1 = i_1) = \sum_{i_0 \in \mathbb{S}} \mathbb{P}(Z_1 = i_1, Z_0 = i_0) \\
= \sum_{i_0 \in \mathbb{S}} \mathbb{P}(Z_1 = i_1 \mid Z_0 = i_0) \mathbb{P}(Z_0 = i_0), \quad i_1 \in \mathbb{S}.$$
(4.1.3)

Example

The random walk

$$S_n := X_1 + X_2 + \dots + X_n, \qquad n \in \mathbb{N}, \tag{4.1.4}$$

considered in Chapter 3, where $(X_n)_{n \ge 1}$ is a sequence of independent Z-valued random increments, is a discrete-time Markov chain with $S = \mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$.

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)

$$\mathbb{P}(S_{n+1} = j \mid S_n = i_n, S_{n-1} = i_{n-1}, \dots, S_1 = i_1)$$

$$= \frac{\mathbb{P}(S_{n+1} = j, S_n = i_n, S_{n-1} = i_{n-1}, \dots, S_1 = i_1)}{\mathbb{P}(S_n = i_n, S_{n-1} = i_{n-1}, \dots, S_1 = i_1)}$$

$$= \frac{\mathbb{P}(S_{n+1} - S_n = j - i_n, S_n - S_{n-1} = i_n - i_{n-1}, \dots, S_2 - S_1 = i_2 - i_1, S_1 = i_1)}{\mathbb{P}(S_n - S_{n-1} = i_n - i_{n-1}, \dots, S_2 - S_1 = i_2 - i_1, S_1 = i_1)}$$

$$= \frac{\mathbb{P}(X_{n+1} = j - i_n, X_n = i_n - i_{n-1}, \dots, X_2 = i_2 - i_1, X_1 = i_1)}{\mathbb{P}(X_n = i_n - i_{n-1}, \dots, X_2 = i_2 - i_1, X_1 = i_1)}$$
(4.1.5)
(4.1.5)

$$= \frac{\mathbb{P}(X_{n+1} = j - i_n)\mathbb{P}(X_n = i_n - i_{n-1}, \dots, X_2 = i_2 - i_1, X_1 = i_1)}{\mathbb{P}(X_n = i_n - i_{n-1}, \dots, X_2 = i_2 - i_1, X_1 = i_1)}$$

$$= \frac{\mathbb{P}(X_{n+1} = j - i_n)}{\mathbb{P}(X_{n+1} = j - i_n)\mathbb{P}(X_n + \dots + X_1 = i_n)}$$

$$= \frac{\mathbb{P}(X_{n+1} = j - i_n, X_n + \dots + X_1 = i_n)}{\mathbb{P}(X_1 + \dots + X_n = i_n)}$$

$$= \frac{\mathbb{P}(X_{n+1} = j - i_n \text{ and } S_n = i_n)}{\mathbb{P}(S_n = i_n)} = \frac{\mathbb{P}(S_{n+1} = j \text{ and } S_n = i_n)}{\mathbb{P}(S_n = i_n)}$$

In addition, the Markov chain $(S_n)_{n \in \mathbb{N}}$ is time homogeneous if the random sequence $(X_n)_{n \ge 1}$ 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, <u>not all Markov chains have independent</u> 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

$$P_{i,j} := \mathbb{P}(Z_1 = j \mid Z_0 = i), \qquad i, j \in \mathbb{S},$$

$$(4.2.1)$$

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 <u>time homogeneous</u>. This data can be encoded into a matrix indexed by $\mathbb{S}^2 = \mathbb{S} \times \mathbb{S}$, called the <u>transition matrix</u> of the
Markov chain:

$$egin{array}{ll} \left[egin{array}{c} P_{i,j} \end{array}
ight]_{i,j\in\mathbb{S}} = \left[egin{array}{c} \mathbb{P}(Z_1=j \mid Z_0=i) \end{array}
ight]_{i,j\in\mathbb{S}}, \end{array}$$

also written on $\mathbb{S} := \mathbb{Z}$ as

$$P = [P_{i,j}]_{i,j\in\mathbb{S}} = \begin{vmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ \cdots & P_{-2,-2} & P_{-2,-1} & P_{-2,0} & P_{-2,1} & P_{-2,2} & \cdots \\ \cdots & P_{-1,-2} & P_{-1,-1} & P_{-1,0} & P_{-1,1} & P_{-1,2} & \cdots \\ \cdots & P_{0,-2} & P_{0,-1} & P_{0,0} & P_{0,1} & P_{0,2} & \cdots \\ \cdots & P_{1,-2} & P_{1,-1} & P_{1,0} & P_{1,1} & P_{1,2} & \cdots \\ \cdots & P_{2,-2} & P_{2,-1} & P_{2,0} & P_{2,1} & P_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{vmatrix}$$

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 <u>number</u> in the matrix, while the final state (j) corresponds to a <u>column number</u>.

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

$$\sum_{j \in S} \mathbb{P}(Z_1 = j \mid Z_0 = i) = \mathbb{P}(\bigcup_{j \in S} \{Z_1 = j\} \mid Z_0 = i) = \mathbb{P}(\Omega) = 1, \quad i \in \mathbb{N},$$
(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 $i \in S$.

Using the matrix notation $P = (P_{i,j})_{i,j \in S}$, and Relation (4.1.1) we find

$$\mathbb{P}(Z_n=i_n,Z_{n-1}=i_{n-1},\ldots,Z_0=i_0)=P_{i_{n-1},i_n}\cdots P_{i_0,i_1}\mathbb{P}(Z_0=i_0),$$

 $i_0, i_1, \ldots, i_n \in \mathbb{S}$, and we rewrite (4.1.3) as

$$\mathbb{P}(Z_1 = i) = \sum_{j \in \mathbb{S}} \mathbb{P}(Z_1 = i \mid Z_0 = j) \mathbb{P}(Z_0 = j) = \sum_{j \in \mathbb{S}} P_{j,i} \mathbb{P}(Z_0 = j), \quad i \in \mathbb{S}.$$
(4.2.3)

A state $k \in S$ is said to be <u>absorbing</u> 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 $[\mathbb{P}(Z_{n+1} = j \mid Z_n = i)]_{i,j \in \mathbb{N}}$ of the chain is written as

$$\left[\begin{array}{cccc} P_{0,0} & P_{0,1} & P_{0,2} & \cdots \\ P_{1,0} & P_{1,1} & P_{1,2} & \cdots \\ P_{2,0} & P_{2,1} & P_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{array}
ight]$$

From (4.2.2) we have

$$\sum_{j \geqslant 0} P_{i,j} = 1,$$

for all $i \in \mathbb{N}$.

In case the Markov chain $(Z_k)_{k \in \mathbb{N}}$ takes values in the finite state space $\mathbb{S} = \{0, 1, \dots, N\}$ its $(N+1) \times (N+1)$ transition matrix.

will simply have the form

$$\left[\begin{array}{ccccc} P_{0,0} & P_{0,1} & P_{0,2} & \cdots & P_{0,N} \ P_{1,0} & P_{1,1} & P_{1,2} & \cdots & P_{1,N} \ P_{2,0} & P_{2,1} & P_{2,2} & \cdots & P_{2,N} \ dots & dot$$

Still on the finite state space $S = \{0, 1, \dots, N\}$, Relation (4.2.3) can be restated in the language of matrix and vector products using the shorthand notation:

$$\eta = \pi P, \tag{4.2.4}$$

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

$$[\eta_0, \eta_1, \dots, \eta_N] = [\pi_0, \dots, \pi_N] imes egin{bmatrix} P_{0,0} & P_{0,1} & P_{0,2} & \cdots & P_{0,N} \ P_{1,0} & P_{1,1} & P_{1,2} & \cdots & P_{1,N} \ P_{2,0} & P_{2,1} & P_{2,2} & \cdots & P_{2,N} \ dots & dots &$$

Invariant vectors

A row vector $\boldsymbol{\pi}$ such that $\boldsymbol{\pi} = \boldsymbol{\pi} \boldsymbol{P}$ is said to be *invariant* or *stationary* by the transition matrix \boldsymbol{P} .

For example, in case the matrix \boldsymbol{P} takes the form

$$P = \left[\begin{array}{cccc} P_{i,j} \end{array}
ight]_{0\leqslant i,j\leqslant N} = \left[egin{array}{ccccc} \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \ \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \ dots & do$$

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 $R_{Y_{19-20}}$

(148

(4.2.5)

4.3 Examples of Markov Chains

() 149

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 $[P_{i,j}]_{i,j\in\mathbb{S}}$ of the unrestricted random walk (4.1.4) is given by

$$i-1 \ i \ i+1$$

 $i-1 \ i \ i+1$
 $\vdots \ \vdots \ \vdots \ \vdots \ \vdots \ \cdot \cdot$
 $\cdots 0 \ p \ 0 \ 0 \ 0 \cdots$
 $\cdots q \ 0 \ p \ 0 \ 0 \cdots$
 $\cdots 0 \ q \ 0 \ p \ 0 \cdots$
 $\cdots 0 \ q \ 0 \ p \ 0 \cdots$
 $\cdots 0 \ q \ 0 \ p \cdots$
 $\cdots 0 \ q \ 0 \ p \cdots$
 $\cdots 0 \ q \ 0 \ p \cdots$
 $\cdots 0 \ q \ 0 \ p \cdots$
 $\cdots 0 \ q \ 0 \ p \cdots$

(4.3.1)

The transition matrix $[P_{i,j}]_{0 \le i,j \le S}$ of the gambling process on $\{0, 1, \ldots, S\}$ with absorbing states (1) and (S) is given by $\bigcirc 150$

iii) Credit rating.

[transition probabilities are expressed in %].

start of a	year	Rating at the end of the year									
	AAA	AA	А	BBB	BB	В	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	
А	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	
В	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	

Rating at the

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%).

$\bigcirc 151$

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 $Z_n \in \{0, 1, ..., N\}$ 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

$$\mathbb{P}(Z_{n+1} = k+1 \mid Z_n = k) = \frac{N-k}{N}, \qquad k = 0, 1, \dots, N-1,$$
(4.3.2)

and

$$\mathbb{P}(Z_{n+1}=k-1 \mid Z_n=k) = rac{k}{N}, \qquad k=1,2,\ldots,N,$$
 (4.3.3) MH3512(4.3.3) MH3512(4.3.3)

with

	0	1	0	• • •	• • •	0	0	
	1/N	0	(N - 1)/N	• • •	•••	0	0	
	0	2/N	0	• • •	• • •	0	0	
	0	0	3/N	• • •	•••	0	0	
P =	•	• • •	·	•.	• •	• • •	:	,
	0	0	•••	•••	3/N	0	0	
	0	0	•••	۰.	0	2/N	0	
	0	0	• • •	• • •	(N-1)/N	0	1/N	
	0	0	• • •	•••	0	1	0	

cf. Exercise 6.7, Exercise 7.3, Problem 7.24 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.

	А	A ♯	В	С	D	E	F	G	G ‡
Α	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
В	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
Е	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



Figure 4.1: Mozart: variations KV 265 for Piano.*

The transitions of Mozart's variations, cf. Figure 4.1 above, on this famous tune \clubsuit have been analyzed to form a transition matrix.[†] Then that transition matrix was used for random melody \clubsuit generation.[‡] Hear also this arrangement \clubsuit and see *here* for details[§] [Pas14], and *here* for recent examples.

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 as in e.g. Problem 4.14 or Problem 5.23. See *here* for an example of the use of Markov chain for random text generation.

(5153)

 $^{^{*}\}mathrm{Click}$ on the figure to play the video (works in Acrobat Reader).

[†]Try *here* if it does not work.

[‡]Try *here* if it does not work.

[§]Try *here* if it does not work.

Other applications of Markov chains include:

- Memory management in computer science,
- Logistics, supply chain management, and waiting queues,
- Modeling of insurance claims (Problem 4.13),
- Board games, e.g. Snakes and Ladders,
- Genetics, cf. the Wright-Fisher model (Problem 5.22).
- Random fields in imaging,
- Artificial intelligence, learning theory and machine learning (Problem 4.14, Problem 5.21, Problem 5.23).

Graph representation

Whenever possible we will represent a Markov chain using a graph, as in the following example with transition matrix

$$P = \begin{bmatrix} 0 & 0.2 & 0.8 & 0 & 0 \\ 0.4 & 0 & 0 & 0.6 & 0 \\ 0.5 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.4 & 0.6 \\ 0 & 0.4 & 0.6 & 0 & 0 \end{bmatrix},$$
(4.3.4)

see Figure 4.2.

Figure 4.2: 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 $\mathbb{P}(Z_{n+1} = j \mid Z_n = i), i, j \in S$, 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 $\mathbb{P}(\cdot | Z_n = i)$ as follows.

i) 2-step transitions. Denoting by S the state space of the process, we have

$$\mathbb{P}(Z_{n+2}=j\mid Z_n=i)=\sum_{l\in\mathbb{S}}\mathbb{P}(Z_{n+2}=j ext{ and } Z_{n+1}=l\mid Z_n=i)$$

^{*}Animated figure (works in Acrobat Reader).

$$\begin{split} &= \sum_{l \in \mathbb{S}} \frac{\mathbb{P}(Z_{n+2} = j, \ Z_{n+1} = l, \ Z_n = i)}{\mathbb{P}(Z_n = i)} \\ &= \sum_{l \in \mathbb{S}} \frac{\mathbb{P}(Z_{n+2} = j, \ Z_{n+1} = l, \ Z_n = i)}{\mathbb{P}(Z_{n+1} = l \text{ and } Z_n = i)} \frac{\mathbb{P}(Z_{n+1} = l \text{ and } Z_n = i)}{\mathbb{P}(Z_n = i)} \\ &= \sum_{l \in \mathbb{S}} \mathbb{P}(Z_{n+2} = j \mid Z_{n+1} = l \text{ and } Z_n = i) \mathbb{P}(Z_{n+1} = l \mid Z_n = i) \\ &= \sum_{l \in \mathbb{S}} \mathbb{P}(Z_{n+2} = j \mid Z_{n+1} = l) \mathbb{P}(Z_{n+1} = l \mid Z_n = i) \\ &= \sum_{l \in \mathbb{S}} P_{i,l} P_{l,j} \\ &= [P^2]_{i,j}, \quad i, j \in \mathbb{S}, \end{split}$$

where we used (4.2.1). In other words, using matrix product notation, we find

ii) **k**-step transitions. More generally we have the following result.

Proposition 4.1

For all $k \in \mathbb{N}$ we have the relation

$$\left[\ \mathbb{P}(Z_{n+k}=j \mid Z_n=i) \
ight]_{i,j\in\mathbb{S}} = \left[\ [P^k]_{i,j} \
ight]_{i,j\in\mathbb{S}} = P^k.$$

() 156

Proof

<u> 👌 15</u>7

For all $k \in \mathbb{N}$, we have

$$\begin{split} \mathbb{P}(Z_{n+k+1} = j \mid Z_n = i) &= \sum_{l \in \mathbb{S}} \mathbb{P}(Z_{n+k+1} = j \text{ and } Z_{n+k} = l \mid Z_n = i) \\ &= \sum_{l \in \mathbb{S}} \frac{\mathbb{P}(Z_{n+k+1} = j, \ Z_{n+k} = l, \ Z_n = i)}{\mathbb{P}(Z_n = i)} \\ &= \sum_{l \in \mathbb{S}} \frac{\mathbb{P}(Z_{n+k+1} = j, \ Z_{n+k} = l, \ Z_n = i)}{\mathbb{P}(Z_{n+k} = l \text{ and } Z_n = i)} \frac{\mathbb{P}(Z_{n+k} = l \text{ and } Z_n = i)}{\mathbb{P}(Z_n = i)} \\ &= \sum_{l \in \mathbb{S}} \mathbb{P}(Z_{n+k+1} = j \mid Z_{n+k} = l \text{ and } Z_n = i) \mathbb{P}(Z_{n+k} = l \mid Z_n = i) \\ &= \sum_{l \in \mathbb{S}} \mathbb{P}(Z_{n+k+1} = j \mid Z_{n+k} = l) \mathbb{P}(Z_{n+k} = l \mid Z_n = i) \\ &= \sum_{l \in \mathbb{S}} \mathbb{P}(Z_{n+k+1} = l \mid Z_n = i) \mathbb{P}(Z_{n+k} = l \mid Z_n = i) \end{split}$$

We have just checked that the family of matrices

$$\left[\ \mathbb{P}(Z_{n+k}=j \mid Z_n=i) \
ight]_{i,j\in\mathbb{S}}, \qquad k \geqslant 1,$$

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 $k \ge 0$ the equality

$$\left[\ \mathbb{P}(Z_{n+k}=j \mid Z_n=i) \
ight]_{i,j\in\mathbb{S}} \ = \ \left[\ [P^k]_{i,j} \
ight]_{i,j\in\mathbb{S}} = P^k$$

holds not only for k = 0 and k = 1, but also for all $k \in \mathbb{N}$.

The matrix product relation

$$P^{m+n} = P^m P^n = P^n P^m,$$
 MH3512 AY19-20

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

$$egin{array}{rl} \mathbb{P}(Z_{n+m}=j\mid Z_0=i) &=& \sum\limits_{l\in\mathbb{S}} \mathbb{P}(Z_m=j\mid Z_0=l) \mathbb{P}(Z_n=l\mid Z_0=i) \ &=& \sum\limits_{l\in\mathbb{S}} \mathbb{P}(Z_n=j\mid Z_0=l) \mathbb{P}(Z_m=l\mid Z_0=i), \end{array}$$

 $i, j \in S$, which is called the *Chapman-Kolmogorov* equation, cf. also the triple (1.2.2).

Example. The gambling process $(Z_n)_{n \ge 0}$.

Taking S = 4 and p = 40%, the transition matrix of the gambling process on $\$ = \{0, 1, \dots, 4\}$ of Chapter 2 reads

$$P = [P_{i,j}]_{0 \le i,j \le 4} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.6 & 0 & 0.4 & 0 & 0 \\ 0 & 0.6 & 0 & 0.4 & 0 \\ 0 & 0 & 0.6 & 0 & 0.4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$
(4.4.1)

and we can check by hand that:

 $P^2 = P imes P$

 $\circ 158$

	$\begin{bmatrix} 1 \end{bmatrix}$	0	0	0	0		$\begin{bmatrix} 1 \end{bmatrix}$	0	0	0	0		1	0	0	0	0	
	0.6	0	0.4	0	0		0.6	0	0.4	0	0		0.6	0.24	0	0.16	0	
=	0	0.6	0	0.4	0	X	0	0.6	0	0.4	0	=	0.36	0	0.48	0	0.16	.
	0	0	0.6	0	0.4		0	0	0.6	0	0.4		0	0.36	0	0.24	0.4	
	0	0	0	0	1		0	0	0	0	1		0	0	0	0	1	

() 159

Exercise: From the above matrix (4.4.1), check that

 $\mathbb{P}(Z_2 = 4 \mid Z_0 = 2) = [P^2]_{2,4} = 0.16,$ $\mathbb{P}(Z_2 = 1 \mid Z_0 = 2) = [P^2]_{2,1} = 0,$ and

$$\mathbb{P}(Z_2=2 \mid Z_0=2) = [P^2]_{2,2} = 0.48.$$

Example. The fifth order transitions of the chain with Markov matrix (4.3.4) can be computed from the fifth matrix power

$$P^5 = \begin{bmatrix} 0.14352 & 0.09600 & 0.25920 & 0.30160 & 0.19968 \\ 0.15840 & 0.10608 & 0.24192 & 0.30400 & 0.18960 \\ 0.17040 & 0.10920 & 0.23280 & 0.30880 & 0.17880 \\ 0.17664 & 0.11520 & 0.22800 & 0.30928 & 0.17088 \\ 0.14904 & 0.09600 & 0.25440 & 0.30520 & 0.19536 \end{bmatrix},$$

cf. e.g. here. Note that for large transition orders (for example 1000 time steps) we get

$$P^{1000} = \begin{bmatrix} 0.16273 & 0.10613 & 0.24056 & 0.30660 & 0.18396 \\ 0.16273 & 0.10613 & 0.24056 & 0.30660 & 0.18396 \\ 0.16273 & 0.10613 & 0.24056 & 0.30660 & 0.18396 \\ 0.16273 & 0.10613 & 0.24056 & 0.30660 & 0.18396 \\ 0.16273 & 0.10613 & 0.24056 & 0.30660 & 0.18396 \end{bmatrix}$$

cf. e.g. here, which suggests a convergence phenomenon in large time for the Markov chain, see Chapter 7 for details.

Example. For the simple random walk of Chapter 3, computing the probability to travel from (1) to 2k = 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 $[P^{20}]_{0,10}$, cf. also Figure 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.* $S = \{0, 1\}$.

To close this chapter we provide a complete study of the two-state Markov chain, whose transition matrix has the form

$$P = \begin{bmatrix} 1-a & a \\ & & \\ b & 1-b \end{bmatrix},$$
(4.5.1)

,

with $a \in [0, 1]$ and $b \in [0, 1]$.



We also have

$$\mathbb{P}(Z_{n+1}=1 \mid Z_n=0)=a, \quad \ \ \mathbb{P}(Z_{n+1}=0 \mid Z_n=0)=1-a,$$

and

$$\mathbb{P}(Z_{n+1}=0 \mid Z_n=1)=b, \quad \mathbb{P}(Z_{n+1}=1 \mid Z_n=1)=1-b.$$

The power

$$P^n = \left[\begin{array}{rrr} 1-a & a \\ & & \\ b & 1-b \end{array} \right]^n$$

of the transition matrix P is computed for all $n \ge 0$ in the next Proposition 4.2. We always exclude the case a = b = 0 since it corresponds to the trivial case where $P = I_d$ is the identity matrix (constant chain).

Proposition 4.2
We have
$$P^{n} = \frac{1}{a+b} \begin{bmatrix} b+a(1-a-b)^{n} & a(1-(1-a-b)^{n}) \\ b(1-(1-a-b)^{n}) & a+b(1-a-b)^{n} \end{bmatrix}, \quad n \in \mathbb{N}.$$

Proof

This result will be proved by a *diagonalization* argument. The matrix \boldsymbol{P} has two eigenvectors^{*a*}

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
 and $\begin{bmatrix} -a \\ b \end{bmatrix}$

with respective eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 1 - a - b$, see for example the commands Eigensystem[1-a,a,b,1-b] and Diagonalize[1-a,a,b,1-b].

Hence \boldsymbol{P} can be written in the *diagonal form*

$$P = M \times D \times M^{-1}, \tag{4.5.2}$$

i.e.

$$P = egin{bmatrix} 1 & -a \ 1 & b \end{bmatrix} imes egin{bmatrix} \lambda_1 & 0 \ 0 & \lambda_2 \end{bmatrix} imes egin{bmatrix} rac{b}{a+b} & rac{a}{a+b} \ -rac{1}{a+b} & rac{1}{a+b} \end{bmatrix}$$

As a consequence of (4.5.2), we have

$$egin{array}{rcl} P^n &=& (M imes D imes M^{-1})^n = (M imes D imes M^{-1})\cdots (M imes D imes M^{-1})\ &=& M imes D imes \cdots imes D imes M^{-1} = M imes D^n imes M^{-1}, \end{array}$$

where

$$D^n = egin{bmatrix} 1 & 0 \ 0 & \lambda_2^n \end{bmatrix}, \qquad n \in \mathbb{N}.$$

hence

$$P^n = egin{bmatrix} 1 & -a \ 1 & b \end{bmatrix} imes egin{bmatrix} 1 & 0 \ 0 & \lambda_2^n \end{bmatrix} imes egin{bmatrix} rac{b}{a+b} & rac{a}{a+b} \ -rac{1}{a+b} & rac{1}{a+b} \end{bmatrix} \ = rac{1}{a+b} egin{bmatrix} b & a \ b & a \end{bmatrix} + rac{\lambda_2^n}{a+b} egin{bmatrix} a & -a \ -b & b \end{bmatrix} \ = rac{1}{a+b} egin{bmatrix} b+a\lambda_2^n & a(1-\lambda_2^n) \ b(1-\lambda_2^n) & a+b\lambda_2^n \end{bmatrix},$$

see also the command MatrixPower[1-a,a,b,1-b,n].

^aPlease refer to MH1201 - Linear Algebra II for more on "eigenvectors, eigenvalues, diagonalization".

(4.5.3)

î D

For an alternative proof of Proposition 4.2, see also Exercise 1.4.1 page 5 of [Nor98] in which \mathbf{P}^n is written as

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 7.18 for a similar analysis.

From the result of Proposition 4.2 we may now compute the probabilities

$$\mathbb{P}(Z_n = 0 \mid Z_0 = 0) = \frac{b + a\lambda_2^n}{a + b}, \quad \mathbb{P}(Z_n = 1 \mid Z_0 = 0) = \frac{a(1 - \lambda_2^n)}{a + b}$$
(4.5.4)

and

$$\mathbb{P}(Z_n = 0 \mid Z_0 = 1) = \frac{b(1 - \lambda_2^n)}{a + b}, \quad \mathbb{P}(Z_n = 1 \mid Z_0 = 1) = \frac{a + b\lambda_2^n}{a + b}.$$
(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.*

$$\mathbb{P}(Z_3=0 \mid Z_0=0) = rac{b+a(1-a-b)^3}{a+b},$$

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:

$$\lim_{n \to \infty} P^n = \lim_{n \to \infty} \left[\begin{array}{cc} \mathbb{P}(Z_n = 0 \mid Z_0 = 0) & \mathbb{P}(Z_n = 1 \mid Z_0 = 0) \\ \mathbb{P}(Z_n = 0 \mid Z_0 = 1) & \mathbb{P}(Z_n = 1 \mid Z_0 = 1) \end{array} \right] = \frac{1}{a+b} \left[\begin{array}{c} b & a \\ b & a \end{array} \right].$$

Note that convergence will be faster when a + b is closer to 1.

MH3512 AY19-20

♦163

ॳ

Hence we have

$$\lim_{n \to \infty} \mathbb{P}(Z_n = 1 \mid Z_0 = 0) = \lim_{n \to \infty} \mathbb{P}(Z_n = 1 \mid Z_0 = 1) = \frac{a}{a+b}$$
(4.5.6)

and

$$\lim_{n \to \infty} \mathbb{P}(Z_n = 0 \mid Z_0 = 0) = \lim_{n \to \infty} \mathbb{P}(Z_n = 0 \mid Z_0 = 1) = \frac{b}{a+b}.$$
(4.5.7)

Consequently,

$$\pi = [\pi_0, \pi_1] := \left[\frac{b}{a+b}, \frac{a}{a+b}\right]$$
(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

the chain is constant and it clearly admits its initial distribution as limiting distribution. In case a = b = 1, we have

$$P = \left[egin{array}{cc} 0 & 1 \ & \ 1 & 0 \end{array}
ight],$$

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 Chapter 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

$$\begin{aligned} \pi P &= \frac{1}{a+b} \begin{bmatrix} b,a \end{bmatrix} \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = \frac{1}{a+b} \begin{bmatrix} b(1-a)+ab \\ ab+a(1-b) \end{bmatrix}^\top \\ &= \frac{1}{a+b} \begin{bmatrix} b,a \end{bmatrix} = \pi, \end{aligned}$$

i.e. $\boldsymbol{\pi}$ is invariant (or stationary) with respect to \boldsymbol{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

$$P = \left[egin{array}{cc} 0 & 1 \ & \ & \ 1 & 0 \end{array}
ight]$$

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 in (B.29), 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 $\{0, 1\}$, cf. (B.29) in the general case.

- 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*, see also *here*. 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

and the chain is indefinitely switching at each time step from one state to the other.

In Figures 4.3 and 4.4 we consider a simulation of the two-state random walk with transition matrix

$$P = \left[egin{array}{cc} 0.8 & 0.2 \ & \ 0.4 & 0.6 \end{array}
ight],$$

i.e. a = 0.2 and b = 0.4. Figure 4.3 represents a sample path $(x_n)_{n=0,1,\dots,100}$ of the chain, while Figure 4.4 represents the sample of the chain, while Figure 4.4 represents the sample of the chain.

average

$$y_n = rac{1}{n+1}(x_0+x_1+\dots+x_n), \qquad n=0,1,\dots,100,$$

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 Chapter 7.



Figure 4.3: Sample path of a two-state chain in continuous time with a = 0.2 and b = 0.4.



Figure 4.4: The proportion of chain values at state 1 tends to 1/3 = a/(a + b).

The source code of the above program written in ${\bf R}$ is given below.*

^{*}Download the corresponding **R** code or the **IPython notebook** that can be run *here*.

```
() 168
```

```
# Dimension of the transition matrix
dim=2
# Parameter definition
a=0.2; b=0.4;
# Definition of the transition matrix
P=matrix(c(1-a,a,b,1-b),nrow=dim,ncol=dim,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(dim,size=1,prob=P[2,])
# Random simulation of Z[j+1] given Z[j]
for (j in seq(1,N)) Z[j+1]=sample(dim,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...")
}
```

We close this chapter with two other sample paths of Markov chains in Figures 4.5 and 4.6 (use the controls to see the animations under Acrobat Reader). In the next Figure 4.5 we check again that the proportion of chain values in the state (1) converges to 1/3 for a two-state Markov chain.

Figure 4.5: Convergence graph for the two-state Markov chain with a = 0.2 and b = 0.4.*

In Figure 4.6 we draw a sample path of a five-state Markov chain.*

Figure 4.6: 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?

^{*}Download the corresponding **R code** or the **IPython notebook** that can be run *here*. [†]Animated figure (works in Acrobat Reader).

Exercise 4.2 Consider the Markov chain $(Z_n)_{n \ge 0}$ with state space $\mathbb{S} = \{1, 2\}$ and transition matrix

$$P = egin{array}{ccc} 1 & 2 \ 0.4 & 0.6 \ 0.8 & 0.2 \end{bmatrix}.$$

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 (B.28), *i.e.*

$$P = \left[\begin{array}{cccc} P_{i,j} \end{array}
ight]_{0\leqslant i,j\leqslant N} = egin{bmatrix} \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \ \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \ arphi & dots & d$$

where $\pi = [\pi_0, \pi_1, \dots, \pi_N] \in [0, 1]^{N+1}$ is a vector such that

$$\pi_0+\pi_1+\cdots+\pi_N=1.$$

- a) Compute the matrix power P^n for all $n \ge 2$.
- b) Show that the vector $\boldsymbol{\pi}$ is an invariant (or stationary) distribution for \boldsymbol{P} .
- c) Show that if $\mathbb{P}(Z_0 = i) = \pi_i$, i = 0, 1, ..., N, then Z_n is independent of Z_k for all $0 \leq k < n$, and $(Z_n)_{n \in \mathbb{N}}$ is an *i.i.d* sequence of random variables with distribution $\pi = [\pi_0, \pi_1, ..., \pi_N]$ over $\{0, 1, ..., N\}$.

Exercise 4.4 Consider a $\{0, 1\}$ -valued "hidden" two-state Markov chain $(X_n)_{n \in \mathbb{N}}$ with transition probability matrix

$$P = egin{bmatrix} P_{0,0} & P_{0,1} \ & P_{1,0} & P_{1,1} \end{bmatrix} = egin{bmatrix} \mathbb{P}(X_1 = 0 \mid X_0 = 0) & \mathbb{P}(X_1 = 1 \mid X_0 = 0) \ & \mathbb{P}(X_1 = 0 \mid X_0 = 1) & \mathbb{P}(X_1 = 1 \mid X_0 = 1) \end{bmatrix},$$

and initial distribution

$$\pi = [\pi_0, \pi_1] = [\mathbb{P}(X_0 = 0), \mathbb{P}(X_0 = 1)].$$

We observe a two-state 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

$$M = \left[egin{array}{cc} m_{0,a} & m_{0,b} \ m_{1,a} & m_{1,b} \end{array}
ight] = \left[egin{array}{cc} \mathbb{P}(O_k = a \mid X_k = 0) & \mathbb{P}(O_k = b \mid X_k = 0) \ \mathbb{P}(O_k = a \mid X_k = 1) & \mathbb{P}(O_k = b \mid X_k = 1) \end{array}
ight],$$

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

$$\mathbb{P}(X_1 = 1 \mid (O_0, O_1) = (a, b))$$
 and $\mathbb{P}(X_1 = 0 \mid (O_0, O_1) = (a, b))$

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 you 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 Z-valued random walk

$$S_n := X_1 + \dots + X_n, \qquad n \in \mathbb{N},$$

whose increments $X_k = S_k - S_{k-1}, k \ge 1$, are recursively defined as follows:

- At time $n = 1, X_1$ is a Bernoulli $\{-1, +1\}$ -valued random variable with

$$\mathbb{P}(X_1 = +1) = p$$
 and $\mathbb{P}(X_1 = -1) = q = 1 - p \in (0, 1).$

- At any subsequent time $n \ge 2$, one randomly draws an integer time index $k \in \{1, 2, \dots, n-1\}$ 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 two-state Markov chain $(X_n)_{n \ge 0}$ with state space $\mathbb{S} = \{0, 1\}$ and transition matrix

$$P = egin{array}{ccc} 0 & 1 \ 1 - a & a \ 1 & b & 1 - b \end{bmatrix},$$

where $a, b \in (0, 1)$, and define a new stochastic process $(Z_n)_{n \ge 1}$ by $Z_n = (X_{n-1}, X_n)$, $n \ge 1$. Argue that $(Z_n)_{n \ge 1}$ is a Markov chain and write down its transition matrix. Start by determining the state space of $(Z_n)_{n \ge 1}$. See Problem 5.23 for the use of this construction in pattern recognition.

Exercise 4.8 Given $p \in [0,1)$, consider the Markov chain $(X_n)_{n \ge 0}$ on the state space $\{0,1,2\}$ having the transition matrix

$$P = egin{array}{cccc} 0 & 1 & 2 \ 0 & p & q & 0 \ 1 & 0 & p & q \ 2 & 0 & 0 & 1 \end{array}
ight
ceil, MH3512 \, AY19-20$$

$$P = \begin{bmatrix} 0 & 1 \\ 1 & a & a \\ 0 & 1 & b \end{bmatrix},$$

with q := 1 - p.

a) Give the probability distribution of the first hitting time

$$T_2:=\infig\{n\geqslant 0\ :\ X_n=2ig\}.$$

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

$$\mathbb{P}(Z=k\mid X_0=1)=inom{k-1}{k-d}(1-p)^dp^{k-d},\qquad k\geqslant d.$$

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

$$\sum_{k \ge 1} k p^{k-1} = \frac{1}{(1-p)^2} \quad \text{and} \quad \sum_{k \ge 2} k(k-1) p^{k-2} = \frac{2}{(1-p)^3}, \qquad 0 \le p < 1$$

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 <u>quit</u> 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 $n \ge 1$.
- 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 $n \ge 2$.

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

$$Z_n := \left\{egin{array}{ll} 0 & ext{if } X_n = 0 ext{ or } X_n = 1, \ X_n & ext{if } X_n = 2 ext{ or } X_n = 3, \end{array}
ight.$$

i.e.

$$Z_n=X_n\mathbb{1}_{\{X_n\in\{2,3\}\}},\qquad n\geqslant 0.$$

a) Compute

$$\mathbb{P}(Z_{n+1} = 2 \mid Z_n = 0 \text{ and } Z_{n-1} = 2)$$

and

$$\mathbb{P}(Z_{n+1} = 2 \mid Z_n = 0 \text{ and } Z_{n-1} = 3),$$

 $n \geqslant 1$.

b) Is $(Z_n)_{n \in \mathbb{N}}$ a Markov chain?

0177

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.



Figure 4.7: Lafenwa intersection.

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 $\pi_L = 11/24$.

Problem 4.13 Phase-type distributions are used in insurance to model the heavy-tailed random claim sizes appearing in reserve and surplus processes.

- a) Given $p \in [0, 1]$, construct the transition matrix of a two-state Markov chain on the state space $\{0, 1\}$ satisfying the following two conditions:
 - i) State (1) is absorbing, i.e. $\mathbb{P}(X_{n+1} = 0 \mid X_n = 0) = 1$, and
 - ii) The first hitting time

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

of state (1) starting from state (1) has the geometric distribution p given by

$$\mathbb{P}(T_0=k\mid X_0=1)=(1-p)p^{k-1}, \qquad k\geqslant 1.$$

- b) Given $p \in [0,1]$ and $d \ge 1$, construct the transition matrix of a d + 1-state Markov chain on the state space $\{0, 1, \ldots, d\}$ satisfying the following two conditions:
 - i) State (1) is absorbing, i.e. $\mathbb{P}(X_{k+1} = 0 \mid X_k = 0) = 1$, and MH3512 AY19-20

ii) The first hitting time T_0 of state 0 starting from state d has the negative binomial distribution

$$\mathbb{P}(T_0=k\mid X_0=d)=inom{k-1}{k-d}(1-p)^dp^{k-d},\qquad k\geqslant d.$$

c) In the sequel we consider a discrete-time Markov chain on $\{0, 1, \ldots, d\}$ having d transient^{*} states $\{1, 2, \ldots, d\}$, and 0 as absorbing state. Show that its transition matrix P takes the form

$$P = \left[egin{array}{cccc} P_{i,j} \end{array}
ight]_{0\leqslant i,j\leqslant d} = egin{bmatrix} 1 & 0 & \cdots & 0 \ lpha_1 & Q_{1,1} & \cdots & Q_{1,d} \ lpha_2 & Q_{2,1} & \cdots & Q_{2,d} \ dots & dots & \ddots & dots \ lpha & Q_{d,1} & \cdots & Q_{d,d} \end{bmatrix} = egin{bmatrix} 1 & 0 \ lpha & Q \end{bmatrix},$$

where α is the column vector $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_d]^{\top}$ and Q is the $d \times d$ matrix

$$Q = egin{bmatrix} Q_{1,1} & \cdots & Q_{1,d} \ dots & \ddots & dots \ Q_{d,1} & \cdots & Q_{d,d} \end{bmatrix},$$

and give the relation between the vector $\boldsymbol{\alpha}$ and the matrix \boldsymbol{Q} using the column vector $\boldsymbol{e} := [1, 1, \dots, 1]^{\top}$.

d) Show by induction on $n \ge 0$ that we have

$$P^{n} = \begin{bmatrix} 1 & 0\\ (I_{d} - Q^{n})e & Q^{n} \end{bmatrix}, \qquad (4.5.9)$$

^{*}Here the transience condition implies that $\mathbb{P}(T_0 < \infty \mid X_0 = i) = 1$ for all i = 1, 2, ..., d, it will be ensured by assuming that $I_d - Q$ is invertible, cf. § 5.4 and § 6.3 for details Y19-20

where I_d is the $d \times d$ identity matrix

$$I_d := egin{bmatrix} 1 & 0 & \cdots & 0 & 0 \ 0 & 1 & \ddots & 0 & 0 \ dots & dots & \ddots & dots & dots \ 0 & 0 & \cdots & 1 & 0 \ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

e) Compute the probability distribution of T_0 using the vector α and the matrix Q^{n-1} , where T_0 is the first hitting time of state (0) starting from state $i \ge 1$.

Hint: Partition the event $\{T_0 = n\}$ as

$$\{T_0=n\}=igcup_{k=1}^d\{X_n=0 ext{ and }X_{n-1}=k\}.$$

f) From now on we assume that the initial distribution of X_0 is given by the *d*-dimensional vector

$$eta = egin{bmatrix} eta_1 \ eta_2 \ dots \ eta_d \end{bmatrix},$$

i.e. $\beta_i = \mathbb{P}(X_0 = i), i = 1, 2, \dots, d$, with $\mathbb{P}(X_0 = 0) = 0$. Compute the probability distribution of T_0 using the vectors α, β , and the matrix Q^{n-1} .

- g) Compute the cumulative distribution function $P(T_0 \leq n)$ of T_0 the vectors β , e, and the matrix Q^n .
- h) Recover the result of Question (g) by rewriting $\mathbb{P}(T_0 \leq n)$ as the probability of not being in any state $i = 1, 2, \ldots, d$ at time n.
- i) Compute the probability generating function

$$G_{T_0}(s):=\sum\limits_{k \geqslant 0} s^k \mathbb{P}(T_0=k)$$
MH3512 AY19-2
of T_0 , assuming the existence of the matrix inverse $(I_d - sQ)^{-1}$ given by the series

$$(I_d-sQ)^{-1}=\sum\limits_{k\geqslant 0}s^kQ^k,\qquad s\in(-1,1].$$

j) Using the probability generating function $s \mapsto G_{T_0}(s)$, compute the first and second moments $E[T_0]$ and $E[T_0^2]$ of T_0 .

Problem 4.14 Hidden Markov models have applications to speech recognition, face recognition, emotion recognition, genomics and biological sequence analysis, sentiment analysis, unsupervised learning, climate change studies, etc. Consider a $\{0, 1\}$ -valued "hidden"



two-state Markov chain $(X_n)_{n\in\mathbb{N}}$ with transition probability matrix

$$P = egin{bmatrix} P_{0,0} & P_{0,1} \ & P_{1,0} & P_{1,1} \end{bmatrix} = egin{bmatrix} \mathbb{P}(X_1 = 0 \mid X_0 = 0) & \mathbb{P}(X_1 = 1 \mid X_0 = 0) \ & \mathbb{P}(X_1 = 0 \mid X_0 = 1) & \mathbb{P}(X_1 = 1 \mid X_0 = 1) \end{bmatrix},$$

and initial distribution

$$\pi = [\pi_0, \pi_1] = [\mathbb{P}(X_0 = 0), \mathbb{P}(X_0 = 1)].$$

We observe a stochastic process $(O_k)_{k\in\mathbb{N}}$ whose state $O_k \in \{a, b, c\}$ at every time $k \in \mathbb{N}$ has a conditional distribution given $X_k \in \{0, 1\}$ denoted by

$$M = egin{bmatrix} m_{0,a} & m_{0,b} & m_{0,c} \ m_{1,a} & m_{1,b} & m_{1,c} \end{bmatrix}$$

() 181

$$= \begin{bmatrix} \mathbb{P}(O_k = a \mid X_k = 0) & \mathbb{P}(O_k = b \mid X_k = 0) & \mathbb{P}(O_k = c \mid X_k = 0) \\ \mathbb{P}(O_k = a \mid X_k = 1) & \mathbb{P}(O_k = b \mid X_k = 1) & \mathbb{P}(O_k = c \mid X_k = 1) \end{bmatrix},$$
(5.182)

called the *emission probability* matrix.

a) Using elements of π , P and M, compute $\mathbb{P}(X_0 = 1, X_1 = 1, X_2 = 0)$ and the probability

$$\mathbb{P}((O_0, O_1, O_2) = (c, a, b) \text{ and } (X_0, X_1, X_2) = (1, 1, 0))$$

of observing the sequence $(O_0, O_1, O_2) = (c, a, b)$ when $(X_0, X_1, X_2) = (1, 1, 0)$.

Hint: By independence, the conditional probability of observing $(O_0, O_1, O_2) = (c, a, b)$ given that $(X_0, X_1, X_2) = (1, 1, 0)$ splits as

$$\mathbb{P}((O_0,O_1,O_2)=(c,a,b)\mid (X_0,X_1,X_2)=(1,1,0))\ =\ \mathbb{P}(O_0=c\mid X_0=1)\mathbb{P}(O_1=a\mid X_1=1)\mathbb{P}(O_2=b\mid X_2=0)$$

b) Find the probability $\mathbb{P}((O_0, O_1, O_2) = (c, a, b))$ that the observed sequence is (c, a, b).

Hint: Use the law of total probability based on all possible values of (X_0, X_1, X_2) .

c) Compute the probabilities

$$\mathbb{P}(X_1 = 1 \mid (O_0, O_1, O_2) = (c, a, b)), \text{ and } \mathbb{P}(X_1 = 0 \mid (O_0, O_1, O_2) = (c, a, b)).$$

From Question (d) to Question (i) below we take $\pi = [\pi_0, \pi_1] := [0.6, 0.4]$ and

$$P := \begin{bmatrix} 0.7 & 0.3 \\ \\ 0.4 & 0.6 \end{bmatrix}, \quad M := \begin{bmatrix} m_{0,a} & m_{0,b} & m_{0,c} \\ m_{1,a} & m_{1,b} & m_{1,c} \end{bmatrix} = \begin{bmatrix} 0.1 & 0.4 & 0.5 \\ 0.7 & 0.2 & 0.1 \end{bmatrix}.$$
(4.5.10)



$$\mathbb{P}((X_0, X_1, X_2) = (x, y, z) ext{ and } (O_0, O_1, O_2) = (c, a, b))$$

for all $x, y, z \in \{0, 1\}$, and identify the most likely sample sequence of values for (X_0, X_1, X_2) .

e) Compute¹ the probability $\mathbb{P}((O_0, O_1, O_2) = (c, a, b))$ that the observed sequence is (c, a, b).

f) Compute¹ the six probabilities

$$\mathbb{P}(X_k=1 \mid (O_0,O_1,O_2)=(c,a,b)), \ \ \mathbb{P}(X_k=0 \mid (O_0,O_1,O_2)=(c,a,b)),$$

k = 0, 1, 2. What is the most likely sequence for (X_0, X_1, X_2) according to this computation?

Based on the knowledge of $\mathbb{P}((X_0, X_1, X_2) = (1, 1, 0) \mid (O_0, O_1, O_2) = (c, a, b))$, one can build an estimator \hat{P} , $\hat{\pi}$, \hat{M} of the model parameters P, π and M, as

$$\begin{cases} \hat{\pi}_{i} := \mathbb{P}(X_{0} = i \mid (O_{0}, O_{1}, O_{2}) = (c, a, b)), \quad (4.5.11a) \\ \\ \hat{P}_{i,j} := \frac{\sum_{t=0}^{N-1} \mathbb{P}(X_{t} = i, X_{t+1} = j \mid (O_{0}, O_{1}, O_{2}) = (c, a, b))) \\ \\ \\ \hat{M}_{i,k} := \frac{\sum_{t=0}^{N-1} \mathbb{P}(X_{t} = i \mid (O_{0}, O_{1}, O_{2}) = (c, a, b))) \\ \\ \\ \hat{M}_{i,k} := \frac{\sum_{t=0}^{N} \mathbb{1}_{\{O_{t} = k\}} \mathbb{P}(X_{t} = i \mid (O_{0}, O_{1}, O_{2}) = (c, a, b))}{\sum_{t=0}^{N} \mathbb{P}(X_{t} = i \mid (O_{0}, O_{1}, O_{2}) = (c, a, b))}, \quad (4.5.11c) \end{cases}$$

with here N = 2. Note that (4.5.11c) averages the number of times the observed state is "k" given that the hidden state is "i", which gives an estimate of the conditional probability $M_{i,k}$.

^{*}Give a numerical value with exactly three significant digits.

g) Compute^{*} the vector estimate $\hat{\pi} = [\hat{\pi}_0, \hat{\pi}_1]$ using (4.5.11a) and the data of Equation (4.5.10).

- h) Compute¹ the matrix estimate \hat{P} using (4.5.11b) and the data of Equation (4.5.10).
- i) Compute¹ the matrix estimate \hat{M} using (4.5.11c) and the data of Equation (4.5.10).
- j) Iterating the estimates (4.5.11a)-(4.5.11c) is computationally intensive, however this procedure admits an efficient recursive implementation via the Baum-Welch algorithm which is based of the EM algorithm.

Imagine an alien trying to analyse an English manuscript without any prior knowledge of English. Using a simple two-state hidden chain $(X_n)_{n \in \mathbb{N}}$ he will try to uncover some *features* of the language, starting with a *binary classification* of the alphabet.

```
install.packages("HMM")
library (HMM)
library (lattice)
text = readChar("my_own_text_file.txt",nchars=10000)
data <- unlist (strsplit (gsub ("[^a-z]", "_", tolower (text)), ""))</pre>
pi = c(0.4, 0.6)
P=t(matrix(c(c(0.6177499, 0.3822501), c(0.8826096, 0.1173904)), nrow=2, ncol=2))
M = t(matrix(c(c(0.037192964, 0.009902360, 0.032833978, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057331132, 0.044882670, 0.057380, 0.044882670, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.05780, 0.057800, 0.057800, 0.057800, 0.057800, 0.057800, 0.057800, 0.057800, 0.057800, 0.057800, 0.05780
0.052143890,0.013665015,0.036187536,0.072293323,0.044793972,0.060008388,
0.004256270,0.024770706,0.053520546,0.014232306,0.046981769,0.053733382,
0.066355203,0.046817817,0.006912535,0.016201697,0.013425499,0.024694447,
0.064902148,0.046170421,0.033586536,0.022203489),
0.0211687193,0.0625229746,0.0039632091,0.0567828864,0.0468108656,0.0168355418,
0.0627882213,0.0286478204,0.0389215263,0.0064318198,0.0001698078,0.0493758725,
0.0652709152,0.0069580806,0.0093043072,0.0028807932,0.0521827110,0.0608822385,
0.0645417465,0.0555249876,0.0576888424)),nrow=27,ncol=2))
model <- initHMM (c("0", "1"),c("_", letters), pi, P, M)
system.time (estimate <- baumWelch (model, data, 100)) # 100 iterations</pre>
xyplot(estimate$hmm$emissionProbs[1,] ~ c(1:27), scales=list(x=list(at=1:27,
labels=c("_", letters))),type="h", lwd=5, xlab="", ylab="")
```

 $^{^{*}\}mathrm{Give}$ a numerical value with exactly three significant digits.

Using a text of your choice and a $\{0, 1\}$ -valued hidden Markov chain $(X_n)_{n \in \mathbb{N}}$, estimate the corresponding emission probability matrix M by running the above R code which relies on the HMM (Hidden Markov Model) package. A text length of $N \simeq 10,000$ characters can be a minimum. The initial values of π , P and M have been set randomly.

What do you conclude?

Instructions:

Specify your choice of document used.

Attach a graph with a discussion of your output.

Suggestions and possible variations:

Try a language different from English.

Increase the state space of $(X_n)_{n \in \mathbb{N}}$ in order to uncover more features of the chosen language.

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 Chapters 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	187	
5.2	Mean Hitting and Absorption Times	195	
5.3	First Return Times	203	
5.4	Mean Number of Returns	211	
Exercises			

5.1 Hitting Probabilities

Let us consider a Markov chain $(Z_n)_{n \in \mathbb{N}}$ with state space S, and let $A \subset S$ denote a subset of S. We are interested in the first time T_A the chain hits the subset A, with

$$T_A := \inf\{n \ge 0 : Z_n \in A\},\tag{5.1.1}$$

with $T_A = 0$ if $Z_0 \in A$ and

$$T_A=+\infty$$
 if $\{n\geqslant 0 \; : \; Z_n\in A\}=\emptyset_2$

i.e. if $Z_n \notin A$ for all $n \in \mathbb{N}$. In case the transition matrix P satisfies

$$P_{k,l} = \mathbb{1}_{\{k=l\}}$$
 MH3512 AY19-20

for all $k, l \in A$, the set A is said to be absorbing.

Similarly to the gambling problem of Chapter 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 $A \subset S$ through state $l \in A$ starting from $k \in S$, 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.

Proposition 5.1
The hitting probabilities
$$g_{l}(k) := \mathbb{P}(Z_{T_{A}} = l \mid Z_{0} = k), \quad k \in S, \ l \in A,$$
satisfy the equation
$$g_{l}(k) = \sum_{m \in S} P_{k,m}g_{l}(m) = P_{k,l} + \sum_{m \in S \setminus A} P_{k,m}g_{l}(m), \quad (5.1.2)$$

$$k \in S \setminus A, \ l \in A, \text{ under the boundary conditions}$$

$$g_{l}(k) = \mathbb{P}(Z_{T_{A}} = l \mid Z_{0} = k) = \mathbb{1}_{\{k = l\}} = \begin{cases} 1 \text{ if } k = l, \\ 0 \text{ if } k \neq l, \end{cases} \quad k \in A, \ l \in S, \end{cases}$$
which hold since $T_{A} = 0$ whenever one starts from $Z_{0} \in A.$

Proof

For all $k \in \mathbb{S} \setminus A$ we have $T_A \ge 1$ given that $Z_0 = k$, hence we can write

$$\begin{array}{lll} g_l(k) &=& \mathbb{P}(Z_{T_A} = l \mid Z_0 = k) \\ &=& \sum\limits_{m \in \mathbb{S}} \mathbb{P}(Z_{T_A} = l \mid Z_1 = m \text{ and } Z_0 = k) \mathbb{P}(Z_1 = m \mid Z_0 = k) \\ &=& \sum\limits_{m \in \mathbb{S}} \mathbb{P}(Z_{T_A} = l \mid Z_1 = m) \mathbb{P}(Z_1 = m \mid Z_0 = k) \\ &=& \sum\limits_{m \in \mathbb{S}} P_{k,m} \mathbb{P}(Z_{T_A} = l \mid Z_1 = m) \\ &=& \sum\limits_{m \in \mathbb{S}} P_{k,m} \mathbb{P}(Z_{T_A} = l \mid Z_0 = m) \\ &=& \sum\limits_{m \in \mathbb{S}} P_{k,m} g_l(m), \qquad k \in \mathbb{S} \setminus A, \quad l \in A, \end{array}$$

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.

Equation (5.1.2) can be rewritten in matrix form as

$$g_l = Pg_l, \qquad l \in A, \tag{5.1.3}$$

where g_l is a column vector, *i.e.*

$$egin{aligned} & \left[egin{aligned} g_l(0) \ dots \ g_l(N) \end{array}
ight] = \left[egin{aligned} P_{0,0} & P_{0,1} & P_{0,2} & \cdots & P_{0,N} \ P_{1,0} & P_{1,1} & P_{1,2} & \cdots & P_{1,N} \ P_{2,0} & P_{2,1} & P_{2,2} & \cdots & P_{2,N} \ dots & dots &$$

MH3512 AY19-20

under the boundary condition

$$g_l(k) = \mathbb{P}(Z_{T_A} = l \mid Z_0 = k) = \mathbb{1}_{\{l\}}(k) = egin{cases} 1, & k = l, \ 0, & k
eq l, \ 0, & k
eq l, \end{cases}$$

for all $k \in A$ and $l \in S$.

In addition, the hitting probabilities $g_l(k) = \mathbb{P}(Z_{T_A} = l \mid Z_0 = k)$ satisfy the condition

$$1 = \mathbb{P}(T_A = \infty \mid Z_0 = k) + \sum_{l \in A} \mathbb{P}(Z_{T_A} = l \mid Z_0 = k) = \mathbb{P}(T_A = \infty \mid Z_0 = k) + \sum_{l \in A} g_l(k),$$
(5.1.4)

for all $k \in S$.

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

$$\mathbb{P}(T_0=\infty\mid Z_0=1)=0.2$$



More generally, if $f: A \longrightarrow \mathbb{R}$ is a function on the domain A, letting

$$g_A(k):=\mathbb{E}\left[f(Z_{T_A})\,ig|\, Z_0=k
ight]=\sum\limits_{l\in A}f(l)\mathbb{P}(Z_{T_A}=l\mid Z_0=k), \hspace{1em} k\in \mathbb{S},$$

by linearity we find the Dirichlet problem

$$(I_d - P)g_A = 0,$$

under the boundary condition

$$g_A(k)=f(k),\qquad k\in A_2$$

see e.g. Theorem 5.3 in [Pri08] for a continuous-time analog.

(190

The next lemma will be used in Chapter 8 on branching Processes.

Lemma 5.2

Assume that state (j) $\in \mathbb{S}$ is absorbing. Then for all (i) $\in \mathbb{S}$ we have

$$\mathbb{P}(T_j < \infty \mid Z_0 = i) = \lim_{n o \infty} \mathbb{P}(Z_n = j \mid Z_0 = i).$$

Proof

We have

$$\{T_j<\infty\}=igcup_{n\geqslant 1}\{Z_n=j\},$$

because the finiteness of T_j means that Z_n becomes equal to (j) for some $n \in \mathbb{N}$. In addition, since $(j) \in S$ 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 we have

$$\alpha_{1} = \mathbb{P}(T_{j} < \infty \mid Z_{0} = i) = \mathbb{P}\left(\bigcup_{n \ge 1} \{Z_{n} = j\} \mid Z_{0} = i\right)$$

$$= \mathbb{P}\left(\lim_{n \to \infty} \{Z_{n} = j\} \mid Z_{0} = i\right) = \lim_{n \to \infty} \mathbb{P}(Z_{n} = j \mid Z_{0} = i).$$
(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

$$P = \begin{bmatrix} Q & R \\ & \\ 0 & I_d \end{bmatrix}, \tag{5.1.6}$$
MH3512 AY19-20

where Q is a square $(r+1) \times (r+1)$ matrix, R is a $(r+1) \times (N-r)$ matrix, and

$$\mathbf{I}_{d} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

is the (N-r) imes (N-r) identity matrix, in which case the states in $\{r+1, r+2, \ldots, N\}$ 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

$$g_l(m) = \mathbb{1}_{\{m=l\}}, \quad l = 0, 1, \dots, N, \quad m = r+1, r+2, \dots, N,$$
 (5.1.7)

hence the equation (5.1.2) can be rewritten as

$$egin{aligned} g_l(k) &= \sum\limits_{m=0}^N P_{k,m} g_l(m) \ &= \sum\limits_{m=0}^r P_{k,m} g_l(m) + \sum\limits_{m=r+1}^N P_{k,m} g_l(m) \ &= \sum\limits_{m=0}^r P_{k,m} g_l(m) + P_{k,l} \ &= \sum\limits_{m=0}^r Q_{k,m} g_l(m) + R_{k,l}, \qquad k=0,1,\ldots,r, \quad l=r+1,\ldots,N, \end{aligned}$$

from (5.1.7) and since $P_{k,l} = R_{k,l}$, $k = 0, 1, \ldots, r$, $l = r + 1, \ldots, N$. Hence we have

$$g_l(k) = \sum\limits_{m=0}^r Q_{k,m} g_l(m) + R_{k,l}, \qquad k = 0, 1, \dots, r, \quad l = r+1, \dots, N.$$

1	Remar	k	5.	3
---	-------	---	----	---

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.

Example

Consider a Markov chain on $\{0,1,2,3\}$ with transition matrix of the form

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a & b & c & d \\ \alpha & \beta & \gamma & \eta \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$
 (5.1.8)

Let $A = \{0, 3\}$ denote the absorbing states of the chain, and let

$$T_{0,3} = \inf\{n \ge 0 \; : \; Z_n = 0 \text{ or } Z_n = 3\}$$

and compute the probabilities

$$g_0(k) = \mathbb{P}(Z_{T_{0,3}} = 0 \mid Z_0 = k)$$

of hitting state 0 first within $\{0,3\}$ starting from k = 0, 1, 2, 3. The chain has the following graph:



MH3512(51499)

Noting that (0) and (3) are absorbing states, and writing the relevant rows of the first step analysis matrix equation g = Pg, we have

$$egin{aligned} g_0(0) &= 1 \ g_0(1) &= a imes 1 + b g_0(1) + c g_0(2) + d imes 0 \ g_0(2) &= lpha imes 1 + eta g_0(1) + \gamma g_0(2) + \eta imes 0 \ g_0(3) &= 0, \end{aligned}$$

i.e.

$$egin{aligned} g_0(0) &= 1 \ g_0(1) &= a + b g_0(1) + c g_0(2) \ g_0(2) &= lpha + eta g_0(1) + \gamma g_0(2) \ g_0(3) &= 0, \end{aligned}$$

which has for solution

$$g_{0}(0) = 1$$

$$g_{0}(1) = \frac{c\alpha + a(1 - \gamma)}{(1 - b)(1 - \gamma) - c\beta}$$

$$g_{0}(2) = \frac{a\beta + \alpha(1 - b)}{(1 - b)(1 - \gamma) - c\beta}$$

$$g_{0}(3) = 0.$$

(5.1.10)

We have $g_l(0) = g_l(3) = 0$ for l = 1, 2, and by a similar analysis, letting

$$g_3(k):=\mathbb{P}(Z_{T_{0,3}}=3\mid Z_0=k), \qquad k=0,1,2,3,$$

we find

$$\left\{egin{array}{l} g_3(0)=0\ g_3(1)=rac{c\eta+d(1-\gamma)}{(1-b)(1-\gamma)-ceta}\ g_3(2)=rac{eta d+\eta(1-b)}{(1-b)(1-\gamma)-ceta}\ g_3(3)=1, \end{array}
ight.$$

and we note that

$$g_0(1) + g_3(1) = rac{clpha + a(1-\gamma)}{(1-b)(1-\gamma) - ceta} + rac{c\eta + d(1-\gamma)}{(1-b)(1-\gamma) - ceta} = 1,$$

since $\alpha + \eta = 1 - \gamma - \beta$ and a + d = 1 - b - c, and similarly

$$g_0(2) + g_3(2) = rac{aeta + lpha(1-b)}{(1-b)(1-\gamma) - ceta} + rac{eta d + \eta(1-b)}{(1-b)(1-\gamma) - ceta} = 1.$$

We also check that in case a = d and $\alpha = \eta$ we have

$$g_0(1)=rac{clpha+a(eta+2lpha)}{(c+2a)(eta+2lpha)-ceta}=rac{clpha+aeta+2alpha}{2clpha+2aeta+4alpha}=g_0(2)=rac{1}{2},$$

and

$$g_0(1) = g_3(1) = g_0(2) = g_3(2) = rac{1}{2}$$

Note that, letting

$$T_0:=\inf\{n\geqslant 0\ :\ Z_n=0\} \hspace{1mm} ext{and}\hspace{1mm} T_3:=\inf\{n\geqslant 0\ :\ Z_n=0\},$$

we also have

$$g_0(k) = \mathbb{P}(Z_{T_{0,3}} = 0 \mid Z_0 = k) = \mathbb{P}(T_0 < \infty \mid Z_0 = k)$$

and

$$g_3(k) = \mathbb{P}(Z_{T_{0,3}} = 3 \mid Z_0 = k) = \mathbb{P}(T_3 < \infty \mid Z_0 = k)$$

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 $A \subset S$ starting from a state $k \in S$. 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 $Z_0 = k \in A$, we have

$$h_A(k) = 0,$$
 for all $k \in A$

MH3512 AY19-20

○195

Proposition 5.4

The mean hitting times

$$h_A(k):=\mathbb{E}[T_A\mid Z_0=k], \qquad k\in\mathbb{S},$$

satisfy the equations

$$h_A(k) \;=\; 1 + \sum\limits_{l \in \mathbb{S}} P_{k,l} h_A(l) = 1 + \sum\limits_{l \in \mathbb{S} \setminus A} P_{k,l} h_A(l), \hspace{1em} k \in \mathbb{S} \setminus A,$$

under the boundary conditions

$$h_A(k) = \mathbb{E}[T_A \mid Z_0 = k] = 0, \qquad k \in A.$$

Proof

For all $k \in S \setminus A$, by first step analysis using the *law of total expectation* applied to the probability measure $\mathbb{P}(\cdot \mid Z_0 = l)$, and the Markov property we have

$$\begin{split} h_A(k) &= \mathbb{E}[T_A \mid Z_0 = k] \\ &= \sum_{l \in \mathbb{S}} \mathbb{E} \left[T_A \mathbb{1}_{\{Z_1 = l\}} \mid Z_0 = k \right] \\ &= \frac{1}{\mathbb{P}(Z_0 = k)} \sum_{l \in \mathbb{S}} \mathbb{E} \left[T_A \mathbb{1}_{\{Z_1 = l\}} \mathbb{1}_{\{Z_0 = k\}} \right] \\ &= \frac{1}{\mathbb{P}(Z_0 = k)} \sum_{l \in \mathbb{S}} \mathbb{E} \left[T_A \mathbb{1}_{\{Z_1 = l \text{ and } Z_0 = k\}} \right] \\ &= \sum_{l \in \mathbb{S}} \mathbb{E} \left[T_A \mid Z_1 = l \text{ and } Z_0 = k \right] \frac{\mathbb{P}(Z_1 = l \text{ and } Z_0 = k)}{\mathbb{P}(Z_0 = k)} \\ &= \sum_{l \in \mathbb{S}} \mathbb{E} [T_A \mid Z_1 = l \text{ and } Z_0 = k] \mathbb{P}(Z_1 = l \mid Z_0 = k) \\ &= \sum_{l \in \mathbb{S}} \mathbb{E} [1 + T_A \mid Z_0 = l] \mathbb{P}(Z_1 = l \mid Z_0 = k) \end{split}$$

MH3512 AY19-20

<u>()</u> 196

$$\begin{split} &= \sum_{l \in \mathbb{S}} (1 + \mathbb{E}[T_A \mid Z_0 = l]) \mathbb{P}(Z_1 = l \mid Z_0 = k) \\ &= \sum_{l \in \mathbb{S}} \mathbb{P}(Z_1 = l \mid Z_0 = k) + \sum_{l \in \mathbb{S}} \mathbb{P}(Z_1 = l \mid Z_0 = k) \mathbb{E}[T_A \mid Z_0 = l] \\ &= 1 + \sum_{l \in \mathbb{S}} \mathbb{P}(Z_1 = l \mid Z_0 = k) \mathbb{E}[T_A \mid Z_0 = l] \\ &= 1 + \sum_{l \in \mathbb{S}} P_{k,l} h_A(l), \quad k \in \mathbb{S} \setminus A, \end{split}$$

with the relation

$$\mathbb{E}[T_A \mid Z_1 = l, \ Z_0 = k] = 1 + \mathbb{E}[T_A \mid Z_0 = l].$$

Hence we have

$$h_A(k) = 1 + \sum_{l \in \mathbb{S}} P_{k,l} h_A(l), \quad k \in \mathbb{S} \setminus A,$$
(5.2.1)

under the boundary conditions

$$h_A(k) = \mathbb{E}[T_A \mid Z_0 = k] = 0, \quad k \in A,$$
 (5.2.2)

the Condition (5.2.2) implies that (5.2.1) becomes

$$h_A(k) = 1 + \sum\limits_{l \in \mathbb{S} ackslash A} P_{k,l} h_A(l), \qquad k \in \mathbb{S} ackslash A.$$

The equations (5.2.1) can be rewritten in matrix form as

$$h_A = egin{bmatrix} 1 \ dots \ 1 \ 1 \end{bmatrix} + Ph_A,$$

by considering only the rows with index $k \in A^c = \mathbb{S} \setminus A$, under the boundary conditions

$$h_A(k)=0, \qquad k\in A.$$
 MH3512 AY19-20

Block triangular transition matrices

When the transition matrix P has the form (5.1.6) and $A = \{r + 1, r + 2, \dots, N\}$, Equation (5.2.1) rewrites as

$$egin{aligned} h_A(k) &= 1 + \sum\limits_{l=0}^N P_{k,l} h_A(l) \ &= 1 + \sum\limits_{l=0}^r P_{k,l} h_A(l) + \sum\limits_{l=r+1}^N P_{k,l} h_A(l) \ &= 1 + \sum\limits_{l=0}^r P_{k,l} h_A(l), \qquad 0 \leqslant k \leqslant r, \end{aligned}$$

since $h_A(l) = 0, l = r + 1, r + 2, ..., N$, *i.e.*

$$h_A(k) = 1 + \sum\limits_{l=0}^r P_{k,l} h_A(l), \qquad 0 \leqslant k \leqslant r,$$

with $h_A(k) = 0, \, k = r + 1, \dots, n$.

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

$$h_{\{0\}}(1) = b \times 1 + (1-b)(1+h_{\{0\}}(1)) = 1 + (1-b)h_{\{0\}}(1),$$
(5.2.3)

with solution

$$h_{\{0\}}(1) = b\sum_{k \geqslant 1} k(1-b)^k = rac{1}{b},$$

and similarly we find

$$h_{\{1\}}(0) = a \sum_{k \geqslant 1} k(1-a)^k = rac{1}{a},$$

with $h_{\{0\}}(0) = h_{\{1\}}(1) = 0$, cf. also (5.3.3) below.

The above can be generalized to derive an equation for an expectation of the form

$$h_A(k):=\mathbb{E}\left[igsim_{i=0}^{T_A}f(Z_i) \ \Big| \ Z_0=k
ight], \qquad k=0,1,\dots,N,$$

where $f(\cdot)$ is a given utility function, as follows:

$$egin{aligned} h_A(k) &= & \mathbb{E}\left[\sum\limits_{i=0}^{T_A} f(Z_i) \ \Big| \ Z_0 = k
ight] \ &= & \sum\limits_{l=0}^{N} P_{k,l}\left(f(k) + \mathbb{E}\left[\sum\limits_{i=1}^{T_A} f(Z_i) \ \Big| \ Z_1 = l
ight]
ight) \ &= & \sum\limits_{l=0}^{N} P_{k,l}f(k) + \sum\limits_{l=0}^{N} P_{k,l} \, \mathbb{E}\left[\sum\limits_{i=1}^{T_A} f(Z_i) \ \Big| \ Z_1 = l
ight] \ &= & f(k) \sum\limits_{l=0}^{N} P_{k,l} + \sum\limits_{l=0}^{r} P_{k,l} \, \mathbb{E}\left[\sum\limits_{i=0}^{T_A} f(Z_i) \ \Big| \ Z_0 = l
ight] \ &= & f(k) + \sum\limits_{l=0}^{r} P_{k,l}h_A(l), \qquad k \in A^c := \{0, 1, \dots, r\}, \end{aligned}$$

with $A:=\{r+1,\ldots,N\}$, hence

$$h_A(k) = f(k) + \sum\limits_{l=0}^r P_{k,l} h_A(l), \qquad k \in A^c = \{0,1,\ldots,r\},$$

which can be rewritten as

$$\begin{bmatrix} h_A(0) \\ \vdots \\ h_A(r) \end{bmatrix} = \begin{bmatrix} f(0) \\ \vdots \\ f(r) \end{bmatrix} + \begin{vmatrix} P_{0,0} & P_{0,1} & P_{0,2} & \cdots & P_{0,r} \\ P_{1,0} & P_{1,1} & P_{1,2} & \cdots & P_{1,r} \\ P_{2,0} & P_{2,1} & P_{2,2} & \cdots & P_{2,r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{r,0} & P_{r,1} & P_{r,2} & \cdots & P_{r,r} \end{bmatrix} \times \begin{bmatrix} h_A(0) \\ \vdots \\ h_A(r) \end{bmatrix},$$

with the boundary condition

$$h_A(k) = f(k), \qquad k \in A = \{r+1, ..., n\},$$

see also Exercise 5.21 and Problem 5.27.

For example:

- When $f = \mathbb{1}_{A^c} = \mathbb{1}_{\{0,1,\dots,r\}}$ is the indicator function over the set A^c , *i.e.*

$$f(Z_i) = \mathbb{1}_{A^c}(Z_i) = egin{cases} 1 & ext{if } Z_i
otin A, \ 0 & ext{if } Z_i
otin A, \end{cases}$$

the quantity $h_A(k)$ coincides with the mean hitting time of the set A starting from the state k. In particular, when $A = \{m\}$ this recovers the equation

$$h_{\{m\}}(k) = 1 + \sum_{\substack{l \in \mathbb{S} \\ l \neq m}} P_{k,l} h_{\{m\}}(l), \qquad k \in \mathbb{S} \setminus \{m\},$$
(5.2.4)

with $h_{\{m\}}(m) = 0$.

- When f is the indicator function $f = \mathbb{1}_{\{l\}}$, *i.e.*

$$f(Z_i) = \mathbbm{1}_{\{l\}}(Z_i) = egin{cases} 1 & ext{if } Z_i = l, \ 0 & ext{if } Z_i
eq l, \end{cases}$$

with $l \in A^c$, the quantity $h_A(k)$ will yield the mean number of visits to state (l) starting from (k) before hitting the set A.

- See Exercise 5.20, Exercise 5.21, and also Problem 5.24 for a complete solution in case f(k) := k and $(Z_k)_{k \ge 0}$ is the gambling process of Chapter 2.

Poisson equation

Taking

$$h_A(k):=\mathbb{E}\left[\sum\limits_{i=0}^{T_A-1}f(Z_i)\ \Big|\ Z_0=k
ight],\qquad k=0,1,\ldots,N,$$

leads to the same first step analysis equation

$$h_A(k) = f(k) + \sum_{m=0}^r P_{k,m} h_A(m), \qquad k \in A^c := \{0, 1, \dots, r\},$$

or

$$(I_d - [P_{k,m}]_{k,m=0,1,...,r})h_A = f_A$$

with the boundary condition

$$h_A(k)=0, \qquad k\in A=\{r+1,\ldots,n\},$$

see e.g. Theorem 5.5 in [Pri08] for a continuous-time analog.

Example

Consider the Markov chain whose transition probability matrix is given by

$$P = \left[egin{array}{ccc} P_{i,j} \end{array}
ight]_{0 \leqslant i,j \leqslant 3} = \left[egin{array}{ccc} lpha & 0 & eta & 0 \ lpha & 0 & 0 & eta \ lpha & eta & 0 & 0 \ lpha & eta & 0 & 0 \ 0 & 0 & 0 & 1 \end{array}
ight],$$

where $\alpha, \beta \ge 0$ 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

$$h_3(k):=\mathbb{E}[T_3\mid Z_0=k]$$

denote the mean (hitting) time needed to reach state (3), after starting from state k = 0, 1, 2, 3. We get

$$egin{aligned} & (h_3(0)=lpha(1+h_3(0))+eta(1+h_3(2))=1+lpha h_3(0)+eta h_3(2)\ & h_3(1)=lpha(1+h_3(0))+eta(1+h_3(3))=1+lpha h_3(0)\ & h_3(2)=lpha(1+h_3(0))+eta(1+h_3(1))=1+lpha h_3(0)+eta h_3(1)\ & h_3(3)=0, \end{aligned}$$

which, using the relation $\alpha = 1 - \beta$, yields

$$h_3(3)=0, \qquad h_3(1)=rac{1}{eta^3}, \qquad h_3(2)=rac{1+eta}{eta^3}, \qquad h_3(0)=rac{1+eta+eta^2}{eta^3},$$

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 gres 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, $(h_3(1), h_3(2), h_3(0))^3$ converge to (1, 2, 3) as β goes to 1, as can be expected.

5.3 First Return Times

Consider now the first <u>return</u> time T_j^r to state $j \in S$, defined by

$$T_j^r:=\inf\{n\geqslant 1\;:\; Z_n=j\},$$

with

$$T_j^r=+\infty ~~ ext{if}~ Z_n
eq j ~ ext{for all}~n\geqslant 1.$$

Note that in contrast with the definition (5.1.1) of the hitting time T_j , the infimum is taken here for $n \ge 1$ 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

$$\mu_j(i) = \mathbb{E}[T_j^r \mid Z_0 = i] \geqslant 1$$

the mean return time to state $j \in S$ after starting from state $i \in S$.

Mean return times can also be computed by first step analysis. We have

$$egin{aligned} \mu_j(i) &= & \mathbb{E}[T_j^r \mid Z_0 = i] \ &= & 1 imes \mathbb{P}(Z_1 = j \mid Z_0 = i) \ &+ & \sum_{\substack{l \in \mathbb{S} \ l \neq j}} \mathbb{P}(Z_1 = l \mid Z_0 = i)(1 + \mathbb{E}[T_j^r \mid Z_0 = l]) \ &= & P_{i,j} + & \sum_{\substack{l \in \mathbb{S} \ l \neq j}} P_{i,l}(1 + \mu_j(l)) \ &= & P_{i,j} + & \sum_{\substack{l \in \mathbb{S} \ l \neq j}} P_{i,l} + & \sum_{\substack{l \in \mathbb{S} \ l \neq j}} P_{i,l} \mu_j(l) \end{aligned}$$

$$= \sum_{l \in \mathbb{S}} P_{i,l} + \sum_{\substack{l \in \mathbb{S} \\ l \neq j}} P_{i,l} \mu_j(l)$$

$$= 1 + \sum_{\substack{l \in \mathbb{S} \\ l \neq j}} P_{i,l} \mu_j(l),$$

$$(5.204)$$

hence

$$\mu_j(i) = 1 + \sum_{\substack{l \in \mathbb{S} \\ l \neq j}} P_{i,l} \mu_j(l), \quad i, j \in \mathbb{S}.$$

$$(5.3.1)$$

Hitting times vs 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 equation (5.2.4) in Section 5.2. In addition, the time T_i^r to return to state (i) is always at least one by construction, hence $\mu_i(i) \ge 1$ cannot vanish, while we always have $h_i(i) = 0$, $i \in S$. On the other hand, by definition we have

$$\mu_i(j) = \mathbb{E}[T_i^r \mid Z_0 = j] = \mathbb{E}[T_i \mid Z_0 = j] = h_i(j),$$

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

$$\mu_{j}(j) = \sum_{l \in S} P_{j,l}(1 + h_{j}(l))$$

$$= P_{j,j} + \sum_{l \neq j} P_{j,l}(1 + h_{j}(l))$$

$$= \sum_{l \in S} P_{j,l} + \sum_{l \neq j} P_{j,l}h_{j}(l)$$

$$= 1 + \sum_{l \neq j} P_{j,l}h_{j}(l), \quad j \in S,$$
(5.3.2)

which is 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), $i, j \in S$.

Examples

 $\bigcirc 205$

The mean return time $\mu_0(i) = \mathbb{E}[T_0^r \mid Z_0 = i]$ to state (1) starting from state (1) $\in \{0, 1\}$ satisfies

$$\left\{egin{array}{l} \mu_0(0) = (1-a) imes 1 + a(1+\mu_0(1)) = 1 + a\mu_0(1) \ \mu_0(1) = b imes 1 + (1-b)(1+\mu_0(1)) = 1 + (1-b)\mu_0(1) \end{array}
ight.$$

which yields

$$\mu_0(0) = 1 + \frac{a}{b} \quad \text{and} \quad \mu_0(1) = h_0(1) = \frac{1}{b},$$
(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 $Z_0 = 0$ is given by

$$f_{0,0}^{(n)} := \mathbb{P}(T_0^r = n \mid Z_0 = 0) = \begin{cases} 0 & \text{if } n = 0, \\ 1 - a & \text{if } n = 1, \\ ab(1-b)^{n-2} & \text{if } n \ge 2, \end{cases}$$
(5.3.4)

hence (5.3.3) can be directly recovered as^{*}

$$egin{aligned} \mu_0(0) &=& \mathbb{E}[T_0^r \mid Z_0 = 0] \ &=& \sum\limits_{n \geqslant 0} n \mathbb{P}(T_0^r = n \mid Z_0 = 0) \ &=& \sum\limits_{n \geqslant 0} n f_{0,0}^{(n)} \ &=& 1 - a + ab \sum\limits_{n \geqslant 2} n(1-b)^{n-2} \ &=& 1 - a + ab \sum\limits_{n \geqslant 0} (n+2)(1-b)^n \end{aligned}$$

*We are using the identities $\sum_{k\geq 0} r^k = (1-r)^{-1}$ and $\sum_{k\geq 1} kr^{k-1} = (1-r)^{-2}$, cf. (A.3) and (A.4).

MH3512 AY19-20 $\,$

$$= 1 - a + ab(1 - b) \sum_{n \ge 0} n(1 - b)^{n-1} + 2ab \sum_{n \ge 0} (1 - b)^n$$

$$= \frac{a + b}{b} = 1 + \frac{a}{b},$$
(5.3.5)

where we used the identity (A.4). Similarly we check that

$$ig(egin{array}{c} \mu_1(0) = 1 + (1-a) \mu_1(0) \ \mu_1(1) = 1 + b \mu_1(0), \end{array}$$

which yields

$$\mu_1(0) = h_1(0) = rac{1}{a} \quad ext{and} \quad \mu_1(1) = 1 + rac{b}{a},$$

and can be directly recovered by

$$\mu_1(1) = 1 - b + ab \sum_{n \ge 0} (n+2)(1-a)^n = \frac{a+b}{a} = 1 + \frac{b}{a},$$
(5.3.6)

as in (5.3.3) and (5.3.5) above, by swapping \boldsymbol{a} with \boldsymbol{b} and state $\boldsymbol{\bigcirc}$ with state $\boldsymbol{\bigcirc}$.

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 $k \ge 1$ exit doors from one compartment, it will choose one of them with probability 1/k, *i.e.* the transition matrix is

 $0 \ 1/2 \ 0 \ 0 \ 0$ 1/21/2P =1/21/2• 1/21/21/21/31/31/3



MH3512 AY19-20

♦ 207

<u>Question</u>: Find the average time to come back to state $\widehat{1}$ after starting from state $\widehat{1}$.

Answer: Letting

$$T_l^r = \inf\{n \geqslant 1 \; : \; Z_n = l\}$$

denote the first return time to state (I), and defining

$$\mu_1(k) := \mathbb{E}[T_1^r \mid Z_0 = k]$$
 MH3512 AY19-20

the mean return time to state (1) starting from (k), we have

$$\begin{cases} \mu_{1}(1) = 1 + \mu_{1}(2) \\ \mu_{1}(2) = \frac{1}{2}(1+0) + \frac{1}{2}(1+\mu_{1}(3)) = 1 + \frac{1}{2}\mu_{1}(3) \\ \mu_{1}(3) = \frac{1}{2}(1+\mu_{1}(2)) + \frac{1}{2}(1+\mu_{1}(6)) = 1 + \frac{1}{2}\mu_{1}(2) + \frac{1}{2}\mu_{1}(6) \\ \mu_{1}(4) = 1 + \mu_{1}(7) \\ \mu_{1}(5) = \frac{1}{2}(1+\mu_{1}(8)) + \frac{1}{2}(1+\mu_{1}(6)) = 1 + \frac{1}{2}\mu_{1}(8) + \frac{1}{2}\mu_{1}(6) \\ \mu_{1}(6) = \frac{1}{2}(1+\mu_{1}(3)) + \frac{1}{2}(1+\mu_{1}(5)) = 1 + \frac{1}{2}\mu_{1}(3) + \frac{1}{2}\mu_{1}(5) \\ \mu_{1}(7) = 1 + \frac{1}{2}\mu_{1}(4) + \frac{1}{2}\mu_{1}(8) = \frac{1}{2}(1+\mu_{1}(4)) + \frac{1}{2}(1+\mu_{1}(8)) \\ \mu_{1}(8) = \frac{1}{3}(1+\mu_{1}(7)) + \frac{1}{3}(1+\mu_{1}(5)) + \frac{1}{3}(1+\mu_{1}(9)) \\ = 1 + \frac{1}{3}(\mu_{1}(7) + \mu_{1}(5) + \mu_{1}(9)) \\ \mu_{1}(9) = 1 + \mu_{1}(8), \end{cases}$$

or

$$\begin{cases} \mu_1(1) = 1 + \mu_1(2), & \mu_1(2) = 1 + \frac{1}{2}\mu_1(3), & \mu_1(3) = 2 + \frac{2}{3}\mu_1(6), \\ \\ \mu_1(4) = 1 + \mu_1(7), & \mu_1(5) = 1 + \frac{1}{2}\mu_1(8) + \frac{1}{2}\mu_1(6), & 0 = 30 + 3\mu_1(8) - 5\mu_1(6), \\ \\ \\ \mu_1(7) = 3 + \mu_1(8), & 0 = 80 + 5\mu_1(6) - 5\mu_1(8), & \mu_1(9) = 1 + \mu_1(8), \end{cases}$$

♦ 209

which yields

$$\mu_1(1) = 16, \ \mu_1(2) = 15, \ \mu_1(3) = 28, \ \mu_1(4) = 59, \ \mu_1(5) = 48,$$

$$\mu_1(6) = 39, \ \mu_1(7) = 58, \ \mu_1(8) = 55, \ \mu_1(9) = 56.$$
(5.3.8)

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.



The next Figure 5.1 represents the mean return times to state 0 according to the initial state on the maze (5.3.7).

ॳ 210



Figure 5.1: Mean return times to state 0 on the maze (5.3.7).

5.4 Mean Number of Returns

Return probabilities

In the sequel we let

$$p_{ij} = \mathbb{P}(T_j^r < \infty \mid Z_0 = i) = \mathbb{P}(Z_n = j \text{ for some } n \ge 1 \mid Z_0 = i)$$

denote the probability of return to state (j) in finite time^{*} 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:

$$p_{ii} = \mathbb{P}(Z_n = i \text{ for some } n \ge 1 \mid Z_0 = i)$$

$$= \mathbb{P}\left(\bigcup_{n\ge 1} \{Z_n = i\} \mid Z_0 = i\right)$$

$$= \sum_{n\ge 1} \mathbb{P}(Z_n = i, Z_{n-1} \ne i, \dots, Z_1 \ne i \mid Z_0 = i)$$

$$= \sum_{n\ge 1} f_{i,i}^{(n)}, \qquad (5.4.1)$$

*When $(i) \neq (j)$, p_{ij} is the probability of *visiting* state (j) in finite time after starting from state (i).

MH3512 AY19-20

ॳ 211

where

$$f_{i,j}^{(n)} := \mathbb{P}(T_j^r = n \mid Z_0 = i) = \mathbb{P}(Z_n = j, Z_{n-1}
eq j, \ldots, Z_1
eq j \mid Z_0 = i),$$

 $i, j \in \mathbb{S}$, is the probability distribution of T_j^r given that $Z_0 = i$, with

$$f_{i,i}^{(0)} = \mathbb{P}(T_i^r = 0 \mid Z_0 = i) = 0.$$

Note that we have

$$f_{i,i}^{(1)} = \mathbb{P}(Z_1 = i \mid Z_0 = i) = P_{i,i}, \qquad i \in \mathbb{S}$$

Convolution equation

By conditioning on the first return time $k \ge 1$, the return time probability distribution $f_{i,i}^{(k)} = \mathbb{P}(T_i^r = k \mid Z_0 = i)$ satisfies the convolution equation

$$egin{aligned} P^n]_{i,i} &= & \mathbb{P}(Z_n = i \mid Z_0 = i) \ &= & \sum\limits_{k=1}^n \mathbb{P}(Z_k = i, Z_{k-1}
eq i, \ldots, Z_1
eq i \mid Z_0 = i) \mathbb{P}(Z_n = i \mid Z_k = i) \ &= & \sum\limits_{k=1}^n \mathbb{P}(Z_k = i, Z_{k-1}
eq i, \ldots, Z_1
eq i \mid Z_0 = i) \mathbb{P}(Z_{n-k} = i \mid Z_0 = i) \ &= & \sum\limits_{k=1}^n f_{i,i}^{(k)} [P^{n-k}]_{i,i}, \end{aligned}$$

which extends the convolution equation (3.4.7) from random walks to the more general setting of Markov chains.

The return probabilities p_{ij} will be used below to compute the average number of returns to a given state, and the distribution $f_{i,j}^{(k)}$, $k \ge 1$, of T_j^r given that $Z_0 = i$ will be useful in Section 6.4 on positive and null recurrence.

Number of returns

Let

$$R_j := \sum_{n \ge 1} \mathbb{1}_{\{Z_n = j\}} \tag{5.4.2}_{\text{MH3512 AY19-20}}$$

denote the number of returns^{*} to state (j) by the chain $(Z_n)_{n \in \mathbb{N}}$. The next proposition shows that, given $\{Z_0 = i\}, R_j$ has \mathbb{I}_a^3 zero-modified geometric distribution with initial mass $1 - p_{ij}$.

Proposition 5.5

The probability distribution of the number of returns R_j to state j given that $\{Z_0 = i\}$ is given by

$$\mathbb{P}(R_j=m\mid Z_0=i)= egin{cases} 1-p_{ij},&m=0,\ p_{ij} imes(p_{jj})^{m-1} imes(1-p_{jj}),&m\geqslant 1, \end{cases}$$

Proof

When the chain never visits state (j) starting from $Z_0 = i$ we have $R_j = 0$, and this happens with probability

$$egin{array}{rl} \mathbb{P}(R_j=0 \mid Z_0=i) &=& \mathbb{P}(T_j^r=\infty \mid Z_0=i) \ &=& 1-\mathbb{P}(T_j^r<\infty \mid Z_0=i) \ &=& 1-p_{ij}. \end{array}$$

Next, when the chain $(Z_n)_{n \in \mathbb{N}}$ makes a number $R_j = m \ge 1$ of visits to state (j) starting from state (i), it makes a first visit to state (j) with probability $\underline{p_{i,j}}$ and then makes $\underline{m-1}$ returns to state (j), each with probability $\underline{p_{jj}}$. After those m visits, it never returns to state (j), and this event occurs with probability $1 - p_{jj}$. Hence, given that $\{Z_0 = i\}$ we have

$$\mathbb{P}(R_j = m \mid Z_0 = i) = egin{cases} p_{ij} imes (p_{jj})^{m-1} imes (1-p_{jj}), & m \geqslant 1, \ 1-p_{ij}, & m = 0, \end{cases}$$

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

$$\mathbb{P}(R_i=m\mid Z_0=i)=(1-p_{ii})(p_{ii})^m, \qquad m\geqslant 0.$$

^{*}Here, R_i is called a number of returns because the time counter is started at n = 1 and excludes the initial state.

ॳ 214

Proposition 5.6

We have

$$\mathbb{P}(R_j < \infty \mid Z_0 = i) = egin{cases} 1 - p_{ij}, & ext{if } p_{jj} = 1, \ 1, & ext{if } p_{jj} < 1. \end{cases}$$

Proof

We note that

$$egin{aligned} \mathbb{P}(R_j < \infty \mid Z_0 = i) &= \mathbb{P}(R_j = 0 \mid Z_0 = i) + \sum\limits_{m \geqslant 1} \mathbb{P}(R_j = m \mid Z_0 = i) \ &= 1 - p_{ij} + p_{ij}(1 - p_{jj}) \sum\limits_{m \geqslant 1} (p_{jj})^{m-1} \ &= \left\{egin{aligned} 1 - p_{ij}, & ext{if } p_{jj} = 1, \ 1, & ext{if } p_{jj} < 1. \end{aligned}
ight. \end{aligned}$$

We also have

$$\mathbb{P}(R_j=\infty \mid Z_0=i) = egin{cases} p_{ij}, & ext{if } p_{jj}=1, \ 0, & ext{if } p_{jj} < 1. \end{cases}$$

In particular if $p_{jj} = 1$, *i.e.* state (j) is recurrent, we have

$$\mathbb{P}(R_j=m\mid Z_0=i)=0, \qquad m\geqslant 1,$$

and in this case,

$$\left\{egin{array}{l} \mathbb{P}(R_j < \infty \mid Z_0 = i) = \mathbb{P}(R_j = 0 \mid Z_0 = i) = 1 - p_{ij}, \ \mathbb{P}(R_j = \infty \mid Z_0 = i) = 1 - \mathbb{P}(R_j < \infty \mid Z_0 = i) = p_{ij}. \end{array}
ight.$$

MH3512 AY19-20

On the other hand, when i = j, by (1.5.13) we find

$$\mathbb{P}(R_{i} < \infty \mid Z_{0} = i) = \sum_{m \geq 0} \mathbb{P}(R_{i} = m \mid Z_{0} = i) \\
= (1 - p_{ii}) \sum_{m \geq 0} (p_{ii})^{m} \\
= \begin{cases} 0, & \text{if } p_{ii} = 1, \\ 1, & \text{if } p_{ii} < 1, \end{cases}$$
(5.4.3)

hence

$$\mathbb{P}(R_{i} = \infty \mid Z_{0} = i) = \begin{cases} 1, & \text{if } p_{ii} = 1, \\ 0, & \text{if } p_{ii} < 1, \end{cases}$$
(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 Chapter 6. By (A.4), when $p_{jj} < 1$ we have $\mathbb{P}(R_j < \infty \mid Z_0 = i) = 1$ and^{*}

$$\mathbb{E}[R_{j} \mid Z_{0} = i] = \sum_{m \ge 0} m \mathbb{P}(R_{j} = m \mid Z_{0} = i)$$
(5.4.5)
$$= (1 - p_{jj}) p_{ij} \sum_{m \ge 1} m(p_{jj})^{m-1}$$
$$= \frac{p_{ij}}{1 - p_{jj}},$$
(5.4.6)

hence

$$\mathbb{E}[R_j \mid Z_0 = i] < \infty \quad ext{if} \quad p_{jj} < 1.$$

If $p_{j,j} = 1$ then $\mathbb{E}[R_j \mid Z_0 = i] = \infty$ unless $p_{i,j} = 0$, in which case $\mathbb{P}(R_j = 0 \mid Z_0 = i) = 1$ and $\mathbb{E}[R_j \mid Z_0 = i] = 0$. In particular, when i = j we find the next proposition.

*We are using the identity
$$\sum_{k \ge 1} kr^{k-1} = (1-r)^{-2}$$
, cf. (A.4).
MH3512 AY19-20

Proposition 5.7

The mean number of returns to state (i) is given by

$$\mathbb{E}[R_i \mid Z_0 = i] = rac{p_{ii}}{1-p_{ii}},$$

and it is finite, *i.e.* $\mathbb{E}[R_i \mid Z_0 = i] < \infty$, if and only if $p_{ii} < 1$.

More generally, by (5.4.2) we can also write

$$\mathbb{E}[R_{j} \mid Z_{0} = i] = \mathbb{E}\left[\sum_{n \ge 1} \mathbb{1}_{\{Z_{n} = j\}} \mid Z_{0} = i\right] \\ = \sum_{n \ge 1} \mathbb{E}[\mathbb{1}_{\{Z_{n} = j\}} \mid Z_{0} = i] \\ = \sum_{n \ge 1} \mathbb{P}(Z_{n} = j \mid Z_{0} = i) \\ = \sum_{n \ge 1} [P^{n}]_{i,j}.$$
(5.4.7)

The above quantity coincides with

$$\mathbb{E}[R_j \mid Z_0 = i] = -\mathbb{1}_{\{i=j\}} + \left[(I_d - P)^{-1}
ight]_{i,j},$$

where by analogy with (A.3), the matrix inverse

$$G:=\sum\limits_{n\geqslant 0}P^n=(I_d-P)^{-1}$$

of $I_d - P$ is called the *potential kernel*, or the *resolvent* of P. Finally, if (m) is the *only* absorbing state, we can also write

$$\mathbb{E}[T_m \mid Z_0 = i] = 1 + \sum_{\substack{j \in \mathbb{S} \ j
eq m}} \mathbb{E}[R_j \mid Z_0 = i] = \sum_{\substack{j \in \mathbb{S} \ j
eq m}} \left[(I_d - P)^{-1}
ight]_{i,j}, \hspace{1em} i
eq m.$$

See [AKS93] for an application of this formula to the Snakes and Ladders game.

MH3512 AY19-20

♦ 216
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

$$\begin{split} \mathbb{P}(X_1 = 0 \mid X_0 = 0) &= 1, \\ \mathbb{P}(X_1 = 0 \mid X_0 = 1) &= 1/2, \\ \mathbb{P}(X_1 = 1 \mid X_0 = 2) &= 1/3, \end{split} \qquad \begin{array}{l} \mathbb{P}(X_1 = 3 \mid X_0 = 3) &= 1, \\ \mathbb{P}(X_1 = 3 \mid X_0 = 1) &= 1/2, \\ \mathbb{P}(X_1 = 3 \mid X_0 = 2) &= 2/3. \end{split}$$

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

$$T_3:=\inf\{n\geqslant 0\;:\;X_n=3\}.$$

c) Letting

$$T_{0,3}:=\inf\{n\geqslant 0\;:\;X_n=0 ext{ or }X_n=3\},$$

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 $(X_n)_{n \ge 0}$ with state space $\mathbb{S} = \{0, 1\}$ and transition matrix

$$P = \left[egin{array}{ccc} 0.4 & 0.6 \ & \ 0.8 & 0.2 \end{array}
ight].$$

Compute the mean duration between two visits to state (1).

Exercise 5.3 Consider the Markov chain $(X_n)_{n \ge 0}$ on $\mathbb{S} = \{0, 1, 2\}$ whose transition probability matrix P is given by

$$P = \begin{array}{ccc} 0 & 1 & 2 \\ 0 & 1 & 0 \\ 1/3 & 0 & 2/3 \\ 2 & 0 & 1 & 0 \end{array} \right].$$

MH3512 AY19-20

- 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 A box contains red balls and green balls. At each time step we pick a ball uniformly at random and without replacement. If the ball is red we lose \$1, and if the ball is green we gain +\$1. The game ends when the box becomes empty. We let f(x, y) denote the value of the game when the game starts with $x \ge 0$ red balls and $y \ge 0$ green balls in the box.

- a) Using first step analysis, derive the finite difference equation satisfied by f(x, y) for $x + y \ge 1$.
- b) Find the boundary conditions f(x,0), $x \ge 1$, and f(0,y), $y \ge 1$.
- c) Solve the equation of Question (a) for f(x, y), x, y = 1, 2, 3.
- d) Find f(x, y) for all $x, y \ge 0$.

Exercise 5.5 Consider the Markov chain with graph



and let

$$T_k^r:=\inf\{n\geqslant 1\;:\;X_n=k\}$$

denote the *return* time to state k = 0, 1, 2, 3.

MH3512 AY19-20

(5.4.8)

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 5.6 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

$$\left[\begin{array}{ccc} P_{i,j} \end{array}
ight]_{0\leqslant i,j\leqslant 3} = \left[egin{array}{cccc} 1 & 0 & 0 & 0 \ 0.3 & 0 & 0.7 & 0 \ 0 & 0.3 & 0 & 0.7 \ 0 & 0 & 0 & 1 \end{array}
ight]$$

- 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 (k), for k = 0, 1, 2, 3.

Exercise 5.7 Consider a random walk with Markov transition matrix given by

$$\left[\begin{array}{cccc} P_{i,j} \end{array}
ight]_{0\leqslant i,j\leqslant 3} = \left[egin{array}{cccc} 0.5 & 0 & 0.5 & 0 \ 0.5 & 0 & 0 & 0.5 \ 0.5 & 0.5 & 0 & 0 \ 0 & 0 & 0 & 1 \end{array}
ight]$$

Compute the average time it takes to reach state ③ given that the chain is started at state ④.

Exercise 5.8 Consider the Markov chain $(X_n)_{n \ge 0}$ on $\{0, 1, 2, 3\}$ whose transition probability matrix P is given by

$$P = \left[egin{array}{ccccc} 1 & 0 & 0 & 0 \ 0.5 & 0 & 0.5 & 0 \ 0 & 0.5 & 0 & 0.5 \ 0 & 0 & 0 & 1 \end{array}
ight]$$

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 5.9 Consider a discrete-time, time-homogeneous Markov chain $(X_n)_{n \in \mathbb{N}}$ on a state space S, and the first hitting time

$$T_A = \inf\{n \geqslant 0 \; : \; Z_n \in A\},$$

of a subset $A \subset$ of 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 $n, m \ge 0, j \in S$, and $(i_k)_{k \in \mathbb{N}} \subset S$ we have

$$\mathbb{P}(X_{T_A+n}=j\mid X_{T_A}=i_0,\ldots,X_0=i_{T_A} ext{ and }T_A<+\infty)=\mathbb{P}(X_n=j\mid X_0=i_0).$$

Exercise 5.10 We consider the simple random walk $(S_n)_{n \in \mathbb{N}}$ of Chapter 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 $(k) \ge 0$ 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 (1) from any state (1) ≥ 0 when q > p.

MH3512 AY19-20

Exercise 5.11 A player tosses a fair six-sided die and records the number appearing on the uppermost face. The die is then tossed a_{221}^{221} 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 5.12 A fish is put into the linear maze as shown, and its state at time n is denoted by $X_n \in \{0, 1, 2, 3, 4, 5\}$:



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

$$T_0 = \inf\{n \geqslant 0 \; : \; X_n = 0\}, \quad ext{and} \quad T_5 = \inf\{n \geqslant 0 \; : \; X_n = 5\},$$

and $g(k) = \mathbb{P}(T_5 < T_0 \mid X_0 = k), k = 0, 1, \dots, 5.$

- 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 (k) it finds the food before getting shocked, for $k = 0, 1, \ldots, 5$.

Exercise 5.13 Starting from a state $m \ge 1$ 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 $h_0(m)$, $m \ge 1$ and give the values of $h_0(0)$ and $h_0(1)$.

c) Show that $h_0(m)$ is given by $h_0(m) = h_0(m-1) + \frac{1}{m}$, $m \ge 1$, and that

$$h_0(m)=\sum\limits_{k=1}^mrac{1}{k},$$

for all $m \in \mathbb{N}$.

Exercise 5.14 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 5.15 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 travelling.

that it will remain trapped in the maze?

Exercise 5.16 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 5.17 Consider the Markov chain on $\{0, 1, 2\}$ with transition matrix

$$\left[egin{array}{cccc} 1/3 & 1/3 & 1/3 \ 1/4 & 3/4 & 0 \ 0 & 0 & 1 \end{array}
ight].$$

- 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 5.18 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

$$egin{aligned} \mathbb{P}(X_{k+1} = x+1, \; Y_{k+1} = y \mid X_k = x, \; Y_k = y) \ &= \; \mathbb{P}(X_{k+1} = x, \; Y_{k+1} = y+1 \mid X_k = x, \; Y_k = y) \ &= \; rac{1}{2}, \qquad k \geqslant 0, \end{aligned}$$
(MH3512 AY19-20)

and let

$$A = ig\{(x,y) \in \mathbb{N} imes \mathbb{N} \; : \; x \geqslant 2, \; y \geqslant 2ig\}$$

Let also

$$T_A:=\inf\{n\geqslant 0 \;:\; X_n\geqslant 2 ext{ and } Y_n\geqslant 2\}$$

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 $x \ge 2$ and $y \ge 2$.
- b) Show that $\mu_A(x, y)$ solves the equation

$$\mu_A(x,y) = 1 + \frac{1}{2}\mu_A(x+1,y) + \frac{1}{2}\mu_A(x,y+1), \quad x,y \in \mathbb{N}.$$
(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 ring toss game, a ring is thrown at two sticks in such a way that each stick has exactly 50% chance to receive the ring. Compute the mean time it takes until both sticks receive least two rings.

Exercise 5.19 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

 $k \ge 0$, and let

$$A:=\{(x,y)\in\mathbb{N} imes\mathbb{N}\ :\ x\geqslant3 ext{ or }y\geqslant3\}$$

Let also

$$T_A:=\infig\{n \geqslant 0 \; : \; (X_n,Y_n) \in Aig\}$$
 MH3512 AY19-20



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 $\bigcirc 225$

$$\mu_A(x,y):=\mathbb{E}ig[T_A\mid (X_0,Y_0)=(x,y)ig], \hspace{1em} (x,y)\in\mathbb{N} imes\mathbb{N}.$$

- 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 $\mu_A(x, y)$ on the domain

$$A^c = ig\{(x,y) \in \mathbb{N} imes \mathbb{N} \; : \; 0 \leqslant x,y \leqslant 3ig\}.$$

- 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 5.20 Let $(X_n)_{n\geq 0}$ be a Markov chain with state space S and transition probability matrix $(P_{ij})_{i,j\in S}$. Our goal is to compute the expected value of the infinite discounted series

$$h(i):=\mathbb{E}\left[\sum\limits_{n\geqslant 0}eta^n c(X_n) \ \Big| \ X_0=i
ight], \qquad i\in \mathbb{S},$$

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

$$h(i) = c(i) + eta \sum\limits_{j \in \mathbb{S}} P_{ij} h(j)$$

for every state (i) \in **S**.

Exercise 5.21 Consider a Markov Decision Process (MDP) on a state space S, with set of actions A and family $(P^a)_{a \in A}$ of transition probability matrices

$$egin{aligned} P: \mathbb{S} imes \mathbb{S} imes \mathbb{A} o [0,1], \ (k,l,a) &\longmapsto P^a_{k,l}. \end{aligned}$$
 MH3512 AY19-20

By first step analysis, derive the *Bellman equation* for the optimal value function

$$V^*(k) = \max_{\pi} \mathbb{E} \left[\sum_{n \geqslant 0} \gamma^n R(X_n) \; \Big| \; X_0 = k
ight],$$

over all *policies* $\pi : \mathbb{S} \to \mathbb{A}$ giving the action chosen at every given state in \mathbb{S} , where $\gamma \in (0, 1)$ is a discount factor and $R : \mathbb{S} \to \mathbb{R}$ is a reward function.

Problem 5.22 Let $(X_n)_{n \in \mathbb{N}}$ be a Markov chain on $\{0, 1, \ldots, N\}$, $N \ge 1$, with transition matrix $P = [P_{i,j}]_{0 \le i,j \le N}$.

a) Consider the hitting times

$$T_0 = \inf\{n \geqslant 0 \; : \; X_n = 0\}, \qquad T_N = \inf\{n \geqslant 0 \; : \; X_n = N\},$$

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

$$g(k) = \sum_{l=0}^{N} P_{k,l}g(l), \quad k = 1, \dots, N-1.$$
 (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

$$P_{k,l} = \mathbb{P}(X_{n+1} = l \mid X_n = k) = inom{N}{l} \left(rac{k}{N}
ight)^l \left(1 - rac{k}{N}
ight)^{N-l},$$

 $k, l = 0, 1, \ldots, N$. 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

$$\mathbb{P}(T_0 < T_N \mid X_0 = k) = rac{N-k}{N}, \qquad k=0,1,\ldots,N.$$

MH3512 AY19-20

♦ 226

e) Let

$$T_{0,N} = \inf\{n \ge 0 : X_n = 0 \text{ or } X_n = N\},$$

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, that the function h satisfies the relation

$$h(k) = 1 + \sum_{l=0}^{N} P_{k,l} h(l), \qquad k = 1, 2, \dots, N-1.$$

g) Assuming that $\underline{N = 3}$, compute

$$h(k) = \mathbb{E}[T_{0,3} \mid X_0 = k], \qquad k = 0, 1, 2, 3.$$

Problem 5.23 Pattern recognition (Exercise 4.7 continued). Consider a sequence $(X_n)_{n\geq 0}$ of *i.i.d.* Bernoulli random variables taking values in a two-letter alphabet $\{a, b\}$, with

$$\mathbb{P}(X_n=a)=p \quad ext{and} \quad \mathbb{P}(X_n=b)=q=1-p, \quad n \geqslant 0,$$

with $0 , and the process <math>(Z_n)_{n \ge 1}$ defined by

$$Z_n:=(X_{n-1},X_n), \qquad n\geqslant 1.$$

- a) Argue that $(Z_n)_{n \ge 1}$ is a Markov chain with four possible states (or words) $\{aa, ab, ba, bb\}$, and write down its 4×4 transition matrix.
- b) Let

$$au_{ab} = \inf\{n \geqslant 1 \; : \; Z_n = (a,b)\}$$

denote the first time of appearance of the pattern "ab" in the sequence (X_0, X_1, X_2, \ldots) . Give the value of

$$G_{ab}(s) := \mathbb{E}[s^{ au_{ab}} \mid Z_1 = (a,b)], \qquad -1 < s < 1.$$

MH3512 AY19-20

c) Consider the probability generating functions

$$G_{aa}(s) := \mathbb{E}[s^{\tau_{ab}} \mid Z_1 = (a, a)], \text{ and } G_{ba}(s) := \mathbb{E}[s^{\tau_{ab}} \mid Z_1 = (b, a)],$$

-1 < s < 1. Using first step analysis, complete the system of equations

$$\begin{cases} G_{aa}(s) = psG_{aa}(s) + qsG_{ab}(s), \\ G_{ba}(s) = ? + ? \end{cases}$$
(5.4.11)

- d) Compute $G_{aa}(s)$ and $G_{ba}(s)$ by solving the system (5.4.11).
- e) Using probability generating functions, compute the averages

$$\mathbb{E}[au_{ab} \mid Z_1 = (a,a)] \quad ext{and} \quad \mathbb{E}[au_{ab} \mid Z_1 = (b,a)].$$

f) Find the average time it takes until we encounter the pattern "ab" in the sequence (X_0, X_1, X_2, \ldots) started with $X_0 = a$.

Problem 5.24 Consider a gambling process $(X_n)_{n \ge 0}$ on the state space $\{0, 1, \ldots, N\}$, with transition probabilities

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

 $k = 1, 2, \dots, N - 1$, with p + q = 1. Let

 $au:=\inf\{n\geqslant 0\;:\;X_n=0 ext{ or }X_n=N\}$

denote the time until the process hits either state (0) or state N, and consider the expectation

$$h(k):=\mathbb{E}\left[\sum\limits_{i=0}^{ au-1}X_i \ \Big| \ X_0=k
ight],$$

of the random sum $\sum_{\mathbf{0} \leq i < \tau} X_i$ of all chain values visited *before* the process hits **0** or *N* after starting from $k = 0, 1, 2, \dots, N$. MH3512 AY19-20

- a) Give the values of h(0) and h(N).*
- b) Show, by first step analysis, that h(k) satisfies the equations

$$h(k) = k + ph(k+1) + qh(k-1), \qquad k = 1, 2, ..., N-1.$$
 (5.4.12)

From now on we take p = q = 1/2.

- c) Find a particular solution of Equation (5.4.12).
- d) Knowing that the solution of the associated homogeneous equation

$$f(k) = \frac{1}{2}f(k+1) + \frac{1}{2}f(k-1), \qquad k = 1, 2, \dots, N-1,$$

takes the form

$$f(k) = C_1 + C_2 k,$$

show that the expectation h(k) solution of (5.4.12) is given by

$$h(k) = k rac{N^2 - k^2}{3}, \qquad k = 0, 1, \dots, N.$$

- e) Compute h(1) when $\underline{N=2}$ and explain why this result makes pathwise sense.
- f) Suppose that you start a business with initial monthly income of \$4K. Every month the income you receive from that business may increase or decrease by \$1K with equal probabilities (1/2, 1/2). You decide to stop that business as soon as your monthly income hits the levels 0 or \$70K, whichever comes first.
 - i) Compute the expected duration of your business (in number of months).[†]
 - ii) Compute your expected accumulated wealth until the month before you stop your business.

*We apply the convention $\sum_{i=0}^{-1} = \sum_{0 \leqslant i < 0} = 0.$ [†]Recall that $\mathbb{E}[T_{0,N} \mid X_0 = k] = k(N-k), k = 0, 1, \dots, N.$

MH3512 AY19-20 $\,$

iii) Compute your expected accumulated wealth under the assumption that your monthly income remains constant equal to \hat{F} over the same mean duration as in Question (f-*i*) above.

iv) Any comment?

Problem 5.25 ([Gus14]). Cellular automata. Given an alphabet $\Sigma = \{a, b\}$ made of two letters we denote by Σ^* the set of all words over $\{a, b\}$, *i.e.* Σ^* is made of all finite sequences of symbols in $\{a, b\}$.

A language \mathcal{L} over $\{a, b\}$ is a collection of (finite) words in Σ^* .

In computer science, an *automaton* given by a function

$$f: \{a,b\} imes \{0,1,\ldots,n\} \longrightarrow \{0,1,\ldots,n\}$$

is reading words of the form $a_1a_2 \cdots a_m \in \mathcal{L}$ by producing a sequence x_1, x_2, \ldots, x_m of integers via the recursion

$$x_1:=f(a_1,x_0),\; x_2:=f(a_2,x_1),\; x_3:=f(a_3,x_2),\ldots,x_m:=f(a_m,x_{m-1}).$$

A word $a_1 a_2 \cdots a_m \in \mathcal{L}$, $m \ge 1$, is said to synchronize the automaton f to state n if we have $x_m = n$. Here, n is regarded as a sink state, also called accepting state.

One says that the automaton f recognizes the language \mathcal{L} if every word $a_1 a_2 \cdots a_m \in \mathcal{L}$ synchronizes the automaton f to state n, *i.e.* satisfies $x_m = n$, starting from any initial state x_0 .

Synchronizing automata are connected to algebra and combinatorics, and they have applications in many areas including robotics, coding theory, network security management, chip design, industrial automation, biocomputing, etc.

a) Let n = 5 and consider the automaton given by the function



f	0	1	2	3	4	5
\boldsymbol{a}	1	2	3	3	1	5
b	0	0	0	4	5	5

Draw a graph for this automaton and find the smallest integers $l, m \ge 1$ such that f can recognize the word $a^l b^m$ starting from any state of the automaton.

b) Determine the language recognized by the automaton and give an example of a five-letter word that cannot synchronize the automaton when it starts from state (4).

Hint: The set of words recognized by the automaton can be written using the notation $\Sigma^* x x x x \Sigma^*$ which represents the concatenations of a word in Σ^* followed by a certain word x x x x x (always the same, to be determined), followed by another word in Σ^* .

c) Let $(X_k)_{k \ge 1}$ be a sequence of *i.i.d.* random variables taking values in $\{a, b\}$, with $\mathbb{P}(X_k = a) = p \in (0, 1)$ and $\mathbb{P}(X_k = b) = q = 1 - p$. Consider the random process $(Y_k)_{k \in \mathbb{N}}$ started at Y_0 , with

$$Y_1 = f(X_1,Y_0), \; Y_2 = f(X_2,Y_1),\ldots,Y_k = f(X_k,Y_{k-1}),\ldots$$

Show that the process $(Y_k)_{k \in \mathbb{N}}$ is a Markov chain on the state space $\{0, 1, 2, 3, 4, 5\}$ and write down its graph and its transition

matrix.

d) Find the average time it takes until the automaton f becomes synchronized by the random words generated from $(X_k)_{k\geq 1}$, starting from the initial state $Y_0 = (0)$.

Hint: Call $h_5(k)$ the average time it takes to reach state (5) starting from state k = 0, 1, 2, 3, 4, 5, and check first that $h_5(4) = ph_5(0)$.

e) (10 points) Let now n = 4 and consider the automaton g defined by

\boldsymbol{g}	0	1	2	3	4
\boldsymbol{a}	0	2	2	1	4
b	0	0	3	4	4

Draw a graph for this automaton.

- f) Find the unique shortest word that synchronizes the automation to state (1) after starting from all states 1, 2, 3, and the unique shortest word that synchronizes the automaton to state (4) after starting from all states 1, 2, 3.
- g) Consider the random process $(Z_k)_{k\in\mathbb{N}}$ started at Z_0 , with

$$Z_1 = g(X_1, Z_0), \; Z_2 = g(X_2, Z_1), \dots, Z_k = g(X_k, Z_{k-1}), \dots$$

Show that the process $(Z_k)_{k \in \mathbb{N}}$ is a Markov chain on the state space $\{0, 1, 2, 3, 4\}$, draw the graph of the chain and write down its transition matrix.

h) Find the probability that the first synchronized word is "abab" when the automaton is started from state (1). Note that according to Question (f), synchronization may here occur through state (0) or through state (4).

Exercise 5.26 Consider the cellular automaton g defined by

\boldsymbol{g}	0	1	2	3	4	
\boldsymbol{a}	0	2	2	1	4	
b	0	0	3	4	4	

This automaton has two sink states (0) and (4), and its graph is given as follows:



a) Find the shortest word that synchronizes this automaton to state (4) after starting from all states 1, 2, 3.

b) Consider an $\{a, b\}$ -valued two-state Markov chain $(X_n)_{n \in \mathbb{N}}$ with transition probability matrix

$$P = \left[egin{array}{cc} 1/2 & 1/2 \ & & \ 1/2 & 1/2 \ & & \ 1/2 & 1/2 \end{array}
ight]$$

and the Markov chain on $(Z_k)_{k\in\mathbb{N}}$ the state space $\{0, 1, 2, 3, 4\}$ started at Z_0 , with

$$Z_1 = g(X_1, Z_0), \; Z_2 = g(X_2, Z_1), \dots, Z_k = g(X_k, Z_{k-1}), \dots$$

Draw the graph of the chain $(Z_k)_{k \in \mathbb{N}}$ and write down its transition probability matrix.

c) Find the probability that the first synchronized word is "aabb" when the automaton is started from state (1).

Problem 5.27 Markov Decision Processes (MDPs) have applications in game theory, robotics, automated control, operations research, information theory, multi-agent systems, swarm intelligence, genetic algorithms, etc, through their use in *reinforcement learning*, and Q-learning, see here for a GridWorld-based algorithmic simulation.

A Markov Decision Process (MDP) with state space S consists in:

- a finite set \mathbb{A} of possible actions,
- a family $(P^{(a)})_{a\in\mathbb{A}}$ of transition probability matrices $(P^{(a)}_{i,j})_{i,j\in\mathbb{S}}$, and
- a policy $\pi: \mathbb{S} \to \mathbb{A}$ giving the action $\pi(k) \in \mathbb{A}$ taken at every given state in $k \in \mathbb{S}$.

When a MDP is in state $X_n = k$ at time n we look up the action $\pi(k) \in \mathbb{A}$ given by the policy π , and we generate the new value X_{n+1} using the transition probabilities $P_{k,\cdot}^{(\pi(k))} = (P_{k,l}^{(\pi(k))})_{l \in \mathbb{S}}$.

0.22 ₽	0.25	0.27 F	0.30 F	0.34 ₽	0.38	0.34	0.30 🛠	0.34	0.38
0.25	0.27	0.30	0.34	0.38	0.42	0.38	0.34 ↔	0.38	0.42
0.22					0.46				0.46
0.20	0.22 ₽	0.25	-0.78		0.52	0.57	0.64	0.57 •••	0.52
0.22 F	0.25	0.27	0.25 ••		0.08 R-1.	-0.36	0.71	0.64	0.57
0.25 F	0.27 F	0.30	0.27 ••		1.20 + R 1.0	0.08	0.79	-0.29	0.52
0.27 🗲	0.30 F*	0.34	0.30		1.08	0.97	0.87	-0.21	0.57
0.31 🗲	0.34 F	0.38	-0.58 R -1.		-0.0	-0.1	0.79	0.71	0.64
0.34	0.38	0.42	0.46	0.52	0.57	0.64	0.7	0.64	0.57
^{0.3} Ł	^{0.34}	0.38	0.42	0.46	0.52	0.57	0.64	0.57	0.52

The Pacman game is a natural application of Markov Decision Processes, see *here*.

The Tetris game can also be modeled as a Markov decision process whose (random) state consists in a given board configuration together with one of seven tile shapes. The actions are naturally chosen among the 40 placement choices for the falling tile, and a new shape $\frac{1}{100}$ shape



is chosen with uniform probability 1/7 at each time step.



The following 10 questions are *interdependent*, and should be treated in sequence.

a) Consider a Markov chain $(X_n)_{n \in \mathbb{N}}$ with state space S and transition matrix $P = (P_{i,j})_{i,j \in S}$. Derive the first step analysis equation for the value function

$$V(k) := \mathbb{E}\left[\sum_{n \ge 0} R(X_n) \mid X_0 = k\right], \quad k \in \mathbb{S},$$
(5.4.13)

defined as the total accumulated reward obtained after starting from state (k), where $R: \mathbb{S} \to \mathbb{R}$ is a reward function.*

 $\bigstar 235$

^{*}We always assume that $R(\cdot)$ and $(X_n)_{n \in \mathbb{N}}$ are such that the series in (5.4.13) are convergent.

b) In Questions (b1), (b2) and (b3) below we consider the deterministic MDP on the state space $\$ = \{1, 2, 3, 4, 5, 6, 7\}$ with actions $\mathbb{A} = \{\downarrow, \rightarrow\}$ and transition probability matrices

	$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$
	0 0 0 0 1 0 0	0 0 1 0 0 0 0
	0 0 0 0 0 1 0	0 0 1 0 0 0 0
$P^{(\downarrow)}:=$	$0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ , \qquad P^{(\to)}:$	$= \left \begin{array}{cccccccc} 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{array} \right ,$
	0 0 0 0 1 0 0	0 0 0 0 0 1 0
	0 0 0 0 0 0 0 1	0 0 0 0 0 0 1
		$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

and the reward function $R:\mathbb{S}\to\mathbb{R}$ given by

$$R(1) = 0, R(2) = -2, R(3) = -1, R(4) = -1, R(5) = -3, R(6) = 5.$$
 (5.4.14)



This MDP can be represented by the following graph, in which the \rightarrow arrows represent the policy choices, while the straight arrows denote Markov transitions. MH3512 AY19-20



b1) Compute the *optimal action-value* function*

$$Q^*(k,\downarrow) := \max_{\pi: \pi(k)=\downarrow} \mathbb{E}\left[\sum_{n \ge 0} R(X_n) \mid X_0 = k\right]$$
(5.4.15)

and

$$Q^*(k, \to) := \max_{\pi : \pi(k) = \to} \mathbb{E}\left[\sum_{n \ge 0} R(X_n) \mid X_0 = k\right], \qquad (5.4.16)$$

starting from state $X_0 = k \in \mathbb{S}$, in the following order: $Q^*(7,\downarrow)$, $Q^*(7,\rightarrow)$, $Q^*(6,\downarrow)$, $Q^*(6,\rightarrow)$, $Q^*(3,\downarrow)$, $Q^*(3,\rightarrow)$, $Q^*(5,\rightarrow)$, $Q^*(2,\downarrow)$, $Q^*(2,\rightarrow)$, $Q^*(4,\rightarrow)$, $Q^*(4,\downarrow)$, $Q^*(1,\downarrow)$, $Q^*(1,\rightarrow)$.

^{*}In the maxima (5.4.15) the action is taken equal to " \downarrow ", resp. " \rightarrow " at the first step only.

b2) Find the *optimal value* function

$$V^*(k):= \max_{\pi} \mathbb{E}\left[\sum\limits_{n \geqslant 0} R(X_n) \; \Big| \; X_0 = k
ight],$$

at all states $k = 1, 2, \ldots, 7$.

Remark: The R package MDPtoolbox can be used to check your results, however, explanations are required.

b3) Find the optimal policy $\pi^* = (\pi^*(1), \pi^*(2), \pi^*(3), \pi^*(4), \pi^*(5))$ of actions leading to the optimal gain starting from any state.^{*}

Remark: The R package MDPtoolbox can be used to check your results, however, explanations are required.

- c) Questions (c1), (c2), (c3) below are for a general MDP.
 - c1) Using first step analysis, show that for a general MDP the *optimal action-value* function^{\dagger}

$$Q^*(k,a) := \max_{\pi: \pi(k)=a} \mathbb{E}\left[\sum_{n \ge 0} R(X_n) \mid X_0 = k\right], \quad k \in \mathbb{S}, \quad a \in \mathbb{A},$$
(5.4.17)

can be written using the transition probability matrix $P^{(a)}$ and the optimal value function $V^*(\cdot)$.[‡]

- c2) How can the optimal policy $\pi^*(k)$ at state $k \in S$ be computed from the optimal action-value function $Q^*(k, a)$?
- c3) By applying first step analysis to the *optimal value* function

$$V^*(k) := \max_{\pi} \mathbb{E}\left[\sum_{n \ge 0} R(X_n) \mid X_0 = k\right], \qquad k \in \mathbb{S},$$
(5.4.18)

derive the so-called *Bellman equation* satisfied by V(k) using the *optimal action-value* function $Q^*(k, a)$.

[‡]We always assume that $R(\cdot)$ and $(X_n)_{n \in \mathbb{N}}$ are such that the series in (5.4.17) and (5.4.18) is convergent.

^{*}The values of $\pi^*(6)$ and $\pi^*(7)$ are not considered because they do not affect the total reward.

[†]In the maxima (5.4.17) the action is taken equal to a at the first step only. After moving to a new state we maximize the future reward according to the best policy choice.

d) In Questions (d1), (d2) and (d3) below we let $p \in [0, 1]$ and consider the stochastic MDP on the state space $S = \{1, 2, 3, 4, 5, 6, 7\}$ with actions $\mathbb{A} = \{\downarrow, \rightarrow\}$ and transition probability matrices

	[0	0	0	1	0	0	0		[0]	1	0	0	0	0	0]
	0	0	0	0	1	0	0		0	0	\boldsymbol{p}	0	${oldsymbol{q}}$	0	0	
	0	0	0	0	0	1	0		0	0	1	0	0	0	0	
$P^{(\downarrow)} :=$	0	0	0	1	0	0	0	$P^{(ightarrow)}:=$	0	0	0	0	1	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	1	
	0	0	0	0	0	0	1		0	0	0	0	0	0	1	

and the reward function (5.4.14). This MDP can be represented by the following graph, in which the \rightarrow arrows represent policy choices, while the straight arrows denote Markov transitions.



d1) Using the argument of Question (c1), compute the *optimal action-value* function^{*}

$$Q^*(k,\downarrow) := \max_{\pi: \pi(k)=\downarrow} \mathbb{E}\left[\sum_{n \ge 0} R(X_n) \mid X_0 = k\right]$$
(5.4.19)

and

$$Q^*(k, \to) := \max_{\pi : \pi(k) = \to} \mathbb{E}\left[\sum_{n \ge 0} R(X_n) \mid X_0 = k\right], \qquad (5.4.20)$$

starting from state $X_0 = k \in S$, in the following order: $Q^*(7,\downarrow), Q^*(7,\rightarrow), Q^*(6,\downarrow), Q^*(6,\rightarrow), Q^*(3,\downarrow), Q^*(3,\rightarrow), Q^*(5,\rightarrow), Q^*(5,\downarrow), Q^*(2,\downarrow), Q^*(4,\rightarrow), Q^*(4,\downarrow), Q^*(1,\downarrow), Q^*(1,\rightarrow).$

Remark: Some values of $Q^*(k,\downarrow)$, $Q^*(k,\rightarrow)$ may now depend on p.

d2) Using the result of Question (d1), find the *optimal value* function

$$V^*(k):= \max_{\pi} \mathbb{E} \left[\sum_{n \geqslant 0} R(X_n) \; \Big| \; X_0 = k
ight],$$

at all states $k = 1, 2, \ldots, 7$, depending on the value of $p \in [0, 1]$.

Remark: The R package MDPtoolbox can be used to check your results, however, explanations are required.

d3) Find the optimal policy $\pi^* = (\pi^*(1), \pi^*(2), \pi^*(3), \pi^*(4), \pi^*(5))$ of actions leading to the optimal gain starting from any state, depending on the value of $p \in [0, 1]$.[†]

Remark: The R package MDPtoolbox can be used to check your results, however, explanations are required.

(240)

^{*}In the maxima (5.4.19) the action is taken equal to " \downarrow ", resp. " \rightarrow " at the first step only.

[†]The values of $\pi^*(6)$ and $\pi^*(7)$ are not considered because they do not affect the total reward.

```
install.packages("MDPtoolbox")
library(MDPtoolbox)
P <- array(0,c(7,7,2))
P[,,1] <- matrix(c(
.,...,,
.,...,), nrow=7, ncol=7, byrow=TRUE)
P[,,2] <- matrix(c(
.,...,,
.,...,), nrow=7, ncol=7, byrow=TRUE)
R <- array(0,c(7,2))
R[,1] <- matrix(c(.,...,), nrow=1, ncol=7, byrow=TRUE)
R[,2] <- R[,1]
mdp_check(P, R)
mdp_value_iteration(P,R,discount=1)
```

241

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	
6.2	Recurrent States	
6.3	Transient States	
6.4	Positive vs Null Recurrence	
6.5	Periodicity and Aperiodicity	
Exer	$rcises \ldots $	

6.1 Communicating States

Definition 6.1

A state $(j) \in S$ is to be <u>accessible</u> from another state $(i) \in S$, and we write $(i) \mapsto (j)$, if there exists a *finite* integer $n \ge 0$ 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 (i) to (j) with non-zero probability in a certain (random) number of steps. We also say that state (i) leads to state (j), and when $i \neq j$ we have

$$\mathbb{P}(T_j^r < \infty \mid X_0 = i) \geqslant \mathbb{P}(T_j^r \leqslant n \mid X_0 = i) \geqslant \mathbb{P}(X_n = j \mid X_0 = i) > 0.$$

MH3512 AY19-20

Remark 6.2

Since $P^0 = I_d$ 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 (i) is always accessible from itself (in zero time steps) even if $P_{i,i} = 0$.

In case (i) \mapsto (j) and (j) \mapsto (i) we say that (i) and (j) communicate^{*} and we write (i) \leftrightarrow (j).

The binary relation " \leftrightarrow " is a called an equivalence relation as it satisfies the following properties:[†]

a) <u>Reflexivity</u>:

For all $i \in \mathbb{S}$ we have (i) \longleftrightarrow (i).

b) <u>Symmetry</u>:

For all $i, j \in S$ we have that $(i) \longleftrightarrow (j)$ is equivalent to $(j) \longleftrightarrow (i)$.

c) <u>Transitivity</u>:

For all $i, j, k \in S$ such that $(i) \longleftrightarrow (j)$ and $(j) \longleftrightarrow (k)$, we have $(i) \longleftrightarrow (k)$.

Proof

It is clear that the relation \longleftrightarrow is reflexive and symmetric. The proof of transitivity can be stated as follows. If $(i) \mapsto (j)$ and

^{*}In graph theory, one says that (i) and (i) are strongly connected.

[†]Please refer to MH1300 - Foundations of Mathematics for more information on equivalence classes.

(j) \mapsto (k), there exists $a \ge 1$ and $b \ge 1$ such that

$$[P^a]_{i,j} > 0, \qquad \quad [P^b]_{j,k} > 0.$$

Next, by (4.1.2), for all $n \ge a + b$ we have

$$P(X_{n} = k \mid X_{0} = i)$$

$$= \sum_{l,m\in\mathbb{S}} \mathbb{P}(X_{n} = k, X_{n-b} = l, X_{a} = m \mid X_{0} = i)$$

$$= \sum_{l,m\in\mathbb{S}} \mathbb{P}(X_{n} = k \mid X_{n-b} = l) \mathbb{P}(X_{n-b} = l \mid X_{a} = m) \mathbb{P}(X_{a} = m \mid X_{0} = i)$$

$$\geq \mathbb{P}(X_{n} = k \mid X_{n-b} = j) \mathbb{P}(X_{n-b} = j \mid X_{a} = j) \mathbb{P}(X_{a} = j \mid X_{0} = i)$$

$$= [P^{a}]_{i,j} [P^{n-a-b}]_{j,j} [P^{b}]_{j,k}$$

$$\geq 0.$$
(6.1.1)

The conclusion follows by taking n = a + b, in which case we have

$$\mathbb{P}(X_n=k\mid X_0=i)\geqslant [P^a]_{i,j}[P^b]_{j,k}>0.$$

The equivalence relation ' \longleftrightarrow '' induces a *partition* of S into disjoint classes A_1, A_2, \ldots, A_m such that $S = A_1 \cup \cdots \cup A_m$, and a) we have (i) \longleftrightarrow (j) for all $i, j \in A_q$, and

b) we have (i) \longleftrightarrow (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 6.3

A Markov chain whose state space is made of a unique communicating class is said to be <u>irreducible</u>, otherwise the chain is said to be <u>reducible</u>.

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)</pre>

Clearly, all states in S communicate when $(X_n)_{n \in \mathbb{N}}$ is irreducible. In case the *i*-th column of a transition matrix P vanishes, *i.e.* $P_{k,i} = 0, i \in S$, then state (i) cannot be reached from any other state and (i) becomes a communicating class on its own, as is the case of state (1) in Exercise 4.10 for $n \ge 2$, 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 (i) 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 " \leftrightarrow ".



The above state space $S = \{0, 1, 2, 3\}$ is partitioned into two communicating classes which are $\{0, 1, 2\}$ and $\{3\}$.

6.2 Recurrent States

Definition 6.4

A state $(i) \in S$ is said to be <u>recurrent</u> if, starting from state (i), the chain will return to state (i) within a finite (random) time, with probability 1, *i.e.*,

$$p_{i,i} := \mathbb{P}(T_i^r < \infty \mid X_0 = i) = \mathbb{P}(X_n = i \text{ for some } n \ge 1 \mid X_0 = i) = 1.$$

$$(6.2.1)$$

The next Proposition 6.5 uses the mean number of returns R_i to state (i) defined in (5.4.2).

Proposition 6.5 For any state (i) \in S, the following statements are equivalent: i) the state (i) \in S is recurrent, *i.e.* $p_{i,i} = 1$, ii) the number of returns to (i) $\in \mathbb{S}$ is a.s.^{*a*} infinite, *i.e.* $\mathbb{P}(R_i = \infty \mid X_0 = i) = 1$, i.e. $\mathbb{P}(R_i < \infty \mid X_0 = i) = 0$, (6.2.2)iii) the mean number of returns to $(i) \in S$ is *infinite*, *i.e.* $\mathbb{E}[R_i \mid X_0 = i] = \infty.$ (6.2.3)iv) we have $\sum\limits_{n\geq 1}f_{i,i}^{(n)}=1,$ (6.2.4)where $f_{i,i}^{(n)} := \mathbb{P}(T_i^r = n \mid X_0 = i), n \ge 1$, is the distribution of T_i^r . ^aalmost surely 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 (1) is recurrent for the random walk of Chapter 3 when p = q = 1/2, while it is not recurrent if $p \neq q$ as by (3.4.14) we have

$$p_{0,0} = \mathbb{P}(T_0 < \infty) = 2\min(p,q).$$
 (6.2.5)

As a consequence of (6.2.3), we have the following result.

A state $i \in S$ is recurrent if and only if

$$\sum\limits_{n \geqslant 1} [P^n]_{i,i} = \infty_{i,i}$$

i.e. the above series *diverges*.

Corollary 6.6

Proof

For all $i, j \in S$, by (5.4.7) we have

$$\mathbb{E}[R_j \mid X_0 = i] = \mathbb{E}\left[\sum_{n \ge 1} \mathbb{1}_{\{X_n = j\}} \mid X_0 = i\right] = \sum_{n \ge 1} \mathbb{E}[\mathbb{1}_{\{X_n = j\}} \mid X_0 = i]$$
$$= \sum_{n \ge 1} \mathbb{P}(X_n = j \mid X_0 = i) = \sum_{n \ge 1} [P^n]_{i,j}, \qquad (6.2.6)$$

as in (5.4.7). To conclude we let j = i and apply (6.2.3).

Corollary 6.6 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 6.7

Let $(j) \in S$ be a recurrent state. Then any state $(i) \in S$ that communicates with state (j) is also recurrent.

Proof

By definition, since (i) \longmapsto (j) and (j) \longmapsto (i), there exists $a \ge 1$ and $b \ge 1$ such that

n

$$[P^a]_{i,j} > 0 \qquad [P^b]_{j,i} > 0,$$

and from (6.1.1) applied with k = i we find

$$egin{array}{ll} \sum\limits_{\geqslant a+b} [P^n]_{i,i} &=& \sum\limits_{n\geqslant a+b} \mathbb{P}(X_n=i\mid X_0=i) \ &\geqslant& [P^a]_{i,j}[P^b]_{j,i}\sum\limits_{n\geqslant a+b} P^{n-a-b}_{j,j} \ &=& [P^a]_{i,j}[P^b]_{j,i}\sum\limits_{n\geqslant 0} [P^n]_{j,j} \ &=& \infty, \end{array}$$

which shows that state (i) is recurrent from Corollary 6.6 and the assumption that state (j) 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 $(i) \in S$ is said to be <u>transient</u> when it is not recurrent, *i.e.*, by (6.2.1),

$$p_{i,i} = \mathbb{P}(T_i^r < \infty \mid X_0 = i) = \mathbb{P}(X_n = i \text{ for some } n \ge 1 \mid X_0 = i) < 1, \tag{6.3.1}$$

or

$$\mathbb{P}(T_i^r = \infty \mid X_0 = i) > 0$$

MH3512 AY19-20

(5248)

(5249)

Proposition 6.8

For any state $(i) \in S$, the following statements are equivalent:

- i) the state (i) \in **S** is transient, *i.e.* $p_{i,i} < 1$,
- ii) the number of returns to $(i) \in S$ is a.s.^{*a*} finite, *i.e.*

$$\mathbb{P}(R_i = \infty \mid X_0 = i) = 0$$
, i.e. $\mathbb{P}(R_i < \infty \mid X_0 = i) = 1$, (6.3.2)

iii) the mean number of returns to $(i) \in S$ is *finite*, *i.e.*

$$\mathbb{E}[R_i \mid X_0 = i] < \infty, \tag{6.3.3}$$

iv) we have

$$\sum_{n \ge 1} f_{i,i}^{(n)} < 1.$$
(6.3.4)

where
$$f_{i,i}^{(n)} := \mathbb{P}(T_i^r = n \mid X_0 = i), n \ge 1$$
, is the distribution of T_i^r .

 a almost surely

Proof

This is a direct consequence of Proposition 6.5 and the definition (6.3.1) of transience. Regarding point (ii) and the Condition (6.3.2) we also note that the state $(i) \in S$ 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 (i) $\in S$ is transient if and only if

$$\mathbb{P}(R_i < \infty \mid X_0 = i) > 0,$$
 MH3512 AY19-2

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 $i \in S$ is finite with a non-zero probability which is necessarily equal to one. As a consequence of Corollary 6.6 we also have the following result.

Corollary 6.9

A state $i \in \mathbb{S}$ is transient if and only if

$$\sum\limits_{n\geqslant 1} [P^n]_{i,i} < \infty$$

i.e. the above series *converges*.

By Corollary 6.9 and the relation

$$\sum\limits_{n \geqslant 0} \left[P^n
ight]_{i,j} = \left[(I_d-P)^{-1}
ight]_{i,j}, \qquad i,j \in \mathbb{S},$$

we find that a chain with finite state space is transient if the matrix $I_d - P$ is invertible, cf. Problem 4.13.

Clearly, any absorbing state is recurrent, and any state that leads to an absorbing state is transient.

In addition, if a state $(i) \in S$ communicates with a transient state (j) then (i) is also transient (otherwise the state (j) would be recurrent by Corollary 6.7). 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 Section 4.5, Relations (4.5.4) and (4.5.5) show that

♦ 250

MH3512 AY19-20

hence state (0) is transient if b = 0 and a > 0, and recurrent otherwise. Similarly we have

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 $g_i(j) = \mathbb{P}(T_i < \infty \mid X_0 = j), i, j \in \{0, 1\}, i.e.$

 $\left\{egin{array}{l} g_0(0) = ag_0(1) + 1 - a \ g_0(1) = b + (1 - b)g_0(1) \ g_1(0) = (1 - a)g_1(0) + a \ g_1(1) = bg_1(0) + 1 - b, \end{array}
ight.$

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 6.10

Let $(X_n)_{n \in \mathbb{N}}$ be a Markov chain with *finite* state space S. Then $(X_n)_{n \in \mathbb{N}}$ has at least one recurrent state.

Proof

Recall that from (5.4.5) we have

$$\mathbb{E}[R_j \mid X_0 = i] = p_{i,j}(1-p_{j,j})\sum\limits_{n \geqslant 1} n(p_{j,j})^{n-1} = rac{p_{i,j}}{1-p_{j,j}}$$

for any states (i), (j) \in S. Assuming that the state (j) \in S is transient we have $p_{j,j} < 1$ by (6.3.1), hence

$$\mathbb{E}[R_j \mid X_0 = i] = \sum\limits_{n \geqslant 1} [P^n]_{i,j} < \infty,$$

by (6.3.3) which implies^{*a*} that

 $\lim_{n
ightarrow\infty}[P^n]_{i,j}=0$

for all transient states $j \in S$. In case all states in S were transient, since S is finite, by the *law of total probability* (4.2.2) we would have

$$0 = \sum_{j \in \mathbb{S}} \lim_{n o \infty} [P^n]_{i,j} = \lim_{n o \infty} \sum_{j \in \mathbb{S}} [P^n]_{i,j} = \lim_{n o \infty} 1 = 1,$$

which is a contradiction. Hence not all states can be transient, and there exists at least one recurrent state.

^{*a*}For any sequence $(a_n)_{n \ge 0}$ of *nonnegative* real numbers, $\sum_{n \ge 0} a_n < \infty$ implies $\lim_{n \to \infty} a_n = 0$.

Exercises:

i) Find which states are transient and recurrent in the chain (5.4.8).


State (3) is clearly recurrent since we have $T_3^r = 1$ with probability one when $X_0 = 3$. State (2) is transient because

$$1 - p_{2,2} = \mathbb{P}(T_2^r = \infty \mid X_0 = 2) = \frac{4}{7} \ge \mathbb{P}(X_1 = 3 \mid X_0 = 2) = 0.5 > 0, \tag{6.3.5}$$

and state (1) is transient because

$$\mathbb{P}(T_1^r = \infty \mid X_0 = 1) = 0.8 \ge \mathbb{P}(X_1 = 3 \mid X_0 = 1) = 0.6,$$
(6.3.6)

see the Exercise 5.5 for the computations of

$$p_{1,1} = \mathbb{P}(T_2^r < \infty \mid X_0 = 2) = 0.8$$

and

$$p_{2,2} = \mathbb{P}(T_2^r < \infty \mid X_0 = 2) = rac{3}{7}.$$

By Corollary 6.7, 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 Chapter 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

$$\mathbb{P}(S_{2n}=0)={\binom{2n}{n}}p^nq^n \quad ext{ and } \quad \mathbb{P}(S_{2n+1}=0)=0, \quad n\in\mathbb{N}.$$

MH3512 AY19-20

♦ 253

Hence

$$egin{aligned} &\sum\limits_{n \geqslant 0} [P^n]_{0,0} \;=\; \sum\limits_{n \geqslant 0} \mathbb{P}(S_n = 0 \mid S_0 = 0) = \sum\limits_{n \geqslant 0} \mathbb{P}(S_{2n} = 0 \mid S_0 = 0) \ &=\; \sum\limits_{n \geqslant 0} {2n \choose n} p^n q^n = H(1) = rac{1}{\sqrt{1 - 4pq}}, \end{aligned}$$

and

$$\mathbb{E}[R_0 \mid S_0 = 0] = \sum\limits_{n \geqslant 1} \mathbb{P}(S_n = 0 \mid S_0 = 0) = rac{1}{\sqrt{1 - 4pq}} - 1,$$

where H(s) is defined in (3.4.8), cf. also Problem 3.11-(f).

Consequently, by Corollary 6.6, all states $(i) \in \mathbb{Z}$ are recurrent when p = q = 1/2, whereas by Corollary 6.9 they are all transient when $p \neq q$, cf. Corollary 6.7.

Alternatively we could reach the same conclusion by directly using (3.4.14) and (6.2.1) which state that

$$\mathbb{P}(T^r_i < \infty \mid X_0 = i) = 2\min(p,q).$$

6.4 Positive vs Null Recurrence

The expected time of return (or mean recurrence time) to a state $(i) \in S$ is given by

$$egin{aligned} \mu_i(i) &:= & \mathbb{E}[T_i^r \mid X_0 = i] \ &= & \sum\limits_{n \geqslant 1} n \mathbb{P}(T_i^r = n \mid X_0 = i) \ &= & \sum\limits_{n \geqslant 1} n f_{i,i}^{(n)}. \end{aligned}$$

Recall that when state (i) 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 (i), 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 6.11

- A *recurrent* state $i \in S$ is said to be:
- a) positive recurrent if the mean return time to (i) is *finite*, *i.e.*

$$\mu_i(i) = \mathbb{E}[T^r_i \mid X_0 = i] < \infty,$$

b) null recurrent if the mean return time to (i) is *infinite*, *i.e.*

$$\mu_i(i) = \mathbb{E}[T^r_i \mid X_0 = i] = \infty .$$

Exercise: Which states are positive/null recurrent in the simple random walk $(S_n)_{n \in \mathbb{N}}$ of Chapter 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 6.12 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 6.12

Assume that the state space S of a Markov chain $(X_n)_{n \in \mathbb{N}}$ is *finite*. Then all recurrent states in S are also positive recurrent.

As a consequence of Definition 6.3, Corollary 6.7, and Theorems 6.10 and 6.12 we have the following corollary.

Corollary 6.13

Let $(X_n)_{n \in \mathbb{N}}$ be an irreducible Markov chain with *finite* state space S. Then *all* states of $(X_n)_{n \in \mathbb{N}}$ are positive recurrent.

4H3512 AY19-2(

(255)

6.5 Periodicity and Aperiodicity

Given a state $i \in S$, consider the sequence

 $\{n \geqslant 1 \; : \; [P^n]_{i,i} > 0\}$

of integers which represent the possible travel times from state (i) to itself.

 Definition 6.14

 The period of the state $i \in S$ is the greatest common divisor^a of the sequence

 $\{n \ge 1 : [P^n]_{i,i} > 0\}.$

 ^aPlease refer to MH1300 - Foundations of Mathematics or MH1301 Discrete Mathematics for more information on greatest common divisors.

A state having period 1 is said to be <u>aperiodic</u>, 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 $i \in S$ is said to be <u>ergodic</u> if it is both positive recurrent and aperiodic.

If $[P^n]_{i,i} = 0$ for all $n \ge 1$ then the set $\{n \ge 1 : [P^n]_{i,i} > 0\}$ is empty and by convention the period of state (i) is defined to be 0. In this case, state (i) is also transient.

Note also that if

$$\{n\geqslant 1 \;:\; [P^n]_{i,i}>0\}$$

contains two distinct numbers that are relatively prime to each other (*i.e.* their greatest common divisor is 1) then state (i) aperiodic.

Proposition 6.15 shows that periodicity is a *class property*, as all states in a given communicating class have same periodicity.

Proposition 6.15

All states that belong to a same communicating class have the same period.

Proof

 $\circ 257$

(See *e.g. here*). Assume that state (i) has period d_i , that (j) communicates with (i), and let $n \in \{m \ge 1 : [P^m]_{j,j} > 0\}$. Since (i) and (j) communicate, there exists $k, l \ge 1$ 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 $d_j \ge d_i$. Exchanging the roles of (i) and (j) we obtain similarly that $d_i \ge d_j$.

A Markov chain is said to be <u>aperiodic</u> 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.



Here we have

$$egin{aligned} \{n \geqslant 1 \ : \ [P^n]_{0,0} > 0\} &= \{2,4,6,8,10,\ldots\}, \ \{n \geqslant 1 \ : \ [P^n]_{1,1} > 0\} &= \{2,4,6,8,10,\ldots\}, \ \{n \geqslant 1 \ : \ [P^n]_{2,2} > 0\} &= \{4,6,8,10,12,\ldots\}, \ \{n \geqslant 1 \ : \ [P^n]_{3,3} > 0\} &= \{4,6,8,10,12,\ldots\}, \end{aligned}$$

hence all states have period **2**, and this is also consequence of Proposition 6.15.

iii) Consider the following chain:



Here we have

 $\{n \geqslant 1 : [P^n]_{0,0} > 0\} = \{4, 5, 6, 7, \ldots\},\$

(6.5.1)

MH3512 AY19-20

$$egin{aligned} \{n \geqslant 1 \ : \ [P^n]_{1,1} > 0\} &= \{4,5,6,7,\ldots\}, \ \{n \geqslant 1 \ : \ [P^n]_{2,2} > 0\} &= \{4,5,6,7,\ldots\}, \ \{n \geqslant 1 \ : \ [P^n]_{3,3} > 0\} &= \{4,5,6,7,\ldots\}, \end{aligned}$$

hence *all* states have period $\mathbf{1}$, see also Proposition 6.15.

iv) Next, consider the modification of (6.5.1):



Here the chain is aperiodic since we have

$$egin{aligned} \{n \geqslant 1 \ : \ [P^n]_{0,0} > 0\} &= \{2,3,4,5,6,7,\ldots\}, \ \ \{n \geqslant 1 \ : \ [P^n]_{1,1} > 0\} &= \{2,3,4,5,6,7,\ldots\}, \ \ \{n \geqslant 1 \ : \ [P^n]_{2,2} > 0\} &= \{3,4,5,6,7,8,\ldots\}, \ \ \ \{n \geqslant 1 \ : \ [P^n]_{3,3} > 0\} &= \{4,6,7,8,9,10,\ldots\}, \end{aligned}$$

hence *all* states have period $\mathbf{1}$.

Exercises:

i) What is the periodicity of the simple random walk $(S_n)_{n \in \mathbb{N}}$ of Chapter 3 on $\mathbb{S} = \mathbb{Z}$?

By (3.3.3) We have

$$[P^{2n}]_{i,i}={\binom{2n}{n}}p^nq^n>0 \quad ext{and} \quad [P^{2n+1}]_{i,i}=0, \qquad n\in\mathbb{N},$$

MH3512 AY19-20

hence

and the chain has period **2**.

ii) Find the periodicity of the chain (5.4.8).



States (0, (1), (2)) and (3) have period 1, hence the chain is aperiodic.

iii) The chain of Figure 4.2 is aperiodic since it is irreducible and state (3) has a returning loop.

Exercises

Exercise 6.1 Consider a Markov chain $(X_n)_{n \ge 0}$ on the state space $\{0, 1, 2, 3\}$, with transition matrix

$$\left[\begin{array}{ccccc} 1/3 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]^{\prime}$$

MH3512 AY19-20

(5260)

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

1/3	1/3	1/3	
1/4	3/4	0	.
0	0	1	

- 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 $(X_n)_{n \ge 0}$ on the state space $\{0, 1, 2, 3, 4\}$, with transition matrix

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

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

$$\left[\begin{array}{cccc} P_{i,j} \end{array}
ight]_{0\leqslant i,j\leqslant 5} = \left[egin{array}{cccccc} 1/2 & 0 & 1/4 & 0 & 0 & 1/4 \ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \ 0 & 0 & 0 & 0 & 1 & 0 \ 0 & 0 & 0 & 0 & 1 & 0 \ 1/6 & 1/2 & 1/6 & 0 & 0 & 1/6 \ 0 & 0 & 1 & 0 & 0 & 0 \ 0 & 0 & 0 & 0 & 0 & 1 \end{array}
ight].$$

- 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

0.8	0	0.2	0	
0	0	1	0	
1	0	0	0	
0.3	0.4	0	0.3	

- 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 $\{0, 1, 2, \dots, N\}$, representing the number of balls in the first box. MH3512 AY19-20 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

$$P_{i,i-1}=rac{1}{lpha+1}, \quad P_{i,i+1}=rac{lpha}{lpha+1}, \qquad i\geqslant 1.$$

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 $\alpha < 1$.

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	265
7.2 Stationary Distributions	268
7.3 Markov Chain Monte Carlo	284
$\operatorname{Exercises}$	288

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 $i, j \in S$, and

ii) they form a probability distribution on S, *i.e.*

$$\sum_{i \in \mathbb{S}} \lim_{n \to \infty} \mathbb{P}(X_n = j \mid X_0 = i) = 1,$$
(7.1.2)

MH3512 AY19-20

for all $i \in S$.

Note that Condition (7.1.2) is always satisfied if the limits (7.1.1) exist and the state space 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

$$[\pi_0, \pi_1] = \left[\frac{b}{a+b}, \frac{a}{a+b}\right],\tag{7.1.3}$$

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

$$(\mu_0(0),\mu_1(1))=\left(1+rac{a}{b},1+rac{b}{a}
ight),$$

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

$$\lim_{n o\infty} \mathbb{P}(X_n=j\mid X_0=i)=rac{1}{\mu_j(j)}, \qquad i,j\in\mathbb{S},$$

MH3512 AY19-20

(7.1.4)

independently of the initial state $i \in S$, where

$$\mu_j(j) = \mathbb{E}[T^r_i \mid X_0 = j] \in [1,\infty]$$

is the mean return time to state $(j) \in S$.

In Theorem 7.2, Condition (i), resp. Condition (ii), is satisfied from Proposition 6.15, resp. from Proposition 6.7, 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, cf. for example Exercise 7.15-(b) below.

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 Chapter 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 $i, j \in S$ 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 Chapter 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 $\mathbb{S} = \{0, 1, \dots, N\}$, whose transition matrix P is regular, *i.e.* there exists $n \ge 1$ 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 $\pi = (\pi_i)_{i=0,1,\dots,N}$ given by

$$\pi_j = \lim_{n \to \infty} \mathbb{P}(X_n = j \mid X_0 = i), \qquad 0 \leq i, j \leq N,$$
(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 6.10.

Proposition 7.4

Let $(X_n)_{n \in \mathbb{N}}$ be a Markov chain with a transient state $(j) \in S$. Then we have

$$\lim_{n o\infty} \mathbb{P}(X_n=j\mid X_0=i)=0,$$

for all $(i) \in S$.

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) page 215, the expected number of returns to (j) starting from state (i) is finite:^{*a*}

$$egin{aligned} \mathbb{E}[R_j \mid X_0 = i] &= \mathbb{E}\left[\sum\limits_{n \geqslant 1} \mathbbm{1}_{\{X_n = j\}} \mid X_0 = i
ight] \ &= \sum\limits_{n \geqslant 1} \mathbb{E}\left[\mathbbm{1}_{\{X_n = j\}} \mid X_0 = i
ight] \ &= \sum\limits_{n \geqslant 1} \mathbb{P}(X_n = j \mid X_0 = i) \ &= rac{p_{ij}}{1 - p_{ij}} < \infty. \end{aligned}$$

The convergence of the above series implies the convergence to $\mathbf{0}$ of its general term, *i.e.*

$$\lim_{n o \infty} \mathbb{P}(X_n = j \mid X_0 = i) = 0$$

for all $(i) \in S$, which is the expected conclusion.

 a The exchange of infinite sums and expectation is valid in particular for nonnegative series.

7.2 Stationary Distributions

Definition 7.5

A probability distribution on S is any a family $\pi = (\pi_i)_{i \in S}$ 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 is said to be <u>stationary</u> 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

$$\mathbb{P}(X_1=i)=\mathbb{P}(X_0=i)=\pi_i, \ \ i\in\mathbb{S},$$

at time 1. This also means that

$$\pi_j = \mathbb{P}(X_1=j) = \sum\limits_{i\in\mathbb{S}} \mathbb{P}(X_1=j\mid X_0=i)\mathbb{P}(X_0=i) = \sum\limits_{i\in\mathbb{S}} \pi_i P_{i,j}, \hspace{1em} j\in\mathbb{S},$$

i.e. the distribution π is stationary if and only if the vector π is invariant (or stationary) by the matrix P, that means

$$\pi = \pi P. \tag{7.2.1}$$

(5269)

Note that in contrast with (5.1.3), the multiplication by P in (7.2.1) is on the <u>right</u> hand side and <u>not</u> on the left. The relation (7.2.1) can be rewritten as the *balance condition*

$$\sum_{i \in \mathbb{S}} \pi_i P_{i,k} = \pi_k = \pi_k \sum_{j \in \mathbb{S}} P_{k,j} = \sum_{j \in \mathbb{S}} \pi_k P_{k,j},$$
(7.2.2)
MH3512 AY19-20

which can be illustrated as follows:



Figure 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

$$egin{aligned} \mathbb{P}(X_{n+1}=j) &= \sum\limits_{i\in\mathbb{S}}\mathbb{P}(X_{n+1}=j\mid X_n=i)\mathbb{P}(X_n=i)\ &= \sum\limits_{i\in\mathbb{S}}P_{i,j}\mathbb{P}(X_n=i) = \sum\limits_{i\in\mathbb{S}}P_{i,j}\pi_i\ &= \ [\pi P]_j = \pi_j, \quad j\in\mathbb{S}, \end{aligned}$$

since the transition matrix of $(X_n)_{n \in \mathbb{N}}$ is time homogeneous, hence

$$\mathbb{P}(X_n = j) = \pi_j, \quad j \in \mathbb{S}, \implies \mathbb{P}(X_{n+1} = j) = \pi_j, \quad j \in \mathbb{S}.$$

By induction on $n \ge 0$, this yields

$$\mathbb{P}(X_n=j)=\pi_j, \qquad j\in\mathbb{S}, \quad n\geqslant 1,$$
 MH3512 AY19-20

i.e. the chain $(X_n)_{n \in \mathbb{N}}$ remains in the same distribution π at all times $n \ge 1$, provided that it has been started with the stationary distribution π at time n = 0.

Proposition 7.7

Assume that $\mathbb{S} = \{0, 1, \dots, N\}$ is finite and that the limits

$$\pi_j^{(i)} := \lim_{n o \infty} \mathbb{P}(X_n = j \mid X_0 = i)$$

exist for all $i, j \in S$, *i.e.* we have

$$\lim_{n o \infty} P^n = \left[egin{array}{cccc} \pi_0^{(0)} & \pi_1^{(0)} & \cdots & \pi_N^{(0)} \ \pi_0^{(1)} & \pi_1^{(1)} & \cdots & \pi_N^{(1)} \ dots & dots & dots & dots \ dots & dots & dots & dots \ \pi_1^{(N)} & \pi_1^{(N)} & \cdots & \pi_N^{(N)} \ \pi_0^{(N)} & \pi_1^{(N)} & \cdots & \pi_N^{(N)} \end{array}
ight].$$

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

$$\pi^{(i)} = \pi^{(i)} P,$$
 (7.2.3)

i.e. $\pi^{(i)}$ is invariant (or stationary) by $P, i = 0, 1, \dots, N$.

Proof

We have

$$egin{aligned} \pi_j^{(i)} &:= \lim_{n o \infty} \mathbb{P}(X_n = j \mid X_0 = i) \ &= \lim_{n o \infty} \mathbb{P}(X_{n+1} = j \mid X_0 = i) \ &= \lim_{n o \infty} \sum_{l \in \mathbb{S}} \mathbb{P}(X_{n+1} = j \mid X_n = l) \mathbb{P}(X_n = l \mid X_0 = i) \ &= \lim_{n o \infty} \sum_{l \in \mathbb{S}} P_{l,j} \mathbb{P}(X_n = l \mid X_0 = i) \ &= \sum_{l \in \mathbb{S}} P_{l,j} \lim_{n o \infty} \mathbb{P}(X_n = l \mid X_0 = i) \end{aligned}$$

MH3512 AY19-20

$$= \sum\limits_{l\in\mathbb{S}}\pi_l^{(i)}P_{l,j}, \qquad i,j\in\mathbb{S},$$

where we exchanged limit and summation because the state space 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), *i.e.*

$$\lim_{n o \infty} P^n = egin{bmatrix} \pi_0 & \pi_1 & \cdots & \pi_N \ \pi_0 & \pi_1 & \cdots & \pi_N \ dots & dots & \ddots & dots \ \pi_0 & \pi_1 & \cdots & \pi_N \end{bmatrix}$$

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

$$\pi_i:=\lim_{n o\infty}\mathbb{P}(X_n=i\mid X_0=j)=\lim_{n o\infty}[P^n]_{j,i}=rac{1}{\mu_i(i)}, \qquad i,j\in \mathbb{S},$$

which also forms a stationary distribution $(\pi_i)_{i\in S} = (1/\mu_i(i))_{i\in S}$, uniquely determined by the equation

 $\pi = \pi P.$

In Theorem 7.8 above, Condition (ii), is satisfied from Proposition 6.15, provided that at least one state is aperiodic, since the charmed irreducible.

See Exercise 7.22 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, ..., n, with n = 7.



(7.2.4)

In view of Theorem 6.12, 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

$$\pi_i := \lim_{n o \infty} \mathbb{P}(X_n = i \mid X_0 = j) = rac{1}{\mu_i(i)}, \qquad i,j \in \mathbb{S},$$

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. 3274

The convergence of the two-state chain to its stationary distribution has been illustrated in Figure 4.5. Before proceeding further we make some comments on the assumptions of Theorems 7.2 and 7.8.

Remarks

- <u>Irreducibility</u>.

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 $S = \{0, 1, 2\}$ 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 S is a finite set it admits at least one communicating class $A \subset S$ that leads to no other class, and it admits a stationary distribution π_A by Corollary 7.11 since it is irreducible, hence a chain with finite state space 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

$$P = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right]$$

is not aperiodic (both states have period $\mathbf{2}$) and it has no limiting distribution because^{*}

$$\mathbb{P}(X_{2n} = 1 \mid X_0) = 1$$
 and $\mathbb{P}(X_{2n+1} = 1 \mid X_0) = 0$, $n \in \mathbb{N}$.

The chain does have an invariant (or stationary) distribution π solution of $\pi = \pi P$, and given by

$$\pi=[\pi_0,\pi_1]=\left[rac{1}{2},rac{1}{2}
ight].$$

- Positive recurrence.

Theorem 7.8, Theorem 7.10 below, and Corollary 7.9 do not apply to the unrestricted random walk $(S_n)_{n \in \mathbb{N}}$ of Chapter 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

$$\pi_{i+1}-\pi_i=rac{p}{q}(\pi_i-\pi_{i-1}),\qquad i\in\mathbb{Z},$$

(5275)

^{*}This two-state chain is a particular case of the circular chain (7.2.4) for n = 2.

As in the direct solution method of page 78, this implies

$$\pi_{i+1}-\pi_i=\left(rac{p}{q}
ight)^i(\pi_1-\pi_0),\qquad i\in\mathbb{N},$$

so that by a telescoping summation argument we have

$$egin{array}{lll} \pi_k &=& \pi_0 + \sum\limits_{i=0}^{k-1} (\pi_{i+1} - \pi_i) \ &=& \pi_0 + (\pi_1 - \pi_0) \sum\limits_{i=0}^{k-1} \left(rac{p}{q}
ight)^i \ &=& \pi_0 + (\pi_1 - \pi_0) rac{1 - (p/q)^k}{1 - p/q}, \qquad k \in \mathbb{N}, \end{array}$$

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

$$egin{aligned} &\lim_{n o \infty} \mathbb{P}(S_{2n} = 2k \mid S_0 = 0) \; = \; \lim_{n o \infty} rac{(2n)!}{(n+k)!(n-k)!} p^{n+k} q^{n-k} \ &\leqslant \; \lim_{n o \infty} rac{(2n)!}{2^{2n} n!^2} \ &= \; \lim_{n o \infty} rac{1}{\sqrt{\pi n}} = 0, \qquad k \in \mathbb{N}, \end{aligned}$$

as in Problem 3.11-(f), 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 $\frac{2}{2}$ 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

$$\pi_i=rac{1}{\mu_i(i)}, \qquad i\in \mathbb{S},$$

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 7.22 is positive recurrent for $\alpha < 1$ only, whereas no stationary distribution exists when $\alpha \ge 1$.

As a consequence of Corollary 6.13 we have the following corollary of Theorem 7.10, which does not require aperiodicity for the stationary distribution to exist.

Corollary 7.11

Let $(X_n)_{n \in \mathbb{N}}$ be an irreducible Markov chain with *finite* state space S. Then the probabilities

$$\pi_k=rac{1}{\mu_k(k)},\qquad k\in \mathbb{S},$$

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 $k \in S$ by Corollaries 6.13 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 $d \ge 2$ 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.

(5278)

a) irreducible + recurrent + aperiodic \implies existence of a limiting distribution,

by Theorem 7.8 we get

b) irreducible + positive recurrent + aperiodic \implies existence of a limiting distribution which is also stationary,

and by Theorem 7.10 we get

c) irreducible + positive recurrent \implies existence of a stationary distribution.

In addition, the limiting or stationary distribution $\pi = (\pi_i)_{i \in S}$ satisfies

$$\pi_i = rac{1}{\mu_i(i)}, \qquad i \in \mathbb{S},$$

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 (i) converges almost surely to $1/\mu_i(i)$, *i.e.*

$$\lim_{n o\infty}rac{1}{n}\sum\limits_{k=1}^n\mathbb{1}_{\{X_k=i\}}=rac{1}{\mu_i(i)},\qquad i\in\mathbb{S}$$

In case $(X_n)_{n \in \mathbb{N}}$ is also positive recurrent, Theorem 7.10 shows that we also have

$$\lim_{n o\infty}rac{1}{n}\sum\limits_{k=1}^n\mathbb{1}_{\{X_k=i\}}=\pi_i,\qquad i\in\mathbb{S},$$

where $(\pi_i)_{i \in S}$ is the stationary distribution of $(X_n)_{n \in \mathbb{N}}$. We refer to Figure 4.5 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

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

٠

MH3512 AY19-20

♦ 280

The equation $\pi = \pi P$ yields

$$\begin{cases} \pi_{1} = \frac{1}{2}\pi_{2} \\ \pi_{2} = \pi_{1} + \frac{1}{2}\pi_{3} \\ \pi_{3} = \frac{1}{2}\pi_{2} + \frac{1}{2}\pi_{6} \\ \pi_{4} = \frac{1}{2}\pi_{7} \\ \pi_{5} = \frac{1}{2}\pi_{6} + \frac{1}{3}\pi_{8} \\ \pi_{6} = \frac{1}{2}\pi_{3} + \frac{1}{2}\pi_{5} \\ \pi_{7} = \pi_{4} + \frac{1}{3}\pi_{8} \\ \pi_{8} = \frac{1}{2}\pi_{5} + \frac{1}{2}\pi_{7} + \pi_{9} \\ \pi_{9} = \frac{1}{3}\pi_{8}, \end{cases}$$

$$\begin{array}{l} \pi_{1} = \frac{1}{2}\pi_{2} \\ \pi_{2} = \pi_{3} \\ \pi_{2} = \pi_{3} \\ \pi_{3} = \pi_{6} \\ \pi_{4} = \frac{1}{2}\pi_{3} \\ \pi_{4} = \frac{1}{2}\pi_{3} \\ \frac{1}{2}\pi_{7} = \frac{1}{3}\pi_{8} \\ \pi_{6} = \pi_{5} \\ \pi_{6} = \pi_{7} \\ \pi_{9} = \frac{1}{3}\pi_{8}, \end{cases}$$

and

$$1 = \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6 + \pi_7 + \pi_8 + \pi_9$$

= $\pi_1 + 2\pi_1 + 2\pi_1 + \pi_1 + 2\pi_1 + 2\pi_1 + 2\pi_1 + 3\pi_1 + \pi_1$
= $16\pi_1$,

hence

$$\pi_1 = rac{1}{16}, \ \pi_2 = rac{2}{16}, \ \pi_3 = rac{2}{16}, \ \pi_4 = rac{1}{16}, \ \pi_5 = rac{2}{16}, \ \pi_6 = rac{2}{16}, \ \pi_7 = rac{2}{16}, \ \pi_8 = rac{3}{16}, \ \pi_9 = rac{1}{16},$$

cf. Figures 7.3 and 7.2 below, and we check that since $\mu_1(1) = 16$ by (5.3.8), we indeed have

$$\pi_1 = rac{1}{\mu_1(1)} = rac{1}{16},$$

MH3512 AY19-20

according to Corollary 7.11.



Figure 7.2: Stationary distribution on the maze (5.3.7).



Figure 7.3: Stationary distribution by state numbering.

The stationary probability distribution of Figures 7.3 and (7.2) can be compared to the proportions of time spent at each state simulated in Figure 7.4 using this Markov chain experiment.

0.191			1			ė	
Time	х			Х		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	-					

Figure 7.4: Simulated stationary distribution.

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:

= [0,1,0,0,0,0,0,0,0;
2,0,1/2,0,0,0,0,0;
1/2,0,0,0,1/2,0,0,0;
0,0,0,0,1,0,0;
0,0,0,0,1/2,0,1/2,0;
0,1/2,0,1/2,0,0,0,0;
0,0,1/2,0,0,0,1/2,0;
0,0,0,1/3,0,1/3,0,1/3;
0,0,0,0,0,1,0]
power(P,1000)
power(P,1001)

♦ 283

we see that

 $\circ 284$

(7.2.5)

(7.2.6)

$$\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 \\ 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 \\ 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 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 3/8 & 0 \\ 1/8 & 0 & 1/4 & 0 & 3/8 &$$

and

which shows that, although $(P^n)_{n \ge 1}$ 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 $\pi = (\pi_i)_{i \in S}$ via a Markov chain that admits π as limiting and stationary distribution. It applies in particular in the setting of huge state spaces S, cf. *e.g.* Problem 7.27-(a).

A Markov chain $(X_k)_{k \in \mathbb{N}}$ with transition matrix P on a state space 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 Figure 7.5.

Note that the detailed balance condition (7.3.1) implies the global balance condition (7.2.2) as, by summation over $i \in S$ in (7.3.1) we have

$$\sum\limits_{i\in\mathbb{S}}\pi_iP_{i,j}=\sum\limits_{i\in\mathbb{S}}\pi_jP_{j,i}=\pi_j\sum\limits_{i\in\mathbb{S}}P_{j,i}=\pi_j, \hspace{1em} j\in\mathbb{S},$$

which shows that $\pi P = \pi$, *i.e.* π is a stationary distribution for P, cf. *e.g.* Problem 7.24-(c).



Figure 7.5: Detailed balance condition (discrete time).

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

$$ilde{P}_{i,j} := P_{i,j} imes \min\left(1, rac{\pi_j P_{j,i}}{\pi_i P_{i,j}}
ight) = \left\{egin{array}{cc} P_{j,i} rac{\pi_j}{\pi_i} & ext{if} & \pi_j P_{j,i} < \pi_i P_{i,j}, \ P_{i,j} & ext{if} & \pi_j P_{j,i} \geqslant \pi_i P_{i,j}, \end{array}
ight.$$

for $i \neq j$, and

$$ilde{P}_{i,i} = 1 - \sum_{k
eq i \in \mathbb{S}} ilde{P}_{i,k} = P_{i,i} + \sum_{i
eq k \in \mathbb{S}} P_{i,k} \left(1 - \min\left(1, rac{\pi_j P_{j,i}}{\pi_i P_{i,j}}
ight)
ight), \quad i \in \mathbb{S}.$$

MH3512 AY19-20

♦ 285

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

$$\pi_i ilde{P}_{i,j} = \left\{ egin{array}{ccc} P_{j,i} \pi_j = \pi_j ilde{P}_{j,i} & ext{if} & \pi_j P_{j,i} < \pi_i P_{i,j}, \ \pi_i P_{i,j} = \pi_j ilde{P}_{j,i} & ext{if} & \pi_j P_{j,i} \geqslant \pi_i P_{i,j}, \end{array}
ight\} = \pi_j ilde{P}_{j,i},$$

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.

Property	Definition
absorbing (state)	$P_{i,i}=1$
recurrent (state)	$\mathbb{P}(T_i^r < \infty \mid X_0 = i) = 1$
transient (state)	$\mathbb{P}(T^r_i < \infty \mid X_0 = i) < 1$
positive recurrent (state)	recurrent and $\mathbb{E}[T^r_i \mid X_0 = i] < \infty$
null recurrent (state)	recurrent and $\mathbb{E}[T^r_i \mid X_0 = i] = \infty$
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 $n \ge 1$
stationary distribution π	obtained from solving $\pi = \pi P$

Table 7.1: Summary of Markov chain properties.

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

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 1/N^2 & 2(N-1)/N^2 & (N-1)^2/N^2 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 2^2/N^2 & 4(N-2)/N^2 & (N-2)^2/N^2 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 3^2/N^2 & 0 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & (N-2)^2/N^2 & 4(N-2)/N^2 & 2^2/N^2 & 0 \\ 0 & 0 & \cdots & 0 & (N-2)^2/N^2 & 4(N-2)/N^2 & 2^2/N^2 & 0 \\ 0 & 0 & \cdots & 0 & 0 & (N-1)^2/N^2 & 2(N-1)/N^2 & 1/N^2 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

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, \dots, N-1. \end{tabular}

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.

♦ 290



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 \boldsymbol{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).*
- 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 *n*-th signal being analysed by the processor. Show that the process $(X_n)_{n \ge 1}$ can be modeled as a Markov chain and determine its transition matrix.
- b) Compute the stationary distribution of $(X_n)_{n \ge 1}$.
- 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?

 $^{^{*}\}mathrm{You}$ may use the symmetry of the problem to simplify the calculations.

Exercise 7.7 Consider a Markov chain $(X_n)_{n \ge 0}$ 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 \bigcirc in the long run.
- d) On the average, how long does it take to reach state 0 after starting from state 2?

Exercise 7.8 Consider the transition probability matrix

$$P = \left[egin{array}{ccc} P_{i,j} \end{array}
ight]_{0 \leqslant i,j \leqslant 3} = egin{array}{cccc} 0 & 1 & 0 & 0 \ 0.1 & 0.4 & 0.2 & 0.3 \ 0.2 & 0.2 & 0.5 & 0.1 \ 0.3 & 0.3 & 0.4 & 0 \end{array}
ight|.$$

- 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. MH3512 AY19-20 Exercise 7.9 Consider the Markov chain with transition probability matrix

$$\left[\left[P_{i,j}
ight]_{0\leqslant i,j\leqslant 3} = \left[egin{array}{cccc} 0 & 1/2 & 0 & 1/2 \ 1/4 & 0 & 3/4 & 0 \ 0 & 1/3 & 0 & 2/3 \ 1/2 & 0 & 1/2 & 0 \end{array}
ight]$$

a) Show that the chain is periodic^{*} 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

$$\mathbb{P}(T=1)=0.1, \hspace{1em} \mathbb{P}(T=2)=0.2, \hspace{1em} \mathbb{P}(T=3)=0.3, \hspace{1em} \mathbb{P}(T=4)=0.4,$$

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

$$P = egin{bmatrix} 0.6 & 0.4 & 0 & 0 & 0 \ 0.3 & 0.7 & 0 & 0 & 0 \ 0.2 & 0 & 0.4 & 0 & 0.4 \ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Find $\lim_{n\to\infty} \mathbb{P}(X_n = A \mid X_0 = C)$.

^{*}A chain is periodic when all states have the same period

Exercise 7.12 Consider a Markov chain $(X_n)_{n \ge 0}$ on the state space $\{0, 1, 2, 3, 4\}$ with transition probability matrix P given by 293

$$P = \begin{bmatrix} 1/3 & 2/3 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1/7 & 6/7 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

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 *n*-th vehicle passing under the bridge, $n \ge 1$.

a) Show that the process $(X_n)_{n \ge 1}$ can be modeled as a Markov chain and write down its transition matrix.

- b) Compute the stationary distribution of $(X_n)_{n \ge 1}$.
- 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 $\mathbb{S} = \{1, 2, \dots, N\}$, 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 ([Gus14]). (Problem 5.25-(d) continued).

a) Is the chain $(Y_k)_{k \in \mathbb{N}}$ reducible? Find its communicating classes.

b) Find the limiting distribution, and the possible stationary distributions of the chain $(Y_k)_{k \in \mathbb{N}}$.

Exercise 7.16 Consider the Markov chain with transition matrix

\boldsymbol{q}	p	0	0	0	
${oldsymbol{q}}$	0	\boldsymbol{p}	0	0	
${oldsymbol{q}}$	0	0	\boldsymbol{p}	0	,
${oldsymbol{q}}$	0	0	0	\boldsymbol{p}	
1	0	0	0	0	
				_	

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 7.17 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 $(X_n)_{n \ge 1}$ on $\{A, B, C, D\}$ and give its transition matrix.

b) Compute the stationary distribution $[\pi_A, \pi_B, \pi_C, \pi_D]$ of $(X_n)_{n \ge 1}$.

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 7.18 Consider the Markov chain with transition probability matrix

 $\left[egin{array}{cccc} 1 & 0 & 0 \ 0 & 1 & 0 \ a & b & c \end{array}
ight],$

with a + b + c = 1.

a) Compute the power P^n for all $n \ge 2$.

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 7.19 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

$$\mathbb{P}(N=k)=eta(1-eta)^{k-1}, \qquad k\geqslant 1,$$

MH3512 AY19-20

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 7.20 Let $(X_n)_{n \in \mathbb{N}}$ be an irreducible aperiodic Markov chain on the finite state space $\mathbb{S} = \{1, 2, \dots, N\}$.

- a) Show that there *exists* a state $i \in \{1, 2, ..., N\}$ 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 $i \in \{1, 2, ..., N\}$ such that the mean return time $\mu_i(i)$ from state i to itself is higher or equal to N, *i.e.* $\mu_i(i) \ge N$.

Exercise 7.21 Consider a Markov chain on the state space $\{1, 2, ..., N\}$. For any $i \in \{2, ..., N-1\}$, the chain has probability $p \in (0, 1)$ to switch from state (i) to state (i+1), and probability q = 1 - p to switch from (i) to (i-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 (N-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 7.22 (Problem 6.9 continued). Let $\alpha > 0$ and consider the Markov chain with state space \mathbb{N} and transition matrix given by

$$P_{i,i-1}=rac{1}{lpha+1}, \hspace{1em} P_{i,i+1}=rac{lpha}{lpha+1}, \hspace{1em} i \geqslant 1.$$

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

$$\pi_k = lpha^{k-1}(1-lpha^2)/2, \qquad k \geqslant 1,$$

where the value of π_0 has to be determined.

- b) Does the chain admit a stationary distribution when $\alpha \ge 1$?
- c) Show that the chain is positive recurrent when $\alpha < 1$.

Exercise 7.23 Consider two discrete-time stochastic processes $(X_n)_{n \in \mathbb{N}}$ and $(Y_n)_{n \in \mathbb{N}}$ on a state space S, such that

 $X_n=Y_n, \qquad n\geqslant au,$

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

 $\mathbb{P}(X_n=x)\leqslant \mathbb{P}(Y_n=x)+\mathbb{P}(au>n) \qquad x\in \mathbb{S}, \quad n\in \mathbb{N}.$

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

$$\sup_{x\in\mathbb{S}}\left|\mathbb{P}(X_n=x)-\mathbb{P}(Y_n=x)
ight|\leqslant\mathbb{P}(au>n),\qquad n\in\mathbb{N}.$$

Problem 7.24 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 $N \ge 1$ and $(X_k)_{k=0,1,\ldots,N}$ a Markov chain with transition matrix P on a state space S, we let

$$Y_k := X_{N-k}, \qquad k = 0, 1, \dots, N,$$

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.*

$$\mathbb{P}(X_k=i)=\pi_i, \qquad i\in S, \quad k=0,1,\ldots,N.$$
 MH3512 AY19-2

Show that the process $(Y_k)_{k=0,1,...,N}$ is a (forward) Markov chain (*i.e.* $(X_k)_{k=0,1,...,N}$ 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 \boldsymbol{P} and $\boldsymbol{\pi}$.

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 $(Y_n)_{n=0,1,\dots,N}$ has the Markov property

$$\mathbb{P}(Y_{n+1} = j \mid Y_n = i_n, \dots, Y_0 = i_0) = \mathbb{P}(Y_{n+1} = j \mid Y_n = i_n).$$

b) We say that $(X_k)_{k=0,1,\dots,N}$ is reversible for π when $(X_k)_{k=0,1,\dots,N}$ and $(Y_k)_{k=0,1,\dots,N}$ 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 <u>reversible</u> 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

$$P_{k_1,k_2}P_{k_2,k_3}\cdots P_{k_{n-1},k_n}P_{k_n,k_1} = P_{k_1,k_n}P_{k_n,k_{n-1}}\cdots P_{k_3,k_2}P_{k_2,k_1}$$
MH3512(7,3,2)

for all sequences $\{k_1, k_2, \ldots, k_n\}$ of states and $n \ge 2$.

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 $\{k_1, k_2, \ldots$ of states, $n \ge 2$, 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 $S = \{0, 1, \dots, M\}$ and that P is the transition matrix

$$P_{i,i+1}=rac{1}{2}-rac{i}{2M}, \qquad P_{i,i}=rac{1}{2}, \qquad P_{i,i-1}=rac{i}{2M}, \qquad 1\leqslant i\leqslant M-1,$$

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

$$\lim_{n o\infty} \mathbb{E}[X_n \mid X_0=i] = rac{M}{2},$$

for all i = 0, 1, ..., M.

MH3512 AY19-20

(5299)

j) Show, using first step analysis and induction on $n \ge 0$, that we have

$$\mathbb{E}\left[X_n-rac{M}{2} \ \Big| \ X_0=i
ight] = \left(i-rac{M}{2}
ight) \left(1-rac{1}{M}
ight)^n, \qquad n \geqslant 0,$$

for all $i = 0, 1, \dots, M$, and that this relation can be used to recover the result of Question (i).

Hint: Letting

$$h_n(i) = \mathbb{E}\left[X_n - rac{M}{2} \ \Big| \ X_0 = i
ight], \qquad n \geqslant 0,$$

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.

Problem 7.25 PageRankTM algorithm ([SvZ09]). A meta search engine tries to provide a single optimized ranking of search results $\{a, b, c, d, e\}$ based on the outputs of 4 different search engines denoted S_1 , S_2 , S_3 , S_4 , a technique known as *rank aggregation*.

Five possible search results *a*, *b*, *c*, *d*, *e* have been respectively ranked as

rank	S_1	$ S_2 $	S_3	$egin{array}{c} S_4 \end{array}$
1	b	c	d	e
2	с	b	e	a
3	d	d	a	d
4	a	e	b	b
5	е	a	с	с

by the search engines S_1 , S_2 , S_3 , S_4 .

a) A state $y \in \{a, b, c, d, e\}$ is said to be better ranked than another state $x \in \{a, b, c, d, e\}$, and we write $x \leq y$, if y ranks higher than x for at least three of the four search engines.

Complete the ranking table

ΥI	a	b	С	d	е
a	Х				
b		Х			
с			Х		
d				Х	
е					Х

with " $x \preceq y$ " or " $x \not \preceq y$ " at position (x, y). The diagonal entries are not relevant.

b) The meta search engine works by constructing a self-improving random sequence $(X_n)_{n \in \mathbb{N}}$ which is supposed to "converge" to the best possible search result based on the data of the four rankings.

Given a search result $X_n = x$ we choose the next search result X_{n+1} by assigning probability 1/5 to each of the search results that are *better ranked* than x. If no search result is better than x, then we keep $X_{n+1} = x$.

What is the state space of the sequence $(X_n)_{n \in \mathbb{N}}$?

Model the process $(X_n)_{n \in \mathbb{N}}$ as a Markov chain and complete the transition matrix

	3/5			
	0			
P =		1/5		
				0
			1/5	

c) Draw the graph of the chain $(X_n)_{n \in \mathbb{N}}$.

Is the chain $(X_n)_{n \in \mathbb{N}}$ reducible?

- d) Does the chain $(X_n)_{n \in \mathbb{N}}$ admit a limiting distribution independent of the initial state?
- e) Does the chain $(X_n)_{n \in \mathbb{N}}$ admit a stationary distribution? Find all stationary distribution(s) of the chain $(X_n)_{n \in \mathbb{N}}$.

f) In PageRankTM-type algorithms, one often chooses to perturb the transition matrix P into the new matrix

with n = 5 here, and $\varepsilon \in (0, 1)$. Show that \tilde{P} is a Markov transition matrix and that the corresponding chain $(\tilde{X}_n)_{n \ge 1}$ is irreducible and aperiodic.

g) Show that \tilde{P} admits a stationary distribution $\tilde{\pi}$ that satisfies

$$ilde{\pi} = \left[rac{arepsilon}{5}, rac{arepsilon}{5}, rac{arepsilon}{5}, rac{arepsilon}{5}
ight] + (1-arepsilon) ilde{\pi} P,$$

and that all probabilities in $\tilde{\pi}$ are greater than $\varepsilon/5$.

- h) Compute the stationary distribution of \tilde{P} .
- i) Provide a ranking of the states $\{a, b, c, d, e\}$ based on the stationary distribution $\tilde{\pi}$.
- j) Compute the mean return times for \tilde{P} , and show that they are all below $5/\epsilon$.

Problem 7.26 ([LPW09]). Convergence to equilibrium. In this problem we derive quantitative bounds for the convergence of a Markov chain to its stationary distribution π . Let P be the transition matrix of a discrete-time Markov chain $(X_n)_{n \in \mathbb{N}}$ on $S = \{1, 2, \ldots, N\}$. Given two probability distributions $\mu = [\mu_1, \mu_2, \ldots, \mu_N]$ and $\nu = [\nu_1, \mu_2, \ldots, \nu_N]$ on $\{1, 2, \ldots, N\}$, the total variation distance between μ and ν is defined as

$$\|\mu -
u\|_{\mathrm{TV}} := rac{1}{2} \sum_{k=1}^{N} |\mu_k -
u_k|.$$
 MH3512 AY19-20

Recall that the vector $\mu P^n = ([\mu P^n]_i)_{i=1,2,...,N}$ denotes the probability distribution of the chain at time $n \in \mathbb{N}$, given it was started with the initial distribution $\mu = [\mu_1, \mu_2, \ldots, \mu_N]$, *i.e.* we have, using matrix product notation,

$$\mathbb{P}(X_n=i) = \sum\limits_{j=1}^N \mathbb{P}(X_n=i \mid X_0=j) \mathbb{P}(X_0=j) = \sum\limits_{j=1}^N \mu_j [P^n]_{j,i} = [\mu P^n]_i,$$

 $i=1,2,\ldots,N$.

- a) Show that for any two probability distributions $\mu = [\mu_1, \mu_2, \dots, \mu_N]$ and $\nu = [\nu_1, \nu_2, \dots, \nu_N]$ on $\{1, 2, \dots, N\}$ we always have $\|\mu \nu\|_{\text{TV}} \leq 1$.
- b) Show that for any two probability distributions $\mu = [\mu_1, \mu_2, \dots, \mu_N]$ and $\nu = [\nu_1, \nu_2, \dots, \nu_N]$ on $\{1, 2, \dots, N\}$ and any Markov transition matrix P we have

$$\|\mu P -
u P\|_{ ext{TV}} \leqslant \|\mu -
u\|_{ ext{TV}}.$$

Hint: Use the triangle inequality

$$\left|\sum\limits_{k=1}^n x_k
ight|\leqslant \sum\limits_{k=1}^n |x_k|, \qquad x_1,x_2,\ldots,x_n\in \mathbb{R}.$$

c) Assume that the chain with transition matrix P admits a stationary distribution $\pi = [\pi_1, \pi_2, \ldots, \pi_N]$. Show that for any probability distribution $\mu = [\mu_1, \mu_2, \ldots, \mu_N]$ we have

$$\|\mu P^{n+1}-\pi\|_{\mathrm{TV}}\leqslant \|\mu P^n-\pi\|_{\mathrm{TV}},\qquad n\in\mathbb{N}.$$

d) Show that the *distance from stationarity*, defined as

$$d(n):=\max_{k=1,2,...,N} \Vert [P^n]_{k,\cdot}-\pi \Vert_{\mathrm{TV}}, \qquad n\in\mathbb{N},$$

satisfies $d(n+1) \leq d(n), n \in \mathbb{N}$.

e) Assume that all entries of \boldsymbol{P} are strictly positive. Explain why the chain is aperiodic and irreducible, and why it admits a limiting and stationary distribution.

In the sequel we assume that P admits an invariant (or stationary) distribution $\pi = [\pi_1, \pi_2, \ldots, \pi_N]$ such that $\pi P = \pi$, and that

$$P_{i,j} \geqslant heta \pi_j, ext{ for all } i,j=1,2,\ldots,N, ext{ MH3512} (7.3.3)_0$$

for some $0 < \theta < 1$. We also let

$$\Pi := egin{bmatrix} \pi_1 & \pi_2 & \pi_3 & \pi_4 & \cdots & \pi_N \ \pi_1 & \pi_2 & \pi_3 & \pi_4 & \cdots & \pi_N \ \pi_1 & \pi_2 & \pi_3 & \pi_4 & \cdots & \pi_N \ dots & d$$

hence (7.3.3) reads $P \ge \theta \Pi$.

f) Show that for all $0 < \theta < 1$ the matrix

$$Q_ heta := rac{1}{1- heta}(P- heta\Pi)$$

is the transition matrix of a Markov chain on $S = \{1, 2, \dots, N\}$.

g) Show by induction on $n \in \mathbb{N}$ that we have

$$P^n-\Pi=(1- heta)^n(Q^n_ heta-\Pi),\qquad n\in\mathbb{N}.$$

h) Show that given any $X_0 = k = 1, 2, \ldots, N$ the total variation distance between the distribution

$$egin{array}{rcl} [P^n]_{k,\cdot} &=& ([P^n]_{k,1},\ldots,[P^n]_{k,N}) \ &=& [\mathbb{P}(X_n=1\mid X_0=k),\ldots,\mathbb{P}(X_n=N\mid X_0=k)] \end{array}$$

of the chain at time n and the stationary distribution $\pi = [\pi_1, \pi_2, \ldots, \pi_N]$ satisfies

$$\|[P^n]_{k,\cdot}-\pi\|_{\mathrm{TV}}\leqslant (1- heta)^n, \qquad n\geqslant 1, \quad k=1,2,\dots,N.$$

Conclude that we have $d(n) \leq (1-\theta)^n, n \geq 1$.

i) Show that the *mixing time* of the chain with transition matrix \boldsymbol{P} , defined as

$$t_{ ext{mix}}:=\min\{n \geqslant 0 \ : \ d(n) \leqslant 1/4\}, ext{MH3512 AY19-20}$$

satisfies

$$t_{ ext{mix}} \leqslant \left[rac{\log 1/4}{\log(1- heta)}
ight].$$

j) Find the optimal value of θ satisfying the condition $P_{i,j} \ge \theta \pi_j$ for all i, j = 1, 2, ..., N for the chain of Exercise 4.12 with N = 3.

Problem 7.27 ([AH07]). The voter model is a particular case of the Ising model. This model has applications to spatial statistics, image analysis and segmentation, opinion studies, urban segregation, language change, metal alloys, magnetic materials, liquid/gas coexistence, phase transitions, plasmas, cell membranes in biophysics, ...

In dimension one, the model is built on the state space $S := \{-1, +1\}^N$ made of elements $z = (z_k)_{1 \le k \le N} \in S$ whose components $z_k \in \{-1, 1\}, k = 1, 2, ..., N$, are called *spins*.

Figure 7.6: Simulation of the voter model with N = 199, p = 0.98, and $z_0 = z_{N+1} = +1$.*

We consider a Markov chain $(Z_n)_{n \in \mathbb{N}}$ on the state space $\mathbb{S} = \{-1, +1\}^N$, whose transitions from an initial configuration $Z_0 = z = (z_k)_{1 \leq k \leq N}$ to a new configuration $Z_1 = \tilde{z} = (\tilde{z}_k)_{1 \leq k \leq N}$ are defined as follows:

First, choose one component z_k in $z = (z_k)_{1 \le k \le N}$ with probability 1/N, k = 1, 2, ..., N, and then consider the following cases:

(i) if
$$(z_{k-1}, z_{k+1}) = (-1, +1)$$
 or $(z_{k-1}, z_{k+1}) = (+1, -1)$

 \Rightarrow flip the sign of z_k , *i.e.* set $\tilde{z}_k := \pm z_k$ with probability 1/2,

^{*}Animated figure (works in Acrobat Reader).

(ii) if $(z_{k-1}, z_{k+1}) = (+1, +1)$: \Rightarrow set $\tilde{z}_k := +1$ with probability p > 0, and $\tilde{z}_k := -1$ with probability q > 0.

(iii) if $(z_{k-1}, z_{k+1}) = (-1, -1)$:

 \Rightarrow set $\tilde{z}_k := -1$ with probability p > 0, and $\tilde{z}_k := +1$ with probability q > 0,

where p + q = 1. The probabilities p and q can be respectively viewed as the probabilities of "agreeing", resp. "disagreeing" with two neighbors sharing the same opinion. The boundary conditions z_0 and z_{N+1} are arbitrarily fixed, and the corresponding instructions can be coded in R as follows:

if (z[k-1]!=z[k+1]) z[k]=sample(c(-1,1), 1,prob=c(0.5,0.5))
if (z[k-1]==1 && z[k+1]==1) z[k]=sample(c(-1,1), 1, prob=c(q,p))
if (z[k-1]==-1 && z[k+1]==-1) z[k]=sample(c(-1,1), 1, prob=c(p,q))

- a) Find the cardinality of the state space S.
- b) Replace the question marks (?) below with the corresponding transition probabilities given that $Z_0 = z = (z_k)_{1 \le k \le N}$: (i) if $(z_{k-1}, z_k, z_{k+1}) = (-1, \pm 1, +1)$ or $(z_{k-1}, z_k, z_{k+1}) = (+1, \pm 1, -1)$, $\mathbb{P}(Z_1 = (z_1, \dots, z_{k-1}, -z_k, z_{k+1}, \dots, z_N) \mid Z_0 = z) := (?)$,

 $\begin{aligned} k &= 1, 2, \dots, N, \\ \text{(ii) if } (z_{k-1}, z_k, z_{k+1}) &= (+1, +1, +1) \text{ or } (z_{k-1}, z_k, z_{k+1}) = (-1, -1, -1), \\ \mathbb{P}(Z_1 &= (z_1, \dots, z_{k-1}, -z_k, z_{k+1}, \dots, z_N) \mid Z_0 = z) := (?), \end{aligned}$

 $\begin{aligned} k &= 1, 2, \dots, N, \\ \text{(iii) if } (z_{k-1}, z_k, z_{k+1}) &= (+, -1, +1) \text{ or } (z_{k-1}, z_k, z_{k+1}) = (-1, +1, -1), \\ \mathbb{P}(Z_1 &= (z_1, \dots, z_{k-1}, -z_k, z_{k+1}, \dots, z_N) \mid Z_0 = z) := (?), \\ k &= 1, 2, \dots, N. \end{aligned}$

(306

c) Show that, based on the three above cases (i)-(ii)-(iii), the transition probabilities of $(Z_n)_{n \in \mathbb{N}}$ take the general form 307

$$\mathbb{P}(Z_1 = \bar{z}^k \mid Z_0 = z) = \frac{1}{N(1 + (p/q)^{z_k(z_{k-1} + z_{k+1})/2})}, \qquad k = 1, 2, \dots, N,$$
(7.3.4)

where $\boldsymbol{z} = (z_1, \ldots, z_N)$ and

$$\bar{z}^k := (z_1, \dots, z_{k-1}, -z_k, z_{k+1}, \dots, z_N)$$
(7.3.5)

denotes the state $z \in S$ after flipping its kth component z_k .

- d) Compute $\mathbb{P}(Z_1 = z \mid Z_0 = z)$ for all $z \in S$ using the complement rule, Relation (7.3.4), and the law of total probability.
- e) Taking N = 3 and setting $\underline{z_0 = z_4 = -1}$, *i.e.* $(z_0, z_1, z_2, z_3, z_4)$ takes the form $(-, \pm, \pm, -)$, complete the missing entries following transition probability matrix P of $(Z_n)_{n \in \mathbb{N}}$ on the state space $\mathbb{S} = \{-, -, -, +, -, -, +, -, +, +, +, +, +, +, +, +\}$:

		+	-+-	-++	+	+-+	++-	+++
	(?)	(?)	q/3	0	q/3	0	0	0
+	(?)	1/2	0	(?)	0	q/3	0	0
-+-	(?)	0	(1+q)/3	(?)	0	0	1/6	0
-++	0	1/6	(?)	(?)	0	0	0	1/6
+	p/3	0	0	0	1/2	(?)	(?)	0
+-+	0	p/3	0	0	p/3	q	0	p/3
++-	0	0	1/6	0	(?)	0	1/2	(?)
+++	0	0	0	1/6	0	q/3	1/6	(?)

Is the chain $(Z_n)_{n \in \mathbb{N}}$ reducible? Why?

f) Does the chain $(Z_n)_{n \in \mathbb{N}}$ admit a limiting distribution? A stationary distribution? Why?

g) Show that if a probability distribution $(\pi_z)_{z\in\mathbb{S}}$ on \mathbb{S} satisfies the relation

$$\frac{\pi_{\bar{z}^k}}{\pi_z} = \frac{\mathbb{P}(Z_1 = \bar{z}^k \mid Z_0 = z)}{\mathbb{P}(Z_1 = z \mid Z_0 = \bar{z}^k)}, \qquad k = 1, 2, \dots, N, \quad z \in \mathbb{S},$$
(7.3.6)

where \bar{z}^k is defined in (7.3.5), then $(\pi_z)_{z\in\mathbb{S}}$ is a stationary distribution for the chain $(Z_n)_{n\in\mathbb{N}}$, *i.e.* we have

$$\Big(\mathbb{P}(Z_0=z)=\pi_z, \ orall \ z\in\mathbb{S}\Big) \quad \Longrightarrow \quad \Big(\mathbb{P}(Z_1=z)=\pi_z, \ orall \ z\in\mathbb{S}\Big).$$

Hint: Start from the law of total probability

$$egin{array}{rll} \mathbb{P}(Z_1=z) &=& \mathbb{P}(Z_1=z \mid Z_0=z) \mathbb{P}(Z_0=z) \ &+ \sum\limits_{k=1}^N \mathbb{P}(Z_1=z \mid Z_0=ar{z}^k) \mathbb{P}(Z_0=ar{z}^k), \end{array}$$

and show, using (7.3.6), that the above equals π_z if $\mathbb{P}(Z_0 = z) = \pi_z$ for all $z \in S$.

h) Show that

$$\frac{\mathbb{P}(Z_1 = \bar{z}^k \mid Z_0 = z)}{\mathbb{P}(Z_1 = z \mid Z_0 = \bar{z}^k)} = \left(\frac{q}{p}\right)^{z_k(z_{k-1} + z_{k+1})/2}, \qquad k = 1, 2, \dots, N, \quad z \in \mathbb{S}.$$
(7.3.7)

Hint: Use Equation (7.3.4) and the relations

$$rac{1+(q/p)}{1+(p/q)} = rac{q}{p} \quad ext{and} \quad rac{1+(p/q)}{1+(q/p)} = rac{p}{q}.$$

i) Show that the probability distribution $(\pi_z)_{z\in\mathbb{S}}$ defined as

$$\pi_z := C_\beta \exp\left(\beta \sum_{l=0}^N z_l z_{l+1}\right), \qquad z \in \mathbb{S},$$
(7.3.8)

is the stationary and limiting distribution of $(Z_n)_{n \in \mathbb{N}}$, where C_β is a normalization constant^{*} and β is to be given in terms of p and q.

♦ 308

^{*}The normalization constant C_{β} does not have to be computed here.

Hint: Using (7.3.7), show that $(\pi_z)_{z\in\mathbb{S}}$ defined in (7.3.8) satisfies (7.3.6).

j) Taking N = 3 and $\underline{z_0 = z_4 = +1}$, *i.e.* $(z_0, z_1, z_2, z_3, z_4)$ takes the form $(+, \pm, \pm, \pm, +)$, compute the limiting distribution of each of the 8 configurations in

$$\mathbb{S} = \{---, --+, -+-, --+, +--, +-+, ++-, +++\},\$$

and find the value of C_{β} .

Figure 7.7: Simulation of the voter model with N = 3, $p = \sqrt{0.75}$, and $z_0 = z_4 = +1$.*

♦ 309

^{*}Animated figure (works in Acrobat Reader).

8 Branching Processes

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	310
8.2 Probability Generating Functions	.
8.3 Extinction Probabilities	318
Exercises	

8.1 Construction and Examples

Consider a time-dependent population made of a number X_n of individuals at generation $n \ge 0$. 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 $k = 1, 2, ..., X_n$ 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

$$X_{n+1} = Y_1 + \dots + Y_{X_n} = \sum_{k=1}^{X_n} Y_k,$$
 (8.1.1)

where the $(Y_k)_{k \ge 1}$ form a sequence of independent, identically distributed, nonnegative integer valued random variables which are assumed to be almost surely finite, *i.e.*

$$\mathbb{P}(Y_k < \infty) = \sum\limits_{n \geqslant 0} \mathbb{P}(Y_k = n) = 1.$$
 MH3512 AY19-20

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

$$X_{n+1} = X_n + Y_1 + \dots + Y_{X_n},$$

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 $S = \mathbb{N}$ and transition matrix given by

$$P = [P_{i,j}]_{i,j\in\mathbb{N}} = \begin{vmatrix} 1 & 0 & 0 & \cdots \\ \mathbb{P}(Y_1 = 0) & \mathbb{P}(Y_1 = 1) & \mathbb{P}(Y_1 = 2) & \cdots \\ P_{2,0} & P_{2,1} & P_{2,2} & \cdots \\ P_{3,0} & P_{3,1} & P_{3,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix}.$$
(8.1.2)

Note that state $(\mathbf{0})$ is absorbing since by construction we always have

$$P_{0,0} = \mathbb{P}(X_{n+1} = 0 \mid X_n = 0) = 1, \qquad n \in \mathbb{N}.$$
 MH3512 AY19-20



Figure 8.1: Example of branching process.

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

$$(Y_k)_{k=1,2,...,X_1} = (Y_1,Y_2,Y_3,Y_4,Y_5,Y_6) = (0,4,1,2,2,3)$$

and

$$egin{array}{rcl} X_2 &=& Y_1+\dots+Y_{X_1} \ &=& Y_1+Y_2+Y_3+Y_4+Y_5+Y_6 \ &=& 0+4+1+2+2+3 \ &=& 12, \end{array}$$

then

$$(Y_k)_{k=1,2,...,X_2} \;=\; (Y_1,Y_2,Y_3,Y_4,Y_5,Y_6,Y_7,Y_8,Y_9,Y_{10},Y_{11},Y_{12})$$

MH3512 AY19-20

312

= (0, 2, 0, 0, 0, 4, 2, 0, 0, 2, 0, 1),

and

$$\begin{aligned} X_3 &= Y_1 + \dots + Y_{X_2} \\ &= Y_1 + Y_2 + Y_3 + Y_4 + Y_5 + Y_6 + Y_7 + Y_8 + Y_9 + Y_{10} + Y_{11} + Y_{12} \\ &= 0 + 2 + 0 + 0 + 0 + 4 + 2 + 0 + 0 + 2 + 0 + 1 \\ &= 11. \end{aligned}$$

The next Figure 8.2 presents another sample tree for the path of a branching process.



Figure 8.2: Sample graph of a branching process.

In Figure 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 polylogarithmics form 1_{19-20}

8.2 Probability Generating Functions

Let now $G_1(s)$ denote the probability generating function of $X_1 = Y_1$, defined as

$$G_1(s):=\mathbb{E}[s^{X_1}\mid X_0=1]=\mathbb{E}[s^{Y_1}]=\sum\limits_{k\geqslant 0}s^k\mathbb{P}(Y_1=k), \hspace{1em} -1\leqslant s\leqslant 1,$$

cf. (1.7.1), denote the probability generating function of the (almost surely finite) random variable $X_1 = Y_1$, with

$$\begin{cases} G_1(0) = \mathbb{P}(Y_1 = 0), \\ G_1(1) = \sum_{n \ge 0} \mathbb{P}(Y_1 = n) = \mathbb{P}(Y_1 < \infty) = 1, \\ \mu := G'_1(1^-) = \sum_{k \ge 0} k \mathbb{P}(Y_1 = k) = \mathbb{E}[X_1 \mid X_0 = 1] = \mathbb{E}[Y_1]. \end{cases}$$
(8.2.1a)

More generally, letting $G_n(s)$ denote the probability generating function of X_n , defined as

$$G_n(s):=\mathbb{E}[s^{X_n}\mid X_0=1]=\sum\limits_{k\geqslant 0}s^k\mathbb{P}(X_n=k\mid X_0=1), \hspace{1em} -1\leqslant s\leqslant 1,$$

 $n \in \mathbb{N}$, we have

$$\begin{cases} G_0(s) = s, & -1 \leq s \leq 1, \\ G_n(0) = \mathbb{P}(X_n = 0 \mid X_0 = 1), & n \in \mathbb{N}, \\ \mu_n := \mathbb{E}[X_n \mid X_0 = 1] = G'_n(1^-) = \sum_{k \geq 0} k \mathbb{P}(X_n = k \mid X_0 = 1), \end{cases}$$
(8.2.2b)

cf. (1.7.3). 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

$$X_n = \sum\limits_{l=1}^k X_n^{(l)}, \qquad n \in \mathbb{N},$$
 MH3512 AY19-24

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

$$egin{aligned} \mathbb{E}\left[s^{X_n} \mid X_0 = k
ight] &= \mathbb{E}\left[s^{X_n^{(1)} + \dots + X_n^{(k)}} \mid X_0^{(1)} = 1, \dots, X_0^{(k)} = 1
ight] \ &= \prod_{l=1}^k \mathbb{E}\left[s^{X_n^{(l)}} \mid X_0^{(l)} = 1
ight] \ &= \left(\mathbb{E}\left[s^{X_n} \mid X_0 = 1
ight]
ight)^k \ &= \left(G_n(s)
ight)^k, \quad -1 \leqslant s \leqslant 1, \quad n \in \mathbb{N}. \end{aligned}$$

The next proposition provides an algorithm for the computation of the probability generating function G_n .

Proposition 8.1

We have the recurrence relation

$$G_{n+1}(s) = G_n(G_1(s)) = G_1(G_n(s)), \quad -1 \le s \le 1, \qquad n \in \mathbb{N}.$$
 (8.2.3)

Proof

By the identity (1.6.11) on random products we have

$$egin{aligned} G_{n+1}(s) &= & \mathbb{E}[s^{X_{n+1}} \mid X_0 = 1] \ &= & \mathbb{E}[s^{Y_1 + \dots + Y_{X_n}} \mid X_0 = 1] \ &= & \mathbb{E}\left[\prod_{l=1}^{X_n} s^{Y_l} \mid X_0 = 1
ight] \ &= & \sum_{k \geqslant 0} \mathbb{E}\left[\prod_{l=1}^{X_n} s^{Y_l} \mid X_n = k
ight] \mathbb{P}(X_n = k \mid X_0 = 1) \ &= & \sum_{k \geqslant 0} \mathbb{E}\left[\prod_{l=1}^k s^{Y_l} \mid X_n = k
ight] \mathbb{P}(X_n = k \mid X_0 = 1) \ &= & \sum_{k \geqslant 0} \mathbb{E}\left[\prod_{l=1}^k s^{Y_l}\right] \mathbb{P}(X_n = k \mid X_0 = 1) \end{aligned}$$

$$egin{aligned} &=& \sum\limits_{k \geqslant 0} \left(\prod\limits_{l=1}^k \mathbb{E}\left[s^{Y_l}
ight]
ight) \mathbb{P}(X_n = k \mid X_0 = 1) \ &=& \sum\limits_{k \geqslant 0} (\mathbb{E}[s^{Y_1}])^k \mathbb{P}(X_n = k \mid X_0 = 1) \ &=& G_n(\mathbb{E}[s^{Y_1}]) \ &=& G_n(G_1(s)), \qquad -1 \leqslant s \leqslant 1. \end{aligned}$$

Instead of (8.2.3) we may also write

$$G_n(s) = G_1(G_1(\cdots(G_1(s),\cdots))), \quad -1 \leqslant s \leqslant 1, \quad (8.2.4)$$

and

$$G_n(s)=G_1(G_{n-1}(s))=G_{n-1}(G_1(s)), \qquad -1\leqslant s\leqslant 1.$$

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

$$\mu_n = \mathbb{E}[X_n \mid X_0 = 1] = (\mathbb{E}[X_1 \mid X_0 = 1])^n = \mu^n, \qquad n \ge 1,$$
(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

$$\mu_n~=~G_n'(1^-$$

$$= \frac{d}{ds} G_1(G_{n-1}(s))_{|s=1}$$

= $G'_{n-1}(1^-)G'_1(G_{n-1}(1^-))$
= $G'_{n-1}(1^-)G'_1(1^-)$
= $\mu \times \mu_{n-1},$

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

$$\mathbb{E}[X_n \mid X_0 = k] = k \mathbb{E}[X_n \mid X_0 = 1] = k \mu^n, \qquad n \geqslant 1,$$

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

$$egin{array}{rcl} \mu &=& \mathbb{E}[Y_1] = \sum\limits_{n \geqslant 1} n \mathbb{P}(Y_1 = n) \ &\geqslant & \mathbb{P}(Y_1 = 1) + 2 \sum\limits_{n \geqslant 2} \mathbb{P}(Y_1 = n) \ &=& \mathbb{P}(Y_1 = 1) + \mathbb{P}(Y_1 \geqslant 2) \ &> & \mathbb{P}(Y_1 \geqslant 1) \ &=& 1, \end{array}$$

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

$$egin{array}{rcl} 1 &>& \mu = \mathbb{E}[Y_1] = \sum\limits_{n \geqslant 1} n \mathbb{P}(Y_1 = n) \ &\geqslant& \sum\limits_{n \geqslant 1} \mathbb{P}(Y_1 = n) \ &=& \mathbb{P}(Y_1 \geqslant 1) \ &=& 1 - \mathbb{P}(Y_0 = 0), \end{array}$$

although the converse is not true in general.

The variance $\sigma_n^2 = \operatorname{Var}[X_n \mid X_0 = 1]$ 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 (1.7.5), where $\sigma^2 = \operatorname{Var}[Y_1]$, which shows by induction that

$$\sigma_n^2 = \operatorname{Var}[X_n \mid X_0 = 1] = \left\{egin{array}{ll} n\sigma^2, & \mu = 1, \ \sigma^2\mu^{n-1}rac{1-\mu^n}{1-\mu} = \sigma^2\sum\limits_{k=0}^{n-1}\mu^{n+k-1}, & \mu
eq 1, \end{array}
ight.$$

 $n \ge 1$ cf. e.g. pages 180-181 of [KT81], and Exercise 8.3-(a) for an application. We also have

$$\operatorname{Var}[X_n \mid X_0 = k] = k \operatorname{Var}[X_n \mid X_0 = 1] = k \sigma_n^2, \qquad k, n \in \mathbb{N},$$

due to Relation (1.6.10) for the variance of a sum of independent random variables.

8.3 Extinction Probabilities

Here we are interested in the time to extinction^{*}

$$T_0:=\inf\{n\geqslant 0\;:\;X_n=0\},$$

^{*}We normally start from $X_0 \ge 1$.

and in the extinction probability

$$lpha_k := \mathbb{P}(T_0 < \infty \mid X_0 = k)$$

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

$$\mathbb{P}(T_0=n\mid X_0=1)=G_n(0)-G_{n-1}(0)=G_1(G_{n-1}(0))-G_{n-1}(0), \ \ n\geqslant 1,$$

with $\mathbb{P}(T_0 = 0 \mid X_0 = 1) = 0$.

\mathbf{Proof}

By the relation $\{X_{n-1} = 0\} \subset \{X_n = 0\}$, we have

$$\left\{T_0=n\right\}=\left\{X_n=0
ight\}\cap\left\{X_{n-1}\geqslant1
ight\}=\left\{X_n=0
ight\}\setminus\left\{X_{n-1}=0
ight\}$$

and

$$egin{aligned} \mathbb{P}(T_0 = n \mid X_0 = 1) &= & \mathbb{P}(\{X_n = 0\} \cap \{X_{n-1} \geqslant 1\} \mid X_0 = 1) \ &= & \mathbb{P}(\{X_n = 0\} \setminus \{X_{n-1} = 0\} \mid X_0 = 1) \ &= & \mathbb{P}(\{X_n = 0\}) - \mathbb{P}(\{X_{n-1} = 0\} \mid X_0 = 1) \ &= & G_n(0) - G_{n-1}(0) \ &= & G_1(G_{n-1}(0)) - G_{n-1}(0), \quad n \geqslant 1, \end{aligned}$$

where we applied Proposition 8.1.

First, we note that by the independence assumption, starting from $X_0 = k \ge 2$ independent individuals, we have

$$\alpha_k = \mathbb{P}(T_0 < \infty \mid X_0 = k) = (\mathbb{P}(T_0 < \infty \mid X_0 = 1))^k = (\alpha_1)^k, \quad k \ge 1.$$
(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 is the

product of the extinction probabilities for each branch, which yields $\alpha_k = (\alpha_1)^k$ since these extinction probabilities are all equal $\overleftrightarrow{\alpha_1}^{320}$ and there are k of them.

The next proposition is a consequence of Lemma 5.2.



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

 $\alpha = G_1(\alpha).$

\mathbf{Proof}

By first step analysis we have

$$egin{array}{rcl} lpha_1 &=& \mathbb{P}(T_0 < \infty \mid X_0 = 1) \ &=& \mathbb{P}(X_1 = 0 \mid X_0 = 1) + \sum\limits_{k \geqslant 1} \mathbb{P}(T_0 < \infty \mid X_1 = k) \mathbb{P}(X_1 = k \mid X_0 = 1) \end{array}$$

$$= \mathbb{P}(Y_1 = 0) + \sum_{k \ge 1} \mathbb{P}(T_0 < \infty \mid X_0 = k) \mathbb{P}(Y_1 = k)$$

$$= \sum_{k \ge 0} (\alpha_1)^k \mathbb{P}(Y_1 = k)$$

$$= G_1(\alpha_1),$$

hence the extinction probability α_1 solves (8.3.2).

Note that from the above proof we find

$$\alpha_1 \ge \mathbb{P}(X_1 = 0 \mid X_0 = 1) = \mathbb{P}(Y_1 = 0),$$
(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

$$lpha=G_1(G_1(lpha)), \quad lpha=G_1(G_1(G_1(lpha))),$$

and more generally

$$\alpha = G_n(\alpha), \qquad n \ge 1, \tag{8.3.4}$$

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

 $lpha_1 := \mathbb{P}(T_0 < \infty \mid X_0 = 1)$

is the smallest solution of the equation $\alpha = G_1(\alpha)$.

Proof

By Lemma 8.4 we have $\alpha_1 = \lim_{n \to \infty} G_n(0)$. Next, we note that the function $s \mapsto G_1(s)$ is increasing because

$$G_1'(s) = \mathbb{E}[Y_1 s^{Y_1 - 1}] > 0, \qquad s \in [0, 1).$$

Hence $s \mapsto G_n(s)$ is also increasing by Proposition 8.1, and for any solution $\alpha \ge 0$ of (8.3.2) we have, by (8.3.4),

$$0\leqslant G_n(0)\leqslant G_n(lpha)=lpha,\qquad n\geqslant 1,$$

and taking the limit in this inequality as n goes to infinity we get

$$0\leqslant lpha_1=\lim_{n
ightarrow\infty}G_n(0)\leqslant lpha,$$

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

$$egin{aligned} lpha &=& \lim_{n o \infty} G_n(lpha) \ &=& \lim_{n o \infty} \left(G_n(0) + \sum\limits_{k \geqslant 1} lpha^k \mathbb{P}(X_n = k \mid X_0 = 1)
ight) \ &\geqslant& \lim_{n o \infty} G_n(0) \ &=& lpha_1. \end{aligned}$$

Therefore α_1 is the smallest solution of (8.3.2).

Since $G_1(0) = \mathbb{P}(Y_1 = 0)$ we have

$$\mathbb{P}(Y_1=0)=G_1(0)\leqslant G_1(lpha_1)=lpha_1,$$

which recovers (8.3.3).

On the other hand, if $\mathbb{P}(Y_1 \ge 1) = 1$ then we have $G_1(0) = 0$, which implies $\alpha_1 = 0$ by Proposition 8.6.

Note that from Lemma 5.2, Proposition 8.4, and (8.3.1), the transition matrix (8.1.2) satisfies

Examples

i) Assume that Y_1 has a Bernoulli distribution with parameter $p \in (0, 1)$, *i.e.*

$$\mathbb{P}(Y_1=1)=p,\qquad \mathbb{P}(Y_1=0)=1-p.$$

In this case the branching process is actually a two-state Markov chain with transition matrix

$$P=\left[egin{array}{cc} 1 & 0 \ 1-p & p \end{array}
ight],$$

and we have

$$G_n(0) = \mathbb{P}(X_n = 0 \mid X_0 = 1) = (1 - p) \sum_{k=0}^{n-1} p^k = 1 - p^n,$$
(8.3.5)

where we used the geometric series (A.2), hence as in (5.1.5) the extinction probability α_1 is given by

$$egin{array}{rcl} lpha_1 &=& \mathbb{P}(T_0 < \infty \mid X_0 = 1) = \mathbb{P}igg(igcup_{n \geqslant 1} \{X_n = 0\} \ \Big| \ X_0 = 1 igg) \ &=& \lim_{n o \infty} \mathbb{P}(X_n = 0 \mid X_0 = 1) \ &=& \lim_{n o \infty} G_n(0) = 1, \end{array}$$

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

$$G_1(s) = \mathbb{E}[s^{Y_1}] = \sum_{k \ge 0} s^k \mathbb{P}(Y_1 = k) = 1 - p + ps,$$
(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 Figure $8_{MH3512 \text{ AY19-20}}$



Figure 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 $n \ge 1$ 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 (1.7.3) or (8.2.2a) as

$$\mathbb{P}(X_n=0 \mid X_0=1) = G_n(0) = 1-p^n.$$

We also have $\mathbb{E}[X_n] = p^n, n \ge 0.$

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

$$egin{array}{rll} G_1(s) &=& \mathbb{E}[s^{Y_1}] = \sum\limits_{k \geqslant 0} s^k \mathbb{P}(Y_1 = k) \ &=& s^0 \mathbb{P}(Y_1 = 0) + s^2 \mathbb{P}(Y_1 = 2) = 1 - p + p s^2. \end{array}$$

MH3512 AY19-20

ॳ 324
♦ 325

We check that the solutions of

$$G_1(lpha)=1-p+plpha^2=lpha,$$

i.e. *

$$p\alpha^{2} - \alpha + q = p(\alpha - 1)(\alpha - q/p) = 0,$$
 (8.3.7)

with q = 1 - p, are given by

$$\left\{\frac{1+\sqrt{1-4pq}}{2p}, \frac{1-\sqrt{1-4pq}}{2p}\right\} = \left\{1, \frac{q}{p}\right\}, \qquad p \in (0, 1].$$
(8.3.8)

Hence the extinction probability is $\alpha_1 = 1$ if $q \ge p$, 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.*

$$\mathbb{P}(Y_1=n)=(1-p)p^n,\qquad n\geqslant 0,$$

with $\mu = \mathbb{E}[Y_1] = p/q$. We have

$$G_1(s) = \mathbb{E}[s^{Y_1}] = \sum_{n \ge 0} s^n \mathbb{P}(Y_1 = n) = (1 - p) \sum_{n \ge 0} p^n s^n = \frac{1 - p}{1 - ps}.$$
(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

$$lpha_1 \;=\; \mathbb{P}(T_0 < \infty \mid X_0 = 1)$$

^{*}Remark that (8.3.7) is identical to the characteristic equation (2.2.15).

$$= \min\left(1, \frac{q}{p}\right) = \begin{cases} \frac{q}{p}, & p \ge 1/2, \text{ (super)critical case,} \\ 1, & p \le 1/2, \text{ (sub)critical case.} \end{cases}$$

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 Figures 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

$$\mathbb{P}(T_0 < \infty \mid X_0 = k) \; = \; \min\left(1, \left(rac{q}{p}
ight)^k
ight) \ = \; egin{cases} \left(rac{q}{p}
ight)^k, & p \geqslant 1/2, \; (ext{super}) ext{critical case}, \ 1, & p \leqslant 1/2, \; (ext{sub}) ext{critical case}. \end{cases}$$

which incidentally coincides with the finite time hitting probability found in (3.4.16) for the simple random walk started from $k \ge 1$.



Figure 8.4: Generating function of Y_1 with p = 3/8 < 1/2 and $\alpha_1 = 1$.

Next, in Figure 8.5 is a graph of the generating function $s \mapsto G_1(s)$ for p = 1/2.

MH3512 AY19-20 $\,$



Figure 8.5: Generating function of Y_1 with p = 1/2 and $\alpha_1 = 1$.

The graph of generating function in Figure 8.6 corresponds to p = 3/4.



Figure 8.6: Generating function of Y_1 with p = 3/4 > 1/2 and $\alpha_1 = q/p = 1/3$.

We also have $\mu_n = (\mathbb{E}[Y_1])^n = (p/q)^n, n \ge 1.$

MH3512 AY19-20

ॳ 327

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

$$\mathbb{P}(Y_1=n)=inom{n+r-1}{r-1}q^rp^n=(n+1)q^rp^n=(n+1)q^2p^n,\qquad n\geqslant 0,$$

with $\underline{r=2}$, cf. (1.5.12).

In this case we have*

$$egin{array}{rll} G_1(s)&=&\mathbb{E}[s^{Y_1}]=\sum\limits_{n\geqslant 0}s^n\mathbb{P}(Y_1=n)\ &=&q^2\sum\limits_{n\geqslant 0}(n+1)p^ns^n=\left(rac{1-p}{1-ps}
ight)^2,\qquad -1\leqslant s\leqslant 1, \end{array}$$

see *here*.[†] 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

$$lpha_1 = rac{1}{2}(3-\sqrt{5}) \simeq 0.382$$

which is the extinction probability, as illustrated in the next Figure 8.7. Here we have $\mathbb{E}[Y_1] = 2$.

*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. [†]We used the identity $\sum_{n \ge 0} (n+1)r^n = (1-r)^{-2}$, cf. (A.4).

MH3512 AY19-20



Figure 8.7: Probability generating function of Y_1 .



Figure 8.8: Simulation of influenza spread.*

The next graph in Figure 8.9 illustrates the extinction of a branching process in finite time when Y_1 has the geometric distribution with

329

^{*}Click on the figure to play the video (works in Acrobat Reader).

p = 1/2, in which case there is extinction within finite time with probability 1.

Figure 8.9: Sample path of a branching process $(X_n)_{n \ge 0}$.*

^{*}Animated figure (works in Acrobat Reader). Download the corresponding **R** code or the **IPython notebook** that can be run *here*.

In Table 8.1 we summarize some questions and their associated solution methods introduced in this chapter and the previous ones 331

How to compute	Method
the expected value $\mathbb{E}[X]$	sum the values of \boldsymbol{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 $\boldsymbol{g(k)}$	solve* $g = Pg$ for $g(k)$.
the mean hitting times $m{h}(m{k})$	$ ext{solve}^* h = \mathbb{1} + Ph ext{ for } h(k).$
the stationary distribution π	solve [†] $\pi = \pi P$ for π .
the extinction probability $oldsymbol{lpha_1}$	solve $G_1(\alpha) = \alpha$ for α and choose the smallest solution.
$\lim_{n o \infty} \left(\left[egin{array}{ccc} 1-a & a \ & b \ & 1-b \end{array} ight] ight)^n$	$\begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{bmatrix}$

Table 8.1: Summary of computing methods.

^{*}Be sure to write only the relevant rows of the system under the appropriate boundary conditions. [†]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 *n*th 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

$$G_n(s) = 1 - \frac{1}{2^n} + \frac{s}{2^n}, \quad s \in \mathbb{R}.$$
 (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 $\boldsymbol{\xi}$ 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 *n*th generation, with $X_0 = 1$.

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

$$\mathbb{P}(Y=0)=c, \quad \mathbb{P}(Y=1)=b, \quad \mathbb{P}(Y=2)=a,$$

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 $(Z_n)_{n \ge 0}$ 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 $\frac{1}{2}$ and $Z_0 = 1$.
- 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 $(X_n)_{n \ge 0}$ is a subcritical branching process, *i.e.* $\mu = \mathbb{E}[Y_1] < 1$, the time to extinction $T_0 := \inf\{n \ge 0 : X_n = 0\}$ 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 $n \geq k$ form an *i.i.d.* sequence with the negative binomial distribution

$$\mathbb{P}(Y_k=n)=(n+1)q^2p^n, \qquad n\geqslant 0, \quad k\geqslant 1,$$

where 0 < q = 1 - p < 1.

a) Compute the probability generating function

$$G_1(s):=\mathbb{E}\left[s^{Y_k}
ight]=\sum\limits_{n\geqslant 0}s^n\mathbb{P}(Y_1=n)$$

of $Y_k, k \ge 1$.

b) Compute the extinction probability $\alpha_1 := \mathbb{P}(T_0 < \infty \mid X_0 = 1)$ of the branching process $(X_n)_{n \ge 0}$ 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

$$\mathbb{P}(Z_1 = 0) = rac{1 - p - q}{1 - p}$$
 and $\mathbb{P}(Z_1 = k) = q p^{k - 1}, \quad k = 1, 2, 3, \dots,$

where $0 \leq p < 1$ and $0 \leq q \leq 1 - p$.

MH3512 AY19-20

- a) Find the probability generating function of Z_1 .
- b) Compute $\mathbb{E}[\mathbf{Z}_1]$.
- 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

$$G_{Z_k}(s) = rac{kp - (kp + p - 1)s}{1 - p + kp - kps}, \qquad -1 < s < 1,$$

for all $k \ge 1$.

Problem 8.11 Consider a branching process with *i.i.d.* offspring sequence $(Y_k)_{k \ge 1}$. 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 $Z = \sum_{k \ge 1} X_k$. 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

$$H(s)=\mathbb{E}\left[s^{Z}
ight], \qquad -1\leqslant s\leqslant 1,$$

denote the generating function of Z.

Show, by first step analysis, that the relation

$$H(s)=G_1(sH(s)), \qquad 0\leqslant s\leqslant 1,$$
 MH3512 AY19-20

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, ..., Z, where $(U_k)_{k \ge 1}$ is an *i.i.d.* sequence of random variables with finite expectation $\mathbb{E}[U]$ and distribution function $F(x) = \mathbb{P}(U \le x)$.

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 $F(x) = 1 e^{-x}$, $x \in \mathbb{R}_+$. Hints and comments on Problem 8.11.
- a) Use the expression (8.2.5) of $\mathbb{E}[X_k]$.
- b) $\mathbb{E}[Z] < \infty$ implies that $Z < \infty$ almost surely.
- c) Given that $X_1 = k$, Z can be decomposed into the sum of k independent population sizes, each of them started with 1 individual.
- d) Compute $G_1(s)$ and μ in this model, and check the condition $\mu < 1$. When computing H(s) you should have to solve a quadratic equation, and to choose the relevant solution out of two possibilities.
- e) The goal of this question is to confirm the result of Question (d) by checking the value of $H(0) = \lim_{s > 0} H(s)$. For this, find the Talyor expansion of $\sqrt{1 4pqs}$ as s tends to 0.
- f) The identity $1 4pq = (q p)^2$ can be useful, and the sign of q p has to be taken into account when computing $\sqrt{1 4pq}$.
- g) Use conditioning on $Z = n, n \in \mathbb{N}$.
- h) The answer will use both H and F.
- i) Compute the value of $\mathbb{E}[U_1]$.

(5337)

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
9.2	Continuous-Time Markov Chains
9.3	Transition Semigroup
9.4	Infinitesimal Generator
9.5	The Two-State Continuous-Time Markov Chain
9.6	Limiting and Stationary Distributions
9.7	The Discrete-Time Embedded Chain
9.8	Mean Absorption Time and Probabilities
Exe	rcises

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 Figure 9.1 represents a sample path of a Poisson process. MH3512 AY19-20



Figure 9.1: Sample path of a Poisson process $(N_t)_{t \in \mathbb{R}_+}$.

We denote by $(T_k)_{k \ge 1}$ 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.*

$$T_k = \inf\{t \in \mathbb{R}_+ \; : \; N_t = k\}, \qquad k \geqslant 1,$$

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 $(T_k)_{k \ge 1}$ as

$$N_t = \sum\limits_{k \geqslant 1} k \mathbb{1}_{[T_k,T_{k+1})}(t) = \sum\limits_{k \geqslant 1} \mathbb{1}_{[T_k,\infty)}(t), \qquad t \in \mathbb{R}_+,$$

MH3512 AY19-20

where

$$\mathbb{1}_{[T_k,\infty)}(t) = \left\{egin{array}{ll} 1 & ext{if} \ t \geqslant T_k, \ 0 & ext{if} \ 0 \leqslant t < T_k, \end{array}
ight.$$

and

$$\mathbb{1}_{[T_k,T_{k+1})}(t) = egin{cases} 1 & ext{if } T_k \leqslant t < T_{k+1}, \qquad k \geqslant 0, \ 0 & ext{if } 0 \leqslant t < T_k ext{ or } t \geqslant T_{k+1}, \qquad k \geqslant 0. \end{cases}$$

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 $0 \leqslant t_0 < t_1 < \cdots < t_n$ and $n \geqslant 1$ the increments

$$N_{t_1} - N_{t_0}, \ldots, \; N_{t_n} - N_{t_{n-1}},$$

over the disjoint time intervals $[t_0, t_1), [t_1, t_2), \ldots, [t_{n-2}, t_{n-1}), [t_{n-1}, t_n]$ 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 $0 \le s \le t$.

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

$$\mathbb{P}(N_t-N_s=k)=\mathrm{e}^{-\lambda(t-s)}rac{(\lambda(t-s))^\kappa}{k!}, \qquad k\in\mathbb{N}, \quad 0\leqslant s\leqslant t,$$

for some constant $\lambda > 0$.

Theorem 9.1 shows in particular that

$$\mathbb{E}[N_t - N_s] = \lambda(t - s)$$
 and $\operatorname{Var}[N_t - N_s] = \lambda(t - s),$

 $0 \leq s \leq t$, cf. Relations (B.4) and (B.5) in the solution of Exercise 1.3-(a).

The parameter $\lambda > 0$ is called the <u>intensity</u> of the process and it can be recovered given from $\mathbb{P}(N_h = 1) = \lambda h e^{-\lambda h}$ as the limit

$$\lambda = \lim_{h \searrow 0} \frac{1}{h} \mathbb{P}(N_h = 1).$$
(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

$$G_t(u):=\mathbb{E}[u^{N_t}], \qquad -1\leqslant u\leqslant 1,$$

satisfies

$$G_t(u):=(G_1(u))^t, \qquad -1\leqslant u\leqslant 1,$$

which implies that

$$G_t(u):=\mathrm{e}^{-tf(u)},\qquad -1\leqslant u\leqslant 1,$$

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

$$f(u)=\lambda imes(1-u),\qquad -1\leqslant u\leqslant 1,$$

where $\lambda > 0$ is given by (9.1.1).

MH3512 AY19-20

In particular, given that $N_0 = 0$, the random variable N_t has a Poisson distribution with parameter λt :

$$\mathbb{P}(N_t=k)=rac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}, \qquad t\in\mathbb{R}_+.$$

From (9.1.1) above we see that^{*}

$$\left\{egin{array}{ll} \mathbb{P}(N_h=0)=\mathrm{e}^{-\lambda h}=1-\lambda h+o(h), & h\searrow 0, \ \mathbb{P}(N_h=1)=h\lambda\mathrm{e}^{-\lambda h}\simeq\lambda h, & h\searrow 0, \end{array}
ight.$$

and more generally that

$$\left(\mathbb{P}(N_{t+h} - N_t = 0) = e^{-\lambda h} = 1 - \lambda h + o(h), \quad h \searrow 0,$$
(9.1.2a)

$$\left\{ \mathbb{P}(N_{t+h} - N_t = 1) = \lambda h e^{-\lambda h} \simeq \lambda h, \quad h \searrow 0, \right.$$
(9.1.2b)

$$\left(\mathbb{P}(N_{t+h} - N_t = 2) = h^2 \frac{\lambda^2}{2} e^{-\lambda h} \simeq h^2 \frac{\lambda^2}{2}, \qquad h \searrow 0,$$
(9.1.2c)

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

$$\sum_{k=0}^{n-1} X_{kh} \simeq \sum_{k=0}^{n-1} (N_{(k+1)h} - N_{kh}) = N_t - N_0 = N_t$$

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

$$\mathbb{P}(N_{t+h}-N_t=k)\simeq h^krac{\lambda^k}{k!}, \qquad h\searrow 0, \qquad t>0.$$

*The notation $f(h) \simeq h^k$ means $\lim_{h \to 0} f(h)/h^k = 1$, and f(h) = o(h) means $\lim_{h \to 0} f(h)/h = 0$.

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

$$\mathbb{P}(T_1 > t) = \mathbb{P}(N_t = 0) = \mathrm{e}^{-\lambda t}, \qquad t \geqslant 0,$$

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

$$\{T_n > t\} = \{N_t < n\}, \qquad t \geqslant 0, \quad n \geqslant 1.$$

Indeed, stating that the *n*-th 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

$$x\longmapsto\lambda^n\mathrm{e}^{-\lambda x}rac{x^{n-1}}{(n-1)!}$$

 $x \in \mathbb{R}_+, n \geqslant 1.$

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 $n \ge 1$, assuming that

$$\mathbb{P}(T_{n-1}>t)=\lambda\int_t^\infty \mathrm{e}^{-\lambda s}rac{(\lambda s)^{n-2}}{(n-2)!}ds,$$

MH3512 AY19-20 $\,$

at the rank n-1 with $n \ge 2$, we obtain

$$egin{aligned} \mathbb{P}(T_n > t) &= \mathbb{P}(T_n > t \geqslant T_{n-1}) + \mathbb{P}(T_{n-1} > t) \ &= \mathbb{P}(N_t = n-1) + \mathbb{P}(T_{n-1} > t) \ &= \mathrm{e}^{-\lambda t} rac{(\lambda t)^{n-1}}{(n-1)!} + \lambda \int_t^\infty \mathrm{e}^{-\lambda s} rac{(\lambda s)^{n-2}}{(n-2)!} ds \ &= \lambda \int_t^\infty \mathrm{e}^{-\lambda s} rac{(\lambda s)^{n-1}}{(n-1)!} ds, \quad t \in \mathbb{R}_+, \end{aligned}$$

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

$au_k = T_{k+1} - T_k, \qquad k \geqslant 1,$

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 5.9 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.*

$$\mathbb{P}(au_0>t_0, au_1>t_1,\ldots, au_n>t_n)=\mathrm{e}^{-\lambda(t_0+t_1+\cdots+t_n)},\qquad t_0,t_1,\ldots,t_n\in\mathbb{R}_+.$$

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</pre>

Similarly, random samples of Poisson process paths can be generated using the following code.

In other words, we have

$$\mathbb{P}(\tau_0 > t_0, \tau_1 > t_1, \dots, \tau_n > t_n) = \mathbb{P}(\tau_0 > t_0) \times \dots \times \mathbb{P}(\tau_n > t_n)$$

$$= \prod_{k=0}^n e^{-\lambda t_k}$$

$$= e^{-\lambda(t_0 + \dots + t_n)},$$
(9.1.3)

for all $t_0, t_1, \ldots, t_n \in \mathbb{R}_+$. In addition, from Proposition 9.2 the sum

$$T_k = au_0 + au_1 + \dots + au_{k-1}, \qquad k \geqslant 1,$$

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

$$\mathbb{E}[au_k] = \lambda \int_0^\infty x \mathrm{e}^{-\lambda x} dx = rac{1}{\lambda},$$

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 Chapter 11.

9.2 Continuous-Time Markov Chains

A S-valued continuous-time stochastic process $(X_t)_{t \in \mathbb{R}_+}$ is said to be <u>Markov</u>, or to have the <u>Markov property</u> 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

$$0 < s_1 < \dots < s_{n-1} < s < t$$

we have

$$\mathbb{P}(X_t = j \mid X_s = i_n, \; X_{s_{n-1}} = i_{n-1}, \dots, X_{s_1} = i_0) = \mathbb{P}(X_t = j \mid X_s = i_n).$$
 MH3512(9.2.91)

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 Section 9.1 is a continuous-time Markov chain because it has independent increments by Condition (i) page 340. 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 S we have (note that $X_0 = 0$ here)

$$\begin{split} \mathbb{P}(X_t = j \mid X_s = i_n, \ X_{s_{n-1}} = i_{n-1}, \dots, X_{s_1} = i_1) \\ &= \frac{\mathbb{P}(X_t = j, X_s = i_n, X_{s_{n-1}} = i_{n-1}, \dots, X_{s_1} = i_1)}{\mathbb{P}(X_s = i_n, \ X_{s_n = 1} = i_{n-1}, \dots, X_{s_2} = i_2, X_{s_1} = i_1)} \\ &= \frac{\mathbb{P}(X_t - X_s = j - i_n, X_s = i_n, \dots, X_{s_2} = i_2, X_{s_1} = i_1)}{\mathbb{P}(X_s = i_n, \dots, X_{s_2} = i_2, X_{s_1} = i_1)} \\ &= \frac{\mathbb{P}(X_t - X_s = j - i_n)\mathbb{P}(X_s = i_n, X_{s_{n-1}} = i_{n-1}, \dots, X_{s_2} = i_2, X_{s_1} = i_1)}{\mathbb{P}(X_s = i_n, X_{s_{n-1}} = i_{n-1}, \dots, X_{s_2} = i_2, X_{s_1} = i_1)} \\ &= \mathbb{P}(X_t - X_s = j - i_n) = \frac{\mathbb{P}(X_t - X_s = j - i_n)\mathbb{P}(X_s = i_n)}{\mathbb{P}(X_s = i_n)} \\ &= \frac{\mathbb{P}(X_t - X_s = j - i_n \text{ and } X_s = i_n)}{\mathbb{P}(X_s = i_n)} \\ &= \frac{\mathbb{P}(X_t = j \text{ and } X_s = i_n)}{\mathbb{P}(X_s = i_n)} = \mathbb{P}(X_t = j \mid X_s = i_n), \end{split}$$

cf. (4.1.5) for the discrete-time version of this argument. Hence, continuous-time processes with *independent increments* are Markov chains. However, <u>not all continuous-time Markov chains have independent increments</u>, 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. MH3512 AY19-20 In other words, a continuous-time Markov chain $(X_t^b)_{t \in \mathbb{R}_+}$ such that^{*}

$$\mathbb{P}(X^b_{t+h}=i+1\mid X^b_t=i) ~=~ \mathbb{P}(X^b_{t+h}-X^b_t=1\mid X^b_t=i) \ \simeq~ \lambda_i h, \qquad h\searrow 0, \quad i\in\mathbb{S},$$

and

$$\mathbb{P}(X_{t+h}^{b} = X_{t}^{b} \mid X_{t}^{b} = i) = \mathbb{P}(X_{t+h}^{b} - X_{t}^{b} = 0 \mid X_{t}^{b} = i) \\
= 1 - \lambda_{i}h + o(h), \quad h \searrow 0, \quad i \in \mathbb{S},$$
(9.2.2)

is called a <u>pure birth process</u> with (possibly) state-dependent birth rates $\lambda_i \ge 0$, $i \in S$, see Figure 9.2. Its inter-jump times $(\tau_k)_{k\ge 0}$ form a sequence of exponential independent random variables with state-dependent parameters.

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 $\lambda_i = \lambda > 0$, $i \in \mathbb{N}$.

^{*}Recall that by definition $f(h) \simeq g(h), h \to 0$, if and only if $\lim_{h\to 0} f(h)/g(h) = 1$.



Figure 9.2: Sample path of a birth process $(X_t^b)_{t \in \mathbb{R}_+}$.

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 (i+1) has an exponential distribution with parameter λ_i . Indeed we have, using the Markov property in continuous time,

$$\begin{split} \mathbb{P}(\tau_{i,i+1} > t+h \mid \tau_{i,i+1} > t \text{ and } X_0^b = i) &= \frac{\mathbb{P}(\tau_{i,i+1} > t+h \mid X_0^b = i)}{\mathbb{P}(\tau_{i,i+1} > t \mid X_0^b = i)} \\ &= \frac{\mathbb{P}(X_{t+h}^b = i \mid X_0^b = i)}{\mathbb{P}(X_t^b = i \mid X_0^b = i)} \\ &= \frac{\mathbb{P}(X_{t+h}^b = i \text{ and } X_0^b = i)\mathbb{P}(X_0^b = i)}{\mathbb{P}(X_t^b = i \text{ and } X_0^b = i)\mathbb{P}(X_0^b = i)} \\ &= \frac{\mathbb{P}(X_{t+h}^b = i \text{ and } X_0^b = i)}{\mathbb{P}(X_t^b = i \text{ and } X_0^b = i)} \end{split}$$
(MH3512 AY19-20

$$= \frac{\mathbb{P}(X_{t+h}^{b} = i, X_{t}^{b} = i, X_{0}^{b} = i)}{\mathbb{P}(X_{t}^{b} = i \text{ and } X_{0}^{b} = i)}$$

$$= \mathbb{P}(X_{t+h}^{b} = i \mid X_{t}^{b} = i \text{ and } X_{0}^{b} = i)$$

$$= \mathbb{P}(X_{t+h}^{b} = i \mid X_{t}^{b} = i)$$

$$= \mathbb{P}(X_{h}^{b} = i \mid X_{0}^{b} = i)$$

$$= \mathbb{P}(\tau_{i,i+1} > h \mid X_{0}^{b} = i)$$

$$= 1 - \lambda_{i}h + o(h),$$
(9.2.3)

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 Chapter 12 for a similar argument.

From (9.2.3) we have

$$\begin{split} \frac{\mathbb{P}(\tau_{i,i+1} > t+h \mid X_0^b = i) - \mathbb{P}(\tau_{i,i+1} > t \mid X_0^b = i)}{h\mathbb{P}(\tau_{i,i+1} > t \mid X_0^b = i)} \\ &= \frac{1}{h} \left(\frac{\mathbb{P}(\tau_{i,i+1} > t+h \mid X_0^b = i)}{\mathbb{P}(\tau_{i,i+1} > t \mid X_0^b = i)} - 1 \right) \\ &= \mathbb{P}(\tau_{i,i+1} > t+h \mid \tau_{i,i+1} > t \text{ and } X_0^b = i) - 1 \\ &\simeq -\lambda_i, \quad h \to 0, \end{split}$$

which can be read as the differential equation

$$rac{d}{dt}\log \mathbb{P}(au_{i,i+1}>t\mid X_0^b=i)=-\lambda_i,$$

where "log" denotes the *natural logarithm* "ln", with solution

$$\mathbb{P}(\tau_{i,i+1} > t \mid X_0^b = i) = e^{-\lambda_i t}, \quad t \in \mathbb{R}_+,$$
(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 (i+1) is given by

$$\mathbb{E}[au_{i,i+1} \mid X^b_0 = i] = rac{1}{\lambda_i}, \qquad i \in \mathbb{S}_{+}$$

see (9.4.10) 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 $(\tau_{j,j+1})_{j \ge i}$ is made of independent random variables which are respectively exponentially distributed with parameters λ_j , $j \ge i$.

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

$$X^b_t = i + \sum\limits_{i < j < \infty} \mathbb{1}_{[T^b_{i,j},\infty)}(t), \qquad t \in \mathbb{R}_+.$$

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 $\lambda > 0$.

Death process

A continuous-time Markov chain $(X_t^d)_{t \in \mathbb{R}_+}$ such that

$$\left\{egin{array}{ll} \mathbb{P}(X^d_{t+h}-X^d_t=-1\mid X^d_t=i)\simeq \mu_ih, & h\searrow 0, & i\in\mathbb{S}, \ \mathbb{P}(X^d_{t+h}-X^d_t=0\mid X^d_t=i)=1-\mu_ih+o(h), & h\searrow 0, & i\in\mathbb{S}, \end{array}
ight.$$

is called a <u>pure death process</u> with (possibly) state-dependent death rates $\mu_i \ge 0$, $i \in S$. Its inter-jump times $(\tau_k)_{k\ge 0}$ form a sequence of exponential independent random variables with state-dependent parameters, see Figure 9.3. MH3512 AY19-20



Figure 9.3: Sample path of a death process $(X_t^d)_{t \in \mathbb{R}_+}$.

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 (i-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 , $j \leq i$, wich

$$\mathbb{P}(au_{j,j-1} > t) = \mathrm{e}^{-\mu_j t}, \qquad t \in \mathbb{R}_+$$

and

$$\mathbb{E}[au_{i,i-1}] = rac{1}{\mu_i}, \qquad i \in \mathbb{S}_{+}$$

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

$$X^d_t = i - \sum\limits_{-\infty < j < i} \mathbb{1}_{[T^d_{i,j},\infty)}(t), \qquad t \in \mathbb{R}_+.$$
MH3512 AY19-20

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 $\mu_n = \lambda \Im_0^2 \overline{0}$, $n \in \mathbb{N}$.

Birth and death process

A continuous-time Markov chain $(X_t)_{t \in \mathbb{R}_+}$ such that, for all $i \in S$,

$$\begin{cases} \mathbb{P}(X_{t+h} - X_t = 1 \mid X_t = i) \simeq \lambda_i h, \quad h \searrow 0, \\ \mathbb{P}(X_{t+h} - X_t = -1 \mid X_t = i) \simeq \mu_i h, \quad h \searrow 0, \text{ and} \\ \mathbb{P}(X_{t+h} - X_t = 0 \mid X_t = i) = 1 - (\lambda_i + \mu_i)h + o(h), \quad h \searrow 0, \end{cases}$$
(9.2.5b)

is called a birth and death process with (possibly) state-dependent birth rates $\lambda_i \ge 0$ and death rates $\mu_i \ge 0$, $i \in S$, see Figure 9.4.9-20



Figure 9.4: Sample path of a birth and death process $(X_t)_{t \in \mathbb{R}_+}$.

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

$$au_i = \min(au_{i,i+1}, au_{i,i-1})$$

i.e. au_i is an exponentially distributed random variable with parameter $\lambda_i + \mu_i$ and

$$\mathbb{E}[au_i] = rac{1}{\lambda_i + \mu_i}.$$
 MH3512 AY19-20

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 354

$$\begin{split} \mathbb{P}(\min(\tau_{i,i+1} \text{ and } \tau_{i,i-1}) > t) &= \mathbb{P}(\tau_{i,i+1} > t \text{ and } \tau_{i,i-1} > t) \\ &= \mathbb{P}(\tau_{i,i+1} > t) \mathbb{P}(\tau_{i,i-1} > t) \\ &= \mathrm{e}^{-t(\lambda_i + \mu_i)}, \quad t \in \mathbb{R}_+, \end{split}$$

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 Chapter 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

$$P_{i,j}(t):=\mathbb{P}(X_{t+s}=j\mid X_s=i), \qquad i,j\in\mathbb{S}, \ \ s,t\in\mathbb{R}_+,$$

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 <u>irreducible</u> 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:

$$\left[\begin{array}{c} P_{i,j}(t) \end{array}
ight]_{i,j\in\mathbb{S}} = \left[\begin{array}{c} \mathbb{P}(X_{t+s}=j \mid X_s=i) \end{array}
ight]_{i,j\in\mathbb{S}}, ext{MH3512 AY19-2}$$

also written as

$$[P_{i,j}(t)]_{i,j\in\mathbb{S}} = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ \cdots & P_{-2,-2}(t) & P_{-2,-1}(t) & P_{-2,0}(t) & P_{-2,1}(t) & P_{-2,2}(t) & \cdots \\ \cdots & P_{-1,-2}(t) & P_{-1,-1}(t) & P_{-1,0}(t) & P_{-1,1}(t) & P_{-1,2}(t) & \cdots \\ \cdots & P_{0,-2}(t) & P_{0,-1}(t) & P_{0,0}(t) & P_{0,1}(t) & P_{0,2}(t) & \cdots \\ \cdots & P_{1,-2}(t) & P_{1,-1}(t) & P_{1,0}(t) & P_{1,1}(t) & P_{1,2}(t) & \cdots \\ \cdots & P_{2,-2}(t) & P_{2,-1}(t) & P_{2,0}(t) & P_{2,1}(t) & P_{2,2}(t) & \cdots \\ \cdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

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

$$\sum_{j \in \mathbb{S}} \mathbb{P}(X_{t+s} = j \mid X_s = i) = 1, \quad i \in \mathbb{S},$$

$$(9.3.1)$$

all <u>rows</u> of the transition matrix semigroup $(P(t))_{t\in\mathbb{R}_+}$ satisfy the condition

$$\sum\limits_{j\in\mathbb{S}}P_{i,j}(t)=1,$$

for $i \in S$. In the sequel we will only consider N-valued Markov process, and in this case the transition semigroup $(P(t))_{t \in \mathbb{R}_+}$ of the Markov process is written as

$$P(t) = [P_{i,j}(t)]_{i,j \in \mathbb{N}} = egin{bmatrix} P_{0,0}(t) & P_{0,1}(t) & P_{0,2}(t) & \cdots \ P_{1,0}(t) & P_{1,1}(t) & P_{1,2}(t) & \cdots \ P_{2,0}(t) & P_{2,1}(t) & P_{2,2}(t) & \cdots \ dots & dots &$$

MH3512 AY19-20

$$\sum\limits_{j \geqslant 0} P_{i,j}(t) = 1,$$

for all $i \in \mathbb{N}$ and $t \in \mathbb{R}_+$.

<u>Exercise</u>: Write down the transition semigroup $[P_{i,j}(t)]_{i,j\in\mathbb{N}}$ of the Poisson process $(N_t)_{t\in\mathbb{R}_+}$.

We can show that

$$\left[\begin{array}{ccc} P_{i,j}(t) \end{array}
ight]_{i,j\in\mathbb{N}} = \left[egin{array}{ccc} \mathrm{e}^{-\lambda t} & \lambda t \mathrm{e}^{-\lambda t} & rac{\lambda^2 t^2}{2} \mathrm{e}^{-\lambda t} & \cdots \ 0 & \mathrm{e}^{-\lambda t} & \lambda t \mathrm{e}^{-\lambda t} & \cdots \ 0 & 0 & \mathrm{e}^{-\lambda t} & \cdots \ dots & dots &$$

Indeed we have

$$P_{i,j}(t) = \mathbb{P}(N_{s+t} = j \mid N_s = i) = \frac{\mathbb{P}(N_{s+t} = j \text{ and } N_s = i)}{\mathbb{P}(N_s = i)}$$

$$= \frac{\mathbb{P}(N_{s+t} - N_s = j - i \text{ and } N_s = i)}{\mathbb{P}(N_s = i)} = \frac{\mathbb{P}(N_{s+t} - N_s = j - i)\mathbb{P}(N_s = i)}{\mathbb{P}(N_s = i)}$$

$$= \mathbb{P}(N_{s+t} - N_s = j - i) = \begin{cases} e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!} & \text{if } j \ge i, \\ 0 & \text{if } j < i. \end{cases}$$

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

♦ 357

$$P(t) = \left[egin{array}{ccccccc} P_{0,0}(t) & P_{0,1}(t) & P_{0,2}(t) & \cdots & P_{0,N}(t) \ P_{1,0}(t) & P_{1,1}(t) & P_{1,2}(t) & \cdots & P_{1,N}(t) \ P_{2,0}(t) & P_{2,1}(t) & P_{2,2}(t) & \cdots & P_{2,N}(t) \ dots & dot$$

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

$$P(s+t) = P(s)P(t) = P(t)P(s),$$

(9.3.2)

which is called the <u>semigroup property</u>.

Proof

Using the Markov property and denoting by 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

$$\begin{aligned} P_{i,j}(t+s) &= \mathbb{P}(X_{t+s} = j \mid X_0 = i) \\ &= \sum_{l \in \mathbb{S}} \mathbb{P}(X_{t+s} = j \text{ and } X_s = l \mid X_0 = i) = \sum_{l \in \mathbb{S}} \frac{\mathbb{P}(X_{t+s} = j, X_s = l, X_0 = i)}{\mathbb{P}(X_0 = i)} \\ &= \sum_{l \in \mathbb{S}} \frac{\mathbb{P}(X_{t+s} = j, X_s = l, X_0 = i)}{\mathbb{P}(X_s = l \text{ and } X_0 = i)} \frac{\mathbb{P}(X_s = l \text{ and } X_0 = i)}{\mathbb{P}(X_0 = i)} \end{aligned}$$

MH3512 AY19-20

$$= \sum_{l \in \mathbb{S}} \mathbb{P}(X_{t+s} = j \mid X_s = l \text{ and } X_0 = i) \mathbb{P}(X_s = l \mid X_0 = i)$$

$$= \sum_{l \in \mathbb{S}} \mathbb{P}(X_{t+s} = j \mid X_s = l) \mathbb{P}(X_s = l \mid X_0 = i) = \sum_{l \in \mathbb{S}} P_{i,l}(s) P_{l,j}(t)$$

$$= [P(s)P(t)]_{i,j},$$

 $i,j \in \mathbb{S}, s,t \in \mathbb{R}_+$. We have shown the relation

$$P_{i,j}(s+t) = \sum\limits_{l \in \mathbb{S}} P_{i,l}(s) P_{l,j}(t),$$

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

$$P(s)P(t) = \begin{bmatrix} e^{-\lambda s} & \lambda s e^{-\lambda s} & \frac{\lambda^2}{2} s^2 e^{-\lambda s} & \cdots \\ 0 & e^{-\lambda s} & \lambda s e^{-\lambda s} & \cdots \\ 0 & 0 & e^{-\lambda s} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \times \begin{bmatrix} e^{-\lambda t} & \lambda t e^{-\lambda t} & \frac{\lambda^2}{2} t^2 e^{-\lambda t} & \cdots \\ 0 & e^{-\lambda t} & \lambda t e^{-\lambda t} & \cdots \\ 0 & 0 & e^{-\lambda t} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

$$= \begin{bmatrix} e^{-\lambda(s+t)} & \lambda(s+t)e^{-\lambda(s+t)} & \frac{\lambda^2}{2}(s+t)^2 e^{-\lambda(s+t)} & \cdots \\ 0 & e^{-\lambda(s+t)} & \lambda(s+t) & e^{-\lambda(s+t)} & \cdots \\ 0 & 0 & e^{-\lambda(s+t)} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$
$$= P(s+t).$$

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. (B.9) 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

$$\begin{split} &[P(s)P(t)]_{i,j} = \sum_{l \geqslant 0} P_{i,l}(s)P_{l,j}(t) = \sum_{l=i}^{j} P_{i,l}(s)P_{l,j}(t) \\ &= e^{-\lambda s - \lambda t} \sum_{l=i}^{j} \frac{(\lambda s)^{l-i}}{(l-i)!} \frac{(\lambda t)^{j-l}}{(j-l)!} = e^{-\lambda s - \lambda t} \frac{1}{(j-i)!} \sum_{l=i}^{j} \binom{j-i}{l-i} (\lambda s)^{l-i} (\lambda t)^{j-l} \\ &= e^{-\lambda s - \lambda t} \frac{1}{(j-i)!} \sum_{l=0}^{j-i} \binom{j-i}{l} (\lambda s)^{l} (\lambda t)^{j-i-l} = e^{-\lambda (s+t)} \frac{1}{(j-i)!} (\lambda s + \lambda t)^{j-i} \\ &= P_{i,j}(s+t), \qquad s, t \in \mathbb{R}_+. \end{split}$$

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 S,

$$P'(t) \;=\; \lim_{h\searrow 0} rac{P(t+h)-P(t)}{h} = \lim_{h\searrow 0} rac{P(t)P(h)-P(t)}{h}$$

MH3512 AY19-20

(5359)

$$= P(t) \lim_{h \searrow 0} \frac{P(h) - P(0)}{h}$$

= $= P(t)Q,$ (9.4.1)

where

$$Q:=P'(0)=\lim_{h\searrow 0}rac{P(h)-P(0)}{h}$$

is called the infinitesimal generator of $(X_t)_{t \in \mathbb{R}_+}$.

When $S = \{0, 1, \dots, N\}$ we will denote by $\lambda_{i,j}, i, j \in S$, the entries of the infinitesimal generator matrix $Q = (\lambda_{i,j})_{i,j \in S}$, *i.e.*

$$Q = \frac{dP(t)}{dt}_{|t=0} = [\lambda_{i,j}]_{0 \leq i,j \leq N} = \begin{bmatrix} \lambda_{0,0} & \lambda_{0,1} & \lambda_{0,2} & \cdots & \lambda_{0,N} \\ \lambda_{1,0} & \lambda_{1,1} & \lambda_{1,2} & \cdots & \lambda_{1,N} \\ \lambda_{2,0} & \lambda_{2,1} & \lambda_{2,2} & \cdots & \lambda_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{N,0} & \lambda_{N,1} & \lambda_{N,2} & \cdots & \lambda_{N,N} \end{bmatrix}.$$
(9.4.2)

Denoting $Q = [\lambda_{i,j}]_{i,j \in \mathbb{S}}$, for all $i \in \mathbb{S}$ we have

$$\sum\limits_{j\in\mathbb{S}}\lambda_{i,j}=\sum\limits_{j\in\mathbb{S}}P_{i,j}'(0)=rac{d}{dt}\sum\limits_{j\in\mathbb{S}}P_{i,j}(t)_{|t=0}=rac{d}{dt}\mathbb{1}_{|t=0}=0,$$

hence the <u>rows</u> of the infinitesimal generator matrix $Q = [\lambda_{i,j}]_{i,j \in S}$ always add up to 0, i.e.

$$\sum\limits_{l\in\mathbb{S}}\lambda_{i,l}=\lambda_{i,i}+\sum\limits_{l
eq i}\lambda_{i,l}=0,$$

or

$$\lambda_{i,i} = -\sum_{l \neq i} \lambda_{i,l}.$$
(9.4.3)
Note that a state (i) such that $\lambda_{i,j} = 0$ for all $j \in S$ is absorbing.

The equation (9.4.1), *i.e.*

$$P'(t) = P(t)Q, \quad t > 0,$$
 (9.4.4)

is called the *forward Kolmogorov equation*, cf. (1.2.2). In a similar way we can show that

$$P'(t) = QP(t), \quad t > 0,$$
 (9.4.5)

which is called the *backward Kolmogorov* equation.

The forward and backward Kolmogorov equations (9.4.4)-(9.4.5) can be solved either using the matrix exponential e^{tQ} defined as

$$\exp(tQ) := \sum_{n \ge 0} \frac{t^n}{n!} Q^n = \mathbf{I}_{\mathbf{d}} + \sum_{n \ge 1} \frac{t^n}{n!} Q^n, \qquad (9.4.6)$$

or by viewing the Kolmogorov equations (9.4.4)-(9.4.5) component by component as systems of differential equations.

In (9.4.6) above, $Q^0 = \mathbf{I_d}$ is the identity matrix, written as

$$\mathrm{I_d} = egin{bmatrix} 1 & 0 & \cdots & 0 & 0 \ 0 & 1 & \cdots & 0 & 0 \ dots & dots & \ddots & dots & dots \ dots & dots & \ddots & dots & dots \ dots & dots & \ddots & dots & dots \ 0 & 0 & \cdots & 1 & 0 \ 0 & 0 & \cdots & 0 & 1 \ \end{pmatrix}$$

when the state space is $S = \{0, 1, \dots, N\}$. Using matrix exponentials, the solution of (9.4.4) is given by

 $P(t)=P(0)\exp(tQ)=\exp(tQ), \ \ t\in \mathbb{R}_+.$

We will often use the first order approximation in $h \to 0$ of

$$P(h) = \exp(hQ) = \mathrm{I_d} + \sum_{n \geqslant 1} rac{h^n}{n!} Q^n = \mathrm{I_d} + hQ + rac{h^2}{2!} Q^2 + rac{h^3}{3!} Q^3 + rac{h^4}{4!} Q^4 + \cdots,$$

given by

$$P(h) = I_d + hQ + o(h), \qquad h \searrow 0, \tag{9.4.7}$$

where o(h) is a function such that $\lim_{h\to 0} o(h)/h = 0$, *i.e.*

$$P(h) = egin{bmatrix} \lambda_{0,0} & \lambda_{0,1} & \lambda_{0,2} & \cdots & \lambda_{0,N} \ \lambda_{1,0} & 1 & \cdots & \lambda_{0,N} \ \lambda_{1,0} & \lambda_{1,1} & \lambda_{1,2} & \cdots & \lambda_{1,N} \ \lambda_{2,0} & \lambda_{2,1} & \lambda_{2,2} & \cdots & \lambda_{2,N} \ \vdots & \vdots & \vdots & \ddots & \vdots \ \lambda_{N,0} & \lambda_{N,1} & \lambda_{N,2} & \cdots & \lambda_{N,N} \ \end{bmatrix} + o(h), \quad h\searrow 0.$$

Relation (9.4.7) yields the transition probabilities over a small time interval of length h > 0, as:

$$\mathbb{P}(X_{t+h}=j\mid X_t=i)=P_{i,j}(h)=egin{cases} \lambda_{i,j}h+o(h), & i
eq j, \quad h\searrow 0,\ 1+\lambda_{i,i}h+o(h), & i=j, \quad h\searrow 0, \end{cases}$$

and by (9.4.3) we also have

$$\mathbb{P}(X_{t+h}=j \mid X_t=i) = P_{i,j}(h) = \begin{cases} \lambda_{i,j}h + o(h), & i \neq j, \quad h \searrow 0, \\ 1 - h \sum_{l \neq i} \lambda_{i,l} + o(h), & i = j, \quad h \searrow 0. \end{cases}$$
(9.4.8)

For example, in the case of a two-state continuous-time Markov chain we have

$$Q = \left[egin{array}{cc} -lpha & lpha \ & & \ eta & -eta \end{array}
ight],$$

with $\alpha, \beta \ge 0$, and

$$P(h) = \mathbf{I}_{d} + hQ + o(h)$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + h \begin{bmatrix} -\alpha & \alpha \\ \beta & -\beta \end{bmatrix} + o(h)$$

$$= \begin{bmatrix} 1 - \alpha h & \alpha h \\ \beta h & 1 - \beta h \end{bmatrix} + o(h), \qquad (9.4.9)$$

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 (1) 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

$$\mathbb{P}(X_{t+h}=j\mid X_t=i)\simeq\lambda_{i,j}h, \qquad h\searrow 0, \quad i
eq j,$$

and

$$\mathbb{P}(X_{t+h} \neq j \mid X_t = i) = 1 - \lambda_{i,j}h + o(h), \quad h \searrow 0, \quad i \neq j,$$

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

$$\mathbb{P}(X_{t+h}=j\mid X_t=i)=\lambda_{i,j}h+o(h), \qquad h\searrow 0, \quad i
eq j,$$

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.*

and we have

$$\mathbb{E}[au_{i,j}] = \lambda_{i,j} \int_0^\infty t \mathrm{e}^{-t\lambda_{i,j}} dt = rac{1}{\lambda_{i,j}}, \qquad i
eq j.$$

When i = j we have

$$\mathbb{P}(X_{t+h}
eq i \mid X_t=i) \simeq h\sum\limits_{l
eq i} \lambda_{i,l}=-\lambda_{i,i}h, \qquad h\searrow 0,$$

and

$$\mathbb{P}(X_{t+h}=i\mid X_t=i)=1-h\sum\limits_{l
eq i}\lambda_{i,l}+o(h)=1+\lambda_{i,i}h+o(h), \qquad h\searrow 0,$$

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.*

$$\mathbb{P}(au_i > t) = \exp\left(-t\sum\limits_{j
eq i} \lambda_{i,j}
ight) = \mathrm{e}^{t\lambda_{i,i}}, \qquad t \in \mathbb{R}_+.$$

In other words, we can also write the time au_i spent in state (i) as

$$au_i = \min_{j \in \mathbb{S} \atop j
eq i} au_{i,j},$$

and this recovers the fact that τ_i is an exponential random variable with parameter $\sum_{j \neq i} \lambda_{i,j}$, since

$$egin{aligned} \mathbb{P}(au_i > t) &= \mathbb{P}\left(\min_{j \in \mathbb{S} \ j
eq i} au_{i,j} > t
ight) \ &= \prod_{\substack{j \in \mathbb{S} \ j
eq i}} \mathbb{P}(au_{i,j} > t) \ &= \exp\left(-t \sum_{j
eq i} oldsymbol{\lambda}_{i,j}
ight) = \mathrm{e}^{t \lambda_{i,i}}, \qquad t \in \mathbb{R}_+. \end{aligned}$$

ॳ 364

cf. (1.5.8) in Chapter 1. In addition we have

$$\mathbb{E}[au_i] = \sum\limits_{l
eq i} \lambda_{i,l} \int_0^\infty t \exp\left(-t \sum\limits_{l
eq i} \lambda_{i,l}
ight) dt = rac{1}{\sum\limits_{l
eq i} \lambda_{i,l}} = -rac{1}{\lambda_{i,i}},$$

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

$$Q = \left[egin{array}{cc} -lpha & lpha \ & \ eta & -eta \end{array}
ight],$$

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 Section 9.5.

ii) Poisson process.

The generator of the Poisson process is given by $\lambda_{i,j} = \mathbb{1}_{\{j=i+1\}}\lambda, i \neq j, i.e.$

MH3512 AY19-20 $\,$

From the relation $P(h) = I_d + hQ + o(h)$ we recover the infinitesimal transition probabilities of the Poisson process as 366

$$\mathbb{P}(N_{t+h}-N_t=1)=\mathbb{P}(N_{t+h}=i+1\mid N_t=i)\simeq\lambda h,$$

 $h\searrow 0, i\in\mathbb{N}$, 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 = \left[egin{array}{cccc} \lambda_{i,j} \end{array}
ight]_{i,j\in\mathbb{N}} = \left[egin{array}{ccccc} -\lambda_0 & \lambda_0 & 0 & \cdots \ 0 & -\lambda_1 & \lambda_1 & \cdots \ 0 & 0 & -\lambda_2 & \cdots \ dots & dots & dots & dots & dots \ dots & dots & dots & dots & dots \ dots & dots & dots \ dots & dots & dots \ dots & dots \ dots & dots \ dots \$$

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.

$$P(h) = egin{bmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots \ 0 & 1 & 0 & \cdots \ 0 & 0 & 1 & \cdots \ \vdots & \vdots & \vdots & \ddots \ \end{bmatrix} + h egin{bmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots \ 0 & -\lambda_1 & \lambda_1 & \cdots \ 0 & 0 & -\lambda_2 & \cdots \ \vdots & \vdots & \vdots & \ddots \ \end{bmatrix} + o(h), \ \ h\searrow 0,$$

we recover the infinitesimal transition probabilities of the pure birth process as

$$\mathbb{P}(X_{t+h}-X_t=1\mid X_t=i)=\mathbb{P}(X_{t+h}=i+1\mid X_t=i)\simeq\lambda_ih,$$

 $h \searrow 0, i \in \mathbb{N}$, 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 $-\mathbb{N} = \{\dots, -2, -1, 0\}$ is

$$Q = \left[egin{array}{cccc} \lambda_{i,j} \end{array}
ight]_{i,j \leqslant 0} = \left[egin{array}{ccccc} \cdots & 0 & \mu_0 & -\mu_0 \ hooppooldow & \mu_1 & -\mu_1 & 0 \ hooppooldow & hooppoo$$

From the relation

$$P(h) = egin{bmatrix} \cdots & 0 & 0 & 1 \ \cdots & 0 & 1 & 0 \ \cdots & 1 & 0 & 0 \ \cdots & \vdots & \vdots & \vdots \end{bmatrix} + h egin{bmatrix} \cdots & \mu_1 & -\mu_1 & 0 \ \cdots & -\mu_2 & 0 & 0 \ \cdots & -\mu_2 & 0 & 0 \ \cdots & \vdots & \vdots & \vdots \end{bmatrix} + o(h), \ \ h\searrow 0,$$

we recover the infinitesimal transition probabilities

$$\mathbb{P}(X_{t+h}-X_t=-1\mid X_t=i)=\mathbb{P}(X_{t+h}=i-1\mid X_t=i)\simeq \mu_i h, \hspace{1em}h\searrow 0,$$

 $i \in \mathbb{S}$, and

$$\mathbb{P}(X_{t+h}=i\mid X_t=i)=\mathbb{P}(X_{t+h}-X_t=0\mid X_t=i)=1-\mu_ih+o(h), \hspace{1em}h\searrow 0,$$

 $i \in \mathbb{S}$, of the pure death process.

v) Birth and death process on $\{0, 1, \ldots, N\}$.

By (9.2.5a)-(9.2.5b) and (9.4.8), the generator of the birth and death process on $\{0, 1, \ldots, N\}$ is

with $\mu_0 = \lambda_N = 0$.

From the Relation (9.4.7) we have

$$P(h) = \mathrm{I_d} + hQ + o(h), \qquad h\searrow 0,$$

and we recover the infinitesimal transition probabilities

$$\mathbb{P}(X_{t+h}-X_t=1\mid X_t=i)\simeq\lambda_i h, \quad h\searrow 0, \quad i=0,1,\ldots,N,$$

and

$$\mathbb{P}(X_{t+h}-X_t=-1\mid X_t=i)\simeq \mu_i h, \quad h\searrow 0, \quad i=0,1,\ldots,N,$$

and

$$\mathbb{P}(X_{t+h}-X_t=0\mid X_t=i)=1-(\lambda_i+\mu_i)h+o(h), \qquad h\searrow 0, \quad i=0,1,\dots,N,$$

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

$$P_{i,i}(t) \geqslant \mathbb{P}(au_i > t) = \mathrm{e}^{-t(\lambda_i + \mu_i)}, \qquad t \in \mathbb{R}_+,$$

and

$$\mathbb{E}[au_i] = rac{1}{\lambda_i + \mu_i}$$

In the case of a pure birth process we find

$$P_{i,i}(t) = \mathbb{P}(au_i > t) = \mathrm{e}^{-t\lambda_i}, \qquad t \in \mathbb{R}_+,$$

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 $S = \{0, 1, ..., N\}$ with $\lambda_i = \lambda$ and $\mu_i = \mu$, i = 1, 2, ..., 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 $S = \{0, 1\}$, in the same way as in Section 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 $\alpha, \beta \ge 0$. The forward Kolmogorov equation (9.4.4) reads

$$P'(t) = P(t) \times \begin{bmatrix} -\alpha & \alpha \\ & \\ \beta & -\beta \end{bmatrix}, \quad t > 0,$$
(9.5.2)

i.e.

$$egin{aligned} P'_{0,0}(t) & P'_{0,1}(t) \ P_{0,1}(t) & P_{0,1}(t) \ P_{1,0}(t) & P'_{1,1}(t) \end{bmatrix} = egin{bmatrix} P_{0,0}(t) & P_{0,1}(t) \ P_{0,1}(t) & P_{0,1}(t) \ P_{0,1}(t) & P_{0,1}(t) \ P_{0,1}(t) & P_{0,1}(t) \ P_{0,1}(t) & P_{0,1}(t) \end{bmatrix} imes egin{bmatrix} -lpha & lpha \ P_{0,0}(t) & P_{0,1}(t) \ P_{0,1}(t) \ P_{0,1}(t) & P_{0,1}(t) \ P_{0,1}($$

$$\left\{egin{array}{ll} P_{0,0}'(t)=-lpha P_{0,0}(t)+eta P_{0,1}(t), & P_{0,1}'(t)=lpha P_{0,0}(t)-eta P_{0,1}(t), \ P_{1,0}'(t)=-lpha P_{1,0}(t)+eta P_{1,1}(t), & P_{1,1}'(t)=lpha P_{1,0}(t)-eta P_{1,1}(t), \end{array}
ight.$$

t > 0, which is a system of four differential equations * with initial condition

$$P(0) = egin{bmatrix} P_{0,0}(0) & P_{0,1}(0) \ & & \ P_{1,0}(0) & P_{1,1}(0) \end{bmatrix} = egin{bmatrix} 1 & 0 \ & & \ 0 & 1 \end{bmatrix} = \mathrm{I}_\mathrm{d}.$$

The solution of the forward Kolmogorov equation (9.5.2) is given by the matrix exponential

$$P(t)=P(0)\exp(tQ)=\exp(tQ)=\exp\left(tiggl(egin{array}{cc} -lpha & lpha\ & \ eta & \ eeta & \ eeta$$

which is computed in the next Proposition 9.6, see also the command MatrixExp[t*-a,a,b,-b].

Proposition 9.6

 $t \in \mathbb{R}_+$.

The solution P(t) of the forward Kolmogorov equation (9.5.2) is given by

$$P(t) = \begin{bmatrix} P_{0,0}(t) & P_{0,1}(t) \\ P_{1,0}(t) & P_{1,1}(t) \end{bmatrix}$$
(9.5.3)
$$= \begin{bmatrix} \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta} e^{-t(\alpha+\beta)} & \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta} e^{-t(\alpha+\beta)} \\ \frac{\beta}{\alpha+\beta} - \frac{\beta}{\alpha+\beta} e^{-t(\alpha+\beta)} & \frac{\alpha}{\alpha+\beta} + \frac{\beta}{\alpha+\beta} e^{-t(\alpha+\beta)} \end{bmatrix},$$

^{*}Please refer to MH3110 - Ordinary Differential Equations for more on first order linear systems of differential equations.

Proof

We will compute the matrix exponential e^{tQ} by the diagonalization technique. The matrix Q has two eigenvectors^{*a*}

1		$\left[-lpha ight]$	
	and		,
1		β	

with respective eigenvalues $\lambda_1 = 0$ and $\lambda_2 = -\alpha - \beta$.^b Hence it can be put in the diagonal form

 $Q = M \times D \times M^{-1}$

as follows:

$$Q = egin{bmatrix} 1 & -lpha \ 1 & eta \end{bmatrix} imes egin{bmatrix} \lambda_1 & 0 \ 0 & \lambda_2 \end{bmatrix} imes egin{bmatrix} rac{eta}{lpha+eta} & rac{lpha}{lpha+eta} \ -rac{1}{lpha+eta} & rac{1}{lpha+eta} \end{bmatrix}.$$

Consequently we have

$$egin{aligned} P(t) &= \exp(tQ) = \sum\limits_{n \geqslant 0} rac{t^n}{n!} Q^n = \sum\limits_{n \geqslant 0} rac{t^n}{n!} (M imes D imes M^{-1})^n \ &= \sum\limits_{n \geqslant 0} rac{t^n}{n!} M imes D^n imes M^{-1} = M imes \left(\sum\limits_{n \geqslant 0} rac{t^n}{n!} D^n
ight) imes M^{-1} \ &= M imes \exp(tD) imes M^{-1} \ &= \left[egin{aligned} 1 & -lpha \\ 1 & eta \end{array}
ight] imes \left[egin{aligned} \mathrm{e}^{t\lambda_1} & 0 \\ 0 & \mathrm{e}^{t\lambda_2} \end{array}
ight] imes \left[egin{aligned} rac{eta}{lpha + eta} & rac{lpha}{lpha + eta} \\ -rac{1}{lpha + eta} & rac{1}{lpha + eta} \end{array}
ight] \end{aligned}$$

<u>(</u>371

$$= \begin{bmatrix} 1 & -\alpha \\ 1 & \beta \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ 0 & e^{-t(\alpha+\beta)} \end{bmatrix} \times \begin{bmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ -\frac{1}{\alpha+\beta} & \frac{1}{\alpha+\beta} \end{bmatrix}$$
$$= \frac{1}{\alpha+\beta} \begin{bmatrix} \beta & \alpha \\ \beta & \alpha \end{bmatrix} + \frac{e^{-t(\alpha+\beta)}}{\alpha+\beta} \begin{bmatrix} \alpha & -\alpha \\ -\beta & \beta \end{bmatrix}$$
$$= \begin{bmatrix} \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta} e^{-t(\alpha+\beta)} & \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta} e^{-t(\alpha+\beta)} \\ \frac{\beta}{\alpha+\beta} - \frac{\beta}{\alpha+\beta} e^{-t(\alpha+\beta)} & \frac{\alpha}{\alpha+\beta} + \frac{\beta}{\alpha+\beta} e^{-t(\alpha+\beta)} \end{bmatrix}$$

•

٠

t > 0.

^{*a*}Please refer to MH1201 - Linear Algebra II for more on eigenvectors, eigenvalues, and diagonalization. ^{*b*}See for example the command Eigensystem[-a,a,b,-b].

From Proposition 9.6 we obtain the probabilities

$$\left\{egin{array}{l} \mathbb{P}(X_h=0\mid X_0=0)=rac{eta+lpha\mathrm{e}^{-(lpha+eta)h}}{lpha+eta}, \ \mathbb{P}(X_h=1\mid X_0=0)=rac{lpha}{lpha+eta}(1-\mathrm{e}^{-(lpha+eta)h}), \ \mathbb{P}(X_h=0\mid X_0=1)=rac{eta}{lpha+eta}(1-e^{-h(lpha+eta)h}), \ \mathbb{P}(X_h=1\mid X_0=1)=rac{lpha+eta\mathrm{e}^{-(lpha+eta)h}}{lpha+eta}, \quad h\in\mathbb{R}_+ \end{array}
ight.$$

MH3512 AY19-20

In other words, (9.5.3) can be rewritten as

$$P(h) = \begin{bmatrix} 1 - \frac{\alpha}{\alpha + \beta} (1 - e^{-(\alpha + \beta)h}) & \frac{\alpha}{\alpha + \beta} (1 - e^{-(\alpha + \beta)h}) \\ \frac{\beta}{\alpha + \beta} (1 - e^{-(\alpha + \beta)h}) & 1 - \frac{\beta}{\alpha + \beta} (1 - e^{-(\alpha + \beta)h}) \end{bmatrix}, \quad h > 0,$$
(9.5.4)

hence, since

$$1 - e^{-(\alpha + \beta)h} \simeq h(\alpha + \beta), \qquad h \searrow 0,$$

the expression (9.5.4) above recovers (9.4.9) as $h\searrow 0$, *i.e.* we have

$$P(h) = egin{bmatrix} 1-hlpha & hlpha \ & & \ & heta & 1-heta \end{bmatrix} + o(h) = \mathrm{I_d} + hQ + o(h), \qquad h\searrow 0,$$

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:

$$\lim_{t o \infty} P(t) = \lim_{t o \infty} \left[egin{array}{ccc} \mathbb{P}(X_t = 0 \mid X_0 = 0) & \mathbb{P}(X_t = 1 \mid X_0 = 0) \ \mathbb{P}(X_t = 0 \mid X_0 = 1) & \mathbb{P}(X_t = 1 \mid X_0 = 1) \end{array}
ight] = rac{1}{lpha + eta} \left[egin{array}{ccc} eta & lpha \ eta & lpha \end{array}
ight],$$

whenever $\alpha > 0$ or $\beta > 0$, whereas if $\alpha = \beta = 0$ we simply have

$$P(t) = egin{bmatrix} 1 & 0 \ & \ 0 & 1 \end{bmatrix} = \mathrm{I}_\mathrm{d}, \qquad t \in \mathbb{R}_+,$$

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

$$\lim_{t \to \infty} \mathbb{P}(X_t = 1 \mid X_0 = 0) = \lim_{t \to \infty} \mathbb{P}(X_t = 1 \mid X_0 = 1) = \frac{\alpha}{\alpha + \beta}$$
MH3512 AY19-20

ॳ 373

and

♦ 374

$$\lim_{t o\infty} \mathbb{P}(X_t=0\mid X_0=0) = \lim_{t o\infty} \mathbb{P}(X_t=0\mid X_0=1) = rac{eta}{lpha+eta}$$

and

$$\pi = [\pi_0, \pi_1] = \left(rac{eta}{lpha + eta}, rac{lpha}{lpha + eta}
ight) = \left(rac{1/lpha}{1/lpha + 1/eta}, rac{1/eta}{1/lpha + 1/eta}
ight)$$

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 $\beta/(\alpha + \beta)$.

Next, we consider a simulation of the two-state continuous Markov chain with infinitesimal generator

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 Figure 9.6 represents the sample average

$$y_t=rac{1}{t}\int_0^t x_s ds, \qquad t\in [0,1],$$

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.^{*}

^{*}Download the corresponding \mathbf{R} code or the **IPython notebook** that can be run *here*.



Figure 9.5: Sample path of a continuous-time two-state chain with $\alpha = 20$ and $\beta = 40$.



Figure 9.6: The proportion of process values in the state 1 converges to $1/3 = \alpha/(\alpha + \beta)$.

ॳ 375

9.6 Limiting and Stationary Distributions

A probability distribution $\pi = (\pi_i)_{i \in S}$ is said to be <u>stationary</u> 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

1

 $\pi Q = 0.$

Proof

Assuming that $\pi Q = 0$, we have

$$egin{aligned} \pi P(t) &= \pi \exp(tQ) = \pi \sum\limits_{n \geqslant 0} rac{t^n}{n!} Q^n \ &= \pi \left(\mathrm{I_d} + \sum\limits_{n \geqslant 1} rac{t^n}{n!} Q^n
ight) \ &= \pi + \sum\limits_{n \geqslant 1} rac{t^n}{n!} \pi Q^n \ &= \pi, \end{aligned}$$

since $\pi Q^n = \pi Q Q^{n-1} = 0$, $n \ge 1$. Conversely, the relation $\pi = \pi P(t)$ shows, by differentiation at t = 0, that

$$0 = \pi P'(0) = \pi Q$$

When $S = \{0, 1, \dots, N\}$ and the generator Q has the form (9.4.2) the relation $\pi Q = 0$ reads

$$\pi_0\lambda_{0,j}+\pi_1\lambda_{1,j}+\dots+\pi_N\lambda_{N,j}=0, \qquad j=0,1,\dots,N,$$

MH3512 AY19-20

i.e.,

$$\sum_{j=0\atop i
eq j}^N \pi_i\lambda_{j,i} = -\pi_j\lambda_{j,j}, \qquad j=0,1,\ldots,N,$$

hence from (9.4.3) we find the balance condition

$$\sum\limits_{i=0\atop i
eq j}^N \pi_i \lambda_{i,j} = \sum\limits_{k=0\atop k
eq j}^N \pi_j \lambda_{j,k},$$

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, ..., N.



Figure 9.7: Global balance condition (continuous time).

Next is the continuous-time analog of Proposition 7.7 in Section 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

$$\pi_j := \lim_{t \to \infty} \mathbb{P}(X_t = j \mid X_0 = i) = \lim_{t \to \infty} P_{i,j}(t), \qquad j \in \mathbb{S},$$

$$(9.6.1)$$

independent of the initial state $i \in S$. 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 $S = \{0, 1, \dots, N\}$ we note that by (9.6.1) since the limiting distribution is independent of the initial state it satisfies

$$egin{aligned} \lim_{t o\infty} P(t) \ &= \ egin{bmatrix} \lim_{t o\infty} P_{0,0}(t) & \cdots & \lim_{t o\infty} P_{0,N}(t) \ dots & dots & \ddots & dots \ dots & dots & \ddots & dots \ \lim_{t o\infty} P_{N,0}(t) & \cdots & \lim_{t o\infty} P_{N,N}(t) \ \end{bmatrix} \ &= \ egin{bmatrix} \pi_0 & \pi_1 & \cdots & \pi_N \ dots & dots & \ddots & dots \ \pi_0 & \pi_1 & \cdots & \pi_N \ dots & dots & \ddots & dots \ \pi \end{bmatrix} = egin{bmatrix} \pi \ dots \ dots \ dots \ \end{pmatrix}, \end{aligned}$$

where $\boldsymbol{\pi}$ is the row vector

$$\pi = [\pi_0, \pi_1, \dots, \pi_n]$$

By the forward Kolmogorov equation (9.4.4) and (9.6.1) we find that the limit of P'(t) exists as $t \to \infty$ since

$$\lim_{t o\infty}P'(t)=\lim_{t o\infty}P(t)Q=\left[egin{array}{c}\pi\dots\ dots\ \pi\end{array}
ight]Q.$$

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

$$P(t) = P(0) + \int_0^t P'(s) ds$$
(9.6.3)

to converge as $t \to \infty$. This shows that

$$\left[egin{array}{c} \pi Q \ dots \ \pi Q \end{array}
ight] = \left[egin{array}{c} \pi \ dots \ dots \ \pi \end{array}
ight] Q = 0$$

MH3512 AY19-20

<u>(† 37</u>8

by (9.4.4), hence we have $\pi Q = 0$, or $\sum_{i \in S} \pi_i \lambda_{i,j} = 0, j \in S$.

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 $P(h) = I_d + hQ + o(h)$ 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 [Lal13], 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 page 9 of [Sig14] 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.*

$$\lim_{t o\infty}P_{i,j}(t)=\lim_{t o\infty}\mathbb{P}(X_t=j\mid X_0=i)=\pi_j, \ \ j\in\mathbb{S},$$

which is independent of the initial state $i \in S$.

Proof

By e.g. Theorem 2.1 in Chapter 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 $\lambda_k \langle u^{(k)}, \pi \rangle_{\mathbb{R}^n} = \langle Q u^{(k)}, \pi \rangle_{\mathbb{R}^n} = \langle u^{(k)}, Q^\top \pi \rangle_{\mathbb{R}^n} = \pi^\top Q u^{(k)} = 0, \ k = 2, \ldots, n$. 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 $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. 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

$$\exp(tD) = ext{diag}(1, ext{e}^{t\lambda_2}, \dots, ext{e}^{t\lambda_n}),$$

and the eigenvalues $\lambda_2, \ldots, \lambda_n$ have to be strictly negative, hence we have

$$egin{aligned} & \lim_{ o \infty} [\mathbb{P}(X_t=j \mid X_0=i)]_{1\leqslant i,j\leqslant n} = \lim_{t o \infty} [M^{-1}\exp(tD)M]_{1\leqslant i,j\leqslant n} \ & = egin{bmatrix} 1 & u_1^{(2)} & \cdots & u_1^{(n)} \ 1 & u_2^{(2)} & \cdots & u_2^{(n)} \ dots & dots & \ddots & dots \ 1 & u_n^{(2)} & \cdots & u_n^{(n)} \ \end{bmatrix} egin{bmatrix} 1 & 0 & \cdots & 0 \ dots & dots & \ddots & dots \ 0 & 0 & \cdots & 0 \ dots & dots & \ddots & dots \ 0 & 0 & \cdots & 0 \ dots & dots & \ddots & dots \ 0 & 0 & \cdots & 0 \ \end{pmatrix} egin{bmatrix} \pi_1 & \cdots & \pi_n \ M_{2,1} & \cdots & M_{2,n} \ dots & dots & \ddots & dots \ M_{n,1} & \cdots & M_{n,n} \ \end{pmatrix} \ & = egin{bmatrix} 1 & 0 & \cdots & 0 \ dots & dots & \ddots & dots \ dots & dots & dots & dots & dots \ dots & dots & dots & dots \ dots & dots & dots & dots \ dots \ dots & dots \ do$$

The discrete-time Ergodic Theorem 7.12 also admits a continuous-time version with a similar proof, stating that if the chain $(X_t)_{t \in \mathbb{R}_{\neq 0}}$

is *irreducible* then the sample average of the number of visits to state (i) converges almost surely to π_i , *i.e.*, (5381)

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t \mathbb{1}_{\{X_t = i\}} dt = \pi_i, \qquad i \in \mathbb{S}.$$
(9.6.4)

Examples

i) Two-state Markov chain.

Consider the two-state Markov chain with infinitesimal generator

$$Q = egin{bmatrix} -lpha & lpha \ & eta \ & eta & -eta \end{bmatrix},$$

the limiting distribution solves $\pi Q = 0$, *i.e.*

$$\left\{egin{array}{l} 0=-lpha\pi_0+eta\pi_1\ 0=lpha\pi_0-eta\pi_1, \end{array}
ight.$$

with $\pi_0 + \pi_1 = 1$, *i.e.*

$$\pi = [\pi_0, \pi_1] = \left[\frac{\beta}{\alpha + \beta}, \frac{\alpha}{\alpha + \beta}\right].$$
(9.6.5)

ii) Birth and death process on \mathbb{N} .

Next, consider the birth and death process on $\mathbb N$ with infinitesimal generator

$$Q = \left[\; \lambda_{i,j} \;
ight]_{i,j \in \mathbb{N}} = egin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 & \cdots \ \mu_1 & -\lambda_1 - \mu_1 & \lambda_1 & 0 & 0 & \cdots \ 0 & \mu_2 & -\lambda_2 - \mu_2 & \lambda_2 & 0 & \cdots \ 0 & 0 & \mu_3 & -\lambda_3 - \mu_3 & \lambda_3 & \cdots \ dots & dots &$$

the stationary distribution solves $\pi Q = 0$, *i.e.*

$$\begin{cases} 0 = -\lambda_0 \pi_0 + \mu_1 \pi_1 \\ 0 = \lambda_0 \pi_0 - (\lambda_1 + \mu_1) \pi_1 + \mu_2 \pi_2 \\ 0 = \lambda_1 \pi_1 - (\lambda_2 + \mu_2) \pi_2 + \mu_3 \pi_3 \\ \vdots \\ 0 = \lambda_{j-1} \pi_{j-1} - (\lambda_j + \mu_j) \pi_j + \mu_{j+1} \pi_{j+1}, \\ \vdots \end{cases}$$

i.e.

$$\begin{cases} \pi_1 = \frac{\lambda_0}{\mu_1} \pi_0 \\ \pi_2 = -\frac{\lambda_0}{\mu_2} \pi_0 + \frac{\lambda_1 + \mu_1}{\mu_2} \pi_1 = -\frac{\lambda_0}{\mu_2} \pi_0 + \frac{\lambda_1 + \mu_1}{\mu_2} \frac{\lambda_0}{\mu_1} \pi_0 = \frac{\lambda_1}{\mu_2} \frac{\lambda_0}{\mu_1} \pi_0 \\ \pi_3 = -\frac{\lambda_1}{\mu_3} \pi_1 + \frac{\lambda_2 + \mu_2}{\mu_3} \pi_2 = -\frac{\lambda_1}{\mu_3} \frac{\lambda_0}{\mu_1} \pi_0 + \frac{\lambda_2 + \mu_2}{\mu_3} \frac{\lambda_1}{\mu_2} \frac{\lambda_0}{\mu_1} \pi_0 = \frac{\lambda_2}{\mu_3} \frac{\lambda_1}{\mu_2} \frac{\lambda_0}{\mu_1} \pi_0 \\ \vdots \\ \pi_{j+1} = \frac{\lambda_j \cdots \lambda_0}{\mu_{j+1} \cdots \mu_1} \pi_0. \\ \vdots \end{cases}$$

Using the convention

$$\lambda_{j-1}\cdots\lambda_0 = \prod_{l=0}^{j-1}\lambda_l = 1$$
 and $\mu_j\cdots\mu_1 = \prod_{l=1}^j\mu_l = 1$

in the case j = 0, we have

$$1 = \sum_{j \ge 0} \pi_j = \pi_0 + \pi_0 \sum_{j \ge 0} \frac{\lambda_j \cdots \lambda_0}{\mu_{j+1} \cdots \mu_1} = \pi_0 + \pi_0 \sum_{j \ge 1} \frac{\lambda_{j-1} \cdots \lambda_0}{\mu_j \cdots \mu_1}$$

382

$$= \ \pi_0 \sum\limits_{i \geqslant 0} rac{\lambda_0 \cdots \lambda_{i-1}}{\mu_1 \cdots \mu_i},$$

hence

$$\pi_0 = rac{1}{\sum\limits_{i \geqslant 0} rac{\lambda_0 \cdots \lambda_{i-1}}{\mu_1 \cdots \mu_i}},$$

and

$$\pi_j = rac{\lambda_0 \cdots \lambda_{j-1}}{\mu_1 \cdots \mu_j \sum\limits_{i \geqslant 0} rac{\lambda_0 \cdots \lambda_{i-1}}{\mu_1 \cdots \mu_i}}, \qquad j \in \mathbb{N}.$$

When $\lambda_i = \lambda$, $i \in \mathbb{N}$, and $\mu_i = \mu$, $i \ge 1$, this gives

$$\pi_j = rac{\lambda^j}{\mu^j \sum\limits_{i \geqslant 0} (\lambda/\mu)^i} = \left(1 - rac{\lambda}{\mu}
ight) \left(rac{\lambda}{\mu}
ight)^j, \qquad j \in \mathbb{N}.$$

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 $\$ = \{0, 1, \dots, N\}$.

The birth and death process on $\mathbb{S} = \{0, 1, \dots, N\}$ has the infinitesimal generator

$$Q = [\lambda_{i,j}]_{0 \leqslant i,j \leqslant N} = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots & \cdots & 0 & 0 & 0 & 0 \\ \mu_1 & -\lambda_1 - \mu_1 & \lambda_1 & \cdots & \cdots & 0 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots & \vdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 & \mu_{N-1} & -\lambda_{N-1} - \mu_{N-1} & \lambda_{N-1} \\ 0 & 0 & 0 & \cdots & 0 & 0 & \mu_N & -\mu_N \end{bmatrix}$$

we can apply (1.6.6) with $\lambda_j = 0, j \ge N$, and $\mu_j = 0, j \ge N + 1$, which yields

$$\pi_j = rac{\lambda_{j-1}\cdots\lambda_0}{\mu_j\cdots\mu_1\sum\limits_{i=0}^Nrac{\lambda_{i-1}\cdots\lambda_0}{\mu_i\cdots\mu_1}}, \qquad j\in\{0,1,\ldots,N\},$$

and coincides with (9.6.5) when N = 1.

When $\lambda_i = \lambda$, $i \in \mathbb{N}$, and $\mu_i = \mu$, $i \ge 1$, this gives

$$\pi_j = rac{1-\lambda/\mu}{1-(\lambda/\mu)^{N+1}} \left(rac{\lambda}{\mu}
ight)^j, \qquad j=0,1,\ldots,N,$$

which is a truncated geometric distribution since $\pi_j = 0$ for all j > N and any $\lambda, \mu > 0$.

ॳ 384

,

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

$$T_1 = \inf\{t > 0 \; : \; X_t
eq X_0\},$$

and

$$T_{n+1}=\inf\{t>T_n\ :\ X_t
eq X_{T_n}\},\qquad n\in\mathbb{N}.$$

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

$$Z_n:=X_{T_n},\qquad n\geqslant 1.$$

The next Figure 9.8 shows the graph of the embedded chain of a birth and death process.*



Figure 9.8: Birth and death process $(X_t)_{t \in \mathbb{R}_+}$ with its embedded chain $(Z_n)_{n \in \mathbb{N}}$.

The results of Chapters 2-8 can now be applied to the discrete-time embedded chain. The next Figure 9.9 represents the discrete-time embedded chain associated to the path of Figure 9.8, in which we have $Z_0 = 0$, $Z_1 = 1$, $Z_2 = 2$, $Z_3 = 3$, $Z_4 = 4$, $Z_5 = 3$, $Z_6 = 4$, $Z_7 = 3$, ...

^{*}Download the corresponding **R** code or the **IPython notebook** that can be run *here*.



Figure 9.9: Discrete-time embedded chain $(Z_n)_{n \in \mathbb{N}}$ based on the path of Figure 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

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{1} \\ \boldsymbol{1} & \boldsymbol{0} \end{bmatrix},\tag{9.7.1}$$

which switches permanently between the states (0) and (1).

In case one of the states $\{0, 1\}$ is absorbing the transition matrix becomes

$$P = \left\{ egin{array}{cccc} 1 & 0 \ 1 & 0 \end{array}
ight], \; \lambda_0 = 0, \; \mu_1 > 0, \ \left[egin{array}{cccc} 0 & 1 \ 0 & 1 \end{array}
ight], \; \lambda_0 > 0, \; \mu_1 = 0, \ \left[egin{array}{cccc} 1 & 0 \ 0 & 1 \end{array}
ight], \; \lambda_0 = 0, \; \mu_1 = 0. \end{array}
ight.$$

MH3512 AY19-20

♦ 386

Birth and death embedded chain

More generally, consider now the birth and death process with infinitesimal generator

$$Q = [\ \lambda_{i,j} \]_{0 \leqslant i,j \leqslant N} = egin{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 & 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 & \vdots & \vdots \ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \ 0 & 0 & 0 & \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 (i + 1) is given by

$$\mathbb{P}(X_{t+h} = i+1 \mid X_t = i \text{ and } X_{t+h} - X_t \neq 0) = \frac{\mathbb{P}(X_{t+h} = i+1 \text{ and } X_t = i)}{\mathbb{P}(X_{t+h} - X_t \neq 0 \text{ and } X_t = i)}$$

$$= \frac{\mathbb{P}(X_{t+h} = i+1 \text{ and } X_t = i)}{\mathbb{P}(X_{t+h} - X_t \neq 0 \mid X_t = i)\mathbb{P}(X_t = i)}$$

$$= \frac{\mathbb{P}(X_{t+h} - X_t \neq 0 \mid X_t = i)}{\mathbb{P}(X_{t+h} - X_t \neq 0 \mid X_t = i)}$$

$$= \frac{\mathbb{P}(X_{t+h} - X_t \neq 0 \mid X_t = i)}{\mathbb{P}(X_{t+h} - X_t \neq 0 \mid X_t = i)}$$

$$\simeq \frac{\lambda_i h}{\lambda_i h + \mu_i h}$$

$$= \frac{\lambda_i}{\lambda_i + \mu_i}, \quad h \searrow 0, \quad i \in S,$$

where we applied (9.4.8), hence the transition matrix of the embedded chain satisfies

$$P_{i,i+1} = \lim_{h \searrow 0} \mathbb{P}(X_{t+h} = i+1 \mid X_t = i \text{ and } X_{t+h} - X_t \neq 0) = \frac{\lambda_i}{\lambda_i + \mu_i}, \quad i \in \mathbb{S}.$$
(9.7.2)

This result can also be obtained from (1.5.9) which states that

$$\mathbb{P}(\tau_{i,i+1} < \tau_{i,i-1}) = \frac{\lambda_i}{\lambda_i + \mu_i}.$$
(9.7.3)

Similarly the probability that a given transition occurs from (i) to (i-1) is

$$\mathbb{P}(X_{t+h} = i-1 \mid X_t = i \text{ and } X_{t+h} - X_t \neq 0) = \frac{\mu_i}{\lambda_i + \mu_i}, \quad h \searrow 0, \quad i \in \mathbb{S},$$

which can also be obtained from (1.5.9) which states that

$$\mathbb{P}(au_{i,i-1} < au_{i,i+1}) = rac{\mu_i}{\lambda_i + \mu_i}.$$

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

$$P = [P_{i,j}]_{i,j \in \mathbb{S}}$$

(388

$$= \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \frac{\mu_1}{\lambda_1 + \mu_1} & 0 & \frac{\lambda_1}{\lambda_1 + \mu_1} & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots & \vdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & \frac{\mu_{N-1}}{\lambda_{N-1} + \mu_{N-1}} & 0 & \frac{\lambda_{N-1}}{\lambda_{N-1} + \mu_{N-1}} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \end{bmatrix},$$

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 (1) and (N) are both absorbing since the birth rate starting from (1) and the death rate starting from (N) are both 0, hence the transition matrix of the embedded chain can be written as

$$P = \begin{bmatrix} P_{i,j} \end{bmatrix}_{0 \leqslant i,j \leqslant N} \\ = \begin{bmatrix} \frac{1}{\lambda_1 + \mu_1} & 0 & \frac{\lambda_1}{\lambda_1 + \mu_1} & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots & \vdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \frac{\mu_{N-1}}{\lambda_{N-1} + \mu_{N-1}} & 0 & \frac{\lambda_{N-1}}{\lambda_{N-1} + \mu_{N-1}} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \end{bmatrix},$$

389

which is the transition matrix of a gambling type process on $\{0, 1, ..., N\}$. When N = 1 this yields $P = I_d$, 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

$$Q = \left[egin{array}{cccc} \lambda_{i,j} \end{array}
ight]_{0\leqslant i,j\leqslant 4} = egin{bmatrix} -10 & 10 & 0 & 0 & 0 \ 10 & -20 & 10 & 0 & 0 \ 0 & 10 & -30 & 20 & 0 \ 0 & 0 & 10 & -40 & 30 \ 0 & 0 & 0 & 20 & -20 \end{bmatrix},$$

the transition matrix of the embedded chain is

$$P = \left[\begin{array}{cccc} P_{i,j} \end{array}
ight]_{0\leqslant i,j\leqslant 4} = \left[egin{array}{cccc} 0 & 1 & 0 & 0 & 0 \ 1/2 & 0 & 1/2 & 0 & 0 \ 0 & 1/3 & 0 & 2/3 & 0 \ 0 & 0 & 1/4 & 0 & 3/4 \ 0 & 0 & 0 & 1 & 0 \end{array}
ight].$$

In case the states 0 and 4 are absorbing, *i.e.*

$$Q = \left[egin{array}{cccc} \lambda_{i,j} \end{array}
ight]_{0 \leqslant i,j \leqslant 4} = \left[egin{array}{ccccc} 0 & 0 & 0 & 0 \ 10 & -20 & 10 & 0 & 0 \ 0 & 10 & -30 & 20 & 0 \ 0 & 0 & 10 & -40 & 30 \ 0 & 0 & 0 & 0 & 0 \end{array}
ight],$$

the transition matrix of the embedded chain becomes

$$P = \left[egin{array}{cccc} P_{i,j} \end{array}
ight]_{0\leqslant i,j\leqslant 4} = \left[egin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \ 1/2 & 0 & 1/2 & 0 & 0 \ 0 & 1/3 & 0 & 2/3 & 0 \ 0 & 0 & 1/4 & 0 & 3/4 \ 0 & 0 & 0 & 0 & 1 \end{array}
ight].$$

More generally, by (9.4.8) and (9.4.3) we could also show that the embedded chain of a continuous-time Markov chain with generator

Q of the form (9.4.2) has the transition matrix

$$P = [P_{i,j}]_{i,j\in\mathbb{S}}$$

$$= \begin{bmatrix} 0 & -\frac{\lambda_{0,1}}{\lambda_{0,0}} & -\frac{\lambda_{0,2}}{\lambda_{0,0}} & \cdots & \cdots & -\frac{\lambda_{0,N-1}}{\lambda_{0,0}} & -\frac{\lambda_{0,N}}{\lambda_{0,0}} \\ -\frac{\lambda_{1,0}}{\lambda_{1,1}} & 0 & -\frac{\lambda_{1,2}}{\lambda_{1,1}} & \cdots & \cdots & -\frac{\lambda_{1,N-1}}{\lambda_{1,1}} & -\frac{\lambda_{1,N}}{\lambda_{1,1}} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -\frac{\lambda_{N-1,0}}{\lambda_{N-1,N-1}} & -\frac{\lambda_{N-1,1}}{\lambda_{N-1,N-1}} & \cdots & \cdots & 0 & -\frac{\lambda_{N-1,N}}{\lambda_{N-1,N-1}} \\ -\frac{\lambda_{N,0}}{\lambda_{N,N}} & -\frac{\lambda_{N,1}}{\lambda_{N,N}} & \cdots & \cdots & -\frac{\lambda_{N,N-1}}{\lambda_{N,N}} & 0 \end{bmatrix},$$

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 (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 (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 (1) starting from state $k \in \{0, 1, \ldots, N\}$. We have the boundary condition $g_0(0)_{\mathbb{H}_{3512}} \mathbf{1}_{\mathbb{H}_{3512}} \mathbf{1}_{\mathbb{$

by first step analysis on the chain $(Z_n)_{n \ge 1}$ we get

$$g_0(k)=rac{\lambda_k}{\lambda_k+\mu_k}g_0(k+1)+rac{\mu_k}{\lambda_k+\mu_k}g_0(k-1),\qquad k=1,2,\ldots,N-1.$$

When the rates $\lambda_k = \lambda$ and $\mu_k = \mu$ are independent of $k \in \{1, 2, \dots, N-1\}$, this equation becomes

$$g_0(k) = pg_0(k+1) + qg_0(k-1), \qquad k = 1, 2, \dots, N-1,$$

which is precisely Equation (2.2.6) for the gambling process with

$$p = rac{\lambda}{\lambda + \mu} \quad ext{and} \quad q = rac{\mu}{\lambda + \mu}$$

When $\lambda_0 = \mu_N = 0$ we have the boundary conditions

$$g_0(0) = 1$$
 and $g_0(N) = 0$

since the state (N) becomes absorbing, and the solution becomes

$$g_0(k)=rac{(\mu/\lambda)^k-(\mu/\lambda)^N}{1-(\mu/\lambda)^N},\qquad k=0,1,\ldots,N,$$

when $\lambda \neq \mu$, according to (2.2.11). When $\lambda \neq \mu$, Relation (2.2.12) shows that

$$g_0(k)=rac{N-k}{N}=1-rac{k}{N},\qquad k=0,1,\ldots,N.$$

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 mean time spent at state (i), given that the next transition is from (i) to (i+1), is equal to

$$\mathbb{E}[au_{i,i+1}] = rac{1}{\lambda_i}, \qquad i=1,2,\ldots,N-1,$$

and the mean time spent at state (i), given that the next transition is from (i) to (i-1), is equal to

$$\mathbb{E}[au_{i,i-1}]=rac{1}{\mu_i}, \qquad i=1,2,\ldots,N-1.$$

We associate a *weighted* graph to the Markov chain $(Z_n)_{n \in \mathbb{N}}$ that includes the average

$$\mathbb{E}[au_i] = rac{1}{\lambda_i + \mu_i}$$

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, ..., N - 1. In the next graph, which is drawn for N = 4, the weights are underlined:

$$\begin{array}{c} \mu_1/(\lambda_1 + \mu_1) \quad \lambda_1/(\lambda_1 + \mu_1) \quad \mu_3/(\lambda_3 + \mu_3) \quad \lambda_3/(\lambda_3 + \mu_3) \\ 0 \\ \hline 1/(\lambda_1 + \mu_1) \\ 0 \\ \mu_1/(\lambda_2 + \mu_2) \\ \mu_2/(\lambda_2 + \mu_2) \\ \mu_2/(\lambda_2 + \mu_2) \\ \end{array}$$

with $\lambda_0 = \mu_4 = 0$.

Proposition 9.10

The mean absorption times

$$h_{0,N}(i)=\mathbb{E}\left[T_{0,N}\mid X_0=i
ight],\qquad i=0,1,\ldots,N,$$

into states $\{0, N\}$ starting from state $i \in \{0, 1, ..., N\}$ satisfy the boundary conditions $h_{0,N}(0) = h_{0,N}(N) = 0$ the first step analysis equation

$$h_{0,N}(i)=rac{1}{\lambda_i+\mu_i}+rac{\mu_i}{\lambda_i+\mu_i}h_{0,N}(i-1)+rac{\lambda_i}{\lambda_i+\mu_i}h_{0,N}(i+1),$$

 $i=1,2,\ldots,N-1$.

Proof

By first step analysis on the discrete-time embedded chain $(Z_n)_{n \ge 1}$ with transition matrix

$$P = [P_{i,j}]_{i,j\in\mathbb{S}} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \frac{\mu_1}{\lambda_1 + \mu_1} & 0 & \frac{\lambda_1}{\lambda_1 + \mu_1} & 0 & \cdots & 0 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots & \vdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \frac{\lambda_{N-1}}{\lambda_{N-1} + \mu_{N-1}} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

we have

$$h_{0,N}(i) = \frac{\mu_i}{\lambda_i + \mu_i} \left(\mathbb{E}[\tau_i] + h_{0,N}(i-1) \right) + \frac{\lambda_i}{\lambda_i + \mu_i} \left(\mathbb{E}[\tau_i] + h_{0,N}(i+1) \right)$$

$$= \frac{\mu_i}{\lambda_i + \mu_i} \left(\frac{1}{\lambda_i + \mu_i} + h_{0,N}(i-1) \right) + \frac{\lambda_i}{\lambda_i + \mu_i} \left(\frac{1}{\lambda_i + \mu_i} + h_{0,N}(i+1) \right),$$

$$(9.8.1)$$

$$i = 1, 2, \dots, N-1.$$

Note that by conditioning on the independent exponential random variables $au_{i,i-1}$ and $au_{i,i+1}$ we can also show that

$$\mathbb{E}[\tau_i \mid \tau_{i,i-1} < \tau_{i,i+1}] = \mathbb{E}[\tau_i \mid \tau_{i,i+1} < \tau_{i,i-1}] = \mathbb{E}[\tau_i] = \frac{1}{\lambda_i + \mu_i},$$

 $i = 1, 2, \dots, N - 1$, cf. (1.7.9) in Exercise 1.4-(a), hence (9.8.1) can be rewritten as

$$h_{0,N}(i) \;=\; rac{\mu_i}{\lambda_i+\mu_i} \left(\mathbb{E}[au_i \mid au_{i,i-1} < au_{i,i+1}] + h_{0,N}(i-1)
ight)$$
 MH3512 AY19-20

<u>♂</u>394

$$+\frac{\lambda_{i}}{\lambda_{i}+\mu_{i}}\left(\mathbb{E}[\tau_{i} \mid \tau_{i,i+1} < \tau_{i,i-1}] + h_{0,N}(i+1)\right).$$
(5.395)

When the rates $\lambda_i = \lambda$ and $\mu_i = \mu$ are independent of $i \in \{1, 2, \dots, N-1\}$, this equation becomes

$$h_{0,N}(i)=rac{1}{\lambda+\mu}+rac{\lambda}{\lambda+\mu}h_{0,N}(i+1)+rac{\mu}{\lambda+\mu}h_{0,N}(i-1),$$

i = 1, 2, ..., N - 1, which is a modification of Equation (2.3.6), by replacing the discrete time step by the average time $1/(\lambda + \mu)$ spent at any state. Rewriting the equation as

$$h_{0,N}(i) = rac{1}{\lambda + \mu} + p h_{0,N}(i+1) + q h_{0,N}(i-1),$$

$$i=1,2,\ldots,N-1,$$
 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

$$p=rac{\lambda}{\lambda+\mu} \quad ext{and} \quad q=rac{\mu}{\lambda+\mu},$$

we find from (2.3.11) that, with $r = q/p = \mu/\lambda$,

$$(\lambda+\mu)h_{0,N}(k)=rac{1}{q-p}\left(k-Nrac{1-r^k}{1-r^N}
ight),$$

i.e.

$$h_{0,N}(k) = \frac{1}{\mu - \lambda} \left(k - N \frac{1 - (\mu/\lambda)^k}{1 - (\mu/\lambda)^N} \right), \qquad k = 0, 1, \dots, N,$$
(9.8.2)

when $\lambda \neq \mu$. In the limit $\lambda \rightarrow \mu$ we find by (2.3.17) that

$$h_{0,N}(k) = rac{1}{2\mu} k(N-k), \qquad k=0,1,\dots,N.$$

This solution is similar to that of the gambling problem with draw Exercise 2.1 as we multiply the solution of the gambling problem $\frac{396}{100}$ 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

$$\frac{\lambda+\mu}{\mu-\lambda}\left(k-N\frac{1-(\mu/\lambda)^k}{1-(\mu/\lambda)^N}\right), \qquad k=0,1,\ldots,N,$$
(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.

How to compute	Method
the infinitesimal generator $Q = (\lambda_{i,j})_{i,j \in S}$	$Q=rac{dP(t)}{dt}_{ert t=0}=P^{\prime}(0).$
the semigroup $(P(t))_{t\in\mathbb{R}_+}$	$egin{aligned} P(t) &= \exp(tQ), t \in \mathbb{R}_+, \ P(h) &= \mathrm{I_d} + hQ + o(h), h \searrow 0. \end{aligned}$
the stationary distribution π	$\mathrm{solve}^*\pi Q=0 ext{ for } \pi.$
the probability distribution of the time $ au_{i,j}$ spent in $i o j$	exponential distribution $(\lambda_{i,j})$.
the probability distribution of the time $ au_i$ spent at state (i)	exponential distribution $\left(\sum_{l\neq i} \lambda_{i,l}\right)$.
$\lim_{t o\infty} \exp\left(t\left[egin{array}{cc} -lpha & lpha\ & & \ eta & -eta\end{array} ight) ight)$	$\begin{bmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{bmatrix}$
hitting probabilities	solve [†] $g = Pg$ for the embedded chain.
mean hitting times	use the embedded chain with weighted links using mean inter-jump times.

Table 9.1: Summary of computing methods.

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 $t \ge 0$.

^{*}Remember that the values of $\pi(k)$ have to add up to 1.

 $^{^\}dagger\mathrm{Be}$ sure to write only the relevant rows of the system under the appropriate boundary conditions.

- a) Explain why $(X_t)_{t \in \mathbb{R}_+}$ is a continuous-time Markov chain with state space $S = \{0, 1, 2, 3\}$, and give its matrix infinites maps 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; \$10.02; \$10.03; \$10.04; \$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

$$\lambda_k = (1+k)^
ho, \qquad k \geqslant 1,$$

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.*

^{*}Recall that a finite-valued random variable may have an infinite mean.

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 $(\tau_0, \tau_1) := (T_1, T_2 - T_1)$.

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. MH3512 AY19-20 a) Find the probability that her waiting time is **0**.

ॳ401

- 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

$$Q = \left[egin{array}{cccc} \square & 0.5 & 0 \ 0.2 & \square & \square \ 0 & \square & -0.4 \end{array}
ight]$$

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 sytem produce per hour in the long run 20

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

$$Q = \left[egin{array}{ccccc} -0.5 & \Box & \Box \ \Box & -0.7 & \Box \ \Box & \Box & -0.4 \end{array}
ight.$$

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

$$egin{bmatrix} -\lambda & \lambda \ \mu & -\mu \end{bmatrix}.$$

Argue that $Z(t) := X_1(t) + X_2(t)$ is a Markov chain on the state space $S = \{0, 1, 2\}$ and determine the transition semigroup P(t) of Z(t).

Exercise 9.17 Consider a two-state discrete-time Markov chain $(\xi_n)_{n \ge 0}$ on $\{0, 1\}$ with transition matrix

$$\begin{bmatrix} 0 & 1 \\ 1 - \alpha & \alpha \end{bmatrix}.$$
(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

$$\lim_{t o\infty} \mathbb{P}(|M_t|\leqslant c)$$

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 $\{0, 1, \dots, N\}$ with transition semigroup $(P(t))_{t \in \mathbb{R}}$ and birth and death rates

$$\lambda_n = (N-n)\lambda, \qquad \mu_n = n\mu, \qquad n = 0, 1, \dots, N.$$
 MH3512 AY19

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

$$\left\{egin{aligned} P_{n,0}'(t) &= -\lambda_0 P_{n,0}(t) + \mu_1 P_{n,1}(t), \ P_{n,k}'(t) &= \lambda_{k-1} P_{n,k-1}(t) - (\lambda_k + \mu_k) P_{n,k}(t) + \mu_{k+1} P_{n,k+1}(t), \ P_{n,N}'(t) &= \lambda_{N-1} P_{n,N-1}(t) - \mu_N P_{n,N}(t), \end{aligned}
ight.$$

$$k=1,2,\ldots,N-1.$$

c) Let

$$G_k(s,t) = \mathbb{E}\left[s^{X_t} \mid X_0 = k
ight] = \sum\limits_{n=0}^N s^n \mathbb{P}(X_t = n \mid X_0 = k) = \sum\limits_{n=0}^N s^n P_{k,n}(t)$$

denote the generating function of X_t given that $X_0 = k \in \{0, 1, ..., N\}$. From the result of Question (b), show that $G_k(s, t)$ satisfies the partial differential equation

$$\frac{\partial G_k}{\partial t}(s,t) = \lambda N(s-1)G_k(s,t) + (\mu + (\lambda - \mu)s - \lambda s^2)\frac{\partial G_k}{\partial s}(s,t), \qquad (9.8.5)$$

with $G_k(s,0) = s^k, \, k = 0, 1, \dots, N.$

d) Verify that the solution of (9.8.5) is given by

$$G_k(s,t)=rac{1}{(\lambda+\mu)^N}(\mu+\lambda s+\mu(s-1)\mathrm{e}^{-(\lambda+\mu)t})^k(\mu+\lambda s-\lambda(s-1)\mathrm{e}^{-(\lambda+\mu)t})^{N-k},$$

 $k=0,1,\ldots,N$.

e) Show that

$$\mathbb{E}[X_t \mid X_0 = k] = \frac{k}{(\lambda + \mu)^N} (\lambda + \mu \mathrm{e}^{-(\lambda + \mu)t}) (\mu + \lambda)^{k-1} (\mu + \lambda)^{N-k}$$
MH3512 AY19-20

$$+\frac{N-k}{(\lambda+\mu)^{N}}(\mu+\lambda)^{k}(\lambda-\lambda e^{-(\lambda+\mu)t})(\mu+\lambda)^{N-k-1}.$$

f) Compute

$$\lim_{t o\infty} \mathbb{E}[X_t \mid X_0 = k]$$

and show that it does not depend on $k \in \{0, 1, \ldots, N\}$.

Problem 9.20 Let T_1, T_2, \ldots be the first two jump times of a Poisson process $(N_t)_{t \in \mathbb{R}_+}$ with intensity λ . Given $f : \mathbb{R} \longrightarrow \mathbb{R}$ an integrable function, show that

$$\mathbb{E}\left[\sum_{k=1}^{N_t} f(T_i)\right] = \lambda \int_0^t f(s) ds.$$
(9.8.6)

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 Chapter 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	406
10.2 Martingales - Definition and Properties	408
10.3 Stopping Times	410
10.4 Ruin Probabilities	415
10.5 Mean Game Duration	419
$\operatorname{Exercises} \ldots \ldots$	423

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 <u>filtration</u> $(\mathcal{F}_n)_{n \in \mathbb{N}}$ generated by a stochastic process $(X_n)_{n \in \mathbb{N}}$ taking its values in a state space S, is the family of σ -algebras

$$\mathcal{F}_n:=\sigma(X_0,X_1,\ldots,X_n),\qquad n\geqslant 0,$$

which denote the collections of events generated by X_0, X_1, \ldots, X_n . Examples of such events include

$$\{X_0 \leqslant a_0, \; X_1 \leqslant a_1, \; \ldots, \; X_n \leqslant a_n\}$$
 MH3512 AY19-20

for a_0, a_1, \ldots, a_n a given fixed sequence of real numbers. Note that we have the inclusion $\mathcal{F}_n \subset \mathcal{F}_{n+1}, n \in \mathbb{N}$.

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

$$S_n := X_1 + X_2 + \dots + X_n, \qquad n \in \mathbb{N},$$

where $(X_k)_{k \ge 1}$ 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 $\mathcal{F}_0 = \{\emptyset, \Omega\}$,

$$\mathcal{F}_1 = ig \{ \emptyset, \{ X_1 = 1 \}, \{ X_1 = -1 \}, \Omega ig \},$$

and

$${\mathcal F}_2=\sigmaig(\{\emptyset,\{X_1=1,X_2=1\},\{X_1=1,X_2=-1\},\{X_1=-1,X_2=1\},\ \{X_1=-1,X_2=-1\},\Omegaig).$$

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.*

$$egin{aligned} \mathcal{G}_2 &:= ig\{ \emptyset, \{X_1 = 1, X_2 = -1\} \cup \{X_1 = -1, X_2 = 1\}, \ &\{X_1 = 1, X_2 = 1\} \cup \{X_1 = -1, X_2 = -1\}, \Omega ig\}, \end{aligned}$$

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 Section 1.6. Given F a random variable with finite mean the conditional expectation $\mathbb{E}[F \mid \mathcal{F}_n]$ refers to

$$\mathbb{E}[F \mid X_0, X_1, \dots, X_n] = \mathbb{E}[F \mid X_0 = k_0, \dots, X_n = k_n]_{k_0 = X_0, \dots, k_n = X_n},$$

given that X_0, X_1, \ldots, X_n are respectively equal to $k_0, k_1, \ldots, k_n \in S$.

MH3512 AY19-20

(5407)

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.* of 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 .

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,^{*a*} discrete-time stochastic process $(Z_n)_{n \in \mathbb{N}}$ is a <u>martingale</u> 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 $\mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] = Z_n, \quad n \in \mathbb{N}.$ (10.2.1) ^{*a*}Integrable means $\mathbb{E}[|Z_n|] < \infty$ for all $n \in \mathbb{N}$.

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

$$\mathbb{E}[M_n \mid \mathcal{F}_k] = M_k, \qquad 0 \leqslant k < n.$$

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

 $\mathbb{E}[Z_n] = \mathbb{E}[Z_0], \qquad n \in \mathbb{N}.$

From the tower property (1.6.8) 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

$$\mathbb{E}[Z_{n+1}] = \mathbb{E}[Z_n] = \mathbb{E}[Z_{n-1}] = \cdots = \mathbb{E}[Z_1] = \mathbb{E}[Z_0], \quad n \in \mathbb{N}.$$

Examples of martingales

1. Any centered^{*} 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

$$egin{array}{rll} \mathbb{E}[S_{n+1} \mid \mathcal{F}_n] &=& \mathbb{E}[S_n \mid \mathcal{F}_n] + \mathbb{E}\left[S_{n+1} - S_n \mid \mathcal{F}_n
ight] \ &=& \mathbb{E}[S_n \mid \mathcal{F}_n] + \mathbb{E}[S_{n+1} - S_n] \ &=& \mathbb{E}[S_n \mid \mathcal{F}_n] = S_n, \qquad n \in \mathbb{N}. \end{array}$$

In addition to being a martingale, a stochastic process $(X_n)_{n \in \mathbb{N}}$ with centered independent increments is also a Markov process, cf. Section 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:

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[\mathbb{E}[F \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n] = \mathbb{E}[F \mid \mathcal{F}_n] = X_n, \quad n \in \mathbb{N},$$
(10.2.2)

^{*}A random variable X_n is said to be <u>centered</u> if $\mathbb{E}[X_n] = 0$.

from the tower property (1.6.8).

The following Figure 10.1 illustrates various estimates $X_n = \mathbb{E}[F \mid \mathcal{F}_n]$ at time n = "Wed", "Thu", "Fri", "Sat", for a random outcome F known at time "Sat", *i.e.* $X_{Wed} = 26$, $X_{Thu} = 28$, $X_{Fri} = 26$, $X_{Sat} = 24$.



Figure 10.1: Updated weather forecast.

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.



A stopping time is a random variable $\tau: \Omega \longrightarrow \mathbb{N}$ such that

 $\{ au>n\}\in\mathcal{F}_n,\qquad n\in\mathbb{N}.$

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

$$\{ au\leqslant n\}\in\mathcal{F}_n,\qquad n\in\mathbb{N},$$

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.



On the other hand, the first time

$$T = \inf \left\{ k \geqslant 0 \; : \; X_k = \max_{l=0,1,...,N} X_l
ight\}$$

MH3512 AY19-20

(10.3.1)

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 <u>stopped process</u> $(Z_{\tau \wedge n})_{n \in \mathbb{N}} = (Z_{\min(\tau,n)})_{n \in \mathbb{N}}$ is defined as $Z_{\tau \wedge n} = Z_{\min(\tau,n)} = \begin{cases} Z_n & \text{if } n < \tau, \\ Z_{\tau} & \text{if } n \geqslant \tau, \end{cases}$

Using indicator functions we may also write

$$Z_{ au \wedge n} = Z_n \mathbb{1}_{\{n < au\}} + Z_ au \mathbb{1}_{\{n \geqslant au\}}, \qquad n \in \mathbb{N}.$$

The following Figure 10.2 is an illustration of the path of a 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 $(\mathcal{F}_n)_{n \in \mathbb{N}}$.



$$M_n = M_0 + \sum_{l=1}^n (M_l - M_{l-1}) = M_0 + \sum_{l \geqslant 1} \mathbb{1}_{\{l \leqslant n\}} (M_l - M_{l-1}),$$

we have

$$M_{ au \wedge n} = M_0 + \sum_{l=1}^{ au \wedge n} (M_l - M_{l-1}) = M_0 + \sum_{l=1}^n \mathbb{1}_{\{l \leqslant au\}} (M_l - M_{l-1}),$$

5413

and for $k \leqslant n$ we find

$$\begin{split} \mathbb{E} \left[M_{\tau \wedge n} \left| \mathcal{F}_k \right] &= M_0 + \sum_{l=1}^n \mathbb{E} \left[\mathbbm{1}_{\{l \leqslant \tau\}} (M_l - M_{l-1}) \left| \mathcal{F}_k \right] \right] \\ &= M_0 + \sum_{l=1}^k \mathbb{E} \left[\mathbbm{1}_{\{l \leqslant \tau\}} (M_l - M_{l-1}) \left| \mathcal{F}_k \right] + \sum_{l=k+1}^n \mathbb{E} \left[\mathbbm{1}_{\{l \leqslant \tau\}} (M_l - M_{l-1}) \right| \mathcal{F}_k \right] \\ &= M_0 + \sum_{l=1}^k (M_l - M_{l-1}) \mathbb{E} \left[\mathbbm{1}_{\{l \leqslant \tau\}} \left| \mathcal{F}_k \right] \right] \\ &+ \sum_{l=k+1}^n \mathbb{E} \left[\mathbbm{1} \left[(M_l - M_{l-1}) \mathbbm{1}_{\{l \leqslant \tau\}} \right] \mathcal{F}_{l-1} \right] \left| \mathcal{F}_k \right] \\ &= M_0 + \sum_{l=1}^k (M_l - M_{l-1}) \mathbbm{1}_{\{l \leqslant \tau\}} \\ &+ \sum_{l=k+1}^n \mathbbm{1} \mathbb{E} \left[\mathbbm{1}_{\{l-1 < \tau\}} \underbrace{\mathbbm{1}_{\{l < T\}}}_{=0} \right] \left| \mathcal{F}_k \right] \\ &= M_0 + \sum_{l=1}^{n} (M_l - M_{l-1}) \\ &= M_0 + \sum_{l=1}^{n} (M_l - M_{l-1}) \\ &= M_{\tau \wedge k}, \end{split}$$

 $k = 0, 1, \ldots, n$, where we used the tower property and the fact that

$$\{ au \geqslant l \} = \{ au > l-1 \} \in \mathcal{F}_{l-1} \subset \mathcal{F}_l \subset \mathcal{F}_k, \qquad 1 \leqslant l \leqslant k.$$

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.* $\tau \leq N$, we have

$$\mathbb{E}[M_{\tau}] = \mathbb{E}[M_{\tau \wedge N}] = \mathbb{E}[M_{\tau \wedge 0}] = \mathbb{E}[M_0].$$
(10.3.2)

As a consequence of (10.3.2), if $(M_n)_{n \in \mathbb{N}}$ is a martingale and $\tau \leq N$ and $\nu \leq N$ are two bounded stopping times bounded by a constant N > 0, we have

$$\mathbb{E}[M_{ au}] = \mathbb{E}[M_{
u}] = \mathbb{E}[M_0].$$
_{MH35}(10,3,3)₀

In case τ is only *a.s.* finite, *i.e.* $\mathbb{P}(\tau < \infty) = 1$, we may also write

$$\mathbb{E}[M_{ au}] = \mathbb{E}\left[\lim_{n o \infty} M_{ au \wedge n}
ight] = \lim_{n o \infty} \mathbb{E}[M_{ au \wedge n}] = \mathbb{E}[M_0],$$

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.* $|M_{\tau \wedge n}| \leq K$ a.s., $n \in \mathbb{N}$, for some constant K > 0.

Analog statements can be proved for submartingales, 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 Chapter 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 $\{0, 1, \dots, B\}$ with independent $\{-1, 1\}$ -valued increments, and

$$\mathbb{P}(S_{n+1} - S_n = +1) = p$$
 and $\mathbb{P}(S_{n+1} - S_n = -1) = q$, $n \in \mathbb{N}$,

as introduced in Section 2.1. Let

$$T_{0,B}:\Omega\longrightarrow\mathbb{N}$$

be the first hitting time of the boundary $\{0, B\}$, defined by

$$au := T_{0,B} := \inf\{n \geqslant 0 \; : \; S_n = B ext{ or } S_n = 0\}.$$

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

$$\{ au > n\} = \{0 < S_0 < B\} \cap \{0 < S_1 < B\} \cap \dots \cap \{0 < S_n < B\}$$

hence $\boldsymbol{\tau}$ is a stopping time.

We will recover the ruin probabilities

$$\mathbb{P}(S_ au=0\mid S_0=k), \qquad k=0,1,\ldots,B,$$

computed in Chapter 2 in three steps, first in the unbiased case p = q = 1/2 (note that the hitting time τ can be shown to be shown t

finite, cf. e.g. the identity (2.2.29)).

ॳ16

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:

$$\mathbb{E}[S_{n+1} - S_n] = 1 \times p + (-1) \times q = 1 \times \frac{1}{2} + (-1) \times \frac{1}{2} = 0, \qquad (10.4.1)$$

hence $(S_n)_{n \in \mathbb{N}}$ is a martingale by Point 1 page 409.

<u>Step 2</u>. 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

$$k = \mathbb{E}[S_0 \mid S_0 = k] = \mathbb{E}[S_{ au \wedge n} \mid S_0 = k], \qquad k = 0, 1, \dots, B.$$

Letting \boldsymbol{n} go to infinity we get

$$\mathbb{E}[S_{ au} \mid S_0 = k] \; = \; \mathbb{E}\left[\lim_{n o \infty} S_{ au \wedge n} \mid S_0 = k
ight] \ = \; \lim_{n o \infty} \mathbb{E}[S_{ au \wedge n} \mid S_0 = k] = k,$$

where the exchange between limit and expectation is justified by the boundedness $|S_{\tau \wedge n}| \leq B$ a.s., $n \in \mathbb{N}$. Hence we have

$$\left\{egin{array}{l} 0 imes \mathbb{P}(S_{ au}=0\mid S_{0}=k)+B imes \mathbb{P}(S_{ au}=B\mid S_{0}=k)=\mathbb{E}[S_{ au}\mid S_{0}=k]=k\ \mathbb{P}(S_{ au}=0\mid S_{0}=k)+\mathbb{P}(S_{ au}=B\mid S_{0}=k)=1, \end{array}
ight.$$

which shows that

$$\mathbb{P}(S_{\tau} = B \mid S_0 = k) = \frac{k}{B}$$
 and $\mathbb{P}(S_{\tau} = 0 \mid S_0 = k) = 1 - \frac{k}{B}$,

 $k = 0, 1, \ldots, B$, 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 417. Theorem 10.6 we need to construct a martingale of a different type. Here we note that the process

$$M_n:=\left(rac{q}{p}
ight)^{S_n},\qquad n\in\mathbb{N},$$

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

$$\begin{split} \mathbb{E}[M_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}\left[\left(\frac{q}{p}\right)^{S_{n+1}} \middle| \mathcal{F}_n\right] = \mathbb{E}\left[\left(\frac{q}{p}\right)^{S_{n+1}-S_n} \left(\frac{q}{p}\right)^{S_n} \middle| \mathcal{F}_n\right] \\ &= \left(\frac{q}{p}\right)^{S_n} \mathbb{E}\left[\left(\frac{q}{p}\right)^{S_{n+1}-S_n} \middle| \mathcal{F}_n\right] \\ &= \left(\frac{q}{p}\right)^{S_n} \mathbb{E}\left[\left(\frac{q}{p}\right)^{S_{n+1}-S_n}\right] \\ &= \left(\frac{q}{p}\right)^{S_n} \left(\frac{q}{p} \mathbb{P}(S_{n+1}-S_n=1) + \left(\frac{q}{p}\right)^{-1} \mathbb{P}(S_{n+1}-S_n=-1)\right) \\ &= \left(\frac{q}{p}\right)^{S_n} \left(p\frac{q}{p} + q\left(\frac{q}{p}\right)^{-1}\right) \\ &= \left(\frac{q}{p}\right)^{S_n} (q+p) = \left(\frac{q}{p}\right)^{S_n} = M_n, \end{split}$$

 $n \in \mathbb{N}$. 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

$$\left(rac{q}{p}
ight)^k = \mathbb{E}[M_0 \mid S_0 = k] = \mathbb{E}[M_n \mid S_0 = k], \qquad k = 0, 1, \dots, B, \quad n \in \mathbb{N}.$$

<u>Step 2</u>. 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

$$\left(rac{q}{p}
ight)^k = \mathbb{E}[M_0 \mid S_0 = k] = \mathbb{E}[M_{ au \wedge n} \mid S_0 = k].$$

Next, letting \boldsymbol{n} go to infinity we find

$$egin{array}{ll} \left(egin{array}{c} q \ p \end{array}
ight)^k &= & \mathbb{E}[M_0 \mid S_0 = k] = \lim_{n o \infty} \mathbb{E}[M_{ au \wedge n} \mid S_0 = k] \ &= & \mathbb{E}\left[\lim_{n o \infty} M_{ au \wedge n} \mid S_0 = k
ight] \ &= & \mathbb{E}[M_ au \mid S_0 = k], \end{array}$$

hence

$$egin{aligned} &\left(rac{q}{p}
ight)^k \;=\; \mathbb{E}[M_{ au} \mid S_0 = k] \ &=\; \left(rac{q}{p}
ight)^B \mathbb{P}\left(M_{ au} = \left(rac{q}{p}
ight)^B \; \Big| \; S_0 = k
ight) + \left(rac{q}{p}
ight)^0 \mathbb{P}\left(M_{ au} = \left(rac{q}{p}
ight)^0 \; \Big| \; S_0 = k
ight) \ &=\; \left(rac{q}{p}
ight)^B \mathbb{P}\left(M_{ au} = \left(rac{q}{p}
ight)^B \; \Big| \; S_0 = k
ight) + \mathbb{P}(M_{ au} = 1 \mid S_0 = k). \end{aligned}$$

Solving the system of equations

$$\left\{ egin{array}{l} \left(rac{q}{p}
ight)^k = \left(rac{q}{p}
ight)^B \mathbb{P}\left(M_ au = \left(rac{q}{p}
ight)^B \ \middle| \ oldsymbol{S}_0 = k
ight) + \mathbb{P}(M_ au = 1 \mid oldsymbol{S}_0 = k) \ \\ \mathbb{P}\left(M_ au = \left(rac{q}{p}
ight)^B \ \middle| \ oldsymbol{S}_0 = k
ight) + \mathbb{P}(M_ au = 1 \mid oldsymbol{S}_0 = k) = 1, \end{array}
ight.$$

ර 419

$$\mathbb{P}(S_{\tau} = B \mid S_0 = k) = \mathbb{P}\left(M_{\tau} = \left(\frac{q}{p}\right)^B \mid S_0 = k\right)$$

$$= \frac{(q/p)^k - 1}{(q/p)^B - 1}, \quad k = 0, 1, \dots, B,$$
(10.4.2)

and

$$egin{aligned} \mathbb{P}(S_{ au} = 0 \mid S_0 = k) &= \mathbb{P}(M_{ au} = 1 \mid S_0 = k) \ &= 1 - rac{(q/p)^k - 1}{(q/p)^B - 1}, \ &= rac{(q/p)^B - (q/p)^k}{(q/p)^B - 1}, \end{aligned}$$

 $k = 0, 1, \dots, B$, 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 Section 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.

<u>Step 1</u>. The process $(S_n^2 - n)_{n \in \mathbb{N}}$ is a martingale.

We have

$$\begin{split} \mathbb{E}[S_{n+1}^2 - (n+1) \mid \mathcal{F}_n] &= \mathbb{E}[(S_n + S_{n+1} - S_n)^2 - (n+1) \mid \mathcal{F}_n] \\ &= \mathbb{E}[S_n^2 + (S_{n+1} - S_n)^2 + 2S_n(S_{n+1} - S_n) - (n+1) \mid \mathcal{F}_n] \\ &= \mathbb{E}[S_n^2 - n - 1 \mid \mathcal{F}_n] + \mathbb{E}[(S_{n+1} - S_n)^2 \mid \mathcal{F}_n] + 2\mathbb{E}[S_n(S_{n+1} - S_n) \mid \mathcal{F}_n] \\ &= S_n^2 - n - 1 + \mathbb{E}[(S_{n+1} - S_n)^2 \mid \mathcal{F}_n] + 2S_n \mathbb{E}[S_{n+1} - S_n \mid \mathcal{F}_n] \end{split}$$

$$= S_n^2 - n - 1 + \mathbb{E}[(S_{n+1} - S_n)^2] + 2S_n \mathbb{E}[S_{n+1} - S_n]$$

= $S_n^2 - n - 1 + \mathbb{E}[(S_{n+1} - S_n)^2]$
= $S_n^2 - n, \quad n \in \mathbb{N}.$

<u>Step 2</u>. The stopped process $(S^2_{\tau \wedge n} - \tau \wedge n)_{n \in \mathbb{N}}$ 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

$$k^2 = \mathbb{E}\left[S_0^2 - 0 \mid S_0 = k
ight] = \mathbb{E}\left[S_{ au \wedge n}^2 - au \wedge n \mid S_0 = k
ight],$$

and after taking the limit as \boldsymbol{n} tends to infinity,

$$egin{aligned} k^2 &= &\lim_{n o \infty} \mathbb{E} \left[S^2_{ au \wedge n} - au \wedge n \mid S_0 = k
ight] \ &= &\lim_{n o \infty} \mathbb{E} \left[S^2_{ au \wedge n} \mid S_0 = k
ight] - \lim_{n o \infty} \mathbb{E} [au \wedge n \mid S_0 = k] \ &= &\mathbb{E} \left[&\lim_{n o \infty} S^2_{ au \wedge n} \mid S_0 = k
ight] - \mathbb{E} \left[&\lim_{n o \infty} au \wedge n \mid S_0 = k
ight] \ &= &\mathbb{E} \left[&\lim_{n o \infty} S^2_{ au \wedge n} - &\lim_{n o \infty} au \wedge n \mid S_0 = k
ight] \ &= &\mathbb{E} \left[S^2_{ au} - au \mid S_0 = k
ight], \end{aligned}$$

since $S^2_{\tau \wedge n} \in [0, B^2]$ for all $n \in \mathbb{N}$ and $n \mapsto \tau \wedge n$ is non-decreasing, and this gives^{*†}

$$egin{aligned} k^2 &= & \mathbb{E}[S_{ au}^2 - au \mid S_0 = k] \ &= & \mathbb{E}[S_{ au}^2 \mid S_0 = k] - \mathbb{E}[au \mid S_0 = k] \ &= & B^2 \mathbb{P}(S_{ au} = B \mid S_0 = k) + 0^2 imes \mathbb{P}(S_{ au} = 0 \mid S_0 = k) - \mathbb{E}[au \mid S_0 = k], \end{aligned}$$

i.e.

$$egin{array}{rll} \mathbb{E}[au \mid S_0 = k] &= B^2 \mathbb{P}(S_ au = B \mid S_0 = k) - k^2 \ &= B^2 rac{k}{B} - k^2 \end{array}$$

^{*}By application of the *dominated convergence* theorem.

[†]By application of the *monotone convergence* theorem.

$$= k(B-k), \qquad \diamondsuit 421$$

 $k = 0, 1, \dots, B$, 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).

<u>Step 2</u>. 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.

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

$$k = \mathbb{E}[S_0 - 0 \mid S_0 = k] = \mathbb{E}[S_{ au \wedge n} - (p - q)(au \wedge n) \mid S_0 = k],$$

and after taking the limit as n goes to infinity,

$$egin{aligned} k &= \lim_{n o \infty} \mathbb{E}[S_{ au \wedge n} - (p-q)(au \wedge n) \mid S_0 = k] \ &= \mathbb{E}\left[\lim_{n o \infty} S_{ au \wedge n} - (p-q)\lim_{n o \infty} au \wedge n \mid S_0 = k
ight] \ &= \mathbb{E}[S_{ au} - (p-q) au \mid S_0 = k], \end{aligned}$$

Probabilities Problem	Unbiased	Biased
Ruin probability	S_n	$\left(rac{oldsymbol{q}}{oldsymbol{p}} ight)^{oldsymbol{S_n}}$
Mean game duration	S_n^2-n	$S_n-(p-q)n$



which gives

$$egin{aligned} k &= & \mathbb{E}[S_{ au} - (p-q) au \mid S_0 = k] \ &= & \mathbb{E}[S_{ au} \mid S_0 = k] - (p-q) \, \mathbb{E}[au \mid S_0 = k] \ &= & B imes \mathbb{P}(S_{ au} = B \mid S_0 = k) + 0 imes \mathbb{P}(S_{ au} = 0 \mid S_0 = k) - (p-q) \, \mathbb{E}[au \mid S_0 = k], \end{aligned}$$

i.e.

$$egin{array}{rll} (p-q) \ \mathbb{E}[au \mid S_0 = k] &= \ B imes \mathbb{P}(S_ au = B \mid S_0 = k) - k \ &= \ B rac{(q/p)^k - 1}{(q/p)^B - 1} - k, \end{array}$$

from (10.4.2), hence

$$\mathbb{E}[au \mid S_0 = k] = rac{1}{p-q} \left(B rac{(q/p)^k - 1}{(q/p)^B - 1} - k
ight), \qquad k = 0, 1, \dots, B,$$

which recovers (2.3.11).

In Table 10.1 we summarize the family of martingales used to treat the above problems.

MH3512 AY19-20

ॳ22

Exercises

Exercise 10.1 Consider a sequence $(X_n)_{n \ge 1}$ of independent Bernoulli random variables with

$$\mathbb{P}(X_n=1)=\mathbb{P}(X_n=-1)=1/2, \qquad n\geqslant 1,$$

and the process $(M_n)_{n\in\mathbb{N}}$ defined by $M_0:=0$ and

$$M_n:=\sum\limits_{k=1}^n 2^{k-1}X_k,\qquad n\geqslant 1,$$

see Figure 10.3.



Figure 10.3: Possible paths of the process $(M_n)_{n \in \mathbb{N}}$.

Note that when $X_1 = X_2 = \cdots = X_{n-1} = -1$ and $X_n = 1$, we have

$$M_n = -\sum_{k=1}^{n-1} 2^{k-1} + 2^{n-1} = -rac{1-2^{n-1}}{1-2} + 2^{n-1} = 1, \qquad n \geqslant 1.$$

- a) Show that the process $(M_n)_{n \in \mathbb{N}}$ is a martingale.
- b) Is the random time

$$au := \inf\{n \geqslant 1 \; : \; M_n = 1\}$$
 MH3512 AY19-20

a stopping time?

c) Consider the stopped process

$$M_{ au \wedge n} := M_n \mathbb{1}_{\{n < au\}} + \mathbb{1}_{\{ au \leqslant n\}} = \left\{egin{array}{c} M_n = 1 - 2^n & ext{if } n < au, \ M_ au = 1 & ext{if } n \geqslant au, \end{array}
ight.$$

 $n \in \mathbb{N}$, see Figure 10.4. Give an interpretation of $(M_{n \wedge \tau})_{n \in \mathbb{N}}$ in terms of betting strategy for a gambler starting a game at $M_0 = 0$.



Figure 10.4: Possible paths of the stopped process $(M_{\tau \wedge n})_{n \in \mathbb{N}}$.

d) Determine the two possible values of $M_{\tau \wedge n}$ and the probability distribution of $M_{\tau \wedge n}$ at any time $n \ge 1$.

e) Show, using the result of Question (d), that we have

$$\mathbb{E}[M_{ au \wedge n}] = 0, \qquad n \in \mathbb{N}$$

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

$$M_n \leqslant \mathbb{E}[M_{n+1} \mid \mathcal{F}_n], \qquad n \geqslant 0.$$

- a) Show that we have $\mathbb{E}[M_{n+1}] \ge \mathbb{E}[M_n]$, $n \ge 0$, *i.e.* a submartingale has a non-decreasing 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

$$A_{n+1}:=A_n+\mathbb{E}[M_{n+1}-M_n\mid\mathcal{F}_n], \hspace{1em} n\geqslant 0,$$

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

$$\mathbb{E}[M_{\sigma}] \leqslant \mathbb{E}[M_{ au}].$$

Hint: Use the Doob stopping time Theorem 10.6 for martingales and (10.3.3).

Exercise 10.3 Consider an asset price $(S_n)_{n=0,1,\dots,N}$ 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

$$\phi\left(rac{S_1+S_2+\dots+S_N}{N}
ight)$$

is upper bounded by the price of the European call option with maturity N, *i.e.* show that

$$\mathbb{E}^*\left[\phi\left(rac{S_1+S_2+\dots+S_N}{N}
ight)
ight]\leqslant\mathbb{E}^*[\phi(S_N)],$$

Hint: Use in the following order:

(i) the convexity inequality

$$\phi\left(rac{x_1+x_2+\dots+x_n}{n}
ight)\leqslantrac{\phi(x_1)+\phi(x_2)+\dots+\phi(x_n)}{n},$$

- (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 stochastic process $(M_n)_{n \in \mathbb{N}}$ is a *sub*martingale if it satisfies

$$M_k \leqslant \mathbb{E}[M_n \mid \mathcal{F}_k], \qquad k=0,1,\dots,n.$$

- a) Show that the expectation $\mathbb{E}[M_n]$ of a submartingale increases with time $n \in \mathbb{N}$.
- b) Consider the random walk given by $S_0 := 0$ and

$$S_n := \sum_{k=1}^n X_k = X_1 + X_2 + \dots + X_n, \qquad n \geqslant 1,$$

where $(X_n)_{n \ge 1}$ is an i.i.d. Bernoulli sequence of $\{0, 1\}$ -valued random variables with $\mathbb{P}(X_n = 1) = p, n \ge 1$. Under which condition on $\alpha \in \mathbb{R}$ is the process $(S_n - \alpha n)_{n \in \mathbb{N}}$ a submartingale? MH3512 AY19-20

Exercise 10.5 Recall that a stochastic process $(M_n)_{n \in \mathbb{N}}$ is a submartingale if it satisfies

$$M_k \leqslant \mathbb{E}[M_n \mid \mathcal{F}_k], \qquad k=0,1,\ldots,n.$$

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 non-decreasing function $\phi(M_n)$ of a submartingale $(M_n)_{n \in \mathbb{N}}$ remains a submartingale.

Problem 10.6

a) Consider $(M_n)_{n \in \mathbb{N}}$ a nonnegative martingale. For any x > 0, let

$$au_x := \inf\{n \geqslant 0 \; : \; M_n \geqslant x\}.$$

Show that the random time τ_x is a stopping time.

b) Show that for all $n \ge 0$ we have

$$\mathbb{P}\left(\max_{k=0,1,\dots,n} M_k \geqslant x\right) \leqslant \frac{\mathbb{E}[M_n]}{x}, \qquad x > 0.$$
(10.5.1)

Hint: Proceed as in the proof of the Markov inequality and use 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 submartingale.

Hint: Use the Doob stopping time theorem for *sub*martingales as in Exercise 10.2-(d).

d) Show that for any $n \ge 0$ we have

$$\mathbb{P}\Big(\max_{k=0,1,...,n}M_k\geqslant x\Big)\leqslant rac{\mathbb{E}[(M_n)^2]}{x^2},\qquad x>0.$$

e) Show that more generally we have

$$\mathbb{P}\Big(\max_{k=0,1,...,n}M_k\geqslant x\Big)\leqslant rac{\mathbb{E}[(M_n)^p]}{x^p},\qquad x>0,$$

for all $n \ge 0$ and $p \ge 1$.

f) Given $(Y_n)_{n \ge 1}$ a sequence of centered independent random variables with same mean $\mathbb{E}[Y_n] = 0$ and variance $\sigma^2 = \operatorname{Var}[Y_n]$, $n \ge 1$, consider the random walk $S_n = Y_1 + Y_2 + \cdots + Y_n$, $n \ge 1$, with $S_0 = 0$.

Show that for all $n \ge 0$ we have

$$\mathbb{P}\Big(\max_{k=0,1,...,n}|S_k|\geqslant x\Big)\leqslant rac{n\sigma^2}{x^2},\qquad x>0.$$

g) Show that for any (not necessarily nonnegative) submartingale we have

$$\mathbb{P}\Big(\max_{k=0,1,...,n}M_k\geqslant x\Big)\leqslant rac{\mathbb{E}[M_n^+]}{x},\qquad x>0,$$

where $z^+ = \max(z, 0), z \in \mathbb{R}$.

h) A stochastic process $(M_n)_{n \in \mathbb{N}}$ is a *super*martingale^{*} if it satisfies

$$\mathbb{E}[M_n \mid \mathcal{F}_k] \leqslant M_k, \qquad k=0,1,\ldots,n.$$

Show that for any *nonnegative super*martingale we have

$$\mathbb{P}\Big(\max_{k=0,1,...,n}M_k\geqslant x\Big)\leqslant rac{\mathbb{E}[M_0]}{x},\qquad x>0.$$

i) Show that for any *nonnegative submartingale* $(M_n)_{n \in \mathbb{N}}$ and any convex non-decreasing nonnegative function ϕ we have

$$\mathbb{P}\Big(\max_{k=0,1,...,n}\phi(M_k)\geqslant x\Big)\leqslant rac{\mathbb{E}[\phi(M_n)]}{x},\qquad x>0.$$

Hint: Consider the stopping time

$$au_x^\phi:=\inf\{n\geqslant 0\;:\; M_n\geqslant x\}.$$

j) Give an example of a nonnegative *super*martingale which is *not* a martingale.

Exercise 10.7 Consider the random walk $(S_n)_{n \in \mathbb{N}}$ on $\{0, 1, \ldots, B\}$ with independent $\{-1, 1\}$ -valued increments, and

$$\mathbb{P}(S_{n+1} - S_n = +1) = p$$
 and $\mathbb{P}(S_{n+1} - S_n = -1) = q$, $n \in \mathbb{N}$,

^{*&}quot;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, page 808.

and the martingale

where $p,q \in (0,1)$ are such that p+q=1. Show that for all $n \ge 0$ and $r \ge 1$ we have

$$\mathbb{P}\Big(\max_{k=0,1,...,n}M_k\geqslant x\Big)\leqslant rac{(p(q/p)^r+q(p/q)^r)^{2n}}{x^r},\qquad x>0.$$

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 Section 9.1. However, in contrast with the previous chapter, no time ordering is *a priori* imposed here on the index set X. Sections 11.4 and 11.5 contain some more advanced results on moments and deviation inequalities for Poisson stochastic integrals.

11.1 Spatial Poisson (1781-1840) Processes	431
11.2 Poisson Stochastic Integrals	433
11.3 Transformations of Poisson Measures	435
11.4 Moments of Poisson Stochastic Integrals	438
11.5 Deviation Inequalities	442
Exercises	445

11.1 Spatial Poisson (1781-1840) Processes

In this section we present the construction of spatial Poisson processes on a space of configurations of $X \subset \mathbb{R}^d$, $d \ge 1$. The set

$$\Omega^X:=ig\{\omega:=(x_i)_{i=1}^N\subset X,\;N\in\mathbb{N}\cup\{\infty\}ig\},$$

is called the space of configurations on $X \subset \mathbb{R}^d$. The next figure illustrates a given configuration $\omega \in \Omega^X$.



Given \boldsymbol{A} a (measurable) subset of \boldsymbol{X} , we let

$$\omega(A)=\#\{x\in\omega\;:\;x\in A\}=\sum\limits_{x\in\omega}\mathbb{1}_A(x)$$

denote the number of configuration points in $\boldsymbol{\omega}$ that are contained in the set \boldsymbol{A} .

Given $\rho: X \longrightarrow \mathbb{R}_+$ a nonnegative function, the Poisson probability measure \mathbb{P}^X_{σ} with intensity $\sigma(dx) = \rho(x)dx$ on X is the only probability measure on Ω^X satisfying

i) For any (measurable) subset \boldsymbol{A} of \boldsymbol{X} such that

$$\sigma(A)=\int_A
ho(x)dx=\int_{\mathbb{R}^d}\mathbb{1}_A(x)
ho(x)dx<\infty,$$

the number $\omega(A)$ of configuration points contained in A is a Poisson random variable with intensity $\sigma(A)$, *i.e.*

$$\mathbb{P}^X_\sigma(\omega\in\Omega^X\ :\ \omega(A)=n)=\mathrm{e}^{-\sigma(A)}rac{(\sigma(A))^n}{n!},\qquad n\in\mathbb{N}$$

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

$$\omega \longmapsto (\omega(A_1), \dots, \omega(A_n)), \qquad \omega \in \Omega^X,$$

is made of independent random variables for all $n \ge 1$.

In the remaining of this chapter we will assume that $\sigma(X) < \infty$ for simplicity.

The Poisson measure \mathbb{P}^X_σ can also be defined as

$$\mathbb{E}_{\mathbb{P}_{\sigma}^{X}}[F] = \mathrm{e}^{-\sigma(X)} \sum_{n \geqslant 0} rac{1}{n!} \int_{X^{n}} f_{n}(x_{1}, x_{2}, \dots, x_{n}) \sigma(dx_{1}) \cdots \sigma(dx_{n})$$
(11.1.1)
$$(11.1.1)$$

$$(11.1.1)$$

$$(11.1.1)$$

$$(11.1.1)$$

$$(11.1.1)$$
for \boldsymbol{F} written as

$$F(\omega) = \sum\limits_{n \geqslant 0} \mathbb{1}_{\{\omega(X) = n\}} f_n(x_1, x_2, \dots, x_n)$$

where f_n is a symmetric integrable function of $\omega = \{x_1, x_2, \dots, x_n\}$ when $\omega(X) = n, n \ge 1$, cf. *e.g.* Proposition 6.1.3 and § 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

$$\mathbb{P}_{\sigma}^{X}(\{x_{1},\ldots,x_{n}\} \subset A \mid \omega(X) = n) = \frac{\mathbb{P}_{\sigma}^{X}(\{x_{1},\ldots,x_{n}\} \subset A \text{ and } \omega(X) = n)}{\mathbb{P}_{\sigma}^{X}(\omega(X) = n)} \\
= \frac{1}{\mathbb{P}_{\sigma}^{X}(\omega(X) = n)} \mathbb{E}_{\mathbb{P}_{\sigma}^{X}}[\mathbb{1}_{\{\omega(X) = n\}}\mathbb{1}_{A^{n}}(x_{1},x_{2},\ldots,x_{n})] \\
= \left(\frac{\sigma(A)}{\sigma(X)}\right)^{n}.$$
(11.1.2)

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

<u>(</u>) 434

Let \boldsymbol{f} be an integrable function on \boldsymbol{X} . We have

$$\mathbb{E}_{\mathbb{P}^X_\sigma}\left[\exp\left(i\int_X f(x)\omega(dx)
ight)
ight]=\exp\left(\int_X (\mathrm{e}^{if(x)}-1)\sigma(dx)
ight).$$

Proof

We assume that $\sigma(X) < \infty$. By (11.1.1) we have

$$egin{aligned} \mathbb{E}_{\mathbb{P}^X_\sigma} \left[\exp\left(i\int_X f(x)\omega(dx)
ight)
ight] \ &= \mathrm{e}^{-\sigma(X)}\sum\limits_{n\geqslant 0}rac{1}{n!}\int_X\cdots\int_X \mathrm{e}^{i(f(x_1)+\cdots+f(x_n))}\sigma(dx_1)\cdots\sigma(dx_n). \ &= \mathrm{e}^{-\sigma(X)}\sum\limits_{n\geqslant 0}rac{1}{n!}\left(\int_X \mathrm{e}^{if(x)}\sigma(dx)
ight)^n \ &= \exp\left(\int_X (\mathrm{e}^{if(x)}-1)\sigma(dx)
ight). \end{aligned}$$

The characteristic function allows us to compute the expectation of $\int_X f(x) \omega(dx)$, as

$$egin{aligned} \mathbb{E}_{\mathbb{P}_{\sigma}^{X}}\left[\int_{X}f(x)\omega(dx)
ight] &= -irac{d}{darepsilon}\mathbb{E}_{\mathbb{P}_{\sigma}^{X}}\left[\exp\left(iarepsilon\int_{X}f(x)\omega(dx)
ight)
ight]_{ertarepsilon=0} \ &= -irac{d}{darepsilon}\exp\left(\int_{X}(\mathrm{e}^{iarepsilon f(x)}-1)\sigma(dx)
ight)_{ertarepsilon=0} \ &= \int_{X}f(x)\sigma(dx), \end{aligned}$$

for f an integrable function on X. As a consequence, the *compensated* Poisson stochastic integral

$$\int_X f(x) \omega(dx) - \int_X f(x) \sigma(dx)$$
 MH3512 AY19-2

is a *centered* random variable, *i.e.* we have

$$\mathbb{E}_{\mathbb{P}^X_\sigma}\left[\int_X f(x) \omega(dx) - \int_X f(x) \sigma(dx)
ight] = 0.$$

The variance can be similarly computed as

$$\mathbb{E}_{\mathbb{P}_{\sigma}^X}\left[\left(\int_X f(x)(\omega(dx)-\sigma(dx))
ight)^2
ight]=\int_X |f(x)|^2\sigma(dx),$$

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

$$\log \mathbb{E}_{\mathbb{P}^X_\sigma} \left[\exp \left(\int_X f(x) \omega(dx)
ight)
ight] = \int_X (\mathrm{e}^{f(x)} - 1) \sigma(dx) = \sum_{n \geqslant 1} rac{1}{n!} \int_X f^n(x) \sigma(dx),$$

shows that the cumulants of $\int_X f(x)\omega(dx)$ are given by

$$\kappa_n = \int_X f^n(x)\sigma(dx), \qquad n \ge 1.$$
(11.2.1)

11.3 Transformations of Poisson Measures

Consider a mapping $au: (X, \sigma) \longrightarrow (Y, \mu)$, and let

$$au_*: \Omega^X \longrightarrow \Omega^Y$$

be the transformed configuration defined by

$$au_*(\omega):=\{ au(x)\ :\ x\in\omega\},\qquad\omega\in\Omega^X,$$

as illustrated in the following figure.



Proposition 11.2

The random configuration

$$egin{array}{rcl} \Omega^X & \colon & \longrightarrow \Omega^Y \ & \omega & \longmapsto au_*(\omega) \end{array}$$

has the Poisson distribution \mathbb{P}^{Y}_{μ} with intensity μ on Y, where μ is defined by

$$\mu(A):=\int_X\mathbbm{1}_A(au(x))\sigma(dx)=\int_X\mathbbm{1}_{ au^{-1}(A)}(x)\sigma(dx)=\sigma(au^{-1}(A)),$$

for \boldsymbol{A} a (measurable) subset of \boldsymbol{X} .

Proof

We have

$$\begin{split} \mathbb{P}_{\sigma}^{X}(\tau_{*}\omega(A) = n) &= \mathbb{P}_{\sigma}^{X}(\omega(\tau^{-1}(A)) = n) \\ &= \mathrm{e}^{-\sigma(\tau^{-1}(A))} \frac{(\sigma(\tau^{-1}(A)))^{n}}{n!} \\ &= \mathrm{e}^{-\mu(A)} \frac{(\mu(A))^{n}}{n!}. \end{split}$$

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

$$\begin{split} \mathbb{P}_{\sigma}^{X}(\{\omega \in \Omega^{X} : \tau_{*}\omega(A_{1}) = k_{1}, \dots, \tau_{*}\omega(A_{n}) = k_{n}\}) \\ &= \prod_{i=1}^{n} \mathbb{P}_{\sigma}^{X}(\{\tau_{*}\omega(A_{i}) = k_{i}\}) \\ &= \prod_{i=1}^{n} \mathbb{P}_{\sigma}^{X}(\{\omega(\tau^{-1}(A_{i})) = k_{i}\}) \\ &= \exp\left(-\sum_{i=1}^{n} \sigma(\tau^{-1}(A_{i}))\right) \prod_{i=1}^{n} \frac{(\sigma(\tau^{-1}(A_{i})))^{k_{i}}}{k_{i}!} \\ &= \exp\left(-\sum_{i=1}^{n} \mu(A_{i})\right) \prod_{i=1}^{n} \frac{(\mu(A_{i}))^{k_{i}}}{k_{i}!} \\ &= \prod_{i=1}^{n} \mathbb{P}_{\mu}^{Y}(\{\omega(A_{i}) = k_{i}\}) \\ &= \mathbb{P}_{\mu}^{Y}(\{\omega(A_{1}) = k_{1}, \dots, \omega(A_{n}) = k_{n}\}). \end{split}$$

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

$$\begin{split} \mathbb{P}_{\sigma}^{X}(\tau_{*}\omega([0,t]) = n) &= \mathbb{P}_{\sigma}^{X}(\omega(\tau^{-1}([0,t])) = n) \\ &= e^{-\sigma(\tau^{-1}([0,t]))} \frac{(\sigma(\tau^{-1}([0,t])))^{n}}{n!} \\ &= e^{-2\lambda t} (2\lambda t)^{n} / n!. \end{split}$$



(5438)

11.4 Moments of Poisson Stochastic Integrals

As a consequence of (11.2.1) and the relation (1.7.8) between cumulants and moments we find that the *n*-th moment of $\int_X f(x)\omega(dx)$ can be written as

$$egin{aligned} \mathbb{E}_{\mathbb{P}_{\sigma}^X}\left[\left(\int_X f(x)\omega(dx)
ight)^n
ight]\ &=\;\sum\limits_{k=0}^nrac{1}{k!}\sum\limits_{d_1+\dots+d_k=n}rac{n!}{d_1!\dots d_k!}\int_X f^{d_1}(x)\sigma(dx)\dots\int_X f^{d_k}(x)\sigma(dx)\ &=\;\sum\limits_{k=1}^n\sum\limits_{B_1^n,B_2^n,\dots,B_k^n}\int_X f^{|B_1^n|}(x)\sigma(dx)\dots\int_X f^{|B_k^n|}(x)\sigma(dx), \end{aligned}$$

where the sum runs over the partitions $B_1^n, B_2^n, \ldots, B_k^n$ of $\{1, \ldots, n\}$ with cardinality $|B_i^n|$. Similarly, we find that the moments of the centered integral $\int_X f(x)\omega(dx) - \int_X f(x)\sigma(dx)$ satisfy

$$\mathbb{E}_{\mathbb{P}_{\sigma}^X}\left[\left(\int_X f(x)\omega(dx)-\int_X f(x)\sigma(dx)
ight)^n
ight] \ = \sum\limits_{k=1}^n \sum\limits_{\substack{B_1^n,B_2^n,...,B_k^n\ |B_1^n| \geqslant 2,...,|B_k^n| \geqslant 2}} \int_X f^{|B_1^n|}(x)\sigma(dx)\cdots \int_X f^{|B_k^n|}(x)\sigma(dx),$$

where the sum runs over the partitions $B_1^n, B_2^n, \ldots, B_k^n$ of $\{1, \ldots, n\}$ of size $|B_k^n| \ge 2$.

By taking $f = \mathbb{1}_A$ we can specialize the above results to the computation of the moments and central moments of the Poisson-random

variable $\omega(A)$ with parameter $\lambda := \sigma(A)$. For the moments we recover the well-known identity

$$\mathbb{E}\left[(\omega(A))^n
ight] = \sum\limits_{k=1}^n \lambda^n S(n,k),$$

where S(n,k) is the Stirling number of the second kind, *i.e.* the number of partitions of $\{1, 2, ..., n\}$ made of k subsets, k = 0, 1, ..., n. Similarly, for the central moments we have

$$\mathbb{E}[(\omega(A) - \sigma(A))^n] = \sum_{k=1}^n \lambda^n S_2(n,k), \qquad (11.4.1)$$

where $S_2(n, k)$ is the number of partitions of $\{1, 2, ..., n\}$ made of k subsets of size at least 2, k = 0, 1, ..., n, cf. [Pri11]. The following version of Mecke's identity [Mec67] allows us to compute the first moment of the first order stochastic integral of a random integrand.

Proposition 11.3

Let $u: X \times \Omega^X \longrightarrow \mathbb{R}$ be a (measurable) process. We have

$$\mathbb{E}_{P^X_\sigma}\left[\int_X u(x,\omega)\omega(dx)
ight] = \mathbb{E}_{P^X_\sigma}\left[\int_X u(x,\omega\cup\{x\})\sigma(dx)
ight],$$

provided that

$$\mathbb{E}_{P^X_\sigma}\left[\int_X |u(x,\omega\cup\{x\})|\sigma(dx)
ight] <\infty.$$

Proof

We take $u(x, \omega)$ written as

$$u(x,\omega) = \sum\limits_{n \geqslant 0} \mathbb{1}_{\{\omega(X)=n\}} f_n(x;x_1,x_2,\ldots,x_n),$$

where $(x_1, x_2, \ldots, x_n) \mapsto f_n(x; x_1, x_2, \ldots, x_n)$ is a symmetric integrable function of $\omega = \{x_1, x_2, \ldots, x_n\}$ when $\omega(X) = n$, for each $n \ge 1$. We have

$$\mathbb{E}_{\mathbb{P}^X_\sigma}\left[\int_X u(x,\omega)\omega(dx)
ight]$$

439

$$\begin{split} &= \mathrm{e}^{-\sigma(X)} \sum_{n \geqslant 0} \frac{1}{n!} \sum_{k=1}^n \int_{X^n} f_n(x_i; x_1, x_2, \dots, x_n) \sigma(dx_1) \cdots \sigma(dx_n) \\ &= \sum_{n \geqslant 1} \frac{\mathrm{e}^{-\sigma(X)}}{(n-1)!} \int_{X^n} f_n(x; x_1, x_2, \dots, x_{i-1}, x, x_i, \dots, x_{n-1}) \sigma(dx) \sigma(dx_1) \cdots \sigma(dx_{n-1}) \\ &= \mathrm{e}^{-\sigma(X)} \sum_{n \geqslant 0} \frac{1}{n!} \int_{X^n} \int_X f_{n+1}(x; x, x_1, x_2, \dots, x_n) \sigma(dx) \sigma(dx_1) \cdots \sigma(dx_n) \\ &= \mathbb{E}_{\mathbb{P}_{\sigma}^X} \left[\int_X u(x, \omega \cup \{x\}) \sigma(dx) \right]. \end{split}$$

Proposition 11.3 can actually be extended to a moment identity, as follows.

Proposition 11.4

([Pri12]) Let $u: X \times \Omega^X \longrightarrow \mathbb{R}$ be a (measurable) process. We have

$$\mathbb{E}_{\mathbb{P}_{\sigma}^X}\left[\left(\int_X u(x,\omega)\omega(dx)
ight)^n
ight] \ = \sum\limits_{k=1}^n \sum\limits_{B_1^n,B_2^n,...,B_k^n} \mathbb{E}_{\mathbb{P}_{\sigma}^X}\left[\int_{X^k} u^{|B_1^n|}(x_1,\omega\cup\{x_1,x_2,\ldots,x_k\})\cdots \ \cdots u^{|B_k^n|}(x_k,\omega\cup\{x_1,x_2,\ldots,x_k\})\sigma(dx_1)\cdots\sigma(dx_k)
ight],$$

where the sum runs over the partitions $B_1^n, B_2^n, \ldots, B_k^n$ of $\{1, \ldots, n\}$, for any $n \ge 1$ such that all terms are integrable.

Proof

We will prove the following slightly more general formula

$$egin{aligned} &\mathbb{E}_{\mathbb{P}_{\sigma}^{X}}\left[F\left(\int_{X}u(x,\omega)\omega(dx)
ight)^{n}
ight]\ &=&\sum\limits_{k=1}^{n}\sum\limits_{B_{1}^{n},B_{2}^{n},...,B_{k}^{n}}\mathbb{E}_{\mathbb{P}_{\sigma}^{X}}\left[\int_{X^{k}}F(\omega\cup\{x_{1},x_{2},\ldots,x_{k}\})u^{|B_{1}^{n}|}(x_{1},\omega\cup\{x_{1},x_{2},\ldots,x_{k}\})\cdots\ &\cdots u^{|B_{k}^{n}|}(x_{k},\omega\cup\{x_{1},x_{2},\ldots,x_{k}\})\sigma(dx_{1})\cdots\sigma(dx_{k})
ight], \end{aligned}$$

MH3512 AY19-20

by induction on $n \ge 1$, for F a sufficiently integrable random variable. Clearly, the formula holds for n = 0 and n = 1. Assuming 1 that it holds at the rank n and using Proposition 11.3 we get

$$\begin{split} \mathbb{F}_{\mathbb{T}_{\sigma}^{X}} \left[F\left(\int_{X} u(x,\omega)\omega(dx) \right)^{n+1} \right] \\ &= \sum_{k=1}^{n} \sum_{B_{1}^{n}, B_{2}^{n}, \dots, B_{k}^{n}} \\ \mathbb{F}_{\mathbb{P}_{\sigma}^{X}} \left[\int_{X^{k}} u^{|B^{n}|}(x_{1}, \omega \cup \{x_{1}, x_{2}, \dots, x_{k}\}) \cdots u^{|B^{n}_{k}|}(x_{k}, \omega \cup \{x_{1}, x_{2}, \dots, x_{k}\}) \\ &\times F(\omega \cup \{x_{1}, x_{2}, \dots, x_{k}\}) \sum_{i=1}^{k} u(x_{i}, \omega \cup \{x_{1}, x_{2}, \dots, x_{k}\})\sigma(dx_{1}) \cdots \sigma(dx_{k}) \right] \\ &+ \sum_{k=1}^{n} \sum_{B_{1}^{n}, B_{2}^{n}, \dots, B_{k}^{n}} \\ \mathbb{E}_{\mathbb{P}_{\sigma}^{X}} \left[\int_{X^{k}} u^{|B^{n}|}(x_{1}, \omega \cup \{x_{1}, x_{2}, \dots, x_{k}\}) \cdots u^{|B^{n}_{k}|}(x_{k}, \omega \cup \{x_{1}, x_{2}, \dots, x_{k}\}) \\ &\times F(\omega \cup \{x_{1}, x_{2}, \dots, x_{k}\}) \int_{X} u(y, \omega \cup \{x_{1}, x_{2}, \dots, x_{k}\})\omega(dy)\sigma(dx_{1}) \cdots \sigma(dx_{k}) \right] \\ &= \sum_{k=1}^{n} \sum_{B_{1}^{n}, B_{2}^{n}, \dots, B_{k}^{n}} \\ \mathbb{E}_{\mathbb{P}_{\sigma}^{X}} \left[\int_{X^{k}} u^{|B^{n}|}(x_{1}, \omega \cup \{x_{1}, x_{2}, \dots, x_{k}\}) \cdots u^{|B^{n}_{k}|}(x_{k}, \omega \cup \{x_{1}, x_{2}, \dots, x_{k}\}) \\ &\times F(\omega \cup \{x_{1}, x_{2}, \dots, x_{k}\}) \sum_{i=1}^{k} u(x_{i}, \omega \cup \{x_{1}, x_{2}, \dots, x_{k}\})\sigma(dx_{1}) \cdots \sigma(dx_{k}) \\ &+ \sum_{k=1}^{n} \sum_{B_{1}^{n}, B_{2}^{n}, \dots, B_{k}^{n}} \\ \mathbb{E}_{\mathbb{P}_{\sigma}^{X}} \left[\int_{X^{k+1}} u^{|B^{n}|}(x_{1}, \omega \cup \{x_{1}, x_{2}, \dots, x_{k}, y\}) u(y, \omega \cup \{x_{1}, x_{2}, \dots, x_{k}, y\})\sigma(dy)\sigma(dx_{1}) \cdots \sigma(dx_{k}) \right] \\ &= \sum_{k=1}^{n+1} \sum_{B_{1}^{n+1}, B_{2}^{n+1}, \dots, B_{k}^{n+1}} \\ \mathbb{E}_{\mathbb{P}_{\sigma}^{X}} \left[\int_{X^{k}} F(\omega \cup \{x_{1}, x_{2}, \dots, x_{k}\}) u^{|B^{n}_{1}^{n}|}(x_{1}, \omega \cup \{x_{1}, x_{2}, \dots, x_{k}\}) \times \cdots \right]$$

$$\cdots imes u^{|B_k^{n+1}|}(x_k, \omega \cup \{x_1, x_2, \dots, x_k\}) \sigma(dx_1) \cdots \sigma(dx_k) \Big],$$

where the sum runs over the partitions $B_1^{n+1}, B_2^{n+1}, \ldots, B_k^{n+1}$ of $\{1, \ldots, n+1\}$.

11.5 Deviation Inequalities

The next proposition is a particular case of results proved for general Poisson functionals in [HP02], [Wu00], cf. also Corollary 6.9.3 of [Pri09]. Such results are useful for the accurate estimation of loss (or outage) probabilities in a wireless network whose users are randomly located within the space X, cf. e.g. [DFM09].

Proposition 11.5
Let
$$f: X \longrightarrow \mathbb{R}$$
 be a (measurable) function such that
i) $f(y) \leq K$, $\sigma(dy)$ -a.e., for some $K > 0$, and
ii) $\int_X |f(y)|^2 \sigma(dy) \leq \alpha^2$, for some $\alpha > 0$.
Then we have
 $\mathbb{P}^X_{\sigma} \left(\int_X f(y) \omega(dy) - \int_X f(y) \sigma(dy) \geq x \right) \leq e^{x/K} \left(1 + \frac{xK}{\alpha^2} \right)^{-\frac{x}{K} - \frac{\alpha^2}{K^2}}$, (11.5.1)
for all $x > 0$.

Proof

We let

$$\int_X f(y) ilde{\omega}(dy) := \int_X f(y) (\omega(dy) - \sigma(dy))$$

denote the compensated Poisson stochastic integral of \boldsymbol{f} , and

$$L(s):=\mathbb{E}_{\mathbb{P}^X_\sigma}\left[\exp\left(s\int_X f(y) ilde{\omega}(dy)
ight)
ight],\qquad s\in\mathbb{R}_+.$$

By Proposition 11.3 we have

$$\begin{split} L'(s) &= \mathbb{E}_{\mathbb{P}_{\sigma}^{X}} \left[\int_{X} f(x) \tilde{\omega}(dx) \exp\left(s \int_{X} f(y) \tilde{\omega}(dy)\right) \right] \\ &= \mathbb{E}_{\mathbb{P}_{\sigma}^{X}} \left[\int_{X} \exp\left(s \int_{X} f(y) \tilde{\omega}(dy)\right) f(x) \tilde{\omega}(dx) \\ &- \int_{X} \exp\left(s \int_{X} f(y) \tilde{\omega}(dy)\right) f(x) \sigma(dx) \right] \\ &= \mathbb{E}_{\mathbb{P}_{\sigma}^{X}} \left[\int_{X} \exp\left(s f(x) + s \int_{X} f(y) \tilde{\omega}(dy)\right) f(x) \sigma(dx) \\ &- \int_{X} \exp\left(s \int_{X} f(y) \tilde{\omega}(dy)\right) f(x) \sigma(dx) \right] \\ &= \mathbb{E}_{\mathbb{P}_{\sigma}^{X}} \left[\int_{X} f(x) (e^{sf(x)} - 1) \sigma(dx) \exp\left(s \int_{X} f(y) \tilde{\omega}(dy)\right) \right] \\ &= \int_{X} f(x) (e^{sf(x)} - 1) \sigma(dx) \mathbb{E}_{\mathbb{P}_{\sigma}^{X}} \left[\exp\left(s \int_{X} f(y) \tilde{\omega}(dy)\right) \right] \\ &= s \int_{X} |f(x)|^{2} \frac{e^{sf(x)} - 1}{sf(x)} \sigma(dx) \mathbb{E}_{\mathbb{P}_{\sigma}^{X}} \left[\exp\left(s \int_{X} f(y) \tilde{\omega}(dy)\right) \right] \\ &\leqslant \frac{e^{sK} - 1}{K} \int_{X} |f(x)|^{2} \sigma(dx) \mathbb{E}_{\mathbb{P}_{\sigma}^{X}} \left[\exp\left(s \int_{X} f(y) \tilde{\omega}(dy)\right) \right] \\ &= \alpha^{2} \frac{e^{sK} - 1}{K} \mathbb{E}_{\mathbb{P}_{\sigma}^{X}} \left[\exp\left(s \int_{X} f(y) \tilde{\omega}(dy)\right) \right] = \alpha^{2} \frac{e^{sK} - 1}{K} L(s), \end{split}$$

which shows that

$$rac{L'(s)}{L(s)}\leqslant h(s):=lpha^2rac{{
m e}^{sK}-1}{K},\qquad s\in {\mathbb R}_+,$$

hence

$$L(t)\leqslant \exp\left(\int_0^t h(s)ds
ight)=\exp\left(lpha^2\int_0^t rac{\mathrm{e}^{sK}-1}{K}ds
ight),\qquad t\in\mathbb{R}_+.$$

Consequently we have

$$\mathbb{E}_{\mathbb{P}^X_\sigma}[F\mathrm{e}^{tF}]\leqslant h(t)\,\mathbb{E}_{\mathbb{P}^X_\sigma}[\mathrm{e}^{tF}],\qquad t\in\mathbb{R}_+$$

Next, the Markov inequality shows that

$$\mathbb{P}^X_\sigma\left(\int_X f(y)(\omega(dy)-\sigma(dy))\geqslant x
ight)=\mathbb{E}_{\mathbb{P}^X_\sigma}\left[\mathbbm{1}_{\{\int_X f(y)(\omega(dy)-\sigma(dy))\geqslant x\}}
ight]$$

MH3512 AY19-20

4 4 1 1

$$egin{aligned} &\leqslant \ \mathrm{e}^{-tx} \, \mathbb{E}_{\mathbb{P}^X_\sigma} \left[\mathbbm{1}_{\{\int_X f(y)(\omega(dy) - \sigma(dy)) \geqslant x\}} \exp\left(t \int_X f(y) \omega(dy)
ight)
ight] \ &\leqslant \ \mathrm{exp} \left(-tx + \int_0^t h(s) ds
ight) \ &\leqslant \ \mathrm{exp} \left(-tx + lpha^2 \int_0^t rac{\mathrm{e}^{sK} - 1}{K} ds
ight) \ &= \ \mathrm{exp} \left(-tx + rac{lpha^2}{K^2} (\mathrm{e}^{tK} - tK - 1)
ight). \end{aligned}$$

Minimizing the above term in t with $t = K^{-1} \log(1 + Kx/\alpha^2)$ shows that

$$\mathbb{P}^X_\sigma\left(\int_X f(y)(\omega(dy)-\sigma(dy))\geqslant x
ight)\leqslant \exp\left(rac{x}{K}-\left(rac{x}{K}+rac{lpha^2}{K^2}
ight)\log\left(1+rac{xK}{lpha^2}
ight)
ight),$$

which yields (11.5.1).

The bound (11.5.1) also shows that

$$\mathbb{P}^X_\sigma\left(\int_X f(y)\omega(dy) - \int_X f(y)\sigma(dy) \geqslant x
ight) \ \leqslant \ \exp\left(-rac{x}{2K}\log\left(1+rac{xK}{lpha^2}
ight)
ight) = \left(1+rac{xK}{lpha^2}
ight)^{-x/2K},$$

for all x > 0.

In case the function $f: X \longrightarrow \mathbb{R}$ is such that

i) $f(y) \leq 0, \, \sigma(dy)$ -a.e., for some $K \in \mathbb{R}$, and

ii) $\int_X |f(y)|^2 \sigma(dy) \leqslant lpha^2,$ MH3512 AY19-20

the above argument can be applied with K = 0, and also shows that

$$\mathbb{P}^X_\sigma\left(\int_X f(y)\omega(dy) - \int_X f(y)\sigma(dy) \geqslant x
ight) \leqslant \exp\left(-rac{x^2}{2lpha^2}
ight),$$

for all x > 0.

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 $(T_k)_{k \ge 1}$. 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 meters.

Exercise 11.3 Some living organisms are distributed in space according to a Poisson process of intensity $\theta = 0.6$ units per mm³. Compute the probability that more than two living organisms are found within a 10 mm³ 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

$$D((x,y),r) = \{(u,v) \in \mathbb{R}^2 \ : \ (x-u)^2 + (y-v)^2 \leqslant r^2\}$$

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 $D((0,0),3) \cup D((x,y),3)$ 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?

Exercise 11.7 Let S_n be a Poisson random variable with parameter λn for all $n \ge 1$, with $\lambda > 0$. Show that the moments of order p of $(S_n - \lambda n)/\sqrt{n}$ satisfy the bound

$$\sup_{n \geqslant 1} \mathbb{E} \left[\left| rac{S_n - \lambda n}{\sqrt{n}}
ight|^p
ight] < C_p$$

where $C_p > 0$ is a finite constant for all $p \ge 1$. *Hint*: Use Relation (11.4.1).

Exercise 11.8 Let $(N_t)_{t \in \mathbb{R}_+}$ denote a standard Poisson process on \mathbb{R}_+ . Given a bounded function $f \in L^1(\mathbb{R}_+)$ we let

$$\int_0^\infty f(y)(dN_y-dy)$$

denote the compensated Poisson stochastic integral of f, and let

$$L(s):=E\left[\exp\left(s\int_0^\infty f(y)(dN_y-dy)
ight)
ight]=\exp\left(\int_0^\infty(\mathrm{e}^{sf(y)}-sf(y)-1)dy
ight),$$

 $s \in \mathbb{R}_+$.

a) Show that we have

$$rac{L'(s)}{L(s)}\leqslant h(s):=lpha^2rac{e^{sK}-1}{K},\qquad s\in\mathbb{R}_+,$$

provided that $f(t) \leq K$, dt-a.e., for some K > 0 and provided in addition that $\int_0^\infty |f(y)|^2 dy \leq \alpha^2$, for some $\alpha > 0$.

b) Show that

$$L(t)\leqslant \exp\left(\int_0^t h(s)ds
ight)=\exp\left(lpha^2\int_0^t rac{e^{sK}-1}{K}ds
ight),\qquad t\in\mathbb{R}_+.$$

c) Show, using Chebyshev's inequality, that

$$\mathbb{P}\left(\int_{0}^{\infty}f(y)(dN_{y}-dy)\geqslant x
ight)\leqslant\mathrm{e}^{-tx}E\left[\exp\left(t\int_{0}^{\infty}f(y)dN_{y}
ight)
ight],$$

and that

$$\mathbb{P}\left(\int_{0}^{\infty}f(y)(dN_{y}-dy)\geqslant x
ight)\leqslant\exp\left(-tx+lpha^{2}\int_{0}^{t}rac{e^{sK}-1}{K}ds
ight).$$

d) By minimization in \boldsymbol{t} , show that

$$\mathbb{P}\left(\int_{0}^{\infty}f(y)dN_{y}-\int_{0}^{\infty}f(y)dy\geqslant x
ight)\leqslant\mathrm{e}^{x/K}\left(1+rac{xK}{lpha^{2}}
ight)^{-x/K-lpha^{2}/K^{2}},$$

for all x > 0, and that

$$\mathbb{P}\left(\int_{0}^{\infty}f(y)dN_{y}-\int_{0}^{\infty}f(y)dy\geqslant x
ight)\leqslant\left(1+rac{xK}{lpha^{2}}
ight)^{-x/2K},$$

for all x > 0.

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	448
12.2 Poisson Process with Time-Dependent Intensity	450
12.3 Mean Time to Failure	452
Exercises	453

12.1 Survival Probabilities

Let $\tau : \Omega \longrightarrow \mathbb{R}_+$ denote (random) the lifetime of an entity, and let $\mathbb{P}(\tau \ge t)$ 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

$$egin{aligned} \mathbb{P}(au > T \mid au > t) &= rac{\mathbb{P}(au > T ext{ and } au > t)}{\mathbb{P}(au > t)} \ &= rac{\mathbb{P}(au > T)}{\mathbb{P}(au > t)}, \qquad 0 \leqslant t \leqslant T. \end{aligned}$$

Let now

$$\lambda(t) := \lim_{h\searrow 0} rac{\mathbb{P}(au < t + h \mid au > t)}{h}, \qquad t \in \mathbb{R}_+,$$

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

$$\begin{split} \lambda(t) &= \lim_{h \searrow 0} \frac{\mathbb{P}(\tau < t+h \mid \tau > t)}{h} \\ &= \frac{1}{\mathbb{P}(\tau > t)} \lim_{h \searrow 0} \frac{\mathbb{P}(\tau < t+h \text{ and } \tau > t)}{h} \\ &= \frac{1}{\mathbb{P}(\tau > t)} \lim_{h \searrow 0} \frac{\mathbb{P}(\tau > t) - \mathbb{P}(\tau > t+h)}{h} \\ &= -\frac{d}{dt} \log \mathbb{P}(\tau > t) \\ &= -\frac{1}{\mathbb{P}(\tau > t)} \frac{d}{dt} \mathbb{P}(\tau > t) \end{split}$$
(12.1.1)
$$&= -\frac{1}{R(t)} \frac{d}{dt} R(t), \end{split}$$

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

$$R(t) = \mathbb{P}(\tau > t) = R(0) \exp\left(-\int_0^t \lambda(u) du\right) = \exp\left(-\int_0^t \lambda(u) du\right), \qquad (12.1.2)$$

 $t \in \mathbb{R}_+$. Hence we have

$$\mathbb{P}(\tau > T \mid \tau > t) = \frac{R(T)}{R(t)} = \exp\left(-\int_t^T \lambda(u) du\right), \qquad t \in [0, T].$$
(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

$$\mathbb{P}(au > T) = \prod_{0 < t < T} \mathbb{P}(au > t + dt \mid au > t) = \prod_{0 < t < T} \exp\left(-\lambda(t) dt
ight),$$

which yields

$$\mathbb{P}(au > t) = \exp\left(-\int_0^t \lambda(s) ds
ight), \qquad t \in \mathbb{R}_+,$$

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

$$\mathbb{P}(au > t) = \mathrm{e}^{-\lambda t}, \qquad t \geqslant 0$$

cf. (1.5.3). Given $(\tau_n)_{n\geq 0}$ a sequence of *i.i.d.* exponentially distributed random variables, letting

$$T_n= au_0+\dots+ au_{n-1},\qquad n\geqslant 1,$$

and

$$N_t = \sum\limits_{n \geqslant 1} \mathbb{1}_{[T_n,\infty)}(t), \qquad t \in \mathbb{R}_+,$$

defines the standard Poisson process with intensity $\lambda > 0$ of Section 9.1 and we have

$$\mathbb{P}(N_t-N_s=k)=\mathrm{e}^{-\lambda(t-s)}rac{(\lambda(t-s))^k}{k!},\qquad k\geqslant 0.$$

The intensity of the Poisson process can in fact made time-dependent. For example under the time change

$$X_t = N_{\int_0^t \lambda(s) ds}$$
 MH3512 AY19-20

where $(\lambda(u))_{u \in \mathbb{R}_+}$ is a deterministic function of time, we have

$$\mathbb{P}(X_t-X_s=k)=rac{\left(\int_s^t\lambda(u)du
ight)^k}{k!}\exp\left(-\int_s^t\lambda(u)du
ight),\qquad k\geqslant 0.$$

In this case we have

$$\mathbb{P}(X_{t+h} - X_t = 0) = e^{-\lambda(t)h} + o(h) = 1 - \lambda(t)h + o(h), \qquad h \searrow 0,$$
(12.2.1)

and

$$\mathbb{P}(X_{t+h} - X_t = 1) = 1 - e^{-\lambda(t)h} + o(h) \simeq \lambda(t)h, \qquad h \searrow 0, \tag{12.2.2}$$

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

$$R(t) = \mathbb{P}(\tau_0 > t) = \mathbb{P}(X_t = 0) = \exp\left(-\int_0^t \lambda(u) du\right), \quad t \ge 0, \quad (12.2.3)$$

hence by (12.1.3) we find

$$egin{array}{ll} \mathbb{P}(X_{t+h}=0\mid X_t=0) &= & \mathbb{P}(au_0>t+h\mid au_0>t) \ &= & \mathrm{e}^{-\lambda(t)h}+o(h) = 1-\lambda(t)h+o(h), \qquad h\searrow 0, \end{array}$$

and

$$\mathbb{P}(X_{t+h} \geqslant 1 \mid X_t = 0) \; = \; \mathbb{P}(au_0 < t+h \mid au_0 > t) = 1 - \mathbb{P}(au_0 > t+h \mid au_0 > t) \ = \; 1 - \mathrm{e}^{-\lambda h} \simeq \lambda(t)h + o(h), \qquad h \searrow 0,$$

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

$$\mathbb{P}(\lambda_{t+h}=\lambda\mid\lambda_t=0)=lpha h, \quad h\searrow 0,$$
 MH3512 AY19-20

and

$$\mathbb{P}(\lambda_{t+h}=0\mid\lambda_t=\lambda)=eta h,\qquad h\searrow 0.$$

In this case the probability distribution of N_t can be explicitly computed, cf. Chapter VI-7 in [KT81].

Renewal processes

A <u>renewal process</u> is a counting process $(N_t)_{t \in \mathbb{R}_+}$ given by

$$N_t = \sum\limits_{k \geqslant 1} k \mathbb{1}_{[T_k, T_{k+1})}(t) = \sum\limits_{k \geqslant 1} \mathbb{1}_{[T_k, \infty)}(t), \qquad t \in \mathbb{R}_+,$$

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

$$\mathbb{E}[\tau] = \int_0^\infty t \frac{d}{dt} \mathbb{P}(\tau < t) dt = -\int_0^\infty t \frac{d}{dt} \mathbb{P}(\tau > t) dt$$

= $-\int_0^\infty t R'(t) dt = \int_0^\infty R(t) dt,$ (12.3.1)

provided that $\lim_{t \searrow 0} tR(t) = 0$. For example when τ has the distribution function (12.2.3) we get

$$\mathbb{E}[au] = \int_0^\infty R(t) dt = \int_0^\infty \exp\left(-\int_0^t \lambda(u) du
ight) dt.$$

In case the function $\lambda(t) = \lambda > 0$ is constant we recover the mean value

$$\mathbb{E}[au] = \int_0^\infty \mathrm{e}^{-\lambda t} dt = rac{1}{\lambda}$$

of the exponential distribution with parameter $\lambda > 0$.

MH3512 AY19-20 $\,$

Exercise

Exercise 12.1 Assume that the random time au has the *Weibull* distribution with probability density

$$f_eta(x)=eta\mathbb{1}_{[0,\infty)}x^{eta-1}\mathrm{e}^{-t^eta},\qquad x\in\mathbb{R},$$

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.

454

Some Useful Identities

Here we present a summary of algebraic identities that are used in this text.

Indicator functions

$$\mathbb{1}_A(x) = \left\{egin{array}{cccc} 1 & ext{if } x \in A, \ & & & \mathbb{1}_{[a,b]}(x) = \left\{egin{array}{ccccc} 1 & ext{if } a \leqslant x \leqslant b, \ & & & & 0 \ & & & 0 \ \end{array}
ight.$$

Binomial coefficients
$$\binom{n}{k} := rac{n!}{(n-k)!k!}, \qquad k = 0, 1, \dots, n.$$

Exponential series

$$e^x = \sum_{n \ge 0} \frac{x^n}{n!}, \qquad x \in \mathbb{R}.$$
 (A.1)

Geometric sum

$$\sum_{k=0}^{n} r^{k} = \frac{1-r^{n+1}}{1-r}, \qquad r \neq 1.$$
(A.2)

456

Geometric series

$$\sum_{k \ge 0} r^k = \frac{1}{1 - r}, \qquad -1 < r < 1.$$
(A.3)

Differentiation of geometric series

$$\sum_{k \ge 1} kr^{k-1} = \frac{\partial}{\partial r} \sum_{k \ge 0} r^k = \frac{\partial}{\partial r} \frac{1}{1-r} = \frac{1}{(1-r)^2}, \qquad -1 < r < 1.$$
(A.4)

Binomial identities

$$\sum_{k=0}^{n} \binom{n}{k} a^{k} b^{n-k} = (a+b)^{n}.$$

$$\sum_{k=0}^{n} \binom{n}{k} = 2^{n}.$$
(A.5)

$$\sum\limits_{k=1}^n k {n \choose k} = n2^{n-1}.$$

$$egin{aligned} &\sum\limits_{k=0}^n k inom{n}{k} a^k b^{n-k} &= \sum\limits_{k=1}^n rac{n!}{(n-k)!(k-1)!} a^k b^{n-k} \ &= \sum\limits_{k=0}^{n-1} rac{n!}{(n-1-k)!k!} a^{k+1} b^{n-1-k} \ &= n \sum\limits_{k=0}^{n-1} inom{n-1}{k} a^{k+1} b^{n-1-k} \ &= na(a+b)^{n-1}, \qquad n \geqslant 1, \end{aligned}$$

$$\sum_{k=0}^{n} k \binom{n}{k} a^{k} b^{n-k} = a \frac{\partial}{\partial a} \sum_{k=0}^{n} \binom{n}{k} a^{k} b^{n-k}$$

$$= a \frac{\partial}{\partial a} (a+b)^{n}$$

$$= na(a+b)^{n-1}, \qquad n \ge 1.$$

Sums of integers

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$
 (A.6)

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$
(A.7)

Taylor expansion

$$(1+x)^{\alpha} = \sum_{k \ge 0} \frac{x^k}{k!} \alpha(\alpha - 1) \times \dots \times (\alpha - (k-1)).$$
(A.8)

Differential equation

The solution of
$$f'(t) = cf(t)$$
 is given by $f(t) = f(0)e^{ct}, \quad t \in \mathbb{R}_+.$ (A.9)

Solutions to the Exercises

Chapter 1 - Probability Background

Exercise 1.1

a) We have

$$\mathbb{P}(X<\infty)=\sum\limits_{k\geqslant 0}\mathbb{P}(X=k)=q\sum\limits_{k\geqslant 0}p^k=rac{q}{1-p},$$

hence

$$\mathbb{P}(X = \infty) = 1 - \mathbb{P}(X < \infty) = 1 - \frac{q}{1 - p} = \frac{1 - p - q}{1 - p}$$

When $0 \leq q < 1 - p$ we have $\mathbb{P}(X = \infty) > 0$, hence $\mathbb{E}[X] = \infty$.

b) We have $\mathbb{P}(Y < \infty) = \mathbb{P}(X < \infty) = 1$ and

$$\mathbb{E}[Y] = \sum\limits_{k \geqslant 0} r^k \mathbb{P}(Y=r^k) = \sum\limits_{k \geqslant 0} r^k \mathbb{P}(X=k) = (1-p) \sum\limits_{k \geqslant 0} (pr)^k,$$

hence $\mathbb{E}[Y] = (1-p)/(1-pr) < \infty$ when r < 1/p, and $\mathbb{E}[Y] = +\infty$ when $r \ge 1/p$.

Exercise 1.2 We write

$$Z = \sum\limits_{k=1}^N X_k$$

where $\mathbb{P}(N = n) = 1/6$, n = 1, 2, ..., 6, and X_k is a Bernoulli random variable with parameter 1/2, k = 1, 2, ..., 6. a) We have

$$\mathbb{E}[Z] \;=\; \mathbb{E}[\mathbb{E}[Z \mid N]] = \sum_{n=1}^{6} \mathbb{E}[Z \mid N = n] \mathbb{P}(N = n)$$

$$= \sum_{n=1}^{6} \mathbb{E} \left[\sum_{k=1}^{N} X_{k} \mid N = n \right] \mathbb{P}(N = n) = \sum_{n=1}^{6} \mathbb{E} \left[\sum_{k=1}^{n} X_{k} \right] \mathbb{P}(N = n)$$

$$= \sum_{n=1}^{6} \sum_{k=1}^{n} \mathbb{E}[X_{k}] \mathbb{P}(N = n) = \frac{1}{6} \sum_{n=1}^{6} \sum_{k=1}^{n} \frac{1}{2} = \frac{1}{2 \times 6} \sum_{n=1}^{6} n$$

$$= \frac{1}{2 \times 6} \times \frac{6 \times (6 + 1)}{2} = \frac{7}{4},$$
(B.1)

where we applied (A.6). Concerning the variance, we have

$$\begin{split} \mathbb{E}[Z^2] &= \mathbb{E}[\mathbb{E}[Z^2 \mid N]] \\ &= \sum_{n=1}^{6} \mathbb{E}[Z^2 \mid N = n] \mathbb{P}(N = n) \\ &= \sum_{n=1}^{6} \mathbb{E}\left[\left(\sum_{k=1}^{n} X_k\right)^2 \mid N = n\right] \mathbb{P}(N = n) \\ &= \sum_{n=1}^{6} \mathbb{E}\left[\sum_{k=1}^{n} X_k \sum_{l=1}^{n} X_l \mid N = n\right] \mathbb{P}(N = n) \\ &= \sum_{n=1}^{6} \mathbb{E}\left[\sum_{k,l=1}^{n} X_k X_l \mid N = n\right] \mathbb{P}(N = n) \\ &= \sum_{n=1}^{6} \mathbb{E}\left[\sum_{k=1}^{n} X_k^2 + \sum_{1 \le k \ne l \le n} X_k X_l\right] \mathbb{P}(N = n) \\ &= \sum_{n=1}^{6} \mathbb{E}\left[\sum_{k=1}^{n} X_k^2\right] \mathbb{P}(N = n) + \sum_{n=1}^{6} \mathbb{E}\left[\sum_{1 \le k \ne l \le n} X_k X_l\right] \mathbb{P}(N = n) \\ &= \sum_{n=1}^{6} \sum_{k=1}^{n} \mathbb{E}\left[X_k^2\right] \mathbb{P}(N = n) + \sum_{n=1}^{6} \sum_{1 \le k \ne l \le n} \mathbb{E}\left[X_k X_l\right] \mathbb{P}(N = n) \\ &= \sum_{n=1}^{6} \sum_{k=1}^{n} \mathbb{E}\left[X_k^2\right] \mathbb{P}(N = n) + \sum_{n=1}^{6} \sum_{1 \le k \ne l \le n} \mathbb{E}\left[X_k X_l\right] \mathbb{P}(N = n) \\ &= \frac{1}{2 \times 6} \sum_{n=1}^{6} n + \frac{1}{6 \times 2^2} \sum_{n=1}^{6} n(n-1) \\ &= \frac{1}{2 \times 6} \sum_{n=1}^{6} n + \frac{1}{6 \times 2^2} \sum_{n=1}^{6} n^2 - \frac{1}{6 \times 2^2} \sum_{n=1}^{6} n \end{split}$$

$$= \frac{1}{6 \times 2^2} \sum_{n=1}^{6} n + \frac{1}{6 \times 2^2} \sum_{n=1}^{6} n^2$$

= $\frac{1}{6 \times 2^2} \frac{6(6+1)}{2} + \frac{1}{6 \times 2^2} \frac{6(6+1)(2 \times 6+1)}{6} = \frac{14}{3},$ (B.2)

where we used (A.7) and (B.1), hence

$$\operatorname{Var}[Z] = \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2 = \frac{14}{3} - \frac{49}{16} = \frac{77}{48}.$$
 (B.3)

Using (1.6.10) and $\operatorname{Var}[X_k] = p(1-p)$ we could also write

$$egin{aligned} &\operatorname{Var}[Z \mid N = n] &= \operatorname{Var}\left[\sum\limits_{k=1}^{N} X_k \mid N = n
ight] \ &= \operatorname{Var}\left[\sum\limits_{k=1}^{n} X_k
ight] \ &= \sum\limits_{k=1}^{n} \operatorname{Var}[X_k] \ &= np(1-p), \end{aligned}$$

which implies

$$\mathbb{E}[Z^2 \mid N = n] = \operatorname{Var}[Z \mid N = n] + (\mathbb{E}[Z \mid N = n])^2 = np(1-p) + n^2p^2,$$

hence

$$egin{array}{rcl} \mathbb{E}\left[Z^2
ight] &=& \sum\limits_{k=1}^6 \mathbb{E}[Z^2 \mid N=n] \mathbb{P}(N=n) \ &=& rac{1}{6} \sum\limits_{k=1}^6 \mathbb{E}[Z^2 \mid N=n] \ &=& rac{1}{6} \sum\limits_{k=1}^6 (np-np^2+n^2p^2), \end{array}$$

and recovers (B.2).

b) Applying (1.3.1) with $B = \{Z = k\}$, we have

$$egin{aligned} \mathbb{P}(Z=l) &=& \mathbb{P}\left(\sum\limits_{k=1}^N X_k = l
ight) \ &=& \sum\limits_{n=1}^6 \mathbb{P}\left(\sum\limits_{k=1}^N X_k = l \ \middle| \ N=n
ight) \mathbb{P}(N=n) \ &=& \sum\limits_{n=1}^6 \mathbb{P}\left(\sum\limits_{k=1}^n X_k = l
ight) \mathbb{P}(N=n). \end{aligned}$$

Next, we notice that the deterministic summation $\sum_{k=1}^{n} X_k$ has a binomial distribution with parameter (n, 1/2), *i.e.*

$$\mathbb{P}\left(\sum_{k=1}^{n} X_k = l\right) = \left\{egin{array}{c} {n \ l} \left(rac{1}{2}
ight)^l \left(rac{1}{2}
ight)^{n-l} & ext{if } l = 0, 1, \dots, n, \ 0 & ext{if } l > n, \end{array}
ight.$$

which yields

$$egin{aligned} \mathbb{P}(Z=0) &=& \sum\limits_{n=1}^6 \mathbb{P}\left(\sum\limits_{k=1}^n X_k = 0
ight) \mathbb{P}(N=n) \ &=& rac{1}{6} \sum\limits_{n=1}^6 \left(rac{1}{2}
ight)^n \ &=& rac{1}{6 imes 2} \sum\limits_{n=0}^5 \left(rac{1}{2}
ight)^n \ &=& rac{1-(1/2)^6}{6}, \end{aligned}$$

and, for $l = 1, 2, \dots, 6$,

$$\mathbb{P}(Z=l) \;=\; \sum\limits_{n=1}^6 \mathbb{P}\left(\sum\limits_{k=1}^n X_k = l
ight) \mathbb{P}(N=n)$$

$$= \frac{1}{6} \sum_{n=l}^{6} \left(\frac{1}{2}\right)^{l} \left(\frac{1}{2}\right)^{n-l} \binom{n}{l}$$
$$= \frac{1}{6} \sum_{n=l}^{6} \binom{n}{l} \left(\frac{1}{2}\right)^{n}, \quad l = 1, 2, \dots, 6,$$

hence

$$\mathbb{P}(Z=l)=rac{1}{6}\sum\limits_{n=\max(1,l)}^{6}\binom{n}{l}\left(rac{1}{2}
ight)^n,\qquad l=0,1,\ldots,6.$$

Note that we have

$$\begin{split} \sum_{l=0}^{6} \mathbb{P}(Z=l) &= \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^{n} + \frac{1}{6} \sum_{l=1}^{6} \sum_{n=l}^{6} \binom{n}{l} \left(\frac{1}{2}\right)^{n} \\ &= \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^{n} + \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^{n} \sum_{l=1}^{n} \binom{n}{l} \\ &= \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^{n} \sum_{l=0}^{n} \binom{n}{l} \\ &= 1, \end{split}$$

where we used the identity $\sum_{l=0}^n \binom{n}{l} = (1+1)^n = 2^n$.

c) We have

$$\begin{split} \mathbb{E}[Z] &= \sum_{l=0}^{6} l \mathbb{P}(Z=l) = \frac{1}{6} \sum_{l=0}^{6} l \sum_{n=l}^{6} \left(\frac{1}{2}\right)^{n} \binom{n}{l} \\ &= \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^{n} \sum_{l=1}^{n} l \binom{n}{l} = \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^{n} \sum_{l=1}^{n} \frac{n!}{(n-l)!(l-1)!} \\ &= \frac{1}{6} \sum_{n=1}^{6} n \left(\frac{1}{2}\right)^{n} \sum_{l=0}^{n-1} \frac{(n-1)!}{(n-1-l)!l!} = \frac{1}{6} \sum_{n=1}^{6} n \left(\frac{1}{2}\right)^{n} \sum_{l=0}^{n-1} \binom{n-1}{l} \\ &= \frac{1}{6 \times 2} \sum_{n=1}^{6} n = \frac{1}{6 \times 2} \frac{6 \times (6+1)}{2} = \frac{7}{4}, \end{split}$$

ॳ62

which recovers (B.1), where we applied (A.5). We also have

$$\begin{split} \mathbb{E}[Z^2] &= \sum_{l=0}^{6} l^2 \mathbb{P}(Z=l) \\ &= \frac{1}{6} \sum_{l=0}^{6} l^2 \sum_{n=l}^{6} \left(\frac{1}{2}\right)^n \binom{n}{l} = \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^n \sum_{l=1}^{n} l^2 \binom{n}{l} \\ &= \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^n \sum_{l=1}^{n} \frac{(l-1+1)n!}{(n-l)!(l-1)!} \\ &= \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^n \sum_{l=1}^{n} \frac{(l-1)n!}{(n-l)!(l-1)!} + \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^n \sum_{l=1}^{n} \frac{n!}{(n-l)!(l-1)!} \\ &= \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^n \sum_{l=2}^{n} \frac{n!}{(n-l)!(l-2)!} + \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^n \sum_{l=1}^{n} \frac{n!}{(n-l)!(l-1)!} \\ &= \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^n n(n-1) \sum_{l=0}^{n-2} \frac{(n-2)!}{(n-2-l)!l!} + \frac{1}{6} \sum_{n=1}^{6} \left(\frac{1}{2}\right)^n \sum_{l=0}^{n-1} \frac{n!}{(n-1-l)!l!} \\ &= \frac{1}{6 \times 2^2} \sum_{n=1}^{6} n(n-1) + \frac{1}{6 \times 2} \sum_{n=1}^{6} n \\ &= \frac{1}{6 \times 2^2} \sum_{n=1}^{6} n^2 - \frac{1}{6 \times 2^2} \sum_{n=1}^{6} n + \frac{1}{6 \times 2} \sum_{n=1}^{6} n \\ &= \frac{1}{6 \times 2^2} \sum_{n=1}^{6} n^2 + \frac{1}{6 \times 2^2} \sum_{n=1}^{6} n \\ &= \frac{1}{6 \times 2^2} \sum_{n=1}^{6} n^2 + \frac{1}{6 \times 2^2} \sum_{n=1}^{6} n \\ &= \frac{1}{6 \times 2^2} \frac{6(6+1)(2 \times 6+1)}{6} + \frac{1}{6 \times 2^2} \frac{6(6+1)}{2} = \frac{14}{3}, \end{split}$$

which recovers (B.3).

Exercise 1.3

a) We assume that the sequence of Bernoulli trials is represented by a family $(X_k)_{k \ge 1}$ of independent Bernoulli random variables with $\operatorname{distribution}$ -0)

$$\mathbb{P}(X_k=1)=p, \qquad \mathbb{P}(X_k=0)=1-p, \qquad k\geqslant 1.$$

We have

$$Z = X_1 + X_2 + \dots + X_N = \sum_{k=1}^N X_k,$$

and, since $\mathbb{E}[X_k] = p$,

$$egin{aligned} \mathbb{E}[Z] &= \mathbb{E}\left[\sum\limits_{k=1}^N X_k
ight] = \sum\limits_{n \geqslant 0} \mathbb{E}\left[\sum\limits_{k=1}^n X_k
ight] \mathbb{P}(N=n) \ &= \sum\limits_{n \geqslant 0} \left(\sum\limits_{k=1}^n \mathbb{E}[X_k]
ight) \mathbb{P}(N=n) = p \sum\limits_{n \geqslant 0} n \mathbb{P}(N=n) = p \mathbb{E}[N]. \end{aligned}$$

Note that N need not have the Poisson distribution for the above equality to hold.

Next, the expectation of the Poisson random variable N with parameter $\lambda > 0$ is given as in (1.6.4) by

$$\mathbb{E}[N] = \sum_{n \ge 0} n \mathbb{P}(N=n) = e^{-\lambda} \sum_{n \ge 0} n \frac{\lambda^n}{n!} = e^{-\lambda} \sum_{n \ge 1} \frac{\lambda^n}{(n-1)!}$$
$$= \lambda e^{-\lambda} \sum_{n \ge 0} \frac{\lambda^n}{n!} = \lambda e^{-\lambda} e^{\lambda} = \lambda,$$
(B.4)

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

$$\mathbb{E}[Z] = p\lambda$$

Concerning the variance we have, since $\mathbb{E}[X_k^2] = p$,

$$egin{aligned} \mathbb{E}[Z^2] &= \mathbb{E}\left[\left(\sum\limits_{k=1}^N X_k
ight)^2
ight] \ &= \sum\limits_{n \geqslant 0} \mathbb{E}\left[\left(\sum\limits_{k=1}^n X_k
ight)^2
ight] \mathbb{P}(N=n) \ &= \sum\limits_{n \geqslant 0} \mathbb{E}\left[\left(\sum\limits_{k=1}^n X_k
ight)\left(\sum\limits_{l=1}^n X_l
ight)^2
ight] \mathbb{P}(N=n) \ &= \sum\limits_{n \geqslant 0} \mathbb{E}\left[\sum\limits_{k,l=1}^n X_k X_l
ight] \mathbb{P}(N=n) \end{aligned}$$

MH3512 AY19-20

ॳ64

$$= \sum_{n \ge 0} \mathbb{E} \left[\left(\sum_{k=1}^n (X_k)^2 + \sum_{1 \le k \ne l \le n} X_k X_l \right) \right] \mathbb{P}(N = n)$$

$$= \sum_{n \ge 0} \left(\mathbb{E} \left[\sum_{k=1}^n (X_k)^2 \right] + \sum_{1 \le k \ne l \le n} \mathbb{E}[X_k] \mathbb{E}[X_l] \right) \mathbb{P}(N = n)$$

$$= \sum_{n \ge 0} \left(np + n(n-1)p^2 \right) \mathbb{P}(N = n)$$

$$= p(1-p) \sum_{n \ge 0} n \mathbb{P}(N = n) + p^2 \sum_{n \ge 0} n^2 \mathbb{P}(N = n)$$

$$= p(1-p) \mathbb{E}[N] + p^2 \mathbb{E}[N^2].$$

Again, the above equality holds without requiring that \boldsymbol{N} has the Poisson distribution.

Next, we have

$$\begin{split} \mathbb{E}[N^2] &= \sum_{n \ge 0} n^2 \mathbb{P}(N=n) = \mathrm{e}^{-\lambda} \sum_{n \ge 0} n^2 \frac{\lambda^n}{n!} \\ &= \mathrm{e}^{-\lambda} \sum_{n \ge 1} n \frac{\lambda^n}{(n-1)!} = \mathrm{e}^{-\lambda} \sum_{n \ge 1} (n-1) \frac{\lambda^n}{(n-1)!} + \mathrm{e}^{-\lambda} \sum_{n \ge 1} \frac{\lambda^n}{(n-1)!} \\ &= \mathrm{e}^{-\lambda} \sum_{n \ge 2} \frac{\lambda^n}{(n-2)!} + \mathrm{e}^{-\lambda} \sum_{n \ge 1} \frac{\lambda^n}{(n-1)!} = \lambda^2 \mathrm{e}^{-\lambda} \sum_{n \ge 0} \frac{\lambda^n}{n!} + \lambda \mathrm{e}^{-\lambda} \sum_{n \ge 0} \frac{\lambda^n}{n!} \\ &= \lambda + \lambda^2, \end{split}$$

hence

$$\operatorname{Var}[N] = \mathbb{E}[N^2] - (\mathbb{E}[N])^2 = \lambda, \tag{B.5}$$

3465

and

$$\begin{aligned} \operatorname{Var}[Z] &= \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2 \\ &= p(1-p) \mathbb{E}[N] + p^2 \mathbb{E}[N^2] - (p \mathbb{E}[N])^2 \\ &= p(1-p) \mathbb{E}[N] + p^2 \operatorname{Var}[N] \\ &= \lambda p(1-p) + \lambda p^2 \\ &= p\lambda. \end{aligned}$$

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

$$\begin{split} \mathbb{P}(Z=l) &= \sum_{n \geq 0} \mathbb{P}\left(\sum_{k=1}^{N} X_{k} = l \mid N = n\right) \mathbb{P}(N=n) \\ &= \sum_{n \geq 0} \mathbb{P}\left(\sum_{k=1}^{n} X_{k} = l\right) \mathbb{P}(N=n) \\ &= e^{-\lambda} \sum_{n \geq l} \binom{n}{l} p^{l} (1-p)^{n-l} \frac{\lambda^{n}}{n!} = \frac{p^{l}}{l!} e^{-\lambda} \sum_{n \geq l} \frac{1}{(n-l)!} (1-p)^{n-l} \lambda^{n} \\ &= \frac{(\lambda p)^{l}}{l!} e^{-\lambda} \sum_{n \geq 0} (1-p)^{n} \frac{\lambda^{n}}{n!} = \frac{(\lambda p)^{l}}{l!} e^{-\lambda} e^{(1-p)\lambda} = \frac{(\lambda p)^{l}}{l!} e^{-p\lambda}, \end{split}$$

hence Z has a Poisson distribution with parameter $p\lambda$. This result is known as *thinning* of the Poisson distribution with parameter $\lambda > 0$.

c) From Question (b), Z is a Poisson random variable with parameter $p\lambda$, hence from (B.4) and (B.5) we have $\mathbb{E}[Z] = \operatorname{Var}[Z] = p\lambda$.

Exercise 1.4 Using the relation $\mathbb{P}(X < Y) = \lambda/(\lambda + \mu)$ that follows from (1.5.9), we can compute the following conditional expectation in the same way as in (1.6.6) of Lemma 1.4:

$$\begin{split} \mathbb{E}[\min(X,Y) \mid X < Y] &= \mathbb{E}[X \mid X < Y] = \frac{1}{\mathbb{P}(X < Y)} \mathbb{E}\left[X \mathbb{1}_{\{X < Y\}}\right] \\ &= \mu(\lambda + \mu) \int_0^\infty e^{-\mu y} \int_0^y x e^{-\lambda x} dx dy \\ &= (\lambda + \mu) \frac{\mu}{\lambda^2} \int_0^\infty e^{-\mu y} (1 - (1 + \lambda y) e^{-\lambda y}) dy \\ &= (\lambda + \mu) \frac{\mu}{\lambda^2} \left(\int_0^\infty e^{-\mu y} dy - \int_0^\infty e^{-\mu y} e^{-\lambda y} dy - \lambda \int_0^\infty e^{-\mu y} y e^{-\lambda y} \right) dy \Big) \\ &= (\lambda + \mu) \frac{\mu}{\lambda^2} \left(\frac{1}{\mu} - \frac{1}{\lambda + \mu} - \frac{\lambda}{(\lambda + \mu)^2} \right) \\ &= (\lambda + \mu) \frac{\mu}{\lambda^2} \left(\frac{(\lambda + \mu)^2}{\mu(\lambda + \mu)^2} - \frac{\mu(\lambda + \mu)}{\mu(\lambda + \mu)^2} - \frac{\lambda \mu}{\mu(\lambda + \mu)^2} \right) \end{split}$$

$$= (\lambda + \mu) \frac{\mu}{\lambda^2} \left(\frac{(\lambda + \mu)^2 - \mu(\lambda + \mu) - \lambda\mu}{\mu(\lambda + \mu)^2} \right)$$

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

Note that the random variable $\min(X, Y)$ is actually independent of the event $\{X < Y\}$. Indeed, for any a > 0 we have

$$\begin{split} \mathbb{P}(\min(X,Y) > a \text{ and } X < Y) &= \lambda \mu \int_{a}^{\infty} e^{-\mu y} \int_{a}^{y} e^{-\lambda x} dx dy \\ &= \mu \int_{a}^{\infty} e^{-\mu y} (e^{-\lambda a} - e^{-\lambda y}) dy \\ &= \mu e^{-\lambda a} \int_{a}^{\infty} e^{-\mu y} dy - \mu \int_{a}^{\infty} e^{-(\lambda+\mu)y} dy \\ &= \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)a} \\ &= \mathbb{P}(X < Y) \mathbb{P}(\min(X,Y) > a). \end{split}$$

Exercise 1.5 Since U is uniformly distributed given L over the interval [0, L], we have

$$f_{U|L=y}(x)=rac{1}{y}\mathbb{1}_{[0,y]}(x), \qquad x\in\mathbb{R}, \hspace{1em} y>0,$$

hence by the definition (1.5.7) of the conditional density $f_{U|L=y}(x)$ we have

$$f_{(U,L)}(x,y) = f_{U|L=y}(x)f_L(y) = \frac{1}{y}\mathbb{1}_{[0,y]}(x)ye^{-y}\mathbb{1}_{[0,\infty)}(y) = \mathbb{1}_{[0,y]}(x)\mathbb{1}_{[0,\infty)}(y)e^{-y}.$$
(B.6)

Next, we want to determine the density function $(x, y) \mapsto f_{(U, L-U)}(x, y)$. For this, for all bounded functions $h : \mathbb{R}^2 \longrightarrow \mathbb{R}$, we rewrite the expectation

$$\mathbb{E}[h(U,L-U)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x,z) f_{(U,L-U)}(x,z) dx dz ext{ MH3512 AY19-20}$$

5468

(B.9)

MH3512 AY19-20

$$\mathbb{E}[h(U, L-U)] = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} h(x, y-x) f_{(U,L)}(x, y) dy \right) dx$$

=
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, z) f_{(U,L)}(x, x+z) dx dz, \qquad (B.8)$$

using the change of variable (x, z) = (x, y - x). Identifying (B.7) with (B.8) yields the relation

$$f_{(U,L-U)}(x,z) = f_{(U,L)}(x,x+z),$$

hence from (B.6) we get

$$egin{aligned} f_{(U,L-U)}(x,z) &=& f_{(U,L)}(x,x+z) \ &=& \mathbbm{1}_{[0,x+z]}(x)\mathbbm{1}_{[0,\infty)}(x+z)\mathrm{e}^{-x-z} \ &=& \mathbbm{1}_{[0,\infty)}(x)\mathbbm{1}_{[0,\infty)}(z)\mathrm{e}^{-x-z}, \end{aligned}$$

since

We refer to Theorem 12.7 page 92 of [JP00] for the general version of the above change of variable formula. Exercise 1.6

a) Assuming that \boldsymbol{X} and \boldsymbol{Y} are independent Poisson random variables with parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, we have

$$\mathbb{P}(X+Y=n) = \sum_{k=0}^{n} \mathbb{P}(X=k \text{ and } X+Y=n)$$

$$= \sum_{k=0}^{n} \mathbb{P}(X=k \text{ and } Y=n-k)$$

$$= \sum_{k=0}^{n} \mathbb{P}(X=k) \mathbb{P}(Y=n-k) = e^{-\lambda-\mu} \sum_{k=0}^{n} \frac{\lambda^{k}}{k!} \frac{\mu^{n-k}}{(n-k)!}$$

$$= e^{-\lambda-\mu} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \lambda^{k} \mu^{n-k} = e^{-\lambda-\mu} \frac{(\lambda+\mu)^{n}}{n!},$$

as
hence X + Y has a Poisson distribution with parameter $\lambda + \mu$. This result can be recovered from the probability generating function (1.7.2) of the Poisson distribution and Relation (1.7.6), as

$$G_{X+Y}(s)=G_X(s)G_Y(s)=\mathrm{e}^{\lambda(s-1)}\mathrm{e}^{\mu(s-1)}=\mathrm{e}^{(\lambda+\mu)(s-1)}, \hspace{1em} -1\leqslant s\leqslant 1.$$

b) We have

$$\mathbb{P}(X = k \mid X + Y = n) = \frac{\mathbb{P}(X = k \text{ and } X + Y = n)}{\mathbb{P}(X + Y = n)} \\
= \frac{\mathbb{P}(X = k \text{ and } Y = n - k)}{\mathbb{P}(X + Y = n)} = \frac{\mathbb{P}(X = k)\mathbb{P}(Y = n - k)}{\mathbb{P}(X + Y = n)} \\
= e^{-\lambda} \frac{\lambda^{k}}{k!} e^{-\mu} \frac{\mu^{n-k}}{(n-k)!} \left(e^{-\lambda - \mu} \frac{(\lambda + \mu)^{n}}{n!} \right)^{-1} \\
= \binom{n}{k} \left(\frac{\lambda}{\lambda + \mu} \right)^{k} \left(\frac{\mu}{\lambda + \mu} \right)^{n-k},$$
(B.10)

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

$$f_\Lambda(x)= heta\mathbb{1}_{[0,\infty)}(x)\mathrm{e}^{- heta x},\qquad x\in\mathbb{R},$$

satisfying

$$\mathbb{P}(\Lambda \leqslant x) = \int_{-\infty}^x f_\Lambda(y) dy = \int_0^x f_\Lambda(y) dy, \qquad x \in \mathbb{R},$$

and

$$d\mathbb{P}(\Lambda\leqslant x)=f_{\Lambda}(x)dx,$$

we have

$$egin{array}{rcl} \mathbb{P}(X=k) &=& \int_0^\infty \mathbb{P}(X=k \mid \Lambda=x) d\mathbb{P}(\Lambda \leqslant x) \ &=& \int_0^\infty \mathbb{P}(X=k \mid \Lambda=x) f_\Lambda(x) dx \end{array}$$

$$egin{aligned} &=& heta \int_0^\infty rac{x^k}{k!} \mathrm{e}^{-(heta+1)x} dx \ &=& rac{ heta}{(heta+1)^{k+1}} \int_0^\infty rac{y^k}{k!} \mathrm{e}^{-y} dy \ &=& rac{ heta}{(heta+1)^{k+1}} \ &=& \left(1-rac{1}{ heta+1}
ight) \left(rac{1}{ heta+1}
ight)^k, \end{aligned}$$

i.e. X and Y are independent random variables with same geometric distribution with parameter $p = 1/(\theta + 1)$. Therefore X + Y has a negative binomial distribution with parameter $(r, p) = (2, 1/(\theta + 1))$, cf. (1.5.12), *i.e.*,

$$\mathbb{P}(X+Y=k)=(k+1)\left(1-rac{1}{ heta+1}
ight)^2\left(rac{1}{ heta+1}
ight)^k,\qquad k\in\mathbb{N}.$$

Hence we have

$$\mathbb{P}(X = k \mid X + Y = n) = \frac{\mathbb{P}(X = k \text{ and } X + Y = n)}{\mathbb{P}(X + Y = n)}$$

$$= \frac{\mathbb{P}(X = k \text{ and } Y = n - k)}{\mathbb{P}(X + Y = n)}$$

$$= \frac{\mathbb{P}(X = k)\mathbb{P}(Y = n - k)}{\mathbb{P}(X + Y = n)}$$

$$= \left(1 - \frac{1}{\theta + 1}\right) \left(\frac{1}{\theta + 1}\right)^k \left(1 - \frac{1}{\theta + 1}\right) \left(\frac{1}{\theta + 1}\right)^{n-k} \frac{1}{n+1} \left(1 - \frac{1}{\theta + 1}\right)^{-2} \left(\frac{1}{\theta + 1}\right)^{-n}$$

$$= \frac{1}{n+1}, \quad k = 0, 1, \dots, n,$$

which shows that the distribution of X given X + Y = n is the discrete uniform distribution on $\{0, 1, \ldots, n\}$.

ॳ470

d) In case **X** and **Y** have the same parameter, *i.e.* $\lambda = \mu$, we have

$$\mathbb{P}(X=k\mid X+Y=n)=inom{n}{k}rac{1}{2^n}, \hspace{1em} k=0,1,\ldots,n,$$

which becomes independent of λ . Hence, when λ is represented by a random variable Λ with probability density $x \mapsto f_{\Lambda}(x)$ on \mathbb{R}_+ , from (B.10) we get

$$egin{aligned} \mathbb{P}(X=k\mid X+Y=n) &= \int_0^\infty \mathbb{P}(X=k\mid X+Y=n ext{ and } \Lambda=x) d\mathbb{P}(\Lambda\leqslant x) \ &= \int_0^\infty \mathbb{P}(X=k\mid X+Y=n ext{ and } \Lambda=x) f_\Lambda(x) dx \ &= \int_0^\infty \mathbb{P}(X=k\mid X+Y=n) f_\Lambda(x) dx \ &= \mathbb{P}(X=k\mid X+Y=n) \int_0^\infty f_\Lambda(x) dx \ &= \mathbb{P}(X=k\mid X+Y=n) = inom{n}{k} rac{1}{2^n}, \quad k=0,1,\ldots,n. \end{aligned}$$

This relation holds regardless of the expression of the probability density f_{Λ} , and in particular when Λ has an exponential distribution with parameter $\theta > 0$.

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

$$\mathbb{P}(C_1=R)=\mathbb{P}(C_1=G)=rac{1}{2},$$

and

$$\mathbb{P}(C_2 = R \text{ and } C_1 = R) = \frac{2}{3}, \qquad \mathbb{P}(C_2 = R \text{ and } C_1 = G) = \frac{1}{3}.$$

On the other hand, we have

$$\mathbb{P}(C_2 = R) = \mathbb{P}(C_2 = R \text{ and } C_1 = R) + \mathbb{P}(C_2 = R \text{ and } C_1 = G)$$

= $\mathbb{P}(C_2 = R \mid C_1 = R)\mathbb{P}(C_1 = R) + \mathbb{P}(C_2 = R \mid C_1 = G)\mathbb{P}(C_1 = G)$
= $\frac{2}{3} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{2} = \frac{1}{2},$

471

$$\mathbb{P}(C_2 = G) = \mathbb{P}(C_2 = G \text{ and } C_1 = R) + \mathbb{P}(C_2 = G \text{ and } C_1 = G) \\
= \mathbb{P}(C_2 = G \mid C_1 = R) \mathbb{P}(C_1 = R) + \mathbb{P}(C_2 = G \mid C_1 = G) \mathbb{P}(C_1 = G) \\
= \frac{1}{3} \times \frac{1}{2} + \frac{2}{3} \times \frac{1}{2} = \frac{1}{2}.$$

Finally, the probability we wish to compute is

$$\mathbb{P}(C_1 = R \mid C_2 = R) = \frac{\mathbb{P}(C_1 = R \text{ and } C_2 = R)}{\mathbb{P}(C_2 = R)}$$

= $\mathbb{P}(C_2 = R \mid C_1 = R) \frac{\mathbb{P}(C_1 = R)}{\mathbb{P}(C_2 = R)} = \frac{2}{3} \times \frac{1/2}{1/2} = \frac{2}{3}.$

Interestingly, we note that although the probabilities of finding a red or green pen at the second stage remain the same as at the first stage $(\mathbb{P}(C_2 = G) = \mathbb{P}(C_1 = G) = 1/2)$, the result obtained at the second stage does provide some information on the outcome at the first stage.

Exercise 1.8

a) The probability that the system operates is

$$\mathbb{P}(X \geqslant 2) = {3 \choose 2} p^2 (1-p) + p^3 = 3p^2 - 2p^3,$$

where X is a binomial random variable with parameter (3, p).

b) The probability that the system operates is

$$egin{aligned} &\int_{0}^{1} \mathbb{P}(X \geqslant 2 \mid p = x) d\mathbb{P}(p \leqslant x) \; = \; \int_{0}^{1} \mathbb{P}(X \geqslant 2 \mid p = x) dx \ &= \; \int_{0}^{1} (3x^2 - 2x^3) dx \ &= \; 3 \int_{0}^{1} x^2 dx - 2 \int_{0}^{1} x^3 dx \ &= \; rac{1}{2}, \end{aligned}$$

similarly to Exercise 1.6-(b).

Chapter 2 - Gambling Problems

Exercise 2.1

a) By first step analysis we have

$$f(k) = (1 - 2r)f(k) + rf(k + 1) + rf(k - 1),$$

which yields the equation

$$f(k) = \frac{1}{2}f(k+1) + \frac{1}{2}f(k-1), \qquad 1 \le k \le S - 1,$$
(B.11)

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 Section 2.2 we know that the general solution of (B.11) has the form

$$f(k)=C_1+C_2k, \qquad k=0,1,\dots,S$$

and after taking into account the boundary conditions we find

$$f(k)=rac{S-k}{S}, \qquad k=0,1,\ldots,S.$$

Intuitively, the game is fair to both players for all values of r, so that the probability of ruin should be the same for all $r \in (0, 1/2)$, which is indeed the case.

c) By first step analysis we have

$$\begin{split} h(k) &= (1-2r)(1+h(k)) + r(1+h(k+1)) + r(1+h(k-1)) \\ &= 1+(1-2r)h(k) + rh(k+1) + rh(k-1), \end{split}$$

hence the equation

$$h(k)=rac{1}{2r}+rac{1}{2}h(k+1)+rac{1}{2}h(k-1), \qquad 1\leqslant k\leqslant S-1,$$

with the boundary conditions h(0) = 0 and h(S) = 0, which is identical to (2.3.6) by changing h(k) into $2r \times h(k)^{4/14}$ (2.3.6). We note that r should be strictly positive, otherwise if r = 0 we will have $h(k) = \infty$ for all k = 1, 2, ..., S - 1 and h(0) = h(S) = 0.

d) After trying a solution of the form $h(k) = Ck^2$ we find

$$Ck^2 = rac{1}{2r} + rac{1}{2}C(k+1)^2 + rac{1}{2}C(k-1)^2,$$

hence C should be equal to C = -1/(2r), hence $k \mapsto -k^2/(2r)$ is a particular solution.

e) Given the hint, the general solution has the form

$$h(k) = C_1 + C_2 k - rac{k^2}{2r}, \qquad k = 0, 1, \dots, S,$$

which gives

$$h(k) = \frac{k(S-k)}{2r}, \qquad k = 0, 1, \dots, S,$$
 (B.12)

after taking into account the boundary conditions.

Remark. The inclusion of draws changes the standard discrete time step length 1 into 1/(2r), which can be interpreted as the mean $\mathbb{E}[\tau] = 1/(2r)$ of the geometric random variable τ representing the time k spent by remaining at the same state, with

$$\mathbb{P}(au=k)=2r(1-2r)^{k-1},\qquad k\geqslant 1.$$

f) Starting from any state $(k) \in \{1, 2, \dots, S-1\}$, the mean duration goes to infinity when r goes to zero.

When r goes to 0 the probability 1 - 2r of a draw increases, therefore the game should take longer. Hence the above answer is compatible with intuition.

Exercise 2.2

a) Taking a = m and b = S in (2.3.18), we find the probability

$$\frac{1 - (p/q)^{S-k}}{1 - (p/q)^{S-m}} = \frac{(q/p)^S - (q/p)^k}{(q/p)^S - (q/p)^m}.$$
(B.13)

b) Taking a = 0 and b = m in (2.3.18), we find the probability

$$1 - \frac{1 - (p/q)^{m-k}}{1 - (p/q)^m} = \frac{1 - (q/p)^k}{1 - (q/p)^m}.$$
(B.14)

c) Taking k = m + 1 in (B.13), we find that the probability of returning to (m) after starting from (m+1) is

$$rac{(q/p)^S-(q/p)^{m+1}}{(q/p)^S-(q/p)^m}.$$

Similarly, taking k = m - 1 in (B.14), we find that the probability of returning to (m) after starting from (m - 1) is

$$\frac{1-(q/p)^{m-1}}{1-(q/p)^m}$$

Therefore, by first step analysis we find that the probability of coming back to state (m) in finite time after starting from $X_0 = m$ is

$$p\frac{1 - (p/q)^{S-(m+1)}}{1 - (p/q)^{S-m}} + q\frac{1 - (q/p)^{m-1}}{1 - (q/p)^m}$$

$$= \frac{p(1 - (p/q)^{S-(m+1)})(1 - (q/p)^m) + q(1 - (q/p)^{m-1})(1 - (p/q)^{S-m})}{(1 - (p/q)^{S-m})(1 - (q/p)^m)}$$

$$= \frac{(p - q(p/q)^{S-m})(1 - (q/p)^m) + (q - p(q/p)^m)(1 - (p/q)^{S-m})}{(1 - (p/q)^{S-m})(1 - (q/p)^m)}$$

$$= \frac{1 - 2p(q/p)^m - 2q(p/q)^{S-m} + (p/q)^{S-2m}}{(1 - (p/q)^{S-m})(1 - (q/p)^m)}$$
MH3512 AY19-20

 $\circ 475$

$$=\frac{1-2p(q/p)^m-2q(p/q)^{S-m}+(p/q)^{S-2m}}{1-(q/p)^m-(p/q)^{S-m}+(p/q)^{S-2m}}, \qquad 1 \le m \le S-1.$$

When S = 2 and m = 1 we find that the above probability is zero, as expected.

d) Using (2.3.11) we can similarly compute the mean time to either come back to m or reach any of the two boundaries $\{0, S\}$, whichever comes first, using first step analysis, as follows:

$$egin{aligned} &1+rac{p}{q-p}\left(1-(S-m)rac{1-q/p}{1-(q/p)^{S-m}}
ight)+rac{q}{q-p}\left(m-1-mrac{1-(q/p)^{m-1}}{1-(q/p)^m}
ight)\ &=-rac{p(S-m)}{q-p}rac{1-q/p}{1-(q/p)^{S-m}}+rac{mq}{q-p}rac{(q/p)^m-(q/p)^{m-1}}{(q/p)^m-1}, &1\leqslant m\leqslant S-1. \end{aligned}$$

When S = 2 and m = 1 we find that the above expression equals 1 as expected.

e) When p = q = 1/2 we would find

$$rac{1}{2}\left(rac{S-m-1}{S-m}
ight)+rac{1}{2}\left(rac{m-1}{m}
ight)=1-rac{S}{2m(S-m)},$$

for the probability of coming back to state (m) in finite time after starting from $X_0 = m$, which gives 1 - 2/S when S = 2m, $m \ge 1$. Similarly, the mean time to either come back to m of reach any of the two boundaries $\{0, S\}$, whichever comes first, is

$$1 + \frac{1}{2}(S - m - 1) + \frac{1}{2}(m - 1) = \frac{S}{2}$$

which does not depend on $m = 1, 2, \ldots, S - 1$.

Exercise 2.3

a) Starting from state (k) we can reach (k+1) in one time step by going up with probability p, or we can go down to (k-1) with probability q, in which case we have to reach (k+1) without hitting state (0) in two stages:

i) first, reach
$$(k)$$
 from $(k-1)$, with probability $\mathbb{P}(\tau_k < \tau_0 \mid X_0 = k-1)$,

ii) then, reach k+1 from k with probability $\mathbb{P}(\tau_{k+1} < \tau_0 \mid X_0 = k)$. MH3512 AY19-20

This gives

$$egin{array}{rll} p_k &=& \mathbb{P}(au_{k+1} < au_0 \mid X_0 = k) \ &=& p + q \mathbb{P}(au_k < au_0 \mid X_0 = k - 1) \mathbb{P}(au_{k+1} < au_0 \mid X_0 = k) \ &=& p + q p_{k-1} p_k, \qquad k = 1, 2, \dots, S-1. \end{array}$$

b) We need to check that (2.3.21) is satisfied when p_k is given by

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

Starting from the right hand side of (2.3.21), this means checking the following:

$$p + q \frac{1 - (q/p)^{k-1}}{1 - (q/p)^k} \frac{1 - (q/p)^k}{1 - (q/p)^{k+1}} = p + q \frac{1 - (q/p)^{k-1}}{1 - (q/p)^{k+1}}$$
$$= \frac{p(1 - (q/p)^{k+1}) + q(1 - (q/p)^{k-1})}{1 - (q/p)^{k+1}}$$
$$= \frac{p + q - p(q/p)^{k+1} - q(q/p)^{k-1}}{1 - (q/p)^{k+1}}$$
$$= \frac{p + q - q(q/p)^k - p(q/p)^k}{1 - (q/p)^{k+1}}$$
$$= \frac{1 - (q/p)^k}{1 - (q/p)^{k+1}}.$$

c) We have

$$egin{aligned} \mathbb{P}(au_S < au_0 \mid X_0 = k) \ &= \mathbb{P}(au_{k+1} < au_0 \mid X_0 = k) \ & imes \mathbb{P}(au_{k+2} < au_0 \mid X_0 = k+1) imes \cdots imes \mathbb{P}(au_S < au_0 \mid X_0 = S-1) \ &= \prod_{l=k}^{S-1} p_l = \prod_{l=k}^{S-1} rac{1-(q/p)^l}{1-(q/p)^{l+1}} = rac{1-(q/p)^k}{1-(q/p)^S}, \quad k=0,1,\dots,S, \end{aligned}$$

MH3512 AY19-20

)

d) In the symmetric case p = q = 1/2, the same reasoning as in part (a) above shows that p_k should satisfy the equation

$$p_k = rac{1}{2} + rac{1}{2} p_{k-1} p_k, \qquad k = 1, 2, \dots, S-1,$$

which rewrites as

$$p_k(2-p_{k-1})=1, \qquad k=1,2,\dots,S-1.$$

We check that the above equation is satisfied by

$$p_k := rac{k}{k+1}, \qquad k=1,2,\ldots,S-1,$$

hence we find again

$$\begin{split} \mathbb{P}(\tau_{S} < \tau_{0} \mid X_{0} = k) \\ &= \mathbb{P}(\tau_{k+1} < \tau_{0} \mid X_{0} = k) \\ &\times \mathbb{P}(\tau_{k+2} < \tau_{0} \mid X_{0} = k+1) \times \dots \times \mathbb{P}(\tau_{S} < \tau_{0} \mid X_{0} = S-1) \\ &= \prod_{l=k}^{S-1} p_{l} \\ &= p_{k} p_{k+1} p_{k+2} \cdots p_{N-1} \\ &= \prod_{l=k}^{S-1} \frac{l}{l+1} \\ &= \frac{k}{S}, \quad k = 0, 1, \dots, S, \end{split}$$

showing that

$$\mathbb{P}(au_0< au_S\mid X_0=k)=1-rac{k}{S},\qquad k=0,1,\dots,S,$$

which recovers (2.2.12).

a) We have

 $f(0)=\mathbb{P}(R_A\mid X_0=0)=q+pq.$

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

b) We have

and

$$egin{array}{rll} f(1) &=& \mathbb{P}(R_A \mid X_0 = 1) \ &=& p\mathbb{P}(R_A \mid X_0 = 2) + q\mathbb{P}(R_A \mid X_0 = 0) \ &=& q^2 + pq^2. \end{array}$$

Exercise 2.5

a) Letting $\varepsilon = r - 1$, *i.e.* $r = 1 + \varepsilon$, we have

$$egin{aligned} h_S(k) &= rac{1}{q-p} \left(k-Srac{1-r^k}{1-r^S}
ight) = rac{1}{parepsilon} \left(k-Srac{1-(1+arepsilon)^k}{1-(1+arepsilon)^S}
ight) \ &= rac{1}{parepsilon} \left(k-Srac{\sum\limits_{i=1}^k inom{k}_i arepsilon^{arepsilon}}{\sum\limits_{i=1}^S inom{S}_i arepsilon^{arepsilon}}
ight) = rac{1}{parepsilon} \left(rac{k\sum\limits_{i=1}^S inom{S}_i arepsilon^S}{\sum\limits_{i=1}^S inom{S}_i arepsilon^{arepsilon}}
ight) \\ &= rac{1}{p} \left(rac{kS(S-1)/2 - Sk(k-1)/2 + k\sum\limits_{i=3}^S inom{S}_i arepsilon^S}{i} arepsilon^{arepsilon-2} - S\sum\limits_{i=3}^k inom{k}_i arepsilon^{arepsilon-2}}
ight) \end{matrix}$$

hence

$$\lim_{r \to 1} \frac{1}{q-p} \left(k - S \frac{1-r^k}{1-r^S} \right) = \lim_{\varepsilon \to 0} \frac{1}{p\varepsilon} \left(k - S \frac{1-(1+\varepsilon)^k}{1-(1+\varepsilon)^S} \right) = k(S-k), \tag{B.15}$$

which recovers (2.3.17).

b) Letting $\varepsilon = 1 - r$, we have

$$f_S(k) = rac{r^S - r^k}{r^S - 1} = rac{(1 + arepsilon)^S - (1 + arepsilon)^k}{(1 + arepsilon)^S - 1}
onumber \ = rac{\sum\limits_{i=0}^S inom{S}_i arepsilon^i - \sum\limits_{i=0}^k inom{k}_i arepsilon^i}{\sum\limits_{i=0}^S inom{S}_i arepsilon^i - 1} = rac{S - k + arepsilon^2 \sum\limits_{i=2}^k inom{k}_i arepsilon^{i-2} - arepsilon^2 \sum\limits_{i=2}^S inom{S}_i arepsilon^{i-2}}{S + \epsilon^2 \sum\limits_{i=2}^S inom{S}_i arepsilon^{i-2}} = rac{S - k + arepsilon^2 \sum\limits_{i=2}^k inom{k}_i arepsilon^{i-2} - arepsilon^2 \sum\limits_{i=2}^S inom{S}_i arepsilon^{i-2}}{S + \epsilon^2 \sum\limits_{i=2}^S inom{S}_i arepsilon^{i-2}}{S + \epsilon^2 \sum\limits_{i=2}^S inom{S}_i arepsilon^{i-2}} = rac{S - k + arepsilon^2 \sum\limits_{i=2}^K inom{S}_i arepsilon^{i-2}}{S + \epsilon^2 \sum\limits_{i=2}^S inom{S}_i arepsilon^{i-2}}{S + \epsilon^2 \sum\limits_{i=2}^S inom{S}_i arepsilon^{i-2}}{S - \epsilon^2$$

hence

$$\lim_{r o 1}rac{r^S-r^k}{r^S-1}=\lim_{arepsilon o 0}rac{(1+arepsilon)^S-(1+arepsilon)^k}{(1+arepsilon)^S-1}=rac{S-k}{S},$$

which recovers (2.2.21).

Exercise 2.6 By first step analysis we now have

$$f(k) = (1 - \alpha - \beta)f(k) + \alpha f(k+1) + \beta f(k-1),$$

which yields

$$f(k)=rac{lpha}{lpha+eta}f(k+1)+rac{eta}{lpha+eta}f(k-1),\qquad 1\leqslant k\leqslant S-1,$$

with f(0) = 1 and f(S) = 0. Therefore this non-symmetric problem with draw can be reduced to a non-symmetric gambling process without draw and probabilities $p = \alpha/(\alpha + \beta)$ and $q = \beta/(\alpha + \beta)$, and solution

$$f(k)=rac{(eta/lpha)^k-(eta/lpha)^S}{1-(eta/lpha)^S},\qquad k=0,1,\ldots,S,$$

obtained by (2.2.11), with $\alpha \neq \beta$. For the mean game duration, by first step analysis we find the equation

$$\begin{split} h(k) &= (1-\alpha-\beta)(1+h(k)) + \alpha(1+h(k+1)) + \beta(1+h(k-1)) \\ &= 1+(1-\alpha-\beta)h(k) + \alpha h(k+1) + \beta h(k-1), \end{split}$$

i.e.

$$h(k)=rac{1}{lpha+eta}+rac{lpha}{lpha+eta}h(k+1)+rac{eta}{lpha+eta}h(k-1), \qquad 1\leqslant k\leqslant S-1,$$

or

$$(\alpha + \beta)h(k) = 1 + p(\alpha + \beta)h(k+1) + q(\alpha + \beta)h(k-1),$$

with $p = \alpha/(\alpha + \beta)$ and $q = \beta/(\alpha + \beta)$ which, using (2.3.11), shows that

$$h(k) = \frac{1}{\beta - \alpha} \left(k - S \frac{1 - (\beta/\alpha)^k}{1 - (\beta/\alpha)^S} \right), \qquad k = 0, 1, \dots, S,$$
(B.16)

when $\alpha \neq \beta$. As in Question (e) of Exercise 2.1, the inclusion of a draw changes the discrete time step length from 1 to $1/(\alpha + \beta)$, which is the mean $\mathbb{E}[\tau]$ of a geometric random variable τ with distribution

$$\mathbb{P}(au=k)=(lpha+eta)(1-lpha-eta)^{k-1}, \qquad k\geqslant 1.$$

When α tends to β , the result of Exercise 2.1 can be recovered as on page 479. Indeed, the limit of (B.16) as $\alpha \rightarrow \beta$ is given as in (B.15) by

$$h(k)=rac{k(S-k)}{2eta},\qquad k=0,1,\ldots,S,$$

which is consistent with (B.12).

Problem 2.7

a) We have

$$g(k) = pg(k+1) + qg(k-1), \qquad k = 1, 2, \dots, S-1,$$
 (B.17)

with

$$g(0) = pg(1) + qg(0)$$
 (B.18)

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.17) and (B.18) 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, ..., S.

MH3512 AY19-20

() 481

c) We have

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

with

$$h(0) = 1 + ph(1) + qh(0)$$

for k = 0, and the boundary condition h(S) = 0.

d) Case $p \neq q$. The solution of the homogeneous equation

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

has the form

$$h(k) = C_1 + C_2(q/p)^k, \qquad k = 1, 2, \dots, S-1,$$

and we can check that $k \mapsto k/(p-q)$ is a particular solution. Hence the general solution of (B.19) has the form

$$h(k) = rac{k}{q-p} + C_1 + C_2 (q/p)^k, \qquad k = 0, 1, \dots, S,$$

with

$$\left\{egin{array}{l} 0=h(S)=rac{S}{q-p}+C_1+C_2(q/p)^S,\ ph(0)=p(C_1+C_2)=1+ph(1)=1+p\left(rac{1}{q-p}+C_1+C_2rac{q}{p}
ight), \end{array}
ight.$$

which yields

$$\left\{ egin{array}{l} C_1 = q rac{(q/p)^S}{(p-q)^2} - rac{S}{q-p}, \ C_2 = -rac{q}{(p-q)^2}, \end{array}
ight.$$

and

$$h(k) = \mathbb{E}[T_S \mid X_0 = k] = rac{S-k}{p-q} + rac{q}{(p-q)^2} ((q/p)^S - (q/p)^k),$$

 $k=0,1,\ldots,S$.

<u>Case p = q = 1/2</u>. The solution of the homogeneous equation is given by

$$h(k) = C_1 + C_2 k, \qquad k = 1, 2, \dots, S-1,$$

and the general solution to (B.19) has the form

$$h(k) = -k^2 + C_1 + C_2 k, \qquad k = 1, 2, \dots, S,$$

with

hence

which yields

$$h(k) = \mathbb{E}[T_S \mid X_0 = k] = (S + k + 1)(S - k), \ \ k = 0, 1, \dots, S.$$

e) When $p \neq q$ we have

$$p_k := \mathbb{P}(T_S < T_0 \mid X_0 = k) = rac{1 - (q/p)^k}{1 - (q/p)^S}, \qquad k = 0, 1, \dots, S,$$

and when p = q = 1/2 we find

$$p_k=rac{k}{S}, \qquad k=0,1,\ldots,S.$$

f) The equality holds because, given that we start from state k+1 at time 1, whether T_S < T₀ or T_S > T₀ does not depend on the past of the process before time 1. In addition it does not matter whether we start from state k+1 at time 1 or at time 0.
g) We have

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

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

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

 $k = 0, 1, \dots, S - 1$. 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

$$\begin{split} \mathbb{P}(X_1 &= k - 1 \mid X_0 = k \text{ and } T_0 < T_S) \\ &= \frac{\mathbb{P}(X_1 = k - 1, \ X_0 = k \text{ and } T_0 < T_S)}{\mathbb{P}(X_0 = k \text{ and } T_0 < T_S)} \\ &= \frac{\mathbb{P}(T_0 < T_S \mid X_1 = k - 1 \text{ and } X_0 = k) \mathbb{P}(X_1 = k - 1 \text{ and } X_0 = k)}{\mathbb{P}(T_0 < T_S \text{ and } X_0 = k)} \\ &= q \frac{\mathbb{P}(T_0 < T_S \mid X_0 = k - 1)}{\mathbb{P}(T_0 < T_S \mid X_0 = k)} \\ &= q \frac{1 - p_{k-1}}{1 - p_k}, \end{split}$$

 $k = 1, 2, \ldots, S - 1$. When $p \neq q$ this yields

$$\mathbb{P}(X_1 = k - 1 \mid X_0 = k ext{ and } T_0 < T_S) = q rac{(q/p)^{k-1} - (q/p)^S}{(q/p)^k - (q/p)^S},$$

MH3512 AY19-20

5484

 $k = 1, 2, \ldots, S - 1$, and when p = q = 1/2 we find

$$\mathbb{P}(X_1 = k - 1 \mid X_0 = k \text{ and } T_0 < T_S) = rac{S + 1 - k}{2(S - k)},$$

 $k = 1, 2, \ldots, S - 1$. Note that this probability is higher than q = 1/2.

i) We find

$$h(k) = 1 + p \frac{p_{k+1}}{p_k} h(k+1) + \left(1 - p \frac{p_{k+1}}{p_k}\right) h(k-1),$$
(B.20)

 $k = 1, 2, \ldots, S - 1$, or, due to the first step equation $p_k = pp_{k+1} + qp_{k-1}$,

$$p_k h(k) = p_k + p p_{k+1} h(k+1) + q p_{k-1} h(k-1), \quad k = 1, 2, \dots, S-1,$$

with the boundary condition h(S) = 0. When p = q = 1/2 we have $p_k = k/S$ by Question (e), hence (B.20) becomes

$$h(k) = 1 + rac{k+1}{2k}h(k+1) + rac{k-1}{2k}h(k-1),$$

 $k=1,2,\ldots,S-1.$

j) We have to solve the equation

$$kh(k) = k + \frac{1}{2}(k+1)h(k+1) + \frac{1}{2}(k-1)h(k-1), \qquad k = 1, 2, \dots, S-1,$$

with the boundary condition h(S) = 0. Letting g(k) := kh(k) we check that g(k) satisfies

$$g(k) = k + \frac{1}{2}g(k+1) + \frac{1}{2}g(k-1), \qquad k = 1, 2, \dots, S-1,$$
 (B.21)

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.21) has the form

$$g(k) = -rac{1}{3}k^3 + C_1 + C_2 k,$$
 MH3512 AY19-20

by the homogeneous solution given in Section 2.3, where C_1 and C_2 are determined by the boundary conditions

$$0=g(0)=C_1$$

and

$$0=g(S)=-rac{1}{3}S^3+C_1+C_2S,$$

i.e. $C_1 = 0$ and $C_2 = S^2/3$. Consequently, we have

$$g(k) = rac{k}{3}(S^2-k^2), \qquad k=0,1,\dots,S,$$

hence we have

$$h(k) = \mathbb{E}[T_S \mid X_0 = k, \ T_S < T_0] = rac{S^2 - k^2}{3}, \qquad k = 1, 2, \dots, S.$$

The conditional expectation h(0) is actually undefined because the event $\{X_0 = 0, T_S < T_0\}$ has probability 0.

Problem 2.8

- a) We have g(0) = 0 and g(S) = g(S+1) = 1.
- b) We have

$$g(k) = pg(k+2) + rg(k) + 2pg(k-1), \qquad 1 \leqslant k \leqslant S-1,$$

and $g(0) = 0$ and $g(S) = g(S+1) = 1.$

c) Trying a solution of the form $g(k) = C\lambda^k$ shows that λ must satisfy

$$\lambda^k = p\lambda^{k+2} + r\lambda^k + 2p\lambda^{k-1}, \qquad 1\leqslant k\leqslant S-1,$$

i.e.

$$\lambda = p\lambda^3 + r\lambda + 2p, \qquad 1\leqslant k\leqslant S-1,$$

hence, using the relation 3p + r = 1 we find

$$\lambda^3 - 3\lambda + 2 = (\lambda^2 - 2\lambda + 1)(\lambda + 2) = (\lambda - 1)^2(\lambda + 2) = 0, \qquad k = 1, 2, \dots, S - 1,$$

with solutions $\lambda \in \{1, -2\}$. Hence we can now try a solution of the form

$$g(k) = C_1 + C_2 k + C_3 (-2)^k$$

with the boundary conditions

$$\left\{egin{aligned} g(0) &= 0 = C_1 + C_3, \ g(S) &= 1 = C_1 + C_2 S + (-2)^S C_3, \ g(S+1) &= 1 = C_1 + C_2 (S+1) - 2 imes (-2)^S C_3, \end{aligned}
ight.$$

i.e.

$$\begin{cases} C_1 = \frac{1}{1 - (3S + 1)(-2)^S} \\ C_2 = -\frac{3(-2)^S}{1 - (3S + 1)(-2)^S} \\ C_3 = -\frac{1}{1 - (3S + 1)(-2)^S}, \end{cases}$$

which yields

$$g(k) = \frac{1}{1 - (3S+1)(-2)^{S}} - \frac{3k(-2)^{S}}{1 - (3S+1)(-2)^{S}} - \frac{(-2)^{k}}{1 - (3S+1)(-2)^{S}} = \frac{1 - 3k(-2)^{S} - (-2)^{k}}{1 - (3S+1)(-2)^{S}}, \quad k = 0, 1, \dots, S + 1.$$
(B.22)

In the graph of Figure S.1 the run probability (B.22) is plotted as a function of k for p = 1/3 and r = 0.



Figure S.1: Ruin probability $f_8(k)$ as a function of $X_0 = k \in [0, 9]$.

- d) Intuitively, this problem corresponds to a fair game for all values of p since the expected gain $\mathbb{E}[X_n X_{n-1}] = 0$ vanishes^{*} for all $n \ge 1$, so that the probability of ruin f(k) should be the same for all $p \in (0, 1/2)$, which is indeed the case in (B.22).
- e) We have h(0) = h(S) = h(S+1) = 0.

f) We find

$$h(k)=1+ph(k+2)+rh(k)+2ph(k-1), \qquad 1\leqslant k\leqslant S-1.$$

g) After checking that $k \mapsto Ck$ cannot be a solution we search for a particular solution of the form $h(k) = Ck^2$, which yields

$$Ck^2 = 1 + pC(k+2)^2 + rCk^2 + 2pC(k-1)^2, \qquad 1 \leqslant k \leqslant S-1,$$

i.e.

$$0 = 1 + pC(2k + 4) + 2pC(-2k + 1), \qquad 1 \leqslant k \leqslant S - 1,$$

or C = -1/(6p).

h) The general solution takes the form

$$h(k)=-rac{k^2}{6p}+C_1+C_2k+C_3(-2)^k$$

♦ 488

^{*}We are in presence of a fair game in which probabilities of going up and down are not equal, while increments are by two units and decrements are by one unit.

with the boundary conditions

$$\left\{egin{aligned} h(0) &= 0 = C_1 + C_3, \ h(S) &= 0 = -rac{S^2}{6p} + C_1 + C_2 S + (-2)^S C_3, \ h(S+1) &= 0 = -rac{(S+1)^2}{6p} + C_1 + C_2 (S+1) - 2 imes (-2)^S C_3, \end{aligned}
ight.$$

i.e.

$$\begin{cases} C_1 = -\frac{S(S+1)}{6p(1-3S(-2)^S - (-2)^S)} \\ C_2 = \frac{S}{6p} + \frac{(S+1)(1-(-2)^S)}{6p(1-3S(-2)^S - (-2)^S)} \\ C_3 = \frac{S(S+1)}{6p(1-3S(-2)^S - (-2)^S)}, \end{cases}$$

which yields

$$\begin{split} h(k) &= -\frac{k^2}{6p} - \frac{S(S+1)}{6p(1-(3S+1)(-2)^S)} + k \left(\frac{S}{6p} + \frac{(S+1)(1-(-2)^S)}{6p(1-(3S+1)(-2)^S)}\right) \\ &+ \frac{S(S+1)(-2)^k}{6p(1-(3S+1)(-2)^S)} \\ &= \frac{k(S-k)}{6p} + \frac{-S(S+1) + k(S+1)(1-(-2)^S) + S(S+1)(-2)^k}{6p(1-(3S+1)(-2)^S)}, \\ &= \frac{k(S-k)}{6p} + \frac{k(S+1)(1-(-2)^S) - S(S+1)(1-(-2)^S)}{6p(1-(3S+1)(-2)^S)} \\ &= \frac{k(S-k)}{6p} + \frac{k(S+1) - S(S+1)}{6p(1-(3S+1)(-2)^S)} + \frac{-k(S+1)(-2)^S + S(S+1)(-2)^k}{6p(1-(3S+1)(-2)^S)} \end{split}$$

5 489

$$=\frac{k(S-k)}{6p} + (S+1)\frac{S(-2)^k - k(-2)^S - (S-k)}{6p(1-(3S+1)(-2)^S)},$$
(B.23)

 $k = 0, 1, \ldots, S + 1$. In the graph of Figure S.2 the mean game duration (B.23) is plotted as a function of k with p = 1/3.



Figure S.2: Mean game duration $h_{20}(k)$ as a function of $X_0 = k \in [0, 20]$.

- i) Starting from a state $k \in \{1, 2, ..., S 1\}$, the mean duration goes to infinity when p goes to zero. Indeed, when p goes to 0 the probability 1 3p of a draw increases and the game should take longer. Hence the above answer is compatible with intuition.
- j) When S tends to infinity, the run probability starting from $k \ge 1$ tends to 0 and the mean game duration tends to infinity.

Problem 2.9

a) We have h(0) = h(1) = 0.

b) Here we need to increment the value of τ by one unit in the first step analysis argument, which yields

$$\begin{split} h(k) &= p \mathbb{E} \left[(1+\tau)^2 \left| X_0 = k+1 \right] + q \mathbb{E} \left[(1+\tau)^2 \left| X_0 = k-1 \right] \right] \\ &= p \mathbb{E} \left[1+2\tau+\tau^2 \left| X_0 = k+1 \right] + q \mathbb{E} \left[1+2\tau+\tau^2 \left| X_0 = k-1 \right] \right] \\ &= 1+p \mathbb{E} \left[2\tau+\tau^2 \left| X_0 = k+1 \right] + q \mathbb{E} \left[2\tau+\tau^2 \left| X_0 = k-1 \right] \right] \\ &= 1+2p \mathbb{E} \left[\tau \left| X_0 = k+1 \right] + 2q \mathbb{E} \left[\tau \left| X_0 = k-1 \right] \right] \end{split}$$
MH3512 AY19-20

$$+p \mathbb{E} \left[\tau^{2} | X_{0} = k+1\right] + q \mathbb{E} \left[\tau^{2} | X_{0} = k-1\right]$$

= $1 + 2p \mathbb{E} \left[\tau | X_{0} = k+1\right] + 2q \mathbb{E} \left[\tau | X_{0} = k-1\right] + ph(k+1) + qh(k-1),$ $\bigcirc 491$

 $1 \leq k \leq S - 1$, where we used the relation p + q = 1.

c) We have

$$\begin{split} h(k) &= 1 + \mathbb{E}\left[\tau \left| X_0 = k+1 \right] + \mathbb{E}\left[\tau \left| X_0 = k-1 \right] + \frac{1}{2}h(k+1) + \frac{1}{2}h(k-1) \right. \\ &= 1 + (S-k-1)(k+1) + (S-k+1)(k-1) + \frac{1}{2}h(k+1) + \frac{1}{2}h(k-1) \\ &= 1 + (S-k-1)k + (S-k-1) + (S-k+1)k - (S-k+1) \\ &+ \frac{1}{2}h(k+1) + \frac{1}{2}h(k-1) \\ &= -1 + 2(S-k)k + \frac{1}{2}h(k+1) + \frac{1}{2}h(k-1), \end{split}$$

 $1\leqslant k\leqslant S-1.$

d) We have

$$h(k)=C_1+C_2k+rac{2k^2}{3}-rac{k^3}{3}(2S-k)$$

with

$$0=h(0)=C_1$$

and

$$0=h(S)=C_2S+rac{2S^2}{3}-rac{S^4}{3}$$

hence

$$C_2 = -rac{2S}{3} + rac{S^3}{3} = rac{S}{3} \left(S^2 - 2
ight)$$

and

$$h(k)=krac{S}{3}ig(S^2-2ig)+rac{2k^2}{3}-rac{k^3}{3}(2S-k).$$

e) We have

$$\begin{split} v(k) &= h(k) - k^2(S-k)^2 \\ &= k \frac{S}{3} \left(S^2 - 2 \right) + \frac{2k^2}{3} - \frac{k^3}{3} (2S-k) - k^2(S-k)^2 \\ &= k \left(\frac{S}{3} \left(S^2 - 2 \right) + \frac{2k}{3} - S \frac{k^2}{3} - \frac{k^2}{3} (S-k) - k(S-k)^2 \right) \\ &= \frac{k}{3} \left(S \left(S^2 - 2 \right) + 2k - Sk^2 - k^2(S-k) - 3k(S-k)^2 \right) \\ &= \frac{k}{3} \left((S-k) \left(S^2 - 2 \right) + Sk(S-k) - k^2(S-k) - 3k(S-k)^2 \right) \\ &= \frac{k}{3} \left((S-k) \left(S^2 - 2 + Sk \right) - k^2(S-k) - 3k(S-k)^2 \right) \\ &= \frac{k(S-k)}{3} \left(k^2 + (S-k)^2 - 2 \right). \end{split}$$

Remark. When k = S/2 we find

$$v(k) \;=\; rac{S^2}{12} \left(S^2/2 - 2
ight)$$

and the standard deviation

$$\sigma(k) ~\simeq~ rac{S^2}{\sqrt{24}} \simeq rac{S^2}{5}.$$

if \boldsymbol{S} is large.

In the graph of Figure S.3 the standard deviation of the game duration τ is plotted as a function of the initial state k = 0 H $_{3.12} \cdot _{AV1} S_{20}$



Figure S.3: Standard deviation of τ as a function of $X_0 = k \in [0, 20]$.

f) We have v(1) = 0, which is consistent with the fact that the game duration is constant equal to one when k = 1 and S = 2.

Chapter 3 - Random Walks

Exercise 3.1

a) We find
$$\begin{pmatrix} 4 \\ 3 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} = 4$$
 paths, as follows.



ॳ93



Figure S.4: Four paths leading from 0 to 2 in four steps.

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 \boldsymbol{n} and \boldsymbol{k} .

i) In case n and k are even, or written as n = 2n' and k = 2k', (3.3.3) shows that

$$egin{aligned} \mathbb{P}(S_n = k \mid S_0 = 0) &= \mathbb{P}(S_{2n'} = 2k' \mid S_0 = 0) \ &= inom{2n'}{n' + k'} p^{n' + k'} q^{n' - k'} \ &= inom{n}{(n+k)/2} p^{(n+k)/2} q^{(n-k)/2}, \qquad -n \leqslant k \leqslant n. \end{aligned}$$

ii) In case n and k are odd, or written as n = 2n' + 1 and k = 2k' + 1, (3.3.4) shows that

$$egin{aligned} \mathbb{P}(S_n = k \mid S_0 = 0) &= \mathbb{P}(S_{2n'+1} = 2k'+1 \mid S_0 = 0) \ &= inom{2n'+1}{n'+k'+1} p^{n'+k'+1} q^{n'-k'} \ &= inom{n}{(n+k)/2} p^{(n+k)/2} q^{(n-k)/2}, \qquad -n \leqslant k \leqslant n. \end{aligned}$$

d) By a first step analysis started at state (0) we have, letting $p_{n,k} := \mathbb{P}(S_n = k)$,

$$\begin{split} p_{n+1,k} &= \mathbb{P}(S_{n+1} = k) \\ &= \mathbb{P}(S_{n+1} = k \mid S_0 = 0) \\ &= \mathbb{P}(S_{n+1} = k \mid S_1 = 1) \mathbb{P}(S_1 = 1 \mid S_0 = 0) \\ &+ \mathbb{P}(S_{n+1} = k \mid S_1 = -1) \mathbb{P}(S_1 = -1 \mid S_0 = 0) \\ &= p \mathbb{P}(S_{n+1} = k \mid S_1 = 1) + q \mathbb{P}(S_{n+1} = k \mid S_1 = -1) \\ &= p \mathbb{P}(S_{n+1} = k - 1 \mid S_1 = 0) + q \mathbb{P}(S_{n+1} = k + 1 \mid S_1 = 0) \\ &= p \mathbb{P}(S_n = k - 1 \mid S_0 = 0) + q \mathbb{P}(S_n = k + 1 \mid S_0 = 0) \\ &= p p_{n,k-1} + q p_{n,k+1}, \end{split}$$

which yields

$$p_{n+1,k} = pp_{n,k-1} + qp_{n,k+1},$$

for all $n \in \mathbb{N}$ and $k \in \mathbb{Z}$. The same equation could be obtained by a *backstep* analysis, by looking at the two possible ways to reach state k at time n + 1 after starting from either k - 1 or k + 1 at time n.

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

$$\begin{split} pp_{n,k-1} + qp_{n,k+1} &= p\binom{n}{(n-1+k)/2} p^{(n-1+k)/2} (1-p)^{(n+1-k)/2} \\ &+ q\binom{n}{(n+1+k)/2} p^{(n+1+k)/2} (1-p)^{(n-1-k)/2} \\ &= \binom{n}{(n-1+k)/2} p^{(n+1+k)/2} q^{(n+1-k)/2} \\ &+ \binom{n}{(n+1+k)/2} p^{(n+1+k)/2} q^{(n+1-k)/2} \\ &= p^{(n+1+k)/2} q^{(n+1-k)/2} \left(\binom{n}{(n-1+k)/2} + \binom{n}{(n+1+k)/2}\right) \end{split}$$

495

$$= p^{(n+1+k)/2} q^{(n+1-k)/2} \left(\frac{n!}{((n+1-k)/2)!((n-1+k)/2)!} + \frac{n!}{((n-1-k)/2)!((n+1+k)/2)!} \right)$$

$$= p^{(n+1+k)/2} q^{(n+1-k)/2} \left(\frac{n!(n+1+k)/2}{((n+1-k)/2)!((n+1+k)/2)!} + \frac{n!(n+1-k)/2}{((n+1-k)/2)!((n+1+k)/2)!} \right)$$

$$= p^{(n+1+k)/2} q^{(n+1-k)/2} \frac{n!(n+1)}{((n+1-k)/2)!((n+1+k)/2)!}$$

$$= p^{(n+1+k)/2} q^{(n+1-k)/2} \binom{n+1}{(n+1+k)/2}$$

$$= p_{n+1,k},$$

which shows that $p_{n,k}$ satisfies Equation (3.4.25). In addition we clearly have

$$p_{0,0} = \mathbb{P}(S_0 = 0) = 1$$
 and $p_{0,k} = \mathbb{P}(S_0 = k) = 0, \quad k \neq 0.$

Exercise 3.2

a) By the space homogeneity of the random walk we have

$$\mathbb{E}[T_{i-1} \mid S_0 = i] = \mathbb{E}[T_0 \mid S_0 = 1], \quad i = 1, 2, \dots, k,$$

hence by adding inductively the average travel times from (i) to (i-1) we find

$$\mathbb{E}[T_0 \mid S_0 = k] = \sum_{i=1}^k \mathbb{E}[T_{i-1} \mid S_0 = i] = k \mathbb{E}[T_0 \mid S_0 = 1].$$

Next, by first step analysis we have

$$\mathbb{E}[T_0 \mid S_0 = 1] \; = \; q imes 1 + p imes (1 + \mathbb{E}[T_0 \mid S_0 = 2]) \ = \; q imes 1 + p imes (1 + 2 \, \mathbb{E}[T_0 \mid S_0 = 1])$$

MH3512 AY19-20

ॳ96

$$= 1 + 2p \mathbb{E}[T_0 \mid S_0 = 1], \qquad (5497)$$

hence

$$\mathbb{E}[T_0 \mid S_0 = 1] = rac{1}{1-2p} = rac{1}{q-p}, \qquad q > p.$$

When $p \ge q$ we find $\mathbb{E}[T_0 \mid S_0 = 1] = +\infty$.

b) By the space homogeneity of the random walk we have

$$\mathbb{P}(T_{i-1} < \infty \mid S_0 = i) = \mathbb{P}(T_0 < \infty \mid S_0 = 1), \quad i = 1, 2, \dots, k,$$

hence by multiplying inductively the hitting probabilities of state (i-1) starting from (i) we find

$$\mathbb{P}(T_0<\infty\mid S_0=k)=\prod\limits_{i=1}^k\mathbb{P}(T_{i-1}<\infty\mid S_0=i)=\left(\mathbb{P}(T_0<\infty\mid S_0=1)
ight)^k.$$

c) By first step analysis we have

$$\mathbb{P}(T_0 < \infty \mid S_0 = 1) = q + p(\mathbb{P}(T_0 < \infty \mid S_0 = 1))^2,$$

which yields the equation

$$p\alpha^2 - \alpha + q = 0$$

for $\alpha := \mathbb{P}(T_0 < \infty \mid S_0 = 1)$, and possible solutions $\alpha \in \{1, q/p\}$. When p > q we find

$$\mathbb{P}(T_0 < \infty \mid S_0 = 1) = rac{q}{p}$$

when $p \leqslant q$ we have $\mathbb{P}(T_0 < \infty \mid S_0 = 1) = 1$. We conclude that in general,

$$\mathbb{P}(T_0 < \infty \mid S_0 = k) = \min\left(1, \left(rac{q}{p}
ight)^k
ight), \qquad k \in \mathbb{N},$$

which coincides with (2.2.13) and (3.4.16).

Exercise 3.3 Recall that by (3.4.27) we have $G_{T_0^r}(s) = 1 - \sqrt{1 - 4pqs^2}, s \in (-1, 1).$

a) We have $\mathbb{P}(T_0^r = 0) = G_{T_0^r}(0) = 0$ and

$$\mathbb{P}(T_0^r < \infty) = G_{T_0^r}(1) = 1 - \sqrt{1 - 4pq} = 2\min(p, q).$$

b) In general we have $\mathbb{P}(T_0^r=n)=rac{1}{n!}G_{T_0^r}^{(n)}(0),\,n\geqslant 0,$ with

$$\left\{egin{array}{l} G_{T_0^r}'(s) &= 4pqs(1-4pqs^2)^{-1/2}, \ G_{T_0^r}''(s) &= 4pq(1-4pqs^2)^{-3/2}, \ G_{T_0^r}''(s) &= 48p^2q^2s(1-4pqs^2)^{-5/2}, \ G_{T_0^r}^{(4)}(s) &= 48p^2q^2(1+16pqs^2)(1-4pqs^2)^{-7/2}, \end{array}
ight.$$

hence

$$\left\{egin{array}{l} \mathbb{P}(T_0^r=1)=G_{T_0^r}'(0)=0,\ \mathbb{P}(T_0^r=2)=G_{T_0^r}''(0)=2pq,\ \mathbb{P}(T_0^r=3)=rac{1}{3!}G_{T_0^r}'''(0)=0,\ \mathbb{P}(T_0^r=4)=rac{1}{4!}G_{T_0^r}^{(4)}(0)=2p^2q^2. \end{array}
ight.$$

c) We have

$$\mathbb{E}[T_0^r \mathbbm{1}_{\{T_0^r < \infty\}}] = G_{T_0^r}'(1) = rac{4pq}{\sqrt{1-4pq}} = rac{4pq}{|p-q|},$$

hence

$$\mathbb{E}[T_0^r \mid T_0^r < \infty] = rac{1}{\mathbb{P}(T_0^r < \infty)} \mathbb{E}[T_0^r \mathbbm{1}_{\{T_0^r < \infty\}}] = 2rac{\max(p,q)}{|p-q|},$$

see (3.4.19).

Exercise 3.4

a) We have the boundary conditions $G_0(s) = 1, G_{\pm\infty}(s) = 0, s \in (-1, 1)$, and

$$egin{array}{rll} G_i(s) &=& \mathbb{E}[s^{T_0} \mid X_0 = i] \ &=& p \, \mathbb{E}[s^{1+T_0} \mid X_0 = i+1] + q \, \mathbb{E}[s^{1+T_0} \mid X_0 = i-1] \ &=& ps \, \mathbb{E}[s^{T_0} \mid X_0 = i+1] + qs \, \mathbb{E}[s^{T_0} \mid X_0 = i-1], \qquad i \geqslant 1, \end{array}$$

hence we have

$$G_i(s) = psG_{i+1}(s) + qsG_{i-1}(s), \quad i \in \mathbb{Z} \setminus \{0\}, \quad s \in [0, 1).$$
(B.24)

b) We note that (B.24) has the same form as Equation (2.2.6), up to a parametrization by $s \in [0, 1)$. Hence, substituting a possible solution of the form $G_i(s) = (\alpha(s))^i$ into (B.24) we find

$$(\alpha(s))^i = ps(\alpha(s))^{i+1} + qs(\alpha(s))^{i-1},$$

which yields the characteristic equation

$$ps(lpha(s))^2-lpha(s)+qs=0, \qquad s\in [0,1),$$

with unknown $\alpha(s)$ and solutions

$$\alpha_{+}(s) = \frac{1 + \sqrt{1 - 4pqs^{2}}}{2ps} = \frac{2qs}{1 - \sqrt{1 - 4pqs^{2}}} > \frac{1}{2ps}, \qquad |s| < 1/(2\sqrt{pq}),$$

and

$$\alpha_{-}(s) = \frac{1 - \sqrt{1 - 4pqs^2}}{2ps} = \frac{2qs}{1 + \sqrt{1 - 4pqs^2}} < 2qs, \qquad |s| < 1/(2\sqrt{pq}),$$

hence the general solution of (B.24) is given by

$$egin{array}{rcl} G_i(s) &=& C_+(lpha_+(s))^i + C_-(lpha_-(s))^i \ &=& C_+\left(rac{1+\sqrt{1-4pqs^2}}{2ps}
ight)^i + C_-\left(rac{1-\sqrt{1-4pqs^2}}{2ps}
ight)^i, & i\in\mathbb{Z}. \end{array}$$

Hence the boundary conditions $G_{\pm\infty}(s) = 0$, or the boundedness condition $|G_i(s)| \leq 1$ for sufficiently small $s \in (-1, 1, 3, 1)$, show

that $C_+ = 0$ and $C_- = 1$ when $i \ge 0$, *i.e.*

$$G_i(s) = \left(\frac{1 - \sqrt{1 - 4pqs^2}}{2ps}\right)^i = \left(\frac{2qs}{1 + \sqrt{1 - 4pqs^2}}\right)^i, \qquad i \ge 0,$$
(B.25)

and $C_+ = 1$ and $C_- = 0$ when $i \leq 0$, *i.e.*

$$G_i(s) = \left(rac{1+\sqrt{1-4pqs^2}}{2ps}
ight)^i = \left(rac{2qs}{1-\sqrt{1-4pqs^2}}
ight)^i, \qquad i \leqslant 0,$$

while by symmetry we also have

$$G_i(s) = \left(\frac{2qs}{1+\sqrt{1-4pqs^2}}\right)^i = \left(\frac{1-\sqrt{1-4pqs^2}}{2ps}\right)^i, \quad i \ge 0,$$
(B.26)

and

$$G_i(s)=\left(rac{2qs}{1-\sqrt{1-4pqs^2}}
ight)^i=\left(rac{1+\sqrt{1-4pqs^2}}{2ps}
ight)^i, \qquad i\leqslant 0.$$

c) Relations (B.25) and (B.26) show from (3.4.14) that

$$\mathbb{P}(T_0 < \infty \mid X_0 = i) = G_i(1) = rac{2q}{1 + \sqrt{1 - 4pq}} = rac{2q}{1 + |p - q|} = \min\left(1, \left(rac{q}{p}
ight)^i
ight),$$

 $i \ge 0$, which recovers (2.2.13) and (3.4.16).

d) When $\boldsymbol{q} \leqslant \boldsymbol{p}$ we have

$$G_i'(s) = -rac{i s^{-i-1}}{(2q)^i} ig(1 - \sqrt{1 - 4pqs^2}ig)^i + rac{4pqi}{(2q)^i} rac{ig(1 - \sqrt{1 - 4pqs^2}ig)^{i-1}}{\sqrt{1 - 4pqs^2}},$$

and when s = 1 this yields

$$G_i'(1) \;=\; -rac{i}{(2q)^i} ig(1-\sqrt{1-4pq}ig)^i + rac{4pqi}{(2q)^i} rac{ig(1-\sqrt{1-4pq}ig)^{i-1}}{\sqrt{1-4pq}} \;,$$

500

 $\bigcirc 501$

$$egin{aligned} &= -rac{i}{(2q)^i} \left(1-(p-q)
ight)^i + rac{4pqi}{(2q)^i} rac{\left(1-(p-q)
ight)^{i-1}}{p-q} \ &= -i + rac{2pi}{p-q} \ &= rac{i}{q-p}, \qquad i \leqslant 0. \end{aligned}$$

e) By first step analysis starting from state 0 we have, when $q \ge p$,

$$\begin{split} \mathbb{E}\left[s^{T_0^r} \mid X_0 = 0\right] &= ps \,\mathbb{E}\left[s^{T_0} \mid X_0 = 1\right] + qs \,\mathbb{E}\left[s^{T_0} \mid X_0 = -1\right] \\ &= ps G_1(s) + qs G_{-1}(s) \\ &= \frac{1 - \sqrt{1 - 4pqs^2}}{2} + \frac{2pqs^2}{1 + \sqrt{1 - 4pqs^2}} \\ &= \frac{(1 - \sqrt{1 - 4pqs^2})(1 + \sqrt{1 - 4pqs^2}) + 4pqs^2}{2(1 + \sqrt{1 - 4pqs^2})} \\ &= \frac{4pqs^2}{1 + \sqrt{1 - 4pqs^2}} \\ &= 1 - \sqrt{1 - 4pqs^2}, \quad s \in [0, 1), \end{split}$$

which recovers (3.4.9), while when $q \leq p$ we find similarly

$$egin{array}{rl} \mathbb{E}\left[s^{T_0^r} \mid X_0 = 0
ight] &= psG_1(s) + qsG_{-1}(s) \ &= rac{2qps^2}{1 + \sqrt{1 - 4pqs^2}} + rac{1 - \sqrt{1 - 4pqs^2}}{2} \ &= 1 - \sqrt{1 - 4pqs^2}, \hspace{1em} s \in [0,1). \end{array}$$

Exercise 3.5 By (3.4.21) we have

$$\mathbb{E}\left[T_0^r\mathbbm{1}_{\{T_0^r<\infty\}}\mid S_0=0
ight]\ =\ \sum\limits_{n\geqslant 1}n\mathbb{P}ig(T_0^r=n\mid S_0=0ig)$$

$$= \sum_{k \ge 1} 2k \mathbb{P} \left(T_0^r = 2k \mid S_0 = 0 \right)$$

$$= \sum_{k \ge 1} \frac{2k}{2k - 1} \binom{2k}{k} (1/4)^k$$

$$= \infty,$$

(B.27)

where the general term of the above series satisfies the equivalence

$$rac{2k}{2k-1} {2k \choose k} (1/4)^k \simeq_{k o \infty} rac{(2k)!}{4^k (k!)^2} \simeq_{k o \infty} \sqrt{rac{1}{2\pi k}}$$

by Stirling's approximation $k! \simeq (k/e)^k \sqrt{2\pi k}$ as k tends to ∞ , from which we conclude to the divergence of the series (B.27).

Exercise 3.6

a) We have

$$egin{array}{rcl} \mathbb{E}\left[M_n
ight] &=& \mathbb{E}\left[\sum\limits_{k=1}^n 2^{k-1} X_k
ight] \ &=& \sum\limits_{k=1}^n 2^{k-1} \mathbb{E}\left[X_k
ight] \ &=& (p-q) \sum\limits_{k=0}^{n-1} 2^k \ &=& (p-q) rac{1-2^n}{1-2} \ &=& (p-q)(2^n-1), \qquad n \geqslant 0 \end{array}$$

b) The two possible values of $M_{\tau \wedge n}$ are 1 and

$$-\sum\limits_{k=1}^n 2^{k-1} = -rac{1-2^n}{1-2} = 1-2^n, \qquad n \geqslant 1.$$

We have

$$\mathbb{P}(M_{n\wedge \tau}=1-2^n)=q^n \quad ext{and} \quad \mathbb{P}(M_{n\wedge \tau}=1)=1-q^n, \qquad n \geqslant 1.$$

- c) The stopped process $(M_{\min(n,\tau)})_{n\in\mathbb{N}}$ represents the wealth of a gambler whose strategy of the gambler is to double the stakes each time he loses, and to quit the game as soon as his gains reach \$1.
- d) We have

$$egin{array}{rll} \mathbb{E}[M_{n\wedge au}] &=& (1-2^n)\mathbb{P}(M_{n\wedge au}=1-2^n)+\mathbb{P}(M_{n\wedge au}=1) \ &=& (1-2^n)q^n+1-q^n \ &=& 1-(2q)^n, \qquad n\geqslant 1. \end{array}$$

In particular, we find $\mathbb{E}[M_{n \wedge \tau}] = 0$ in the fair game with p = q = 1/2.

Exercise 3.7

- a) We have $\mathbb{P}(T^{(m)} < m) = 0$, $\mathbb{P}(T^{(m)} = m) = p^m$, $\mathbb{P}(T^{(m)} = m + 1) = qp^m$, and $\mathbb{P}(T^{(m)} = m + 2) = q^2p^m + pqp^m = qp^m$.
- b) The idea is to start by start by flipping a coin and to observe the number k of consecutive "1" until we get the first "0".
 - 1) If k = m then the game ends, and this happens with probability $\mathbb{P}(T^{(m)} = m) = p^m$.
 - 2) If k < m, the sequence of "1" is broken and we need to start again at time k + 1. This happens with probability $p^k q$ and we need to factor in the power s^{k+1} where k + 1 is the number of time steps until we reach the first "0", and restart the counter $T^{(m)}$.

Therefore we have

$$egin{aligned} G_{T^{(m)}}(s) &= s^m \mathbb{P}(T^{(m)}=m) + \sum\limits_{k=0}^{m-1} qp^k \, \mathbb{E}[s^{k+1+T^{(m)}}] \ &= p^m s^m + \sum\limits_{k=0}^{m-1} p^k qs^{k+1} \, \mathbb{E}[s^{T^{(m)}}] \ &= p^m s^m + \sum\limits_{k=0}^{m-1} p^k qs^{k+1} G_{T^{(m)}}(s), \qquad s \in (-1,1). \end{aligned}$$

c) We have

$$G_{T^{(m)}}(s) \left(1 - \sum\limits_{k=0}^{m-1} p^k q s^{k+1}
ight) = p^m s^m,$$
 MH3512 AY19-20

504

$$egin{aligned} G_{T^{(m)}}(s) &= rac{p^m s^m}{1 - \sum_{k=0}^{m-1} p^k q s^{k+1}} \ &= rac{p^m s^m}{1 - rac{q s (1-(ps)^m}{1-ps}} \ &= rac{p^m s^m (1-ps)}{1-ps - qs (1-(ps)^m)} \ &= rac{p^m s^m (1-ps)}{1-s + q p^m s^{m+1}}, \quad s \in (-1,) \end{aligned}$$

We note that

$$\mathbb{P}ig(T^{(m)} < \inftyig) = G_{T^{(m)}}(1) = (1-p)rac{p^m}{qp^m} = 1.$$

1].

d) Differentiating in (3.4.28) with respect to s, we find

$$G_{T^{(m)}}'(s) = \sum\limits_{k=0}^{m-1} p^k q(k+1) s^k G_{T^{(m)}}(s) + \sum\limits_{k=0}^{m-1} p^k q s^{k+1} G_{T^{(m)}}'(s) + m p^m s^{m-1},$$

 $s\in(-1,1),$ hence, since $G_{T^{(m)}}(1)=1$ because $\mathbb{P}ig(T^{(m)}<\inftyig)=1,$ we have

$$egin{aligned} G_{T^{(m)}}'(1) &= (1-p)\sum\limits_{k=0}^{m-1}p^k(k+1)+mp^m+G_{T^{(m)}}'(1)\sum\limits_{k=0}^{m-1}p^kq \ &= rac{1-p^m}{1-p}+qG_{T^{(m)}}'(1)\sum\limits_{k=0}^{m-1}p^k \ &= rac{1-p^m}{1-p}+qG_{T^{(m)}}'(1)rac{1-p^m}{1-p} \ &= rac{1-p^m}{1-p}+G_{T^{(m)}}'(1)(1-p^m), \end{aligned}$$
which yields

$$\mathbb{E}\left[T^{(m)}
ight] = G'_{T^{(m)}}(1) = rac{1-p^m}{p^m(1-p)} = rac{1-(1/p)^m}{p(1-1/p)} = \sum_{k=1}^m rac{1}{p^k}$$

For example, for an unbiased coin with p = 1/2 the mean time until the first winning streak of length m is

$$\mathbb{E}\left[T^{(m)}\right] = \sum_{k=1}^{m} \frac{1}{(1/2)^k} = \sum_{k=1}^{m} 2^k = 2\frac{1-2^m}{1-2} = 2(2^m-1).$$

Exercise 3.8

a) Applying (3.4.29) with m = n - 1 and k = 0, we find that this number is $\binom{2n-2}{n-1}$.

b) Applying (3.4.29) with
$$m = n - 1$$
 and $k = -1$, we find that this number is $\binom{2n-2}{n-2} = \binom{2n-2}{n}$.

c) On the graph below, the blue path joining $S_1 = 1$ to $S_{2n-1} = 1$ by crossing (1) is associated to a unique red path joining $S_1 = 1$ to $S_{2n-1} = -1$.

MH3512 AY19-20

505



d) This number is
$$\binom{2n-2}{n-2} = \binom{2n-2}{n}$$
.

e) This number is

$$\binom{2n-2}{n-1} - \binom{2n-2}{n-2} = \frac{(2n-2)!}{(n-1)!(n-1)!} - \frac{(2n-2)!}{(n-2)!n!} \\ = \frac{(n^2 - n(n-1))(2n-2)!}{n!n!} \\ = \frac{(2n-2)!}{(n-1)!n!}.$$

f) According to the following graph, to each of the $\frac{(2n-2)!}{(n-1)!n!}$ blue paths joining $S_1 = 1$ to $S_{2n-2} = 1$ without crossing (1) between time 1 and time 2n - 1 we can associate a red path joining $S_1 = -1$ to $S_{2n-2} = -1$ without crossing (1).



Adding the number of paths joining $S_1 = 1$ to $S_{2n-2} = 1$ without crossing (1) between time 1 and time 2n - 1 to the number of paths joining $S_1 = -1$ to $S_{2n-2} = -1$ without crossing (1) between time 1 and time 2n - 1, we get the total to the number of paths joining $S_0 = 0$ to $S_{2n} = 0$ without crossing (1), between time 0 and time 2n, as follows:

$$2rac{(2n-2)!}{(n-1)!n!} = rac{2n(2n-2)!}{n!n!} = rac{1}{2n-1} {2n \choose n}.$$
 MH3512 AY19-20

We note that the convolution equation (3.4.3) is also satisfied in the following form:

$$egin{array}{rcl} rac{1}{2n-1}inom{2n}{n}&=&\sum\limits_{k=0}^{n-1}rac{1}{2n-2k-1}inom{2n-2k}{n-k}inom{2k}{k}\ &=&\sum\limits_{k=0}^{n-1}rac{1}{2k-1}inom{2k}{k}inom{2n-2k}{n-k}, &n\geqslant 1, \end{array}$$

see *e.g.* here.

Problem 3.9

a) We have

$$\mathbb{P}(S_{2n}=2k)=inom{2n}{n+k}p^{n+k}q^{n-k}, \quad -n\leqslant k\leqslant n.$$

b) We partition the event $\{S_{2n} = 0\}$ into

$$\{S_{2n}=0\}=igcup_{k=1}^{2n}ig\{S_1
eq 0,\ldots,S_{2k-1}
eq 0,\ S_{2k}=0ig\},\qquad n\geqslant 1,$$

according to all possible times $2k = 2, 4, \ldots, 2n$ of *first* return to state \bigcirc *before* time 2n, see Figure S.5.



♦ 509

Figure S.5: Last return to state 0 at time k = 10.

Then we have

$$\begin{split} \mathbb{P}(S_{2n} = 0) &= \sum_{r=1}^{n} \mathbb{P}\big(S_2 \neq 0, \dots, S_{2r-1} \neq 0, \ S_{2r} = 0, \ S_{2n} = 0\big) \\ &= \sum_{r=1}^{n} \mathbb{P}\big(S_{2n} = 0 \mid S_{2r} = 0, \ S_{2r-1} \neq 0, \dots, \ S_2 \neq 0\big) \\ &\times \mathbb{P}\big(S_2 \neq 0, \dots, S_{2r-1} \neq 0, \ S_{2r} = 0\big) \\ &= \sum_{r=1}^{n} \mathbb{P}\big(S_{2n} = 0 \mid S_{2r} = 0\big) \mathbb{P}\big(T_0 = 2r\big) \\ &= \sum_{k=1}^{n} \mathbb{P}\big(S_{2n-2r} = 0\big) \mathbb{P}\big(T_0 = 2r\big), \qquad n \geqslant 1. \end{split}$$

c) The idea of the proof is to note that after starting from $S_0 = 0$, one may move up with probability 1/2, in which case $T_0 = 2r$ time steps strictly above 0 will be counted from time 0 until time T_0 , after which the remaining 2r - 2k time steps will be counted from time T_0 until time 2n. On the other hand, if one moves down with probability 1/2, zero time step strictly above 0 will be counted from time 1/2, after which the remaining 2k time steps strictly above 2 will be counted from time $T_0 = 2r$, after which the remaining 2k time steps strictly above zero will be counted from time $T_0 = 2r$, after which the remaining 2k time steps strictly above zero will be counted from time $T_0 = 2r$, after which the remaining 2k time steps strictly above zero will be counted from time $T_0 = 2r$.

until time 2n. Hence we have

$$\begin{split} \mathbb{P}(T_{2n}^{+} = 2k) &= \sum_{r=1}^{n} \mathbb{P}(S_{0} = 0, T_{0} = 2r, T_{2n}^{+} = 2k) \\ &= \sum_{r=1}^{n} \mathbb{P}(S_{0} = 0, S_{1} = 1, T_{0} = 2r, T_{2n}^{+} = 2k) \\ &+ \sum_{r=1}^{n} \mathbb{P}(S_{0} = 0, S_{1} = -1, T_{0} = 2r, T_{2n}^{+} = 2k) \\ &= \sum_{r=1}^{k} \mathbb{P}(S_{0} = 0, S_{1} = 1, T_{0} = 2r) \mathbb{P}(T_{2n}^{+} = 2k \mid S_{1} = 1, T_{0} = 2r) \\ &+ \sum_{r=1}^{n-k} \mathbb{P}(S_{0} = 0, S_{1} = -1, T_{0} = 2r) \mathbb{P}(T_{2n}^{+} = 2k \mid S_{1} = -1, T_{0} = 2r) \\ &= \sum_{r=1}^{k} \mathbb{P}(S_{0} = 0, S_{1} = -1, T_{0} = 2r) \mathbb{P}(T_{2n-2r}^{+} = 2k - 2r) \\ &+ \sum_{r=1}^{n-k} \mathbb{P}(S_{0} = 0, S_{1} = -1, T_{0} = 2r) \mathbb{P}(T_{2n-2r}^{+} = 2k) \\ &= \frac{1}{2} \sum_{r=1}^{k} \mathbb{P}(T_{0} = 2r) \mathbb{P}(T_{2n-2r}^{+} = 2k - 2r) + \frac{1}{2} \sum_{r=1}^{n-k} \mathbb{P}(T_{0} = 2r) \mathbb{P}(T_{2n-2r}^{+} = 2k), \end{split}$$

 $n \geqslant 1$.

d) We check that, when

$$\mathbb{P}ig(T^+_{2n-2r}=2k-2rig)=2^{-(2n-2r)}igg(rac{2k-2r}{k-r}ig)igg(rac{2n-2k}{n-k}ig)$$

and

$$\mathbb{P}ig(T^+_{2n-2r}=2kig)=2^{-(2n-2r)}igg(2k\kig)ig(2n-2r-2k\n-r-kig),$$

we have

$$egin{aligned} &rac{1}{2}\sum\limits_{r=1}^k \mathbb{P}ig(T_0=2rig)\mathbb{P}ig(T_{2n-2r}^+=2k-2rig)+rac{1}{2}\sum\limits_{r=1}^{n-k}\mathbb{P}ig(T_0=2rig)\mathbb{P}ig(T_{2n-2r}^+=2kig)\ &= rac{1}{2}\sum\limits_{r=1}^k \mathbb{P}ig(T_0=2rig)2^{-2n+2r}igg(rac{2k-2r}{k-r}ig)igg(rac{2n-2k}{n-k}igg) \end{aligned}$$

510

$$\begin{split} &+ \frac{1}{2} \sum_{r=1}^{n-k} 2^{-2n+2r} \mathbb{P}(T_0 = 2r) \binom{2k}{k} \binom{2n-2r-2k}{n-r-k} \\ &= \frac{1}{2} 2^{-2n} \binom{2n-2k}{n-k} 2^{2k} \sum_{r=1}^{k} \mathbb{P}(T_0 = 2r) \frac{1}{2^{2(k-r)}} \binom{2k-2r}{k-r} \\ &+ \frac{1}{2} 2^{-2n} \binom{2k}{k} 2^{2(n-k)} \sum_{r=1}^{n-k} \mathbb{P}(T_0 = 2r) \frac{1}{2^{2(n-k-r)}} \binom{2n-2r-2k}{n-r-k} \\ &= \frac{1}{2} 2^{-2(n-k)} \binom{2n-2k}{n-k} \sum_{r=1}^{k} \mathbb{P}(T_0 = 2r) \mathbb{P}(S_{2k-2r} = 0) \\ &+ \frac{1}{2} \binom{2k}{k} 2^{-2k} \sum_{r=1}^{n-k} \mathbb{P}(T_0 = 2r) \mathbb{P}(S_{2n-2k+2r} = 0) \\ &= \frac{1}{2} 2^{-2(n-k)} \binom{2n-2k}{n-k} \mathbb{P}(S_{2k} = 0) + \frac{1}{2} 2^{-2k} \binom{2k}{k} \mathbb{P}(S_{2n-2k} = 0) \\ &= \frac{1}{2} 2^{-2n} \binom{2n-2k}{n-k} \binom{2k}{k} + \frac{1}{2} 2^{-2n} \binom{2k}{k} \binom{2n-2k}{n-k} \\ &= 0 \end{split}$$

e) We have

$$egin{aligned} \mathbb{P}ig(T_{2n}^+ = 2kig) &= 2^{-2n}inom{2k}{k}inom{2n-2k}{n-k}\ &= 2^{-2n}inom{2k!}{k!^2}inom{(2n-2k)!}{(n-k)!^2}\ &\cong 2^{-2n}inom{(2k/e)^{2k}\sqrt{4\pi k}}{(k/e)^{2k}2\pi k}inom{((2n-2k)/e)^{(2n-2k)}\sqrt{2\pi(2n-2k)}}{((n-k)/e)^{(2n-2k)}2\pi(n-k)}\ &= rac{1}{\pi\sqrt{k(n-k)}}, \qquad k,n-k o\infty. \end{aligned}$$

Next, we compute the limit

$$\begin{split} \lim_{n \to \infty} \mathbb{P}\big(T_{2n}^+/2n \leqslant x\big) &= \lim_{n \to \infty} \sum_{k=0}^{nx} \mathbb{P}\big(T_{2n}^+/2n = k/n\big) \\ &= \lim_{n \to \infty} \sum_{0 \leqslant k/n \leqslant x} 2^{-2n} \binom{2k}{k} \binom{2n-2k}{n-k} \\ &\simeq \frac{1}{\pi} \lim_{n \to \infty} \frac{1}{n} \sum_{0 \leqslant k/n \leqslant x} \frac{1}{\sqrt{k(1-k/n)/n}} \\ &= \frac{1}{\pi} \int_0^x \frac{1}{\sqrt{t(1-t)}} dt \\ &= \frac{1}{2} + \frac{\arcsin(2x-1)}{\pi} \\ &= \frac{2}{\pi} \arcsin\sqrt{x}, \qquad x \in [0,1], \end{split}$$

which yields the arcsine distribution.

Problem 3.10

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

$$\left\lfloor \inf_{k=0,1,...,n} S_k, \; \sup_{k=0,1,...,n} S_k
ight
vert,$$

which has

$$R_n = 1 + \left(\sup_{k=0,1,...,n} S_k
ight) - \left(\inf_{k=0,1,...,n} S_k
ight)$$

elements. In addition we have $R_0 = 1$ and $R_1 = 2$.

b) At each time step $k \ge 1$ 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

$$\{R_k - R_{k-1} = 1\} = \{S_k
eq S_0, \ S_k
eq S_1, \dots, \ S_k
eq S_{k-1}\},$$
 MH3512 AY19-20

hence, applying the probability \mathbb{P} to both sides, we get

$$\mathbb{P}(R_k-R_{k-1}=1)=\mathbb{P}(S_k-S_0
eq 0,\ S_k-S_1
eq 0,\ldots,\ S_k-S_{k-1}
eq 0).$$

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

$$egin{aligned} \mathbb{P}(R_k-R_{k-1}=1) &= \mathbb{P}(S_k-S_0
eq 0,\ S_k-S_1
eq 0,\ldots,\ S_k-S_{k-1}
eq 0) \ &= \mathbb{P}(X_1+\cdots+X_k
eq 0,\ X_2+\cdots+X_k
eq 0,\ldots,\ X_k
eq 0) \ &= \mathbb{P}(X_1
eq 0,\ X_1+X_2
eq 0,\ldots,\ X_1+\cdots+X_k
eq 0), \end{aligned}$$

for all $k \ge 1$, since the sequence $(X_k)_{k\ge 1}$ is made of independent and identically distributed random variables.

d) We have the telescoping sum

$$egin{array}{rcl} R_n &=& R_0 + (R_1 - R_0) + \dots + (R_n - R_{n-1}) \ &=& R_0 + \sum\limits_{k=1}^n (R_k - R_{k-1}), & n \in \mathbb{N}. \end{array}$$

e) By (1.2.4) we have

$$\mathbb{P}(T_0=\infty)=\mathbb{P}\left(igcap_{k\geqslant 1}\{T_0>k\}
ight)=\lim_{k
ightarrow\infty}\mathbb{P}(T_0>k),$$

since

$$\{T_0>k+1\}\Longrightarrow\{T_0>k\},\qquad k\geqslant 1,$$

i.e. $({T_0 > k})_{k \ge 1}$ is a decreasing sequence of events.

f) Noting that $R_k - R_{k-1} \in \{0,1\}$ is a Bernoulli random variable with

$$\mathbb{E}[R_k - R_{k-1}] = \mathbb{P}(R_k - R_{k-1} = 1),$$
 MH3512 AY19-2

we find that

$$\begin{split} \mathbb{E}[R_n] &= \mathbb{E}\left[R_0 + \sum_{k=1}^n (R_k - R_{k-1})\right] = R_0 + \sum_{k=1}^n \mathbb{E}[R_k - R_{k-1}] \\ &= R_0 + \sum_{k=1}^n \mathbb{P}(R_k - R_{k-1} = 1) \\ &= R_0 + \sum_{k=1}^n \mathbb{P}(X_1 \neq 0, X_1 + X_2 \neq 0, \dots, X_1 + \dots + X_k \neq 0) \\ &= R_0 + \sum_{k=1}^n \mathbb{P}(S_1 \neq 0, S_2 \neq 0, \dots, S_k \neq 0) = R_0 + \sum_{k=1}^n \mathbb{P}(T_0 > k) \\ &= 1 + \sum_{k=1}^n \mathbb{P}(T_0 > k) = \mathbb{P}(T_0 > 0) + \sum_{k=1}^n \mathbb{P}(T_0 > k) = \sum_{k=0}^n \mathbb{P}(T_0 > k). \end{split}$$

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 $N \ge 1$ such that

$$|\mathbb{P}(T_0=\infty)-\mathbb{P}(T_0>k)|$$

Hence for $n \ge N$ we have

$$egin{aligned} & \left|\mathbb{P}(T_0=\infty)-rac{1}{n}\sum\limits_{k=1}^n\mathbb{P}(T_0>k)
ight|=\left|rac{1}{n}\sum\limits_{k=1}^n(\mathbb{P}(T_0=\infty)-\mathbb{P}(T_0>k))
ight|\ &\leqslant\ rac{1}{n}\sum\limits_{k=1}^n|\mathbb{P}(T_0=\infty)-\mathbb{P}(T_0>k)|+rac{1}{n}\sum\limits_{k=N+1}^n|\mathbb{P}(T_0=\infty)-\mathbb{P}(T_0>k)|\ &\leqslant\ rac{N}{n}+rac{n-N}{n}arepsilon\ &\leqslant\ rac{N}{n}+arepsilon. \end{aligned}$$

ර 514

Then, choosing $N_0 \ge 1$ such that $(N+1)/n \le \varepsilon$ for $n \ge N_0$, we get

$$\left|\mathbb{P}(T_0=\infty)-rac{1}{n}\,\mathbb{E}[R_n]
ight|\leqslant rac{1}{n}+\left|\mathbb{P}(T_0=\infty)-rac{1}{n}\sum\limits_{k=1}^n\mathbb{P}(T_0>k)
ight|\leqslant 2arepsilon,$$

 $n \ge N_0$, which concludes the proof.

Alternatively, the answer to that question can be derived by applying the Cesàro theorem, which states that in general we have

$$\lim_{n o \infty} rac{1}{n} \sum_{k=0}^n a_k = a$$

when the sequence $(a_k)_{k\in\mathbb{N}}$ has the limit a, by taking $a_k = \mathbb{P}(T_0 > k), k \in \mathbb{N}$, since we have $\lim_{k\to\infty} \mathbb{P}(T_0 > k) = \mathbb{P}(T_0 = \infty)$.

h) From Relation (3.4.15) in Section 3.4 we have

$$\mathbb{P}(T_0=+\infty)=|p-q|,$$

hence by the result of Question (g) we get

$$\lim_{n o\infty}rac{1}{n}\,\mathbb{E}[R_n]=|p-q|,$$

when $p \neq q$, and

$$\lim_{n o\infty}rac{1}{n}\mathbb{E}[R_n]=0,$$

when p = q = 1/2.

Problem 3.11 The question of recurrence of the d-dimensional random walk has been first solved in [Pol21], cf. [CMS07] for the solution proposed in this problem.

a) We partition the event $\{S_n = \vec{0}\}$ into

$$\{S_n=ec{0}\} = igcup_{k=2}^n \{S_{n-k}=ec{0}, \; S_{n-k+1}
eq ec{0}, \ldots, S_{n-1}
eq ec{0}, \; S_n=ec{0}\}, \; \; n \geqslant 1,$$

according to the time of *last* return to state $\vec{0}$ before time n, with $\mathbb{P}(\{S_1 = \vec{0}\}) = 0$ since we are starting from $S_0 = \vec{0}_{\text{H3}}$ then we

$$\begin{split} \mathbb{P}(S_n = \vec{0}) &:= \mathbb{P}(S_n = \vec{0} \mid S_0 = \vec{0}) \\ &= \sum_{k=2}^n \mathbb{P}(S_{n-k} = \vec{0}, \ S_{n-k+1} \neq \vec{0}, \dots, S_{n-1} \neq \vec{0}, \ S_n = \vec{0} \mid S_0 = \vec{0}) \\ &= \sum_{k=2}^n \mathbb{P}(S_{n-k+1} \neq \vec{0}, \dots, S_{n-1} \neq \vec{0}, \ S_n = \vec{0} \mid S_{n-k} = \vec{0}, \ S_0 = \vec{0}) \\ &\qquad \times \mathbb{P}(S_{n-k} = \vec{0} \mid S_0 = \vec{0}) \\ &= \sum_{k=2}^n \mathbb{P}(S_1 \neq \vec{0}, \dots, S_{k-1} \neq \vec{0}, \ S_k = \vec{0} \mid S_0 = \vec{0}) \mathbb{P}(S_{n-k} = \vec{0} \mid S_0 = \vec{0}) \\ &= \sum_{k=2}^n \mathbb{P}(T_{\vec{0}}^r = k \mid S_0 = \vec{0}) \mathbb{P}(S_{n-k} = \vec{0} \mid S_0 = \vec{0}) \\ &= \sum_{k=2}^n \mathbb{P}(S_{n-k} = \vec{0}) \mathbb{P}(T_{\vec{0}}^r = k), \qquad n \ge 1. \end{split}$$

b) We have

$$\sum_{n=1}^{m} \mathbb{P}(S_n = \vec{0}) = \sum_{n=1}^{m} \sum_{k=2}^{n} \mathbb{P}(T_{\vec{0}}^r = k) \mathbb{P}(S_{n-k} = \vec{0})$$
$$= \sum_{k=2}^{m} \sum_{n=k}^{m} \mathbb{P}(T_{\vec{0}}^r = k) \mathbb{P}(S_{n-k} = \vec{0})$$
$$= \sum_{k=2}^{m} \mathbb{P}(T_{\vec{0}}^r = k) \sum_{l=0}^{m-k} \mathbb{P}(S_l = \vec{0})$$
$$\leqslant \sum_{k=2}^{m} \mathbb{P}(T_{\vec{0}}^r = k) \sum_{l=0}^{m} \mathbb{P}(S_l = \vec{0})$$
$$= \left(\sum_{n=0}^{m} \mathbb{P}(S_n = \vec{0})\right) \left(\sum_{n=2}^{m} \mathbb{P}(T_{\vec{0}}^r = n)\right)$$

c) By letting m tend to ∞ in (3.4.30) we get

$$\mathbb{P}ig(T^r_{ec 0} < \inftyig) = \sum\limits_{n \geqslant 2} \mathbb{P}ig(T^r_{ec 0} = nig) \geqslant 1,$$
 MH3512 AY19-20

•

ॳ 516

which allows us to conclude. Note that the sum of the series $\sum_{n \ge 0} \mathbb{P}(S_n = \vec{0})$ actually represents the average number of visits to state $\vec{0}$.

d) We have

$$\begin{split} \sum_{n=1}^{2m} \mathbb{P}\big(S_n = \vec{0}\big) &= \sum_{n=2}^{2m} \mathbb{P}\big(T_{\vec{0}}^r = n\big) \sum_{l=0}^{2m-n} \mathbb{P}\big(S_l = \vec{0}\big) \\ &\geqslant \sum_{n=2}^{m} \mathbb{P}\big(T_{\vec{0}}^r = n\big) \sum_{l=0}^{2m-n} \mathbb{P}\big(S_l = \vec{0}\big) \\ &\geqslant \sum_{n=2}^{m} \mathbb{P}\big(T_{\vec{0}}^r = n\big) \sum_{l=0}^{m} \mathbb{P}\big(S_l = \vec{0}\big). \end{split}$$

e) Letting m tend to $+\infty$ in (3.4.31) we find

$$\mathbb{P}ig(T^r_{ec{0}} < \inftyig) = \sum_{n \geqslant 2} \mathbb{P}ig(T^r_{ec{0}} = nig) \leqslant rac{\sum\limits_{n \geqslant 2} \mathbb{P}ig(S_n = ec{0}ig)}{\sum\limits_{n \geqslant 0} \mathbb{P}ig(S_n = ec{0}ig)} = 1 - rac{1}{\sum\limits_{n \geqslant 0} \mathbb{P}ig(S_n = ec{0}ig)} < 1.$$

f) When $\underline{d=1}$ we have

$$\mathbb{P}(S_{2n}=0) = rac{1}{2^{2n}} {2n \choose n} = rac{(2n)!}{2^{2n}(n!)^2} \simeq_{n o \infty} rac{1}{\sqrt{\pi n}},$$

by Stirling's approximation, hence

$$\sum\limits_{n \geqslant 0} \mathbb{P}(S_n = 0) = \infty.$$

and the result of Question (c) shows that $\mathbb{P}(T_0^r < \infty) = 1$.

- g) We note that the random walk can return to state $\vec{0}$ in 2n time steps by
 - k forward steps in the direction e_1 ,
 - k backward steps in the direction $-e_1$,
 - n-k forward steps in the direction e_2 ,
 - n-k backward steps in the direction $-e_2$,

where k ranges from 0 to 2n. For each k = 0, 1, ..., n the number of ways to arrange those four types of moves among 2n time steps is the multinomial coefficient

$$egin{pmatrix} 2n \ k,k,n-k,n-k \end{pmatrix} = rac{(2n!)}{k!k!(n-k)!(n-k)!},$$

hence, since every sequence of 2n moves occur with the same probability $(1/4)^{2n}$, by summation over $k = 0, 1, \ldots, n$ we find

$$egin{aligned} \mathbb{P}ig(S_{2n} = ec{0}ig) &= \left(rac{1}{4}
ight)^{2n}\sum\limits_{k=0}^{n}rac{(2n!)}{(k!)^2((n-k)!)^2} \ &= rac{(2n)!}{4^{2n}(n!)^2}\sum\limits_{k=0}^{n}ig(rac{n}{k}ig)^2 \ &= rac{(2n)!}{4^{2n}(n!)^2}ig(rac{2n}{n}ig) \ &= rac{((2n)!)^2}{4^{2n}(n!)^2} \simeq_{n o \infty}rac{1}{2\pi n}, \end{aligned}$$

where we used Stirling's approximation, and this yields

$$\sum_{n \geqslant 0} \mathbb{P}(S_n = \vec{0}) = \infty.$$

To conclude, the result of Question (c) shows again that $\mathbb{P}(T^r_{\vec{0}} < \infty) = 1$.

h) In order to come back to $\vec{0}$ we need to take i_1 forward steps in the direction e_1 and i_1 backward steps in the direction e_1 , and similarly for i_2, \ldots, i_d . The number of ways to arrange such paths is given by the multinomial coefficient

$$igg(rac{2n}{(i_1,i_1,i_2,i_2,\ldots,i_d,i_d} igg) = rac{(2n)!}{(i_1!)^2\cdots(i_d!)^2},$$

and by summation over all possible indices $i_1, i_2, \ldots, i_d \ge 0$ satisfying $i_1 + \cdots + i_d = n$ and multiplying by the probability $(1/(2d))^{2n}$ of each path we find

$$\mathbb{P}ig| S_{2n} = ec{0}ig) \; = \; rac{1}{(2d)^{2n}} \sum_{i_1 + \cdots + i_d = n top i_1, i_2, \ldots, i_d \geqslant 0} igg(rac{2n}{i_1, i_1, i_2, i_2, \ldots, i_d, i_d} igg) \; .$$

$$= \; rac{1}{(2d)^{2n}} \sum_{i_1+\dots+i_d=n top i_1,i_2,\dots,i_d \geqslant 0} rac{(2n)!}{(i_1!)^2 \cdots (i_d!)^2}.$$

i) By the hint provided that we have

$$egin{aligned} &\sum_{n \geqslant 1} \mathbb{P}ig(S_{2n} = ec{0}ig) \; = \; \sum_{n \geqslant 1} rac{1}{(2d)^{2n}} inom{2n}{n!} \sum_{\substack{i_1 + \cdots + i_d = n \ i_1, i_2, \dots, i_d \geqslant 0}} rac{(n!)^2}{(i_1!)^2 \cdots (i_d!)^2} \ &\leqslant \; \sum_{n \geqslant 1} rac{1}{(2d)^{2n}} inom{2n}{n!} rac{n!}{(a_n!)^d (a_n + 1)^{b_n}} \sum_{\substack{i_1 + \cdots + i_d = n \ i_1, i_2, \dots, i_d \geqslant 0}} rac{n!}{i_1! \cdots i_d!} \ &\leqslant \; \sum_{n \geqslant 1} rac{1}{(2d)^{2n}} inom{2n}{n!} rac{n!d^n}{(a_n!)^d a_n^{b_n}} = \sum_{n \geqslant 1} rac{(2n)!}{2^{2n} d^n n! (a_n!)^d a_n^{b_n}}. \end{aligned}$$

j) For sufficiently large \boldsymbol{n} , by the Stirling approximation we have

$$\begin{aligned} \frac{(2n)!}{2^{2n}d^n n!(a_n!)^d a_n^{b_n}} &\simeq \frac{(2n/e)^{2n}\sqrt{4\pi n}}{2^{2n}d^n (n/e)^n \sqrt{2\pi n} ((a_n/e)^{a_n}\sqrt{2\pi a_n})^d a_n^{b_n}} \\ &= \frac{\sqrt{2}}{(2\pi)^{d/2}} \frac{n^n}{\mathrm{e}^{b_n} (a_n d)^n a_n^{d/2}} \\ &= \frac{\sqrt{2}}{(2\pi)^{d/2}} \frac{(1-b_n/n)^{-n}}{\mathrm{e}^{b_n} a_n^{d/2}} \\ &\leqslant \frac{\sqrt{2}d^{d/2}}{(2\pi)^{d/2}} \frac{(1-(d-1)/n)^{-n}}{(a_n d)^{d/2}} \\ &\simeq \frac{\sqrt{2}d^{d/2}\mathrm{e}^{d-1}}{(2\pi)^{d/2}} \frac{1}{n^{d/2}}, \end{aligned}$$

since $a_n d \simeq n$ as n goes to infinity from the relation $a_n d/n = 1 - b_n/n$.

k) The result of Question (j) shows that

$$\sum_{n \geqslant 0} \mathbb{P}(S_n = \vec{0}) < \infty,$$
 MH3512 AY19-20

519

hence $\mathbb{P}(T_{\vec{0}}^r = \infty) > 0$ by the result of Question (e), and the random walk is *not* recurrent when $\underline{d \geq 3}$.

Problem 3.12

- a) The cookie random walk does not have the Markov property when $p \neq 1/2$ because in this case the transition probabilities at a given state may depend on the past behavior of the chain starting from time 1. On the other hand, the cookie random walk has the Markov property when p = 1/2 because in this case it coincides with the usual symmetric random walk with independent increments.
- b) Knowing that the probability for a symmetric random walk to reach state (x+1) before hitting state (0) starting from (k) is k/(x+1), by first step analysis we find the probability

$$p+(1-p)rac{x-1}{x+1}=1-rac{2q}{x+1},\qquad x\geqslant 1.$$

c) We have $\mathbb{P}(\tau_1 < \tau_0 \mid S_0 = 0) = 1/2$ and by the (strong) Markov property, by reasoning inductively on the transitions from state (0) to state (1), then from state (2) to state (2), etc, up to state (x), we find

$$egin{aligned} \mathbb{P}(au_x < au_0 \mid S_0 = 0) &= \prod_{l=0}^{x-1} \mathbb{P}(au_{l+1} < au_0 \mid S_0 = l) \ &= \mathbb{P}(au_1 < au_0 \mid S_0 = 0) \prod_{l=1}^{x-1} \mathbb{P}(au_{l+1} < au_0 \mid S_0 = l) \ &= rac{1}{2} \prod_{l=1}^{x-1} \left(1 - rac{2q}{l+1}
ight) \ &= rac{1}{2} \prod_{l=2}^{x} \left(1 - rac{2q}{l}
ight), \qquad x \geqslant 1. \end{aligned}$$

d) Using the inequality $\log(1+z) \leq z$ for z > -1, we have

$$egin{array}{l} \sum\limits_{l=2}^x \log\left(1-rac{2q}{l}
ight) &\leqslant & \int_2^x \log\left(1-rac{2q}{y}
ight) dy \ &\leqslant & -2q\int_2^x rac{1}{y} dy \end{array}$$

MH3512 AY19-20 $\,$

 $= -2q\lograc{x}{2},$

hence

$$egin{aligned} \mathbb{P}(au_x < au_0 \mid S_0 = 0) &= \; rac{1}{2} \exp\left(\sum\limits_{l=2}^x \left(1 - rac{2q}{l}
ight)
ight) \ &\leqslant \; \exp\left(-2q\lograc{x}{2}
ight) \leqslant \left(rac{x}{2}
ight)^{-2q}, \end{aligned}$$

 $x \geqslant 2$.

e) We have

$$egin{aligned} \mathbb{P}(au_0 < \infty \mid S_0 = 0) & \geqslant & \mathbb{P}\left(igcup_{x \geqslant 1} au_0 < au_x \mid S_0 = 0
ight) \ &= & \lim_{x o \infty} \mathbb{P}(au_0 < au_x \mid S_0 = 0) \ &= & \lim_{x o \infty} (1 - \mathbb{P}(au_x \leqslant au_0 \mid S_0 = 0)) \ &= & \lim_{x o \infty} (1 - \mathbb{P}(au_x < au_0 \mid S_0 = 0)) \ &\geqslant & 1 - \lim_{x o \infty} (x/2)^{-2q} \ &= & 1. \end{aligned}$$

Remark. We note that $\mathbb{P}(\tau_x < \infty \mid S_0 = 0) = 1$, as can be seen by bounding from Question (a) of Problem 2.7 with rebound. Note also that $\mathbb{P}(B \cap A) = \mathbb{P}(B)$ when $\mathbb{P}(A) = 1$.

f) This question can be solved in three possible ways.

i) By first step analysis. We have

$$\mathbb{E}[au_1 \mid S_0 = 0] = rac{1}{2} imes 1 + rac{1}{2}(1 + \mathbb{E}[au_1 \mid S_0 = 0]),$$

hence

$$\mathbb{E}[au_1 \mid S_0 = 0] = 2.$$
 MH3512 AY19-20

ii) By pathwise analysis. We have

$$\mathbb{E}[au_1 \mid S_0 = 0] = \sum\limits_{k \geqslant 1} rac{k}{2^k} = rac{1}{2} \sum\limits_{k \geqslant 1} rac{k}{2^{k-1}} = rac{1/2}{(1-1/2)^2} = 2.$$

iii) By applying the result of Question (d) of Problem 2.7 page 100 with S = 1 and k = 0, which shows that

$$\mathbb{E}[au_1 \mid S_0 = 0] = (S+k+1)(S-k) = 2.$$

g) Applying the result of Question (d) of Problem 2.7 together with first step analysis, we find the mean time

$$p + q(1 + (x + 1 + (x - 1) + 1)(x + 1 - (x - 1))) = p + q(1 + 4x + 2) = 1 + q(4x + 2).$$

h) We have

$$egin{aligned} \mathbb{E}[au_x \mid S_0 = 0] &= 2 + \sum\limits_{k=1}^{x-1} (1 + q(4k+2)) \ &= 2 + (1+2q)(x-1) + 4q \sum\limits_{k=1}^{x-1} k \ &= 2 + (1+2q)(x-1) + 2qx(x-1) \ &= 2 + (1+2q)x - (1+2q) + 2qx^2 - 2qx \ &= 1 - 2q + x + 2qx^2, \quad x \geqslant 1. \end{aligned}$$

i) Using the notation of Problem 2.7 we have

$$\mathbb{P}(S_1 = x + 1 \mid S_0 = x \text{ and } \tau_{x+1} < \tau_0) = p \frac{\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x + 1)}{\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x)}$$
$$= \frac{p}{\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x)}$$

and

$$\mathbb{P}(S_1 = x - 1 \mid S_0 = x, au_{x+1} < au_0) \; = \; q rac{\mathbb{P}(au_{x+1} < au_0 \mid S_0 = x - 1)}{\mathbb{P}(au_{x+1} < au_0 \mid S_0 = x)}$$

 $= \frac{q(x-1)/(x+1)}{\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x)},$ (523)

as we have $\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x+1) = 1$ and

$$\mathbb{P}(au_{x+1} < au_0 \mid S_0 = x-1) = rac{x-1}{x+1},$$

because the random walk becomes symmetric when started from state x-1. Next, we note that, according to Question (b), $\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x)$ can be computed as

$$egin{array}{rl} \mathbb{P}(au_{x+1} < au_0 \mid S_0 = x) &= p \mathbb{P}(au_{x+1} < au_0 \mid S_1 = x+1) \ &+ q \mathbb{P}(au_{x+1} < au_0 \mid S_1 = x-1) \ &= p + q rac{x-1}{x+1}, \end{array}$$

and in this case we get

$$\mathbb{P}(S_1 = x + 1 \mid S_0 = x \text{ and } \tau_{x+1} < \tau_0) = \frac{p}{p + q(x-1)/(x+1)} = \frac{p}{1 - 2q/(x+1)}$$

and

$$\mathbb{P}(S_1 = x - 1 \mid S_0 = x \text{ and } \tau_{x+1} < \tau_0) = \frac{q(x - 1)/(x + 1)}{p + q(x - 1)/(x + 1)} = \frac{q(x - 1)/(x + 1)}{1 - 2q/(x + 1)}$$

j) The mean time is given by

$$\begin{split} \mathbb{E}[\tau_{x+1} \mid S_0 = x, \tau_{x+1} < \tau_0] &= 1 + \frac{((x+1)^2 - (x-1)^2)}{3} \frac{q(x-1)/(x+1)}{1 - 2q/(x+1)} \\ &= 1 + \frac{4qx(x-1)/(x+1)}{3(1 - 2q/(x+1))} \\ &= 1 + \frac{4qx(x-1)}{3(x+1 - 2q)}, \end{split}$$

which yields 1 + (2x - 2)/3 when p = q = 1/2.

ॳ 524

Remark. Using the above result we can also compute the average time to reach state \hat{x} with $x \ge 2$ after starting from state $\hat{1}$ and given one does not hit $\hat{0}$ by the summation

$$\begin{split} \mathbb{E}[\tau_x \mid S_0 &= 1, \ \tau_x < \tau_0] = \sum_{k=2}^x \mathbb{E}[\tau_k \mid S_0 = k - 1, \tau_k < \tau_0] \\ &= x - 1 + \frac{4q}{3} \sum_{k=2}^x \frac{(k-1)(k-2)}{k-2q} \\ &= x - 1 + \frac{2qx(x-1)}{3} - \frac{8p\tilde{q}}{3} \sum_{k=1}^{x-1} \frac{k}{k+1-2q} \\ &= x - 1 + \frac{2qx(x-1)}{3} - \frac{8pq}{3} (x-1) + \frac{8pq}{3} (p-q) \sum_{k=1}^{x-1} \frac{1}{k+1-2q} \\ &= x - 1 + \frac{4q}{3} \left(\frac{x(x-1)}{2} - 2p(x-1) + 2p(p-q) \sum_{k=1}^{x-1} \frac{1}{k+1-2q} \right). \end{split}$$

When p = q = 1/2 we recover the known expression

$$\mathbb{E}[au_x \mid S_0 = 1, \; au_x < au_0] = x - 1 + rac{2}{3} \sum\limits_{k=1}^{x-2} k \ = x - 1 + rac{(x-1)(x-2)}{3} \ = rac{x^2 - 1}{3}, \quad x \geqslant 2,$$

cf. Question (j) of Problem 2.7.

Problem 3.13

a) The probability $\mathbb{P}(X = 0)$ that the random walk eats no cookies before hitting the origin is the probability of going directly from (1) to (0) in one time step, which is 1/2.

The probability $\mathbb{P}(X = 1)$ that the random walk eats exactly *one* cookie before hitting the origin is the probability of first moving

from 0 to 1 in one time step and then back to 0 in one time step, that is $q \times (1/2) = q/2$.

In general we have

$$egin{aligned} \mathbb{P}(X=x) &= \mathbb{P}(au_x < au_0 \mid S_0 = 0) - \mathbb{P}(au_{x+1} < au_0 \mid S_0 = 0) \ &= rac{1}{2} \prod\limits_{l=2}^x \left(1 - rac{2q}{l}
ight) - rac{1}{2} \prod\limits_{l=2}^{x+1} \left(1 - rac{2q}{l}
ight), \ &= rac{1}{2} \left(1 - \left(1 - rac{2q}{x+1}
ight)
ight) \prod\limits_{l=2}^x \left(1 - rac{2q}{l}
ight) \ &= rac{q}{x+1} \prod\limits_{l=2}^x \left(1 - rac{2q}{l}
ight). \end{aligned}$$

b) We have

$$\mathbb{E}[X] = \sum\limits_{x \geqslant 0} x \mathbb{P}(X=x) = q \sum\limits_{x \geqslant 0} rac{x}{x+1} \prod\limits_{l=2}^x \left(1-rac{2q}{l}
ight),$$

hence

$$qc_q\sum\limits_{x\geqslant 0}rac{x}{(x+1)x^{2q}}\leqslant \mathbb{E}[X]\leqslant qC_q\sum\limits_{x\geqslant 0}rac{x}{(x+1)x^{2q}},$$

and $\mathbb{E}[X]$ is finite if and only if 2q > 1.

Remark. One could show in addition that the mean return time to 0 is always infinite, see [AR05].

Chapter 4 - Discrete-Time Markov Chains

Exercise 4.1

a) This process is a Markov chain because its increments $Z_{n+1} - Z_n = 2(S_{n+1} - S_n)$, $n \in \mathbb{N}$, are independent random variables. b) We have

$$egin{array}{rcl} Z_{n+1} &=& (S_{n+1})^2 \ &=& (S_{n+1}-S_n+S_n)^2 \end{array} \end{array} egin{array}{rcl} & ext{MH3512 AY19-2} \end{array}$$

$$= (S_n)^2 + 2(S_{n+1} - S_n)S_n + (S_{n+1} - S_n)^2$$

$$= (S_n)^2 + 2(S_{n+1} - S_n)S_n + 1$$

$$= Z_n + 2X_{n+1}S_n + 1,$$

(526)

where $(X_n)_{n \ge 1} = (S_n - S_{n-1})_{n \ge 1}$ denotes the independent increments of $(S_n)_{n \ge 0}$, and the increment of $(Z_n)_{n \in \mathbb{N}}$ is

$$Z_{n+1}-Z_n=2X_{n+1}S_n+1,\qquad n\in\mathbb{N}.$$

In the symmetric case $\mathbb{P}(X_{n+1} = \pm 1) = 1/2$, the process $(Z_n)_{n \ge 0}$ is a Markov chain because $X_{n+1}S_n$ takes the values

$$\{S_n,-S_n\}=\{\sqrt{Z_n},-\sqrt{Z_n}\}$$

with probabilities 1/2, hence the probability distribution of $X_{n+1}S_n$ is always (1/2, 1/2) over the two values $\{\sqrt{Z_n}, -\sqrt{Z_n}\}$, and the distribution of Z_{n+1} is fully determined given the value of $Z_n = (S_n)^2$.

In the non symmetric case, with

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

the process $(Z_n)_{n \ge 0}$ will not be a Markov chain because when $S_n > 0$ we have

$$(S_n,-S_n)=(\sqrt{Z_n},-\sqrt{Z_n})$$

and the probability distribution of $X_{n+1}S_n$ over $(\sqrt{Z_n}, -\sqrt{Z_n})$ is (p,q). On the other hand, when $S_n < 0$ we have

$$(S_n,-S_n)=(-\sqrt{Z_n},\sqrt{Z_n})$$

and the probability distribution of $X_{n+1}S_n$ over $(\sqrt{Z_n}, -\sqrt{Z_n})$ becomes (q, p) which differs from (p, q) since $p \neq q$. Hence the sign of S_n is needed and the knowledge of $Z_n = |S_n|$ is no longer sufficient in order to determine the distribution of Z_{n+1} given Z_n , therefore $(Z_n)_{n\geq 0}$ is not a Markov chain.

$$\begin{split} \mathbb{P}(X_7 &= 1 \text{ and } X_5 = 2 \mid X_4 = 1 \text{ and } X_3 = 2) \\ &= \mathbb{P}(X_7 = 1 \text{ and } X_5 = 2 \mid X_4 = 1) \\ &= \frac{\mathbb{P}(X_7 = 1, \ X_5 = 2, \ X_4 = 1)}{\mathbb{P}(X_4 = 1)} \\ &= \frac{\mathbb{P}(X_7 = 1, \ X_5 = 2, \ X_4 = 1)}{\mathbb{P}(X_5 = 2 \text{ and } X_4 = 1)} \frac{\mathbb{P}(X_5 = 2 \text{ and } X_4 = 1)}{\mathbb{P}(X_4 = 1)} \\ &= \mathbb{P}(X_7 = 1 \mid X_5 = 2 \text{ and } X_4 = 1) \mathbb{P}(X_5 = 2 \mid X_4 = 1) \\ &= 0.6 \times \mathbb{P}(X_7 = 1 \mid X_5 = 2). \end{split}$$

Next, we note that

$$P^2 = egin{bmatrix} 0.4 & 0.6 \ & & \ 0.8 & 0.2 \end{bmatrix} imes egin{bmatrix} 0.4 & 0.6 \ & & \ 0.8 & 0.2 \end{bmatrix} = egin{bmatrix} 0.64 & 0.36 \ & & \ 0.48 & 0.52 \end{bmatrix},$$

hence $\mathbb{P}(X_7 = 1 \mid X_5 = 2) = 0.48$, and we conclude that

$$\mathbb{P}(X_7 = 1 \text{ and } X_5 = 2 \mid X_4 = 1 \text{ and } X_3 = 2) = 0.6 \times 0.48 = 0.288.$$

b) We have

$$\mathbb{E}[X_2 \mid X_1 = 1] = 1 \times \mathbb{P}(X_2 = 1 \mid X_1 = 1) + 2 \times \mathbb{P}(X_2 = 2 \mid X_1 = 1) = 0.4 + 2 \times 0.6 = 1.6.$$

(B.28)

$$P^{n} = P = \begin{bmatrix} \pi_{0} & \pi_{1} & \pi_{2} & \pi_{3} & \cdots & \pi_{N} \\ \pi_{0} & \pi_{1} & \pi_{2} & \pi_{3} & \cdots & \pi_{N} \\ \pi_{0} & \pi_{1} & \pi_{2} & \pi_{3} & \cdots & \pi_{N} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \pi_{0} & \pi_{1} & \pi_{2} & \pi_{3} & \cdots & \pi_{N} \end{bmatrix}$$

for all $n \ge 1$.

b) The vector $\boldsymbol{\pi}$ is an invariant (or stationary) distribution for \boldsymbol{P} because it satisfies

$$[\pi_0, \pi_1, \dots, \pi_N] = [\pi_0, \dots, \pi_N] imes egin{bmatrix} \pi_0 & \pi_1 & \pi_2 & \cdots & \pi_N \ \pi_0 & \pi_1 & \pi_2 & \cdots & \pi_N \ dots & dot$$

as in (4.2.5).

c) Since $P^n = P$, $n \ge 1$, we have

$$\mathbb{P}(Z_n = i \mid Z_k = j) = \mathbb{P}(Z_n = i) = \pi_i, \hspace{1em} i, j = 0, 1, \dots, N, \hspace{1em} 0 \leqslant k < n,$$

hence by the Markov property we find

$$\mathbb{P}(Z_n = i \text{ and } Z_k = j) = \mathbb{P}(Z_n = i \mid Z_k = j) \mathbb{P}(Z_k = j)$$

$$= \mathbb{P}(Z_n = i) \mathbb{P}(Z_k = j),$$
(B.29)
(B.29)
(B.29)
(B.29)

,

i, j = 0, 1, ..., N, which shows that Z_n is independent of Z_k for $0 \leq k < n$. Similarly, using induction and the Markov property, or (4.1.1), shows that

$$egin{aligned} \mathbb{P}(Z_n = i_n, \ Z_{n-1} = i_{n-1}, \dots, Z_0 = i_0) \ &= \ \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1}) \cdots \mathbb{P}(Z_1 = i_1 \mid Z_0 = i_0) \mathbb{P}(Z_0 = i_0) \ &= \ \mathbb{P}(Z_n = i_n) \mathbb{P}(Z_{n-1} = i_{n-1}) \cdots \mathbb{P}(Z_0 = i_0) \ &= \ \pi_{i_n} \pi_{i_{n-1}} \cdots \pi_{i_0}, \end{aligned}$$

 $i_0, i_1, \ldots, i_n \in \{0, 1, \ldots, N\}$, which shows that $(Z_n)_{n \in \mathbb{N}}$ is an *i.i.d* sequence of random variables with distribution π over $\{0, 1, \ldots, N\}$.

Exercise 4.4

a) We have

$$\mathbb{P}((O_0, O_1) = (a, b) \text{ and } (X_0, X_1) = (1, 1))$$

= $\mathbb{P}((O_0, O_1) = (a, b) \mid (X_0, X_1) = (1, 1))\mathbb{P}((X_0, X_1) = (1, 1))$
= $\mathbb{P}(O_0 = a \mid X_0 = 1)\mathbb{P}(O_1 = b \mid X_1 = 1)\mathbb{P}((X_0, X_1) = (1, 1))$
= $\pi_1 P_{1,1} m_{1,a} m_{1,b},$

since

$$\mathbb{P}((X_0,X_1)=(1,1))=\pi_1P_{1,1}.$$

b) We have

$$\begin{split} \mathbb{P}((O_0,O_1) &= (a,b)) \\ &= \sum_{x,y \in \{0,1\}} \mathbb{P}((O_0,O_1) = (a,b) \text{ and } (X_0,X_1) = (x,y)) \\ &= \sum_{x,y \in \{0,1\}} \pi_x P_{x,y} m_{x,a} m_{y,b} \\ &= \pi_0 P_{0,0} m_{0,a} m_{0,b} + \pi_0 P_{0,1} m_{0,a} m_{1,b} + \pi_1 P_{1,0} m_{1,a} m_{0,b} + \pi_1 P_{1,1} m_{1,a} m_{1,b}. \end{split}$$

c) We have

$$ig\{X_1=1ig\} \ = \ ig\{(X_0,X_1)=(0,1)ig\}ig\cupig\{(X_0,X_1)=(0,1)ig\}$$

$$igcup_{x\in\{0,1\}} igcup_{x\in\{0,1\}} igl(X_0,X_1)=(1,1)igr\} igl(X_0,X_1)=(1,1)igr\} = igl\{(X_0,X_1)=(x,1)igr\} = igl\{(X_0,X_1)=(0,1)igr\} igrt\{(X_0,X_1)=(1,1)igr\},$$

where the above union is a partition, hence

$$\begin{split} & \mathbb{P}(X_1 = 1 \mid (O_0, O_1) = (a, b)) \\ &= \sum_{x \in \{0,1\}} \mathbb{P}((X_0, X_1) = (x, 1) \mid (O_0, O_1) = (a, b)) \\ &= \frac{1}{\mathbb{P}((O_0, O_1) = (a, b))} \sum_{x \in \{0,1\}} \mathbb{P}((X_0, X_1) = (x, 1) \text{ and } (O_0, O_1) = (a, b)) \\ &= \frac{1}{\mathbb{P}((O_0, O_1) = (a, b))} \sum_{x \in \{0,1\}} \pi_x P_{x,1} m_{x,a} m_{1,b} \\ &= \frac{1}{\mathbb{P}((O_0, O_1) = (a, b))} \left(\pi_0 P_{0,1} m_{0,a} m_{1,b} + \pi_1 P_{1,1} m_{1,a} m_{1,b} \right). \end{split}$$

Exercise 4.5

- a) The resulting system cannot be a Markov chain because the behavior of the chain at time n + 1 may depend on the past of the chain from time 0 to time n 1. Indeed, the fact that have visited a given state between time 0 to time n 1 will affect the probability of visiting that state at time n.
- b) The process $(Z_n)_{n \in \mathbb{N}}$ cannot be a Markov chain because when e.g. $Z_n = 25$ the distribution of the next value of the chain at time n + 1 is not uniquely determined from the data of $Z_n = 25$. Indeed, it may depend on how the state 25 has been reached, e.g. through (5, 0) or (4, 3), and this information will influence the next transition probabilities, which can then can depend on past chain values from time 0. Other examples include (99, 101) and (11, 141) whose sums of squares equal 20002.

Exercise 4.6 The Elephant Random Walk $(S_n)_{n \in \mathbb{N}}$ does *not* have the Markov property because the distribution of S_{n+1} given the value of S_n depends on extra information that relies on the past data of (X_1, \ldots, X_{n-1}) .

Exercise 4.7 The state space of $(Z_n)_{n \ge 1}$ is $\{(0,0), (0,1), (1,0), (1,1)\}$. Based on $Z_n = (X_{n-1}, X_n)$, the distribution of $Z_{n+1} = (X_n, X_{n+1})$ at time n + 1 is fully determined from the data of X_n and the transition matrix of $(X_n)_{n \in \mathbb{N}}$ hence $(Z_n)_{n \ge 1}$ is a (0,1)-valued Markov chain and its transition matrix is given by

531

Exercise 4.8

a) The probability distribution of T_2 is negative binomial with parameters (p, 2), *i.e.*

$$\mathbb{P}(T_2=k \mid X_0=0) = {k-1 \choose k-2} (1-p)^2 p^{k-2}, \qquad k \geqslant 2.$$

b) We have

$$\begin{split} \mathbb{E}[T_2 \mid X_0 = 0] &= \sum_{k \geqslant 2} k \mathbb{P}(T_2 = k \mid X_0 = 0) \\ &= (1 - p)^2 \sum_{k \geqslant 2} k \binom{k - 1}{k - 2} p^{k - 2} \\ &= (1 - p)^2 \sum_{k \geqslant 2} k (k - 1) p^{k - 2} \\ &= (1 - p)^2 \frac{\partial^2}{\partial p^2} \sum_{k \geqslant 0} p^k \\ &= (1 - p)^2 \frac{\partial^2}{\partial p^2} \frac{1}{1 - p} \\ &= \frac{2}{1 - p} \\ &= \frac{2}{q}. \end{split}$$

Remark. We could also recover the values $\mathbb{E}[T_1 \mid X_0 = 0]$ and $\mathbb{E}[T_2 \mid X_0 = 1]$ from the first step analysis equations

$$\mathbb{E}[T_2 \mid X_0 = 1] = p(1 + \mathbb{E}[T_2 \mid X_0 = 1]) + q imes 1.$$

and

$$\mathbb{E}[T_2 \mid X_0 = 0] = p(1 + \mathbb{E}[T_2 \mid X_0 = 0]) + q imes (1 + \mathbb{E}[T_2 \mid X_0 = 1]) \,,$$

which yield

$$\mathbb{E}[T_2 \mid X_0 = 1] = \frac{1}{q}$$
 and $\mathbb{E}[T_2 \mid X_0 = 0] = \frac{2}{q}$

Exercise 4.9 When N = 5, the transition matrix of this chain is

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1/25 & 8/25 & 16/25 & 0 & 0 & 0 \\ 0 & 4/25 & 12/25 & 9/25 & 0 & 0 \\ 0 & 0 & 9/25 & 12/25 & 4/25 & 0 \\ 0 & 0 & 0 & 16/25 & 8/25 & 1/25 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

For general $N \ge 1$, the transition matrix reads

$$P = \begin{vmatrix} 0 & 1 & 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ 1/N^2 & 2(N-1)/N^2 & (N-1)^2/N^2 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 2^2/N^2 & 4(N-2)/N^2 & (N-2)^2/N^2 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 3^2/N^2 & 0 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 3^2/N^2 & 0 & 0 \\ 0 & 0 & \cdots & 0 & (N-2)^2/N^2 & 4(N-2)/N^2 & 2^2/N^2 & 0 \\ 0 & 0 & \cdots & 0 & 0 & (N-1)^2/N^2 & 2(N-1)/N^2 & 1/N^2 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 1 & M_{3512} \text{ AY19-20} \end{vmatrix}$$

and the matrix entries are given in general as

$$\left\{egin{array}{l} P_{k,k-1} = \mathbb{P}(X_1 = k-1 \mid X_0 = k) = rac{k^2}{N^2}, \ P_{k,k} = \mathbb{P}(X_1 = k \mid X_0 = k) = rac{2k(N-k)^2}{N^2}, \ P_{k,k+1} = \mathbb{P}(X_1 = k+1 \mid X_0 = k) = rac{(N-k)^2}{N^2}, \end{array}
ight.$$

 $k=1,2,\ldots,N-1$, with $P_{0,1}=P_{N,N-1}=1$. We check that

$$P_{k,k-1} + P_{k,k} + P_{k,k+1} = rac{k^2 + 2k(N-k)^2 + (N-k)^2}{N^2} = 1, \quad k = 1, 2, \dots, N-1.$$

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

$$P = \left[\begin{array}{ccccc} P_{i,j} \end{array}
ight]_{i,j \in \mathbb{N}} = \left[egin{array}{ccccc} 1 & 0 & 0 & 0 & 0 & 0 & \cdots & \ q & 0 & p & 0 & 0 & 0 & \cdots & \ q & 0 & 0 & p & 0 & 0 & \cdots & \ q & 0 & 0 & 0 & p & 0 & \cdots & \ q & 0 & 0 & 0 & 0 & p & \cdots & \ dots & dots &$$

After \boldsymbol{n} time steps we have

$$\mathbb{P}(S_n=n+1\mid S_0=1)=p^n, \qquad n\geqslant 1,$$

and, by decomposing over n possible paths made of l steps up and one step down, $l = 0, 1, \ldots, n-1$, we find

$$\mathbb{P}(S_n = 0 \mid S_0 = 1) = \sum_{l=0}^{n-1} q p^l = q \frac{1-p^n}{1-p}$$

♦ 533

$$= 1 - p^{n}$$

= 1 - $\mathbb{P}(S_n = n + 1 \mid S_0 = 1), \quad n \ge 1.$

More generally, we have

$$\mathbb{P}(S_n=n+k\mid S_0=k)=p^n, \qquad k,n\geqslant 1,$$

and

$$egin{aligned} \mathbb{P}(S_n = 0 \mid S_0 = k) &= q \sum\limits_{l=0}^{n-1} p^l \ &= q rac{1-p^n}{1-p} \ &= 1-p^n \ &= 1-\mathbb{P}(S_n = n+k \mid S_0 = k), \qquad k, n \geqslant 1, \end{aligned}$$

by (A.2), hence \mathbf{P}^n is given by

$$P^{n} = [[P^{n}]_{i,j}]_{i,j \in \mathbb{N}} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 - p^{n} & 0 & \cdots & 0 & p^{n} & 0 & 0 & 0 & \cdots \\ 1 - p^{n} & 0 & \cdots & 0 & 0 & p^{n} & 0 & 0 & \cdots \\ 1 - p^{n} & 0 & \cdots & 0 & 0 & 0 & p^{n} & 0 & \cdots \\ 1 - p^{n} & 0 & \cdots & 0 & 0 & 0 & p^{n} & 0 & \cdots \\ 1 - p^{n} & 0 & \cdots & 0 & 0 & 0 & 0 & p^{n} & \cdots \\ \vdots & \ddots \end{bmatrix},$$
(B.30)

in which the n columns n^o 2 to n + 1 are identically 0. We can also check by matrix multiplication that this relation is consistent

with $P^{n+1} = P \times P^n$, $n \ge 1$, *i.e.*

from the relation $1 - p + p(1 - p^n) = 1 - p^{n+1}$, $n \in \mathbb{N}$.

b) In this case the transition matrix \boldsymbol{P} becomes

$$P = \left[\begin{array}{ccccc} P_{i,j} \end{array}
ight]_{i,j \in \mathbb{N}} = \left[egin{array}{cccc} q & p & 0 & 0 & 0 & 0 & \cdots & \ q & 0 & p & 0 & 0 & 0 & \cdots & \ q & 0 & 0 & p & 0 & 0 & \cdots & \ q & 0 & 0 & 0 & p & 0 & \cdots & \ q & 0 & 0 & 0 & 0 & p & \cdots & \ dots & dots &$$

with

(B.31)

$$P^2 = egin{bmatrix} q & qp & p^2 & 0 & 0 & 0 & \cdots \ q & qp & 0 & p^2 & 0 & 0 & \cdots \ q & qp & 0 & 0 & p^2 & 0 & \cdots \ q & qp & 0 & 0 & 0 & p^2 & \cdots \ q & qp & 0 & 0 & 0 & 0 & \cdots \ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \ \end{bmatrix},$$

and more generally, by induction on $n \geqslant 2$ we find

Exercise 4.11

a) We have

$$\mathbb{P}(Z_{n+1} = 2 \mid Z_n = 0 \text{ and } Z_{n-1} = 2) = \frac{\mathbb{P}(Z_{n+1} = 2, Z_n = 0, Z_{n-1} = 2)}{\mathbb{P}(Z_n = 0 \text{ and } Z_{n-1} = 2)}$$

$$= \frac{\mathbb{P}(X_{n+1} = 2, X_n \in \{0, 1\}, X_{n-1} = 2)}{\mathbb{P}(X_n \in \{0, 1\} \text{ and } X_{n-1} = 2)}$$

$$= \frac{\mathbb{P}(X_{n+1} = 2, X_n = 0, X_{n-1} = 2)}{\mathbb{P}(X_n \in \{0, 1\} \text{ and } X_{n-1} = 2)} + \frac{\mathbb{P}(X_{n+1} = 2, X_n = 1, X_{n-1} = 2)}{\mathbb{P}(X_n \in \{0, 1\} \text{ and } X_{n-1} = 2)}$$

$$= 0,$$

 $n \ge 1$. Next, we have

$$\mathbb{P}(Z_{n+1}=2 \mid Z_n=0 \text{ and } Z_{n-1}=3) = rac{\mathbb{P}(Z_{n+1}=2, \ Z_n=0, \ Z_{n-1}=3)}{\mathbb{P}(Z_n=0 \text{ and } Z_{n-1}=3)}$$

$$= \frac{\mathbb{P}(X_{n+1} = 2, X_n \in \{0, 1\}, X_{n-1} = 3)}{\mathbb{P}(X_n \in \{0, 1\} \text{ and } X_{n-1} = 3)}$$

$$= \frac{\mathbb{P}(X_{n+1} = 2, X_n = 0, X_{n-1} = 3)}{\mathbb{P}(X_n \in \{0, 1\} \text{ and } X_{n-1} = 3)} + \frac{\mathbb{P}(X_{n+1} = 2, X_n = 1, X_{n-1} = 3)}{\mathbb{P}(X_n \in \{0, 1\} \text{ and } X_{n-1} = 3)}$$

$$= \frac{\mathbb{P}(X_{n+1} = 2, X_n = 0, X_{n-1} = 3)}{\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3) + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)}$$

$$+ \frac{\mathbb{P}(X_{n+1} = 2, X_n = 1, X_{n-1} = 3)}{\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3) + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)}$$

$$= \frac{\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3) + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)}{\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3) + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)}$$

$$= \frac{\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3) + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)}{\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3) + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)}$$

$$= \frac{\mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3) + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)}{\mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3) + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)}$$

$$= \frac{\mathbb{P}(X_{n+1} = 2 \mid X_n = 0)}{1 + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)/\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3)}$$

$$= \frac{\mathbb{P}(X_{n+1} = 2 \mid X_n = 0)}{1 + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)/\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3)}$$

$$= \frac{\mathbb{P}(X_{n+1} = 2 \mid X_n = 0)}{1 + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)/\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3)}$$

 $n \geqslant 1$.

b) The process $(Z_n)_{n \in \mathbb{N}}$ is *not* a Markov chain because by Question (a) we have

$$0 = \mathbb{P}(Z_{n+1} = 2 \mid Z_n = 0 \text{ and } Z_{n-1} = 2)$$

$$\neq \mathbb{P}(Z_{n+1} = 2 \mid Z_n = 0 \text{ and } Z_{n-1} = 3)$$

$$= \frac{1}{4}.$$

MH3512 AY19-20

537

Exercise 4.12



a) We find

 $360\eta_L + 505\eta_M + 640\eta_H = 360 \times 50\% + 505 \times 40\% + 640 \times 10\% \simeq 446.$

b) We find

$$\begin{split} \eta P &= \left[50\%, 40\%, 10\% \right] \times \left[\begin{array}{ccc} 2/3 & 1/6 & 1/6 \\ 1/3 & 1/2 & 1/6 \\ 1/6 & 2/3 & 1/6 \end{array} \right] \\ &= \left[\frac{2 \times 0.5}{3} + \frac{0.4}{3} + \frac{0.1}{6}, \frac{0.5}{6} + \frac{0.4}{2} + \frac{2 \times 0.1}{3}, \frac{0.5}{6} + \frac{0.4}{6} + \frac{0.1}{6} \right] \\ &= \left[\frac{29}{60}, \frac{21}{60}, \frac{10}{60} \right] \\ &\simeq \left[48.3\%, 35\%, 16.7\% \right]. \end{split}$$

c) We find

$$360\eta_L + 505\eta_M + 640\eta_H = 360 \times \frac{29}{60} + 505 \times \frac{21}{60} + 640 \times \frac{10}{60} \simeq 457.$$

d) The equation $\pi = \pi P$ reads

$$[\pi_L,\pi_M,\pi_H] = [\pi_L,\pi_M,\pi_H] imes P = [\pi_L,\pi_M,\pi_H] imes egin{bmatrix} 8/12 & 2/12 & 2/12 \ 4/12 & 6/12 & 2/12 \ 2/12 & 8/12 & 2/12 \ 2/12 & 8/12 & 2/12 \ \end{bmatrix},$$

and it can be rewritten as

$$\pi(P-I_d)=\pi P-\pi=0,$$

i.e.,

$$egin{array}{rll} \pi(P-I_d) &=& [\pi_L,\pi_M,\pi_H](P-I_d) \ &=& [\pi_L,\pi_M,\pi_H] imes egin{bmatrix} -4/12 & 2/12 & 2/12 \ 4/12 & -6/12 & 2/12 \ 2/12 & 8/12 & -10/12 \end{array}$$

$$= \begin{bmatrix} 0\\0\\0\end{bmatrix},$$

which can be rewritten as

$$\left\{ egin{array}{l} \pi_L - 3\pi_M + 4\pi_H = 0 \ \pi_L + \pi_M - 5\pi_H = 0. \end{array}
ight.$$

Combining the above with the additional condition

$$\pi_L + \pi_M + \pi_H = 1,$$

we arrive at

$$[\pi_L, \pi_M, \pi_H] = \left[rac{11}{24}, rac{9}{24}, rac{4}{24}
ight] \simeq [45.8\%, 37.5\%, 16.7\%].$$

Problem 4.13

a) We have

$$P = \left[egin{array}{cc} 1 & 0 \ q & p \end{array}
ight],$$

with q := 1 - p.

b) We have

$$P = egin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \ q & p & 0 & \cdots & 0 & 0 \ 0 & q & p & \cdots & 0 & 0 \ dots & dots & \ddots & dots & dots$$

with q := 1 - p.

c) Clearly, the first row of P has to be $[1, 0, \ldots, 0]$ because state 0 is absorbing, and the remaining of the matrix can take the form

 $[\alpha, Q]$. In addition, every row of the $d \times (d+1)$ matrix $[\alpha, Q]$ has to add up to one, *i.e.*

$$lpha_k + \sum\limits_{l=1}^d Q_{k,l} = 1, \qquad k=1,2,\ldots,d,$$

which can be rewritten as

$$(I_d - Q)e = egin{bmatrix} 1 - Q_{1,1} & -Q_{1,2} & \cdots & -Q_{1,d} \ -Q_{1,1} & 1 - Q_{1,2} & \cdots & -Q_{1,d} \ dots & dots & \ddots & dots \ dots & dots & \ddots & dots \ dots & dots & dots & dots \ dots & dots & dots \ dots & dots & dots \ \ dots \ \ dots \ \ \ \dots \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$$

i.e.

$$\alpha = (I_d - Q)e.$$

d) Clearly, the conclusion holds for n = 0, and also at the rank n = 1 since $\alpha = (I_d - Q)e$. Next, we assume that the relation (4.5.9) holds at the rank $n \ge 0$. In this case, we have
$$= \begin{bmatrix} 1 & 0\\ (I_d - Q^{n+1})e & Q^{n+1} \end{bmatrix},$$
 $\circlearrowright 541$

since

$$lpha + Q(I_d - Q^n)e = (I_d - Q)e + (Q - Q^{n+1})e = (I_d - Q^{n+1})e.$$

e) We have

$$egin{aligned} \mathbb{P}(T_0 = n \mid X_0 = i) &= \mathbb{P}\left(igcup_{k=1}^d \left\{X_{n-1} = k, \ T_0 = n
ight\} \mid X_0 = i
ight) \ &= \sum\limits_{k=1}^d \mathbb{P}(X_{n-1} = k \mid X_0 = i)\mathbb{P}(X_n = 0 \mid X_{n-1} = k) \ &= \sum\limits_{k=1}^d [P^{n-1}]_{i,k} P_{k,0} \ &= \sum\limits_{k=1}^d lpha_k [Q^{n-1}]_{i,k} \ &= [Q^{n-1}lpha]_i, \qquad n \geqslant 1. \end{aligned}$$

f) We have

$$egin{aligned} \mathbb{P}(T_0=n) &=& \sum\limits_{i=1}^d \mathbb{P}(X_0=i)\mathbb{P}(T_0=n\mid X_0=i) \ &=& \sum\limits_{i=1}^d \sum\limits_{k=1}^d eta_i lpha_k Q_{i,k}^{n-1} \ &=& eta^ op Q^{n-1} lpha, \qquad n \geqslant 1. \end{aligned}$$

g) We have

$$egin{array}{rcl} \mathbb{P}(T_0 \leqslant n) &=& \sum\limits_{k=1}^n \mathbb{P}(T_0=k) \ &=& \sum\limits_{k=1}^n eta^ op Q^{k-1} lpha \end{array}$$

$$egin{array}{rll} &=& eta^{ op}(I_d-Q^n)(I_d-Q)^{-1}lpha \ &=& eta^{ op}(I_d-Q^n)e \ &=& 1-eta^{ op}Q^ne, \qquad n\geqslant 1. \end{array}$$

Alternatively, using the relation $\alpha = (I_d - Q)e$ and a telescopic sum, we recover this result as

$$egin{aligned} \mathbb{P}(T_0 \leqslant n) &= \sum\limits_{k=1}^n \mathbb{P}(T_0 = k) \ &= \sum\limits_{k=1}^n eta^ op Q^{k-1} lpha \ &= \sum\limits_{k=1}^n eta^ op Q^{k-1} (I_d - Q) e \ &= \sum\limits_{k=0}^{n-1} eta^ op Q^k e - \sum\limits_{k=1}^n eta^ op Q^k e \ &= eta^ op e \ &= eta^ op e - eta^ op Q^n e \ &= 1 - eta^ op Q^n e, \quad n \geqslant 1. \end{aligned}$$

h) We have

$$egin{aligned} \mathbb{P}(T_0 \leqslant n) &= 1 - \sum\limits_{k=1}^d \mathbb{P}(X_n = k) \ &= 1 - \sum\limits_{k=1}^d \sum\limits_{i=1}^d eta_i \mathbb{P}(X_n = k \mid X_0 = i) \ &= 1 - \sum\limits_{k=1}^d \sum\limits_{i=1}^d eta_i Q_{i,k}^n \ &= 1 - eta^ op Q^n e, \qquad n \geqslant 1. \end{aligned}$$

Alternatively we could also write

$$egin{array}{rcl} \mathbb{P}(T_0\leqslant n)&=&\mathbb{P}(X_n=0)\ &=&\sum\limits_{i=1}^deta_i\mathbb{P}(X_n=0\mid X_0=i) \end{array}$$

ර 543

$$egin{aligned} &=& \sum\limits_{i=1}^d eta_i [P^n]_{i,0} \ &=& \sum\limits_{i=1}^d eta_i [(I_d-Q^n)e]_i \ &=& \sum\limits_{i=1}^d eta_i - \sum\limits_{i=1}^d eta_i [Q^ne]_i \ &=& 1-eta^ op Q^n e, \qquad n \geqslant 1. \end{aligned}$$

i) We have

$$egin{aligned} G_{T_0}(s) &=& \sum\limits_{k \geqslant 0} s^k \mathbb{P}(T_0 = k) \ &=& \mathbb{P}(X_0 = 0) + \sum\limits_{k \geqslant 1} s^k eta^ op Q^{k-1} lpha \ &=& s \sum\limits_{k \geqslant 0} s^k eta^ op Q^k lpha \ &=& s eta^ op (I_d - sQ)^{-1} lpha \ &=& s eta^ op (I_d - sQ)^{-1} (I_d - Q) e. \end{aligned}$$

We note that

$$\mathbb{P}(T_0 < \infty) = G_{T_0}(1) = \beta^{ op}(I_d - Q)^{-1}(I_d - Q)e = \beta^{ op}e = 1,$$

hence state 0 is reached in finite time with probability one.

j) We have

$$G_{T_0}'(s)=eta^ op(I_d-sQ)^{-1}lpha+seta^ op Q(I_d-sQ)^{-2}lpha,$$

hence

$$egin{aligned} \mathbb{E}[T_0] &= G_{T_0}'(1) \ &= eta^ op (I_d - Q)^{-1} lpha + eta^ op Q (I_d - Q)^{-2} lpha \ &= eta^ op (I_d - Q) (I_d - Q)^{-2} lpha + eta^ op Q (I_d - Q)^{-2} lpha \ &= eta^ op (I_d - Q)^{-2} lpha \ &= eta^ op (I_d - Q)^{-2} lpha \ &= eta^ op (I_d - Q)^{-1} e. \end{aligned}$$

We also have

$$G_{T_0}''(s) = eta^ op Q (I_d - sQ)^{-2} lpha + eta^ op Q (I_d - sQ)^{-2} lpha + 2seta^ op Q^2 (I_d - sQ)^{-3} lpha,$$

hence

$$egin{array}{rll} \mathbb{E}[T_0(T_0-1)] &=& G_{T_0}''(1) \ &=& 2eta^ op Q(I_d-Q)^{-2}lpha+2eta^ op Q^2(I_d-Q)^{-3}lpha \ &=& 2eta^ op Q(I_d-Q)^{-3}lpha, \ &=& 2eta^ op Q(I_d-Q)^{-2}e, \end{array}$$

hence

$$egin{array}{rll} \mathbb{E}[T_0^2] &=& \mathbb{E}[T_0(T_0-1)] + \mathbb{E}[T_0] \ &=& 2eta^ op Q(I_d-Q)^{-2}e + eta^ op (I_d-Q)^{-1}e \ &=& 2eta^ op Q(I_d-Q)^{-2}e + eta^ op (I_d-Q)(I_d-Q)^{-2}e \ &=& eta^ op (I_d+Q)(I_d-Q)^{-2}e. \end{array}$$

More generally, by (1.7.4) we could also compute the *factorial moment*

$$\mathbb{E}[T_0(T_0-1)\cdots(T_0-k+1)] = G_{T_0}^{(k)}(1) = k! eta^ op Q^{k-1}(I_d-Q)^{-k} e,$$

for all $k \ge 1$.

Problem 4.14

a) We have

$$\begin{split} \mathbb{P}((O_0, O_1, O_2) &= (c, a, b) \text{ and } (X_0, X_1, X_2) = (1, 1, 0)) \\ &= \mathbb{P}((O_0, O_1, O_2) = (c, a, b) \mid (X_0, X_1, X_2) = (1, 1, 0)) \mathbb{P}((X_0, X_1, X_2) = (1, 1, 0)) \\ &= \mathbb{P}(O_0 = c \mid X_0 = 1) \mathbb{P}(O_1 = a \mid X_1 = 1) \mathbb{P}(O_2 = b \mid X_2 = 0) \mathbb{P}((X_0, X_1, X_2) = (1, 1, 0)) \\ &= \pi_1 P_{1,1} P_{1,0} m_{1,c} m_{1,a} m_{0,b}, \end{split}$$

since

$$\mathbb{P}((X_0, X_1, X_2) = (1, 1, 0)) = \pi_1 P_{1,1} P_{1,0}.$$
 MH3512 AY19-20

544

b) We have

$$\begin{split} \mathbb{P}((O_0,O_1,O_2) &= (c,a,b)) \\ &= \sum_{x,y,z \in \{0,1\}} \mathbb{P}((O_0,O_1,O_2) = (c,a,b) \text{ and } (X_0,X_1,X_2) = (x,y,z)) \\ &= \sum_{x,y,z \in \{0,1\}} \pi_x P_{x,y} P_{y,z} m_{x,c} m_{y,a} m_{z,b}. \end{split}$$

c) We have

$$egin{aligned} &\{X_1=1\} \;=\; ig\{(X_0,X_1,X_2)=(0,1,0)ig\}ig\cupig\{(X_0,X_1,X_2)=(0,1,1)ig\}\ &\cupig\{(X_0,X_1,X_2)=(1,1,0)ig\}ig\cupig\{(X_0,X_1,X_2)=(1,1,1)ig\}\ &=\; igcup_{x,z\in\{0,1\}}ig\{(X_0,X_1,X_2)=(x,1,z)ig\}, \end{aligned}$$

where the above union is a partition, hence

$$\begin{split} \mathbb{P}(X_1 &= 1 \mid (O_0, O_1, O_2) = (c, a, b)) \\ &= \sum_{x,z \in \{0,1\}} \mathbb{P}((X_0, X_1, X_2) = (x, 1, z) \mid (O_0, O_1, O_2) = (c, a, b)) \\ &= \frac{1}{\mathbb{P}((O_0, O_1, O_2) = (c, a, b))} \\ &\times \sum_{x,z \in \{0,1\}} \mathbb{P}((X_0, X_1, X_2) = (x, 1, z) \text{ and } (O_0, O_1, O_2) = (c, a, b)) \\ &= \frac{1}{\mathbb{P}((O_0, O_1, O_2) = (c, a, b))} \sum_{x,z \in \{0,1\}} \pi_x P_{x,1} P_{1,z} m_{x,c} m_{1,b} m_{z,c}. \end{split}$$

d) By the result of Question (a) we find

$$\begin{cases} \mathbb{P}((X_0, X_1, X_2) = (0, 0, 0) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.00588, \\ \mathbb{P}((X_0, X_1, X_2) = (0, 0, 1) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.00126, \\ \mathbb{P}((X_0, X_1, X_2) = (0, 1, 0) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.0101, \\ \mathbb{P}((X_0, X_1, X_2) = (0, 1, 1) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.00756, \\ \mathbb{P}((X_0, X_1, X_2) = (1, 0, 0) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.000448, \\ \mathbb{P}((X_0, X_1, X_2) = (1, 0, 1) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.000960, \\ \mathbb{P}((X_0, X_1, X_2) = (1, 1, 0) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.00269, \\ \mathbb{P}((X_0, X_1, X_2) = (1, 1, 1) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.00202. \end{cases}$$

From the above computation we deduce that the most likely sample sequence of values for (X_0, X_1, X_2) given the observation $(O_0, O_1, O_2) = (c, a, b)$ is $(X_0, X_1, X_2) = (0, 1, 0)$, with the probability

$$\mathbb{P}((X_0, X_1, X_2) = (0, 1, 0) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.0101.$$

e) By the result of Question (b) we find

$$\mathbb{P}((O_0,O_1,O_2)=(c,a,b))=0.030028\simeq 3\%.$$

f) By proceeding as in Question (c) we find

$$\begin{split} \mathbb{P}(X_0 &= 0 \mid (O_0, O_1, O_2) = (c, a, b)) = 0.825, \\ \mathbb{P}(X_0 &= 1 \mid (O_0, O_1, O_2) = (c, a, b)) = 0.175, \\ \mathbb{P}(X_1 &= 0 \mid (O_0, O_1, O_2) = (c, a, b)) = 0.256, \\ \mathbb{P}(X_1 &= 1 \mid (O_0, O_1, O_2) = (c, a, b)) = 0.744, \\ \mathbb{P}(X_2 &= 0 \mid (O_0, O_1, O_2) = (c, a, b)) = 0.636, \\ \mathbb{P}(X_2 &= 1 \mid (O_0, O_1, O_2) = (c, a, b)) = 0.364. \end{split}$$

According to the above estimates, the most likely sequence for (X_0, X_1, X_2) given the observation $(O_0, O_1, O_2) = (c, a, b)$ is (0, 1, 0).

g) We find
$$\hat{\pi} = [\hat{\pi}_0, \hat{\pi}_1] = [0.825, 0.175]$$

h) We find

$$\hat{P} = \left[egin{array}{ccc} 0.415 & 0.585 \ 0.482 & 0.518 \end{array}
ight]$$

i) We find

$$\hat{M} = \left[egin{array}{cccccc} 0.149 & 0.370 & 0.481 \ 0.580 & 0.284 & 0.136 \end{array}
ight].$$

j) The estimates of the matrix M obtained from the R code can be plotted as follows:





Figure S.7: Plot of estimate $\frac{1}{a,e,i,u,o}$ are more frequently associated to the state 0 of the hidden chain $(X_n)_{n \in \mathbb{N}}$. The vowels "a, e, i, o, u", together with the spacing character "_" total 93% of emission probabilities from state $\mathbb{Q}_{\mathcal{A}}$

and the combined probabilities of vowels from state (1) is only 6.2×10^{-9} %.

Human intervention can be nevertheless required in order to set a probability threshold that can distinguish vowels from consonants, e.g. to separate "u" from "t". The classification effect is enhanced in the following Figure S.8 that plots $\eta \mapsto (M_{0,\eta}/M_{0,n})(M_{1,n})(M_{1,n})(M_{1,n})(M_{1,n})(M_{1,n}))$ by combining the information available in the two rows of the emission matrix M, showing that "y" recovers its "semi-vowel" status.



Figure S.8: Plot of $\eta \mapsto (M_{0,\eta}/M_{0,"_"})((M_{1,"_"} - M_{1,\eta})/M_{1,"_"})^2$.

Note that the graphs of Figures S.6 and S.7 do *not* represent a frequency analysis. A frequency analysis of letters can be represented as the histogram of Figure S.9 using the command

text = readChar("brown.txt",nchars=10000)
data <- unlist (strsplit (gsub ("[^a-z]", "_", tolower (text)), ""))
barplot(col = rainbow(30), table(data), cex.names=0.7)</pre>

with the following output:

The command

estimate\$hmm\$transProbs[1,]

yields the transition probability estimate

$$\hat{P} = \left[egin{array}{ccc} 0 & 1 \ 0.1906356 & 0.8093644 \end{array}
ight]$$



Figure S.9: Frequency analysis of alphabet letters.

for the hidden chain $(X_n)_{n \in \mathbb{N}}$.

Note that \hat{P} is not the transition matrix of vowels vs consonants. For example of the word "universities" contains eleven letter transitions, including:

five vowel-consonant transitions,

one vowel-vowel transition,

four consonant-vowel transitions,

one consonant-consonant transitions,

which would yield the transition probability estimate

$$\left[egin{array}{cc} 5/6 & 1/6 \ 4/5 & 1/5 \end{array}
ight],$$

assuming the alphabet has *already* been partitioned. Such a matrix can be estimated on the whole text, from the code

```
x <- unlist (strsplit (gsub ("[^a-z]", "", tolower (text)), ""))
y <- unlist (strsplit (gsub ("[^a,e,i,o,u]", "2", tolower (x)), "")))
z <- as.numeric(noquote(unlist (strsplit (gsub ("[a,e,i,o,u]", "1",y), ""))))
p <- matrix(nrow = 2, ncol = 2, 0)
for (t in 1:(length(z) - 1)) p[z[t], z[t + 1]] <- p[z[t], z[t + 1]] + 1
for (i in 1:2) p[i, ] <- p[i, ] / sum(p[i, ])</pre>
```

 \mathbf{as}

$\left. \begin{array}{ccc} 0.1424749 & 0.8575251 \\ 0.5360502 & 0.4639498 \end{array} \right],$

which means that inside the text, a vowel is followed by a consonant for 85.7% of the time, while a consonant is followed by a vowel for 53% of the time.

The Baum-Welch algorithm does more than a simple frequency/transition analysis, as it can estimate the emission probability matrix M, which can be used to partition the alphabet. However, the algorithm is not making a one-to-one association between states (0, 1) of $(X_n)_{n \in \mathbb{N}}$ and letters; the association is only probabilistic and expressed through the estimate \hat{M} of the emission matrix.

Using a three-state model shows a more definite identification of vowels from state 3 in Figure S.10, and a special weight given to the letters h and t from state 1 in Figure S.11.



Figure S.10: Plot of estimate\$hmm\$emissionProbs[3,].

551



Figure S.11: Plot of estimate\$hmm\$emissionProbs[1,].

Chapter 5 - First Step Analysis

Exercise 5.1

a) The chain has the transition matrix

$$\left[\begin{array}{ccc} P_{i,j} \end{array}
ight]_{i,j\in\mathbb{N}} = \left[egin{array}{cccc} 1 & 0 & 0 & 0 \ 1/2 & 0 & 1/2 & 0 \ 0 & 1/3 & 0 & 2/3 \ 0 & 0 & 0 & 1 \end{array}
ight],$$

and the graph



b) Compute $\alpha := \mathbb{P}(T_3 < \infty \mid X_0 = 1)$ and $\beta := \mathbb{P}(T_3 < \infty \mid X_0 = 2)$, where

$$T_3:=\inf\{n \geqslant 0 \; : \; X_n=3\}.$$
 MH3512 AY19-20

<u>Answer</u>: By first step analysis, we have

$$\left\{egin{array}{l} lpha = \mathbb{P}(T_3 < \infty \mid X_0 = 1) = rac{1}{2} imes 0 + rac{1}{2}eta \ eta = \mathbb{P}(T_3 < \infty \mid X_0 = 2) = rac{1}{3}lpha + rac{2}{3} imes 1, \end{array}
ight.$$

hence

$$lpha = \mathbb{P}(T_3 < \infty \mid X_0 = 1) = rac{2}{5} \ ext{ and } \ eta = \mathbb{P}(T_3 < \infty \mid X_0 = 2) = rac{4}{5}.$$

c) By first step analysis we have

$$egin{aligned} & \left[\mathbbm{E}[T_{0,3} \mid X_0 = 1] \ = rac{1}{2} \left(1 + \mathbbm{E}[T_{0,3} \mid X_0 = 0]
ight) + rac{1}{2} \left(1 + \mathbbm{E}[T_{0,3} \mid X_0 = 2]
ight) \ & = 1 + rac{1}{2} \mathbbm{E}[T_{0,3} \mid X_0 = 2], \ & \mathbbm{E}[T_{0,3} \mid X_0 = 2] \ = 1 + rac{1}{3} \mathbbm{E}[T_{0,3} \mid X_0 = 1] + rac{2}{3} \left(1 + \mathbbm{E}[T_{0,3} \mid X_0 = 3]
ight) \ & = 1 + rac{1}{3} \mathbbm{E}[T_{0,3} \mid X_0 = 1], \ & \mathbbm{E}[T_{0,3} \mid X_0 = 1], \end{aligned}$$

i.e.

$$\mathbb{E}[T_{0,3} \mid X_0 = 1] = rac{3}{2} + rac{1}{6} \mathbb{E}[T_{0,3} \mid X_0 = 1]$$

 $\mathbb{E}[T_{0,3} \mid X_0 = 2] = rac{4}{3} + rac{1}{6} \mathbb{E}[T_{0,3} \mid X_0 = 2],$

hence

$$\mathbb{E}[T_{0,3} \mid X_0 = 1] = rac{9}{5},$$
 $\mathbb{E}[T_{0,3} \mid X_0 = 2] = rac{8}{5}.$

Exercise 5.2 By interpreting the mean duration between two visits to state (1) as the mean return time $\mu_1(1)$ from state (1) to tate (1) we find

$$\mu_1(1) = 1 + rac{b}{a} = 1 + rac{0.8}{0.6} = rac{7}{3}$$

from e.g. (5.3.6). This mean time can also be recovered by pathwise analysis using the mean 1/a of the geometric distribution on $\{1, 2, 3, \ldots\}$ with parameter $a \in (0, 1]$, as

$$(1-b) imes 1+b imes \left(1+rac{1}{a}
ight)=1+rac{b}{a}=rac{7}{3}$$

Exercise 5.3

a) The chain has the following graph:



Noting that state 0 is absorbing, by first step analysis we have

$$\left\{egin{array}{l} g_0(0) = 1 \ g_0(1) = rac{1}{3}g_0(0) + rac{2}{3}g_0(2) \ g_0(2) = g_0(1), \end{array}
ight.$$

which has for solution

$$g_0(0) = g_0(1) = g_0(2) = 1.$$
 MH3512 AY19-20

b) By first step analysis we have

$$\left\{egin{array}{l} h_0(0)=0\ h_0(1)=1+rac{1}{3}h_0(0)+rac{2}{3}h_0(2)\ h_0(2)=1+h_0(1), \end{array}
ight.$$

which has for solution

$$h_0(0)=0, \ \ h_0(1)=5, \ \ h_0(2)=6.$$

Exercise 5.4

a) The finite difference equation satisfied by f(x, y), $x + y \ge 1$, is given by

$$f(x,y) = \frac{x}{x+y}(f(x-1,y)-1) + \frac{y}{x+y}(f(x,y-1)+1).$$

b) The boundary conditions are given by

f(x,0)=-x and f(0,y)=0, $x,y \geqslant 1.$

$$\begin{cases} f(1,1) = \frac{1}{2}(f(0,1)-1) + \frac{1}{2}(f(1,0)+1) = 0, \\ f(1,2) = \frac{1}{3}(f(0,2)-1) + \frac{2}{3}(f(1,1)+1) = 1, \\ f(2,2) = \frac{1}{2}(f(1,2)-1) + \frac{1}{2}(f(2,1)+1) = 0, \\ f(1,3) = \frac{1}{4}(f(0,3)-1) + \frac{3}{4}(f(1,2)+1) = 2, \\ f(2,3) = \frac{2}{5}(f(1,3)-1) + \frac{3}{5}(f(2,2)+1) = 1, \\ f(3,3) = \frac{1}{2}(f(2,3)-1) + \frac{1}{2}(f(3,2)+1) = 0. \end{cases}$$

d) We check that f(x,y) := y - x solves the finite difference equation

$$\begin{aligned} &\frac{x}{x+y}(f(x-1,y)-1) + \frac{y}{x+y}(f(x,y-1)+1) \\ &= \frac{x}{x+y}(y-(x-1)-1) + \frac{y}{x+y}(y-1-x+1) \\ &= \frac{x}{x+y}(y-x) + \frac{y}{x+y}(y-x) \\ &= y-x \\ &= f(x,y), \end{aligned}$$

with the correct boundary conditions.

Exercise 5.5

c)

hence

and

$$\left\{ egin{array}{l} p_{0,2} = 0.2 imes p_{1,2} + 0.8 \ p_{1,2} = 0.1 imes p_{1,2} + 0.3 imes p_{0,2} + 0.6 imes p_{3,2} \ p_{2,2} = 0.5 imes p_{0,2} + 0.5 imes p_{3,2} \ p_{3,2} = 0, \end{array}
ight.$$

$$\left\{ egin{array}{l} p_{0,2} = 0.2 imes p_{1,2} + 0.8 \ p_{1,2} = rac{p_{0,2}}{3} \ p_{2,2} = rac{p_{0,2}}{2} \ p_{3,2} = 0, \end{array}
ight.$$

$$\left\{egin{array}{l} p_{0,2}=rac{6}{7} \ p_{1,2}=rac{2}{7} \ p_{2,2}=rac{3}{7} \ p_{3,2}=0. \end{array}
ight.$$

We note that

$$\mathbb{P}(T_2^r = \infty \mid X_0 = 2) = 1 - p_{2,2} = rac{4}{7} \geqslant \mathbb{P}(X_1 = 3 \mid X_0 = 2) = 0.5,$$

~

as seen in (6.3.5).

b) By first step analysis, we have

$$\left\{\begin{array}{l} p_{0,1}=0.2+0.8\times p_{2,1}\\ p_{1,1}=0.1+0.3\times p_{0,1}+0.6\times p_{3,1}\\ p_{2,1}=0.5\times p_{0,1}+0.5\times p_{3,1}\\ p_{3,1}=0, \end{array}\right.$$
MH3512 AY19-20

hence

558

$$\left\{egin{array}{l} p_{0,1}=0.2+0.8 imes p_{2,1}=0.2+0.4 imes p_{0,1}\ p_{1,1}=0.1+0.3 imes p_{0,1}\ p_{2,1}=rac{p_{0,1}}{2}\ p_{3,1}=0, \end{array}
ight. \left\{egin{array}{l} p_{0,1}=rac{1}{3}\ p_{1,1}=0.2\ p_{2,1}=rac{1}{6}\ p_{3,1}=0. \end{array}
ight.$$

and

We note that

$$\mathbb{P}(T_1^r = \infty \mid X_0 = 1) = 1 - p_{1,1} = 0.8 \geqslant \mathbb{P}(X_1 = 3 \mid X_0 = 1) = 0.6,$$

as seen in (6.3.6).

Exercise 5.6 This exercise is a particular case of the Example of Section 5.1, by taking a := 0.3, b := 0, c := 0.7, d := 0, $\alpha := 0$, $\beta := 0.3$, $\gamma := 0$, $\eta := 0.7$, and the chain has the graph



a) States 0 and 3 are absorbing.

b) By first step analysis we find the equations

$$\left\{egin{array}{l} g_0(0) = 1 \ g_0(1) = 0.3 + 0.7 g_0(2) \ g_0(2) = 0.3 g_0(1) \ g_0(3) = 0, \end{array}
ight.$$

hence

$$g_0(1) = rac{0.3}{1-0.7 imes 0.3} \quad ext{and} \quad g_0(2) = rac{0.3 imes 0.3}{1-0.7 imes 0.3},$$

which is consistent with the answer

$$\left\{egin{array}{l} g_0(1)=rac{a}{1-eta c} \ g_0(2)=rac{aeta}{1-eta c} \end{array}
ight.$$

obtained from (5.1.10). Note that this chain is actually a particular case of the gambling process of Chapter 2 with S = 3, hence by (2.2.11) we can also write

$$g_0(1) = \mathbb{P}(R_A \mid X_0 = 1) = rac{(3/7) - (3/7)^3}{1 - (3/7)^3} = rac{(7/3)^{3-1} - 1}{(7/3)^3 - 1} \simeq 0.3797$$

c) By first step analysis we find the equations

$$egin{aligned} & h_1(0) = 1 + h_1(0) \ & h_1(1) = 0 \ & h_1(2) = 0.3(1 + h_1(1)) + 0.7(1 + h_1(3)) = 1 + 0.7h_1(3) \ & h_1(3) = 1 + h_1(3), \end{aligned}$$

which admit the only solution $h_1(1) = 0$ and $h_1(0) = h_1(2) = h_1(2) = +\infty$.

Exercise 5.7 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 starting from state k = 0, 1, 2, 3. We have

$$\left\{egin{array}{l} h_3(0) = 1 + rac{1}{2}h_3(0) + rac{1}{2}h_3(2) \ h_3(1) = 1 + rac{1}{2}h_3(0) \ h_3(2) = 1 + rac{1}{2}h_3(0) + rac{1}{2}h_3(1) \ h_3(3) = 0, \end{array}
ight.$$

which yields

$$h_3(3)=0, \qquad h_3(1)=8, \qquad h_3(2)=12, \qquad h_3(0)=14.$$

We check that $h_3(3) < h_3(1) < h_3(2) < h_3(0)$, as can be expected from the graph.

Exercise 5.8 This exercise has some similarities with the gambling problem of Chapter 2, with the maze problems of Section 5.3, and with Exercise 5.12 below.

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 have

$$\left\{egin{array}{l} g_0(0) = 1 \ g_0(1) = rac{1}{2} + rac{1}{2}g_0(2) \ g_0(2) = rac{1}{2}g_0(1) \ g_0(3) = 0, \end{array}
ight.$$

which has for solution

$$g_0(0)=1, \ \ g_0(1)=rac{2}{3}, \ \ g_0(2)=rac{1}{3}, \ \ g_0(3)=0.$$

c) By first step analysis we have

$$\left\{egin{array}{l} h_{0,3}(0)=0\ h_{0,3}(1)=1+rac{1}{2}h_{0,3}(2)\ h_{0,3}(2)=1+rac{1}{2}h_{0,3}(1)\ h_{0,3}(3)=0, \end{array}
ight.$$

which has for solution

$$h_{0,3}(0)=0, \ \ h_{0,3}(1)=2, \ \ h_{0,3}(2)=2, \ \ h_{0,3}(3)=0.$$

Exercise 5.9 Using the law of total probability with respect to the partition $\Omega = \bigcup_{m \ge 0} \{T_A = m\}$, we have, provided that $i_0 \in A$, 562

$$\begin{split} \mathbb{P}(X_{T_A+n} = j \mid X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty) \\ &= \sum_{m \geqslant 0} \mathbb{P}(X_{T_A+n} = j \text{ and } T_A = m \mid X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty) \\ &= \sum_{m \geqslant 0} \frac{\mathbb{P}(X_{T_A+n} = j, T_A = m, X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty)}{\mathbb{P}(X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty) \\ &= \sum_{m \geqslant 0} \frac{\mathbb{P}(X_{n+m} = j, T_A = m, X_m = i_0, \dots, X_0 = i_m)}{\mathbb{P}(X_{T_A} = i_0, \dots, X_0 = i_m, X_m \in A, X_{m-1} \notin A, \dots, X_0 \notin A)} \\ &= \sum_{m \geqslant 0} \frac{\mathbb{P}(X_{n+m} = j, X_m = i_0, \dots, X_0 = i_m, X_m \in A, X_{m-1} \notin A, \dots, X_0 \notin A)}{\mathbb{P}(X_{T_A} = i_0, \dots, X_0 = i_m, X_m \in A, X_{m-1} \notin A, \dots, X_0 \notin A)} \\ &= \sum_{m \geqslant 0} \frac{\mathbb{P}(X_{n+m} = j, X_m = i_0, \dots, X_0 = i_m, X_m \in A, X_{m-1} \notin A, \dots, X_0 \notin A)}{\mathbb{P}(X_m = i_0, \dots, X_0 = i_m, X_m \in A, X_{m-1} \notin A, \dots, X_0 \notin A)} \\ &= \sum_{m \geqslant 0} \mathbb{P}(X_{n+m} = j \mid X_m = i_0, \dots, X_0 = i_m, X_m \in A, X_{m-1} \notin A, \dots, X_0 \notin A) \\ &= \sum_{m \geqslant 0} \mathbb{P}(X_{n+m} = j \mid X_m = i_0, \dots, X_0 = i_m, X_m \in A, X_{m-1} \notin A, \dots, X_0 \notin A) \\ &= \sum_{m \geqslant 0} \mathbb{P}(X_{n+m} = j \mid X_m = i_0, \dots, X_0 = i_m, X_m \in A, X_{m-1} \notin A, \dots, X_0 \notin A) \\ &= \sum_{m \geqslant 0} \mathbb{P}(X_{n+m} = j \mid X_m = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty) \\ &= \sum_{m \geqslant 0} \mathbb{P}(X_{n+m} = j \mid X_m = i_0 \text{ and } X_m \in A, X_{m-1} \notin A, \dots, X_0 \notin A) \\ &= \sum_{m \geqslant 0} \mathbb{P}(X_{n+m} = j \mid X_m = i_0 \text{ and } X_m \in A, X_{m-1} \notin A, \dots, X_0 \notin A) \\ &= \sum_{m \geqslant 0} \mathbb{P}(X_{n+m} = j \mid X_m = i_0) \\ &= \sum_{m \geqslant 0} \mathbb{P}(X_{n+m} = j \mid X_m = i_0) \\ &= \sum_{m \geqslant 0} \mathbb{P}(X_{n+m} = j \mid X_m = i_0) \\ &= \sum_{m \geqslant 0} \mathbb{P}(X_{n+m} = j \mid X_m = i_0) \\ &= \mathbb{P}(X_n = j_0 \dots, X_0 = i_m \text{ and } X_m \in A, X_{m-1} \notin A, \dots, X_0 \notin A) \\ &= \mathbb{P}(X_n = j \mid X_0 = i_0) \sum_{m \geqslant 0} \frac{\mathbb{P}(X_m = i_0, \dots, X_0 = i_m \text{ and } X_m \in A, X_{m-1} \notin A, \dots, X_0 \notin A) \\ \mathbb{P}(X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty) \end{aligned}$$

$$= \mathbb{P}(X_n = j \mid X_0 = i_0) \sum_{m \ge 0} \frac{\mathbb{P}(X_m = i_0, \dots, X_0 = i_m \text{ and } T_A = m)}{\mathbb{P}(X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty)}$$

= $\mathbb{P}(X_n = j \mid X_0 = i_0) \sum_{m \ge 0} \frac{\mathbb{P}(X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A = m)}{\mathbb{P}(X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty)}$
= $\mathbb{P}(X_n = j \mid X_0 = i_0),$

for all $n, m \ge 0, j \in \mathbb{S}$, and $(i_k)_{k \in \mathbb{N}} \subset \mathbb{S}$.

Exercise 5.10

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

$$f(k)=pf(k+1)+qf(k-1),\qquad k\geqslant 1,$$

which is (2.2.6), and has the general solution

$$f(k) = C_1 + C_2 r^k, \qquad k \in \mathbb{N}, \tag{B.32}$$

where $\mathbf{r} = \mathbf{q}/\mathbf{p}$, by (2.2.16), see also the command RSolve[f[k]=pf[k+1]+(1-p)f[k-1],f[k],k].

- i) In case $q \ge p$, f(k) would tend to (positive or negative) infinity if $C_2 \ne 0$, 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 k becomes 0 in the limit as k tends to infinity, *i.e.* we have

$$\lim_{k o\infty}f(k)=\lim_{k o\infty}\mathbb{P}(T_0<\infty\mid S_0=k)=0,\qquad k\in\mathbb{N},$$

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 $k \ge 0$.

Alternatively, denoting by T_k the hitting time of state (k) we could show the decomposition

$$egin{array}{rll} f(k+1) &=& \mathbb{P}(T_0 < \infty \mid S_0 = k+1) \ &=& \mathbb{P}(T_k < \infty \mid S_0 = k+1) imes \mathbb{P}(T_0 < \infty \mid S_0 = k) \ &=& \mathbb{P}(T_0 < \infty \mid S_0 = 1) imes \mathbb{P}(T_0 < \infty \mid S_0 = k) \end{array}$$

MH3512 AY19-20

563

$$= f(1) \times f(k), \quad k \in \mathbb{N},$$
 $\circlearrowright 564$

based on the Markov property, *i.e.* the probability of going down from k + 1 to 0 in a finite time should be the product of the probability to go down one step in finite time, multiplied by the probability of going down k steps in a finite time. This shows that

$$f(k) = f(0)(f(1))^k = (f(1))^k, \qquad k \geqslant 0,$$

hence we have $C_1 = 0$, $C_2 = 1$, and f(1) = r = q/p, which yields

$$f(k) = \left(rac{q}{p}
ight)^k, \qquad k \in \mathbb{N},$$

and recovers (2.2.13) and (3.4.16), cf. also (B.25) and (B.26).

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

$$egin{array}{rll} h(k) &=& p(1+h(k+1))+q(1+h(k-1)) \ &=& 1+ph(k+1)+qh(k-1), \qquad k\geqslant 1, \end{array}$$

which is (2.3.6) and has a general solution of the form

$$h(k) = C_1 + C_2 r^k + \frac{1}{q-p} k, \quad k \in \mathbb{N},$$
 (B.33)

by (2.3.9). Next, we note that by the Markov property we should have the decomposition

$$egin{array}{rcl} h(k+1) &=& \mathbb{E}[T_0 \mid S_0 = k+1] \ &=& \mathbb{E}[T_0 \mid S_0 = 1] + \mathbb{E}[T_0 \mid S_0 = k] \ &=& h(1) + h(k), \qquad k \in \mathbb{N}, \end{array}$$

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

$$h(k) = h(0) + kh(1) = kh(1),$$
 MH3512 AY19-20

hence by (B.33) we have $C_1 = C_2 = 0$, h(1) = 1/(q - p), and

$$h(k)=rac{k}{q-p}, \qquad k\in \mathbb{N}.$$

In case $q \leq p$ the above argument would yield a negative value for h(k), which is impossible, and h(k) has to be infinite for all $k \geq 1$.

Exercise 5.11 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:

$$13 = 3 + 4 + 1 + 5$$
, or $13 = 1 + 6 + 3 + 3$, or $13 = 1 + 1 + 2 + 1 + 3 + 1 + 4$, 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

$$X_n = \sum\limits_{k=1}^n \xi_k, \qquad n \geqslant 0,$$

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 $(X_k)_{k=0,1,\dots,n}$ 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 $(X_n)_{n \ge 0}$ is actually a random walk with independent increments ξ_1, ξ_2, \ldots .

MH3512 AY19-20

(5565)

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 $k \ge 14$, and the probabilities g_k can be easily computed for $k = 8, \ldots, 13$ as

$$\begin{array}{rcl} g_{13} &=& \mathbb{P}(X_{T_A}=13 \mid X_0=13)=1, \\ g_{12} &=& \mathbb{P}(X_{T_A}=13 \mid X_0=12)=0, \\ g_{11} &=& \mathbb{P}(X_{T_A}=13 \mid X_0=11)=0, \\ g_{10} &=& \mathbb{P}(X_{T_A}=13 \mid X_0=10)=\frac{1}{6}, \\ g_{9} &=& \mathbb{P}(X_{T_A}=13 \mid X_0=9)=\frac{1}{6}+\frac{1}{6}\times\frac{1}{6}=\frac{7}{36}, \end{array}$$

$$g_8 = \mathbb{P}(X_{T_A} = 13 \mid X_0 = 8) = \frac{1}{6} + 2 \times \frac{1}{6} \times \frac{1}{6} + \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \frac{49}{216},$$

where for the computation of g_8 we note that there are only 4 ways to reach state (13) from state (8) without hitting states (11) or (12), by the four combinations

13 = 8 + 5, 13 = 8 + 1 + 4, 13 = 8 + 2 + 3 13 = 8 + 1 + 1 + 3.

Clearly, things become easily complicated for $k \leq 7$, and we will rely on first step analysis to continue the calculation. We have

$$g(k)=rac{1}{6}\sum\limits_{i=1}^{6}g_{k+i},\qquad k\in\mathbb{N},$$

i.e.

$$g_{10} = \frac{1}{6}g_{13}$$

$$g_{9} = \frac{1}{6}(g_{10} + g_{13})$$

$$g_{8} = \frac{1}{6}(g_{9} + g_{10} + g_{13})$$

$$g_{7} = \frac{1}{6}(g_{8} + g_{9} + g_{10} + g_{13})$$

$$g_{6} = \frac{1}{6}(g_{7} + g_{8} + g_{9} + g_{10})$$

$$g_{5} = \frac{1}{6}(g_{6} + g_{7} + g_{8} + g_{9} + g_{10})$$

$$g_{4} = \frac{1}{6}(g_{5} + g_{6} + g_{7} + g_{8} + g_{9} + g_{10})$$

$$g_{3} = \frac{1}{6}(g_{4} + g_{5} + g_{6} + g_{7} + g_{8} + g_{9})$$

$$g_{2} = \frac{1}{6}(g_{3} + g_{4} + g_{5} + g_{6} + g_{7} + g_{8})$$

$$g_{1} = \frac{1}{6}(g_{2} + g_{3} + g_{4} + g_{5} + g_{6} + g_{7})$$

MH3512 AY19-20

567

$$g_0 \;=\; rac{1}{6}(g_1+g_2+g_3+g_4+g_5+g_6).$$

In order to solve this system of equations we rewrite it as

i.e.

$$g_0 = rac{7^{10} - 7^6 imes 6^4 - 4 imes 7^3 imes 6^6}{6^{11}} \simeq 0.181892636.$$

The total number of paths leading to 13 in the above fashion is 928. More generally, the numbers of such paths starting from k = 0, 1, ..., 10 can be computed by induction according to the following table:

♦ 568

k	0	1	2	3	4	5	6	7	8	9	10
number of paths	928	492	248	125	63	32	16	8	4	2	1

Table 14.1: Number of paths starting from k = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

In Figure S.12 we enumerate the corresponding possible paths.

Figure S.12: Graph of 928 paths leading from 0 to state 13.*

In addition we could compute the number of paths leading to 13 in k rounds with k = 3, 4, 5, 6, 7, 8, 9, 10, 11 as in the following table:

k	3	4	5	6	7	8	9	10	11
number of paths	18	88	191	246	209	120	45	10	1

Table 14.2: Number of paths starting taking k rounds, k = 3, 4, 5, 6, 7, 8, 9, 10, 11.

♦ 569

^{*}Animated figure (works in Acrobat Reader).

This would also allow us to recover the value of g_0 as

$$g_0 = \frac{18}{6^3} + \frac{88}{6^4} + \frac{191}{6^5} + \frac{246}{6^6} + \frac{209}{6^7} + \frac{120}{6^8} + \frac{45}{6^9} + \frac{10}{6^{10}} + \frac{1}{6^{11}} \\ = \frac{65990113}{362797056} \simeq 0.181892636,$$

see e.g. here.

Exercise 5.12

a) The transition matrix is given by

Γ×	Х	Х	Х	Х	×]
q	0	p	0	0	0
0	${m q}$	0	p	0	0
0	0	${m q}$	0	p	0
0	0	0	${m q}$	0	p
L ×	Х	Х	Х	Х	×

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

$$g(k) = q \times g(k-1) + p \times g(k+1), \qquad 1 \leq k \leq 4.$$
(B.34)

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.34) rewritten as

$$g(k)=rac{1}{2} imes g(k-1)+rac{1}{2} imes g(k+1), \qquad 1\leqslant k\leqslant 4.$$

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

$$g(k)=rac{k}{5}, \qquad k=0,1,\ldots,5.$$

Exercise 5.13

a) The transition matrix is given by

Γ	1	0	0	0	0	0	• • • -	
	1 1	0	0	0	0	0		
	1/0	1/0	0	0	0	0	•••	
	1/2	1/2	0	0	0	0	•••	
	1/3	1/3	1/3	0	0	0	• • •	
	1/4	1/4	1/4	1/4	0	0	• • •	•
	1/5	1/5	1/5	1/5	1/5	0	•••	
	•	•	•	•	•	•	•	
	:	:	:	:	:	:	••	

Note that an arbitrary choice has been made for the first line (*i.e.* state 0 is absorbing), however other choices would not change the answer to the question.

b) We have

$$h_0(m) = \sum\limits_{k=0}^{m-1} rac{1}{m} (1+h_0(k)) = 1 + rac{1}{m} \sum\limits_{k=0}^{m-1} h_0(k), \qquad m \geqslant 1,$$

and $h_0(0) = 0, h_0(1) = 1.$

c) We have

$$\begin{split} h_0(m) &= 1 + \frac{1}{m} \sum_{k=0}^{m-1} h_0(k) \\ &= 1 + \frac{1}{m} h_0(m-1) + \frac{m-1}{m} \sum_{k=0}^{m-2} \frac{h_0(k)}{m-1} \\ &= 1 + \frac{1}{m} h_0(m-1) - \frac{m-1}{m} + \frac{m-1}{m} \left(1 + \sum_{k=0}^{m-2} \frac{h_0(k)}{m-1} \right) \\ &= 1 + \frac{1}{m} h_0(m-1) - \frac{m-1}{m} + \frac{m-1}{m} h_0(m-1) \\ &= 1 - \frac{m-1}{m} + \frac{1}{m} h_0(m-1) + \frac{m-1}{m} h_0(m-1) \\ &= h_0(m-1) + \frac{1}{m}, \qquad m \ge 1, \end{split}$$

hence

$$egin{array}{rll} h_0(m) &=& h_0(m-1)+rac{1}{m} \ &=& h_0(m-2)+rac{1}{m-1}+rac{1}{m}=\sum\limits_{k=1}^mrac{1}{k}, \qquad m\geqslant 1. \end{array}$$

We can also note that from the Markov property we have

$$h_0(m) = h_{m-1}(m) + h_0(m-1),$$

where the mean time $h_{m-1}(m)$ from state m to state m-1 is equal to 1/m since state m-1 is always reached in one step from state m, with probability 1/m.

Exercise 5.14

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

$$P = \begin{bmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

٠

c) By first step analysis, the average time $h_3(i)$ from state (i) to state (3) solves

$$\begin{cases} h_3(0) = \frac{1}{3}(1+h_3(0)) + \frac{1}{3}(1+h_3(1)) + \frac{1}{3} \\ h_3(1) = 1 + h_3(2) \\ h_3(2) = 1 + h_3(0) \\ h_3(3) = 0, \end{cases}$$

which yields

$$h_3(0) = rac{1}{3}(1+h_3(0)) + rac{1}{3}(3+h_3(0)) + rac{1}{3},$$

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. The result $h_3(0) = 4$ can also be obtained directly by writing the equations

$$egin{aligned} & f_{3}(0) = rac{1}{3}(1+h_{3}(0)) + rac{1}{3}(1+h_{3}(1)) + rac{1}{3} imes 0 \ & h_{3}(1) = 1+h_{3}(2) \ & h_{3}(2) = 1+h_{3}(0). \end{aligned}$$

The difficulty in this exercise is that the Markov chain itself is not specified and one had to first design a suitable Markov model, sometimes involving a Markov chain with weighted links. On the other hand, the equation h = 1 + Ph cannot be directly used when the links have different weights starting from state (1).

Alternative solutions:

(i) The problem can be solved in a simpler way with only **3** states, by adding a weight corresponding to the travel time (underlined)

on each link:



Here the average time $h_2(i)$ from state (i) to state (2) solves

$$\begin{cases} h_2(0) = \frac{1}{3}(\underline{1} + h_2(0)) + \frac{1}{3}(\underline{1} + h_2(1)) + \frac{1}{3} \times \underline{0} \\ \\ h_2(1) = \underline{2} + h_2(0) \\ \\ h_2(2) = 0, \end{cases}$$

which yields

$$h_2(0) = rac{1}{3}(1+h_2(0)) + rac{1}{3}(3+h_2(0)) + rac{1}{3} imes 0,$$

i.e. $h_2(0) = 4$, which is consistent with the answer to Question (c) above since the time from the exit to the outside is here taken equal to 0.

(ii) The answer can even be further simplified using the following graph, which no longer uses the notion of Markov chain, and in which each link is weighted by a travel time:



By first step analysis the average time $t_1(0)$ to travel from state (0) to state (1) is directly given by

$$t_1(0) = rac{1}{3}(1+t_1(0)) + rac{1}{3}(3+t_1(0)) + rac{1}{3} imes 0,$$

i.e. $t_1(0) = 4$.

(iii) The problem could also be solved using 4 transient states and one absorbing state, based on the following graph:



Exercise 5.15 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 5.14 and its solution for a more detailed analysi of a similar problem.

Exercise 5.16 The chain has the following graph:



a) Let us compute

$$g_0(k) = \mathbb{P}(T_0 < \infty \mid X_0 = k) = \mathbb{P}(X_{T_{\{0,3\}}} = 0 \mid X_0 = k), \qquad k = 0, 1, 2, 3.$$

MH3512 AY19-20

575

Since states 0 and 3 are absorbing, by first step analysis we have

$$egin{aligned} g_0(0) &= 1 \ g_0(1) &= 0.1 imes g_0(0) + 0.6 imes g_0(1) + 0.1 imes g_0(2) + 0.2 imes g_0(3) \ g_0(2) &= 0.2 imes g_0(0) + 0.3 imes g_0(1) + 0.4 imes g_0(2) + 0.1 imes g_0(3) \ g_0(3) &= 0, \end{aligned}$$

i.e.

$$\left\{egin{aligned} g_0(0) &= 1 \ g_0(1) &= 0.1 + 0.6 imes g_0(1) + 0.1 imes g_0(2) \ g_0(2) &= 0.2 + 0.3 imes g_0(1) + 0.4 imes g_0(2) \ g_0(3) &= 0, \end{aligned}
ight.$$

which has for solution

$$g_0(0)=1, \hspace{0.3cm} g_0(1)=rac{8}{21}, \hspace{0.3cm} g_0(2)=rac{11}{21}, \hspace{0.3cm} g_0(3)=0,$$

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

$$\left\{egin{aligned} h_{0,3}(0) &= 0\ h_{0,3}(1) &= 0.1 imes 1 + 0.6 imes (1 + h_{0,3}(1)) + 0.1 imes (1 + h_{0,3}(2)) + 0.2 imes (1 + h_{0,3}(3))\ h_{0,3}(2) &= 0.2 imes 1 + 0.4 imes (1 + h_{0,3}(1)) + 0.3 imes (1 + h_{0,3}(2)) + 0.1 imes (1 + h_{0,3}(3))\ h_{0,3}(3) &= 0, \end{array}
ight.$$

576
$$egin{aligned} & h_{0,3}(0) = 0 \ & h_{0,3}(1) = 1 + 0.6 imes h_{0,3}(1) + 0.1 imes h_{0,3}(2) \ & h_{0,3}(2) = 1 + 0.4 imes h_{0,3}(1) + 0.3 imes h_{0,3}(2) \ & h_{0,3}(3) = 0, \end{aligned}$$

which has for solution

$$h_{0,3}(0)=0, \hspace{1em} h_{0,3}(1)=rac{10}{3}, \hspace{1em} h_{0,3}(2)=rac{10}{3}, \hspace{1em} h_{0,3}(3)=0.$$

Note that the relation $h_{0,3}(1) = h_{0,3}(2)$ can be guessed from the symmetry of the problem.

Exercise 5.17

a) Let $g_2(k) = \operatorname{I\!P}(T_2 < \infty \mid X_0 = k), k = 0, 1, 2$. We have $g_2(2) = 1$ and by first step analysis,

$$\left\{ egin{array}{l} g_2(0) = rac{1}{3}g_2(0) + rac{1}{3}g_2(1) + rac{1}{3} \ g_2(1) = rac{1}{4}g_2(0) + rac{3}{4}g_2(1), \end{array}
ight.$$

with $g_2(2) = 1$ since state (2) is absorbing. This shows that

$$\left\{ egin{array}{l} rac{2}{3}g_2(0) = rac{1}{3}g_2(1) + rac{1}{3} \ g_2(1) = g_2(0), \end{array}
ight.$$

hence $\mathbb{P}(T_2 < \infty \mid X_0 = 1) = g_2(0) = g_2(1) = g_2(2) = 1.$

Next, let $g_1(k) = \mathbb{P}(T_1^r < \infty \mid X_0 = k), k = 0, 1, 2$. We have $g_1(2) = 0$ and by first step analysis we have

$$egin{aligned} f g_1(0) &= rac{1}{3}g_1(0) + rac{1}{3} \ g_1(1) &= rac{1}{4}g_1(0) + rac{3}{4} \ g_1(2) &= 0, \end{aligned}$$

hence

$$\left\{egin{array}{l} g_1(0)=rac{1}{2}\ g_1(1)=rac{7}{8}\ g_1(2)=0, \end{array}
ight.$$

and $\operatorname{IP}(T_1^r < \infty \mid X_0 = 1) = g_1(1) = 7/8.$

b) We have $\mathbb{E}[T_1^r \mid X_0 = 1] = +\infty \times \mathbb{P}(T_1^r = \infty \mid X_0 = 1) = +\infty$ because $\mathbb{P}(T_1^r = \infty \mid X_0 = 1) > 0$ as state (1) is transient, cf. Proposition 7.4.

On the other hand, let $h_2(k) = \mathbb{E}[T_2 < \infty \mid X_0 = k], k = 0, 1, 2$. We have $h_2(2) = 0$ and by first step analysis we have

$$\begin{cases} h_2(0) = 1 + \frac{1}{3}h_2(0) + \frac{1}{3}h_2(1) \\\\ h_2(1) = 1 + \frac{1}{4}h_2(0) + \frac{3}{4}h_2(1) \\\\ h_2(2) = 0, \end{cases}$$
MH3512 AY19-20

(578)

hence

hence

and $\mathbb{E}[T_2^r < \infty \mid X_0 = 1] = h_2(1) = 11.$

Exercise 5.18

a) When $X_0 = x \geqslant 2$ and $Y_0 = y \geqslant 2$ we have $T_A = 0$, hence

$$\mu_A(x,y):=\mathbb{E}[T_A<\infty\mid X_0=x,\;Y_0=y]=0,\qquad x\geqslant 2,\;\;\;y\geqslant 2.$$

b) This equation is obtained by first step analysis, noting that we can only move up to to the right with probability 1/2 in both cases.

c) We note that $\mu_A(x,y) = \mu_A(x,y+1)$ for $y \geqslant 2$ and

$$\mu_A(1,y) = 1 + rac{1}{2} \mu_A(2,y) + rac{1}{2} \mu_A(1,y+1) = 1 + rac{1}{2} \mu_A(1,y), \hspace{1em} y \geqslant 2,$$

hence $\mu_A(1, y) = 2$ for all $y \ge 2$. We also have

$$\mu_A(0,y) = 1 + rac{1}{2} \mu_A(1,y) + rac{1}{2} \mu_A(0,y+1) = 2 + rac{1}{2} \mu_A(0,y), \hspace{1em} y \geqslant 2,$$

hence $\mu_A(0,y) = 4$, $y \ge 2$. By symmetry we also have $\mu_A(x,1) = 2$ and $\mu_A(x,0) = 4$ for all $x \ge 2$.

These results can also be recovered using pathwise analysis as

$$\mu_A(1,y) = \sum\limits_{k \geqslant 1} rac{k}{2^k} = rac{1}{2} \sum\limits_{k \geqslant 0} rac{k}{2^{k-1}} = rac{1}{2(1-1/2)^2} = 2, \hspace{1em} y \geqslant 2,$$

which yields similarly $\mu_A(x,1) = 2$ for all $x \ge 2$. Repeating this argument once also leads to $\mu_A(x,0) = \mu_A(0,y) = 4$ for all $x, y \ge 2$.

d) We have

$$\begin{split} \mu_A(1,1) &= 1 + \frac{1}{2}\mu_A(2,1) + \frac{1}{2}\mu_A(1,2) = 3, \\ \mu_A(0,1) &= 1 + \frac{1}{2}\mu_A(1,1) + \frac{1}{2}\mu_A(0,2) = \frac{9}{2}, \\ \mu_A(1,0) &= 1 + \frac{1}{2}\mu_A(2,0) + \frac{1}{2}\mu_A(1,1) = \frac{9}{2}, \\ \mu_A(0,0) &= 1 + \frac{1}{2}\mu_A(1,0) + \frac{1}{2}\mu_A(0,1) = \frac{11}{2}, \end{split}$$

hence the mean time it takes until both cans contain at least \$2 is $\mu_A(0,0) = 11/2$.

MH3512 AY19-20

580

4	4	2	0	0	0
3	4	2	0	0	0
2	4	2	0	0	0
1	9/2	3	2	2	2
0	11/2	9/2	4	4	4
	0	1	2	3	4

Table 14.3: Values of $\mu_A(x, y)$ with N = 2 and the set A in blue.

Exercise 5.19

- a) We have $\mu_A(x,y) = 0$ for all $(x,y) \in A$.
- b) We have

$$\mu_A(x,y) = 1 + \frac{1}{2}\mu_A(x+1,y) + \frac{1}{2}\mu_A(x,y+1), \qquad 0 \le x, y \le 3.$$
(B.35)

c) We have

$$\begin{split} \mu_A(2,2) &= 1 + \frac{1}{2}\mu_A(3,2) + \frac{1}{2}\mu_A(2,3) = 1, \\ \mu_A(1,2) &= 1 + \frac{1}{2}\mu_A(2,2) + \frac{1}{2}\mu_A(1,3) = \frac{3}{2}, \\ \mu_A(2,1) &= 1 + \frac{1}{2}\mu_A(2,2) + \frac{1}{2}\mu_A(3,1) = \frac{3}{2}, \\ \mu_A(0,2) &= 1 + \frac{1}{2}\mu_A(1,2) + \frac{1}{2}\mu_A(0,3) = \frac{7}{4}, \\ \mu_A(2,0) &= 1 + \frac{1}{2}\mu_A(2,1) + \frac{1}{2}\mu_A(3,0) = \frac{7}{4}, \\ \mu_A(1,1) &= 1 + \frac{1}{2}\mu_A(2,1) + \frac{1}{2}\mu_A(1,2) = \frac{5}{2}, \\ \mu_A(0,1) &= 1 + \frac{1}{2}\mu_A(1,1) + \frac{1}{2}\mu_A(0,2) = \frac{25}{8}, \\ \mu_A(1,0) &= 1 + \frac{1}{2}\mu_A(1,0) + \frac{1}{2}\mu_A(0,1) = \frac{33}{8}. \end{split}$$

d) The mean number of rounds is $\mu_A(0,0) = 33/8 = 4.125$.

^{*}Animated figure (works in Acrobat Reader).

4	0	0	0	0	0
3	0	0	0	0	0
2	7/4	3/2	1	0	0
1	25/8	5/2	3/2	0	0
0	33/8	25/8	7/4	0	0
	0	1	2	3	4

Table 14.4: Values of $\mu_A(x, y)$ with N = 3 and the set A in blue.

(583

Exercise 5.20 We have

$$egin{aligned} h(k) &= & \mathbb{E}\left[\sum\limits_{i \geqslant 0} eta^i c(X_i) \; \Big| \; X_0 = k
ight] \ &= & \mathbb{E}\left[c(X_0) \; \Big| \; X_0 = k
ight] + \mathbb{E}\left[\sum\limits_{i \geqslant 1} eta^i c(X_i) \; \Big| \; X_0 = k
ight] \ &= & c(k) + \sum\limits_{j \in S} P_{k,j} \, \mathbb{E}\left[\sum\limits_{i \geqslant 1} eta^i c(X_i) \; \Big| \; X_1 = j
ight] \ &= & c(k) + eta \sum\limits_{j \in S} P_{k,j} \, \mathbb{E}\left[\sum\limits_{i \geqslant 0} eta^i c(X_i) \; \Big| \; X_0 = j
ight] \ &= & c(k) + eta \sum\limits_{j \in S} P_{k,j} h(j), \qquad k \in S. \end{aligned}$$

However, this type of equation may be difficult to solve in general. We refer to Problem 5.24 for a particular case with explicit solution.

Exercise 5.21 We have

$$egin{aligned} V^*(k) &= \max_\pi \mathbb{E}\left[\sum\limits_{n \geqslant 0} \gamma^n R(X_n) \mid X_0 = k
ight] \ &= \max_\pi \mathbb{E}\left[R(X_0) + \sum\limits_{n \geqslant 1} \gamma^n R(X_n) \mid X_0 = k
ight] \ &= \max_\pi \mathbb{E}\left[R(k) + \gamma \sum\limits_{n \geqslant 1} \gamma^{n-1} R(X_n) \mid X_0 = k
ight] \ &= \max_\pi \left(R(k) + \mathbb{E}\left[\gamma \sum\limits_{n \geqslant 1} \gamma^{n-1} R(X_n) \mid X_0 = k
ight]
ight) \ &= R(k) + \gamma \max_\pi \mathbb{E}\left[\sum\limits_{n \geqslant 1} \gamma^{n-1} R(X_n) \mid X_0 = k
ight] \ &= R(k) + \gamma \max_{a \in \mathcal{A}} \sum\limits_{l \in \mathbb{S}} P^a_{k,l} \max_\pi \mathbb{E}\left[\sum\limits_{n \geqslant 1} \gamma^{n-1} R(X_n) \mid X_1 = k
ight] \end{aligned}$$

l

^{*}Animated figure (works in Acrobat Reader).

585

l

$$egin{aligned} &= R(k) + \gamma \max_{a \in \mathcal{A}} \sum\limits_{l \in \mathbb{S}} P^a_{k,l} \max_{\pi} \mathbb{E} \left[\sum\limits_{n \geqslant 0} \gamma^n R(X_{n+1}) \; \Big| \; X_1 = \ &= R(k) + \gamma \max_{a \in \mathcal{A}} \sum\limits_{l \in \mathbb{S}} P^a_{k,l} \mathbb{E} \left[\sum\limits_{n \geqslant 0} \gamma^n R(X_n) \; \Big| \; X_0 = l
ight] \ &= R(k) + \gamma \max_{a \in \mathcal{A}} \sum\limits_{l \in \mathbb{S}} P^a_{k,l} V^*(l), \qquad k \in \mathbb{S}. \end{aligned}$$

The optimal policy $\pi_k^* \in \mathbb{A}$ starting from state $k \in \mathbb{S}$ is given by

$$\pi_k^* = \operatorname{argmax}_{a \in \mathbb{A}} \sum\limits_{l \in \mathbb{S}} P_{k,l}^a V^*(l).$$

Problem 5.22

a) The boundary conditions g(0) and g(N) are given by g(0) = 1 and g(N) = 0.

b) We have

$$egin{aligned} g(k) &= & \mathbb{P}(T_0 < T_N \mid X_0 = k) = \sum\limits_{l=0}^N \mathbb{P}(T_0 < T_N \mid X_1 = l) \mathbb{P}(X_1 = l \mid X_0 = k) \ &= & \sum\limits_{l=0}^N \mathbb{P}(T_0 < T_N \mid X_1 = l) P_{k,l} \ &= & \sum\limits_{l=0}^N \mathbb{P}(T_0 < T_N \mid X_0 = l) P_{k,l} \ &= & \sum\limits_{l=0}^N g(l) P_{k,l}, \qquad k = 0, 1, \dots, N. \end{aligned}$$

c) We find

$$\left[\begin{array}{cccc} P_{i,j} \end{array}
ight]_{0\leqslant i,j\leqslant 3} = \left[egin{array}{cccc} 1 & 0 & 0 & 0 \ (2/3)^3 & (2/3)^2 & 2/3^2 & 1/3^3 \ 1/3^3 & 2/3^2 & (2/3)^2 & (2/3)^3 \ 0 & 0 & 0 & 1 \end{array}
ight]$$

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

$$\begin{split} \sum_{l=0}^{N} g(l) P_{k,l} &= \sum_{l=0}^{N} \frac{N!}{(N-l)!l!} \left(\frac{k}{N}\right)^{l} \left(1 - \frac{k}{N}\right)^{N-l} \frac{N-l}{N} \\ &= \sum_{l=0}^{N-1} \frac{N!}{(N-l)!l!} \left(\frac{k}{N}\right)^{l} \left(1 - \frac{k}{N}\right)^{N-l} \frac{N-l}{N} \\ &= \sum_{l=0}^{N-1} \frac{(N-1)!}{(N-l-1)!l!} \left(\frac{k}{N}\right)^{l} \left(1 - \frac{k}{N}\right)^{N-l} \\ &= \left(1 - \frac{k}{N}\right) \sum_{l=0}^{N-1} \binom{N-1}{l} \left(\frac{k}{N}\right)^{l} \left(1 - \frac{k}{N}\right)^{N-1-l} \\ &= \left(1 - \frac{k}{N}\right) \left(\frac{k}{N} + 1 - \frac{k}{N}\right)^{N-1} \\ &= \frac{N-k}{N} \\ &= g(k), \qquad k = 0, 1, \dots, N, \end{split}$$

which allows us to conclude by uniqueness of the solution given two boundary conditions, cf. Exercise 5.10 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 (1) and (N) are absorbing. AY19-20

$$h(k) = \mathbb{E}[T_{0,N} | X_0 = k]$$

$$= \sum_{l=0}^{N} (1 + \mathbb{E}[T_{0,N} | X_1 = l]) \mathbb{P}(X_1 = l | X_0 = k)$$

$$= \sum_{l=0}^{N} (1 + \mathbb{E}[T_{0,N} | X_1 = l]) P_{k,l} = \sum_{l=0}^{N} P_{k,l} + \sum_{l=0}^{N} \mathbb{E}[T_{0,N} | X_1 = l] P_{k,l}$$

$$= 1 + \sum_{l=1}^{N-1} \mathbb{E}[T_{0,N} | X_1 = l] P_{k,l} = 1 + \sum_{l=1}^{N-1} h(l) P_{k,l}, \qquad (B.36)$$

 $k=1,2,\ldots,N-1.$

g) In this case, the Equation (B.36) 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 5.23

a) As in the solution of Exercise 4.7, the transition matrix is given by

	aa	ab	ba	bb	
aa	p	q	0	0	
ab	0	0	$oldsymbol{p}$	\boldsymbol{q}	
ba	p	\boldsymbol{q}	0	0	.
bb	0	0	p	q	

b) We have $au_{ab} = 1$ with probability one, hence

$$G_{ab}(s)=\mathbb{E}[s\mid Z_1=(a,b)]=s.$$

c) We find

$$\left\{ egin{array}{l} G_{aa}(s) = psG_{aa}(s) + qsG_{ab}(s), \ G_{ba}(s) = psG_{aa}(s) + qsG_{ab}(s). \end{array}
ight.$$

d) We have

$$\left\{egin{array}{l} G_{aa}(s)=psG_{aa}(s)+qs^2,\ G_{ba}(s)=psG_{aa}(s)+qs^2, \end{array}
ight.$$

hence

$$G_{aa}(s) = G_{ba}(s) = rac{pqs^3}{1-ps} + qs^2 = rac{qs^2}{1-ps}, \qquad s \in (-1,1).$$

We note that

$$egin{aligned} \mathbb{P}(au_{ab} < \infty \mid Z_1 = (a,a)) &= \mathbb{P}(au_{ab} < \infty \mid Z_1 = (b,a)) \ &= G_{ba}(1^-) \ &= \lim_{s
earrow 1} G_{ba}(s) \ &= \lim_{s
earrow 1} rac{qs^2}{1-ps} \end{aligned}$$

589

$$= \frac{q}{1-p}$$
$$= 1.$$

e) We have

$$egin{array}{rll} \mathbb{E}[au_{ab} \mid Z_1 = (a,a)] &= \mathbb{E}[au_{ab} \mid Z_1 = (b,a)] \ &= G_{ba}'(1) = G_{aa}'(1) \ &= rac{2q}{1-p} + rac{pq}{(1-p)^2} = 2 + rac{p}{q}. \end{array}$$

f) This average time is

$$p \, \mathbb{E}[au_{ab} \mid Z_1 = (a,a)] + q \, \mathbb{E}[au_{ab} \mid Z_1 = (a,b)] = p \left(2 + rac{p}{q}
ight) + q = 1 + rac{p}{q}.$$

Problem 5.24

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

$$h(k) = \mathbb{E}\left[\sum_{i=0}^{\tau-1} X_{i} \mid X_{0} = k\right]$$

$$= \mathbb{E}\left[X_{0} \mid X_{0} = k\right] + \mathbb{E}\left[\sum_{i=1}^{\tau-1} X_{i} \mid X_{0} = k\right]$$

$$= k + p \mathbb{E}\left[\sum_{i=1}^{\tau-1} X_{i} \mid X_{1} = k + 1\right] + q \mathbb{E}\left[\sum_{i=1}^{\tau-1} X_{i} \mid X_{1} = k - 1\right]$$

$$= k + p \mathbb{E}\left[\sum_{i=0}^{\tau-2} X_{i+1} \mid X_{1} = k + 1\right] + q \mathbb{E}\left[\sum_{i=0}^{\tau-2} X_{i+1} \mid X_{1} = k - 1\right]$$

$$= k + p \mathbb{E}\left[\sum_{i=0}^{\tau-1} X_{i} \mid X_{0} = k + 1\right] + q \mathbb{E}\left[\sum_{i=0}^{\tau-1} X_{i} \mid X_{0} = k - 1\right]$$

$$= k + ph(k + 1) + qh(k - 1), \quad 1 \le k \le N - 1,$$
(B.37)

where we used the fact that $\tau - 1$ in (B.37) becomes τ in (B.38).

When passing from

to

$$p \mathbb{E} \left[\sum_{i=1}^{ au-1} X_i \ \Big| \ X_1 = k+1
ight] + q \mathbb{E} \left[\sum_{i=1}^{ au-1} X_i \ \Big| \ X_1 = k-1
ight]$$
 $p \mathbb{E} \left[\sum_{i=0}^{ au-1} X_i \ \Big| \ X_0 = k+1
ight] + q \mathbb{E} \left[\sum_{i=0}^{ au-1} X_i \ \Big| \ X_0 = k-1
ight],$

we are shifting the time summation index from i to i - 1, and after that the summation starts from 0 instead of 1. Nevertheless the summation remains up to $\tau - 1$ at the second step because in both cases, τ denotes the first hitting time of 0 or N. Writing $\tau - 2$ in the second step would be wrong because the first hitting time of 0 or N within the expectation is τ (not $\tau - 1$).

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.

Note that when trying successively h(k) = C, h(k) = Ck, $h(k) = Ck^2$ and $h(k) = Ck^3$, the quantity C has to be a constant.

Here, the index k is a variable, not a constant. Finding for example

$$C = 3k$$

can only be wrong because 3k is not a constant. On the other hand, finding C = 2 (say) would make sense because 2 is a constant. Note that N also has the status of a constant (parameter) here.

d) The general solution has the form

$$h(k) = -rac{k^3}{3} + C_1 + C_2 k,$$
 MH3512 AY19-20

and the boundary conditions show that

$$\left\{egin{array}{l} 0=h(0)=C_1,\ 0=h(N)=-rac{N^3}{3}+C_1+C_2N, \end{array}
ight.$$

hence $C_1 = 0, C_2 = N^2/3$, and

$$h(k) = -\frac{k^3}{3} + N^2 \frac{k}{3} = \frac{k}{3} (N^2 - k^2) = k(N - k) \frac{N + k}{3}, \qquad k = 0, 1, \dots, N.$$
(B.39)

- 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.39) we find

$$h(4) = \frac{4}{3}(70^2 - 4^2) = $6512$$
K = \$6.512M.

iii) In that case we find

$4K \times 264 = 1056K = 1.056M.$

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.

However, this problem has been solved in the fair game case and without time discount, which may not be realistic. A more thorough analysis should consider

$$h(k) = \mathbb{E} \left[\sum\limits_{i=0}^{ au-1} eta^i X_i \ \Big| \ X_0 = k
ight]$$

with discount factor $\beta \in (0, 1)$ in the case where p < q, with additional computational difficulties.

MH3512 AY19-20

591

Note that the probability of hitting 0 first (*i.e.* ending with a bankruptcy) is given by

$$\mathbb{P}(X_{T_{0,N}}=0 \mid X_0=k) = 1-rac{k}{N} = \%94$$

when k = 4 and N = 70.

Problem 5.25 This problem is based on a simplified version of questions considered in [Gus14]. a) We find



The smallest integers are l = 3, m = 2 and the word is a^3b^2 .

- b) This language can be denoted by $\Sigma^* a^3 b^2 \Sigma^*$. An example of five-letter word is *aabbb*.
- c) The process $(Y_k)_{k \in \mathbb{N}}$ is clearly a Markov chain because given Y_k , the distribution of $Y_{k+1} := f(X_{k+1}, Y_k)$ is independent of Y_0, \ldots, Y_{k-1} . The graph of the chain $(Y_k)_{k \in \mathbb{N}}$ is



The transition matrix of the chain $(Y_k)_{k\in\mathbb{N}}$ is

$$\left[\begin{array}{c} P_{i,j} \end{array}
ight]_{0\leqslant i,j\leqslant 5} = \left[egin{array}{cccc} q & p & 0 & 0 & 0 & 0 \ q & 0 & p & 0 & 0 & 0 \ q & 0 & 0 & p & q & 0 \ 0 & 0 & 0 & p & q & 0 \ 0 & p & 0 & 0 & 0 & q \ 0 & 0 & 0 & 0 & 0 & 1 \end{array}
ight].$$

MH3512 AY19-20

○ 592

d) Denoting by $h_5(k)$ the time it takes to reach state (5) starting from state k = 0, 1, 2, 3, 4, 5, we have the equations

$$egin{aligned} h_5(0) &= 1 + qh_5(0) + ph_5(1) \ h_5(1) &= 1 + qh_5(0) + ph_5(2) \ h_5(2) &= 1 + qh_5(0) + ph_5(3) \ h_5(3) &= 1 + ph_5(3) + qh_5(4) \ h_5(4) &= 1 + ph_5(1) + qh_5(5) \ h_5(5) &= 0, \end{aligned}$$

i.e.

i.e.

$$egin{aligned} ph_5(0) &= 1 + ph_5(1) \ h_5(1) &= 1 + qh_5(0) + ph_5(2) \ h_5(2) &= 1 + qh_5(0) + ph_5(3) \ qh_5(3) &= 1 + qh_5(4) = 1 + qph_5(0) \ h_5(4) &= 1 + ph_5(1) = ph_5(0) \ h_5(5) &= 0, \end{aligned}$$

$$\begin{cases} h_5(0) = \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + h_5(3) \\ h_5(1) = \frac{1}{p^2} + \frac{1}{p^3} + h_5(3) \\ h_5(2) = \frac{1}{p^3} + h_5(3) \\ h_5(3) = \frac{1}{q} + ph_5(0) \\ h_5(4) = ph_5(0) \\ h_5(5) = 0, \end{cases}$$

(593

594

$$\begin{cases} h_5(0) = \frac{q(p^2 + p + 1) + p^3}{p^3 q^2} = \frac{1}{p^3 q^2} \\ h_5(1) = \frac{1}{p^3 q^2} - \frac{1}{p} \\ h_5(2) = \frac{1}{p^3 q^2} - \frac{1}{p} - \frac{1}{p^2} \\ h_5(3) = \frac{1}{q} + \frac{1}{p^2 q^2} \\ h_5(4) = \frac{1}{p^2 q^2} \\ h_5(5) = 0. \end{cases}$$



i.e.



- f) The words are abab and aabb.
- g) The graph of the chain $(Z_k)_{k\in\mathbb{N}}$ is



The transition matrix of the chain $(Z_k)_{k\in\mathbb{N}}$ is

$$\left[\begin{array}{cccc} P_{i,j} \end{array}
ight]_{0\leqslant i,j\leqslant 4} = \left[egin{array}{cccc} 1 & 0 & 0 & 0 & 0 \ q & 0 & p & 0 & 0 \ 0 & 0 & p & q & 0 \ 0 & p & 0 & 0 & q \ 0 & 0 & 0 & 0 & 1 \end{array}
ight].$$

h) Denoting by $g_0(k)$ the probability that state (1) is reached first starting from state k = 0, 1, 2, 3, 4, we have the equations 595

$$egin{aligned} g_0(0) &= 1 \ g_0(1) &= qg_0(0) + pg_0(2) = q + pg_0(2) \ g_0(2) &= pg_0(2) + qg_0(3) \ g_0(3) &= pg_0(1) + qg_0(4) = pg_0(1) \ g_0(4) &= 0, \end{aligned}$$

i.e.

$$egin{aligned} g_0(0) &= 1 \ g_0(1) &= q + pg_0(2) \ g_0(2) &= pg_0(2) + qg_0(3) = pg_0(2) + qpg_0(1) \ g_0(3) &= pg_0(1) \ g_0(4) &= 0, \end{aligned}$$

i.e.

$$\left\{egin{array}{l} g_0(0) = 1 \ g_0(1) = q + p^2 g_0(1) \ g_0(2) = p g_0(1) \ g_0(3) = p g_0(1) \ g_0(4) = 0, \end{array}
ight.$$

i.e.

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = \frac{q}{1-p^2} = \frac{1}{1+p} \\ g_0(2) = \frac{pq}{1-p^2} = \frac{p}{1+p} \\ g_0(3) = \frac{pq}{1-p^2} = \frac{p}{1+p} \\ g_0(4) = 0. \end{cases}$$

The probability $g_0(1)$ can also be computed by pathwise analysis and a geometric series, as

$$g_0(1) = 1 - pq \sum_{k \geqslant 0} p^{2k} = 1 - rac{pq}{1 - p^2} = 1 - rac{p}{1 + p} = rac{1}{1 + p}$$

Now, starting from state (1) one may move directly to state (0) with probability q, in which case the first synchronized word is "b". MH3512 AY19-20

not "abab". For this reason we need to subtract q from $g_0(1)$, and the probability that the first synchronized word is "abab" starting from state (1) is $1/(1+p) - (1-p) = p^2/(1+p)$.

Note that the above computations apply only when $p \in [0, 1)$. In case p = 1 the problem admits a trivial solution since the word "abab" will never occur.

Exercise 5.26

- a) The unique shortest word that synchronizes to state (4) starting from all states 1, 2, 3 is "*aabb*".
- b) The process $(Z_k)_{k \in \mathbb{N}}$ is a Markov chain on the state space $\{0, 1, 2, 3, 4\}$, with the following graph:



The transition matrix of the chain $(Z_k)_{k\in\mathbb{N}}$ is

$$\left[\begin{array}{cccc} P_{i,j} \end{array}
ight]_{0\leqslant i,j\leqslant 4} = egin{bmatrix} 1 & 0 & 0 & 0 & 0 \ 1/2 & 0 & 1/2 & 0 & 0 \ 0 & 0 & 1/2 & 1/2 & 0 \ 0 & 1/2 & 0 & 0 & 1/2 \ 0 & 0 & 0 & 0 & 1 \ \end{bmatrix}$$

٠

c) Denoting by $g_4(k)$ the probability that state (4) is reached first starting from state k = 0, 1, 2, 3, 4, we have the equations 597

$$egin{aligned} g_4(0) &= 0 \ g_4(1) &= rac{1}{2}g_4(0) + rac{1}{2}g_4(2) &= rac{1}{2}g_4(2) \ g_4(2) &= rac{1}{2}g_4(2) + rac{1}{2}g_4(3) \ g_4(3) &= rac{1}{2}g_4(1) + rac{1}{2}g_4(4) &= rac{1}{2}g_4(1) + rac{1}{2}g_4(4) = 1, \end{aligned}$$

with the solution

$$\left\{egin{array}{l} g_4(0)=0\ g_4(1)=rac{1}{3}\ g_4(2)=rac{2}{3}\ g_4(3)=rac{2}{3}\ g_4(4)=1. \end{array}
ight.$$

Hence the probability that the first synchronized word is "aabb" when the automaton is started from state 1 is 1/3.

Problem 5.27

a) We have

$$V(k) \;=\; \mathbb{E}\left[\sum\limits_{n \geqslant 0} R(X_n) \;\Big|\; X_0 = k
ight]$$
MH3512 AY19-2

$$egin{aligned} &= \sum \limits_{m \in \mathbb{S}} P_{k,m} \left(R(k) + \mathbb{E} \left[\sum \limits_{n \geqslant 1} R(X_n) \; \Big| \; X_1 = k
ight]
ight) \ &= \sum \limits_{m \in \mathbb{S}} P_{k,m} R(k) + \sum \limits_{m \in \mathbb{S}} P_{k,m} \mathbb{E} \left[\sum \limits_{n \geqslant 1} R(X_n) \; \Big| \; X_1 = m
ight] \ &= R(k) \sum \limits_{m \in \mathbb{S}} P_{k,m} + \sum \limits_{m \in \mathbb{S}} P_{k,m} \mathbb{E} \left[\sum \limits_{n \geqslant 0} R(X_n) \; \Big| \; X_0 = m
ight] \ &= R(k) + \sum \limits_{m \in \mathbb{S}} P_{k,m} V(m), \qquad k \in \mathbb{S}. \end{aligned}$$

b) b1) We have

$$Q^*(7,\downarrow) = 0, \quad Q^*(7,\rightarrow) = 0, \quad Q^*(6,\downarrow) = 5, \quad Q^*(6,\rightarrow) = 5,$$

and $Q^*(3,\downarrow) = 4$. Regarding $Q^*(3,\rightarrow)$, we have

$$Q^*(3,
ightarrow)=-1+\maxig(Q^*(3,\downarrow),Q^*(3,
ightarrow)ig),$$

which implies

$$Q^*(3,
ightarrow) < Q^*(3,\downarrow),$$

hence

$$Q^*(3,
ightarrow)=-1+Q^*(3,\downarrow)=3.$$

Similarly, we find

$$\left\{ egin{array}{ll} Q^*(5,\downarrow)=-1, & Q^*(5,
ightarrow)=2, \ Q^*(2,\downarrow)=0, & Q^*(2,
ightarrow)=2, \ Q^*(4,\downarrow)=0, & Q^*(4,
ightarrow)=1, \ Q^*(1,\downarrow)=1, & Q^*(1,
ightarrow)=2. \end{array}
ight.$$

We can solve by backward optimization (or dynamic programming):

598



b2) At every state (\mathbf{k}) we have

 $V^*(k) = \maxig(Q^*(k,\downarrow),Q^*(k,
ightarrow)ig),$

hence

$$\left\{egin{array}{l} V^*(7) = 0, \ V^*(6) = 5, \ V^*(3) = 4, \ V^*(5) = 2, \ V^*(2) = 2, \ V^*(4) = 1, \ V^*(1) = 2. \end{array}
ight.$$

We obtain the following backward optimization tree:



b3) We find

$$\pi^* = (\pi^*(1), \pi^*(2), \pi^*(3), \pi^*(4), \pi^*(5), \pi^*(6), \pi^*(7)) = (
ightarrow,
ightarrow,
ightarrow,
ightarrow,
ightarrow),$$

which is consistent with the MDPtoolbox output:

```
install.packages("MDPtoolbox")
library(MDPtoolbox)
P <- array(0, c(7, 7, 2))
0,0,0,0,1,0,0,
                  0,0,0,0,0,1,0,
                  0,0,0,1,0,0,0,
                  0,0,0,0,1,0,0,
                  0,0,0,0,0,0,1,
                  0,0,0,0,0,0,1), nrow=7, ncol=7, byrow=TRUE)
P[,,2] <- matrix(c(0,1,0,0,0,0,0,
                  0,0,1,0,0,0,0,
                  0,0,1,0,0,0,0,
                  0,0,0,0,1,0,0,
                  0,0,0,0,0,1,0,
                  0,0,0,0,0,0,1,
                  0,0,0,0,0,0,1), nrow=7, ncol=7, byrow=TRUE)
R \le array(0, c(7, 2))
R[,1] <- matrix(c(0, -2, -1, -1, -3, 5, 0), nrow=1, ncol=7, byrow=TRUE)
R[,2] <- R[,1]
mdp_check(P, R)
mdp value iteration(P,R,discount=1)
$V
[1] 2 2 4 1 2 5 0
$policy
[1] 2 2 1 2 2 1 1
```

c) c1) We have

$$egin{aligned} Q^*(k,a) &:= \max_{\pi\,:\,\pi(k)=a} \mathbb{E}\left[\sum\limits_{n\geqslant 0} R(X_n) \ \Big| \ X_0 = k
ight] \ &= \max_{\pi\,:\,\pi(k)=a} \mathbb{E}\left[R(X_0) + \sum\limits_{n\geqslant 1} R(X_n) \ \Big| \ X_0 = k
ight] \ &= \max_{\pi\,:\,\pi(k)=a} \mathbb{E}\left[R(k) + \sum\limits_{n\geqslant 1} R(X_n) \ \Big| \ X_0 = k
ight] \end{aligned}$$

MH3512 AY19-20

601

$$egin{aligned} &= \sum \limits_{l \in \mathbb{S}} P_{k,l}^{(a)} \left(R(k) + \max_{\pi} \mathbb{E} \left[\sum \limits_{n \geqslant 1} R(X_n) \; \Big| \; X_1 = l
ight]
ight) \ &= \; R(k) + \sum \limits_{l \in \mathbb{S}} P_{k,l}^{(a)} \max_{\pi} \mathbb{E} \left[\sum \limits_{n \geqslant 0} R(X_{n+1}) \; \Big| \; X_1 = l
ight] \ &= \; R(k) + \sum \limits_{l \in \mathbb{S}} P_{k,l}^{(a)} \max_{\pi} \mathbb{E} \left[\sum \limits_{n \geqslant 0} R(X_n) \; \Big| \; X_0 = l
ight] \ &= \; R(k) + \sum \limits_{l \in \mathbb{S}} P_{k,l}^{(a)} V^*(l), \qquad k \in \mathbb{S}. \end{aligned}$$

Similarly, the value function $Q^{\pi}(k,a)$ for a given policy π by setting the first action at state (k) to a satisfies the equation

$$Q^{\pi}(k,a) = R(k) + \sum_{l \in S} P_{k,l}^{(a)} V^{\pi}(l), \qquad k \in S.$$
(B.40)

c2) The optimal policy $\pi_k^* \in \mathbb{A}$ at state $k \in \mathbb{S}$ is given by

$$\pi_k^* = \mathrm{argmax}_{a \in \mathbb{A}} Q^*(k,a) = \mathrm{argmax}_{a \in \mathbb{A}} \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} V^*(l).$$

c3) We have

$$egin{aligned} V^*(k) &=& \max_\pi \mathbb{E}\left[\sum\limits_{n \geqslant 0} R(X_n) \; \Big| \; X_0 = k
ight] \ &=& \max_{a \in \mathcal{A}} \max_{\pi \; : \; \pi(k) = a} \mathbb{E}\left[\sum\limits_{n \geqslant 0} R(X_n) \; \Big| \; X_0 = k
ight] \ &=& \max_{a \in \mathcal{A}} Q^*(k,a) \ &=& \max_{a \in \mathcal{A}} \left(R(k) + \sum\limits_{l \in \mathbb{S}} P_{k,l}^{(a)} V^*(l)
ight) \ &=& R(k) + \max_{a \in \mathcal{A}} \sum\limits_{l \in \mathbb{S}} P_{k,l}^{(a)} V^*(l), \quad k \in \mathbb{S}. \end{aligned}$$

MH3512 AY19-20

ॳ 602

This also implies

$$V^{*}(k) = R(k) + \max_{a \in \mathcal{A}} \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} V^{*}(l)$$

=
$$\max_{a \in \mathcal{A}} \left(R(k) + \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} V^{*}(l) \right)$$

=
$$\max_{a \in \mathcal{A}} Q^{*}(k, a), \quad k \in \mathbb{S}.$$
 (B.41)

Similarly, the value function $V^{\pi}(k)$ for a given policy satisfies the equation

$$V^{\pi}(k) = R(k) + \sum_{l \in \mathbb{S}} P_{k,l}^{(\pi(k))} V^{\pi}(l) \qquad k \in \mathbb{S},$$
(B.42)

and the Bellman equation for $Q^*(k,a)$ can also be rewritten as

$$Q^*(k,a) = R(k) + \sum\limits_{l\in \mathbb{S}} P^{(a)}_{k,l} \max\limits_{b\in \mathcal{A}} Q^*(k,b), \qquad k\in \mathbb{S}.$$

In order to solve the MDP problem, starting from an arbitrary initial policy choice π , one can:

- i) Compute V^{π} by applying (B.42) iteratively.
- ii) Deduce the value of $Q^{\pi}(k, a)$ for every state k and action a from (B.40).

iii) For every state (k) and action a, based on (B.41), choose to update π with $\pi(k) := a$ if $Q^{\pi}(k, a) > V^{\pi}(k)$.

iv) Repeat the above iteratively.

d) d1) Similarly to Question (b1), we have

We also have $Q^*(2,\downarrow) = 0$ and

$$egin{aligned} Q^*(2, o) &= -2 + p \max \left(Q^*(3, \downarrow), Q^*(3, o)
ight) + q \max \left(Q^*(5, o), Q^*(5, o)
ight) \ &= -2 + p Q^*(3, \downarrow) + q Q^*(5, o) \ &= -2 + 4p + 2q = 2p, \end{aligned}$$

and

$$Q^*(4,\downarrow)=0, \quad Q^*(4,
ightarrow)=1, \quad Q^*(1,\downarrow)=1, \quad Q^*(1,
ightarrow)=Q^*(2,
ightarrow)=2p.$$

In other words, we have the following backward optimization (or dynamic programming) graph:



$$V^*(k) = \maxig(Q^*(k, \downarrow), Q^*(k,
ightarrowig),$$

MH3512 AY19-20

ॳ 604

hence



$$\left\{egin{array}{l} V^*(7) = 0, \ V^*(6) = 5, \ V^*(5) = 2, \ V^*(4) = 1, \ V^*(3) = 4, \ V^*(2) = 2p, \ V^*(1) = \max(2p,1). \end{array}
ight.$$

. .

d3) When p = 0 we find

$$\pi^*=(\pi^*(1),\pi^*(2),\pi^*(3),\pi^*(4),\pi^*(5),\pi^*(6),\pi^*(7))=(\downarrow,\downarrow,\downarrow,
ightarrow,\downarrow,
ightarrow,\downarrow)$$

which is consistent with the MDPtoolbox output:

\$V							
[1]	1	0	4	1	2	5	0
\$po	oli	icy	7				
[1]	1	1	1	2	2	1	1



Figure S.15: Optimal value function with p = 0.

When 0 we obtain

$$\pi^*=(\pi^*(1),\pi^*(2),\pi^*(3),\pi^*(4),\pi^*(5),\pi^*(6),\pi^*(7))=(\downarrow,
ightarrow,\downarrow,
ightarrow,\downarrow,
ightarrow,\downarrow)$$

which is consistent with the MDP toolbox output, here with p = 0.25:

\$V [1] 1.0 0.5 4.0 1.0 2.0 5.0 0.0 \$policy [1] 1 2 1 2 2 1 1 **♦** 606



Figure S.16: Optimal value function with 0 .

When p = 1/2 we find

MH3512 AY19-20

ॳ 607



Figure S.17: Optimal value function with p = 1/2.

which is consistent with the MDPtoolbox output, with p = 0.5:

\$V	
[1] 1 1 4 1 2 5 0	
\$policy	

When 1/2 we obtain

 $\pi^*=(\pi^*(1),\pi^*(1),\pi^*_3,\pi^*_4,\pi^*_5,\pi^*_6,\pi^*_7)=(
ightarrow,
ightarrow,
ightarrow,
ightarrow,
ightarrow).$

MH3512 AY19-20

♦ 608



Figure S.18: Optimal value function with 1/2 .

which is also consistent with the following MDP toolbox output, here with p = 0.75:

```
library(MDPtoolbox)
P <- array(0, c(7, 7, 2))
p=0.75
q=1-p
P[,,1] <- matrix(c(0,0,0,1,0,0,0,
                  0,0,0,0,1,0,0,
                  0,0,0,0,0,1,0,
                  0,0,0,1,0,0,0,
                  0,0,0,0,1,0,0,
                  0,0,0,0,0,0,1,
                  0,0,0,0,0,0,1), nrow=7, ncol=7, byrow=TRUE)
0,0,p,0,q,0,0,
                  0,0,1,0,0,0,0,
                  0,0,0,0,1,0,0,
                  0,0,0,0,0,1,0,
                  0,0,0,0,0,0,1,
                  0,0,0,0,0,0,1), nrow=7, ncol=7, byrow=TRUE)
R <- array(0, c(7, 2))
R[,1] <- matrix(c(0, -2, -1, -1, -3, 5, 0), nrow=1, ncol=7, byrow=TRUE)
R[,2] <- R[,1]
mdp check(P, R)
mdp_value_iteration(P,R,discount=1)
$V
[1] 1.5 1.5 4.0 1.0 2.0 5.0 0.0
$policy
[1] 2 2 1 2 2 1 1
```

Chapter 6 - Classification of States

\$10

a) The graph of the chain is



This Markov chain is reducible because its state space can be partitioned into two communicating classes as $S = \{0\} \cup \{1, 2, 3\}$.

- b) State (0) has period 1 and states (1), (2), (3) have period 3.
- c) We have

$$p_{0,0} = \mathbb{P}(T_0 < \infty \mid X_0 = 0) = \mathbb{P}(T_0 = 1 \mid X_0 = 0) = rac{1}{3},$$

and

$$\mathbb{P}(T_0 = \infty \mid X_0 = 0) = 1 - \mathbb{P}(T_0 < \infty \mid X_0 = 0) = rac{2}{3}.$$

We also have

$$egin{aligned} \mathbb{P}(R_0 < \infty \mid X_0 = 0) &= & \mathbb{P}(T_0 = \infty \mid X_0 = 0) \sum\limits_{n \geqslant 1} \left(\mathbb{P}(T_0 < \infty \mid X_0 = 0)
ight)^n \ &= & rac{2}{3} \sum\limits_{n \geqslant 1} \left(rac{1}{3}
ight)^n = 1. \end{aligned}$$

d) There are no absorbing states, state (0) is transient, and states (1), (2), (3) are recurrent by Corollary 6.7. State (0) is transient since $\mathbb{P}(R_0 < \infty \mid X_0 = 0) = 1$, as expected according to (5.4.3).

Exercise 6.2

- a) The chain is reducible and its communicating classes are $\{0, 1\}$ and $\{2\}$.
- b) States 0 and 1 are transient and the state 2 is (positive) recurrent and absorbing.
- c) All states have period **1**.

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

$$\mathbb{P}(T_0^r=\infty|X_0=0)\geqslantrac{1}{4}>0,$$

and the remaining states (1), (2), (3) are also transient because they communicate with the transient state (0), cf. Corollary 6.7. By pathwise or first step analysis we can actually check that

$$\mathbb{P}(T_0^r = \infty | X_0 = 0) = rac{1}{4} (\mathbb{P}(T_0^r = \infty | X_0 = 1) + \mathbb{P}(T_0^r = \infty | X_0 = 2) + \mathbb{P}(T_0^r = \infty | X_0 = 3)) = rac{3}{4}.$$

d) The Markov chain is reducible because its state space $S = \{0, 1, 2, 3, 4\}$ can be partitioned into two communicating classes $\{0, 1, 2, 3\}$ and $\{4\}$. MH3512 AY19-20


- 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



This chain is reducible, with communicating classes $\{0, 2\}$, $\{1\}$, and $\{3\}$.

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 6.10 and Corollary 6.7, and they are also positive recurrent since the state space is finite. There are no absorbing states.

Exercise 6.6

- a) The communicating classes are $\{0\}, \{1, 2\}, \{3, 4, 5\}, ,$
- b) the transient states are (1), (2), (3), (4), (5).
- c) the recurrent state is \bigcirc .
- d) the positive recurrent state is 0.
- e) the period of state (0) is 1, the period of states (1) and (2) is 2, and the period of states (3), (4), (5) is 3.

Exercise 6.7

a) This question refers to the Ehrenfest chain, cf. Example (iv) page 151. By (4.3.2) and (4.3.3) the transition matrix of this chain takes the form

0	1	0	0	• • •	• • •	0	0
1/N	0	(N-1)/N	0	•••	•••	0	0
0	2/N	0	(N-2)/N	•••	• • •	0	0
0	0	3/N	0	• • •	•••	0	0
•	• •	• •	·	•.	• •	• •	:
• •	• •	• •	•.	•.	• •	• •	:
0	0	• • •	• •	0	3/N	0	0
0	0	• • •	0	(N-2)/N	0	2/N	0
0	0	• • •	0	0	(N-1)/N	0	1/N
0	0	• • •	0	0	0	1	0

Note that the random choice made is that of a ball among N balls, not a box among two boxes.

b) All states have period **2** and are recurrent. There is no transient state.

Exercise 6.8

a) For the Markov chain of Exercise 4.10-(a) we find that state 0 is (positive) recurrent as it is absorbing, while every state k is transient when $k \ge 1$ since $\mathbb{P}(T_k^r < \infty \mid X_0 = k) = 0, k \ge 1$. State 0 is aperiodic, and all other states have period_M43512 AY19-20

b) For the success runs chain of Exercise 4.10-(b), we have

$$\mathbb{P}(T_0^r < \infty \mid X_0) = q \sum\limits_{n \geqslant 0} p^n = rac{q}{1-p} = 1,$$

hence state 0 is recurrent. Since the chain is irreducible, it is recurrent by Corollary 6.7. State 0 is also positive recurrent as we have

$$\mathbb{E}[T_{0}^{r} \mid X_{0}] = q \sum\limits_{n \geqslant 1} n p^{n-1} = rac{q}{(1-p)^{2}} = rac{1}{q}.$$

Next, by first step analysis we have

$$egin{aligned} \mathbb{E}[T_k^r \mid X_0 = 0] &= \sum\limits_{i=1}^k \mathbb{E}[T_i^r \mid X_0 = i-1] \ &= \sum\limits_{i=1}^k ig(q(1 + \mathbb{E}[T_{i-1}^r \mid X_0 = 0]) + p) \ &= \sum\limits_{i=1}^k ig(1 + q \, \mathbb{E}[T_{i-1}^r \mid X_0 = 0]) \ &= k + q \sum\limits_{i=1}^k \mathbb{E}[T_{i-1}^r \mid X_0 = 0], \qquad k \geqslant 1, \end{aligned}$$

which shows by induction that $\mathbb{E}[T_k^r \mid X_0 = 0] < \infty, k \ge 0$. On the other hand, we have

$$egin{array}{rl} \mathbb{E}[T_k^r \mid X_0 = k] &= 1 + q \, \mathbb{E}[T_k^r \mid X_0 = 0] + p \, \mathbb{E}[T_k^r \mid X_0 = k + 1] \ &= 1 + q \, \mathbb{E}[T_k^r \mid X_0 = 0] + p \, \mathbb{E}[T_k^r \mid X_0 = k], \hspace{1em} k \geqslant 0, \end{array}$$

since the starting point does not matter as long as we start from $k \ge 1$, hence

$$\mathbb{E}[T_k^r \mid X_0=k] = rac{1}{q} + \mathbb{E}[T_k^r \mid X_0=0], \qquad k \geqslant 1,$$

which shows that $\mathbb{E}[T_k^r \mid X_k = 0] < \infty, k \ge 0$, therefore the success runs Markov chain is also positive recurrent.

0615

Problem 6.9 By the same argument as in Exercise 3.2, by first step analysis we have

$$egin{aligned} \mathbb{E}[T_0^r \mid X_0 = 1] &= rac{1}{lpha + 1} imes 1 + rac{lpha}{lpha + 1} imes (1 + \mathbb{E}[T_0^r \mid X_0 = 2]) \ &= rac{1}{lpha + 1} imes 1 + rac{lpha}{lpha + 1} imes (1 + 2 \, \mathbb{E}[T_0^r \mid X_0 = 1]) \ &= 1 + rac{2lpha}{lpha + 1} \, \mathbb{E}[T_0^r \mid X_0 = 1], \end{aligned}$$

hence

$$\mathbb{E}[T_0^r \mid X_0 = 1] = \frac{\alpha + 1}{1 - \alpha}, \qquad \alpha < 1.$$

By the space homogeneity of the chain, we deduce that

$$egin{aligned} \mathbb{E}[T_0^r \mid X_0 = k] &= \sum\limits_{i=1}^k \mathbb{E}[T_{i-1}^r \mid X_0 = i] \ &= k \, \mathbb{E}[T_0^r \mid X_0 = 1] \ &= k rac{lpha + 1}{1 - lpha}, \qquad k \geqslant 1, \quad lpha < 1, \end{aligned}$$

and

$$\mathbb{E}[T_0^r \mid X_0 = 0] = 1 + \mathbb{E}[T_0^r \mid X_0 = 1] = 1 + rac{lpha+1}{1-lpha} = rac{2}{1-lpha},$$

hence state 0 is positive recurrent if $\alpha < 1$. On the other hand, we have

$$\begin{split} \mathbb{E}[T_1^r \mid X_0 = 1] &= \frac{1}{\alpha + 1} \times 2 + \frac{\alpha}{\alpha + 1} \times (1 + \mathbb{E}[T_1^r \mid X_0 = 2]) \\ &= 1 + \frac{1}{\alpha + 1} + \frac{\alpha}{\alpha + 1} \mathbb{E}[T_0^r \mid X_0 = 1] \\ &= 1 + \frac{1}{\alpha + 1} + \frac{\alpha}{1 - \alpha} \\ &= \frac{2}{1 - \alpha^2}. \end{split}$$

ॳ 616

Similarly to the above, we have $\mathbb{E}[T_1^r \mid X_0 = 0] = 1$ and, for $k \ge 1$,

$$egin{aligned} \mathbb{E}[T_{k+1}^r \mid X_0 = k] &= rac{lpha}{lpha + 1} imes 1 + rac{1}{lpha + 1} imes (1 + \mathbb{E}[T_{k+1}^r \mid X_0 = k - 1]) \ &= 1 + rac{1}{lpha + 1} imes (\mathbb{E}[T_k^r \mid X_0 = k - 1] + \mathbb{E}[T_{k+1}^r \mid X_0 = k]), \end{aligned}$$

hence

$$\mathbb{E}[T_{k+1}^r \mid X_0 = k] = rac{lpha+1+\mathbb{E}[T_k^r \mid X_0 = k-1]}{lpha}, \qquad k \geqslant 1,$$

which can be solved as

$$\mathbb{E}[T_{k+1}^r \mid X_0 = k] = rac{2lpha^{-k} - lpha - 1}{1 - lpha}, \qquad k \geqslant 0,$$

see e.g. the command RSolve[f[k+1]==1+1/a+f[k]/a, f[0] == 1, f[k], k], which shows that

$$\begin{split} \mathbb{E}[T_k^r \mid X_0 = k] \\ &= \frac{1}{\alpha + 1} \left(1 + \mathbb{E}[T_k^r \mid X_0 = k - 1] \right) + \frac{\alpha}{\alpha + 1} \left(1 + \mathbb{E}[T_k^r \mid X_0 = k + 1] \right) \\ &= 1 + \frac{1}{\alpha + 1} \mathbb{E}[T_k^r \mid X_0 = k - 1] + \frac{\alpha}{\alpha + 1} \mathbb{E}[T_k^r \mid X_0 = k + 1] \\ &= 1 + \frac{1}{\alpha + 1} \mathbb{E}[T_k^r \mid X_0 = k - 1] + \frac{\alpha}{1 - \alpha} \\ &= \frac{1}{1 - \alpha} + \frac{1}{\alpha + 1} \mathbb{E}[T_k^r \mid X_0 = k - 1] \\ &= \frac{2\alpha^{-k}}{1 - \alpha^2}, \\ &< \infty, \end{split}$$

therefore the chain is positive recurrent if and only if $\alpha < 1$.

Chapter 7 - Long-Run Behavior of Markov Chains

Exercise 7.1

a) The chain is reducible, state (1) is aperiodic and all other states have period 0. On the other hand, we check that state (1) is positive recurrent and all other states are transient, cf. Exercise 6.8-(a). Therefore, the assumptions of Theorems 7.2, 7.8 and 7.10 are not satisfied.

b) Writing $\boldsymbol{\pi} = \boldsymbol{\pi} \boldsymbol{P}$, *i.e.*

$$[\pi_0,\pi_1,\pi_2,\pi_3,\cdots] = [\pi_0,\pi_1,\pi_2,\pi_3,\cdots] imes egin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \ q & 0 & p & 0 & 0 & \cdots \ q & 0 & 0 & p & 0 & \cdots \ q & 0 & 0 & 0 & p & \cdots \ q & 0 & 0 & 0 & p & \cdots \ q & 0 & 0 & 0 & 0 & \cdots \ \vdots & 0 & 0 & 0 & 0 & \ddots \ \end{bmatrix},$$

shows that

$$\begin{cases} \pi_0 = \pi_0 + q(\pi_1 + \pi_2 + \cdots) = \pi_0 + q(1 - \pi_0) = q + p\pi_0, \\ \pi_1 = 0, \\ \pi_2 = p\pi_1, \\ \pi_3 = p\pi_2, \\ \vdots \end{cases}$$

hence $\pi_k = 0, k \ge 1$, and the stationary distribution is given by

$$\pi = [1, 0, 0, 0, \ldots].$$

c) Taking the limit as n tends to infinity in (B.30), we find

$$\lim_{n o \infty} P^n = egin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \ 1 & 0 & 0 & 0 & 0 & \cdots \ 1 & 0 & 0 & 0 & 0 & \cdots \ 1 & 0 & 0 & 0 & 0 & \cdots \ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

MH3512 AY19-20

(€ 618

hence the limiting distribution coincides with the stationary distribution π obtained in (b).

MH3512 AY19-20

Exercise 7.2

- a) The chain irreducible, aperiodic, recurrent and positive recurrent by Exercise 6.8-(c). Therefore, the assumptions of Theorems 7.2, 7.8 and 7.10 are all satisfied and the chain admits a limiting distribution which coincides with its stationary distribution.
- b) Writing $\boldsymbol{\pi} = \boldsymbol{\pi} \boldsymbol{P}$, *i.e.*

$$[\pi_0,\pi_1,\pi_2,\pi_3,\cdots] = [\pi_0,\pi_1,\pi_2,\pi_3,\cdots] imes egin{bmatrix} q & p & 0 & 0 & 0 & \cdots \ q & 0 & p & 0 & 0 & \cdots \ q & 0 & 0 & p & 0 & \cdots \ q & 0 & 0 & 0 & p & \cdots \ q & 0 & 0 & 0 & 0 & \cdots \ \vdots & 0 & 0 & 0 & 0 & \ddots \ \end{bmatrix},$$

shows that

$$\left\{ egin{array}{l} \pi_1 = p \pi_0, \ \pi_2 = p \pi_1, \ \pi_3 = p \pi_2, \ dots \end{array}
ight.$$

hence by induction we get $\pi_k = p^k \pi_0, \, k \geqslant 1$. By the condition $\sum_{k \ge 0} \pi_k = 1$ we find

$$1=\sum\limits_{k\geqslant 0}\pi_k=\pi_0\sum\limits_{k\geqslant 0}p^k=rac{\pi_0}{1-p},$$

which shows that $\pi_0 = q$. We conclude that π is the geometric distribution

$$\pi_k = (1-p)p^k, \qquad k \in \mathbb{N}$$
 .

For the chain of Exercise 4.10-(a) we find $\pi_1 = 0$ hence the stationary distribution is given by $\pi = (1, 0, 0, \ldots)$.

c) Taking the limit as n tends to infinity in (B.31), we find, when p < 1,

$$\lim_{n \to \infty} P^{n} = \begin{bmatrix} q & qp & qp^{2} & qp^{3} & qp^{4} & \cdots \\ q & qp & qp^{2} & qp^{3} & qp^{4} & \cdots \\ q & qp & qp^{2} & qp^{3} & qp^{4} & \cdots \\ q & qp & qp^{2} & qp^{3} & qp^{4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

hence the chain admits the geometric distribution with parameter p as a limiting distribution, which coincides with the stationary distribution π obtained in (b).

Exercise 7.3

- a) The Ehrenfest chain is irreducible and positive recurrent, and has period 2, therefore it satisfies the conditions of Theorem 7.10 and it admits a stationary distribution.
- b) The stationary distribution of the Ehrenfest chain satisfies the equation $\pi P = \pi$, which reads

	0	1	0	0	• • •	• • •	0	0
	1/N	0	(N-1)/N	0	• • •	• • •	0	0
	0	2/N	0	(N-2)/N	• • •	• • •	0	0
	0	0	3/N	0	• • •	• • •	0	0
r 1	•	• •	• •	•••	•	• •	• •	• •
$[\pi_0,\pi_1,\pi_2,\ldots] imes$	• •	• •	• •	•••	•••	• • •	• •	• •
	0	0	•••	• •	0	3/N	0	0
	0	0	• • •	0	(N-2)/N	0	2/N	0
	0	0	• • •	0	0	(N - 1)/N	0	1/N
	0	0	• • •	0	0	0	J MH351	2 AY9-20

$$= [\pi_0, \pi_1, \pi_2, \ldots] = \pi, \qquad \circlearrowright 621$$

cf. Exercise 6.7. This yields

$$\begin{cases} \pi_0 = \frac{\pi_1}{N} \\ \pi_1 = \pi_0 + \frac{2}{N} \pi_2, \\ \pi_2 = \pi_1 \frac{N-1}{N} + \frac{3}{N} \pi_3, \\ \pi_3 = \pi_2 \frac{N-2}{N} + \frac{4}{N} \pi_4, \\ \vdots \end{cases}$$

from which we deduce

$$\left\{egin{array}{l} \pi_0 = rac{\pi_1}{N} \ \pi_1 = rac{2}{N-1} \pi_2, \ \pi_2 = rac{3}{N-2} \pi_3, \ \pi_3 = rac{4}{N-3} \pi_4, \ dots \end{array}
ight.$$

hence

$$\pi_0 = rac{1}{N} imes rac{2}{N-1} imes rac{3}{N-2} imes \cdots imes rac{k}{N-k+1} \pi_k = rac{k!(N-k)!}{N!} \pi_k,
onumber \ \pi_k = inom{N}{k} \pi_0, \qquad k=0,\ldots,N.$$

or

In addition we have

$$1 = \sum\limits_{k=0}^{N} \pi_k = \pi_0 \sum\limits_{k=0}^{N} {N \choose k} = (1+1)^n \pi_0 = 2^N \pi_0,$$

hence $\pi_0 = 2^{-N}$ and $(\pi_k)_{k=0,1,\dots,N}$ is the binomial distribution with parameter (N, 1/2), *i.e.*

$$\pi_k = \binom{N}{k} \left(\frac{1}{2}\right)^N, \qquad k = 0, 1, \dots, N.$$
(B.43)

Conversely, from (B.43) we can recover

$$\begin{split} \sum_{l=0}^{N} \pi_{l} P_{l,k} &= \pi_{k-1} \frac{N - (k-1)}{N} + \pi_{k+1} \frac{k+1}{N} \\ &= \binom{N}{k-1} \left(\frac{1}{2}\right)^{N} \frac{N - k + 1}{N} + \binom{N}{k+1} \left(\frac{1}{2}\right)^{N} \frac{k+1}{N} \\ &= \left(\frac{1}{2}\right)^{N} \frac{(N-1)!k}{(N-k)!k!} + \left(\frac{1}{2}\right)^{N} \frac{(N-1)!(N-k)}{(N-k)!k!} \\ &= \left(\frac{1}{2}\right)^{N} \binom{N}{k} \\ &= \pi_{k}, \quad k = 1, 2, \dots, N-1, \end{split}$$

and similarly for k = 0 and k = N, showing that $\pi = \pi P$.

c) The powers of the transition matrix of the Ehrenfest chain can be computed using the following R code:

```
install.packages("expm")
library(expm)
Ehrenfest <-
function(n) {
  States <- c(0, seq(1,n))
  TPM <- matrix(0,nrow=length(States),ncol=length(States),dimnames=
                 list(seq(0,n),seq(0,n)))
  tran_prob <- function(i,n) {</pre>
    tranRow <- rep(0,n+1)
   if(i==0) tranRow[2] < -1
   if(i==n) tranRow[(n+1)-1] < 1
    if(i!=0 & i!=n) {
     i=i+1
      tranRow[j-1] <- i/(n)
     tranRow[j+1] < -1-i/(n)
    }
    return(tranRow)
  }
  for(j in 0:(n))TPM[j+1,] <- tran_prob(j,n)
  return(TPM)
}
P = Ehrenfest(4)
P%^%1000
P%^%1001
P%^%1002
```

which shows that, although $(P^n)_{n \ge 1}$ admits two converging subsequences, $\lim_{n \to \infty} P^n$ does not exist. Note that the conditions of Theorems 7.8 and 7.10 are not satisfied here.

523

Exercise 7.4

a) The Bernoulli-Laplace chain is irreducible, aperiodic and positive recurrent. Therefore it satisfies all the assumptions of Theorems 7.2,
 7.8 and 7.10, the chain admits a limiting distribution which coincides with its stationary distribution.

b) The equation $\pi = \pi P$ reads

$$\begin{cases} \pi_0 = \frac{\pi_1}{N^2} \\ \pi_1 = \pi_0 + \frac{2(N-1)}{N^2} \pi_1 + \frac{2^2}{N^2} \pi_2, \\ \pi_2 = \frac{(N-1)^2}{N^2} \pi_1 + \frac{4(N-2)}{N^2} \pi_2 + \frac{3^2}{N^2} \pi_3, \\ \pi_3 = \frac{(N-2)^2}{N^2} \pi_1 + \frac{6(N-3)}{N^2} \pi_2 + \frac{3^2}{N^2} \pi_3, \\ \vdots \end{cases}$$

from which we deduce

$$\left\{ egin{array}{l} \pi_{0} = rac{\pi_{1}}{N^{2}} \ \pi_{1} = rac{2^{2}}{(N-1)^{2}} \pi_{2}, \ \pi_{2} = rac{3^{2}}{(N-2)^{2}} \pi_{3}, \ \pi_{3} = rac{4^{2}}{(N-3)^{2}} \pi_{4}, \ dots \end{array}
ight.$$

hence

$$egin{aligned} \pi_0 &= rac{1}{N^2} imes rac{2^2}{(N-1)^2} imes rac{3^2}{(N-2)^2} imes \cdots imes rac{k^2}{(N-k+1)^2} \pi_k = rac{k!^2 (N-k)!^2}{N!^2} \pi_k, \ \pi_k &= inom{k} = inom{N}{k}^2 \pi_0, \qquad k=0,\ldots,N. \end{aligned}$$

or

In addition we have

$$1=\sum\limits_{k=0}^N \pi_k=\pi_0\sum\limits_{k=0}^N {N \choose k}^2={2N \choose N}\pi_0,$$

hence $\pi_0 = 1/\binom{2N}{N}$ and $(\pi_k)_{k=0,1,...,N}$ is the probability distribution given by

$$\pi_k = rac{inom{N}{k}^2}{inom{2N}{N}}, \qquad k=0,1,\ldots,N.$$

c) The Bernoulli-Laplace chain admits a limiting distribution which coincides with its stationary distribution by Theorem 7.8.

Exercise 7.5

a) We have

0	0.5	0	0.5	
0.5	0	0.5	0	
0	0.5	0	0.5	•
0.5	0	0.5	0	

b) By first step analysis, we have

$$\left\{egin{array}{l} \mu_0(0) = 1 + rac{1}{2} \mu_0(1) + rac{1}{2} \mu_0(3) \ \mu_0(1) = 1 + rac{1}{2} \mu_0(2) \ \mu_0(2) = 1 + rac{1}{2} \mu_0(1) + rac{1}{2} \mu_0(3) \ \mu_0(3) = 1 + rac{1}{2} \mu_0(2). \end{array}
ight.$$

625

By symmetry of the maze we have $\mu_0(1) = \mu_0(3)$, hence

$$\left\{egin{array}{l} \mu_0(0)=\mu_0(2)=1+\mu_0(1)\ \mu_0(1)=\mu_0(3)=1+rac{1}{2}\mu_0(2) \end{array}
ight.$$

and

$$\mu_0(0) = 4, \quad \mu_0(1) = 3, \quad \mu_0(2) = 4, \quad \mu_0(3) = 3$$

The symmetry of the problem shows that we have, $\mu_0(1) = \mu_0(3)$, which greatly simplifies the calculations. One has to pay attention to the fact that $\mu_0(0)$ is a return time, *not* a hitting time, so that $\mu_0(0)$ cannot be equal to zero.

c) Clearly, the probability distribution

$$(\pi_0,\pi_1,\pi_2,\pi_3)=\left(rac{1}{4},rac{1}{4},rac{1}{4},rac{1}{4}
ight)$$

is invariant and satisfies the condition $\pi = \pi P$, see also Exercise 7.14.

Recall that the family $(\pi_0, \pi_1, \pi_2, \pi_3)$ is a *(probability) distribution* if and only if

(i) $\pi_0 \ge 0$, $\pi_1 \ge 0$, $\pi_2 \ge 0$, $\pi_3 \ge 0$,

and

(ii)

$$\pi_0 + \pi_1 + \pi_2 + \pi_3 = 1.$$

In the present situation we have

$$\pi_0=\pi_1=\pi_2=\pi_3$$

and

$$\pi_0 + \pi_1 + \pi_2 + \pi_3 = 1 = 4\pi_0$$

which necessarily yields

$$\pi_0=\pi_1=\pi_2=\pi_3=rac{1}{4}=rac{1}{\mu_0(0)}.$$

MH3512 AY19-20 $\,$

ॳ 626

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 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 $\pi = \pi P$, *i.e.*

$$\left\{egin{array}{l} \pi_D = (1-a)\pi_D + b\pi_N = rac{1}{4}\pi_D + rac{1}{4}\pi_N \ \pi_N = a\pi_D + (1-b)\pi_N = rac{3}{4}\pi_D + rac{3}{4}\pi_N \end{array}
ight.$$

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 \widehat{N} starting from state \widehat{D} satisfies

$$\mu_N(D) = (1-a)(1+\mu_N(D)) + a \qquad (B.44)$$

hence $\mu_N(D) = 1/a = 4/3$.

Additional comments:

(i) The stationary distribution $[\pi_D, \pi_N]$ satisfies

$$[\pi_D,\pi_N] = \left[rac{\mu_N(D)}{\mu_D(N)+\mu_N(D)},rac{\mu_D(N)}{\mu_D(N)+\mu_N(D)}
ight],$$

as the ratios of times spent in states D and N, since $\mu_N(D) = 1/a$ and $\mu_D(N) = 1/b$.

(ii) The value of $\mu_N(D)$ may also be recovered as

$$\mu_N(D) = rac{3}{4} \sum\limits_{k \geqslant 1} k \left(rac{1}{4}
ight)^{k-1} = rac{3}{4} rac{1}{(1-1/4)^2} = rac{4}{3},$$

cf. (A.4).

e) The average time $\mu_D(N)$ to reach state \bigcirc starting from state \bigtriangledown satisfies

$$\mu_D(N) = (1-b)(1+\mu_D(N)) + b \tag{B.45}$$

hence $\mu_N(D) = 1/b = 4$.

Additional comments:

(i) The value of $\mu_D(N)$ may also be recovered as

$$\mu_D(N) = rac{1}{4} \sum_{k \geqslant 1} k \left(rac{3}{4}
ight)^{k-1} = rac{1}{4} rac{1}{(1-3/4)^2} = 4,$$

cf. (A.4).

(ii) The values of μ_D^D and μ_N^N can also be computed from $\mu_D^D = 1/\pi_D$ and $\mu_N^N = 1/\pi_N$.

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 6.10 and Corollary 6.7.

c) It suffices to consider the subchain on $\{0,1,2\}$ with transition matrix

$$ilde{P} = egin{bmatrix} 0 & 1 & 0 \ 0.2 & 0 & 0.8 \ 0.3 & 0 & 0.7 \end{bmatrix},$$

and to solve $\pi = \pi \tilde{P}$, *i.e.*

$$\left\{egin{array}{l} \pi_0 = 0.2\pi_1 + 0.3\pi_2 \ \pi_1 = \pi_0 \ \pi_2 = 0.8\pi_1 + 0.7\pi_2 \end{array}
ight.$$

which yields $\pi_1 = \pi_0$ and $0.3\pi_2 = 0.8\pi_1 = 0.8\pi_0$, with

$$1 = \pi_0 + \pi_1 + \pi_2 = 2\pi_0 + \frac{8}{3}\pi_0,$$

i.e.

$$\pi_0=rac{3}{14}, \ \ \pi_1=rac{3}{14}, \ \ \pi_2=rac{4}{7},$$

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 (1) starting from state (k), we have

$$\left\{egin{array}{l} h_0(0)=0\ h_0(1)=0.2(1+h_0(0))+0.8(1+h_0(2))\ h_0(2)=0.3(1+h_0(0))+0.7(1+h_0(2))\ h_0(3)=0.4(1+h_0(0))+0.6(1+h_0(1)), \end{array}
ight.$$

$$\begin{pmatrix} h_0(0) = 0, \\ h_0(1) = 1 + 0.8h_0(2), \\ h_0(2) = 1 + 0.7h_0(2), \\ h_0(3) = 1 + 0.6h_0(1), \end{cases}$$
 or
$$\begin{cases} h_0(0) = 0, \\ h_0(1) = 1 + 0.8h_0(2), \\ 0.3h_0(2) = 1, \\ h_0(3) = 1 + 0.6h_0(1), \end{cases}$$

hence

$$h_0(0)=0, \ \ h_0(1)=rac{11}{3}, \ \ h_0(2)=rac{10}{3}, \ \ h_0(3)=rac{16}{5},$$

and the mean time to reach state (1) 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

$$h_0(2) = 0.3 \sum\limits_{k \geqslant 1} k(0.7)^{k-1} = rac{0.3}{(1-0.7)^2} = rac{10}{3}.$$

Note that the value of $h_0(2)$ could also be computed by restriction to the sub-chain $\{0, 1, 2\}$, by solving

$$\left\{egin{array}{l} h_0(0)=0\ h_0(1)=0.2(1+h_0(0))+0.8(1+h_0(2))\ h_0(2)=0.3(1+h_0(0))+0.7(1+h_0(2)). \end{array}
ight.$$

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

$$\pi_0 = c imes 161, \qquad \pi_1 = c imes 460, \qquad \pi_2 = c imes 320, \qquad \pi_3 = c imes 170,$$

see *here*. 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

$$c = \frac{1}{161 + 460 + 320 + 170} = \frac{1}{1111},$$

hence

$$\pi_0 = \frac{161}{1111}, \quad \pi_1 = \frac{460}{1111}, \quad \pi_2 = \frac{320}{1111}, \quad \pi_3 = \frac{170}{1111}.$$
 (B.46)

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

$$\left\{egin{array}{ll} \mu_0(0)&=1+\mu_0(1),\ \mu_0(1)&=0.1+0.4(1+\mu_0(1))+0.2(1+\mu_0(2))+0.3(1+\mu_0(3))\ &=1+0.4\mu_0(1)+0.2\mu_0(2)+0.3\mu_0(3),\ \mu_0(2)&=0.2+0.2(1+\mu_0(1))+0.5(1+\mu_0(2))+0.1(1+\mu_0(3))\ &=1+0.2\mu_0(1)+0.5\mu_0(2)+0.1\mu_0(3),\ \mu_0(3)&=1+0.3\mu_0(1)+0.4\mu_0(2), \end{array}
ight.$$

hence

$$\mu_0(1) = rac{950}{161}, \qquad \mu_0(2) = rac{860}{161}, \qquad \mu_0(3) = rac{790}{161},$$

see *here* or *here*. 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

$$\mu_0(0) = 1 + \mu_0(1) = 1 + rac{950}{161} = rac{161 + 950}{161} = rac{1111}{161},$$

hence the relation $\pi_0 = 1/\mu_0(0)$ is satisfied from (B.46).

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 6.12. By Proposition 6.15, 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. This equation reads

 $\left\{ egin{array}{l} \pi_0 = rac{1}{4} \pi_1 + rac{1}{2} \pi_3 \ \pi_1 = rac{1}{2} \pi_0 + rac{1}{3} \pi_2 \ \pi_2 = rac{3}{4} \pi_1 + rac{1}{2} \pi_3 \ \pi_3 = rac{1}{2} \pi_0 + rac{2}{3} \pi_2, \end{array}
ight.$

hence

$$\pi_1 = \pi_0, \quad \pi_2 = 3\pi_0/2, \quad \pi_3 = 3\pi_0/2,$$

see *here*, under the condition

 $\pi_0 + \pi_1 + \pi_2 + \pi_3 = 1,$

which yields the solution

$$\pi_0=2c, \ \ \pi_1=2c, \ \ \pi_2=3c, \ \ \pi_3=3c,$$

and c = 1/10, *i.e*

$$\pi_0=20\%, \ \ \pi_1=20\%, \ \ \pi_2=30\%, \ \ \pi_3=30\%.$$

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 $X_n = k \ge 2$ units of time until failure, we know with certainty that at the next time step n + 1 there will remain $X_{n-1} = k - 1 \ge 1$ units of time until failure. Hence at any time $n \ge 1$ we have

$$X_n = 4 \implies X_{n+1} = 3 \implies X_{n+2} = 2 \implies X_{n+3} = 1$$

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

$$\mathbb{P}(X_{n+1}=k \mid X_n=1) = \mathbb{P}(T=k), \qquad k=1,2,3,4,$$
 MH3512 AY19-2

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

$$\lim_{n o\infty} \mathbb{P}(X_n=1)$$

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

$$\pi_1 = \lim_{n o \infty} \mathbb{P}(X_n = 1),$$

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:

$$\left\{egin{array}{ll} \pi_1 = 0.1 \pi_1 + \pi_2 \ \pi_2 = 0.2 \pi_1 + \pi_3 \ \pi_3 = 0.3 \pi_1 + \pi_4 \ \pi_4 = 0.4 \pi_1. \end{array}
ight.$$

hence

$$\pi_2=0.9\pi_1, \qquad \pi_3=0.7\pi_1, \qquad \pi_4=0.4\pi_1,$$

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

$$\pi_1 = rac{1}{3}, \quad \pi_2 = rac{9}{30}, \quad \pi_3 = rac{7}{30}, \quad \pi_4 = rac{4}{30}.$$

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

$$T_{\{A,B\}} = \inf\{n \geqslant 0 \; : \; X_n \in \{A,B\}\}$$

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

$$\mathbb{P}(T_{\{A,B\}}<\infty\mid X_0=C)=rac{1}{3},$$

which can also be recovered using a geometric series as

$$\mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C) = 0.2 \sum_{n \geqslant 0} (0.4)^n = rac{0.2}{1 - 0.4} = rac{1}{3}.$$

On the other hand, $\{A, B\}$ is a closed two-state chain with transition matrix

$$\left[egin{array}{ccc} 1-a & a \ b & 1-b \end{array}
ight] = \left[egin{array}{ccc} 0.6 & 0.4 \ 0.3 & 0.7 \end{array}
ight],$$

hence, starting from any state within $\{A, B\}$, the long run probability of being in A is given by

$$\begin{split} \lim_{n \to \infty} \mathbb{P}(X_n &= A \mid X_0 \in \{A, B\}) \\ &= \lim_{n \to \infty} \mathbb{P}(X_n = A \mid X_0 = A) \frac{\mathbb{P}(X_0 = A)}{\mathbb{P}(X_0 \in \{A, B\})} \\ &+ \lim_{n \to \infty} \mathbb{P}(X_n = A \mid X_0 = B) \frac{\mathbb{P}(X_0 = B)}{\mathbb{P}(X_0 \in \{A, B\})} \\ &= \frac{b}{a + b} \left(\frac{\mathbb{P}(X_0 = A)}{\mathbb{P}(X_0 \in \{A, B\})} + \frac{\mathbb{P}(X_0 = B)}{\mathbb{P}(X_0 \in \{A, B\})} \right) \\ &= \frac{0.3}{0.3 + 0.4} = \frac{3}{7}. \end{split}$$

Since

$$\{X_n=A\}\subset\{T_{A,B}\leqslant n\}\subset\{T_{A,B}<\infty\},\qquad n\in\mathbb{N},$$

we conclude that

$$lpha \ := \ \lim_{n o \infty} \mathbb{P}(X_n = A \mid X_0 = C)$$

$$= \lim_{n \to \infty} \mathbb{P}(T_{A,B} < \infty \text{ and } X_n = A \mid X_0 = C)$$

$$= \lim_{n \to \infty} \mathbb{P}(T_{A,B} < \infty \text{ and } X_{n+T_{A,B}} = A \mid X_0 = C)$$

$$= \mathbb{P}(T_{A,B} < \infty \mid X_0 = C) \lim_{n \to \infty} \mathbb{P}(X_{n+T_{A,B}} = A \mid T_{A,B} < \infty \text{ and } X_0 = C)$$

$$= \mathbb{P}(T_{A,B} < \infty \mid X_0 = C)$$

$$\times \lim_{n \to \infty} \mathbb{P}(X_{n+T_{A,B}} = A \mid T_{A,B} < \infty, X_{T_{A,B}} \in \{A, B\}, X_0 = C)$$

$$= \mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C) \lim_{n \to \infty} \mathbb{P}(X_n = A \mid X_0 \in \{A, B\})$$

$$= \frac{1}{3} \times \frac{3}{7}$$

$$= \frac{1}{7},$$

where we used the strong Markov property, cf. Exercise 5.9.

The value α of this probability could also be obtained in a faster way by the following more intuitive first step analysis:

$$lpha = 0.4 imes lpha + 0.2 imes rac{0.3}{0.7},$$

which recovers $\alpha = 1/7$, although this way of reasoning is a bit too fast to be truly recommended.

On the other hand, the numerical computation of the 20th power of P shows that

$$P^{20}=egin{bmatrix} 0.4285 & 0.5714 & 0 & 0 & 0\ 0.4285 & 0.5714 & 0 & 0 & 0\ 0.4285 & 0.5714 & 0 & 0 & 0\ 0.1428 & 0.1904 & 1.099510^{-8} & 0 & 0.66666\ 0.2500 & 0.3333 & 1.099510^{-8} & 1.048510^{-14} & 0.4166\ 0 & 0 & 0 & 0 & 1 \ \end{bmatrix},$$

which recovers

$$\lim_{n o\infty} \mathbb{P}(X_n=A\mid X_0=C)=rac{1}{7}\simeq 0.142857$$

MH3512 AY19-20

() 636

up to 6 decimals, and one can reasonably conjecture that the limiting distribution of the chain is given by

$$\lim_{n \to \infty} P^n = \begin{vmatrix} 3/7 & 4/7 & 0 & 0 & 0 \\ 3/7 & 4/7 & 0 & 0 & 0 \\ 1/7 & 4/21 & 0 & 0 & 2/3 \\ 1/4 & 1/3 & 0 & 0 & 5/12 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix},$$
 (B.47)

which could also be recovered term by term using the above method. From this matrix one also sees clearly that C and D are transient states since they correspond to vanishing columns in the above matrix. From (B.47) we also note that $\lim_{n\to\infty} \mathbb{P}(X_n = i \mid X_0 = j)$ is dependent on the initial state (j), due to the fact that the chain is not irreducible and Theorems 7.2 and 7.8 cannot be applied except if the initial state belongs to the communicating class $\{A, B\}$, for which (3/7, 4/7) is a limiting distribution independent of the initial state. We also check that every limiting distribution in (B.47) is a stationary distribution solution of $\pi = \pi P$, cf. Proposition 7.7.

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.

MH3512 AY19-20

ॳ 637

$$\left[\begin{array}{rr}4/5&1/5\\3/4&1/4\end{array}\right]$$

b) The stationary distribution $\pi = (\pi_C, \pi_T)$ is solution of $\pi = \pi P$, *i.e.*

$$\left\{ egin{array}{l} \pi_{C} = rac{4}{5}\pi_{C} + rac{3}{4}\pi_{T} \ \pi_{T} = rac{1}{5}\pi_{C} + rac{1}{4}\pi_{T} \end{array}
ight.$$

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 \widehat{T} starting from \widehat{C} and \widehat{T} , respecticely. By first step analysis we have

which has for solution $\mu_T(C) = 5$ and $\mu_T(T) = 19/4$. Consequently, it takes on average 19/4 = 4.75 vehicles after a truck until the next truck is seen under the bridge, and we check that the relation $\pi_D = 1/\mu_T(T)$ holds.

Exercise 7.14

a) We note that, due to the symmetry of \boldsymbol{P} , the vector

$$\pi = [\pi_1, \pi_2, \dots, \pi_N] := [1/N, \dots, 1/N]$$
 MH3512 AY19-20

is an invariant probability distribution for the chain, since we can check the balance condition

$$egin{array}{rcl} [\pi P]_j &=& \sum\limits_{i=1}^N \pi_i P_{i,j} \ &=& rac{1}{N} \sum\limits_{i=1}^N P_{j,i} \ &=& rac{1}{N} = \pi_j, \qquad j=1,2,\ldots,N. \end{array}$$

b) When N = 2 the chain can only have the transition matrix

$$P = \left[egin{array}{cc} 0 & 1 \ & \ 1 & 0 \end{array}
ight]$$

which clearly has period 2. When $N \ge 3$ the chain is aperiodic because it is irreducible and one can travel from (0) to (0) in 2 steps via the path $(0 \rightarrow (1 \rightarrow (0))$, and in 3 steps via the path $(0 \rightarrow (1 \rightarrow (2)))$.

Exercise 7.15 (Problem 5.25 continued).

- a) The chain is reducible and its communicating classes are $\{0, 1, 2, 3, 4\}$ and $\{5\}$.
- b) The limiting distribution is (0, 0, 0, 0, 0, 1) independently of the initial state because the states $\{0, 1, 2, 3, 4\}$ are transient (cf. Proposition 7.4) and state (5) is absorbing. This means that

$$\lim_{n o \infty} P^n = egin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \ 0 & 0 & 0 & 0 & 0 & 1 \ 0 & 0 & 0 & 0 & 0 & 1 \ 0 & 0 & 0 & 0 & 0 & 1 \ 0 & 0 & 0 & 0 & 0 & 1 \ 0 & 0 & 0 & 0 & 0 & 1 \ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

which would be difficult to recover by a direct computation of P^n .

For the stationary distribution, the equation $\pi = \pi P$ reads

$$egin{aligned} \pi_0 &= q\pi_0 + q\pi_1 + q\pi_2 \ \pi_1 &= p\pi_0 + p\pi_4 \ \pi_2 &= p\pi_1 \ \pi_3 &= p\pi_2 + p\pi_3 \ \pi_4 &= p\pi_3 \ \pi_5 &= q\pi_4 + \pi_5, \end{aligned}$$

i.e.

$$\left\{ egin{array}{l} p\pi_0 = q\pi_1 + q\pi_2 \ \pi_1 = p\pi_0 + p\pi_4 \ \pi_2 = p\pi_1 \ q\pi_3 = p\pi_2 \ \pi_4 = p\pi_3 \ \pi_4 = 0, \end{array}
ight.$$

hence $(\pi_0, \pi_1, \pi_2, \pi_3, \pi_4, \pi_5) = (0, 0, 0, 0, 0, 1)$, which coincides with the limiting distribution.

Note that the relation $\pi_i = 1/\mu_i(i)$ still holds for i = 0, 1, 2, 3, 4, 5, although not all of the assumptions of Theorems 7.2, 7.8 and 7.8 (notably the irreducibility condition) are satisfied here.

Exercise 7.16

a) We solve the system of equations

$$\left\{egin{array}{l} \pi_0 = q(\pi_0 + \pi_1 + \pi_2 + \pi_3) + \pi_4 = q + p \pi_4 \ \pi_1 = p \pi_0 \ \pi_2 = p \pi_1 = p^2 \pi_0 \ \pi_3 = p \pi_2 = p^3 \pi_0 \ \pi_4 = p \pi_3 = p^4 \pi_0, \end{array}
ight.$$

which yields

$$1=\pi_0+\pi_1+\pi_2+\pi_3+\pi_4=\pi_0(1+p+p^2+p^3+p^4),$$

MH3512 AY19-20

♦ 640

and

$$\left\{ egin{array}{l} \pi_0 = rac{1}{1+p+p_p^2+p^3+p^4} \ \pi_1 = rac{p^2}{1+p+p^2+p^3+p^4} \ \pi_2 = rac{p^2}{1+p+p^2+p^3+p^4} \ \pi_3 = rac{p^3}{1+p+p^2+p^3+p^4} \ \pi_4 = rac{p^4}{1+p+p^2+p^3+p^4}. \end{array}
ight.$$

b) Since the chain is irreducible and aperiodic with finite state space, its limiting distribution coincides with its stationary distribution.

Exercise 7.17

a) The transition matrix \boldsymbol{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

$$\pi P = [\pi_A, \pi_B, \pi_C, \pi_D] imes egin{bmatrix} 0 & 1/2 & 0 & 1/2 \ 1/3 & 0 & 1/3 & 1/3 \ 0 & 1 & 0 & 0 \ 1/2 & 1/2 & 0 & 0 \ \end{bmatrix} = egin{bmatrix} rac{1}{3} \pi_B + rac{1}{2} \pi_D \ rac{1}{2} \pi_A + \pi_C + rac{1}{2} \pi_D \ rac{1}{3} \pi_B \ rac{1}{3} \pi_B \ rac{1}{2} \pi_A + rac{1}{3} \pi_B \ rac{1}{2} \pi_A \ rac{1}{2} \pi_A + rac{1}{3} \pi_B \ rac{1}{2} \pi_A \ rac{1}{2} \pi_A + rac{1}{3} \pi_A \ rac{1}{2} \pi_A \ rac{1$$

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

$$\begin{cases} \mu_D(A) = \frac{1}{2} + \frac{1}{2} \left(1 + \mu_D(B) \right) = 1 + \frac{1}{2} \mu_D(B) \\ \mu_D(B) = \frac{1}{3} + \frac{1}{3} \left(1 + \mu_D(A) \right) + \frac{1}{3} \left(1 + \mu_D(C) \right) = 1 + \frac{1}{3} (\mu_D(A) + \mu_D(C)) \\ \mu_D(C) = 1 + \mu_D(B) \\ \mu_D(D) = \frac{1}{2} \left(1 + \mu_D(A) \right) + \frac{1}{2} \left(1 + \mu_D(B) \right) = 1 + \frac{1}{2} (\mu_D(A) + \mu_D(B)), \end{cases}$$

which has for solution

$$\mu_D(A) = rac{8}{3}, \ \ \mu_D(B) = rac{10}{3}, \ \ \mu_D(C) = rac{13}{3}, \ \ \mu_D(D) = 4.$$

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 7.18 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

$$P^2 = egin{bmatrix} 1 & 0 & 0 \ 0 & 1 & 0 \ a(1+c) & b(1+c) & c^2 \end{bmatrix}$$
 $P^3 = egin{bmatrix} 1 & 0 & 0 \ a(1+c+c^2) & b(1+c+c^2) & c^3 \end{bmatrix}$

and

we infer that P^n takes the general form

$$P^n = \left[egin{array}{cccc} 1 & 0 & 0 \ 0 & 1 & 0 \ a_n & b_n & c_n \end{array}
ight],$$

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

$$P^{n+1} = egin{bmatrix} 1 & 0 & 0 \ 0 & 1 & 0 \ a_{n+1} & b_{n+1} & c_{n+1} \end{bmatrix} = egin{bmatrix} 1 & 0 & 0 \ 0 & 1 & 0 \ a & b & c \end{bmatrix} imes egin{bmatrix} 1 & 0 & 0 \ 0 & 1 & 0 \ a_n & b_n & c_n \end{bmatrix}$$

shows that we have the recurrence relations

$$\left\{egin{array}{ll} a_{n+1}=a+ca_n,\ b_{n+1}=b+cb_n,\ c_{n+1}=c imes c_n,\ \end{array}
ight. egin{array}{ll} a_n=a+ac+\cdots+ac^{n-1}=arac{1-c^n}{1-c},\ b_n=b+bc+\cdots+bc^{n-1}=brac{1-c^n}{1-c},\ c_{n+1}=c imes c_n,\ \end{array}
ight.$$

hence

$$P^n = egin{bmatrix} 1 & 0 & 0 \ 0 & 1 & 0 \ a rac{1-c^n}{1-c} & b rac{1-c^n}{1-c} & c^n \end{bmatrix}.$$

b) From the structure of \mathbf{P}^n it follows that the chain admits a limiting distribution

$$\lim_{n o \infty} P^n = egin{bmatrix} 1 & 0 & 0 \ 0 & 1 & 0 \ a & b \ 1-c & 0 \end{bmatrix}$$

٠

which is dependent of the initial state, provided that c < 1. The limiting probabilities

$$\lim_{n o\infty} \mathbb{P}(X_n=0\mid X_0=2)=rac{a}{1-c},$$

resp.

$$\lim_{n o\infty} \mathbb{P}(X_n=1 \mid X_0=2) = rac{b}{1-c},$$

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. Here again, Theorem 7.8 does not apply for the same reasons as in Question (b) and the limiting and stationary distributions may differ.

Exercise 7.19

a) The process $(X_n)_{n \in \mathbb{N}}$ is a two-state Markov chain on $\{0, 1\}$ with transition matrix

$$\left[egin{array}{cc} lpha & eta \ p & q \end{array}
ight]$$

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 $\pi_1 = \frac{\beta}{p+\beta}$.

Exercise 7.20

a) Since the chain is irreducible and aperiodic with finite state space it admits a unique stationary distribution $\pi = (\pi_1, \pi_2, \dots, \pi_N)$ with

$$\pi_i=rac{1}{\mu_i(i)}, \qquad i=1,2,\dots,N,$$

by Corollary 7.9. Since $\pi = (\pi_1, \pi_2, \ldots, \pi_N)$ is a probability distribution we have $\pi_1 + \pi_2 + \cdots + \pi_N = 1$ hence there exists $i \in \{1, 2, \ldots, N\}$ with $\pi_i = 1/\mu_i(i) \ge 1/N$, *i.e.* $\mu_i(i) \le N$, otherwise the sum $\pi_1 + \pi_2 + \cdots + \pi_N$ would be lower than Δpne_{20}

b) Similarly, the condition $\pi_1 + \pi_2 + \cdots + \pi_N = 1$ shows that there must exist $i \in \{1, 2, \ldots, N\}$ with $\pi_i = 1/\mu_i(i) \leq 1/N$, $begin{subarray}{l} begin{subarray}{l} begin{subaray}$

Exercise 7.21

a) The $N \times N$ transition matrix of the chain is

	$\left[egin{array}{c} q \\ q \end{array} ight]$	$egin{array}{c} p \ 0 \end{array}$	$0 \ p$	•••	•••	0 0	0 0	0 0	
	0	${m q}$	0	•.	•.	0	0	0	
P =	0	0	${m q}$	•.	•.	0	0	0	.
	:	:	÷	•.	•.	÷	÷	÷	
	0	0	0	••.	••.	0	p	0	
	0	0	0	•••	• • •	${oldsymbol{q}}$	0	$oldsymbol{p}$	
	0	0	0	• • •	• • •	0	${m q}$	\boldsymbol{p}	

- 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

$$\pi_2 = \frac{p}{q} \pi_1$$
 and $\pi_N = \frac{p}{q} \pi_{N-1}, \qquad k = 1, 2, \dots, N-1,$

and

$$p(\pi_k - \pi_{k-1}) = q(\pi_{k+1} - \pi_k)$$
 $k = 2, 3, \dots, N-1.$

We check that π_k given by

$$\pi_k = rac{p^{k-1}}{q^{k-1}} \pi_1, \qquad k=1,2,\ldots,N,$$
 MH3512 AY19-2

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

$$\pi_k = rac{1-p/q}{1-(p/q)^N} rac{p^{k-1}}{q^{k-1}}, \qquad k=1,2,\ldots,N,$$

provided that 0 . When <math>p = q = 1/2 we find that the uniform distribution

$$\pi_k=rac{1}{N}, \qquad k=1,2,\ldots,N,$$

is stationary. When p = 0 the stationary distribution is $\mathbb{1}_{\{0\}} = [1, 0, \dots, 0, 0]$, and when p = 1 it is $\mathbb{1}_{\{N\}} = [0, 0, \dots, 0, 1]$.

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. It can be easily checked that this coincidence also occurs here for p = 0 and p = 1, although in those cases the chain is not irreducible and not aperiodic.

Exercise 7.22 We note that the chain is irreducible and has period 2.

a) By Problem 6.9 the chain is positive recurrent when $\alpha < 1$, hence by Theorem 7.10 it admits a stationary distribution π . The relation $\pi = \pi P$ reads

$$\pi = [\pi_0, \pi_1, \pi_2, \ldots] = [\pi_0, \pi_1, \pi_2, \ldots] imes egin{pmatrix} 0 & 1 & 0 & 0 & \cdots \ rac{1}{lpha+1} & 0 & rac{lpha}{lpha+1} & 0 & \cdots \ 0 & rac{1}{lpha+1} & 0 & rac{lpha}{lpha+1} & \cdots \ rac{1}{lpha+1} & 0 & rac{lpha}{lpha+1} & \cdots \ rac{1}{arepsilon} & rac{1}{lpha+1} & rac{1}{lpha+1} & rac{1}{lpha+1} & \cdots \ rac{1}{arepsilon} & rac{1}{arepsilon+1} & \cdots \ rac{1}{arepsilon} & rac{1}{arepsilon} & arepsilon} & rac{1}{arepsilon} & rac{1}{arepsilon} & rac{1}{arepsilon} & rac{1}{a$$

hence

$$\pi_0 = \frac{\pi_1}{\alpha + 1}, \qquad \pi_1 = \pi_0 + \frac{\pi_2}{\alpha + 1}, \qquad \pi_2 = \frac{\alpha \pi_1}{\alpha + 1} + \frac{\pi_3}{\alpha + 1}, \dots$$

i.e.

$$\pi_0=rac{\pi_2}{lpha(lpha+1)}, \qquad \pi_1=rac{\pi_2}{lpha}, \qquad \pi_2=rac{\pi_3}{lpha}, \dots$$
 MH3512 AY19-

from which we can deduce that

$$\pi_k=\pi_0lpha^{k-1}(1+lpha),\qquad k\geqslant 1.$$

From the condition $\sum_{k \ge 0} \pi_k = 1$ we find

$$1 = \pi_0 + \pi_0(1 + \alpha) \sum_{k \geqslant 1} \alpha^{k-1} = \pi_0 + \pi_0 rac{1 + lpha}{1 - lpha},$$

hence

$$\pi_0=rac{1-lpha}{2} \hspace{0.1 in} ext{and} \hspace{0.1 in} \pi_k=lpha^{k-1}rac{1-lpha^2}{2}, \qquad k\geqslant 1.$$

More generally, from the relation $\pi = \pi P$ we can also write

$$\pi_k = rac{lpha \pi_{k-1}}{lpha + 1} + rac{\pi_{k+1}}{lpha + 1}, \qquad k \geqslant 2,$$

which can be written as

$$\pi_k = q\pi_{k-1} + p\pi_{k+1}, \qquad k \ge 2, \tag{B.48}$$

where $p = 1/(\alpha + 1)$ and $q = \alpha/(\alpha + 1)$. We look for a solution of (B.48) of the form $\pi_k = C(q/p)^k = C\alpha^k$ as in (2.2.14), $k \ge 1$, which clearly satisfies (B.48) written as

$$rac{lpha lpha^{k-1}}{lpha+1}+rac{lpha^{k+1}}{lpha+1}=lpha^k, \qquad k\geqslant 2.$$

When $\alpha < 1$ the value of *C* can be found from the condition

$$1 = \sum_{k \ge 0} \pi_k$$

= $\pi_0 + \sum_{k \ge 2} \pi_k$
= $\frac{\pi_2}{\alpha(\alpha + 1)} + \sum_{k \ge 1} \pi_k$
= $C \frac{\alpha}{\alpha + 1} + C \frac{\alpha}{1 - \alpha}$

$$= C rac{2lpha}{1-lpha^2}$$

hence

$$C = rac{1-lpha^2}{2lpha}$$

and

$$\pi_0=rac{1-lpha}{2}, \hspace{1em} \pi_k=lpha^{k-1}rac{1-lpha^2}{2}, \hspace{1em} k\geqslant 1.$$

- b) When $\alpha > 1$ the series $\sum_{k \ge 1} \pi_k$ does not converge so there cannot be a stationary distribution. In addition, π_0 would be negative if $\alpha > 1$. When $\alpha = 1$ we find $\pi_k = 0$, $k \in \mathbb{N}$, which cannot be a probability distribution as well (note that from Question (a) the solution of $\pi = \pi P$ is unique).
- c) From the solution of Problem 6.9, we check that

$$\mathbb{E}[T_0 \mid X_0=0]=rac{1}{1-lpha}=rac{1}{\pi_0},$$

and

$$\mathbb{E}[T_k \mid X_0 = k] = rac{2lpha^{k-1}}{1-lpha^2} = rac{1}{\pi_k} < \infty, \qquad k \geqslant 1,$$

as expected from Theorem 7.10.

Exercise 7.23

a) We have

$$egin{array}{rl} \mathbb{P}(X_n=x)&=&\mathbb{P}(\{X_n=x\}\cap\{ au\leqslant n\})+\mathbb{P}(\{X_n=x\}\cap\{ au>n\})\ &=&\mathbb{P}(\{Y_n=x\}\cap\{ au\leqslant n\})+\mathbb{P}(\{X_n=x\}\cap\{ au>n\})\ &\leqslant&\mathbb{P}(Y_n=x)+\mathbb{P}(au>n). \end{array}$$

b) Similarly to Question (a), we have

$$egin{array}{rl} \mathbb{P}(Y_n=x) &=& \mathbb{P}(\{Y_n=x\}\cap\{ au\leqslant n\})+\mathbb{P}(\{Y_n=x\}\cap\{ au>n\}) \ &=& \mathbb{P}(\{X_n=x\}\cap\{ au\leqslant n\})+\mathbb{P}(\{Y_n=x\}\cap\{ au>n\}) \end{array}$$
$$\leqslant \ \mathbb{P}(X_n=x)+\mathbb{P}(au>n),$$

hence

$$-\mathbb{P}(au>n)\leqslant \mathbb{P}(X_n=x)-\mathbb{P}(Y_n=x)\leqslant \mathbb{P}(au>n),$$

which leads to the conclusion.

Problem 7.24

a) We have

$$\mathbb{P}(Y_{n+1} = j \mid Y_n = i) = \frac{\mathbb{P}(Y_{n+1} = j \text{ and } Y_n = i)}{\mathbb{P}(Y_n = i)} \\
= \frac{\mathbb{P}(Y_{n-1} = j)}{\mathbb{P}(Y_n = i)} \frac{\mathbb{P}(Y_n = i \text{ and } Y_{n+1} = j)}{\mathbb{P}(Y_{n-1} = j)} \\
= \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i)} \frac{\mathbb{P}(X_{N-n} = i \text{ and } X_{N-n-1} = j)}{\mathbb{P}(X_{N-n-1} = j)} \\
= \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i)} \frac{\mathbb{P}(X_{N-n} = i \text{ and } X_{N-n-1} = j)}{\mathbb{P}(X_{N-n-1} = j)} \\
= \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n-1} = j)} \mathbb{P}(X_{N-n} = i \mid X_{N-n-1} = j) = \frac{\pi_j}{\pi_i} P_{j,i}.$$

On the other hand, we have

$$\begin{split} \mathbb{P}(Y_{n+1} = j \mid Y_n = i_n, \dots, Y_0 = i_0) &= \frac{\mathbb{P}(Y_{n+1} = j, Y_n = i_n, \dots, Y_0 = i_0)}{\mathbb{P}(Y_n = i_n, \dots, Y_0 = i_0)} \\ &= \frac{\mathbb{P}(X_{N-n-1} = j, X_{N-n} = i_n, \dots, X_N = i_0)}{\mathbb{P}(X_{N-n} = i_n, \dots, X_N = i_0)} \\ &= \mathbb{P}(X_{N-n-1} = j \text{ and } X_{N-n} = i_n) \\ &\times \frac{\mathbb{P}(X_{N-n+1} = i_{n-1}, \dots, X_N = i_0 \mid X_{N-n-1} = j, X_{N-n} = i_n)}{\mathbb{P}(X_{N-n+1} = i_{n-1}, \dots, X_N = i_0)} \end{split}$$

ॳ9

$$= \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i_n)} \frac{\mathbb{P}(X_{N-n-1} = j \text{ and } X_{N-n} = i_n)}{\mathbb{P}(X_{N-n-1} = j)}$$

$$\times \frac{\mathbb{P}(X_{N-n+1} = i_{n-1}, \dots, X_N = i_0 \mid X_{N-n-1} = j, X_{N-n} = i_n)}{\mathbb{P}(X_{N-n+1} = i_{n-1}, \dots, X_N = i_0 \mid X_{N-n} = i_n)}$$

$$= \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i_n)} \mathbb{P}(X_{N-n} = i_n \mid X_{N-n-1} = j,)$$

$$\times \frac{\mathbb{P}(X_{N-n+1} = i_{n-1}, \dots, X_N = i_0 \mid X_{N-n-1} = j)}{\mathbb{P}(X_{N-n+1} = i_{n-1}, \dots, X_N = i_0 \mid X_{N-n} = i_n)}$$

$$= \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i_n)} \mathbb{P}(X_{N-n} = i_n \mid X_{N-n-1} = j) = \frac{\pi_j}{\pi_{i_n}} P_{j,i_n},$$

and this shows that

$$\mathbb{P}(Y_{n+1}=j \mid Y_n=i_n,\ldots,Y_0=i_0) = \mathbb{P}(Y_{n+1}=j \mid Y_n=i_n) = rac{\pi_j}{\pi_{i_n}} P_{j,i_n},$$

i.e. the time-reversed process $(Y_n)_{n=0,1,\dots,N}$ has the Markov property.

b) We find
$$P_{i,j} = \frac{\pi_i}{\pi_j} P_{j,i},$$
(B.49)

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

$$\pi_j = \sum\limits_i \pi_j P_{j,i} = \sum\limits_i \pi_i P_{i,j} = [\pi P]_j.$$

d) According to the detailed balance condition (B.49), we have

$$egin{aligned} P_{k_1,k_2}P_{k_2,k_3}\cdots P_{k_n,k_1}&=&P_{k_n,k_1}\prod\limits_{i=1}^{n-1}P_{k_i,k_{i+1}}=P_{k_n,k_1}\prod\limits_{i=1}^{n-1}rac{\pi_{k_{i+1}}}{\pi_{k_i}}P_{k_{i+1},k_i}\ &=&rac{\pi_{k_n}}{\pi_{k_1}}P_{k_n,k_1}\prod\limits_{i=1}^{n-1}P_{k_{i+1},k_i}=P_{k_1,k_n}\prod\limits_{i=1}^{n-1}P_{k_{i+1},k_i}, \end{aligned}$$

MH3512 AY19-20

() 650

holds for all sequences $\{k_1, k_2, \ldots, k_n\}$ of states and $n \ge 2$.

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

$$[P^{n-1}]_{i,j} = \sum_{k_2,\dots,k_{n-1}} P_{i,k_2} P_{k_2,k_3} \cdots P_{k_{n-1},j_2}$$

we get

$$egin{array}{rll} [m{P}^{n-1}]_{k_1,k_n} P_{k_n,k_1} &=& P_{k_n,k_1} \sum\limits_{k_2,\ldots,k_{n-1}} P_{k_1,k_2} P_{k_2,k_3} \cdots P_{k_{n-1},k_n} \ &=& P_{k_1,k_n} \sum\limits_{k_2,\ldots,k_{n-1}} P_{k_n,k_{n-1}} \cdots P_{k_3,k_2} P_{k_2,k_1} \ &=& P_{k_1,k_n} [m{P}^{n-1}]_{k_n,k_1}. \end{array}$$

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

$$\pi_i P_{i,i+1} = \pi_i \left(rac{1}{2} - rac{i}{2M}
ight) = \pi_{i+1} P_{i+1,i} = \pi_{i+1} rac{i+1}{2M},$$

hence

$$rac{\pi_{i+1}}{\pi_i} = rac{1-i/M}{(i+1)/M} = rac{M-i}{i+1},$$
 MH3512 AY19-20

which shows that

$$\pi_i = rac{(M-i+1)}{i} rac{(M-i+2)}{(i-1)} \cdots rac{(M-1)}{2} rac{M}{1} \pi_0 = rac{M!}{i!(M-i)!} \pi_0 = \pi_0 inom{M}{i},$$

 $i=0,1,\ldots,M,$ where the constant $\pi_0>0$ is given by

$$1 = \sum\limits_i \pi_i = \pi_0 \sum\limits_{i=0}^M inom{M}{i} = \pi_0 2^M,$$

hence $\pi_0 = 2^{-M}$ and

$$\pi_i = rac{1}{2^M} inom{M}{i}, \qquad i=0,1,\ldots,M.$$

g) We have

$$\begin{split} &[\pi P]_{i} = P_{i+1,i}\pi_{i+1} + P_{i,i}\pi_{i} + P_{i-1,i}\pi_{i-1} \\ &= \frac{1}{2^{M}}\frac{i+1}{2M}\binom{M}{i+1} + \frac{1}{2^{M}}\left(\frac{1}{2} - \frac{i-1}{2M}\right)\binom{M}{i-1} + \frac{1}{2^{M}} \times \frac{1}{2}\binom{M}{i} \\ &= \frac{1}{2^{M}}\frac{i+1}{2M}\frac{M!}{(i+1)!(M-i-1)!} + \frac{1}{2^{M}}\frac{M-i+1}{2M}\frac{M!}{(i-1)!(M-i+1)!} \\ &+ \frac{1}{2^{M} \times 2}\frac{M!}{i!(M-i)!} \\ &= \frac{1}{2}\left(\frac{1}{2^{M}}\frac{(M-1)!}{i!(M-i-1)!} + \frac{1}{2^{M}}\frac{(M-1)!}{(i-1)!(M-i)!}\right) + \frac{1}{2^{M+1}}\binom{M}{i} \\ &= \frac{1}{2^{M+1}}\binom{M-1}{i} + \frac{1}{2^{M+1}}\binom{M-1}{i-1} + \frac{1}{2^{M+1}}\binom{M}{i} \\ &= \frac{1}{2^{M}}\binom{M}{i}, \end{split}$$

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 .

 $\circ 652$

$$egin{aligned} &\lim_{n o\infty} \mathbb{E}[X_n \mid X_0=i] = \lim_{n o\infty} \sum\limits_{j=0}^M j \mathbb{P}(X_n=j \mid X_0=i) \ &= \sum\limits_{j=0}^M j \lim_{n o\infty} \mathbb{P}(X_n=j \mid X_0=i) = \sum\limits_{j=0}^M j \pi_j \ &= rac{1}{2^M} \sum\limits_{j=0}^M j inom{M}{j} = rac{1}{2^M} \sum\limits_{j=0}^M j rac{M!}{j!(M-j)!} \ &= rac{M}{2^M} \sum\limits_{j=1}^M rac{(M-1)!}{(j-1)!(M-j)!} = rac{M}{2^M} \sum\limits_{j=0}^{M-1} rac{(M-1)!}{j!(M-1-j)!} = rac{M}{2}, \end{aligned}$$

independently of $i = 0, 1, \ldots, M$.

j) Clearly, the relation

$$\mathbb{E}\left[X_0 - rac{M}{2} \ \Big| \ X_0 = i
ight] = \left(i - rac{M}{2}
ight)$$

holds when n = 0. Next, assuming that the relation holds at the rank $n \ge 0$ we have

$$\begin{split} h(i) &= \mathbb{E} \left[X_{n+1} - \frac{M}{2} \mid X_0 = i \right] \\ &= P_{i,i+1} \mathbb{E} \left[X_{n+1} - \frac{M}{2} \mid X_1 = i + 1 \right] + P_{i,i} \mathbb{E} \left[X_{n+1} - \frac{M}{2} \mid X_1 = i \right] \\ &+ P_{i,i-1} \mathbb{E} \left[X_{n+1} - \frac{M}{2} \mid X_1 = i - 1 \right] \\ &= \left(\frac{1}{2} - \frac{i}{2M} \right) \mathbb{E} \left[X_{n+1} - \frac{M}{2} \mid X_1 = i + 1 \right] + \frac{1}{2} \mathbb{E} \left[X_{n+1} - \frac{M}{2} \mid X_1 = i \right] \\ &+ \frac{i}{2M} \mathbb{E} \left[X_{n+1} - \frac{M}{2} \mid X_1 = i - 1 \right] \\ &= \left(\frac{1}{2} - \frac{i}{2M} \right) \mathbb{E} \left[X_n - \frac{M}{2} \mid X_0 = i + 1 \right] + \frac{1}{2} \mathbb{E} \left[X_n - \frac{M}{2} \mid X_0 = i \right] \end{split}$$

$$egin{aligned} &+rac{i}{2M}\,\mathbb{E}\left[X_n-rac{M}{2}\,\Big|\,X_0=i-1
ight]\ &=\left(rac{1}{2}-rac{i}{2M}
ight)\left(i+1-rac{M}{2}
ight)\left(1-rac{1}{M}
ight)^n+rac{1}{2}\left(i-rac{M}{2}
ight)\left(1-rac{1}{M}
ight)^n\ &+rac{i}{2M}\left(i-1-rac{M}{2}
ight)\left(1-rac{1}{M}
ight)^n=\left(i-rac{M}{2}
ight)\left(1-rac{1}{M}
ight)^{n+1},\qquad n\geqslant 0, \end{aligned}$$

for all i = 0, 1, ..., M.

Taking the limit as \boldsymbol{n} goes to infinity we get

$$\lim_{n o \infty} \mathbb{E} \left[X_n - rac{M}{2} \ \Big| \ X_0 = i
ight] = \lim_{n o \infty} \left(i - rac{M}{2}
ight) \left(1 - rac{1}{M}
ight)^n = 0,$$

hence

$$\lim_{n o\infty} \mathbb{E}[X_n] = rac{M}{2},$$

for all $i = 0, 1, \ldots, M$, which recovers the result of Question (i).

Problem 7.25

a) The ranking table is completed as follows:

\preceq	a	b	с	d	е
a	Х	\neq	⊭	lΥ	١
b	\neq	Х	≠	∡	\neq
с	$\not\perp$	ΥI	Х	⊉	∡
d	$\not\perp$	$\not\perp$	≠	Х	\neq
е	$\not\perp$	∡	≠	\leq	Х

 $\bigcirc 654$

b) The state space of the chain $(X_n)_{n \in \mathbb{N}}$ is (a, b, c, d, e) and its transition matrix is

$$P = \begin{bmatrix} 3/5 & 0 & 0 & 1/5 & 1/5 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1/5 & 4/5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/5 & 4/5 \end{bmatrix}$$

c) The chain is clearly reducible:



d) Starting from state (a), (d) or (e), the limiting distribution is (0, 0, 0, 1, 0), starting from state (b) or (c), the limiting distribution is (0, 1, 0, 0, 0), so that although the chain admits limiting distributions, it does <u>not</u> admit a limiting distribution independent of the initial state.

More precisely it can be checked that the power P^n of order $n \ge 1$ of the transition matrix P takes the form

$$P^n = egin{bmatrix} (3/5)^n & 0 & 0 & 1-(4/5)^n & a_n \ 0 & 1 & 0 & 0 & 0 \ 0 & 1-(4/5)^n & (4/5)^n & 0 & 0 \ 0 & 0 & 0 & 1 & 0 \ 0 & 0 & 0 & 1-(4/5)^n & (4/5)^n \end{bmatrix}$$

where $a_n = (4/5)^n - (3/5)^n$ since the sum of probabilities over the rows of P^n is equal to 1, hence

$$\lim_{n \to \infty} P^n = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

e) The equation

 $\pi=\pi P$

reads

$$\begin{aligned} \pi &= \left[\pi_a, \pi_b, \pi_c, \pi_d, \pi_e \right] \\ &= \left. \pi \begin{bmatrix} 3/5 & 0 & 0 & 1/5 & 1/5 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1/5 & 4/5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/5 & 4/5 \end{bmatrix} \\ &= \left[3\pi_a/5, \pi_b + \pi_c/5, 4\pi_c/5, \pi_a/5 + \pi_d + \pi_e/5, \pi_a/5 + 4\pi_e/5 \right], \end{aligned}$$

i.e.

$$[0, 0, 0, 0, 0] = [-2\pi_a/5, \pi_c/5, -\pi_c/5, \pi_a/5 + \pi_e/5, \pi_a/5 - \pi_e/5],$$

or

 $\left\{egin{array}{l} \pi_a=0\ \pi_c=0\ \pi_e=0. \end{array}
ight.$

As a consequence, any probability distribution of the form

$$\pi = [\pi_a, \pi_b, \pi_c, \pi_d, \pi_e] = [0, p, 0, 1 - p, 0],$$

with $p \in [0, 1]$, will be a stationary distribution for the chain with matrix P. The stationary distribution is not unique here because the chain is reducible. MH3512 AY19-20

- f) All rows in the matrix \tilde{P} clearly add up to 1, so \tilde{P} is a Markov transition matrix. On the other hand, all states become accessible from each other so that the new chain is irreducible and all states have period 1.
- g) Since the chain is irreducible, aperiodic and has a finite state space, we know by Corollary 7.9 that it admits a unique stationary distribution $\tilde{\pi}$. The equation $\tilde{\pi} = \tilde{\pi}\tilde{P}$ reads

h) The equation

$$ilde{\pi} = \left[rac{arepsilon}{5}, rac{arepsilon}{5}, rac{arepsilon}{5}, rac{arepsilon}{5}
ight] + (1-arepsilon) ilde{\pi}P$$

reads

i.e.

$$\begin{split} [\pi_a, \pi_b, \pi_c, \pi_d, \pi_e] &= \left[\frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}\right] + (1-\varepsilon)\tilde{\pi} \begin{bmatrix} 3/5 & 0 & 0 & 1/5 & 1/5 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1/5 & 4/5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/5 & 4/5 \end{bmatrix} \\ &= \left[\frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}\right] \\ &+ (1-\varepsilon) \left[3\pi_a/5, \pi_b + \pi_c/5, 4\pi_c/5, \pi_a/5 + \pi_d + \pi_e/5, \pi_a/5 + 4\pi_e/5 \right], \end{split}$$

$$egin{aligned} & [0,0,0,0,0]\ &= & [arepsilon+3(1-arepsilon)\pi_a-5\pi_a,arepsilon+5(1-arepsilon)\pi_b-5\pi_b+(1-arepsilon)\pi_c,arepsilon+4(1-arepsilon)\pi_c-5\pi_c,\ &arepsilon+(1-arepsilon)\pi_a+5(1-arepsilon)\pi_d-5\pi_d+(1-arepsilon)\pi_e,arepsilon+(1-arepsilon)\pi_a+4(1-arepsilon)\pi_e-5\pi_e]\,, \end{aligned}$$

658

$$\left\{egin{array}{l}arepsilon - (2+3arepsilon)\pi_a = 0\ arepsilon - 5arepsilon\pi_b + (1-arepsilon)\pi_c = 0\ arepsilon - \pi_c(1+4arepsilon) = 0\ arepsilon + (1-arepsilon)\pi_a - 5arepsilon\pi_d + (1-arepsilon)\pi_e = 0\ arepsilon + (1-arepsilon)\pi_a - (1+4arepsilon)\pi_e = 0, \end{array}
ight.$$

i.e.

i.e.

$$\left\{ egin{array}{l} \pi_a = rac{arepsilon}{2+3arepsilon}, \ \pi_b = rac{2+3arepsilon}{5(1+4arepsilon)}, \ \pi_c = rac{arepsilon}{1+4arepsilon}, \ \pi_d = rac{3+2arepsilon}{5(1+4arepsilon)}, \ \pi_e = rac{arepsilon(3+2arepsilon)}{(1+4arepsilon)(2+3arepsilon)}. \end{array}
ight.$$

We also check that

$$\begin{aligned} \pi_a + \pi_b + \pi_c + \pi_d + \pi_e \\ &= \frac{\varepsilon}{2+3\varepsilon} + \frac{2+3\varepsilon}{5(1+4\varepsilon)} + \frac{\varepsilon}{1+4\varepsilon} + \frac{3+2\varepsilon}{5(1+4\varepsilon)} + \frac{\varepsilon(3+2\varepsilon)}{(1+4\varepsilon)(2+3\varepsilon)} \\ &= \frac{5\varepsilon(1+4\varepsilon)}{5(2+3\varepsilon)(1+4\varepsilon)} + \frac{(2+3\varepsilon)^2}{5(1+4\varepsilon)(2+3\varepsilon)} + \frac{5\varepsilon(2+3\varepsilon)}{5(1+4\varepsilon)(2+3\varepsilon)} \\ &+ \frac{(3+2\varepsilon)(2+3\varepsilon)}{5(2+3\varepsilon)(1+4\varepsilon)} + \frac{5\varepsilon(3+2\varepsilon)}{5(1+4\varepsilon)(2+3\varepsilon)} \end{aligned}$$

$$\circ 659$$

$$= \frac{5\varepsilon(1+4\varepsilon) + (2+3\varepsilon)(5+10\varepsilon) + 5\varepsilon(3+2\varepsilon)}{5(1+4\varepsilon)(2+3\varepsilon)}$$
$$= 1.$$

i) We note that

 $\pi_a < \pi_c < \pi_e < \pi_b < \pi_d,$

hence we will rank the states as

rank	state
1	d
2	b
3	e
4	с
5	a

based on the idea that the most visited states should rank higher. In the graph of Figure S.19 the stationary distribution is plotted as a function of $\varepsilon \in [0, 1]$.



Figure S.19: Stationary distribution as a function of $\varepsilon \in [0, 1]$.

$$egin{aligned} \mu_a(a) &= 3+rac{2}{arepsilon}\ \mu_b(b) &= rac{5(1+4arepsilon)}{2+3arepsilon}\ \mu_c(c) &= 4+rac{1}{arepsilon}\ \mu_d(d) &= rac{5(1+4arepsilon)}{3+2arepsilon}\ \mu_e(e) &= rac{(1+4arepsilon)(2+3arepsilon)}{arepsilon(3+2arepsilon)} \end{aligned}$$

In the graph of Figure S.20 the mean return times are plotted as a function of $\varepsilon \in [0, 1]$. A commonly used value in the literature is $\varepsilon = 1/7$.



Figure S.20: Mean return times as functions of $\varepsilon \in [0, 1]$.

We note that the ranking of states is clearer for smaller values of ε . In particular ε cannot be be chosen too large, for example taking $\varepsilon = 1$ makes all mean return times and equal and corresponds to a uniform stationary distribution. However the mean return times can be higher and hence the simulations can take longer for small values of ε .

a) For any two probability distributions $\mu = [\mu_1, \mu_2, \dots, \mu_N]$ and $\nu = [\nu_1, \nu_2, \dots, \nu_N]$ on $\{1, 2, \dots, N\}$ we have

ॳ 661

$$egin{array}{rll} \|\mu-
u\|_{ ext{TV}} &=& rac{1}{2}\sum\limits_{k=1}^{N}|\mu_k-
u_k| \ &\leqslant& rac{1}{2}\sum\limits_{k=1}^{N}(\mu_k+
u_k) \ &=& rac{1}{2}\sum\limits_{k=1}^{N}\mu_k+rac{1}{2}\sum\limits_{k=1}^{N}
u_k \ &=& 1. \end{array}$$

$$egin{aligned} &|\mu P -
u P \|_{\mathrm{TV}} \;=\; rac{1}{2} \sum\limits_{j=1}^{N} |[\mu P]_j - [
u P]_j| \ &=\; rac{1}{2} \sum\limits_{j=1}^{N} \left| \sum\limits_{i=1}^{n} \mu_i P_{i,j} - \sum\limits_{i=1}^{n}
u_i P_{i,j}
ight| \ &\leqslant\; rac{1}{2} \sum\limits_{j=1}^{N} \sum\limits_{i=1}^{n} P_{i,j} |\mu_i -
u_i| \ &=\; rac{1}{2} \sum\limits_{i=1}^{n} |\mu_i -
u_i| \sum\limits_{j=1}^{N} P_{i,j} \ &=\; rac{1}{2} \sum\limits_{i=1}^{n} |\mu_i -
u_i|. \end{aligned}$$

c) Replacing μ and ν with μP^n and π in the result of Question (b) we find

$$egin{array}{lll} \|\mu P^{n+1} - \pi\|_{ ext{TV}} &= \|(\mu P^n) P - \pi P\|_{ ext{TV}} \ &\leqslant \ \|\mu P^n - \pi\|_{ ext{TV}}. \end{array}$$

d) Letting $k \in \{1, 2, \dots, N\}$ and taking

$$\mu:=(0,\ldots,0, \stackrel{1}{\underset{k}{1}}, 0,\ldots,0)$$

we have $\mu P^{n+1} = [P^{n+1}]_{k,.}$ and by Question (c) we find

$$egin{aligned} \| [P^{n+1}]_{k,\cdot} - \pi \|_{ ext{TV}} &= & \| \mu P^{n+1} - \pi P \|_{ ext{TV}} \ &\leqslant & \| \mu P^n - \pi \|_{ ext{TV}} \ &= & \| [P^n]_{k,\cdot} - \pi \|_{ ext{TV}}. \end{aligned}$$

Taking the maximum over $k = 1, 2, \ldots, N$ in the above inequality yields

$$d(n+1) = \max_{k=1,2,...,N} \| [P^{n+1}]_{k,\cdot} - \pi \|_{ ext{TV}} \leqslant \max_{k=1,2,...,N} \| [P^n]_{k,\cdot} - \pi \|_{ ext{TV}} = d(n),$$

 $n \in \mathbb{N}$.

- e) The chain is irreducible because all states can communicate in one time step since $P_{i,j} > 0$, $1 \leq i, j \leq N$. In addition the chain is aperiodic as all states have period one, given that $P_{i,i} > 0$, i = 1, 2, ..., N. Since the state space is finite, Corollary 6.2 shows that all states are positive recurrent, hence by Theorem 7.2 the chain admits a limiting and a stationary distribution that are equal.
- f) We note that $\boldsymbol{Q}_{\boldsymbol{\theta}}$ can be written as

$$\begin{split} Q_{\theta} &= \left[\begin{array}{ccc} [Q_{\theta}]_{i,j} \end{array} \right]_{1 \leqslant i,j \leqslant N} \\ &= \left[\begin{array}{cccc} [Q_{\theta}]_{1,1} & [Q_{\theta}]_{1,2} & \cdots & [Q_{\theta}]_{1,N} \\ [Q_{\theta}]_{2,1} & [Q_{\theta}]_{2,2} & \cdots & [Q_{\theta}]_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ [Q_{\theta}]_{N,1} & [Q_{\theta}]_{N,2} & \cdots & [Q_{\theta}]_{N,N} \end{array} \right] \\ &= \left[\begin{array}{cccc} \frac{1}{1-\theta}(P_{1,1} - \theta\pi_{1}) & \frac{1}{1-\theta}(P_{1,2} - \theta\pi_{2}) & \cdots & \frac{1}{1-\theta}(P_{1,N} - \theta\pi_{N}) \\ \frac{1}{1-\theta}(P_{2,1} - \theta\pi_{1}) & \frac{1}{1-\theta}(P_{2,2} - \theta\pi_{2}) & \cdots & \frac{1}{1-\theta}(P_{2,N} - \theta\pi_{N}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{1-\theta}(P_{N,1} - \theta\pi_{1}) & \frac{1}{1-\theta}(P_{N,2} - \theta\pi_{2}) & \cdots & \frac{1}{1-\theta}(P_{N,N} - \theta\pi_{N}) \end{array} \right] \end{split}$$

♦ 662

Clearly, all entries of $oldsymbol{Q}_{oldsymbol{ heta}}$ are nonnegative due to the condition

$$P_{i,j} \geqslant heta \pi_j, \qquad i,j=1,2,\ldots,N.$$

In addition, for all $i=1,2,\ldots,N$ we have

$$egin{aligned} &\sum\limits_{j=1}^{N} [Q_{ heta}]_{i,j} \; = \; rac{1}{1- heta} \sum\limits_{j=1}^{N} (P_{i,j} - heta \Pi_{i,j}) \ &= \; rac{1}{1- heta} \sum\limits_{j=1}^{N} (P_{i,j} - heta \pi_{j}) \ &= \; rac{1}{1- heta} \sum\limits_{j=1}^{N} P_{i,j} - rac{ heta}{1- heta} \sum\limits_{j=1}^{N} \pi_{j} \ &= \; rac{1}{1- heta} - rac{ heta}{1- heta} \ &= \; 1, \qquad 0 < heta < 1, \end{aligned}$$

and we conclude that Q_{θ} is a Markov transition matrix.

g) Clearly, the property holds for n = 1 by the definition of Q_{θ} . Next, assume that

$$P^n = \Pi + (1- heta)^n \left(Q_{ heta}^n - \Pi\right)$$

for some $n \ge 1$. Noting that the condition $\pi P = \pi$ implies $\Pi P = \Pi$, we have

$$\begin{split} P^{n+1} &= \left(\Pi + (1-\theta)^n (Q_{\theta}^n - \Pi) \right) P \\ &= \Pi P + (1-\theta)^n Q_{\theta}^n P - (1-\theta)^n \Pi P \\ &= \Pi + (1-\theta)^n Q_{\theta}^n P - (1-\theta)^n \Pi \\ &= \Pi + (1-\theta)^n Q_{\theta}^n \big(\Pi + (1-\theta) (Q_{\theta} - \Pi) \big) - (1-\theta)^n \Pi \\ &= \Pi + \theta (1-\theta)^n Q_{\theta}^n \Pi + (1-\theta)^{n+1} Q_{\theta}^{n+1} - (1-\theta)^n \Pi. \end{split}$$

Next, we note that since Q_{θ} is a Markov transition matrix by Question (f) we have $Q_{\theta}\Pi = \Pi$, in other words we have $P\Pi = \Pi^2 = \Pi$, and

$$Q_ heta\Pi=rac{1}{1- heta}\left(P- heta\Pi
ight)\Pi=rac{1}{1- heta}ig(P\Pi- heta\Pi^2ig)=rac{1}{1- heta}\left(\Pi- heta\Piig)=\Pi,$$

and more generally $Q_{\theta}^{n}\Pi = \Pi$, $n \ge 1$, hence

$$\begin{split} P^{n+1} &= \Pi + \theta (1-\theta)^n Q_{\theta}^n \Pi + (1-\theta)^{n+1} Q_{\theta}^{n+1} - (1-\theta)^n \Pi \\ &= \Pi + \theta (1-\theta)^n \Pi + (1-\theta)^{n+1} Q_{\theta}^{n+1} - (1-\theta)^n \Pi \\ &= \Pi + (1-\theta)^{n+1} Q_{\theta}^{n+1} - (1-\theta)^{n+1} \Pi \\ &= \Pi + (1-\theta)^{n+1} (Q_{\theta}^{n+1} - \Pi). \end{split}$$

h) Let $k \in \{1, 2, \dots, N\}$. By Question (g) we have

$$egin{aligned} \|[P^n]_{k,\cdot} - \pi\|_{\mathrm{TV}} &= \|[P^n]_{k,\cdot} - \Pi_{k,\cdot}\|_{\mathrm{TV}} \ &= rac{1}{2}\sum\limits_{j=1}^N |[P^n]_{k,j} - \pi_j| \ &= rac{1}{2}\sum\limits_{j=1}^N |(1- heta)^n [Q^n_ heta]_{k,j} - (1- heta)^n \pi_j \ &= rac{(1- heta)^n}{2}\sum\limits_{j=1}^N |[Q^n_ heta]_{k,j} - \pi_j| \ &= (1- heta)^n \|[Q^n_ heta]_{k,\cdot} - \pi\|_{\mathrm{TV}} \ &\leqslant (1- heta)^n, \quad n \in \mathbb{N}, \end{aligned}$$

where we applied the result of Question (a), since $\Pi_{k,.} = \pi$ is a probability distribution and the same holds for $[Q_{\theta}^n]_{k,.}$ for all k = 1, 2, ..., N by Question (f).

The relation

$$\|[P^n]_{k,\cdot}-\pi\|_{\mathrm{TV}}=(1- heta)^n\|[Q^n_ heta]_{k,\cdot}-\pi\|_{\mathrm{TV}},\qquad n\in\mathbb{N},$$

also shows that, in total variation distance, at each time step the chain associated to P converges faster (by a factor $1 - \theta$) to π than the chain associated to Q_{θ} .

Finally, we find

$$d(n) = \max_{k=1,2,...,N} \| [P^n]_{k,\cdot} - \pi \|_{\mathrm{TV}} \leqslant (1- heta)^n, \qquad n \geqslant 0.$$

i) If $t_{\text{mix}} = 0$ the inequality is clearly satisfied, so that we can suppose that $t_{\text{mix}} \ge 1$. By the definition of t_{mix} and the result of

Question (h) we have

$$rac{1}{4} < d(t_{ ext{mix}}-1) \leqslant (1- heta)^{t_{ ext{mix}}-1},$$

hence

$$\lograc{1}{4} < \log d(t_{\mathrm{mix}}-1) \leqslant \log \left((1- heta)^{t_{\mathrm{mix}}-1}
ight) = (t_{\mathrm{mix}}-1)\log(1- heta),$$

and

$$t_{ ext{mix}} - 1 \leqslant rac{\log d(t_{ ext{mix}} - 1)}{\log(1 - heta)} < rac{\log 1/4}{\log(1 - heta)}$$

Hence we have

$$t_{\rm mix} < 1 + \frac{\log 1/4}{\log(1-\theta)},$$

which yields

$$t_{
m mix} < 1 + \left[rac{\log 1/4}{\log(1- heta)}
ight],$$

and finally

$$t_{ ext{mix}} \leqslant \left[rac{\log 1/4}{\log(1- heta)}
ight].$$

j) Given the transition matrix

$$P = \begin{bmatrix} 2/3 & 1/6 & 1/6 \\ 1/3 & 1/2 & 1/6 \\ 1/6 & 2/3 & 1/6 \end{bmatrix}$$

and its stationary distribution

$$\pi = [\pi_1, \pi_2, \pi_3] = [11/24, 9/24, 4/24]$$

we check that in order to satisfy all nine conditions $P_{i,j} \ge \theta \pi_j$, i, j = 1, 2, 3, the value of θ should be in the range [0, 4/11]. The optimal value of θ is the one that minimizes the bound $\left[\frac{\log 1/4}{\log(1-\theta)}\right]$, *i.e.* $\theta = 4/11$, and

$$t_{ ext{mix}} \leqslant \left[rac{\log 1/4}{\log 7/11}
ight] = \left\lceil 3.067
ight
ceil = 4.$$



Figure S.21: Graphs of distance to stationarity d(n) and upper bound $(1 - \theta)^n$.

We check from the above graph that the actual value of the mixing time is $t_{\text{mix}} = 2$. The value of d(0) is the maximum distance between π and all deterministic initial distributions starting from states k = 1, 2, ..., N.

Remark. We have shown that the conditions $\pi P = \pi$ and $P_{i,j} \ge \theta \pi_j$, i, j = 1, 2, ..., N, for some $\theta \in (0, 1)$, define a unique (stationary) distribution π which is also a limiting distribution independent of the initial state. This is the case in particular when $P_{i,j} > 0$, i, j = 1, 2, ..., N, in which case the chain is irreducible and aperiodic, and admits a unique limiting and stationary distribution. More generally, the result holds when P is *regular*, *i.e.* when there exists $n \ge 1$ such that $[P^n]_{i,j} > 0$ for all i, j = 1, 2, ..., N, cf. § 4.3-4.5 of [LPW09].

Below is the Matlab/Octave code used to generate Figure S.21.

P = [2/3,1/6,1/6; 1/3,1/2,1/6; 1/6,2/3,1/6;] pi = [11/24,9/24,4/25]theta = 4/11
for n = 1:11 y(n)=n-1; u(n)=0.25; $z(n)=(1-theta)^{(n-1)};$

MH3512 AY19-20

(5666)

```
distance(n) = 0;
for k = 1:3
d = mpower(P,n-1)(k,1:3) - pi;
dist=0;
for i = 1:3
dist = dist + 0.5 * abs(d(i));
end
distance(n) = max(distance(n), dist);
end
end
graphics_toolkit("gnuplot");
plot(y,distance,'-bo','LineWidth',8,y,z,'-ro','LineWidth',8,y,u,'-k',
      'LineWidth',8)
legend('d(n)', '(1-\lambda theta)^n')
set (gca, 'xtick', 1:10)
set (gca, 'ytick', 0:0.1:1)
grid on
xlabel('time steps n')
ylabel('distance')
pause
```

Problem 7.27

- a) The cardinality of \mathbb{S} is 2^N .
- b) We note that $z_k(z_{k-1} + z_{k+1})/2$ can only take the three possible values -1, 0, +1, and treat all cases separately.
- c) We have

$$egin{aligned} \mathbb{P}(Z_1 = z \mid Z_0 = z) &= 1 - \sum\limits_{k=1}^N \mathbb{P}(Z_1 = ar{z}^k \mid Z_0 = z) \ &= 1 - rac{1}{N} \sum\limits_{k=1}^N rac{1}{1 + (p/q)^{z_k(z_{k-1} + z_{k+1})/2}}, \qquad z \in \mathbb{S}. \end{aligned}$$

d) The transition probability matrix P of $(Z_n)_{n \in \mathbb{N}}$ is given by

The chain is irreducible because starting from any configuration $z = (z_k)_{1 \leq k \leq N} \in \mathbb{S}$ we can reach any other configuration $\hat{z} = (\hat{z}_k)_{1 \leq k \leq N} \in \mathbb{S}$ in a finite number of time steps. For this, count the number of spins in $z = (z_k)_{1 \leq k \leq N}$ that differ from the spins in $\hat{z} = (\hat{z}_k)_{1 \leq k \leq N}$ and flip them one by one until we reach $\hat{z} = (\hat{z}_k)_{1 \leq k \leq N}$.

Alternatively, we could also enumerate all possible 2^N configurations by flipping one spin at a time, starting from $z = (+1, +1, \ldots, + until we reach z = (-1, -1, \ldots, -1)$. When N = 3 the chain has the following graph:



From (7.3.4) we note the property

$$\mathbb{P}(Z_1 = \bar{z}^k \mid Z_0 = z) + \mathbb{P}(Z_1 = z \mid Z_0 = \bar{z}^k)$$

$$= \frac{1}{N(1 + (p/q)^{z_k(z_{k-1} + z_{k+1})/2})} + \frac{1}{N(1 + (p/q)^{-z_k(z_{k-1} + z_{k+1})/2})},$$

$$= \frac{1}{N}, \quad k = 1, 2, \dots, N.$$
MH3512 AY19-20

e) The chain is irreducible by Question (d), and it has a finite state space of cardinality 2^N by Question (a), hence it is positive recurrent by Corollary 6.13. In addition it is aperiodic since every state has a returning loop because

$$\mathbb{P}(Z_1=z\mid Z_0=z)\geqslant\min(p,q)>0, \qquad z\in\mathbb{S}.$$

Hence by e.g. Theorem 7.2 it admits a unique stationary distribution which coincides with its limiting distribution.

f) Under (7.3.6) we have

$$egin{aligned} \mathbb{P}(Z_1 = z) &= \mathbb{P}(Z_1 = z \mid Z_0 = z) \pi_z + \sum\limits_{k=1}^N \mathbb{P}(Z_1 = z \mid Z_0 = ar{z}^k) \pi_{ar{z}^k} \ &= \pi_z \mathbb{P}(Z_1 = z \mid Z_0 = z) + \pi_z \sum\limits_{k=1}^N \mathbb{P}(Z_1 = ar{z}^k \mid Z_0 = z) \ &= \pi_z \left(\mathbb{P}(Z_1 = z \mid Z_0 = z) + \sum\limits_{k=1}^N \mathbb{P}(Z_1 = ar{z}^k \mid Z_0 = z)
ight) \ &= \pi_z, \end{aligned}$$

hence $(\pi_z)_{z\in\mathbb{S}}$ is a stationary distribution.

g) By (7.3.4) we have

$$\begin{split} \frac{\mathbb{P}(Z_1 = \bar{z}^k \mid Z_0 = z)}{\mathbb{P}(Z_1 = z \mid Z_0 = \bar{z}^k)} &= \frac{1 + (p/q)^{\bar{z}^k(\bar{z}^k_{k-1} + \bar{z}^k_{k+1})/2}}{1 + (p/q)^{z_k(z_{k-1} + z_{k+1})/2}} \\ &= \frac{1 + (p/q)^{-z_k(z_{k-1} + z_{k+1})/2}}{1 + (p/q)^{z_k(z_{k-1} + z_{k+1})/2}} \\ &= \frac{q^{z_k(z_{k-1} + z_{k+1})/2}(1 + (q/p)^{z_k(z_{k-1} + z_{k+1})/2})}{q^{z_k(z_{k-1} + z_{k+1})/2} + p^{z_k(z_{k-1} + z_{k+1})/2}} \\ &= \frac{(q/p)^{z_k(z_{k-1} + z_{k+1})/2}(p^{z_k(z_{k-1} + z_{k+1})/2} + q^{z_k(z_{k-1} + z_{k+1})/2})}{q^{z_k(z_{k-1} + z_{k+1})/2} + p^{z_k(z_{k-1} + z_{k+1})/2}} \\ &= \left(\frac{q}{p}\right)^{z_k(z_{k-1} + z_{k+1})/2}, \quad k = 1, 2, \dots, N. \end{split}$$

h) For all $z \in S$ we have

$$egin{aligned} \pi_{ar{z}^k} &= \ C_eta \exp\left(eta \sum\limits_{l=0}^{k-2} z_l z_{l+1} - eta z_{k-1} z_k - eta z_k z_{k+1} + eta \sum\limits_{l=k+1}^N z_l z_{l+1}
ight) \ &= \ C_eta \exp\left(-2eta z_k (z_{k-1} + z_{k+1}) + eta \sum\limits_{l=0}^N z_l z_{l+1}
ight) \ &= \ \pi_z \mathrm{e}^{-2eta z_k (z_{k-1} + z_{k+1})} \ &= \ \left(rac{q}{p}
ight)^{z_k (z_{k-1} + z_{k+1})/2} \pi_z, \qquad k=1,2,\ldots,N, \end{aligned}$$

and the *inverse temperature* $\boldsymbol{\beta}$ is given by

 $eta = rac{1}{4}\lograc{p}{q},$ $p = rac{1}{1+e^{-4eta}}.$

The constant C_{β} is chosen so that

 $\sum\limits_{z\in\mathbb{S}}\pi_z=C_{eta}\sum\limits_{z\in\mathbb{S}}\exp\left(eta\sum\limits_{l=0}^Nz_lz_{l+1}
ight)=1,$

i.e.

i.e.

$$C_eta = \left(\sum\limits_{z \in \mathbb{S}} \exp\left(eta \sum\limits_{l=0}^N z_l z_{l+1}
ight)
ight)^{-1}.$$

The stationary distribution $(\pi_z)_{z\in\mathbb{S}}$ is known as the *Boltzmann distribution*.

i) We have

and from the relation

$$\pi_{---} + \pi_{--+} + \pi_{-+-} + \pi_{-++} + \pi_{+--} + \pi_{++-} + \pi_{++-} = 1,$$

we find

$$egin{aligned} C_{eta} &=& rac{1}{e^{4eta} + \mathrm{e}^{-4eta} + 6} \ &=& rac{1}{4\cosh^2(2eta) + 4} \ &=& rac{pq}{q^2 + p^2 + 6pq} \ &=& rac{1}{6 + p/q + q/p} \ &=& rac{pq}{1 + 4pq}. \end{aligned}$$

We note that when p > 1/2 the configuration "+++" has the highest probability p^2 , while "-+-" has the lowest probability q^2 in the long run, due to the presence of two "opinion leaders" $z_0 = +1$ and $z_4 = +1$ who will not change their minds.

We can also compute the probabilities of having more "+" than "-" in the long run, as

$$\pi_{-++}+\pi_{+-+}+\pi_{++-}+\pi_{+++}=rac{p(1+2q)}{1+4pq},$$

while the probability of having more "-" than "+" is

$$\pi_{---} + \pi_{--+} + \pi_{-+-} + \pi_{+--} = rac{q(1+2p)}{1+4pq}.$$



Figure S.22: Probability of a majority of "+" in the long run as a function of $p \in [0, 1]$.

Clearly, the end result is influenced by the boundary conditions $z_0 = z_4 = +1$.

For another example, taking $z_0 = -1$ and $z_4 = +1$, we have

$$\circ 674$$

where

$$C_eta = rac{1}{4\sqrt{p/q}+4\sqrt{q/p}} = rac{\sqrt{pq}}{4}.$$

The probabilities of having more "+" than "-" in the long run are

$$\pi_{-++} + \pi_{+-+} + \pi_{++-} + \pi_{+++} = rac{1}{2}$$

while the probability of having more "-" than "+" is also

$$\pi_{---} + \pi_{--+} + \pi_{-+-} + \pi_{+--} = \frac{1}{2}$$

The study undertaken in [BM18] shows that $\beta = 0.853$ in the Ising model for a Facebook friendship-network with N = 333 nodes and 2519 edges, in which spin values refer to gender, *i.e.*

$$p=rac{1}{1+e^{-4eta}}=0.968077467.$$

Chapter 8 - Branching Processes

Exercise 8.1

a) We have

$$egin{aligned} \mathbb{P}(X_2 > 0) &= 1 - \mathbb{P}(X_2 = 0) \ &= 1 - \left(rac{1}{5} + rac{3}{5} imes rac{1}{5} + rac{1}{5} imes rac{1}{5} imes rac{1}{5}
ight) \ &= 1 - rac{41}{125} = rac{84}{125}. \end{aligned}$$

b) We have

$$\mathbb{P}(X_2 = 1) = \mathbb{P}(X_2 = 1 \text{ and } X_1 = 0) + \mathbb{P}(X_2 = 1 \text{ and } X_1 = 1) \\
+ \mathbb{P}(X_2 = 1 \text{ and } X_1 = 2) \\
= \mathbb{P}(X_2 = 1 \mid X_1 = 0) \mathbb{P}(X_1 = 0) \\
+ \mathbb{P}(X_2 = 1 \mid X_1 = 1) \mathbb{P}(X_1 = 1) \\
+ \mathbb{P}(X_2 = 1 \mid X_1 = 2) \mathbb{P}(X_1 = 2) \\
= 0 \times \frac{1}{5} + \frac{3}{5} \times \frac{3}{5} + 2 \times \frac{1}{5} \times \frac{3}{5} \times \frac{1}{5} \\
= \frac{51}{125}.$$

We could also compute

$$\mathbb{P}(X_2 = 2) = \mathbb{P}(X_2 = 2 \text{ and } X_1 = 0) + \mathbb{P}(X_2 = 2 \text{ and } X_1 = 1) \\ + \mathbb{P}(X_2 = 2 \text{ and } X_1 = 2) \\ = \mathbb{P}(X_2 = 2 \mid X_1 = 0) \mathbb{P}(X_1 = 0) \\ + \mathbb{P}(X_2 = 2 \mid X_1 = 1) \mathbb{P}(X_1 = 1) \\ + \mathbb{P}(X_2 = 2 \mid X_1 = 2) \mathbb{P}(X_1 = 2)$$

ර 675

$$= 0 \times \frac{1}{5} + \frac{1}{5} \times \frac{3}{5} + \times \frac{1}{5} \times \left(\frac{3}{5}\right)^2 + 2 \times \frac{1}{5} \times \frac{1}{5} \times \frac{1}{5} \\ = \frac{26}{125},$$

and

$$\mathbb{P}(X_2 = 3) = \mathbb{P}(X_2 = 3 \text{ and } X_1 = 0) + \mathbb{P}(X_2 = 3 \text{ and } X_1 = 1) \\
+ \mathbb{P}(X_2 = 3 \text{ and } X_1 = 2) \\
= \mathbb{P}(X_2 = 3 \mid X_1 = 0) \mathbb{P}(X_1 = 0) \\
+ \mathbb{P}(X_2 = 3 \mid X_1 = 1) \mathbb{P}(X_1 = 1) \\
+ \mathbb{P}(X_2 = 3 \mid X_1 = 2) \mathbb{P}(X_1 = 2) \\
= 2 \times \frac{1}{5} \times \frac{1}{5} \times \frac{3}{5} \\
= \frac{6}{125},$$

and

$$\mathbb{P}(X_2 = 4) = \mathbb{P}(X_2 = 4 \text{ and } X_1 = 0) + \mathbb{P}(X_2 = 4 \text{ and } X_1 = 1) \\
+ \mathbb{P}(X_2 = 4 \text{ and } X_1 = 2) \\
= \mathbb{P}(X_2 = 4 \mid X_1 = 0) \mathbb{P}(X_1 = 0) \\
+ \mathbb{P}(X_2 = 4 \mid X_1 = 1) \mathbb{P}(X_1 = 1) \\
+ \mathbb{P}(X_2 = 4 \mid X_1 = 2) \mathbb{P}(X_1 = 2) \\
= \frac{1}{5} \times \frac{1}{5} \times \frac{1}{5} \\
= \frac{1}{125},$$

which recovers

$$\mathbb{P}(X_2 \ge 1) = rac{51 + 26 + 6 + 1}{125} = rac{84}{125}.$$

MH3512 AY19-20

ॳ 676

$$\mathbb{P}(X_1 = 2 \mid X_2 = 1) = \frac{\mathbb{P}(X_1 = 2 \text{ and } X_2 = 1)}{\mathbb{P}(X_2 = 1)}$$
$$= \mathbb{P}(X_2 = 1 \mid X_1 = 2) \frac{\mathbb{P}(X_1 = 2)}{\mathbb{P}(X_2 = 1)}$$
$$= 2 \times \frac{1}{5} \times \frac{3}{5} \times \frac{1}{5} \times \frac{125}{51} = \frac{2}{17}.$$

We can also compute

$$\mathbb{P}(X_1 = 1 \mid X_2 = 1) = \mathbb{P}(X_2 = 1 \mid X_1 = 1) \frac{\mathbb{P}(X_1 = 1)}{\mathbb{P}(X_2 = 1)}$$
$$= \frac{3}{5} \times \frac{3/5}{51/125}$$
$$= \frac{15}{17},$$

which allows us to check that

$$\mathbb{P}(X_1 = 2 \mid X_2 = 1) + \mathbb{P}(X_1 = 1 \mid X_2 = 1) = \frac{2}{17} + \frac{15}{17} = 1,$$

since $\mathbb{P}(X_1 = 0 \mid X_2 = 1) = 0$.

Exercise 8.2

a) We have

$$G_1(s) = \mathbb{E}[s^Y] = s^0 \mathbb{P}(Y=0) + s^1 \mathbb{P}(Y=1) = rac{1}{2} + rac{1}{2}s, \qquad s \in \mathbb{R}.$$

b) We prove this statement by induction. Clearly it holds at the order 1. Next, assuming that (8.3.10) holds at the order $n \ge 1$ we get

$$egin{array}{rll} G_{n+1}(s) &=& G_1(G_n(s)) = G_1\left(1-rac{1}{2^n}+rac{s}{2^n}
ight) \ &=& rac{1}{2}+rac{1}{2}\left(1-rac{1}{2^n}+rac{s}{2^n}
ight) = 1-rac{1}{2^{n+1}}+rac{s}{2^{n+1}}. \end{array}$$

ॳ 677

Additional comments:

(i) We may also directly note that

$$\mathbb{P}(X_n = 1 \mid X_0 = 1) = \mathbb{P}(Y_1 = 1, Y_2 = 1, \dots, Y_n = 1) = (\mathbb{P}(Y = 1))^n = \frac{1}{2^n},$$

hence

$$\mathbb{P}(X_n=0\mid X_0=1)=1-\mathbb{P}(X_n=1\mid X_0=1)=1-rac{1}{2^n},$$

and

$$G_n(s) = \mathbb{P}(X_n = 0 \mid X_0 = 1) + s \mathbb{P}(X_n = 1 \mid X_0 = 1) = 1 - rac{1}{2^n} + rac{s}{2^n}$$

(ii) It is also possible to write

$$G_n(s) = \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^n} + \frac{s}{2^n} = \frac{s}{2^n} + \frac{1}{2} \sum_{k=0}^{n-1} \frac{1}{2^k}$$
$$= \frac{1}{2^n} s + \frac{1}{2} \times \frac{1 - (1/2)^n}{1 - 1/2} = 1 - \frac{1}{2^n} + \frac{1}{2^n} s,$$

although this is not recommended here.

(iii) It is **wrong** (why?) to write

$$G_n(s) = \mathbb{E}[s^{Y_1 + \dots + Y_{X_n}}] = A\!\!\mathbb{E}[s^{Y_1}] \,/ \cdot \!/ \cdot I\!\!\mathbb{E}[s^{Y_{X_n}}]$$

Note that the left hand side is a well-defined deterministic number, while the right hand side is not well-defined as a random product.

- (iv) For $n \ge 2$ we do **not** have $G_n(s) = (G_1(s))^n$.
- c) We have

$$\mathbb{P}(X_n=0 \mid X_0=1) = G_n(0) = 1 - rac{1}{2^n}$$

Additional comments:

(i) We may also directly write

$$\mathbb{P}(X_n=0 \mid X_0=1) \; = \; 1 - \mathbb{P}(X_n=1 \mid X_0=1)$$

♦679

On the other hand, we do **not** have

$$\mathbb{P}(X_n=0\mid X_0=1)=\langle \mathbb{P}(X_1
eq/0
angle)^n,$$

since the events $\{X_n = 0\}$ and $\{Y_1 = 0, Y_2 = 0, \dots, Y_n = 0\}$ are **not** equivalent, more precisely we only have

$$\{Y_1=0,Y_2=0,\ldots,Y_n=0\} \subsetneqq \{X_n=0\},$$

hence

$$rac{1}{2^n} \leqslant \mathbb{P}(X_n=0 \mid X_0=1) = 1-rac{1}{2^n}, \qquad n \geqslant 1.$$

(ii) The probability $\beta_n := \mathbb{P}(X_n = 0 \mid X_0 = 1)$ is not solution of $G_n(\beta_n) = \beta_n$. It is easy to check that the equality

$$G_n\left(1-rac{1}{2^n}
ight)=1
eq 1/rac{1}{2^n}$$

does *not* hold for $n \ge 1$.

(iii) In fact, $\{X_n = 0\}$ means that extinction occurs at time n, or has already occurred *before* time n.

d) We have

$$\mathbb{E}[X_n \mid X_0 = 1] = G'_n(s)_{|s=1} = (\mathbb{E}[Y_1])^n = rac{1}{2^n}.$$

Additional comment:

In this simple setting we could also write

$$egin{aligned} \mathbb{E}[X_n \mid X_0 = 1] &= 0 imes \mathbb{P}(X_n = 0 \mid X_0 = 1) + 1 imes \mathbb{P}(X_n = 1 \mid X_0 = 1) \ &= \mathbb{P}(X_n = 1 \mid X_0 = 1) \ &= rac{1}{2^n}. \end{aligned}$$

e) The extinction probability α is solution of $G_1(\alpha) = \alpha$, *i.e.*

$$\alpha = \frac{1}{2} + \frac{1}{2}\alpha,$$

with unique solution $\alpha = 1$.

Additional comment:

Since the sequence of events $(\{X_n = 0\})_{n \ge 1}$ is increasing, we also have

$$lpha = \mathbb{P}\left(igcup_{n \geqslant 1}\{X_n = 0\}
ight) = \lim_{n o \infty} \mathbb{P}(\{X_n = 0\}) = \lim_{n o \infty} 1 - rac{1}{2^n} = 1.$$

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

$$egin{aligned} G_2(s) &= G_1(G_1(s)) \ &= G_1(0.2 + 0.5s + 0.3s^2) \ &= 0.2 + 0.5(0.2 + 0.5s + 0.3s^2) + 0.3(0.2 + 0.5s + 0.3s^2)^2 \ &= 0.312 + 0.31s + 0.261s^2 + 0.09s^3 + 0.027s^4, \end{aligned}$$

with

$$G_2^\prime(s) = 0.31 + 0.522s + 0.27s^2 + 0.108s^3$$

and

$$G_2''(s) = 0.522 + 0.54s + 0.324s^2,$$

hence

ॳ 681

$$G'_2(1) = (G'_1(1))^2 = (1.1)^2 = 1.21$$
 and $G''_2(1) = 1.386$,

and

$$\mathbb{E}[X_2^2] = G_2''(1) + G_2'(1) = 1.386 + 1.21 = 2.596.$$

By (1.7.5) this yields

$$\operatorname{Var}[X_2] = 2.596 - (1.21)^2.$$

b) We have

$$G_2(s) = 0.312 + 0.31s + 0.261s^2 + 0.09s^3 + 0.027s^4,$$

hence

$$\mathbb{P}(X_2=0)=0.312, \ \ \mathbb{P}(X_2=1)=0.31, \ \ \mathbb{P}(X_2=2)=0.261,$$

and

$$\mathbb{P}(X_2 = 3) = 0.09, \quad \mathbb{P}(X_2 = 4) = 0.027.$$

c) We have

$$\begin{split} \mathbb{P}(X_4 = 0) &= G_4(0) \\ &= G_2(G_2(0)) \\ &= 0.312 + 0.31G_2(0) + 0.261(G_2(0))^2 + 0.09(G_2(0))^3 + 0.027(G_2(0))^4 \\ &= 0.312 + 0.31 \times 0.312 + 0.261 \times 0.312^2 + 0.09 \times 0.312^3 + 0.027 \times 0.312^4 \\ &\simeq 0.44314. \end{split}$$

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 $\pmb{\alpha}$ solves the equation

$$lpha = G_1(lpha) = 0.2 + 0.5 lpha + 0.3 lpha^2,$$

i.e.

$$0.3lpha^2 - 0.5lpha + 0.2 = 0.3(lpha - 1)(lpha - 2/3) = 0,$$

hence $\alpha = 2/3$.

♦ 682

Exercise 8.4

a) We have

$$G_1(s)=\mathbb{P}(Y=0)+s\mathbb{P}(Y=1)+s^2\mathbb{P}(Y=2)=as^2+bs+c,\qquad s\in\mathbb{R}.$$

b) Letting X_n denote the number of individuals in the population at generation $n \ge 0$, 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 8.5

a) We have

$$G_Y(s) = \mathbb{P}(Y=0) + s\mathbb{P}(Y=1) + s^2\mathbb{P}(Y=2) = q^2 + 2spq + s^2p^2 = (q+ps)^2,$$

b) The extinction probability satisfies the equation $\alpha = G_Y(\alpha)$, *i.e.*

$$\alpha = q^2 + 2\alpha pq + \alpha^2 p^2 = (q + \alpha p)^2,$$

with solutions $\alpha = 1$ and $\alpha = q^2/p^2$, hence the extinction probability is

$$\min\left(1,rac{q^2}{p^2}
ight) = \left\{egin{array}{cc} 1 & ext{if } p \leqslant q, \ \ rac{q^2}{p^2} & ext{if } p \geqslant q. \end{array}
ight.$$

c) Since $\{Z_1 = 0\} \subset \{Z_2 = 0\}$, this probability is

$$\mathbb{P}(Z_2 = 0 \text{ and } Z_1 \ge 1) = \mathbb{P}(Z_2 = 0 \cap \{Z_1 = 0\}^c)$$

= $\mathbb{P}(Z_2 = 0) - \mathbb{P}(Z_1 = 0)$
= $G_Y(G_Y(0)) - G_Y(0)$
= $2pq^3 + q^2p^4$.

Alternatively, it can be recovered pathwise as

$$2pq imes q^2+p^2 imes q^2 imes q^2=2pq^3+p^2q^4.$$

d) When $q \leq p$ the extinction probability is q^2/p^2 starting from $Z_0 = 1$, hence when Z_0 has a Poisson distribution with parameter $\lambda > 0$ this extinction probability is

$$\mathbb{P}(T_0 < \infty) = \sum_{n \ge 0} \mathbb{P}(T_0 < \infty ext{ and } Z_0 = n)$$

= $\mathbb{P}(Z_0 = n) \sum_{n \ge 0} \mathbb{P}(T_0 < \infty \mid Z_0 = n)$
= $\sum_{n \ge 0} \left(\frac{q^2}{p^2}\right)^n \mathbb{P}(Z_0 = n)$

ॳ 683

$$\begin{array}{rcl} &=& \mathrm{e}^{-\lambda}\sum\limits_{n \geqslant 0} \frac{1}{n!} \left(\frac{\lambda q^2}{p^2} \right)^n \\ &=& \mathrm{e}^{\lambda (q^2/p^2 - 1)} \\ &=& \mathrm{e}^{\lambda (q^2 - p^2)/p^2} \\ &=& \mathrm{e}^{\lambda (q - p)/p^2}, \end{array}$$

whereas when $q \ge p$ the extinction probability becomes

$$\sum\limits_{n \geqslant 0} \mathbb{P}(Z_0 = n) = \mathrm{e}^{-\lambda} \sum\limits_{n \geqslant 0} rac{\lambda^n}{n!} = 1.$$

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

$$\begin{aligned} \frac{1}{4} \times \left(\frac{1}{4}\right)^2 \times \cdots \times \left(\frac{1}{4}\right)^{2^{n-1}} &= \prod_{k=0}^{n-1} \left(\frac{1}{4}\right)^{2^k} \\ &= \left(\frac{1}{4}\right)^{\sum_{k=0}^{n-1} 2^k} \\ &= \left(\frac{1}{4}\right)^{2^n-1}, \qquad n \ge 0. \end{aligned}$$

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

$$\mathbb{P}(Y=0)=rac{1}{12}, \qquad \mathbb{P}(Y=1)=rac{2}{3}, \qquad \mathbb{P}(Y=2)=rac{1}{4},$$

which has the generating function

$$egin{aligned} G_1(s) &= & \mathbb{P}(Y=0) + s \mathbb{P}(Y=1) + s^2 \mathbb{P}(Y=2) \ &= & rac{1}{12} + rac{2s}{3} + rac{s^2}{4} = rac{1}{12}(1+8s+3s^2), \end{aligned}$$
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

$$\frac{1}{3} \times \left(\frac{1}{3}\right)^2 \times \cdots \times \left(\frac{1}{3}\right)^{2^{n-1}} = \prod_{k=0}^{n-1} \left(\frac{1}{3}\right)^{2^k} = \left(\frac{1}{3}\right)^{2^{n-1}},$$

 $n \ge 0$. The probability distribution

$$\mathbb{P}(Y=0) = \frac{1}{6}, \quad \mathbb{P}(Y=1) = \frac{2}{2}, \quad \mathbb{P}(Y=2) = \frac{1}{3},$$

of the number \boldsymbol{Y} of red cells has the generating function

$$egin{array}{rcl} G_1(s) &=& {
m I\!P}(Y=0) + s {
m I\!P}(Y=1) + s^2 {
m I\!P}(Y=2) \ &=& rac{1}{6} + rac{s}{2} + rac{s^2}{3} = rac{1}{12}(2+6s+4s^2), \end{array}$$

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 8.7 Using the relation^{*}

$$\mathbb{E}[T_0 \mid X_0 = k] \leqslant k \, \mathbb{E}[T_0 \mid X_0 = 1], \qquad k \in \mathbb{N},$$

we have

$$egin{array}{rl} \mathbb{E}[T_0 \mid X_0 = 1] &=& \sum\limits_{k \geqslant 0} \mathbb{P}(Y_1 = k) ig(1 + \mathbb{E}[T_0 \mid X_0 = k]ig) \ &=& \sum\limits_{k \geqslant 0} \mathbb{P}(Y_1 = k) + \sum\limits_{k \geqslant 0} \mathbb{P}(Y_1 = k) \mathbb{E}[T_0 \mid X_0 = k] \end{array}$$

*This inequality follows from the relation $\mathbb{E}[\max(X_1,\ldots,X_n)] \leq \mathbb{E}[X_1+\cdots+X_n].$

() 686

$$egin{array}{rll} &=& 1+\sum\limits_{k \geqslant 0} \mathbb{P}(Y_1=k) \ \mathbb{E}[T_0 \mid X_0=k] \ &\leqslant& 1+\mathbb{E}[T_0 \mid X_0=1] \sum\limits_{k \geqslant 0} k \mathbb{P}(Y_1=k) \ &=& 1+\mathbb{E}[T_0 \mid X_0=1] \ \mathbb{E}[Y_1], \end{array}$$

hence

$$\mathbb{E}[T_0 \mid X_0 = 1] \leqslant rac{1}{1 - \mathbb{E}[Y_1]} < \infty.$$

Exercise 8.8 This is a particular case of Example (iv) on page 328.

a) We have

$$G_1(s) = q^2 \sum_{n \geqslant 1} n(ps)^{n-1} = rac{q^2}{(1-ps)^2}.$$

b) The equation $G_1(\alpha) = \alpha$ reads

$$lpha = rac{q^2}{(1-plpha)^2},$$

i.e.

$$p^2\alpha^3 - 2p\alpha^2 + \alpha - q^2 = 0,$$

which is known to admit $\alpha = 1$ for solution, *i.e.*

$$(\alpha - 1)(p^2\alpha^2 - (1 - q^2)\alpha + q^2) = (\alpha - 1)(p^2\alpha^2 - p(1 + q)\alpha + q^2) = 0,$$

whose smallest solution

$$lpha = rac{p(1+q) - \sqrt{p^2(1+q)^2 - 4p^2q^2}}{2} = prac{2-p - \sqrt{4p - 3p^2}}{2}$$

is the extinction probability.

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

$$G_X(s)=\mathbb{E}[s^X]=\sum\limits_{k\geqslant 0}s^k\mathbb{P}(X=k)=rac{1}{2}\sum\limits_{k\geqslant 0}(s/2)^k=rac{1}{2-s},$$

 $-1 < s \leq 1$.

c) The probability we are looking for is

$$\mathbb{P}(X_3=0 \mid X_0=0) = G_X(G_X(G_X(0))) = rac{1}{2-rac{1}{2-1/2}} = rac{3}{4}.$$

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

$$G_Z^{(3)}(s) = G_Z(G_Z(G_Z(s))) = rac{37}{64} + rac{27s}{64}$$

at s = 0, where G_Z is the probability generating function

$$G_Z(s)=rac{1}{4}+rac{3s}{4}$$

Exercise 8.10

a) We have

$$egin{aligned} G_{Z_1}(s) &=& 1-rac{q}{1-p}+qs\sum\limits_{k\geqslant 1}(ps)^{k-1}\ &=& 1-rac{q}{1-p}+rac{qs}{1-ps}\ &=& rac{1-p-q}{1-p}+rac{qs}{1-ps}. \end{aligned}$$

MH3512 AY19-20

ॳ 687

b) We have

$$G'_{Z_1}(s) = rac{q}{1-ps} + rac{pqs}{(1-ps)^2},$$

hence

$$\mathbb{E}[Z_1] = G'_{Z_1}(1) = rac{q}{1-p} + rac{pq}{(1-p)^2} = rac{q(1-p)+pq}{(1-p)^2} = rac{q}{(1-p)^2}.$$

- c) The critical value of q is $q = (1-p)^2$.
- d) When $q = (1-p)^2$ we have

$$G_{Z_1}(s) = p + rac{(1-p)^2 s}{1-ps}, \qquad -1 < s < 1.$$

The relation is clearly satisfied at the rank k = 1. Next, assuming that the relation

$$G_{Z_k}(s) = rac{kp - (kp + p - 1)s}{1 - p + kp - kps}, \qquad -1 < s < 1,$$

holds at the rank $k \ge 1$, we have

$$\begin{aligned} G_{Z_{k+1}}(s) &= G_{Z_k}(G_{Z_1}(s)) \\ &= \frac{kp - (kp + p - 1)G_{Z_1}(s)}{1 - p + kp - kpG_{Z_1}(s)}, \\ &= \frac{kp - (kp + p - 1)\left(p + \frac{(1 - p)^2 s}{1 - ps}\right)}{1 - p + kp - kp\left(p + \frac{(1 - p)^2 s}{1 - ps}\right)} \\ &= \frac{kp(1 - ps) - (kp + p - 1)\left(p(1 - ps) + (1 - p)^2 s\right)}{(1 - p + kp)(1 - ps) - kp\left(p(1 - ps) + (1 - p)^2 s\right)} \\ &= \frac{kp(1 - ps) - (kp + p - 1)\left(p + (1 - 2p)s\right)}{(1 - p + kp)(1 - ps) - kp\left(p + (1 - 2p)s\right)} \\ &= \frac{(1 - p)(k + 1)p - (1 - p)((k + 2)p - 1)s}{(1 - p)(kp + 1) - (1 - p)(k + 1)ps} \end{aligned}$$

$$= \frac{(k+1)p - ((k+1)p + p - 1)s}{1 - p + (k+1)p - (k+1)ps}, \qquad -1 < s < 1.$$

Problem 8.11

a) We have

$$\mathbb{E}[Z_n] = \sum\limits_{k=1}^n \mathbb{E}[X_k] = \sum\limits_{k=1}^n \mu^k = \mu \sum\limits_{k=0}^{n-1} \mu^k = \mu rac{1-\mu}{1-\mu}^n, \qquad n \in \mathbb{N}.$$

b) We have

$$\mathbb{E}[Z] = \mathbb{E}\left[\sum_{k \geqslant 1} X_k
ight] = \sum_{k \geqslant 1} \mathbb{E}[X_k] = \mu \sum_{k \geqslant 0} \mu^k = rac{\mu}{1-\mu}, \qquad n \in \mathbb{N},$$

provided that $\mu < 1$.

c) We have

$$\begin{split} H(s) &= \mathbb{E} \left[s^{Z} \mid X_{0} = 1 \right] \\ &= s^{0} \mathbb{P}(Y_{1} = 0) + \sum_{k \ge 1} \mathbb{E} \left[s^{Z} \mid X_{1} = k \right] \mathbb{P}(Y_{1} = k) \\ &= \mathbb{P}(Y_{1} = 0) + \sum_{k \ge 1} \left(\mathbb{E} \left[s^{Z} \mid X_{1} = 1 \right] \right)^{k} \mathbb{P}(Y_{1} = k) \\ &= \mathbb{P}(Y_{1} = 0) + \sum_{k \ge 1} \left(\mathbb{E} \left[s^{1 + \sum_{i \ge 2} X_{i}} \mid X_{1} = 1 \right] \right)^{k} \mathbb{P}(Y_{1} = k) \\ &= \mathbb{P}(Y_{1} = 0) + \sum_{k \ge 1} \left(\mathbb{E} \left[s^{1 + \sum_{i \ge 1} X_{i}} \mid X_{0} = 1 \right] \right)^{k} \mathbb{P}(Y_{1} = k) \\ &= \mathbb{P}(Y_{1} = 0) + \sum_{k \ge 1} \left(\mathbb{E} \left[s^{1 + Z} \mid X_{0} = 1 \right] \right)^{k} \mathbb{P}(Y_{1} = k) \\ &= \sum_{k \ge 0} \left(s \mathbb{E} \left[s^{Z} \mid X_{0} = 1 \right] \right)^{k} \mathbb{P}(Y_{1} = k) \\ &= \sum_{k \ge 0} (s \mathbb{E} \left[s^{Z} \mid X_{0} = 1 \right] \right)^{k} \mathbb{P}(Y_{1} = k) \\ &= G_{1}(sH(s)). \end{split}$$

d) We have

$$H(s) = G_1(sH(s)) = rac{1-p}{1-psH(s)},$$

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

$$\lim_{s\searrow 0^+} H(s) = \lim_{s\searrow 0^+} rac{1-(1-2pqs)}{2ps} = q = \mathbb{P}(Z=0) = \mathbb{P}(Y_1=0) = H(0).$$

Alternatively, L'Hospital's rule can be used to compute the limit of H(s) expressed as a ratio.

f) We have

$$H'(s) = \frac{pq}{ps\sqrt{1-4pqs}} - \frac{1-\sqrt{1-4pqs}}{2ps^2},$$

and

$$\begin{array}{ll} H'(1) &=& \displaystyle \frac{pq}{p\sqrt{1-4pq}} - \frac{1-\sqrt{1-4pq}}{2p} = \frac{pq}{p(q-p)} - \frac{1-(q-p)}{2p} \\ &=& \displaystyle \frac{q}{q-p} - 1 = \displaystyle \frac{p}{q-p} = \displaystyle \frac{\mu}{1-\mu}, \end{array}$$

with $\mu = p/q$ for p < 1/2, which shows that

$$\mathbb{E}[Z] = rac{\mu}{1-\mu}$$

and recovers the result of Question (b).

$$egin{aligned} \mathbb{E}\left[\sum\limits_{k=1}^{Z}U_k
ight] &= \sum\limits_{n\geqslant 0}\mathbb{E}\left[\sum\limits_{k=1}^{Z}U_k
ight|Z=n
ight]\mathbb{P}(Z=n) \ &= \sum\limits_{n\geqslant 0}\mathbb{E}\left[\sum\limits_{k=1}^{n}U_k
ight]\mathbb{P}(Z=n) = \sum\limits_{n\geqslant 0}n\,\mathbb{E}[U_1]\mathbb{P}(Z=n) \ &= \mathbb{E}[U_1]\,\mathbb{E}[Z] = \mathbb{E}[U_1]rac{\mu}{1-\mu}. \end{aligned}$$

h) We have

$$egin{aligned} \mathbb{P}(U_k < x, \ k = 1, 2, \dots, Z) &= \sum\limits_{n \geqslant 0} \mathbb{P}(U_k < x, \ k = 1, 2, \dots, Z \mid Z = n) \mathbb{P}(Z = n) \ &= \sum\limits_{n \geqslant 0} \mathbb{P}(U_k < x, \ k = 1, 2, \dots, n) \mathbb{P}(Z = n) \ &= \sum\limits_{n \geqslant 0} \mathbb{P}(U_1 < x) imes \dots imes \mathbb{P}(U_n < x) \mathbb{P}(Z = n) \ &= \sum\limits_{n \geqslant 0} (\mathbb{P}(U_1 < x))^n \mathbb{P}(Z = n) = \sum\limits_{n \geqslant 0} (F(x))^n \mathbb{P}(Z = n) = H(F(x)), \end{aligned}$$

under the convention that the condition is satisfied by default when Z = 0.

Remark. We could also compute the same probability given that $Z \ge 1$, and this would give

$$\begin{split} \mathbb{P}(U_k < x, \ k = 1, 2, \dots, Z \mid Z \geqslant 1) \\ &= \frac{1}{\mathbb{P}(Z \geqslant 1)} \sum_{n \geqslant 1} \mathbb{P}(U_k < x, \ k = 1, 2, \dots, Z \mid Z = n) \mathbb{P}(Z = n) \\ &= \frac{1}{\mathbb{P}(Z \geqslant 1)} \sum_{n \geqslant 1} \mathbb{P}(U_k < x, \ k = 1, 2, \dots, n) \mathbb{P}(Z = n) \\ &= \frac{1}{\mathbb{P}(Z \geqslant 1)} \sum_{n \geqslant 1} \mathbb{P}(U_1 < x)^n \mathbb{P}(Z = n) = \sum_{n \geqslant 1} F(x)^n \frac{\mathbb{P}(Z = n)}{\mathbb{P}(Z \geqslant 1)} \\ &= \frac{1}{\mathbb{P}(Z \geqslant 1)} (H(F(x)) - \mathbb{P}(Z = 0)) = \frac{1}{p} (H(F(x)) - q). \end{split}$$

ර 691

i) We have

$$\mathbb{E}\left[\sum\limits_{k=1}^{Z} U_k
ight] = \mathbb{E}[U_1]rac{\mu}{1-\mu} = rac{\mu}{1-\mu} = rac{p}{q-p}$$

We find

$$egin{aligned} \mathbb{P}(U_k < x, \; k = 1, 2, \dots, Z) \; &= \; H(F(x)) \ &= \; H(1 - e^{-x}) \ &= \; rac{1 - \sqrt{1 - 4pq(1 - e^{-x})}}{2p(1 - e^{-x})}. \end{aligned}$$

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

$$\begin{split} \mathbb{P}(\min(\tau_1,\tau_2,\ldots,\tau_k)>t) &= \mathbb{P}(\tau_1>t,\tau_2>t,\ldots,\tau_k>t) \\ &= \mathbb{P}(\tau_1>t)\mathbb{P}(\tau_2>t)\cdots\mathbb{P}(\tau_k>t) = (\mathrm{e}^{-\mu t})^k = \mathrm{e}^{-k\mu t}, \end{split}$$

 $t \in \mathbb{R}_+$, 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

$$Q = egin{bmatrix} -\lambda & \lambda & 0 & 0 & 0 & 0 \ \mu & -\mu - \lambda & \lambda & 0 & 0 & 0 \ 0 & 2\mu & -2\mu - \lambda & \lambda & 0 & 0 \ 0 & 0 & 3\mu & -3\mu - \lambda & \lambda & 0 \ 0 & 0 & 0 & 4\mu & -4\mu - \lambda & \lambda \ 0 & 0 & 0 & 0 & 5\mu & -5\mu \end{bmatrix},$$

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 $\pi Q = 0$, *i.e.*

$$\begin{cases} 0 = -\lambda \pi_{0} + \mu \pi_{1} \\ 0 = \lambda \pi_{0} - (\lambda + \mu) \pi_{1} + 2\mu \pi_{2} \\ 0 = \lambda \pi_{1} - (\lambda + 2\mu) \pi_{2} + 3\mu \pi_{3} \\ 0 = \lambda \pi_{2} - (\lambda + 3\mu) \pi_{3} + 4\mu \pi_{4} \\ 0 = \lambda \pi_{3} - (\lambda + 4\mu) \pi_{4} + 5\mu \pi_{5} \\ 0 = \lambda \pi_{4} - 5\mu \pi_{5}, \end{cases}$$
 which gives
$$\begin{cases} 0 = -\lambda \pi_{0} + \mu \pi_{1} \\ 0 = -\lambda \pi_{1} + 2\mu \pi_{2} \\ 0 = -\lambda \pi_{2} + 3\mu \pi_{3} \\ 0 = -\lambda \pi_{3} + 4\mu \pi_{4} \\ 0 = -\lambda \pi_{4} + 5\mu \pi_{5}, \end{cases}$$

hence

$$\pi_1 = rac{\lambda}{\mu} \pi_0, \quad \pi_2 = rac{\lambda}{2\mu} \pi_1, \quad \pi_3 = rac{\lambda}{3\mu} \pi_2, \quad \pi_4 = rac{\lambda}{4\mu} \pi_3, \quad \pi_5 = rac{\lambda}{5\mu} \pi_4,$$
 $\pi_1 = rac{\lambda}{\mu} \pi_0, \quad \pi_2 = rac{\lambda^2}{2\mu^2} \pi_0, \quad \pi_3 = rac{\lambda^3}{3!\mu^3} \pi_0, \quad \pi_4 = rac{\lambda}{4!\mu^4} \pi_0, \quad \pi_5 = rac{\lambda}{5!\mu^5} \pi_0,$

i.e.

which is a truncated Poisson distribution with

$$\pi_0+rac{\lambda}{\mu}\pi_0+rac{\lambda^2}{2\mu^2}\pi_0+rac{\lambda^3}{3!\mu^3}\pi_0+rac{\lambda^4}{4!\mu^4}\pi_0+rac{\lambda^5}{5!\mu^5}\pi_0=1,$$

hence

$$egin{aligned} \pi_0 &= rac{1}{1+rac{\lambda}{\mu}+rac{\lambda^2}{2\mu^2}+rac{\lambda^3}{3!\mu^3}+rac{\lambda^4}{4!\mu^4}+rac{\lambda^5}{5!\mu^5}} \ &= rac{\mu^5}{\mu^5+\lambda\mu^4+\lambda^2\mu^3/2+\lambda^3\mu^2/3!+\lambda^4\mu/4!+\lambda^5/5!}. \end{aligned}$$

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

$$\pi_5 = rac{\lambda^5}{120 \mu^5 + 120 \lambda \mu^4 + 60 \lambda^2 \mu^3 + 20 \lambda^3 \mu^2 + 5 \lambda^4 \mu + \lambda^5}$$

Note that of at most two machines can be under repair, the infinitesimal generator Q of $(X_t)_{t \in \mathbb{R}_+}$ will become

$$Q = egin{bmatrix} -2\lambda & 2\lambda & 0 & 0 & 0 & 0 \ \mu & -\mu - 2\lambda & 2\lambda & 0 & 0 & 0 \ 0 & 2\mu & -2\mu - 2\lambda & 2\lambda & 0 & 0 \ 0 & 0 & 3\mu & -3\mu - 2\lambda & 2\lambda & 0 \ 0 & 0 & 0 & 4\mu & -4\mu - \lambda & \lambda \ 0 & 0 & 0 & 0 & 5\mu & -5\mu \end{bmatrix}$$

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

$$\mathbb{P}(\tau_k^R > t) = \mathrm{e}^{-\lambda_R t},$$

where N_t^R denotes a Poisson process with intensity λ_R .

MH3512 AY19-20 $\,$

$$egin{aligned} \mathbb{P}(N^W_t\leqslant 3) &= \mathbb{P}(N^W_t=0) + \mathbb{P}(N^W_t=1) + \mathbb{P}(N^W_t=2) + \mathbb{P}(N^W_t=3) \ &= \mathrm{e}^{-\lambda_W t} + \lambda_W t \mathrm{e}^{-\lambda_W t} + \mathrm{e}^{-\lambda_W t} rac{\lambda^2_W t^2}{2} + \mathrm{e}^{-\lambda_W t} rac{\lambda^3_W t^3}{6}, \end{aligned}$$

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.

Note that the difference $N_t^R - N_t^W$ between the number N_t^R of "read" consultations and the number N_t^W of "write" consultations is a birth and death process with state-independent birth and death rates λ_R and λ_W .

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

$$\begin{split} \mathbb{P}(N_t^R = k \mid N_t^R + N_t^W = n) &= \frac{\mathbb{P}(N_t^R = k \text{ and } N_t^R + N_t^W = n)}{\mathbb{P}(N_t^R + N_t^W = n)} \\ &= \frac{\mathbb{P}(N_t^R = k \text{ and } N_t^W = n - k)}{\mathbb{P}(N_t^R + N_t^W = n)} = \frac{\mathbb{P}(N_t^R = k)\mathbb{P}(N_t^W = n - k)}{\mathbb{P}(N_t^R + N_t^W = n)} \\ &= e^{-\lambda_R t} \frac{(\lambda_R t)^k}{k!} e^{-\lambda_W t} \frac{(\lambda_W t)^{n-k}}{(n-k)!} \left(e^{-\lambda_W t - \lambda_R t} \frac{(\lambda_W + \lambda_R)^n t^n}{n!} \right)^{-1} \\ &= \binom{n}{k} \left(\frac{\lambda_R}{\lambda_R + \lambda_W} \right)^k \left(\frac{\lambda_W}{\lambda_R + \lambda_W} \right)^{n-k}, \quad k = 0, 1, \dots, n, \end{split}$$

cf. (B.10) in the solution of Exercise 1.6.

a) The number X_t of machines operating at time t is a birth and death process on $\{0, 1, 2\}$ with infinitesimal generator

$$Q = \left[egin{array}{ccc} -\lambda & \lambda & 0 \ \mu & -(\lambda+\mu) & \lambda \ 0 & 2\mu & -2\mu \end{array}
ight].$$

The stationary distribution $\pi = (\pi_0, \pi_1, \pi_2)$ is solution of $\pi Q = 0$, *i.e.*

$$\left\{egin{array}{l} 0 = -\lambda \pi_0 + \mu \pi_1 \ 0 = \lambda \pi_0 - (\lambda + \mu) \pi_1 + 2 \mu \pi_2 \ 0 = \lambda \pi_1 - 2 \mu \pi_2 \end{array}
ight.$$

under the condition $\pi_0 + \pi_1 + \pi_2 = 1$, which yields

$$(\pi_0,\pi_1,\pi_2)=\left(rac{2\mu^2}{2\mu^2+2\lambda\mu+\lambda^2},rac{2\lambda\mu}{2\mu^2+2\lambda\mu+\lambda^2},rac{\lambda^2}{2\mu^2+2\lambda\mu+\lambda^2}
ight),$$

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

$$egin{aligned} \mathbb{E}[T_0 \mid X_0 = 1] \ &= \mathbb{P}(X_\lambda < X_\mu) imes \left(rac{1}{\mu} + \mathbb{E}[T_0 \mid X_0 = 1]
ight) + \mathbb{P}(X_\mu < X_\lambda) imes rac{1}{\mu} \ &= rac{1}{\mu} + \mathbb{P}(X_\lambda < X_\mu) imes \mathbb{E}[T_0 \mid X_0 = 1] \ &= rac{1}{\mu} + rac{\lambda}{\lambda + \mu} \mathbb{E}[T_0 \mid X_0 = 1], \end{aligned}$$

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 first step_analysis

argument is similar to the one used in the solution of Exercise 9.17. This yields

$$\mathbb{E}[T_0 \mid X_0 = 1] = rac{\lambda + \mu}{\mu^2},$$

hence the corresponding rate is $\mu^2/(\lambda + \mu)$ and the infinitesimal generator of the chain becomes

$$Q = \left[egin{array}{ccc} -\lambda & \lambda \ rac{1}{\mathbb{E}[T_0 \mid X_0 = 1]} & -rac{1}{\mathbb{E}[T_0 \mid X_0 = 1]} \end{array}
ight] = \left[egin{array}{ccc} -\lambda & \lambda \ rac{\mu^2}{\lambda + \mu} & -rac{\mu^2}{\lambda + \mu} \end{array}
ight].$$

The stationary distribution $\pi = (\pi_0, \pi_1)$ is solution of $\pi Q = 0$, *i.e.*

$$\left\{egin{array}{l} 0=-\lambda\pi_{0}+\pi_{1}rac{\mu^{2}}{\lambda+\mu}\ 0=\lambda\pi_{0}-\pi_{1}rac{\mu^{2}}{\lambda+\mu} \end{array}
ight.$$

under the condition $\pi_0 + \pi_1 = 1$, which yields

$$(\pi_0,\pi_1)=\left(rac{\mu^2}{\mu^2+\lambda\mu+\lambda^2},rac{\lambda\mu+\lambda^2}{\mu^2+\lambda\mu+\lambda^2}
ight),$$

i.e. the probability that no machine is operating when $\lambda = \mu = 1$ is $\pi_0 = 1/3$.

Alternatively, this question can be solved by considering a birth and death process $(X_t)_{t \in \mathbb{R}_+}$ on $\{0, 1, 2\}$ with infinitesimal generator

$$Q = egin{bmatrix} -\lambda & \lambda & 0 \ \mu & -(\lambda + \mu) & \lambda \ 0 & \mu & -\mu \end{bmatrix},$$
 MH3512 AY19-20

for which

- $X_t = 0 \iff$ no machine is working,
- $X_t = 1 \iff$ one machine is working and the other is under repair,
- $X_t = 2 \iff$ one machine is working and the other one is waiting.

In this case, $\pi Q = 0$ yields

$$(\pi_0,\pi_1,\pi_2)=\left(rac{\mu^2}{\mu^2+\lambda\mu+\lambda^2},rac{\lambda\mu}{\mu^2+\lambda\mu+\lambda^2},rac{\lambda^2}{\mu^2+\lambda\mu+\lambda^2}
ight),$$

hence

$$\pi_0 = \frac{1}{3}, \qquad \pi_1 = \frac{1}{3}, \qquad \pi_2 = \frac{1}{3},$$

when $\lambda = \mu$.

Exercise 9.4

a) The process $(X_t)_{t \in \mathbb{R}_+}$ is a continuous-time Markov chain due to the Poisson arrival of customers in the queue. The state space of the chain $(X_t)_{t \in \mathbb{R}_+}$ is $\mathbb{S} = \{0, 1, 2, 3\}$ since there cannot be more than 3 people in the queue as the cable car departs immediately as soon as there are more 4 people in the queue. The infinitesimal generator of the chain is given by

$$Q = egin{bmatrix} -\lambda & \lambda & 0 & 0 \ 0 & -\lambda & \lambda & 0 \ 0 & 0 & -\lambda & \lambda \ \lambda & 0 & 0 & -\lambda \end{bmatrix}$$

- c) This mean time is the sum

$$\mathbb{E}[au_0] + \mathbb{E}[au_1] + \mathbb{E}[au_2] + \mathbb{E}[au_3] = rac{4}{\lambda}$$

of the means of four exponential random variables with parameter $\lambda > 0$, hence it equals $4/\lambda$.

Exercise 9.5 [MT15]

a) The process $(L_t)_{t \in \mathbb{R}_+}$ is Markov from its construction using Poisson processes and the memoryless property of exponential random times. The matrix of the infinitesimal generator is

-5λ	5λ	0	0	0	0	•••]	
$\mu + heta$	$-\mu - heta - 5\lambda$	5λ	0	0	0	•••	
0	$2\mu + heta$	$-2\mu- heta-5\lambda$	5λ	0	0	•••	
0	0	$3\mu + heta$	$-3\mu - heta - 5\lambda$	5λ	0	•••	•
•	•	•	•	•	•	•	
•	•	•	•	:	:	••	

b) Given that the mean number of issued order until time t is λt and the mean number of cancelled order until time t is μt , we deduce the relation $\mu t = 0.95\lambda t$, *i.e.* $\mu = 0.95\lambda$.

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 $\lambda_k = (1+k)^{\rho}, k \ge 1$. 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 $\sum_{k\ge 1} \tau_k$. It is known that $\sum_{k\ge 1} \tau_k < \infty$

almost surely if the expectation $\mathbb{E}\left[\sum_{k\geq 1} \tau_k\right]$ is finite, and in this situation the crack grows to infinity within a finite time. We have

$$\mathbb{E}\left[\sum\limits_{k\geqslant 1} au_k
ight] = \sum\limits_{k\geqslant 1}\mathbb{E}[au_k] = \sum\limits_{k\geqslant 1}rac{1}{\lambda_k} = \sum\limits_{k\geqslant 1}rac{1}{(1+k)^
ho}$$

By comparison with the integral of the function $x \mapsto 1/(1+x)^{
ho}$ we get

$$\mathbb{E}\left[\sum\limits_{k\geqslant 1} au_k
ight] ~=~ \sum\limits_{k\geqslant 1}rac{1}{(1+k)^
ho} \ \leqslant~~ \sum\limits_{k\geqslant 1}\int_{k-1}^krac{1}{(1+x)^
ho}dx \ \leqslant~~ \int_0^\inftyrac{1}{(1+x)^
ho}dx \ =~~ rac{1}{1-
ho}\left[(1+x)^{1-
ho}
ight]_0^\infty$$

$$= \frac{1}{\rho - 1} < \infty,$$

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

$$\mathbb{E}\left[\sum\limits_{k\geqslant 1} au_k
ight] \ = \ \sum\limits_{k\geqslant 1}rac{1}{(1+k)^
ho} \ \geqslant \ \sum\limits_{k\geqslant 1}\int_k^{k+1}rac{1}{(1+x)^
ho}dx \ \geqslant \ \int_1^\inftyrac{1}{(1+x)^
ho}dx \ = \ \infty,$$

hence the mean time for the crack to grow to infinite length is infinite when $\rho \leq 1$.

Remark. This problem could also be treated in discrete time, assuming that τ_k is the (random) crack length increase at each step. In this case, the relation

$$\mathbb{E}\left[\sum\limits_{k \geqslant 1} au_k
ight] < \infty$$

for $\rho > 1$ shows that growth of the crack length to infinity cannot occur in this case.

Exercise 9.7

a) This time is the expected value of the third jump time T_3 , *i.e.*

$$\mathbb{E}[T_3] = \mathbb{E}[au_0] + \mathbb{E}[au_1] + \mathbb{E}[au_2] = rac{3}{\lambda} = 30 ext{ minutes}.$$

b) This probability is

$$\mathbb{P}(N_{60} < 3) = \mathbb{P}(N_{60} = 0) + \mathbb{P}(N_{60} = 1) + \mathbb{P}(N_{60} = 2) \ = e^{-60\lambda}(1 + 60\lambda + (60\lambda)^2/2) \ = 25 e^{-6} \simeq 0.062.$$

Exercise 9.8

5701

a) By the independence of increments of the Poisson process $(N_t)_{t\in\mathbb{R}_+}$ we find

$$\mathbb{P}(N_{t_3} = 5 \mid N_{t_1} = 1) = \frac{\mathbb{P}(N_{t_3} = 5 \text{ and } N_{t_1} = 1)}{\mathbb{P}(N_{t_1} = 1)}$$

$$= \frac{\mathbb{P}(N_{t_3} - N_{t_1} = 4 \text{ and } N_{t_1} = 1)}{\mathbb{P}(N_{t_1} = 1)} = \frac{\mathbb{P}(N_{t_3} - N_{t_1} = 4)\mathbb{P}(N_{t_1} = 1)}{\mathbb{P}(N_{t_1} = 1)}$$

$$= \mathbb{P}(N_{t_3} - N_{t_1} = 4) = \frac{(\lambda(t_3 - t_1))^4}{4!} e^{-\lambda(t_3 - t_1)}.$$

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

$$\begin{split} &\mathbb{E}[N_{t_1}N_{t_4}(N_{t_3}-N_{t_2})] \\ &= \mathbb{E}[N_{t_1}(N_{t_4}-N_{t_3})(N_{t_3}-N_{t_2})] + \mathbb{E}[N_{t_1}(N_{t_3}-N_{t_2})(N_{t_3}-N_{t_2})] \\ &+ \mathbb{E}[N_{t_1}(N_{t_2}-N_{t_1})(N_{t_3}-N_{t_2})] + \mathbb{E}[N_{t_1}N_{t_1}(N_{t_3}-N_{t_2})] \\ &= \mathbb{E}[N_{t_1}] \mathbb{E}[N_{t_4}-N_{t_3}] \mathbb{E}[N_{t_3}-N_{t_2}] + \mathbb{E}[N_{t_1}] \mathbb{E}[(N_{t_3}-N_{t_2})^2] \\ &+ \mathbb{E}[N_{t_1}] \mathbb{E}[N_{t_2}-N_{t_1}] \mathbb{E}[N_{t_3}-N_{t_2}] + \mathbb{E}[N_{t_1}^2] \mathbb{E}[N_{t_3}-N_{t_2}] \\ &= \lambda^3 t_1(t_4-t_3)(t_3-t_2) + \lambda^2 t_1(t_3-t_2)(1+\lambda(t_3-t_2))) \\ &+ \lambda^3 t_1(t_2-t_1)(t_3-t_2) + \lambda^2 t_1(1+\lambda t_1)(t_3-t_2). \end{split}$$

c) We have $\{T_2 > t\} = \{N_t \leqslant 1\}, t \in \mathbb{R}_+$, hence

$$\mathbb{E}[N_{t_2} \mid T_2 > t_1] = \mathbb{E}[N_{t_2} \mid N_{t_1} \leqslant 1] = rac{1}{\mathbb{P}(N_{t_1} \leqslant 1)} \mathbb{E}[N_{t_2}\mathbb{1}_{\{N_{t_1} \leqslant 1\}}],$$

by (1.6.6). Now, using the independence of increments between $N_{t_2} - N_{t_1}$ and N_{t_1} , we have

$$\mathbb{E}[N_{t_2}\mathbb{1}_{\{N_{t_1}\leqslant 1\}}] = \mathbb{E}[(N_{t_2} - N_{t_1})\mathbb{1}_{\{N_{t_1}\leqslant 1\}}] + \mathbb{E}[N_{t_1}\mathbb{1}_{\{N_{t_1}\leqslant 1\}}]$$

$$= \mathbb{E}[N_{t_2} - N_{t_1}] \mathbb{E}[\mathbb{1}_{\{N_{t_1} \leqslant 1\}}] + \mathbb{E}[N_{t_1}\mathbb{1}_{\{N_{t_1} \leqslant 1\}}] \\ = \mathbb{E}[N_{t_2} - N_{t_1}]\mathbb{P}(N_{t_1} \leqslant 1) + \mathbb{E}[N_{t_1}\mathbb{1}_{\{N_{t_1} \leqslant 1\}}],$$

hence

$$\begin{split} \mathbb{E}[N_{t_2} \mid T_2 > t_1] &= \mathbb{E}[N_{t_2} \mid N_{t_1} \leqslant 1] = \frac{1}{\mathbb{P}(N_{t_1} \leqslant 1)} \mathbb{E}[N_{t_2} \mathbb{1}_{\{N_{t_1} \leqslant 1\}}] \\ &= \frac{1}{\mathbb{P}(N_{t_1} \leqslant 1)} \mathbb{E}[(N_{t_2} - N_{t_1} + N_{t_1}) \mathbb{1}_{\{N_{t_1} \leqslant 1\}}] \\ &= \frac{1}{\mathbb{P}(N_{t_1} \leqslant 1)} \mathbb{E}[(N_{t_2} - N_{t_1}) \mathbb{1}_{\{N_{t_1} \leqslant 1\}}] + \frac{1}{\mathbb{P}(N_{t_1} \leqslant 1)} \mathbb{E}[N_{t_1} \mathbb{1}_{\{N_{t_1} \leqslant 1\}}] \\ &= \mathbb{E}[N_{t_2} - N_{t_1}] + \frac{1}{\mathbb{P}(N_{t_1} \leqslant 1)} \mathbb{E}[N_{t_1} \mathbb{1}_{\{N_{t_1} \leqslant 1\}}] \\ &= \mathbb{E}[N_{t_2} - N_{t_1}] + \frac{1}{\mathbb{P}(N_{t_1} \leqslant 1)} (0 \times \mathbb{P}(N_{t_1} = 0) + 1 \times \mathbb{P}(N_{t_1} = 1))) \\ &= \mathbb{E}[N_{t_2} - N_{t_1}] + \frac{\mathbb{P}(N_{t_1} = 1)}{\mathbb{P}(N_{t_1} \leqslant 1)} = \lambda(t_2 - t_1) + \frac{\lambda t_1 \mathrm{e}^{-\lambda t_1}}{\mathrm{e}^{-\lambda t_1} + \lambda t_1 \mathrm{e}^{-\lambda t_1}}. \end{split}$$

Exercise 9.9 The generator of the process is given by

$$Q = egin{bmatrix} -\lambda_0 & \lambda_0 & 0 \ \mu_1 & -\lambda_1 - \mu_1 & \lambda_1 \ 0 & \mu_2 & -\mu_2 \end{bmatrix} = egin{bmatrix} -2lpha & 2lpha & 0 \ eta & -lpha - eta & lpha \ 0 & 2eta & -2eta \end{bmatrix}.$$

Writing the equation $\pi Q = 0$ shows that

$$\pi_1=rac{2lpha}{eta}\pi_0 \ \ ext{and} \ \ \pi_2=rac{lpha}{2eta}\pi_1,$$

and the condition

$$\pi_0 + \pi_1 + \pi_2 = 1$$

shows that

$$\pi_0 = \left(1 + rac{lpha}{eta}
ight)^{-2}$$
MH3512 AY19-20

ॳ702

and

MH3512 AY19-20

$$\pi_0 = \left(rac{eta}{lpha+eta}
ight)^2, \quad \pi_1 = rac{2lphaeta}{(lpha+eta)^2} \quad ext{and} \quad \pi_2 = \left(rac{lpha}{lpha+eta}
ight)^2,$$

which is a binomial distribution with parameter $(2, p) = (2, \alpha/(\alpha + \beta))$, *i.e.*

$$\pi_0 = p^2, \quad \pi_1 = \binom{2}{1} p(1-p) \quad \text{and} \quad \pi_2 = (1-p)^2.$$

Exercise 9.10 This is an extension of Exercise 9.9. The generator of the process is given by

$$Q = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & \cdots & 0 & 0 & 0 \\ \mu_1 & -\lambda_1 - \mu_1 & \lambda_1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \mu_{N-1} & -\lambda_{N-1} - \mu_{N-1} & \lambda_{N-1} \\ 0 & 0 & 0 & \cdots & 0 & \mu_N & -\mu_N \end{bmatrix}$$
$$= \begin{bmatrix} -\alpha N & \alpha N & 0 & \cdots & 0 & 0 & 0 \\ \beta & -\alpha (N-1) - \beta & \alpha (N-1) & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \beta (N-1) & -\alpha - \beta (N-1) & \alpha \\ 0 & 0 & 0 & \cdots & 0 & \beta N & -\beta N \end{bmatrix}.$$

Writing the equation $\pi Q = 0$ shows that we have $-\alpha N \pi_0 + \beta \pi_1 = 0$,

$$\beta(k+1)\pi_{k+1} - (\alpha(N-k) + \beta k)\pi_k + \alpha(N-(k-1))\pi_{k-1} = 0, \quad k = 1, 2, \dots, N-1,$$

and $\alpha \pi_{N-1} - \beta N \pi_N = 0$, from which we deduce the recurrence relation

$$\pi_{k+1}=rac{lpha}{eta}rac{N-k}{k+1}\pi_k,\qquad k=0,1,\ldots,N-1,$$

and by induction on $k=1,2,\ldots,N$ we find $\pi_1=N\pi_0lpha/eta$ and

$$\pi_2=rac{lpha}{eta}rac{(N-1)}{2}\pi_1=\left(rac{lpha}{eta}
ight)^2rac{N(N-1)}{1 imes 2}\pi_0, \hspace{1em} \pi_3=\left(rac{lpha}{eta}
ight)^3rac{N(N-1)(N-2)}{1 imes 2 imes 3}\pi_0,$$

hence

$$egin{aligned} \pi_k &=& \left(rac{lpha}{eta}
ight)^k rac{N(N-1)\cdots(N-k+1)}{k!} \pi_0 \ &=& \left(rac{lpha}{eta}
ight)^k rac{N!}{(N-k)!k!} \pi_0 = \left(rac{lpha}{eta}
ight)^k inom{N}{k} \pi_0, \end{aligned}$$

 $k = 0, 1, \dots, N$. The condition

$$\pi_0+\pi_1+\dots+\pi_N=1$$

shows that

$$1=\pi_0\sum\limits_{k=0}^N \left(rac{lpha}{eta}
ight)^k rac{N!}{(N-k)!k!}=\left(1+rac{lpha}{eta}
ight)^N \pi_0,$$

hence

$$\pi_0 = \left(1+rac{lpha}{eta}
ight)^{-N}$$

and we have

$$egin{aligned} \pi_k &= \left(1+rac{lpha}{eta}
ight)^{-N} \left(rac{lpha}{eta}
ight)^k rac{N!}{(N-k)!k!} \ &= \left(rac{lpha}{lpha+eta}
ight)^k \left(rac{eta}{lpha+eta}
ight)^{N-k} rac{N!}{(N-k)!k!}, \qquad k=0,1,\ldots,N, \end{aligned}$$

hence the stationary distribution π is a binomial distribution with parameter $(N, p) = (N, \alpha/(\alpha + \beta))$.

hence the forward Kolmogorov equation P'(t) = P(t)Q reads

$$\begin{bmatrix} P_{0,0}'(t) & P_{0,1}'(t) & P_{0,2}'(t) & P_{0,3}'(t) & \cdots \\ P_{1,0}'(t) & P_{1,1}'(t) & P_{1,2}'(t) & P_{1,3}'(t) & \cdots \\ P_{2,0}'(t) & P_{2,1}'(t) & P_{2,2}'(t) & P_{2,3}'(t) & \cdots \\ P_{3,0}'(t) & P_{3,1}'(t) & P_{3,2}'(t) & P_{3,3}(t) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \\ = \begin{bmatrix} P_{0,0}(t) & P_{0,1}(t) & P_{0,2}(t) & P_{0,3}(t) & \cdots \\ P_{1,0}(t) & P_{1,1}(t) & P_{1,2}(t) & P_{1,3}(t) & \cdots \\ P_{2,0}(t) & P_{2,1}(t) & P_{2,2}(t) & P_{2,3}(t) & \cdots \\ P_{3,0}(t) & P_{3,1}(t) & P_{3,2}(t) & P_{3,3}(t) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \times \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & \cdots \\ 0 & -3 & 3 & 0 & 0 & \cdots \\ 0 & 0 & -2 & 2 & 0 & \cdots \\ 0 & 0 & 0 & -5 & 5 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix},$$

which yields

$$\left\{egin{array}{l} P_{0,0}'(t) = -P_{0,0}(t), \ P_{0,1}'(t) = P_{0,0}(t) - 3P_{0,1}(t), \ P_{0,2}'(t) = 3P_{0,1}(t) - 2P_{0,2}(t), \ P_{0,3}'(t) = 2P_{0,2}(t) - 5P_{0,3}(t). \end{array}
ight.$$

The first equation is solved by (A.9) as

$$P_{0,0}(t)=P_{0,0}(0){
m e}^{-t}={
m e}^{-t}, \qquad t\in {\mathbb R}_+,$$

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

$$P_{0,1}'(t) = e^{-t} - 3P_{0,1}(t), \tag{B.50}$$

its associated homogeneous equation is $P'_{0,1}(t) = -3P_{0,1}(t)$, which has $t \mapsto C_1 e^{-3t}$ for homogeneous solution. By substituting the particular solution $t \mapsto C_2 e^{-t}$ in (B.50) we get

$$-C_2 \mathrm{e}^{-t} = \mathrm{e}^{-t} - 3C_2 \mathrm{e}^{-t},$$

i.e. $C_2 = 1/2$, hence the general solution

$$P_{0,1}(t) = rac{1}{2} \mathrm{e}^{-t} + C_1 \mathrm{e}^{-3t},$$

and from the initial condition $P_{0,1}(0) = 0$ we get $C_1 = -1/2$, *i.e.*

$$P_{0,1}(t) = rac{1}{2}(\mathrm{e}^{-t} - \mathrm{e}^{-3t}), \qquad t \in \mathbb{R}_+.$$

The remaining equations can be solved similarly by searching for a suitable particular solution. For

$$P_{0,2}'(t) = rac{3}{2}(\mathrm{e}^{-t} - \mathrm{e}^{-3t}) - 2P_{0,2}(t),$$

we find, searching for a particular solution of the form $t \mapsto ae^{-t} + be^{-3t}$,

$$P_{0,2}(t) = \frac{3}{2} e^{-3t} (1 - e^t)^2, \qquad t \in \mathbb{R}_+,$$
(B.51)

see *here*, and for

 $P_{0,3}'(t) = 3e^{-3t}(1 - e^t)^2 - 5P_{0,3}(t),$

we find

$$P_{0,3}(t) = \frac{1}{4} e^{-5t} (e^t - 1)^3 (1 + 3e^t), \qquad t \in \mathbb{R}_+,$$
(B.52)

see *here*. Note that using the *backward* Kolmogorov equation P'(t) = QP(t) can lead to more complicated calculations.

Probabilistic approach

The above results may be recovered by a probabilistic approach, with the change of variable z = x + y, as

$$\begin{split} P_{0,1}(t) &= \mathbb{P}(X_t = 1 \mid X_0 = 0) \\ &= \mathbb{P}(\tau_0 < t \text{ and } \tau_0 + \tau_1 > t) \\ &= \mathbb{P}(\tau_0 < t < \tau_0 + \tau_1) \\ &= \lambda_0 \lambda_1 \int_{\{(x,y) : \ 0 < x < t, \ x + y > t\}} e^{-\lambda_0 x} e^{-\lambda_1 y} dx dy \\ &= \lambda_0 \lambda_1 \int_0^t \int_t^\infty e^{-\lambda_0 x} e^{-\lambda_1 (z - x)} dx dz \\ &= \lambda_0 \lambda_1 \int_0^t \int_t^\infty e^{-(\lambda_0 - \lambda_1) x} e^{-\lambda_1 z} dx dz \\ &= e^{-\lambda_1 t} \lambda_0 \int_0^t e^{-(\lambda_0 - \lambda_1) x} dx \\ &= \frac{\lambda_0}{\lambda_0 - \lambda_1} (e^{-\lambda_1 t} - e^{-\lambda_0 t}) \\ &= \frac{1}{2} (e^{-t} - e^{-3t}), \qquad t \in \mathbb{R}_+. \end{split}$$

As for $P_{0,2}(t)$ it also reads

$$P_{0,2}(t) = \mathbb{P}(X_t = 2 \mid X_0 = 0) = \mathbb{P}(\tau_0 + \tau_1 < t \text{ and } \tau_0 + \tau_1 + \tau_2 > t),$$

and (B.51) can be recovered via a triple integral. Similarly, (B.52) can be approached by a quadruple integral, etc. Exercise 9.12 Noting that the two events

$$\{T_1>t,\;T_2>t+s\}=\{X_t=0,\;0\leqslant X_{t+s}\leqslant 1\}$$

coincides for all $s, t \in \mathbb{R}_+$, we find that

$$\mathbb{P}(T_1 > t \text{ and } T_2 > t + s) = \mathbb{P}(X_t = 0 \text{ and } X_{t+s} \in \{0, 1\} \mid X_0 = 0)$$

= $\mathbb{P}(X_{t+s} \in \{0, 1\} \mid X_t = 0) \mathbb{P}(X_t = 0 \mid X_0 = 0)$
= $\mathbb{P}(X_s \in \{0, 1\} \mid X_0 = 0) \mathbb{P}(X_t = 0 \mid X_0 = 0)$
= $(\mathbb{P}(X_s = 0 \mid X_0 = 0) + \mathbb{P}(X_s = 1 \mid X_0 = 0)) \mathbb{P}(X_t = 0 \mid X_0 = 0)$

MH3512 AY19-20

(B.53)

$$= P_{0,0}(t)(P_{0,0}(s) + P_{0,1}(s)).$$

5708

Next, we note that we have

$$P_{0,0}(t)=\mathrm{e}^{-\lambda_0 t},\qquad t\in\mathbb{R}_+,$$

and, from Relation (B.53) above,

$$P_{0,1}(t) = rac{\lambda_0}{\lambda_1 - \lambda_0} \left(\mathrm{e}^{-\lambda_0 t} - \mathrm{e}^{-\lambda_1 t}
ight), \qquad t \in \mathbb{R}_+,$$

hence

$$\begin{split} \mathbb{P}(T_1 > t \text{ and } T_2 > t + s) &= \mathrm{e}^{-\lambda_0(t+s)} + \frac{\lambda_0}{\lambda_1 - \lambda_0} \left(\mathrm{e}^{-\lambda_0(t+s)} - \mathrm{e}^{-\lambda_0 t - \lambda_1 s} \right) \\ &= \frac{\lambda_1}{\lambda_1 - \lambda_0} \mathrm{e}^{-\lambda_0(t+s)} - \frac{\lambda_0}{\lambda_1 - \lambda_0} \mathrm{e}^{-\lambda_0 t - \lambda_1 s}. \end{split}$$

Then, since

$$\mathbb{P}(T_1 > x ext{ and } T_2 > y) = \int_x^\infty \int_y^\infty f_{(T_1,T_2)}(u,v) du dv,$$

by (1.5.4) we get

$$egin{aligned} f_{(T_1,T_2)}(x,y) &= rac{\partial^2}{\partial y \partial x} \mathbb{P}(T_1 > x ext{ and } T_2 > y) \ &= rac{\partial^2}{\partial y \partial x} \left(rac{\lambda_1}{\lambda_1 - \lambda_0} \mathrm{e}^{-\lambda_0 y} - rac{\lambda_0}{\lambda_1 - \lambda_0} \mathrm{e}^{-\lambda_0 x - \lambda_1 (y - x)}
ight) \ &= -\lambda_0 rac{\partial}{\partial y} \mathrm{e}^{-\lambda_0 x - \lambda_1 (y - x)} = \lambda_0 \lambda_1 \mathrm{e}^{-\lambda_0 x - \lambda_1 (y - x)}, \end{aligned}$$

provided that $y \ge x \ge 0$. When $x > y \ge 0$ 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

$$f_{(au_0, au_1)}(s,t)=f_{(T_1,T_2)}(s,s+t)=\lambda_0\lambda_1\mathrm{e}^{-\lambda_0s-\lambda_1t},\quad s,t\in\mathbb{R}_+,$$

which shows that τ_0 , τ_1 are two independent exponentially distributed random variables with parameters λ_0 and λ_1 , respectively. 709

Exercise 9.13 Let $(N_t)_{t \in \mathbb{R}_+}$ denote a Poisson process with intensity $\lambda > 0$.

a) This probability is equal to

$$\mathbb{P}(N_T=0)=\mathbb{P}(au_0>T)=\mathrm{e}^{-\lambda T}$$

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 (1.6.12))

$$egin{aligned} \mathbb{E}[au_0 \mid au_0 < T] &= rac{1}{\mathbb{P}(au_0 < T)} \mathbb{E}[au_0 \mathbbm{1}_{\{ au_0 < T\}} \ &= rac{\lambda}{1-\mathrm{e}^{-\lambda T}} \int_0^T x \mathrm{e}^{-\lambda x} dx \ &= rac{1-(1+\lambda T)\mathrm{e}^{-\lambda T}}{\lambda(1-\mathrm{e}^{-\lambda T})} \end{aligned}$$

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

$$t = 0 \times \mathbb{P}(N_T = 0) + (\mathbb{E}[\tau_0 \mid \tau_0 < T] + t) \times \mathbb{P}(\tau_0 \leqslant T)$$

= $\frac{1 - (1 + \lambda T)e^{-\lambda T}}{\lambda} + t(1 - e^{-\lambda T})$ (B.54)

with unknown \boldsymbol{t} , and solution

$$t = rac{e^{\lambda T} - 1 - \lambda T}{\lambda}.$$

Alternatively, we could also rewrite the above equation (B.54) as

$$t = \mathbb{E}[0 imes \mathbbm{1}_{\{ au_0 \geqslant T\}} + (au_0 + t) \mathbbm{1}_{\{ au_0 < T\}}] = rac{1 - (1 + \lambda T) \mathrm{e}^{-\lambda T}}{\lambda} + t(1 - \mathrm{e}^{-\lambda T}).$$

c) Denoting by t the mean time until she finishes crossing the street we have, by first step analysis in continuous time,

$$egin{aligned} t &= & \mathbb{E}\left[T\mathbbm{1}_{\{T < au_0\}} + (au_0 + t)\mathbbm{1}_{\{T \geqslant au_0\}}
ight] \ &= & \lambda T \int_T^\infty \mathrm{e}^{-\lambda s} ds + \lambda \int_0^T (s+t) \mathrm{e}^{-\lambda s} ds \ &= & T\mathbb{P}(au_0 > T) + \lambda \int_0^T s \mathrm{e}^{-\lambda s} ds + t\mathbb{P}(au_0 < T) \ &= & T \mathrm{e}^{-\lambda T} + rac{1 - \mathrm{e}^{-\lambda T}(1+\lambda T)}{\lambda} + t(1 - \mathrm{e}^{-\lambda T}), \end{aligned}$$

which yields

$$t = \frac{e^{\lambda T} - 1}{\lambda}.$$

Alternatively, we can write

$$egin{aligned} t &= T \mathbb{P}(au_0 > T) + ig(\mathbb{E}[au_0 \mid au_0 > T] + tig) \mathbb{P}(au_0 > T) \ &= T \mathbb{P}(au_0 > T) + \mathbb{E}[au_0 \mathbb{1}_{\{ au_0 > T\}}] + t \mathbb{P}(au_0 > T) \ &= T \mathbb{P}(au_0 > T) + \lambda \int_0^T s \mathrm{e}^{-\lambda s} ds + t \mathbb{P}(au_0 < T), \end{aligned}$$

which yields the same result.

d) In this case, T becomes an independent exponentially distributed random variable with parameter $\mu > 0$, hence we can write

$$t = \mathbb{E}\left[\frac{e^{\lambda T} - 1}{\lambda}\right]$$
$$= \mu \int_0^\infty \frac{e^{\lambda u} - 1}{\lambda} e^{-\mu u} du$$
$$= \frac{\mu}{\lambda} \int_0^\infty e^{(\lambda - \mu)u} du - \frac{\mu}{\lambda} \int_0^\infty e^{-\mu u} du$$
$$= \frac{\mu}{\lambda(\mu - \lambda)} - \frac{1}{\lambda}$$
$$= \frac{1}{\mu - \lambda}$$

MH3512 AY19-20

5710

if $\mu > \lambda$, with $t = +\infty$ if $\mu \leq \lambda$.

Exercise 9.14

- a) This probability is the probability that an exponential random variable with parameter μ is lower than T, *i.e.* $\mathbb{P}(\tau_0 > T) = e^{-\mu T}$.
- b) Denoting by $t = \mathbb{E}[W]$ the mean time until the machine breaks down we have, by first step analysis in continuous time,

$$egin{aligned} t &= & \mathbb{E}\left[T\mathbbm{1}_{\{T < au_0\}} + (au_0 + t)\mathbbm{1}_{\{T \geqslant au_0\}}
ight] \ &= & \mu T \int_T^\infty \mathrm{e}^{-\mu s} ds + \mu \int_0^T (s+t) \mathrm{e}^{-\mu s} ds \ &= & T\mathbb{P}(au_0 > T) + \mu \int_0^T s \mathrm{e}^{-\mu s} ds + t\mathbb{P}(au_0 < T) \ &= & T\mathrm{e}^{-\mu T} + rac{1 - \mathrm{e}^{-\mu T}(1+\mu T)}{\mu} + t(1 - \mathrm{e}^{-\mu T}), \end{aligned}$$

which yields

$$t = \mathbb{E}[W] = rac{e^{\mu T} - 1}{\mu}.$$

Alternatively, we can write

$$egin{aligned} t &= T \mathbb{P}(au_0 > T) + ig(\mathbb{E} \left[au_0 \mid au_0 > T
ight] + t ig) \mathbb{P}(au_0 > T) \ &= T \mathbb{P}(au_0 > T) + \mathbb{E} \left[au_0 \mathbbm{1}_{\{ au_0 > T\}}
ight] + t \mathbb{P}(au_0 > T) \ &= T \mathbb{P}(au_0 > T) + \mu \int_0^T s \mathrm{e}^{-\mu s} ds + t \mathbb{P}(au_0 < T), \end{aligned}$$

which yields the same result.

c) This proportion is

$$\frac{\mathbb{E}[W]}{\mathbb{E}[W]+1/\lambda} = \frac{\mathrm{e}^{\mu T}-1}{e^{\mu T}-1+\mu/\lambda}.$$

The state of the machine can be modeled using a two-state Markov chain with infinitesimal generator

$$Q = \left[egin{array}{cc} -\lambda & \lambda \ rac{1}{\mathbb{E}[W]} & -rac{1}{\mathbb{E}[W]} \end{array}
ight] = \left[egin{array}{cc} -\lambda & \lambda \ \mu \ rac{\mu}{\mathrm{e}^{\mu T}-1} & rac{\mu}{1-\mathrm{e}^{\mu T}} \end{array}
ight],$$

and the above proportion can be rewritten using the stationary distribution

$$\begin{split} [\pi_0,\pi_1] &= \left[\frac{1/\lambda}{1/\lambda + (\mathrm{e}^{\mu T}-1)/\mu}, \frac{(\mathrm{e}^{\mu T}-1)/\mu}{1/\lambda + (\mathrm{e}^{\mu T}-1)/\mu}\right] \\ &= \left[\frac{\mu}{\mu + \lambda (\mathrm{e}^{\mu T}-1)}, \frac{\lambda (\mathrm{e}^{\mu T}-1)}{\mu + \lambda (\mathrm{e}^{\mu T}-1)}\right] \\ &= \left[\frac{1/\lambda}{\mathrm{E}[W]+1/\lambda}, \frac{\mathrm{E}[W]}{\mathrm{E}[W]+1/\lambda}\right]. \end{split}$$

Exercise 9.15

a) The generator Q of $(X_t)_{t \in \mathbb{R}_+}$ is given by

$$Q = \left[egin{array}{cccc} -0.5 & 0.5 & 0 \ 0.2 & -0.45 & 0.25 \ 0 & 0.4 & -0.4 \end{array}
ight].$$

b) Solving for $\pi Q = 0$ we have

$$egin{aligned} \pi Q &= & [\pi_0,\pi_1,\pi_2] imes egin{bmatrix} -0.5 & 0.5 & 0 \ 0.2 & -0.45 & 0.25 \ 0 & 0.4 & -0.4 \end{bmatrix} \ &= & egin{bmatrix} -0.5 imes \pi_0 + 0.2 imes \pi_1 \ 0.5 imes \pi_0 - 0.45 imes \pi_1 + 0.4 imes \pi_2 \ 0.25 imes \pi_1 - 0.4 imes \pi_2 \end{bmatrix}^T = [0,0,0], \end{aligned}$$

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

$$0 \times \pi_0 + 1 \times \pi_1 + 2 \times \pi_2 = \frac{40}{81} + \frac{50}{81} = \frac{90}{81}.$$

d) We find

$$100 imes rac{90}{81} = rac{1000}{9}.$$
 MH3512 AY19-20

e) We have

$$Q = \left[egin{array}{cccc} -0.5 & 0.5 & 0 \ 0.2 & -0.7 & 0.5 \ 0 & 0.4 & -0.4 \end{array}
ight],$$

and solving $\pi Q = 0$ shows that

$$[\pi_0, \pi_1, \pi_2] = \left[\frac{0.32}{2.12}, \frac{0.8}{2.12}, \frac{1}{2.12} \right] = \left[0.15094, 0.37736, 0.47170 \right].$$

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

$$Q = \left[egin{array}{cc} -\lambda & \lambda \ \mu & -\mu \end{array}
ight].$$

The infinitesimal generator of $Z(t) := X_1(t) + X_2(t)$ is given by

$$egin{bmatrix} -2\lambda & 2\lambda & 0 \ \mu & -\lambda-\mu & \lambda \ 0 & 2\mu & -2\mu \end{bmatrix},$$

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).

Note that the generator of the sum $(Z(t))_{t \in \mathbb{R}_+}$ of $(X_1(t))_{t \in \mathbb{R}_+}$ and $(X_2(t))_{t \in \mathbb{R}_+}$ is not the sum of the matrix generators of $(X_1(t))_{t \in \mathbb{R}_+}$ and $(X_2(t))_{t \in \mathbb{R}_+}$. Recall that by Proposition 9.6, the semi-groups of $X_1(t)$ and $X_2(t)$ are given by

$$\begin{bmatrix} \mathbb{P}(X_{1}(t) = 0 \mid X_{1}(0) = 0) & \mathbb{P}(X_{1}(t) = 1 \mid X_{1}(0) = 0) \\ \mathbb{P}(X_{1}(t) = 0 \mid X_{1}(0) = 1) & \mathbb{P}(X_{1}(t) = 1 \mid X_{1}(0) = 1) \end{bmatrix}$$
$$= \begin{bmatrix} \mathbb{P}(X_{2}(t) = 0 \mid X_{2}(0) = 0) & \mathbb{P}(X_{2}(t) = 1 \mid X_{2}(0) = 0) \\ \mathbb{P}(X_{2}(t) = 0 \mid X_{2}(0) = 1) & \mathbb{P}(X_{2}(t) = 1 \mid X_{2}(0) = 1) \end{bmatrix}$$
MH3512 AY19-20

$$= \begin{bmatrix} \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-t(\lambda+\mu)} & \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu} e^{-t(\lambda+\mu)} \\ \frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu} e^{-t(\lambda+\mu)} & \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} e^{-t(\lambda+\mu)} \end{bmatrix}$$

As for the transition semi-group of Z(t), we have

$$\begin{aligned} P_{0,0}(t) &= \mathbb{P}(Z(t) = 0 \mid Z(0) = 0) \\ &= \mathbb{P}(X_1(t) = 0 \text{ and } X_2(t) = 0 \mid X_1(0) = 0 \text{ and } X_2(0) = 0) \\ &= \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 0) \\ &= (\mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0))^2 \\ &= \left(\frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-t(\lambda + \mu)}\right)^2. \end{aligned}$$

٠

For $P_{0,1}(t)$ we have

$$\begin{aligned} P_{0,1}(t) &= \mathbb{P}(Z(t) = 1 \mid Z(0) = 0) \\ &= \mathbb{P}(X_1(t) = 0 \text{ and } X_2(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 0) \\ &+ \mathbb{P}(X_1(t) = 1 \text{ and } X_2(t) = 0 \mid X_1(0) = 0 \text{ and } X_2(0) = 0) \\ &= \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 0) \\ &+ \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 1 \mid X_2(0) = 0) \\ &= 2\mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 0). \end{aligned}$$

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. We have

$$\mathbb{P}(Z(t) = 1 \mid Z(0) = 1) \\
= \mathbb{P}(Z(t) = 1 \mid \{X_1(0) = 0 \text{ and } X_2(0) = 1\} \cup \{X_1(0) = 1 \text{ and } X_2(0) = 0\}) \\
= \frac{\mathbb{P}(\{Z(t) = 1\} \cap (\{X_1(0) = 0 \text{ and } X_2(0) = 1\} \cup \{X_1(0) = 1 \text{ and } X_2(0) = 0\}))}{\mathbb{P}(\{X_1(0) = 0 \text{ and } X_2(0) = 1\} \cup \{X_1(0) = 1 \text{ and } X_2(0) = 0\})}$$

$$= \frac{\mathbb{P}(\{Z(t) = 1\} \cap \{X_1(0) = 0 \text{ and } X_2(0) = 1\})}{\mathbb{P}(Z(0) = 1)} + \frac{\mathbb{P}(\{Z(t) = 1\} \cap \{X_1(0) = 1 \text{ and } X_2(0) = 0\})}{\mathbb{P}(Z(0) = 1)}$$

$$= \mathbb{P}(Z(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1) \frac{\mathbb{P}(X_1(0) = 0 \text{ and } X_2(0) = 1)}{\mathbb{P}(Z(0) = 1)} + \mathbb{P}(Z(t) = 1 \mid X_1(0) = 1 \text{ and } X_2(0) = 0) \frac{\mathbb{P}(X_1(0) = 1 \text{ and } X_2(0) = 0)}{\mathbb{P}(Z(0) = 1)}$$

$$= \mathbb{P}(Z(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1) \frac{\mathbb{P}(X_1(0) = 0 \text{ and } X_2(0) = 1)}{\mathbb{P}(Z(0) = 1)} + \mathbb{P}(Z(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1) \frac{\mathbb{P}(X_1(0) = 1 \text{ and } X_2(0) = 1)}{\mathbb{P}(Z(0) = 1)}$$

$$= \mathbb{P}(Z(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1) \frac{\mathbb{P}(X_1(0) = 1 \text{ and } X_2(0) = 0)}{\mathbb{P}(Z(0) = 1)}$$

The above equality

$$\mathbb{P}(Z(t) = 1 \mid X_1(0) = 1 \text{ and } X_2(0) = 0) = \mathbb{P}(Z(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1)$$

is justified by the fact that the transition probabilities of $(Z(t))_{t \in \mathbb{R}_+}$ depend only on $Z(0) = X_1(0) + X_2(0) = 0 + 1 = 1 + 0 = 1$.

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

$$P_{1,1}(t) = \mathbb{P}(Z(t) = 1 \mid Z(0) = 1)$$

= $\mathbb{P}(X_1(t) = 0 \text{ and } X_2(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1)$
+ $\mathbb{P}(X_1(t) = 1 \text{ and } X_2(t) = 0 \mid X_1(0) = 0 \text{ and } X_2(0) = 1)$
= $\mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0)\mathbb{P}(X_2(t) = 1 \mid X_2(0) = 1)$
+ $\mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0)\mathbb{P}(X_2(t) = 0 \mid X_2(0) = 1).$

Note that in the above calculation we chose to represent the event $\{Z(0) = 1\}$ by $\{X_1(0) = 0 \text{ and } X_2(0) = 1\}$, however making the other choice $\{X_1(0) = 1 \text{ and } X_2(0) = 0\}$ would lead to the same result because $X_1(t)$ and $X_2(t)$ have same infinitesimal generators.

♂715

Concerning $P_{1,0}(t)$ we have

$$P_{1,0}(t) = \mathbb{P}(Z(t) = 0 \mid Z(0) = 1)$$

= $\mathbb{P}(X_1(t) = 0 \text{ and } X_2(t) = 0 \mid X_1(0) = 0 \text{ and } X_2(0) = 1).$

On the other hand, we have

$$P_{1,2}(t) = \mathbb{P}(Z(t) = 2 \mid Z(0) = 1)$$

= $\mathbb{P}(X_1(t) = 1 \text{ and } X_2(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1)$
= $\mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0)\mathbb{P}(X_2(t) = 1 \mid X_2(0) = 1).$

We check that

$$\begin{split} P_{1,0}(t) + P_{1,1}(t) + P_{1,2}(t) &= \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 1) \\ &+ \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 1 \mid X_2(0) = 1) \\ &+ \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 1 \mid X_2(0) = 1) \\ &= (\mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0) + \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0)) \\ &\times (\mathbb{P}(X_2(t) = 1 \mid X_2(0) = 1) + \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 1)) \\ &= 1. \end{split}$$

Note that for Z(t) to be Markov, the processes $X_1(t)$ and $X_2(t)$ should have same infinitesimal generators. For example, if $X_1(t)$ and $X_2(t)$ have different transition rates, then starting from Z(t) = 1 we need the information whether $X_1(t) = 1$ or $X_2(t) = 1$ in order to determine what will be the next transition rate. However, the knowledge of Z(t) = 1 is not sufficient for this.

Altogether, there are $3 \times 3 = 9$ transition probabilities to compute since the chain Z(t) has 3 states $\{0, 1, 2\}$, and the remaining computations are left to the reader.

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_{20}

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

$$egin{aligned} \mathbb{E}\left[T_0^r \mid X_0 = 1
ight] &= lpha \left(rac{1}{\lambda} + \mathbb{E}\left[T_0^r \mid X_0 = 1
ight]
ight) + (1-lpha) imes \left(rac{1}{\lambda} + 0
ight) \ &= rac{1}{\lambda} + lpha \mathbb{E}\left[T_0^r \mid X_0 = 1
ight], \end{aligned}$$

hence

$$\mathbb{E}\left[T_0^r \mid X_0 = 1\right] = \frac{1}{\lambda(1 - \alpha)} \tag{B.55}$$

and

$$\mathbb{E}\left[T_0^r \mid X_0=0
ight] = rac{1}{\lambda} + 1 imes \mathbb{E}\left[T_0^r \mid X_0=1
ight] = rac{2-lpha}{\lambda(1-lpha)}.$$

Note that (B.55) 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

$$\mathbb{E}\left[T_1^r \mid X_0 = 1
ight] \;=\; rac{1}{\lambda} + lpha \mathbb{E}\left[T_1^r \mid X_0 = 1
ight] + (1 - lpha) \mathbb{E}\left[T_1^r \mid X_0 = 0
ight] \ =\; rac{2 - lpha}{\lambda} + lpha \mathbb{E}\left[T_1^r \mid X_0 = 1
ight],$$

since $\mathbb{E}\left[T_1^r \mid X_0 = 0\right] = 1/\lambda$, hence

$$\mathbb{E}\left[T_1^r \mid X_0 = 1
ight] = rac{2-lpha}{\lambda(1-lpha)}.$$

c) This continuous-time first step analysis argument is similar to the one used in the solution of Exercise 9.3. Since

$$\mathbb{E}\left[T_0^r \mid X_0 = 1
ight] = rac{1}{\lambda(1-lpha)},$$
 MH3512 AY19-20

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

$$rac{1}{\mathbb{E}\left[T_0^r \mid X_0 = 1
ight]} = \lambda(1-lpha),$$

and the infinitesimal generator Q of X_t is

$$egin{bmatrix} -\lambda & \lambda \ 1 & -rac{1}{\mathbb{E}\left[T_0^r \mid X_0=1
ight]} & -rac{\lambda}{\mathbb{E}\left[T_0^r \mid X_0=1
ight]} \end{bmatrix} = egin{bmatrix} -\lambda & \lambda \ (1-lpha)\lambda & -(1-lpha)\lambda \end{bmatrix}.$$

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

$$egin{aligned} & \left(N_{t_n}^1+N_{t_n}^2-(N_{t_{n-1}}^1+N_{t_{n-1}}^2),\ldots,N_{t_2}^1+N_{t_2}^2-(N_{t_1}^1+N_{t_1}^2)
ight)\ &=\left(N_{t_n}^1-N_{t_{n-1}}^1+N_{t_n}^2-N_{t_{n-1}}^2,\ldots,N_{t_2}^1-N_{t_1}^1+N_{t_2}^2-N_{t_1}^2
ight) \end{aligned}$$

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

$$N_{t_{n-1}}^1 - N_{t_{n-2}}^1, \dots, N_{t_2}^1 - N_{t_1}^1,$$

and of

$$N_{t_n}^2 - N_{t_{n-1}}^2, \dots, N_{t_2}^2 - N_{t_1}^2,$$
 MH3512 AY19-20

hence it is also independent of

$$N_{t_{n-1}}^1 - N_{t_{n-2}}^1 + N_{t_{n-1}}^2 - N_{t_{n-2}}^2, \dots, N_{t_2}^1 - N_{t_1}^1 + N_{t_2}^2 - N_{t_1}^2$$

Similarly it follows that $N_{t_n}^2 - N_{t_{n-1}}^2$ is independent of

$$N_{t_{n-1}}^1 - N_{t_{n-2}}^1 + N_{t_{n-1}}^2 - N_{t_{n-2}}^2, \dots, N_{t_2}^1 - N_{t_1}^1 + N_{t_2}^2 - N_{t_1}^2,$$

hence $N_{t_n}^1 + N_{t_n}^2 - (N_{t_{n-1}}^1 + N_{t_{n-1}}^2)$ is independent of

$$N_{t_{n-1}}^1 - N_{t_{n-2}}^1 + N_{t_{n-1}}^2 - N_{t_{n-2}}^2, \dots, N_{t_2}^1 - N_{t_1}^1 + N_{t_2}^2 - N_{t_1}^2$$

This shows the required mutual independence by induction on $n \ge 1$.

(iii) The process $(N_t^1 + N_t^2)_{t \in \mathbb{R}_+}$ has stationary increments.

We note that the distributions of the random variables $N_{t+h}^1 - N_{s+h}^1$ and $N_{t+h}^2 - N_{s+h}^2$ do not depend on $h \in \mathbb{R}_+$, hence by the law of total probability we check that

$$egin{aligned} \mathbb{P}(N_{t+h}^1+N_{t+h}^2-(N_{s+h}^1+N_{s+h}^2)&=n)\ &=&\sum\limits_{k=0}^n \mathbb{P}(N_{t+h}^1-N_{s+h}^1&=k)\mathbb{P}(N_{t+h}^2-N_{s+h}^2&=n-k) \end{aligned}$$

is independent of $h \in \mathbb{R}_+$.

The intensity of $N_t^1 + N_t^2$ 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

$$egin{aligned} \mathbb{P}(M_{t+h}-M_t=n) &= \mathbb{P}(N_{t+h}^1-N_{t+h}^2-(N_{s+h}^1-N_{s+h}^2)=n) \ &= \mathbb{P}(N_{t+h}^1-N_{s+h}^1-(N_{t+h}^2-N_{s+h}^2)=n) \ &= \sum\limits_{k \geqslant 0} \mathbb{P}(N_{t+h}^1-N_{s+h}^1=n+k) \mathbb{P}(N_{t+h}^2-N_{s+h}^2=k) \end{aligned}$$

which is independent of $h \in \mathbb{R}_+$ since the distributions of $N_{t+h}^1 - N_{s+h}^1$ and $N_{t+h}^2 - N_{s+h}^2$ are independent of $h \in \mathbb{R}_+$, 720

c) For $n \in \mathbb{N}$ we have

$$egin{aligned} \mathbb{P}(M_t = n) &= \mathbb{P}(N_t^1 - N_t^2 = n) \ &= \sum\limits_{k \geqslant \max(0, -n)} \mathbb{P}(N_t^1 = n + k) \mathbb{P}(N_t^2 = k) \ &= \mathrm{e}^{-(\lambda_1 + \lambda_2)t} \sum\limits_{k \geqslant \max(0, -n)} rac{\lambda_1^{n+k} \lambda_2^k t^{n+2k}}{k! (n+k)!} \ &= \left(rac{\lambda_1}{\lambda_2}
ight)^{n/2} \mathrm{e}^{-(\lambda_1 + \lambda_2)t} \sum\limits_{k \geqslant \max(0, -n)} rac{(t\sqrt{\lambda_1 \lambda_2})^{n+2k}}{(n+k)!k!} \ &= \left(rac{\lambda_1}{\lambda_2}
ight)^{n/2} \mathrm{e}^{-(\lambda_1 + \lambda_2)t} I_{|n|}(2t\sqrt{\lambda_1 \lambda_2}), \end{aligned}$$

where

$$I_n(x)=\sum\limits_{k\geqslant 0}rac{(x/2)^{n+2k}}{k!(n+k)!},\qquad x>0,$$

is the modified Bessel function with parameter $n \ge 0$. When $n \le 0$, by exchanging λ_1 and λ_2 we get

$$egin{aligned} \mathbb{P}(M_t = n) &= \mathbb{P}(-M_t = -n) \ &= \left(rac{\lambda_2}{\lambda_1}
ight)^{-n/2} \mathrm{e}^{-(\lambda_1 + \lambda_2)t} I_{-n}(2t\sqrt{\lambda_1\lambda_2}) \ &= \left(rac{\lambda_1}{\lambda_2}
ight)^{n/2} \mathrm{e}^{-(\lambda_1 + \lambda_2)t} I_{-n}(2t\sqrt{\lambda_1\lambda_2}), \end{aligned}$$

hence in the general case we have

$$\mathbb{P}(M_t=n)=\left(rac{\lambda_1}{\lambda_2}
ight)^{n/2}\mathrm{e}^{-(\lambda_1+\lambda_2)t}I_{|n|}(2t\sqrt{\lambda_1\lambda_2}),\qquad n\in\mathbb{Z},$$

which is known as the Skellam distribution.
This also shows that the semigroup P(t) of the birth and death process with state-independent birth and death rates λ_1 and λ_2^1 satisfies

$$egin{aligned} P_{i,j}(t) &=& \mathbb{P}(M_t=j\mid M_0=i) \ &=& \mathbb{P}(M_t=j-i) \ &=& \left(rac{\lambda_1}{\lambda_2}
ight)^{(j-i)/2} \mathrm{e}^{-(\lambda_1+\lambda_2)t} I_{|j-i|}(2t\sqrt{\lambda_1\lambda_2}), \end{aligned}$$

 $i, j \in \mathbb{Z}, t \in \mathbb{R}_+.$

When $\lambda_1 = \lambda_2 = \lambda$ we find

$$\mathbb{P}(M_t=n)=\mathrm{e}^{-2\lambda t}I_{|n|}(2\lambda t).$$

d) From the bound^{*}

$$I_{|n|}(y) < C_n y^{|n|} \mathrm{e}^y, \qquad y>1,$$

we get

$$egin{aligned} \mathbb{P}(M_t = n) &\leqslant \ C_n \mathrm{e}^{-(\lambda_1 + \lambda_2)t} \left(rac{\lambda_1}{\lambda_2}
ight)^{n/2} (2t\sqrt{\lambda_1\lambda_2})^{|n|} \mathrm{e}^{2t\sqrt{\lambda_1\lambda_2}} \ &= \ C_n \mathrm{e}^{-t(\sqrt{\lambda_1} - \sqrt{\lambda_2})^2} \left(rac{\lambda_1}{\lambda_2}
ight)^{n/2} (2t\sqrt{\lambda_1\lambda_2})^{|n|}, \end{aligned}$$

which tends to 0 as t goes to infinity when $\lambda_1 \neq \lambda_2$.

Hence we have †

$$\lim_{t \to \infty} \mathbb{P}(|M_t| < c) = \sum_{-c < k < c} \lim_{t \to \infty} \mathbb{P}(|M_t| = k) = 0, \quad c > 0.$$
(B.56)

Remark. There exists a shorter proof by a probabilistic argument (not required here), which is displayed below, cf. [BN96].

^{*}See e.g. Theorem 2.1 of [Laf91] for a proof of this inequality.

[†]Treating the case $\lambda_1 = \lambda_2$ is more complicated and not required.

Noting that since $\mathbb{E}[M_t] = t(\lambda_1 - \lambda_2)$, for t large enough we have, depending on the sign of $\lambda_1 - \lambda_2$,

$$c-\mathbb{E}[M_t]\leqslant -rac{1}{2}\,\mathbb{E}[M_t] \quad ext{or} \quad -c-\mathbb{E}[M_t]\geqslant rac{1}{2}\,\mathbb{E}[M_t],$$

we get

$$-c \leqslant M_t \leqslant c \quad \Longrightarrow \quad M_t - \mathbb{E}[M_t] \leqslant -rac{1}{2} \mathbb{E}[M_t] \quad ext{or} \quad M_t - \mathbb{E}[M_t] \geqslant rac{1}{2} \mathbb{E}[M_t],$$

hence by Chebyshev's inequality we have

$$egin{aligned} \mathbb{P}(|M_t|\leqslant c) &\leqslant & \mathbb{P}\left(|M_t-\mathbb{E}[M_t]|\geqslant rac{1}{2}\,\mathbb{E}[M_t]
ight) \ &\leqslant & 4rac{ ext{Var}[M_t]}{(\mathbb{E}[M_t])^2} = rac{4}{t}rac{\lambda_1+\lambda_2}{(\lambda_1-\lambda_2)^2}, \end{aligned}$$

which tends to 0 as t tends to infinity, provided that $\lambda_1 \neq \lambda_2$. A probabilistic proof is also available in case $\lambda_1 = \lambda_2$, using the central limit theorem.

e) When $M_t \ge 0$, M_t represents the number of waiting customers. When $M_t \le 0$, $-M_t$ represents the number of waiting drivers.

Relation (B.56) 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.

♂722

$$Q = egin{bmatrix} -N\lambda & N\lambda & 0 & \cdots & \cdots & 0 & 0 & 0 \ \mu & -\mu - (N-1)\lambda & (N-1)\lambda & \cdots & \cdots & 0 & 0 & 0 \ dots & dots &$$

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

$$rac{\partial G_k}{\partial t}(s,t) = \sum\limits_{n=0}^N s^n P_{k,n}'(t),$$

and use the expression

$$rac{\partial G_k}{\partial s}(s,t) = \sum\limits_{n=1}^N n s^{n-1} P_{k,n}'(t).$$

d) We have

$$\begin{split} \lambda N(s-1)G_k(s,t) &+ (\mu + (\lambda - \mu)s - \lambda s^2) \frac{\partial G_k}{\partial s}(s,t) - \frac{\partial G_k}{\partial t}(s,t) \\ &= -(s-1)(\lambda + \mu)(N-k)\left((s-1)\mu \mathrm{e}^{-(\lambda+\mu)t} + \lambda s + \mu\right)^k \lambda \left(-(s-1)\lambda \mathrm{e}^{-(\lambda+\mu)t} + \lambda s + \mu\right)^{N-k-1} \mathrm{e}^{-(\lambda+\mu)t} \\ &+ (s-1)(\lambda + \mu)k\mu \left((s-1)\mu \mathrm{e}^{-(\lambda+\mu)t} + \lambda s + \mu\right)^{k-1} \left(-(s-1)\lambda \mathrm{e}^{-(\lambda+\mu)t} + \lambda s + \mu\right)^{N-k} \mathrm{e}^{-(\lambda+\mu)t} \\ &+ (s-1)\left((s-1)\mu \mathrm{e}^{-(\lambda+\mu)t} + \lambda s + \mu\right)^k N\lambda \left(-(s-1)\lambda \mathrm{e}^{-(\lambda+\mu)t} + \lambda s + \mu\right)^{N-k} \\ &+ \left(\lambda s^2 - (\lambda - \mu)s - \mu\right) (N-k) \left(\lambda \mathrm{e}^{-(\lambda+\mu)t} - \lambda\right) \left((s-1)\mu \mathrm{e}^{-(\lambda+\mu)t} + \lambda s + \mu\right)^k \left(-(s-1)\lambda \mathrm{e}^{-(\lambda+\mu)t} + \lambda s + \mu\right)^{N-k-1} \\ &- \left(\mu \mathrm{e}^{-(\lambda+\mu)t} + \lambda\right) k \left((s-1)\mu \mathrm{e}^{-(\lambda+\mu)t} + \lambda s + \mu\right)^{k-1} \left(-(s-1)\lambda \mathrm{e}^{-(\lambda+\mu)t} + \lambda s + \mu\right)^{N-k} \left(\lambda s^2 - (\lambda - \mu)s - \mu\right) \\ &= 0. \end{split}$$

e) This expression follows from the relation

$$\mathbb{E}[X_t \mid X_0 = k] = rac{\partial G_k}{\partial s} (s,t)_{|s=1}$$

ॳ723

and the result of Question (d).

♂724

f) We have

$$egin{aligned} &\lim_{t o\infty} \mathbb{E}[X_t\mid X_0=k] \;=\; k rac{\lambda(\lambda+\mu)^{k-1}}{(\lambda+\mu)^N} (\mu+\lambda)^{N-k} \ &+ (N-k) rac{(\mu+\lambda)\lambda^{k-1}}{(\lambda+\mu)^N} \lambda^{N-k} = rac{N\lambda}{\lambda+\mu}. \end{aligned}$$

Problem 9.20 First, we note that in case

$$f(x)=\mathbb{1}_{[a,b]}(x), \qquad 0\leqslant a\leqslant b\leqslant t,$$

we have, by direct counting,

$$\sum\limits_{k=1}^{N_t} f(T_k) = \sum\limits_{k=1}^{N_t} \mathbb{1}_{[a,b]}(T_k) = N_b - N_a.$$

Hence

$$\mathbb{E}\left[\sum\limits_{k=1}^{N_t}f(T_k)
ight] = \mathbb{E}[N_b - N_a] = \lambda(b-a) = \lambda\int_0^t\mathbb{1}_{[a,b]}(s)ds = \lambda\int_0^tf(s)ds,$$
for $f(x) = \mathbb{1}_{[a,b]}(x).$

hence (9.8.6) is proved for $f(x) = \mathbb{1}_{[a,b]}(x)$.

Next, we check by linearity that the (9.8.6) extends to a linear combination of indicator functions of the form

$$f(x) = \sum\limits_{k=1}^n lpha_k \mathbb{1}_{[a_{k-1},a_k]}(x), \qquad 0 \leqslant a_0 < a_1 < \cdots < a_n < t.$$

The difficult part is to do the extension from the linear combinations of indicator functions to "any" integrable function $f:[0,t] \longrightarrow \mathbb{R}$. This requires the knowledge of measure theory. For a different proof using the exponentially distributed jump times of the Poisson process, see Proposition 2.3.4 in [Pri09].

Let $g_n, n \ge 1$, be defined as

$$g_n(t_1, t_2, \dots, t_n) = \sum_{k=1}^n \mathbb{1}_{\{t_{k-1} < t < t_k\}}(f(t_1) + \dots + f(t_{k-1})) + \mathbb{1}_{\{t_n < t\}}(f(t_1) + \dots + f(t_n)),$$
 MH3512 AY19-20

Then

$$\begin{split} \mathbb{E}\left[\sum_{k=1}^{\min(n,N_{t})} f(T_{k})\right] &= \mathbb{E}[g(T_{1},T_{2},\ldots,T_{n})] \\ &= \lambda^{n} \sum_{k=1}^{n} \int_{0}^{\infty} e^{-\lambda t_{n}} \int_{0}^{t_{n}} \cdots \int_{0}^{t_{2}} \mathbb{1}_{\{t_{k-1} < t < t_{k}\}} (f(t_{1}) + \cdots + f(t_{k-1})) dt_{1} \cdots dt_{n} \\ &+ \lambda^{n} \int_{0}^{t} e^{-\lambda t_{n}} \int_{0}^{t_{n}} \cdots \int_{0}^{t_{2}} (f(t_{1}) + \cdots + f(t_{n})) dt_{1} \cdots dt_{n} \\ &= \lambda^{n} \sum_{k=1}^{n} \int_{t}^{\infty} e^{-\lambda t_{n}} \frac{(t_{n} - t)^{n-k}}{(n - k - 1)!} dt_{n} \\ &\times \int_{0}^{t} \int_{0}^{t_{k-1}} \cdots \int_{0}^{t_{2}} (f(t_{1}) + \cdots + f(t_{k-1})) dt_{1} \cdots dt_{k-1} \\ &+ \lambda^{n} \int_{0}^{t} e^{-\lambda t_{n}} \int_{0}^{t_{n}} \cdots \int_{0}^{t_{2}} (f(t_{1}) + \cdots + f(t_{n})) dt_{1} \cdots dt_{n} \\ &= e^{-\lambda t} \sum_{k=1}^{n} \lambda^{k} \int_{0}^{\infty} e^{-\lambda t_{n}} \frac{(\lambda t)^{n-k}}{(n - k - 1)!} dt \\ &\times \int_{0}^{t} \int_{0}^{t_{k-1}} \cdots \int_{0}^{t_{2}} (f(t_{1}) + \cdots + f(t_{k-1})) dt_{1} \cdots dt_{k-1} \\ &+ \lambda^{n} \int_{0}^{t} e^{-\lambda t_{n}} \int_{0}^{t_{n-1}} \cdots \int_{0}^{t_{2}} (f(t_{1}) + \cdots + f(t_{k-1})) dt_{1} \cdots dt_{k-1} \\ &+ \lambda^{n} \int_{0}^{t} e^{-\lambda t_{n}} \int_{0}^{t_{k-1}} \cdots \int_{0}^{t_{2}} (f(t_{1}) + \cdots + f(t_{k-1})) dt_{1} \cdots dt_{k-1} \\ &+ e^{-\lambda t} \sum_{k \ge n} \lambda^{k} \int_{0}^{t} \int_{0}^{t_{k-1}} \cdots \int_{0}^{t_{2}} (f(t_{1}) + \cdots + f(t_{k-1})) dt_{1} \cdots dt_{k-1} \\ &+ e^{-\lambda t} \sum_{k \ge n} \lambda^{k} \int_{0}^{t} \int_{0}^{t_{k-1}} \cdots \int_{0}^{t_{2}} (f(t_{1}) + \cdots + f(t_{k-1})) dt_{1} \cdots dt_{k-1} \\ &+ e^{-\lambda t} \sum_{k \ge n} \lambda^{k} \int_{0}^{t} \int_{0}^{t_{k-1}} \cdots \int_{0}^{t_{2}} (f(t_{1}) + \cdots + f(t_{k-1})) dt_{1} \cdots dt_{k-1} \\ &+ e^{-\lambda t} \sum_{k \ge n} \lambda^{k} \int_{0}^{t} \int_{0}^{t_{k-1}} \cdots \int_{0}^{t_{2}} (f(t_{1}) + \cdots + f(t_{k-1})) dt_{1} \cdots dt_{k-1} \\ &+ e^{-\lambda t} \sum_{k \ge n} \lambda^{k} \int_{0}^{t} \int_{0}^{t_{k-1}} \cdots \int_{0}^{t_{2}} (f(t_{1}) + \cdots + f(t_{k-1})) dt_{1} \cdots dt_{k-1} \\ &+ e^{-\lambda t} \sum_{k \ge n} \lambda^{k} \int_{0}^{t} \int_{0}^{t} \int_{t_{n}}^{t_{k}} \cdots \int_{t_{n}}^{t_{n+2}} dt_{n+1} \cdots dt_{k} \end{aligned}$$

() 725

$$\begin{split} & \times \int_0^{t_{n+1}} \cdots \int_0^{t_2} (f(t_1) + \cdots + f(t_n)) dt_1 \cdots dt_n \\ &= \mathrm{e}^{-\lambda t} \sum_{k \geqslant 0} \lambda^k \int_0^t \int_0^{t_k} \cdots \int_0^{t_2} (f(t_1) + \cdots + f(t_{\min(k,n)})) dt_1 \cdots dt_k \\ &= \mathrm{e}^{-\lambda t} \sum_{k \geqslant 0} \frac{\lambda^k}{k!} \int_0^t \cdots \int_0^t (f(t_1) + \cdots + f(t_{\min(k,n)})) dt_1 \cdots dt_k. \end{split}$$

Hence as \boldsymbol{n} goes to infinity,

$$egin{aligned} &\mathbb{E}\left[\sum\limits_{k=1}^{N_t}f(T_k)
ight] = \lim_{n o\infty}\mathbb{E}\left[\sum\limits_{k=1}^{\min(n,N_t)}f(T_k)
ight] \ &= \lim_{n o\infty}\mathrm{e}^{-\lambda t}\sum\limits_{k\geqslant 0}rac{\lambda^k}{k!}\int_0^t\cdots\int_0^t(f(t_1)+\cdots+f(t_{\min(k,n)}))dt_1\cdots dt_k \ &= \mathrm{e}^{-\lambda t}\sum\limits_{k\geqslant 0}rac{\lambda^k}{k!}\int_0^t\cdots\int_0^t(f(t_1)+\cdots+f(t_k))dt_1\cdots dt_k \ &= \lambda\mathrm{e}^{-\lambda t}\int_0^tf(s)ds\sum\limits_{k\geqslant 1}rac{(\lambda t)^{k-1}}{(k-1)!} = \lambda\int_0^tf(s)ds. \end{aligned}$$

Chapter 10 - Discrete-Time Martingales

Exercise 10.1

a) We have

$$egin{array}{rll} \mathbb{E}\left[M_{n+1} \mid \mathcal{F}_n
ight] &=& \mathbb{E}\left[\sum\limits_{k=1}^{n+1} 2^{k-1} X_k \mid \mathcal{F}_n
ight] \ &=& \mathbb{E}\left[\sum\limits_{k=1}^n 2^{k-1} X_k \mid \mathcal{F}_n
ight] + \mathbb{E}\left[2^n X_{n+1} \mid \mathcal{F}_n
ight] \ &=& \sum\limits_{k=1}^n 2^{k-1} X_k + 2^n \, \mathbb{E}[X_{n+1}] \ &=& M_n, \qquad n \geqslant 0. \end{array}$$

MH3512 AY19-20

ॳ726

- b) This random time is a hitting time, so it is a stopping time.
- c) The strategy of the gambler is to double the stakes each time he loses (double down strategy), and to quit the game as soon as his gains reach **\$1**.
- d) The two possible values of $M_{\tau \wedge n}$ are 1 and

$$-\sum\limits_{k=1}^n 2^{k-1} = -rac{1-2^n}{1-2} = 1-2^n, \qquad n \geqslant 1.$$

We have

$$\mathbb{P}(M_{n \wedge \tau} = 1 - 2^n) = 2^{-n}$$
 and $\mathbb{P}(M_{n \wedge \tau} = 1) = 1 - 2^{-n}, \quad n \ge 1.$

e) We have

$$egin{array}{rll} \mathbb{E}[M_{n\wedge au}] &=& (1-2^n)\mathbb{P}(M_{n\wedge au}=1-2^n)+\mathbb{P}(M_{n\wedge au}=1) \ &=& (1-2^n)2^{-n}+(1-2^{-n}) \ &=& 0, \qquad n\geqslant 1. \end{array}$$

f) The stopping time theorem directly states that

$$\mathbb{E}[M_{n\wedge au}] = \mathbb{E}[M_0] = 0.$$

Exercise 10.2

a) From the tower property of conditional expectations we have:

$$\mathbb{E}[M_{n+1}] = \mathbb{E}[\mathbb{E}[M_{n+1} \mid \mathcal{F}_n]] \geqslant \mathbb{E}[M_n], \qquad n \geqslant 0.$$

b) If $(Z_n)_{n \in \mathbb{N}}$ is a stochastic process with independent increments have negative expectation, we have

$$egin{array}{rll} \mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] &=& \mathbb{E}[Z_n \mid \mathcal{F}_n] + \mathbb{E}\left[Z_{n+1} - Z_n \mid \mathcal{F}_n
ight] \ &=& \mathbb{E}[Z_n \mid \mathcal{F}_n] + \mathbb{E}[Z_{n+1} - Z_n] \ &\leqslant& \mathbb{E}[Z_n \mid \mathcal{F}_n] = Z_n, \qquad n \geqslant 0. \end{array}$$

c) We let $A_0 := 0, A_{n+1} := A_n + \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n], n \ge 0$, and

$$N_n := M_n - A_n, \qquad n \in \mathbb{N}. \tag{B.57}$$

(i) For all $n \in \mathbb{N}$, we have

$$\begin{split} \mathbb{E}[N_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}[M_{n+1} - A_{n+1} \mid \mathcal{F}_n] \\ &= \mathbb{E}[M_{n+1} - A_n - \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] \mid \mathcal{F}_n] \\ &= \mathbb{E}[M_{n+1} - A_n \mid \mathcal{F}_n] - \mathbb{E}[\mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] \mid \mathcal{F}_n] \\ &= \mathbb{E}[M_{n+1} - A_n \mid \mathcal{F}_n] - \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] \\ &= -\mathbb{E}[A_n \mid \mathcal{F}_n] + \mathbb{E}[M_n \mid \mathcal{F}_n] = M_n - A_n = N_n, \end{split}$$

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

$$egin{array}{rcl} A_{n+1}-A_n&=&\mathbb{E}[M_{n+1}-M_n\mid\mathcal{F}_n]\ &=&\mathbb{E}[M_{n+1}\mid\mathcal{F}_n]-\mathbb{E}[M_n\mid\mathcal{F}_n]\ &=&\mathbb{E}[M_{n+1}\mid\mathcal{F}_n]-M_n\geqslant 0, \end{array}$$

since $(M_n)_{n \in \mathbb{N}}$ is a submartingale.

- (iii) By induction we have $A_{n+1} = A_n + \mathbb{E}[M_{n+1} M_n | \mathcal{F}_n], n \in \mathbb{N}$, which is \mathcal{F}_n -measurable if A_n is \mathcal{F}_{n-1} -measurable, $n \ge 1$.
- (iv) This property is obtained by construction in (B.57).
- 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

$$egin{array}{rcl} \mathbb{E}[M_\sigma] &=& \mathbb{E}[N_\sigma] + \mathbb{E}[A_\sigma] \ &=& \mathbb{E}[N_ au] + \mathbb{E}[A_\sigma] \ &\leqslant& \mathbb{E}[N_ au] + \mathbb{E}[A_ au] \ &\leqslant& \mathbb{E}[N_ au] + \mathbb{E}[A_ au] \ &=& \mathbb{E}[M_ au], \end{array}$$

by (10.3.3), since $(M_n)_{n \in \mathbb{N}}$ is a martingale and $(A_n)_{n \in \mathbb{N}}$ is non-decreasing.

 $\bigcirc 728$

Exercise 10.3 We have

$$\mathbb{E}^{*}\left[\phi\left(\frac{S_{1}+\dots+S_{N}}{N}\right)\right] \leqslant \mathbb{E}^{*}\left[\frac{\phi(S_{1})+\dots+\phi(S_{N})}{N}\right] \qquad \text{since } \phi \text{ is convex},$$

$$= \frac{\mathbb{E}^{*}[\phi(S_{1})]+\dots+\mathbb{E}^{*}[\phi(S_{N})]}{N}$$

$$= \frac{\mathbb{E}^{*}[\phi(\mathbb{E}^{*}[S_{N} \mid \mathcal{F}_{1}])]+\dots+\mathbb{E}^{*}[\phi(\mathbb{E}^{*}[S_{N} \mid \mathcal{F}_{N}])]}{N} \qquad \text{because } (S_{n})_{n \in \mathbb{N}} \text{ is a martingale},$$

$$\leqslant \frac{\mathbb{E}^{*}[\mathbb{E}^{*}[\phi(S_{N}) \mid \mathcal{F}_{1}]]+\dots+\mathbb{E}^{*}[\mathbb{E}^{*}[\phi(S_{N}) \mid \mathcal{F}_{N}]]}{N} \qquad \text{by Jensen's inequality},$$

$$= \frac{\mathbb{E}^{*}[\phi(S_{N})]+\dots+\mathbb{E}^{*}[\phi(S_{N})]}{N} \qquad \text{by the tower property},$$

$$= \mathbb{E}^{*}[\phi(S_{N})].$$

Exercise 10.4

a) We have

$$\mathbb{E}[M_n] \leqslant \mathbb{E}[\mathbb{E}[M_{n+1} \mid \mathcal{F}_n]] = \mathbb{E}[M_{n+1}], \qquad n \in \mathbb{N}.$$

b) Writing

$$S_n - lpha n = \underbrace{(X_1 + X_2 + \dots + X_n - np)}_{ ext{martingale}} + n(p - lpha)$$

as the sum of a martingale (a stochastic process with centered independent increments) and $n(p-\alpha)$, we conclude that $(S_n)_{n\in\mathbb{N}}$ is a *sub*martingale if and only if $p \ge \alpha$. Indeed, we have

$$egin{array}{rll} \mathbb{E}[S_n-lpha n\mid \mathcal{F}_k] &=& \mathbb{E}[X_1+X_2+\dots+X_n-np\mid \mathcal{F}_k]+n(p-lpha) \ &=& X_1+X_2+\dots+X_k-kp+n(p-lpha) \ &=& X_1+X_2+\dots+X_k-klpha+(n-k)(p-lpha) \ &\geqslant& S_k-klpha, \qquad k=0,1,\dots,n, \end{array}$$

 $\text{ if and only if } p \geqslant \alpha. \\$

Exercise 10.5

a) We have

$$\phi(M_k) = \phi\left(\mathbb{E}[M_n \mid \mathcal{F}_k]
ight) \leqslant \mathbb{E}[\phi(M_n) \mid \mathcal{F}_k], \qquad k = 0, 1, \dots, n$$

() 729

b) We have

$\phi(M_k)\leqslant \phi\left(\mathbb{E}[M_n\mid \mathcal{F}_k] ight)\leqslant \mathbb{E}[\phi(M_n)\mid \mathcal{F}_k], \qquad k=0,1,\dots,n.$

Problem 10.6

a) We have

$$\{ au_x > n\} = igcap_{k=0}^n \{M_k < x\}.$$

On the other hand, for all k = 0, 1, ..., n we have $\{M_k < x\} \in \mathcal{F}_k \subset \mathcal{F}_n$, hence $\{\tau_x > n\} \in \mathcal{F}_n$ by stability of σ -algebras by intersection, cf. (1.1.1).

b) We have

$$\begin{split} x \mathbb{P}\Big(\max_{k=0,1,\dots,n} M_k \geqslant x\Big) &= x \mathbb{P}(\tau_x \leqslant n) \\ &= x \mathbb{E}[\mathbb{1}_{\{\tau_x \leqslant n\}}] \\ &\leqslant \mathbb{E}\left[M_{\tau_x \wedge n} \mathbb{1}_{\{\tau_x \leqslant n\}}\right] \\ &\leqslant \mathbb{E}[M_{\tau_x \wedge n}] \\ &= \mathbb{E}[M_n], \end{split} \tag{B.59}$$

where we used the condition $M_{\tau_x \wedge n} \ge 0$ from (B.58) to (B.59), hence

$$\mathbb{P}\left(\max_{k=0,1,\dots,n} M_k \geqslant x\right) \leqslant \frac{\mathbb{E}[M_n]}{x}, \qquad x > 0.$$
(B.61)

Remark. The nonnegativity of $(M_n)_{n \in \mathbb{N}}$ is used to reach (B.59), and the Doob stopping time theorem is used to conclude to (B.60).

c) When $(M_n)_{n \in \mathbb{N}}$ is a *sub*martingale we have

 $\mathbb{E}[M_{ au_x\wedge n}]\leqslant \mathbb{E}[M_n],$

by the Doob stopping time theorem for *sub*martingales, *cf.* Exercise 10.2-(d), hence we can still conclude to (B.60) and (B.61) above.

d) Since $x \mapsto x^2$ is a convex function, $((M_n)^2)_{n \in \mathbb{N}}$ is a *sub*martingale by Question (a), hence by Question (c) we have

$$\mathbb{P}\Big(\max_{k=0,1,...,n}M_k\geqslant x\Big)=\mathbb{P}\Big(\max_{k=0,1,...,n}(M_k)^2\geqslant x^2\Big)\leqslant rac{\mathbb{E}[(M_n)^2]}{x^2},\qquad x>0.$$

e) Similarly to Question (d), $x \mapsto x^p$ is a convex function for all $p \ge 1$ hence $((M_n)^2)_{n \in \mathbb{N}}$ is a submartingale by Question (a), and by Question (c) we find

$$\mathbb{P}\Big(\max_{k=0,1,...,n}M_k\geqslant x\Big)=\mathbb{P}\Big(\max_{k=0,1,...,n}(M_k)^p\geqslant x^p\Big)\leqslant rac{\mathbb{E}[(M_n)^p]}{x^p},\qquad x>0$$

f) We note that $(S_n)_{n \in \mathbb{N}}$ is a martingale because it has centered and independent increments, with

$$\mathbb{E}[(S_n)^2] = \operatorname{Var}[S_n] = n \operatorname{Var}[Y_1] = n\sigma^2,$$

hence by Question (d) we have

$$\mathbb{P}\Big(\max_{k=0,1,...,n}S_k\geqslant x\Big)\leqslant rac{\mathbb{E}[(S_n)^2]}{x^2}=rac{n\sigma^2}{x^2},\qquad x>0.$$

g) When $(M_n)_{n \in \mathbb{N}}$ is a (not necessarily nonnegative) *sub*martingale we can modify the answer to Question (b) using the Doob stopping time theorem for *super*martingales, as follows:

$$egin{aligned} x \mathbb{P}\Big(\max_{k=0,1,...,n} M_k \geqslant x\Big) &= x \, \mathbb{E}[\mathbbm{1}_{\{ au_x \leqslant n\}}] \ &\leqslant & \mathbb{E}\left[M_{ au_x \wedge n} \mathbbm{1}_{\{ au_x \leqslant n\}}
ight] \ &\leqslant & \mathbb{E}[(M_{ au_x \wedge n})^+] \ &\leqslant & \mathbb{E}[(M_n)^+], \end{aligned}$$

since $((M_k)^+)_{k\in\mathbb{N}}$ is a submartingale because $x\mapsto x^+$ is a non-decreasing convex function, cf. Question (b), hence

$$\mathbb{P}\Big(\max_{k=0,1,...,n}M_k\geqslant x\Big)\leqslant rac{\mathbb{E}[(M_n)^+]}{x},\qquad x>0.$$

h) We have

$$egin{aligned} x \mathbb{P}\Big(\max_{k=0,1,...,n} M_k \geqslant x\Big) &= x \mathbb{P}(au_x \leqslant n) \ &= x \, \mathbb{E}[\mathbbm{1}_{\{ au_x \leqslant n\}}] \ &\leqslant & \mathbb{E}\left[M_{ au_x \wedge n} \mathbbm{1}_{\{ au_x \leqslant n\}}
ight] \ &\leqslant & \mathbb{E}[M_{ au_x \wedge n}] \end{aligned}$$

 $\leqslant \mathbb{E}[M_0].$

i) We have

$$egin{aligned} x \mathbb{P}\Big(\max_{k=0,1,...,n} \phi(M_k) \geqslant x\Big) &= x \, \mathbb{E}\left[\mathbbm{1}_{\left\{\max_{k=0,1,...,n} \phi(M_k) \geqslant x
ight\}}
ight] \ &= x \, \mathbb{E}\left[\mathbbm{1}_{\left\{ au_x^\phi \leqslant n
ight\}}
ight] \ &= \mathbb{E}\left[\phi\left(M_{ au_x^\phi \wedge n}
ight)\mathbbm{1}_{\left\{ au_x^\phi \leqslant n
ight\}}
ight] \ &= \mathbb{E}\left[\phi(M_{ au_x^\phi \wedge n})
ight] \ &= \mathbb{E}\left[\phi\left(M_{ au_x^\phi \wedge n}
ight)
ight] \ &= \mathbb{E}[\phi\left(M_{ au_x^\phi \wedge n}
ight)
ight], \end{aligned}$$

where the last inequality follows from Exercise 10.2-(d), since both $\tau_x^{\phi} \wedge n$ and n are stopping times.

j) Consider for example any nonnegative martingale such as $M_n = (p/q)^{S_n}$ where $S_n = X_1 + \cdots + X_n$ and $(X_k)_{k \ge 1}$ is a sequence of independent identically distributed Bernoulli random variables with $p = \mathbb{P}(X_k = 1)$ and $q = 1 - p = \mathbb{P}(X_k = -1)$, $k \ge 1$. Then $Z_n := e^{-n}M_n$ will be a *super*martingale since

$$\mathbb{E}[Z_n \mid \mathcal{F}_k] = \mathrm{e}^{-n} \, \mathbb{E}[M_n \mid \mathcal{F}_k] = \mathrm{e}^{-n} M_k \leqslant \mathrm{e}^{-k} M_k = Z_k, \hspace{1em} k = 0, 1, \dots, n.$$

Exercise 10.7 By (3.3.3) we have

$$\begin{split} \mathbb{E}[(M_{2n})^r] &= \mathbb{E}\left[(q/p)^{rS_{2n}}\right] \\ &= \sum_{k=-n}^n (q/p)^{2kr} \mathbb{P}(S_{2n} = 2k) \\ &= \sum_{k=-n}^n \binom{2n}{n+k} (q/p)^{2kr} p^{n+k} q^{n-k} \\ &= \sum_{k=0}^{2n} \binom{2n}{k} (q/p)^{2(k-n)r} p^k q^{2n-k} \\ &= (q/p)^{-2nr} q^{2n} \sum_{k=0}^{2n} \binom{2n}{k} ((q/p)^{2r-1})^k \\ &= (p(q/p)^r + q(p/q)^r)^{2n} \end{split}$$

$$= \prod_{k=1}^{2n} \mathbb{E}\left[(q/p)^{(S_{k+1}-S_k)r} \right], \qquad \diamondsuit 733$$

hence by Problem 10.6-(e), for all $n \ge 0$ and $r \ge 1$ we have

$$egin{aligned} \mathbb{P}\Big(\max_{k=0,1,...,n}M_k\geqslant x\Big)&\leqslant&rac{\mathbb{E}[(M_n)^r]}{x^r}\ &=&rac{(p(q/p)^r+q(p/q)^r)^{2n}}{x^r},\qquad x>0. \end{aligned}$$



Figure S.23: Supremum deviation probability with n = 7 and p = 0.4.

Chapter 11 - Spatial Poisson Processes

Exercise 11.1 We note that N_t can be constructed as $N_t(\omega) = \omega([0,t])$, and by (11.1.2) we find that given $\{N_T = k\}$, the distribution of (T_1, T_2, \ldots, T_k) is uniform on

$$\{(t_1,t_2,\ldots,t_k) \; : \; 0\leqslant t_1\leqslant \cdots \leqslant t_k\leqslant T\}.$$

This yields

$$\mathbb{E}[T_1 \mid N_T = 2] = rac{2}{T^2} \int_0^T \int_0^y x dx dy = rac{1}{T^2} \int_0^T y^2 dy = rac{T}{3},$$

and

$$\mathbb{E}[T_2 \mid N_T = 2] = rac{2}{T^2} \int_0^T y \int_0^y dx dy = rac{2}{T^2} \int_0^T y^2 dy = rac{2T}{3},$$

and we check that

$$\mathbb{E}[T_1+T_2\mid N_T=2]=T=rac{2}{T}\int_0^Txdx,$$

^{*}The animation works in Acrobat Reader on the entire pdf file.

Figure S.24: Martingale supremum as a function of time.^{*}

which is consistent with the fact that the (unordered) locations of the two jump times are inform of [0, T] given that $N_T = 2$, cf. (11.1.2) page 433.

On the other hand, by the memoryless property of the standard Poisson process we have

$$\mathbb{E}[T_3 \mid N_T=2]=T+\mathbb{E}[T_1 \mid N_T=0]=T+rac{1}{\lambda}.$$

Exercise 11.2 The probability that there are 10 events within a circle of radius **3** meters is

$$e^{-9\pi\lambda} \frac{(9\pi\lambda)^{10}}{10!} = e^{-9\pi/2} \frac{(9\pi/2)^{10}}{10!} \simeq 0.0637.$$

Exercise 11.3 The probability that more than two living organisms are in this measured volume is

$$egin{aligned} \mathbb{P}(N \geqslant 3) &= 1 - \mathbb{P}(N \leqslant 2) = 1 - \mathrm{e}^{-10 heta} \left(1 + 10 heta + rac{(10 heta)^2}{2}
ight) \ &= 1 - \mathrm{e}^{-6} \left(1 + 6 + rac{6^2}{2}
ight) = 1 - 25 \, \mathrm{e}^{-6} \simeq 0.938. \end{aligned}$$

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

$$\mathbb{P}(X_A \ge 1 \text{ and } X_B \ge 1) = \mathbb{P}(X_A \ge 1) \mathbb{P}(X_B \ge 1) \\ = (1 - \mathbb{P}(X_A = 0))(1 - \mathbb{P}(X_B = 0)) \\ = (1 - e^{-0.5})^2 \simeq 0.2212.$$

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.*

$$\mathbb{P}(X_N=k)=inom{N}{k}inom{\lambda}{N}^kinom{1-rac{\lambda}{N}}^{N-k},\qquad k=0,1,\ldots,N,$$

and we find

$$\lim_{N o\infty} \mathbb{P}(X_N=k) = rac{\lambda^k}{k!} \lim_{N o\infty} \left(\left(1-rac{\lambda}{N}
ight)^{N-k} \prod_{i=0}^{k-1} rac{N-i}{N}
ight) = \mathrm{e}^{-\lambda} rac{\lambda^k}{k!},$$

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

$$e^{-9\pi/2} \frac{(9\pi/2)^{10}}{10!}.$$

b) This probability is

$$e^{-9\pi/2} \frac{(9\pi/2)^5}{5!} \times e^{-9\pi/2} \frac{(9\pi/2)^3}{3!}$$

c) This probability is

₫ 737

$$e^{-9\pi} \frac{(9\pi)^8}{8!}.$$

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.10) in the solution of Exercise 1.6 and Exercise 9.2-(d), and we find the probability

$$\binom{5}{3}(1/4)^3(3/4)^2 = rac{45}{512} \simeq 0.08789$$

Exercise 11.7 ([WSL12]) By the moment identity Relation (11.4.1) we have

$$egin{aligned} \mathbb{E}\left[\left|rac{S_n-\lambda n}{\sqrt{n}}
ight|^p
ight] &= \left.n^{-p/2}\sum\limits_{k=0}^p(n\lambda)^kS_2(p,k) = n^{-p/2}\sum\limits_{k=0}^{p/2}(n\lambda)^kS_2(p,k)
ight. \ &= \left.n^{-p/2}\sum\limits_{k=0}^{p/2}(n\lambda)^kp^k = n^{-p/2}rac{(np\lambda)^{1+p/2}-1}{np\lambda-1}
ight. \ &\leqslant \left.rac{(p\lambda)^{1+p/2}}{p\lambda-1/n} < C_p, \end{aligned}$$

where $C_p > 0$ is a finite constant.

Exercise 11.8

a) We have

$$egin{aligned} L'(s) &= \int_X f(x)(\mathrm{e}^{sf(x)}-1)\sigma(dx) \, \mathbb{E}_{\mathbb{P}^X_\sigma}\left[\exp\left(s\int_0^\infty f(y)(dN_y-dy)
ight)
ight] \ &= s\int_X |f(x)|^2 rac{\mathrm{e}^{sf(x)}-1}{sf(x)}\sigma(dx) \, \mathbb{E}_{\mathbb{P}^X_\sigma}\left[\exp\left(s\int_0^\infty f(y)(dN_y-dy)
ight)
ight] \ &\leqslant \, rac{\mathrm{e}^{sK}-1}{K}\int_X |f(x)|^2\sigma(dx) \, \mathbb{E}_{\mathbb{P}^X_\sigma}\left[\exp\left(s\int_0^\infty f(y)(dN_y-dy)
ight)
ight] \end{aligned}$$

$$= lpha^2 rac{\mathrm{e}^{sK}-1}{K} \mathbb{E}_{\mathbb{P}^X_\sigma} \left[\exp \Big(s \int_0^\infty f(y) (dN_y - dy) \Big)
ight] = lpha^2 rac{\mathrm{e}^{sK}-1}{K} L(s),$$

which shows that

$$rac{L'(s)}{L(s)}\leqslant h(s):=lpha^2rac{{
m e}^{sK}-1}{K},\qquad s\in {\mathbb R}_+.$$

b) We have

$$egin{aligned} \log L(t) &= \ \log L(0) + \int_0^t d \log L(s) \ &\leqslant \ \int_0^t rac{L'(s)}{L(s)} ds \ &\leqslant \ \int_0^t h(s) ds, \end{aligned}$$

hence

$$L(t)\leqslant \exp\left(\int_0^t h(s)ds
ight)=\exp\left(lpha^2\int_0^t rac{\mathrm{e}^{sK}-1}{K}ds
ight),\qquad t\in\mathbb{R}_+.$$

c) By the Chebyshev inequality we have

$$egin{aligned} \mathbb{P}^X_\sigma\left(\int_0^\infty f(y)(dN_y-dy)\geqslant x
ight) &= \mathbb{E}_{\mathbb{P}^X_\sigma}\left[\mathbbm{1}_{\{\int_0^\infty f(y)(dN_y-dy)\geqslant x\}}
ight] \ &\leqslant \ \mathrm{e}^{-tx}\,\mathbb{E}_{\mathbb{P}^X_\sigma}\left[\mathbbm{1}_{\{\int_0^\infty f(y)(dN_y-dy)\geqslant x\}}\exp\left(t\int_0^\infty f(y)dN_y
ight)
ight] \ &\leqslant \ \mathrm{exp}\left(-tx+ \sum_{\sigma}^t\left[\exp\left(t\int_0^\infty f(y)dN_y
ight)
ight] \ &\leqslant \ \exp\left(-tx+ +\int_0^t h(s)ds
ight) \ &\leqslant \ \exp\left(-tx+ lpha^2\int_0^trac{\mathrm{e}^{sK}-1}{K}ds
ight) \ &= \ \exp\left(-tx+ rac{lpha^2}{K^2}(\mathrm{e}^{tK}-tK-1)
ight), \end{aligned}$$

MH3512 AY19-20

() 738

which also yields

$$\mathbb{P}\left(\int_{0}^{\infty}f(y)(dN_{y}-dy)\geqslant x
ight)\leqslant\exp\left(-tx+lpha^{2}\int_{0}^{t}rac{e^{sK}-1}{K}ds
ight).$$

d) Minimizing the above term in t with $t = K^{-1} \log(1 + Kx/\alpha^2)$ shows that

$$egin{aligned} \mathbb{P}^X_\sigma\left(\int_0^\infty f(y)(dN_y-dy)\geqslant x
ight) &\leqslant & \exp\left(rac{x}{K}-\left(rac{x}{K}+rac{lpha^2}{K^2}
ight)\log\left(1+rac{xK}{lpha^2}
ight)
ight) \ &\leqslant & \exp\left(-rac{x}{2K}\log\left(1+rac{xK}{lpha^2}
ight)
ight) \ &= & \left(1+rac{xK}{lpha^2}
ight)^{-x/2K}. \end{aligned}$$

Chapter 12 - Reliability and Renewal Processes

Exercise 12.1

a) We have

$$F_eta(t) = \mathbb{P}(au < t) = \int_0^t f_eta(x) dx = eta \int_0^t x^{eta - 1} \mathrm{e}^{-x^eta} dx = -\left[\mathrm{e}^{-x^eta}
ight]_0^t = 1 - \mathrm{e}^{-t^eta},$$

 $t \in \mathbb{R}_+$.

b) We have

$$R(t)=\mathbb{P}(au>t)=1-F_eta(t)=\mathrm{e}^{-t^eta},\qquad t\in\mathbb{R}_+.$$

c) We have

$$\lambda(t) = -rac{d}{dt}\log R(t) = eta t^{eta - 1} \qquad t \in \mathbb{R}_+.$$

d) By (12.3.1) we have

$$\mathbb{E}[au] = \int_0^\infty R(t) dt = \int_0^\infty \mathrm{e}^{-t^eta} dt$$

In particular this yields $\mathbb{E}[\tau] = \sqrt{\pi/2}$ when $\beta = 2$.

5739

Index

Q-learning, 234 σ -algebra, 23 complement rule, 26absorbing set, 188 state, 149, 188, 256, 287 absorptionprobability, 391 time, 195, 392 accepting state, 230 accessible state, 243 action-value function, 237 adjacency matrix, 145 algorithm Baum-Welch, 185, 551 *EM*, **185** aperiodic chain, 257 state, 256, 266, 272, 274, 287 aperiodicity, 256, 274 backward Kolmogorov equation, 361 backward optimization, 598 balance condition, 269, 377, 639 Baum-Welch algorithm, 185, 551

Bellman equation, 226, 238

Bernoulli distribution, 62 random walk, 106 Bernoulli distribution, 40 Bernoulli-Laplace chain, 175, 289 Bessel function, 720 binary classification, 185 binomial coefficient, 455distribution, 40, 110 identity, 456birth and death process, 352, 383 birth process, 346 Boltzmann distribution, 308, 671 boundary condition, 71, 86 branching process, 310 subcritical, 318 supercritical, 317 Brownian motion, 15

Cauchy distribution, 36 process, 18 cellular automata, 230 centered random variable, 409 central moments, 439

Cesàro mean, 132 characteristic equation, 75, 103 function, 433 class communicating, 244 recurrent, 246, 248 class property, 247, 250, 256 classification of states, 242code Matlab, 283, 634 communicating class, 244state, 242, 244 compound Poisson process, 17 conditional expectation, 46probability, 28 conditional expectation, 407conditioning, 28continuous-time Markov chain, 338, 345 convolution equation, 112, 134, 212 cookie, 136, 139, 520, 524 counting process, 338 Cox process, 451critical (branching process), 317 cumulants, 60, 435cumulative distribution function, 34 joint, 37death process, 350 density function, 33 marginal, 38detailed balance, 285, 298, 308

deviation inequalities, 442diagonalization, 161, 371 Dirichlet problem, 190 discrete distribution, 40discrete-time Markov chain, 141 martingale, 408distance from stationarity, 303 distribution Bernoulli, 40 binomial, 40, 110 Boltzmann, 308, 671 Cauchy, 36discrete, 40Erlang, 343exponential, 35gamma, <u>36</u> Gaussian, 35 geometric, 41 zero-modified, 213 invariant, 148, 270, 272, 275 limiting, 163, 164, 265, 266, 374, 376 lognormal, <u>36</u> marginal, 50negative binomial, 41, 179, 328, 335, 470 Pascal, 41phase-type, 178 Poisson, 41 power tail, 313Skellam, 720 stationary, 268, 284, 299, 376, 398 uniform, 35Weibull, 453dominated convergence, 420

Doob-Meyer decomposition, 426 double down, 727 dynamic programming, 225, 234, 597, 598 Ehrenfest chain, 151, 288, 299, 620 elephant random walk, 174 EM algorithm, 185 embedded chain, 385 emission probabilities, 172, 182 equation characteristic, 75, 103 convolution, 112, 134, 212 homogeneous, 90, 96 Kolmogorov (backward), 361 Kolmogorov (forward), 361, 369, 378, 404 partial differential, 404 equivalence relation, 243 ergodic state, 256, 287 theorem continuous time, 374, 381 discrete time, 167, 279 Erlang distribution, 343 ERW, 174event, 23excited random walk, 136, 139, 520, 524 expectation, 42, 331conditional, 46exponential distribution, 35, 63, 337, 344, 363, 364, 397, 450 exponential series, 455 extinction probability, 318, 331 factorial moment, 59, 544 failure rate function, 449

failure time, 452 filtration, 406 first step analysis, 70, 88, 96, 155, 187, 203, 226, 291, 320, 336, 392, 473, 495, 567, 574, 585, 635, 696 forward Kolmogorov equation, 361, 369, 378, 404 front running, 399

gambling process, 65, 158 gamma distribution, 36 gamma function, 36 gamma process, 17 Gaussian distribution, 35 generating function, 56, 114, 323, 404 geometric distribution, 41 zero-modified, 213 series, 69, 206, 456 sum, 455 GridWorld, 234

hidden Markov, 172, 181, 544 high-frequency trading, 399 hitting probability, 187, 331 time, 195 HMM, 185 homogeneous equation, 90, 96

increment independent, 144, 340, 346, 409, 426 stationary, 340 independence, 28, 30, 33, 39, 41, 48, 59, 466 independent increments, 144, 340, 346, 409, 426 indicator function, 31, 455

infimum, 41, 85 infinitesimal generator, 360, 398 information flow, 407intensity parameter, 341 invariant distribution, 148, 270, 272, 275 inverse Gaussian process, 19 temperature, 308 IPython notebook, 167, 170, 330, 374, 385 irreducible, 244, 266, 272, 277, 287, 354 Ising model, 305 joint cumulative distribution function, 37 probability density function, 37 kernel potential, 216 Kolmogorov equation backward, 361 forward, 361, 369, 378, 404 lawof total expectation, 50, 196 of total probability, 26, 30, 50, 143, 145, 155, 188, 252, 357 layering, 399 lazy random walk, 96 limiting distribution, 163, 164, 265, 266, 374, 376 Bernoulli-Laplace chain, 289 Ehrenfest chain, 288 success runs chain, 288 loqnormal distribution, 36 marginal density, 38

distribution, 50Markov decision process, 225, 234, 597 hidden, 172, 181, 544 property, 141, 196, 345 strong, 220, 338, 344, 350, 636 Markov chain, 141 Bernoulli-Laplace, 175, 289 continuous time, 338, 345 discrete time, 141 embedded, 385irreducible, 244, 266, 272, 277, 287 Monte Carlo, 285 recurrent, 246, 266 reducible, 244 regular, 266, 287 reversible, 285, 298 success runs, 176, 264, 288 two-state, 160, 192, 198, 205, 250, 266, 272, 323, 360, 369, 374, 386, 402, 451 martingale, 423, 726 discrete time, 408submartingale, 415, 426, 727 supermartingale, 731 Matlab, 283, 634 matrilineality, 335 maze, 206, 221 MCMC, 285 MDP, 225, 234, 597 MDPtoolbox, 238, 240mean game duration, 85, 419 *hitting time*, *331*, *398* number of returns, 116, 211 MH3512 AY19-20

♦ 743

recurrence time, 254return time, 203 time to failure, 452memoryless property, 349, 449 Metropolis algorithm, 285 mixing time, 304model hidden Markov, 172, 181, 544 moment, 60, 435generating function, 433 monotone convergence, 420natural logarithm, 138, 349 negative binomial distribution, 41, 179, 328, 335, 470 inverse Gaussian process, 20 null recurrent, 255, 287 number of returns, 116, 211 Octave, 283, 634 order book, 399 $PageRank^{TM}, 302$ partial differential equation, 404 partition, 30, 438 Pascal distribution, 41 Pascal triangle, 652 patrilineality, 335 pattern recognition, 227 periodicity, 256 phase-type distribution, 178 Poisson central moments, 439 compound process, 17 cumulants, 435distribution, 41, 63moments, 435process, 16, 338, 356, 400, 431, 450 transformation, 436 stochastic integral, 433

Poisson moments, 439

policy, 226, 234 polylogarithmic, 313population genetics, 226 positive recurrence, 255, 272, 277, 287 Bernoulli-Laplace chain, 289 Ehrenfest chain, 288 success runs chain, 288 potential kernel, 216 power tail distribution, 313 probability absorption, 391 conditional, 28 density function, 33 joint, 37distribution, 33, 268 extinction, 318 generating function, 56, 314, 469 measure, 26ruin, 68, 415 space, 21 problemDirichlet, 190 process *birth*, **346** birth and death, 352, 383 branching, 310 Cauchy, 18 compound Poisson, 17 counting, 338 *Cox*, <u>451</u> death, 350*qambling*, 65, 158 qamma, 17inverse Gaussian, 19 negative inverse Gaussian, 20 Poisson, 16, 338, 356, 400, 431, 450 renewal. 452spatial Poisson, 431 stable, 18 stopped, 412 variance gamma, 19 Python code, 167, 170, 330, 374, 385

R code, 44, 48, 167, 170, 185, 245, 306, 330, 344, 374, 385, 622 R package HMM, 185 markovchain, 245 MDPtoolbox, 238, 240 **ॳ**744

randomvariable, 30walk, 106, 134, 143, 515 random sum, 52, 463 random walk elephant, 174excited, 136, 139, 520, 524 range process, 132 rank aggregation, 300rebound, 100recurrence, 134, 515 recurrent, 246class, 246, 248 null, 255 positive, 255random walk, 134 state, 246, 266, 287 reducible, 244reflected path, 128, 503, 505 reflection principle, 128, 503, 505 reflexive (relation), 243 regular transition matrix, 266, 287, 666 reinforcement learning, 234 relation equivalence, 243reflexive, 243 symmetric, 243 transitive, 243 reliability function, 449 renewal process, 452 resolvent, 216 return probabilities, 211 time, 110, 203 reversibility, 297 condition. 285reversible Markov chain, 285, 298 ruin probability, 68, 415 search engine, 300, 302 second chance, 100 semigroup, 354property, 357sink state, 230 Skellam distribution, 720 snake

> and ladders, 154, 216 game, 173

spatial Poisson process, 431 spin, 305 spoofing, 399 St. Petersburg paradox, 45, 92, 121 stable process, 18 stateabsorbing, 188, 256, 287 accepting, 230 accessible, 243 aperiodic, 256, 266, 272, 274, 287 communicating, 242, 244 ergodic, 256, 287 null recurrent, 255, 287 positive recurrent, 255, 272, 277, 287 recurrent, 246, 266, 287 sink, 230 transient, 248, 287 stationary distribution, 268, 284, 376, 398 increments, 340stationary distribution, 299 Bernoulli-Laplace chain, 289 Ehrenfest chain, 288 success runs chain, 288 stickiness, 166 Stirling approximation, 130, 135, 276 number. 439stochastic dynamic programming, 225, 234, 597 stopped process, 412 stopping time, 410theorem, 413streak (winning), 127 strong Markov property, 220, 338, 344, 350, 636 strongly connected (graph theory), 243 subcritical, 318 success runs Markov chain, 176, 264, 288 supercritical, 317 survival probability, 448 symmetric (relation), 243 synchronizing automata, 230

telescoping sum, 79, 131, 132, 276 thinning, 62, 466 time homogeneous, 68, 144, 270, 354 time series, 14 tower property, 50, 54, 408, 409 trading, 399 transience, 248transient state, 248, 287 transitionmatrix, 144semigroup, 354transitive (relation), 243 two-state Markov chain, 160, 192, 198, 205, 250, 266, 272, 323, 360, 369, 374, 386, 402, 451 uniform distribution, 35unsupervised learning, 181 utility function, 198, 225 value function, 238 variance, 51, 59 gamma process, 19 voter model, 305Weibull distribution, 453 winning streaks, 127 Wright-Fisher model, 226 zero-modified distribution, 213

Author index

ॳ747

Abiola, S. 177 Agapie, A. 305 Althoen, S.C. 216 Anděl, J. 103 Antal, T. 136, 139 Asmussen, S. 255 Bagchi, S. 313 Benjamini, I. 136 Bhattacharya, B.B. 674 Bosq, D. 277, 278, 341, 721 Champion, W.L. 515 Decreusefond, L. 442 Devore, J.L. 21 Doob, J.L. 413, 426 Elevoowo, I.O. 177 Ferraz, E. 442 Gusev, V.V. 230, 294, 592 Höns, R. 305 Houdré, C. 442 Hudecová, Š. 103 Iribarren, J.L. 313 Jacod, J. 21, 468 Karlin, S. 266, 267, 272, 318, 379, 452 Kijima, M. 255

King, L. 216

Laforgia, A. 721 Lalley, S.P. 379 Lee, D.D. 737 Levin, D.A. 267, 302, 660

Martins, P. 442 Mecke, J. 439 Meyer, P.A. 426 Mills, T.M. 515 Moro, E. 313 Mukherjee, S. 674 Muni Toke, I. 399, 699

Nguyen, H.T. 277, 278, 341, 721 Norris, J.R. 163, 279, 344

Olaleye, T. 177

Pasquali, C. 153 Peres, Y. 267, 302, 660 Pitman, J. 21 Polya, G. 515 Protter, P. 21, 468

Redner, S. 136, 139

Salako, M.O. 177 Schalekamp, F. 300 Schilling, K. 216 Schütz, G.M. 174 Sellke, S.H. 313 Shroff, N.B. 313 Sigman, K. 379 Smith, S.J. 515 Sowunmi, F.A. 177 Stocker, A.A. 737

Taylor, H.M. 266, 267, 272, 318, 379, 452

Trimper, S. 174

van Zuylen, A. 300

Wang, Z. 737 Wilmer, E.L. 267, 302, 660 Wilson, D.B. 136 Wu, L.M. 442

Bibliography

- [AH07] A. Agapie and R. Höns. Analysis of a voter model. Math. Rep. (Bucur.), 9(59)(2):135–145, 2007. (Cited on page 305.)
- [AH12] J. Anděl and Š. Hudecová. Variance of the game duration in the gambler's ruin problem. *Statist. Probab. Lett.*, 82(9):1750–1754, 2012. (Cited on page 103.)
- [AKS93] S.C. Althoen, L. King, and K. Schilling. How long is a game of Snakes and Ladders? *The Mathematical Gazette*, 77(478):71–76, 1993. (Cited on page 216.)
- [AR05] T. Antal and S. Redner. The excited random walk in one dimension. J. Phys. A, 38(12):2555-2577, 2005. (Cited on pages 136, 139, and 525.)
- [Asm03] S. Asmussen. Applied probability and queues, volume 51 of Applications of Mathematics (New York). Springer-Verlag, New York, second edition, 2003. Stochastic Modelling and Applied Probability. (Cited on page 255.)
- [BM18] B.B. Bhattacharya and S. Mukherjee. Inference in Ising models, 2018. (Cited on page 674.)
- [BN96] D. Bosq and H.T. Nguyen. A Course in Stochastic Processes: Stochastic Models and Statistical Inference. Mathematical and Statistical Methods. Kluwer, 1996. (Cited on pages 277, 341, and 721.)
- [BW03] I. Benjamini and D.B. Wilson. Excited random walk. *Electron. Comm. Probab.*, 8:86–92 (electronic), 2003. (Cited on page 136.)
- [Çın11] Erhan Çınlar. Probability and stochastics, volume 261 of Graduate Texts in Mathematics. Springer, New York, 2011. (Cited on page 33.)
- [CMS07] W.L. Champion, T.M. Mills, and S.J. Smith. Lost in space. Math. Scientist, 32:88–96, 2007. (Cited on page 515.)
- [Dev03] J. L. Devore. *Probability and Statistics for Engineering and the Sciences*. Duxbury Press, sixth edition, 2003. (Cited on page 21.)

- [DFM09] L. Decreusefond, E. Ferraz, and P. Martins. Upper bound of loss probability for the dimensioning of OFDMA systems with multi class randomly located users. In *WiOpt, workshop SPASWIN 2009. Seoul, South Korea*, 2009. (Cited on page 442.)
- [Doo53] J.L. Doob. Stochastic processes. John Wiley & Sons Inc., New York, 1953. (Cited on page 429.)
- [Doo84] J.L. Doob. Classical potential theory and its probabilistic counterpart. Springer-Verlag, Berlin, 1984. (Cited on page 429.)
- [Dur99] R. Durrett. Essentials of Stochastic Processes. Springer, 1999.
- [Gus14] V.V. Gusev. Synchronizing automata with random inputs. In *Developments in Language Theory*, volume 8633 of *Lecture Notes in Computer Science*, pages 68–75. Springer, New York, 2014. (Cited on pages 230, 294, and 592.)
- [HP02] C. Houdré and N. Privault. Concentration and deviation inequalities in infinite dimensions via covariance representations. Bernoulli, 8(6):697–720, 2002. (Cited on page 442.)
- [IM11] J.L. Iribarren and E. Moro. Branching dynamics of viral information spreading. Phys. Rev. E, 84:046116, 13, 2011. (Cited on page 313.)
- [JP00] J. Jacod and P. Protter. *Probability essentials*. Springer-Verlag, Berlin, 2000. (Cited on pages 21 and 468.)
- [JS01] P.W. Jones and P. Smith. *Stochastic Processes: An Introduction*. Arnold Texts in Statistics. Hodder Arnold, 2001.
- [Kij97] M. Kijima. Markov processes for stochastic modeling. Stochastic Modeling Series. Chapman & Hall, London, 1997. (Cited on page 255.)
- [KT81] S. Karlin and H.M Taylor. A Second Course in Stochastic Processes. Academic Press Inc., New York, 1981. (Cited on pages 266, 267, 272, 318, 379, and 452.)
- [Laf91] A. Laforgia. Bounds for modified Bessel functions. J. Comput. Appl. Math., 34(3):263–267, 1991. (Cited on page 721.)
- [Lal13] S.P. Lalley. Continuous time Markov chains. http://galton.uchicago.edu/~lalley/Courses/313/ContinuousTime.pdf, 2013. Accessed: 2015-12-10. (Cited on page 379.)
- [LPW09] D.A. Levin, Y. Peres, and E.L. Wilmer. Markov Chains and Mixing Times. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson. (Cited on pages 267, 302, 660, and 666.)
- [Luk70] E. Lukacs. *Characteristic Functions*. Hafner Publishing Co., New York, 1970. Second edition, revised and enlarged. (Cited on page 61.)

- [Mec67] J. Mecke. Stationäre zufällige Masse auf lokalkompakten Abelschen Gruppen. Z. Wahrscheinlichkeitstheorie Verw. @eb., 9:36–58, 1967. (Cited on page 439.)
- [MT15] I. Muni Toke. The order book as a queueing system: average depth and influence of the size of limit orders. *Quantitative Finance*, 15(5):795–808, 2015. (Cited on pages 399 and 699.)
- [Nor98] J.R. Norris. *Markov Chains*, volume 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. Reprint of 1997 original. (Cited on pages 163, 279, and 344.)
- [OSA⁺09] T. Olaleye, F.A. Sowunmi, S. Abiola, M.O. Salako, and I.O. Eleyoowo. A Markov chain approach to the dynamics of vehicular traffic characteristics in Abeokuta metropolis. *Research Journals of Applied Sciences, Engineering and Technology*, 1(3):160–166, 2009. (Cited on page 177.)
- [Pas14] C. Pasquali. On Markov Chain Music. The Pasqualian Mathematics, Poetry and Music. http://thepasqualian.com/?p= 1831, 2014. Accessed: 2014-12-08. (Cited on page 153.)
- [Pit99] J. Pitman. *Probability*. Springer, 1999. (Cited on page 21.)
- [Pol21] G. Polya. Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffrend die Irrfahrt im Strassenetz. Mathematische Annalen, 84:149–160, 1921. (Cited on page 515.)
- [Pri08] N. Privault. Potential theory in classical probability. In U. Franz and M. Schürmann, editors, *Quantum potential theory*, volume 1954 of *Lecture Notes in Math.*, pages 3–59. Springer, Berlin, 2008. (Cited on pages 190 and 201.)
- [Pri09] N. Privault. Stochastic analysis in discrete and continuous settings with normal martingales, volume 1982 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2009. (Cited on pages 433, 442, and 724.)
- [Pri11] N. Privault. Generalized Bell polynomials and the combinatorics of Poisson central moments. *Electron. J. Combin.*, 18(1):Research Paper 54, 10, 2011. (Cited on page 439.)
- [Pri12] N. Privault. Moments of Poisson stochastic integrals with random integrands. *Probability and Mathematical Statistics*, 32(2):227–239, 2012. (Cited on page 440.)
- [Sig14] K. Sigman. Continuous-time Markov chains. http://www.columbia.edu/~ks20/stochastic-I/stochastic-I-CTMC.pdf, 2014. Accessed: 2014-12-09. (Cited on page 379.)
- [SSB08] S.H. Sellke, N.B. Shroff, and S. Bagchi. Modeling and automated containment of worms. *IEEE Transactions on dependable* and secure computing, 5(2):71–86, 2008. (Cited on page 313.) _{MH3512 AY19-20}

- [ST08] G.M. Schütz and S. Trimper. Elephants can always remember: Exact long-range memory effects in a non-Markovian random walk. *Phys. Rev. E (3)*, 2008. (Cited on page 174.)
- [Ste01] J.M. Steele. Stochastic Calculus and Financial Applications, volume 45 of Applications of Mathematics. Springer-Verlag, New York, 2001.
- [SvZ09] F. Schalekamp and A. van Zuylen. Rank aggregation: Together we're strong. In 2009 Proceedings of the Eleventh Workshop on Algorithm Engineering and Experiments (ALENEX09), pages 38–51. SIAM, 2009. (Cited on page 300.)
- [WSL12] Z. Wang, A.A. Stocker, and D.D. Lee. Optimal neural tuning curves for arbitrary stimulus distributions: Discrimax, infomax and minimum l^p loss. In Advances in Neural Information Processing Systems (NIPS 2012), volume 25, pages 1–9, 2012. (Cited on page 737.)
- [Wu00] L.M. Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probab. Theory Related Fields*, 118(3):427–438, 2000. (Cited on page 442.)