

Nicolas Privault

# Understanding Markov Chains

Examples and Applications

Third Edition

Solutions Manual

October 22, 2026



# Contents

|                                         |            |
|-----------------------------------------|------------|
| <b>Solutions to the Exercises</b> ..... | <b>1</b>   |
| Chapter 1 .....                         | 1          |
| Chapter 2 .....                         | 14         |
| Chapter 3 .....                         | 31         |
| Chapter 4 .....                         | 51         |
| Chapter 5 .....                         | 72         |
| Chapter 6 .....                         | 112        |
| Chapter 7 .....                         | 118        |
| Chapter 8 .....                         | 159        |
| Chapter 9 .....                         | 175        |
| Chapter 10 .....                        | 204        |
| Chapter 11 .....                        | 216        |
| Chapter 12 .....                        | 221        |
| <b>References</b> .....                 | <b>223</b> |



# Solutions to the Exercises

## Chapter 1 - Probability Background

Exercise 1.1 Assuming that (1.2.2) holds, by the complement rule and using the relation

$$\sum_{\emptyset \neq I \subset J} (-1)^{|I|+1} = 1 - \sum_{I \subset J} (-1)^{|I|} = 1$$

we have

$$\begin{aligned} \mathbb{P} \left( \bigcap_{j \in J} A_j \right) &= \mathbb{P} \left( \left( \bigcup_{j \in J} A_j^c \right)^c \right) \\ &= 1 - \mathbb{P} \left( \bigcup_{j \in J} A_j^c \right) \\ &= 1 - \sum_{I \subset J} (-1)^{|I|+1} \mathbb{P} \left( \bigcap_{i \in I} A_i^c \right) \\ &= 1 - \sum_{\emptyset \neq I \subset J} (-1)^{|I|+1} \mathbb{P} \left( \bigcap_{i \in I} A_i^c \right) \\ &= \sum_{I \subset J} (-1)^{|I|+1} \left( 1 - \mathbb{P} \left( \bigcap_{i \in I} A_i^c \right) \right) \\ &= \sum_{I \subset J} (-1)^{|I|+1} \mathbb{P} \left( \bigcup_{i \in I} A_i \right), \end{aligned}$$

which proves (1.2.3).

## Exercise 1.2

a) We have

$$\mathbb{P}(X < \infty) = \sum_{k \geq 0} \mathbb{P}(X = k) = q \sum_{k \geq 0} p^k = \frac{q}{1-p},$$

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

$$\mathbb{E}[Y] = \sum_{k \geq 0} r^k \mathbb{P}(Y = r^k) = \sum_{k \geq 0} r^k \mathbb{P}(X = k) = (1-p) \sum_{k \geq 0} (pr)^k,$$

hence  $\mathbb{E}[Y] = (1-p)/(1-pr) < \infty$  when  $r < 1/p$ , and  $\mathbb{E}[Y] = +\infty$  when  $r \geq 1/p$ .

Exercise 1.3 We write

$$Z = \sum_{k=1}^N X_k$$

where  $\mathbb{P}(N = n) = 1/6$ ,  $n = 1, 2, \dots, 6$ , and  $X_k$  is a Bernoulli random variable with parameter  $1/2$ ,  $k = 1, 2, \dots, 6$ .

a) We have

$$\begin{aligned} \mathbb{E}[Z] &= \mathbb{E}[\mathbb{E}[Z \mid N]] = \sum_{n=1}^6 \mathbb{E}[Z \mid N = n] \mathbb{P}(N = n) \\ &= \sum_{n=1}^6 \mathbb{E} \left[ \sum_{k=1}^N X_k \mid N = n \right] \mathbb{P}(N = n) = \sum_{n=1}^6 \mathbb{E} \left[ \sum_{k=1}^n X_k \right] \mathbb{P}(N = n) \\ &= \sum_{n=1}^6 \sum_{k=1}^n \mathbb{E}[X_k] \mathbb{P}(N = n) = \frac{1}{6} \sum_{n=1}^6 \sum_{k=1}^n \frac{1}{2} = \frac{1}{2 \times 6} \sum_{n=1}^6 n \\ &= \frac{1}{2 \times 6} \times \frac{6 \times (6+1)}{2} = \frac{7}{4}, \end{aligned} \tag{S.1}$$

where we applied (A.6). Concerning the variance, we have

$$\begin{aligned} \mathbb{E}[Z^2] &= \mathbb{E}[\mathbb{E}[Z^2 \mid N]] \\ &= \sum_{n=1}^6 \mathbb{E}[Z^2 \mid N = n] \mathbb{P}(N = n) \end{aligned}$$

$$\begin{aligned}
&= \sum_{n=1}^6 \mathbb{E} \left[ \left( \sum_{k=1}^n X_k \right)^2 \mid N = n \right] \mathbb{P}(N = n) \\
&= \sum_{n=1}^6 \mathbb{E} \left[ \sum_{k=1}^n X_k \sum_{l=1}^n X_l \mid N = n \right] \mathbb{P}(N = n) \\
&= \sum_{n=1}^6 \mathbb{E} \left[ \sum_{k,l=1}^n X_k X_l \mid N = n \right] \mathbb{P}(N = n) \\
&= \sum_{n=1}^6 \mathbb{E} \left[ \sum_{k=1}^n X_k^2 + \sum_{1 \leq k \neq l \leq n} X_k X_l \right] \mathbb{P}(N = n) \\
&= \sum_{n=1}^6 \mathbb{E} \left[ \sum_{k=1}^n X_k^2 \right] \mathbb{P}(N = n) + \sum_{n=1}^6 \mathbb{E} \left[ \sum_{1 \leq k \neq l \leq n} X_k X_l \right] \mathbb{P}(N = n) \\
&= \sum_{n=1}^6 \sum_{k=1}^n \mathbb{E} [X_k^2] \mathbb{P}(N = n) + \sum_{n=1}^6 \sum_{1 \leq k \neq l \leq n} \mathbb{E} [X_k X_l] \mathbb{P}(N = n) \\
&= \sum_{n=1}^6 \sum_{k=1}^n \mathbb{E} [X_k^2] \mathbb{P}(N = n) + \sum_{n=1}^6 \sum_{1 \leq k \neq l \leq n} \mathbb{E} [X_k] \mathbb{E} [X_l] \mathbb{P}(N = n) \\
&= \frac{1}{2 \times 6} \sum_{n=1}^6 n + \frac{1}{6 \times 2^2} \sum_{n=1}^6 n(n-1) \\
&= \frac{1}{2 \times 6} \sum_{n=1}^6 n + \frac{1}{6 \times 2^2} \sum_{n=1}^6 n^2 - \frac{1}{6 \times 2^2} \sum_{n=1}^6 n \\
&= \frac{1}{6 \times 2^2} \sum_{n=1}^6 n + \frac{1}{6 \times 2^2} \sum_{n=1}^6 n^2 \\
&= \frac{1}{6 \times 2^2} \frac{6(6+1)}{2} + \frac{1}{6 \times 2^2} \frac{6(6+1)(2 \times 6 + 1)}{6} = \frac{14}{3}, \tag{S.2}
\end{aligned}$$

where we used (A.7) and (S.1), hence

$$\text{Var}[Z] = \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2 = \frac{14}{3} - \frac{49}{16} = \frac{77}{48}. \tag{S.3}$$

Using (1.6.10) and  $\text{Var}[X_k] = p(1-p)$  we could also write

$$\begin{aligned}
\text{Var}[Z \mid N = n] &= \text{Var} \left[ \sum_{k=1}^N X_k \mid N = n \right] \\
&= \text{Var} \left[ \sum_{k=1}^n X_k \right]
\end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^n \text{Var}[X_k] \\
&= np(1-p),
\end{aligned}$$

which implies

$$\mathbb{E}[Z^2 \mid N = n] = \text{Var}[Z \mid N = n] + (\mathbb{E}[Z \mid N = n])^2 = np(1-p) + n^2 p^2,$$

hence

$$\begin{aligned}
\mathbb{E}[Z^2] &= \sum_{k=1}^6 \mathbb{E}[Z^2 \mid N = n] \mathbb{P}(N = n) \\
&= \frac{1}{6} \sum_{k=1}^6 \mathbb{E}[Z^2 \mid N = n] \\
&= \frac{1}{6} \sum_{k=1}^6 (np - np^2 + n^2 p^2),
\end{aligned}$$

and recovers (S.2).

b) Applying (1.3.1) with  $B = \{Z = k\}$ , we have

$$\begin{aligned}
\mathbb{P}(Z = l) &= \mathbb{P}\left(\sum_{k=1}^N X_k = l\right) \\
&= \sum_{n=1}^6 \mathbb{P}\left(\sum_{k=1}^N X_k = l \mid N = n\right) \mathbb{P}(N = n) \\
&= \sum_{n=1}^6 \mathbb{P}\left(\sum_{k=1}^n X_k = l\right) \mathbb{P}(N = n).
\end{aligned}$$

Next, we notice that the deterministic summation  $\sum_{k=1}^n X_k$  has a binomial distribution with parameter  $(n, 1/2)$ , *i.e.*

$$\mathbb{P}\left(\sum_{k=1}^n X_k = l\right) = \begin{cases} \binom{n}{l} \left(\frac{1}{2}\right)^l \left(\frac{1}{2}\right)^{n-l} & \text{if } l = 0, 1, \dots, n, \\ 0 & \text{if } l > n, \end{cases}$$

which yields

$$\mathbb{P}(Z = 0) = \sum_{n=1}^6 \mathbb{P}\left(\sum_{k=1}^n X_k = 0\right) \mathbb{P}(N = n)$$

$$\begin{aligned}
&= \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n \\
&= \frac{1}{6 \times 2} \sum_{n=0}^5 \left(\frac{1}{2}\right)^n \\
&= \frac{1 - (1/2)^6}{6},
\end{aligned}$$

and, for  $l = 1, 2, \dots, 6$ ,

$$\begin{aligned}
\mathbb{P}(Z = l) &= \sum_{n=1}^6 \mathbb{P}\left(\sum_{k=1}^n X_k = l\right) \mathbb{P}(N = n) \\
&= \frac{1}{6} \sum_{n=l}^6 \left(\frac{1}{2}\right)^l \left(\frac{1}{2}\right)^{n-l} \binom{n}{l} \\
&= \frac{1}{6} \sum_{n=l}^6 \binom{n}{l} \left(\frac{1}{2}\right)^n, \quad l = 1, 2, \dots, 6,
\end{aligned}$$

hence

$$\mathbb{P}(Z = l) = \frac{1}{6} \sum_{n=\max(1,l)}^6 \binom{n}{l} \left(\frac{1}{2}\right)^n, \quad l = 0, 1, \dots, 6.$$

Note that we have

$$\begin{aligned}
\sum_{l=0}^6 \mathbb{P}(Z = l) &= \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n + \frac{1}{6} \sum_{l=1}^6 \sum_{n=l}^6 \binom{n}{l} \left(\frac{1}{2}\right)^n \\
&= \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n + \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n \sum_{l=1}^n \binom{n}{l} \\
&= \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n \sum_{l=0}^n \binom{n}{l} \\
&= 1,
\end{aligned}$$

where we used the identity  $\sum_{l=0}^n \binom{n}{l} = (1+1)^n = 2^n$ .

c) We have

$$\mathbb{E}[Z] = \sum_{l=0}^6 l \mathbb{P}(Z = l) = \frac{1}{6} \sum_{l=0}^6 l \sum_{n=l}^6 \left(\frac{1}{2}\right)^n \binom{n}{l}$$

$$\begin{aligned}
&= \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n \sum_{l=1}^n l \binom{n}{l} = \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n \sum_{l=1}^n \frac{n!}{(n-l)!(l-1)!} \\
&= \frac{1}{6} \sum_{n=1}^6 n \left(\frac{1}{2}\right)^n \sum_{l=0}^{n-1} \frac{(n-1)!}{(n-1-l)!l!} = \frac{1}{6} \sum_{n=1}^6 n \left(\frac{1}{2}\right)^n \sum_{l=0}^{n-1} \binom{n-1}{l} \\
&= \frac{1}{6 \times 2} \sum_{n=1}^6 n = \frac{1}{6 \times 2} \frac{6 \times (6+1)}{2} = \frac{7}{4},
\end{aligned}$$

which recovers (S.1), where we applied (A.5). We also have

$$\begin{aligned}
\mathbb{E}[Z^2] &= \sum_{l=0}^6 l^2 \mathbb{P}(Z=l) \\
&= \frac{1}{6} \sum_{l=0}^6 l^2 \sum_{n=l}^6 \left(\frac{1}{2}\right)^n \binom{n}{l} = \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n \sum_{l=1}^n l^2 \binom{n}{l} \\
&= \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n \sum_{l=1}^n \frac{(l-1+1)n!}{(n-l)!(l-1)!} \\
&= \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n \sum_{l=1}^n \frac{(l-1)n!}{(n-l)!(l-1)!} + \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n \sum_{l=1}^n \frac{n!}{(n-l)!(l-1)!} \\
&= \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n \sum_{l=2}^n \frac{n!}{(n-l)!(l-2)!} + \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n \sum_{l=1}^n \frac{n!}{(n-l)!(l-1)!} \\
&= \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n n(n-1) \sum_{l=0}^{n-2} \frac{(n-2)!}{(n-2-l)!l!} + \frac{1}{6} \sum_{n=1}^6 \left(\frac{1}{2}\right)^n \sum_{l=0}^{n-1} \frac{n!}{(n-1-l)!l!} \\
&= \frac{1}{6 \times 2^2} \sum_{n=1}^6 n(n-1) + \frac{1}{6 \times 2} \sum_{n=1}^6 n \\
&= \frac{1}{6 \times 2^2} \sum_{n=1}^6 n^2 - \frac{1}{6 \times 2^2} \sum_{n=1}^6 n + \frac{1}{6 \times 2} \sum_{n=1}^6 n \\
&= \frac{1}{6 \times 2^2} \sum_{n=1}^6 n^2 + \frac{1}{6 \times 2^2} \sum_{n=1}^6 n \\
&= \frac{1}{6 \times 2^2} \frac{6(6+1)(2 \times 6+1)}{6} + \frac{1}{6 \times 2^2} \frac{6(6+1)}{2} = \frac{14}{3},
\end{aligned}$$

which recovers (S.3).

#### Exercise 1.4

- a) We assume that the sequence of Bernoulli trials is represented by a family  $(X_k)_{k \geq 1}$  of independent Bernoulli random variables with

distribution

$$\mathbb{P}(X_k = 1) = p, \quad \mathbb{P}(X_k = 0) = 1 - p, \quad k \geq 1.$$

We have

$$Z = X_1 + X_2 + \cdots + X_N = \sum_{k=1}^N X_k,$$

and, since  $\mathbb{E}[X_k] = p$ ,

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

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

$$\begin{aligned} \mathbb{E}[N] &= \sum_{n \geq 0} n \mathbb{P}(N = n) = e^{-\lambda} \sum_{n \geq 0} n \frac{\lambda^n}{n!} = e^{-\lambda} \sum_{n \geq 1} \frac{\lambda^n}{(n-1)!} \\ &= \lambda e^{-\lambda} \sum_{n \geq 0} \frac{\lambda^n}{n!} = \lambda e^{-\lambda} e^{\lambda} = \lambda, \end{aligned} \tag{S.4}$$

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

$$\mathbb{E}[Z] = p\lambda.$$

Concerning the variance we have, since  $\mathbb{E}[X_k^2] = p$ ,

$$\begin{aligned} \mathbb{E}[Z^2] &= \mathbb{E}\left[\left(\sum_{k=1}^N X_k\right)^2\right] \\ &= \sum_{n \geq 0} \mathbb{E}\left[\left(\sum_{k=1}^n X_k\right)^2\right] \mathbb{P}(N = n) \\ &= \sum_{n \geq 0} \mathbb{E}\left[\left(\sum_{k=1}^n X_k\right)\left(\sum_{l=1}^n X_l\right)\right] \mathbb{P}(N = n) \\ &= \sum_{n \geq 0} \mathbb{E}\left[\sum_{k,l=1}^n X_k X_l\right] \mathbb{P}(N = n) \end{aligned}$$

$$\begin{aligned}
&= \sum_{n \geq 0} \mathbb{E} \left[ \left( \sum_{k=1}^n (X_k)^2 + \sum_{1 \leq k \neq l \leq n} X_k X_l \right) \right] \mathbb{P}(N = n) \\
&= \sum_{n \geq 0} \left( \mathbb{E} \left[ \sum_{k=1}^n (X_k)^2 \right] + \sum_{1 \leq k \neq l \leq n} \mathbb{E}[X_k] \mathbb{E}[X_l] \right) \mathbb{P}(N = n) \\
&= \sum_{n \geq 0} (np + n(n-1)p^2) \mathbb{P}(N = n) \\
&= p(1-p) \sum_{n \geq 0} n \mathbb{P}(N = n) + p^2 \sum_{n \geq 0} n^2 \mathbb{P}(N = n) \\
&= p(1-p) \mathbb{E}[N] + p^2 \mathbb{E}[N^2].
\end{aligned}$$

Again, the above equality holds without requiring that  $N$  has the Poisson distribution.

Next, we have

$$\begin{aligned}
\mathbb{E}[N^2] &= \sum_{n \geq 0} n^2 \mathbb{P}(N = n) = e^{-\lambda} \sum_{n \geq 0} n^2 \frac{\lambda^n}{n!} \\
&= e^{-\lambda} \sum_{n \geq 1} n \frac{\lambda^n}{(n-1)!} = e^{-\lambda} \sum_{n \geq 1} (n-1) \frac{\lambda^n}{(n-1)!} + e^{-\lambda} \sum_{n \geq 1} \frac{\lambda^n}{(n-1)!} \\
&= e^{-\lambda} \sum_{n \geq 2} \frac{\lambda^n}{(n-2)!} + e^{-\lambda} \sum_{n \geq 1} \frac{\lambda^n}{(n-1)!} = \lambda^2 e^{-\lambda} \sum_{n \geq 0} \frac{\lambda^n}{n!} + \lambda e^{-\lambda} \sum_{n \geq 0} \frac{\lambda^n}{n!} \\
&= \lambda + \lambda^2,
\end{aligned}$$

hence

$$\text{Var}[N] = \mathbb{E}[N^2] - (\mathbb{E}[N])^2 = \lambda, \quad (\text{S.5})$$

and

$$\begin{aligned}
\text{Var}[Z] &= \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2 \\
&= p(1-p) \mathbb{E}[N] + p^2 \mathbb{E}[N^2] - (p \mathbb{E}[N])^2 \\
&= p(1-p) \mathbb{E}[N] + p^2 \text{Var}[N] \\
&= \lambda p(1-p) + \lambda p^2 \\
&= p\lambda.
\end{aligned}$$

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

$$\mathbb{P}(Z = l) = \sum_{n \geq 0} \mathbb{P} \left( \sum_{k=1}^N X_k = l \mid N = n \right) \mathbb{P}(N = n)$$

$$\begin{aligned}
&= \sum_{n \geq 0} \mathbb{P} \left( \sum_{k=1}^n X_k = l \right) \mathbb{P}(N = n) \\
&= e^{-\lambda} \sum_{n \geq l} \binom{n}{l} p^l (1-p)^{n-l} \frac{\lambda^n}{n!} = \frac{p^l}{l!} e^{-\lambda} \sum_{n \geq l} \frac{1}{(n-l)!} (1-p)^{n-l} \lambda^n \\
&= \frac{(\lambda p)^l}{l!} e^{-\lambda} \sum_{n \geq 0} (1-p)^n \frac{\lambda^n}{n!} = \frac{(\lambda p)^l}{l!} e^{-\lambda} e^{(1-p)\lambda} = \frac{(\lambda p)^l}{l!} e^{-p\lambda},
\end{aligned}$$

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 (S.4) and (S.5) we have  $\mathbb{E}[Z] = \text{Var}[Z] = p\lambda$ .

**Exercise 1.5** 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 1.10:

$$\begin{aligned}
\mathbb{E}[\min(X, Y) \mid X < Y] &= \mathbb{E}[X \mid X < Y] = \frac{1}{\mathbb{P}(X < Y)} \mathbb{E}[X \mathbb{1}_{\{X < Y\}}] \\
&= \mu(\lambda + \mu) \int_0^\infty e^{-\mu y} \int_0^y x e^{-\lambda x} dx dy \\
&= (\lambda + \mu) \frac{\mu}{\lambda^2} \int_0^\infty e^{-\mu y} (1 - (1 + \lambda y) e^{-\lambda y}) dy \\
&= (\lambda + \mu) \frac{\mu}{\lambda^2} \left( \int_0^\infty e^{-\mu y} dy - \int_0^\infty e^{-\mu y} e^{-\lambda y} dy - \lambda \int_0^\infty e^{-\mu y} y e^{-\lambda y} dy \right) \\
&= (\lambda + \mu) \frac{\mu}{\lambda^2} \left( \frac{1}{\mu} - \frac{1}{\lambda + \mu} - \frac{\lambda}{(\lambda + \mu)^2} \right) \\
&= (\lambda + \mu) \frac{\mu}{\lambda^2} \left( \frac{(\lambda + \mu)^2}{\mu(\lambda + \mu)^2} - \frac{\mu(\lambda + \mu)}{\mu(\lambda + \mu)^2} - \frac{\lambda\mu}{\mu(\lambda + \mu)^2} \right) \\
&= (\lambda + \mu) \frac{\mu}{\lambda^2} \left( \frac{(\lambda + \mu)^2 - \mu(\lambda + \mu) - \lambda\mu}{\mu(\lambda + \mu)^2} \right) \\
&= \frac{1}{\lambda + \mu} = \mathbb{E}[\min(X, Y)].
\end{aligned}$$

Note that (1.8.10) could also be obtained as in (1.6.7) or from point (iv) page 42, since the random variable  $\min(X, Y)$  is actually independent of the event  $\{X < Y\}$ . Indeed, for any  $a > 0$  we have

$$\begin{aligned}
\mathbb{P}(\min(X, Y) > a \text{ and } X < Y) &= \lambda\mu \int_a^\infty e^{-\mu y} \int_a^y e^{-\lambda x} dx dy \\
&= \mu \int_a^\infty e^{-\mu y} (e^{-\lambda a} - e^{-\lambda y}) dy \\
&= \mu e^{-\lambda a} \int_a^\infty e^{-\mu y} dy - \mu \int_a^\infty e^{-(\lambda + \mu)y} dy
\end{aligned}$$

$$\begin{aligned}
&= \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)a} \\
&= \mathbb{P}(X < Y) \mathbb{P}(\min(X, Y) > a).
\end{aligned}$$

Exercise 1.6 Since  $U$  is uniformly distributed given  $L$  over the interval  $[0, L]$ , we have

$$\varphi_{U|L=y}(x) = \frac{1}{y} \mathbb{1}_{[0,y]}(x), \quad x \in \mathbb{R}, \quad y > 0,$$

hence by the definition (1.5.7) of the conditional density  $\varphi_{U|L=y}(x)$  we have

$$\begin{aligned}
\varphi_{(U,L)}(x, y) &= \varphi_{U|L=y}(x) \varphi_L(y) \\
&= \frac{1}{y} \mathbb{1}_{[0,y]}(x) y e^{-y} \mathbb{1}_{[0,\infty)}(y) \\
&= \mathbb{1}_{[0,y]}(x) \mathbb{1}_{[0,\infty)}(y) e^{-y}.
\end{aligned} \tag{S.6}$$

Next, we want to determine the density function  $(x, y) \mapsto \varphi_{(U,L-U)}(x, y)$ . For this, for all bounded functions  $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ , we rewrite the expectation

$$\mathbb{E}[h(U, L - U)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, z) \varphi_{(U,L-U)}(x, z) dx dz \tag{S.7}$$

as

$$\begin{aligned}
\mathbb{E}[h(U, L - U)] &= \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} h(x, y - x) \varphi_{(U,L)}(x, y) dy \right) dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, z) \varphi_{(U,L)}(x, x + z) dx dz,
\end{aligned} \tag{S.8}$$

using the change of variable  $(x, z) = (x, y - x)$ . Identifying (S.7) with (S.8) yields the relation

$$\varphi_{(U,L-U)}(x, z) = \varphi_{(U,L)}(x, x + z),$$

hence from (S.6) we get

$$\begin{aligned}
\varphi_{(U,L-U)}(x, z) &= \varphi_{(U,L)}(x, x + z) \\
&= \mathbb{1}_{[0,x+z]}(x) \mathbb{1}_{[0,\infty)}(x + z) e^{-x-z} \\
&= \mathbb{1}_{[0,\infty)}(x) \mathbb{1}_{[0,\infty)}(z) e^{-x-z},
\end{aligned}$$

since

$$\begin{aligned}
\{0 \leq x \leq x + z \ \& \ x + z \geq 0\} &\iff \{0 \leq x \leq x + z\} \\
&\iff \{x \geq 0 \ \& \ z \geq 0\}.
\end{aligned}$$

We refer to Theorem 12.7 page 92 of [Jacod and Protter \(2000\)](#) for the general version of the above change of variable formula.

### Exercise 1.7

- a) Assuming that  $X$  and  $Y$  are independent Poisson random variables with parameters  $\lambda$  and  $\mu$ , we have

$$\begin{aligned}\mathbb{P}(X + Y = n) &= \sum_{k=0}^n \mathbb{P}(X = k \text{ and } X + Y = n) \\ &= \sum_{k=0}^n \mathbb{P}(X = k \text{ and } Y = n - k) \\ &= \sum_{k=0}^n \mathbb{P}(X = k) \mathbb{P}(Y = n - k) = e^{-\lambda-\mu} \sum_{k=0}^n \frac{\lambda^k}{k!} \frac{\mu^{n-k}}{(n-k)!} \\ &= e^{-\lambda-\mu} \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} \lambda^k \mu^{n-k} = e^{-\lambda-\mu} \frac{(\lambda + \mu)^n}{n!}, \quad (\text{S.9})\end{aligned}$$

hence  $X + Y$  has a Poisson distribution with parameter  $\lambda + \mu$ . This result can be recovered from the probability generating function (1.7.4) of the Poisson distribution and Relation (1.7.9), as

$$G_{X+Y}(s) = G_X(s)G_Y(s) = e^{\lambda(s-1)}e^{\mu(s-1)} = e^{(\lambda+\mu)(s-1)}, \quad -1 \leq s \leq 1.$$

- b) We have

$$\begin{aligned}\mathbb{P}(X = k \mid X + Y = n) &= \frac{\mathbb{P}(X = k \text{ and } X + Y = n)}{\mathbb{P}(X + Y = n)} \\ &= \frac{\mathbb{P}(X = k \text{ and } Y = n - k)}{\mathbb{P}(X + Y = n)} = \frac{\mathbb{P}(X = k) \mathbb{P}(Y = n - k)}{\mathbb{P}(X + Y = n)} \\ &= e^{-\lambda} \frac{\lambda^k}{k!} e^{-\mu} \frac{\mu^{n-k}}{(n-k)!} \left( e^{-\lambda-\mu} \frac{(\lambda + \mu)^n}{n!} \right)^{-1} \\ &= \binom{n}{k} \left( \frac{\lambda}{\lambda + \mu} \right)^k \left( \frac{\mu}{\lambda + \mu} \right)^{n-k}, \quad (\text{S.10})\end{aligned}$$

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

$$\varphi_\Lambda(x) = \theta \mathbb{1}_{[0, \infty)}(x) e^{-\theta x}, \quad x \in \mathbb{R},$$

satisfying

$$\mathbb{P}(A \leq x) = \int_{-\infty}^x \varphi_\Lambda(y) dy = \int_0^x \varphi_\Lambda(y) dy, \quad x \in \mathbb{R},$$

and

$$d\mathbb{P}(\Lambda \leq x) = \varphi_\Lambda(x)dx,$$

we have

$$\begin{aligned}\mathbb{P}(X = k) &= \int_0^\infty \mathbb{P}(X = k \mid \Lambda = x) d\mathbb{P}(\Lambda \leq x) \\ &= \int_0^\infty \mathbb{P}(X = k \mid \Lambda = x) \varphi_\Lambda(x) dx \\ &= \theta \int_0^\infty \frac{x^k}{k!} e^{-(\theta+1)x} dx \\ &= \frac{\theta}{(\theta+1)^{k+1}} \int_0^\infty \frac{y^k}{k!} e^{-y} dy \\ &= \frac{\theta}{(\theta+1)^{k+1}} \\ &= \left(1 - \frac{1}{\theta+1}\right) \left(\frac{1}{\theta+1}\right)^k,\end{aligned}$$

*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.*,

$$\mathbb{P}(X+Y = k) = (k+1) \left(1 - \frac{1}{\theta+1}\right)^2 \left(\frac{1}{\theta+1}\right)^k, \quad k \in \mathbb{N}.$$

Hence we have

$$\begin{aligned}\mathbb{P}(X = k \mid X+Y = n) &= \frac{\mathbb{P}(X = k \text{ and } X+Y = n)}{\mathbb{P}(X+Y = n)} \\ &= \frac{\mathbb{P}(X = k \text{ and } Y = n-k)}{\mathbb{P}(X+Y = n)} \\ &= \frac{\mathbb{P}(X = k)\mathbb{P}(Y = n-k)}{\mathbb{P}(X+Y = n)} \\ &= \left(1 - \frac{1}{\theta+1}\right) \left(\frac{1}{\theta+1}\right)^k \left(1 - \frac{1}{\theta+1}\right) \left(\frac{1}{\theta+1}\right)^{n-k} \\ &\quad \times \frac{1}{n+1} \left(1 - \frac{1}{\theta+1}\right)^{-2} \left(\frac{1}{\theta+1}\right)^{-n} \\ &= \frac{1}{n+1}, \quad k = 0, 1, \dots, n,\end{aligned}$$

which shows that the distribution of  $X$  given  $X+Y = n$  is the discrete uniform distribution on  $\{0, 1, \dots, n\}$ .

d) In case  $X$  and  $Y$  have the same parameter, *i.e.*  $\lambda = \mu$ , we have

$$\mathbb{P}(X = k \mid X + Y = n) = \binom{n}{k} \frac{1}{2^n}, \quad k = 0, 1, \dots, n,$$

which becomes independent of  $\lambda$ . Hence, when  $\lambda$  is represented by a random variable  $\Lambda$  with probability density  $x \mapsto \varphi_\Lambda(x)$  on  $\mathbb{R}_+$ , from (S.10) we get

$$\begin{aligned} \mathbb{P}(X = k \mid X + Y = n) &= \int_0^\infty \mathbb{P}(X = k \mid X + Y = n \text{ and } \Lambda = x) d\mathbb{P}(\Lambda \leq x) \\ &= \int_0^\infty \mathbb{P}(X = k \mid X + Y = n \text{ and } \Lambda = x) \varphi_\Lambda(x) dx \\ &= \int_0^\infty \mathbb{P}(X = k \mid X + Y = n) \varphi_\Lambda(x) dx \\ &= \mathbb{P}(X = k \mid X + Y = n) \int_0^\infty \varphi_\Lambda(x) dx \\ &= \mathbb{P}(X = k \mid X + Y = n) = \binom{n}{k} \frac{1}{2^n}, \quad k = 0, 1, \dots, n. \end{aligned}$$

This relation holds regardless of the expression of the probability density  $\varphi_\Lambda$ , and in particular when  $\Lambda$  has an exponential distribution with parameter  $\theta > 0$ .

**Exercise 1.8** 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

$$\mathbb{P}(C_1 = R) = \mathbb{P}(C_1 = G) = \frac{1}{2},$$

and

$$\mathbb{P}(C_2 = R \text{ and } C_1 = R) = \frac{2}{3}, \quad \mathbb{P}(C_2 = R \text{ and } C_1 = G) = \frac{1}{3}.$$

On the other hand, we have

$$\begin{aligned} \mathbb{P}(C_2 = R) &= \mathbb{P}(C_2 = R \text{ and } C_1 = R) + \mathbb{P}(C_2 = R \text{ and } C_1 = G) \\ &= \mathbb{P}(C_2 = R \mid C_1 = R) \mathbb{P}(C_1 = R) + \mathbb{P}(C_2 = R \mid C_1 = G) \mathbb{P}(C_1 = G) \\ &= \frac{2}{3} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{2} = \frac{1}{2}, \end{aligned}$$

and

$$\begin{aligned} \mathbb{P}(C_2 = G) &= \mathbb{P}(C_2 = G \text{ and } C_1 = R) + \mathbb{P}(C_2 = G \text{ and } C_1 = G) \\ &= \mathbb{P}(C_2 = G \mid C_1 = R) \mathbb{P}(C_1 = R) + \mathbb{P}(C_2 = G \mid C_1 = G) \mathbb{P}(C_1 = G) \\ &= \frac{1}{3} \times \frac{1}{2} + \frac{2}{3} \times \frac{1}{2} = \frac{1}{2}. \end{aligned}$$

Finally, the probability we wish to compute is

$$\begin{aligned}\mathbb{P}(C_1 = R \mid C_2 = R) &= \frac{\mathbb{P}(C_1 = R \text{ and } C_2 = R)}{\mathbb{P}(C_2 = R)} \\ &= \mathbb{P}(C_2 = R \mid C_1 = R) \frac{\mathbb{P}(C_1 = R)}{\mathbb{P}(C_2 = R)} = \frac{2}{3} \times \frac{1/2}{1/2} = \frac{2}{3}.\end{aligned}$$

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 1.9

a) The probability that the system operates is given by

$$\mathbb{P}(X \geq 2) = \binom{3}{2} p^2 (1-p) + p^3 = 3p^2 - 2p^3,$$

where  $X$  is a binomial random variable with parameter  $(3, p)$ .

b) The probability that the system operates is given by

$$\begin{aligned}\int_0^1 \mathbb{P}(X \geq 2 \mid p = x) d\mathbb{P}(p \leq x) &= \int_0^1 \mathbb{P}(X \geq 2 \mid p = x) dx \\ &= \int_0^1 (3x^2 - 2x^3) dx \\ &= 3 \int_0^1 x^2 dx - 2 \int_0^1 x^3 dx \\ &= \frac{1}{2},\end{aligned}$$

similarly to Exercise 1.7-(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

$$f(k) = \frac{1}{2}f(k+1) + \frac{1}{2}f(k-1), \quad 1 \leq k \leq S-1, \quad (\text{S.11})$$

with the boundary conditions  $f(0) = 1$  and  $f(S) = 0$ , which is identical to Equation (2.2.19).

We refer to this equation as the *homogeneous equation*.

- b) According to the result of (2.2.20) in Section 2.2 we know that the general solution of (S.11) has the form

$$f(k) = C_1 + C_2 k, \quad k = 0, 1, \dots, S$$

and after taking into account the boundary conditions we find

$$f(k) = \frac{S-k}{S}, \quad k = 0, 1, \dots, S.$$

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

$$\begin{aligned} h(k) &= (1-2r)(1+h(k)) + r(1+h(k+1)) + r(1+h(k-1)) \\ &= 1 + (1-2r)h(k) + rh(k+1) + rh(k-1), \end{aligned}$$

hence the equation

$$h(k) = \frac{1}{2r} + \frac{1}{2}h(k+1) + \frac{1}{2}h(k-1), \quad 1 \leq k \leq S-1,$$

with the boundary conditions  $h(0) = 0$  and  $h(S) = 0$ , which is identical to (2.3.7a) by changing  $h(k)$  into  $2r \times h(k)$  in (2.3.7a). We note that  $r$  should be strictly positive, otherwise if  $r = 0$  we will have  $h(k) = \infty$  for all  $k = 1, 2, \dots, S-1$  and  $h(0) = h(S) = 0$ .

- d) After trying a solution of the form  $h(k) = Ck^2$  we find

$$Ck^2 = \frac{1}{2r} + \frac{1}{2}C(k+1)^2 + \frac{1}{2}C(k-1)^2,$$

hence  $C$  should be equal to  $C = -1/(2r)$ , hence  $k \mapsto -k^2/(2r)$  is a particular solution.

- e) Given the hint, the general solution has the form

$$h(k) = C_1 + C_2 k - \frac{k^2}{2r}, \quad k = 0, 1, \dots, S,$$

which gives

$$h(k) = \frac{k(S-k)}{2r}, \quad k = 0, 1, \dots, S, \quad (\text{S.12})$$

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 geometrically distributed random variable  $\tau$  representing the time  $k$  spent by remaining at the same state, with

$$\mathbb{P}(\tau = k) = 2r(1 - 2r)^{k-1}, \quad k \geq 1.$$

- f) Starting from any state  $\textcircled{k} \in \{1, 2, \dots, S-1\}$ , 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 (2.3.20), we find the probability

$$\frac{1 - (p/q)^{S-k}}{1 - (p/q)^{S-m}} = \frac{(q/p)^S - (q/p)^k}{(q/p)^S - (q/p)^m}. \quad (\text{S.13})$$

- b) Taking  $a = 0$  and  $b = m$  in (2.3.20), we find the probability

$$1 - \frac{1 - (p/q)^{m-k}}{1 - (p/q)^m} = \frac{1 - (q/p)^k}{1 - (q/p)^m}. \quad (\text{S.14})$$

- c) Taking  $k = m+1$  in (S.13), we find that the probability of returning to  $\textcircled{m}$  after starting from  $\boxed{m+1}$  is

$$\frac{(q/p)^S - (q/p)^{m+1}}{(q/p)^S - (q/p)^m}.$$

Similarly, taking  $k = m-1$  in (S.14), we find that the probability of returning to  $\textcircled{m}$  after starting from  $\boxed{m-1}$  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  $\textcircled{m}$  in finite time after starting from  $X_0 = m$  is

$$\begin{aligned} & p \frac{1 - (p/q)^{S-(m+1)}}{1 - (p/q)^{S-m}} + q \frac{1 - (q/p)^{m-1}}{1 - (q/p)^m} \\ &= \frac{p(1 - (p/q)^{S-(m+1)})(1 - (q/p)^m) + q(1 - (q/p)^{m-1})(1 - (p/q)^{S-m})}{(1 - (p/q)^{S-m})(1 - (q/p)^m)} \\ &= \frac{(p - q(p/q)^{S-m})(1 - (q/p)^m) + (q - p(q/p)^m)(1 - (p/q)^{S-m})}{(1 - (p/q)^{S-m})(1 - (q/p)^m)} \end{aligned}$$

$$\begin{aligned}
&= \frac{1 - 2p(q/p)^m - 2q(p/q)^{S-m} + (p/q)^{S-2m}}{(1 - (p/q)^{S-m})(1 - (q/p)^m)} \\
&= \frac{1 - 2p(q/p)^m - 2q(p/q)^{S-m} + (p/q)^{S-2m}}{1 - (q/p)^m - (p/q)^{S-m} + (p/q)^{S-2m}}, \quad 1 \leq m \leq S-1.
\end{aligned}$$

When  $S = 2$  and  $m = 1$  we find that the above probability is zero, as expected.

- d) Using (2.3.12) 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:

$$\begin{aligned}
&1 + \frac{p}{q-p} \left( 1 - (S-m) \frac{1-q/p}{1-(q/p)^{S-m}} \right) + \frac{q}{q-p} \left( m-1 - m \frac{1-(q/p)^{m-1}}{1-(q/p)^m} \right) \\
&= -\frac{p(S-m)}{q-p} \frac{1-q/p}{1-(q/p)^{S-m}} + \frac{mq}{q-p} \frac{(q/p)^m - (q/p)^{m-1}}{(q/p)^m - 1}, \quad 1 \leq m \leq S-1.
\end{aligned}$$

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

$$\frac{1}{2} \left( \frac{S-m-1}{S-m} \right) + \frac{1}{2} \left( \frac{m-1}{m} \right) = 1 - \frac{S}{2m(S-m)},$$

for the probability of coming back to state  $\textcircled{m}$  in finite time after starting from  $X_0 = m$ , which gives  $1 - 2/S$  when  $S = 2m$ ,  $m \geq 1$ . Similarly, the mean time to either come back to  $m$  or to 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  $m = 1, 2, \dots, S-1$ .

### Exercise 2.3

- a) Starting from state  $\textcircled{k}$  we can reach  $\textcircled{k+1}$  in one time step by going up with probability  $p$ , or we can go down to  $\textcircled{k-1}$  with probability  $q$ , in which case we have to reach  $\textcircled{k+1}$  without hitting state  $\textcircled{0}$  in two stages:

- i) first, reach  $\textcircled{k}$  from  $\textcircled{k-1}$ , with probability  $\mathbb{P}(\tau_k < \tau_0 \mid X_0 = k-1)$ ,
- ii) then, reach  $\textcircled{k+1}$  from  $\textcircled{k}$  with probability  $\mathbb{P}(\tau_{k+1} < \tau_0 \mid X_0 = k)$ .

This gives

$$\begin{aligned}
p_k &= \mathbb{P}(\tau_{k+1} < \tau_0 \mid X_0 = k) \\
&= p + q\mathbb{P}(\tau_k < \tau_0 \mid X_0 = k-1)\mathbb{P}(\tau_{k+1} < \tau_0 \mid X_0 = k) \\
&= p + qp_{k-1}p_k, \quad k = 1, 2, \dots, S-1.
\end{aligned}$$

b) We need to check that (2.3.22) is satisfied when  $p_k$  is given by

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

Starting from the right-hand side of (2.3.22), this means checking the following:

$$\begin{aligned}
p + q \frac{1 - (q/p)^{k-1}}{1 - (q/p)^k} \frac{1 - (q/p)^k}{1 - (q/p)^{k+1}} &= p + q \frac{1 - (q/p)^{k-1}}{1 - (q/p)^{k+1}} \\
&= \frac{p(1 - (q/p)^{k+1}) + q(1 - (q/p)^{k-1})}{1 - (q/p)^{k+1}} \\
&= \frac{p + q - p(q/p)^{k+1} - q(q/p)^{k-1}}{1 - (q/p)^{k+1}} \\
&= \frac{p + q - q(q/p)^k - p(q/p)^k}{1 - (q/p)^{k+1}} \\
&= \frac{1 - (q/p)^k}{1 - (q/p)^{k+1}}.
\end{aligned}$$

c) We have

$$\begin{aligned}
&\mathbb{P}(\tau_S < \tau_0 \mid X_0 = k) \\
&= \mathbb{P}(\tau_{k+1} < \tau_0 \mid X_0 = k) \\
&\quad \times \mathbb{P}(\tau_{k+2} < \tau_0 \mid X_0 = k+1) \times \dots \times \mathbb{P}(\tau_S < \tau_0 \mid X_0 = S-1) \\
&= \prod_{l=k}^{S-1} p_l = \prod_{l=k}^{S-1} \frac{1 - (q/p)^l}{1 - (q/p)^{l+1}} = \frac{1 - (q/p)^k}{1 - (q/p)^S}, \quad k = 0, 1, \dots, S,
\end{aligned}$$

which recovers (2.2.28).

d) In the symmetric case  $p = q = 1/2$ , the same reasoning as in Question (a) above shows that  $p_k$  should satisfy the equation

$$p_k = \frac{1}{2} + \frac{1}{2}p_{k-1}p_k, \quad k = 1, 2, \dots, S-1,$$

which rewrites as

$$p_k(2 - p_{k-1}) = 1, \quad k = 1, 2, \dots, S-1.$$

We check that the above equation is satisfied by

$$p_k := \frac{k}{k+1}, \quad k = 1, 2, \dots, S-1,$$

hence we find again

$$\begin{aligned} & \mathbb{P}(\tau_S < \tau_0 \mid X_0 = k) \\ &= \mathbb{P}(\tau_{k+1} < \tau_0 \mid X_0 = k) \\ & \quad \times \mathbb{P}(\tau_{k+2} < \tau_0 \mid X_0 = k+1) \times \cdots \times \mathbb{P}(\tau_S < \tau_0 \mid X_0 = S-1) \\ &= \prod_{l=k}^{S-1} p_l \\ &= p_k p_{k+1} p_{k+2} \cdots p_{S-1} \\ &= \prod_{l=k}^{S-1} \frac{l}{l+1} \\ &= \frac{k}{S}, \quad k = 0, 1, \dots, S, \end{aligned}$$

showing that

$$\mathbb{P}(\tau_0 < \tau_S \mid X_0 = k) = 1 - \frac{k}{S}, \quad k = 0, 1, \dots, S,$$

which recovers (2.2.13).

#### Exercise 2.4

a) We have

$$f(0) = \mathbb{P}(R_A \mid X_0 = 0) = q + pq.$$

b) We have

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

and

$$\begin{aligned} f(1) &= \mathbb{P}(R_A \mid X_0 = 1) \\ &= p\mathbb{P}(R_A \mid X_0 = 2) + q\mathbb{P}(R_A \mid X_0 = 0) \\ &= q^2 + pq^2. \end{aligned}$$

#### Exercise 2.5

a) Letting  $\varepsilon = r - 1$ , i.e.  $r = 1 + \varepsilon$ , we have

$$h_S(k) = \frac{1}{q-p} \left( k - S \frac{1-r^k}{1-r^S} \right) = \frac{1}{p\varepsilon} \left( k - S \frac{1-(1+\varepsilon)^k}{1-(1+\varepsilon)^S} \right)$$

$$\begin{aligned}
&= \frac{1}{p\varepsilon} \left( \frac{k - S \frac{\sum_{i=1}^k \binom{k}{i} \varepsilon^i}{\sum_{i=1}^S \binom{S}{i} \varepsilon^i}}{\sum_{i=1}^S \binom{S}{i} \varepsilon^i} \right) = \frac{1}{p\varepsilon} \left( \frac{k \sum_{i=1}^S \binom{S}{i} \varepsilon^i - S \sum_{i=1}^k \binom{k}{i} \varepsilon^i}{\sum_{i=1}^S \binom{S}{i} \varepsilon^i} \right) \\
&= \frac{1}{p} \left( \frac{kS(S-1)/2 - Sk(k-1)/2 + k \sum_{i=3}^S \binom{S}{i} \varepsilon^{i-2} - S \sum_{i=3}^k \binom{k}{i} \varepsilon^{i-2}}{S + \varepsilon \sum_{i=2}^S \binom{S}{i} \varepsilon^{i-2}} \right)
\end{aligned}$$

hence

$$\lim_{r \rightarrow 1} \frac{1}{q-p} \left( k - S \frac{1-r^k}{1-r^S} \right) = \lim_{\varepsilon \rightarrow 0} \frac{1}{p\varepsilon} \left( k - S \frac{1-(1+\varepsilon)^k}{1-(1+\varepsilon)^S} \right) = k(S-k), \quad (\text{S.15})$$

which recovers (2.3.19).

b) Letting  $\varepsilon = 1 - r$ , we have

$$\begin{aligned}
f_S(k) &= \frac{r^S - r^k}{r^S - 1} = \frac{(1+\varepsilon)^S - (1+\varepsilon)^k}{(1+\varepsilon)^S - 1} \\
&= \frac{\sum_{i=0}^S \binom{S}{i} \varepsilon^i - \sum_{i=0}^k \binom{k}{i} \varepsilon^i}{\sum_{i=0}^S \binom{S}{i} \varepsilon^i - 1} = \frac{S - k + \varepsilon^2 \sum_{i=2}^k \binom{k}{i} \varepsilon^{i-2} - \varepsilon^2 \sum_{i=2}^S \binom{S}{i} \varepsilon^{i-2}}{S + \varepsilon^2 \sum_{i=2}^S \binom{S}{i} \varepsilon^{i-2}}
\end{aligned}$$

hence

$$\lim_{r \rightarrow 1} \frac{r^S - r^k}{r^S - 1} = \lim_{\varepsilon \rightarrow 0} \frac{(1+\varepsilon)^S - (1+\varepsilon)^k}{(1+\varepsilon)^S - 1} = \frac{S-k}{S},$$

which recovers (2.2.22).

**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

$$f(k) = \frac{\alpha}{\alpha + \beta} f(k+1) + \frac{\beta}{\alpha + \beta} f(k-1), \quad 1 \leq k \leq S-1,$$

with  $f(0) = 1$  and  $f(S) = 0$ . Therefore this non-symmetric problem with draw can be reduced to a non-symmetric gambling process with-

out draw and probabilities  $p = \alpha/(\alpha + \beta)$  and  $q = \beta/(\alpha + \beta)$ , and solution

$$f(k) = \frac{(\beta/\alpha)^k - (\beta/\alpha)^S}{1 - (\beta/\alpha)^S}, \quad k = 0, 1, \dots, S,$$

obtained by (2.2.12), with  $\alpha \neq \beta$ . For the mean game duration, by first step analysis we find the equation

$$\begin{aligned} h(k) &= (1 - \alpha - \beta)(1 + h(k)) + \alpha(1 + h(k+1)) + \beta(1 + h(k-1)) \\ &= 1 + (1 - \alpha - \beta)h(k) + \alpha h(k+1) + \beta h(k-1), \end{aligned}$$

i.e.

$$h(k) = \frac{1}{\alpha + \beta} + \frac{\alpha}{\alpha + \beta}h(k+1) + \frac{\beta}{\alpha + \beta}h(k-1), \quad 1 \leq k \leq S-1,$$

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 (2.3.12), shows that

$$h(k) = \frac{1}{\beta - \alpha} \left( k - S \frac{1 - (\beta/\alpha)^k}{1 - (\beta/\alpha)^S} \right), \quad k = 0, 1, \dots, S, \quad (\text{S.16})$$

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 geometrically distributed random variable  $\tau$  with distribution

$$\mathbb{P}(\tau = k) = (\alpha + \beta)(1 - \alpha - \beta)^{k-1}, \quad k \geq 1.$$

When  $\alpha$  tends to  $\beta$ , the result of Exercise 2.1 can be recovered as on page 20. Indeed, the limit of (S.16) as  $\alpha \rightarrow \beta$  is given as in (S.15) by

$$h(k) = \frac{k(S-k)}{2\beta}, \quad k = 0, 1, \dots, S,$$

which is consistent with (S.12).

## Problem 2.7

a) We have

$$pg(k) = pg(k+1) + qg(k-1), \quad k = 1, 2, \dots, S-1, \quad (\text{S.17})$$

with

$$g(0) = pg(1) + qg(0) \quad (\text{S.18})$$

for  $k = 0$ , and the boundary condition  $g(S) = 1$ .

- b) We observe that the constant function  $g(k) = C$  is solution of both (S.17) and (S.18) and the boundary condition  $g(S) = 1$  yields  $C = 1$ , hence

$$g(k) = \mathbb{P}(W \mid X_0 = k) = 1$$

for all  $k = 0, 1, \dots, S$ .

- c) We have

$$h(k) = 1 + ph(k+1) + qh(k-1), \quad k = 1, 2, \dots, S-1, \quad (\text{S.19})$$

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

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

has the form

$$h(k) = C_1 + C_2(q/p)^k, \quad k = 1, 2, \dots, S-1,$$

and we can check that  $k \mapsto k/(p-q)$  is a particular solution. Hence the general solution of (S.19) has the form

$$h(k) = \frac{k}{q-p} + C_1 + C_2(q/p)^k, \quad k = 0, 1, \dots, S,$$

with

$$\begin{cases} 0 = h(S) = \frac{S}{q-p} + C_1 + C_2(q/p)^S, \\ ph(0) = p(C_1 + C_2) = 1 + ph(1) = 1 + p \left( \frac{1}{q-p} + C_1 + C_2 \frac{q}{p} \right), \end{cases}$$

which yields

$$\begin{cases} C_1 = q \frac{(q/p)^S}{(p-q)^2} - \frac{S}{q-p}, \\ C_2 = -\frac{q}{(p-q)^2}, \end{cases}$$

and

$$h(k) = \mathbb{E}[T_S \mid X_0 = k] = \frac{S-k}{p-q} + \frac{q}{(p-q)^2} ((q/p)^S - (q/p)^k),$$

$$k = 0, 1, \dots, S.$$

Case  $p = q = 1/2$ . The solution of the homogeneous equation is given by

$$h(k) = C_1 + C_2 k, \quad k = 1, 2, \dots, S-1,$$

and the general solution of (S.19) has the form

$$h(k) = -k^2 + C_1 + C_2 k, \quad k = 1, 2, \dots, S,$$

with

$$\begin{cases} 0 = h(S) = -S^2 + C_1 + C_2 S, \\ \frac{h(0)}{2} = \frac{C_1}{2} = 1 + \frac{h(1)}{2} = 1 + \frac{C_1 + C_2 - 1}{2}, \end{cases}$$

hence

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

which yields

$$h(k) = \mathbb{E}[T_S \mid X_0 = k] = (S + k + 1)(S - k), \quad k = 0, 1, \dots, S.$$

e) When  $p \neq q$  we have

$$p_k := \mathbb{P}(T_S < T_0 \mid X_0 = k) = \frac{1 - (q/p)^k}{1 - (q/p)^S}, \quad k = 0, 1, \dots, S,$$

and when  $p = q = 1/2$  we find

$$p_k = \frac{k}{S}, \quad k = 0, 1, \dots, S.$$

f) The equality holds because, given that we start from state  $\boxed{k+1}$  at time 1, whether  $T_S < T_0$  or  $T_S > T_0$  does not depend on the past of the process before time 1. In addition it does not matter whether we start from state  $\boxed{k+1}$  at time 1 or at time 0.

g) We have

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

$k = 0, 1, \dots, S-1$ . By the result of Question (e), when  $p \neq q$  we find

$$\mathbb{P}(X_1 = k + 1 \mid X_0 = k \text{ and } T_S < T_0) = p \frac{1 - (q/p)^{k+1}}{1 - (q/p)^k},$$

$k = 1, 2, \dots, S - 1$ , and in case  $p = q = 1/2$  we get

$$\mathbb{P}(X_1 = k + 1 \mid X_0 = k \text{ and } T_S < T_0) = \frac{k + 1}{2k},$$

$k = 1, 2, \dots, S - 1$ . Note that this probability is higher than  $p = 1/2$ .

h) Similarly, we have

$$\begin{aligned} \mathbb{P}(X_1 = k - 1 \mid X_0 = k \text{ and } T_0 < T_S) &= \frac{\mathbb{P}(X_1 = k - 1, X_0 = k \text{ and } T_0 < T_S)}{\mathbb{P}(X_0 = k \text{ and } T_0 < T_S)} \\ &= \frac{\mathbb{P}(T_0 < T_S \mid X_1 = k - 1 \text{ and } X_0 = k) \mathbb{P}(X_1 = k - 1 \text{ and } X_0 = k)}{\mathbb{P}(T_0 < T_S \text{ and } X_0 = k)} \\ &= q \frac{\mathbb{P}(T_0 < T_S \mid X_0 = k - 1)}{\mathbb{P}(T_0 < T_S \mid X_0 = k)} \\ &= q \frac{1 - p_{k-1}}{1 - p_k}, \end{aligned}$$

$k = 1, 2, \dots, S - 1$ . When  $p \neq q$  this yields

$$\mathbb{P}(X_1 = k - 1 \mid X_0 = k \text{ and } T_0 < T_S) = q \frac{(q/p)^{k-1} - (q/p)^S}{(q/p)^k - (q/p)^S},$$

$k = 1, 2, \dots, S - 1$ , and when  $p = q = 1/2$  we find

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

$k = 1, 2, \dots, S - 1$ . Note that this probability is higher than  $q = 1/2$ .

i) We find

$$h(k) = 1 + p \frac{p_{k+1}}{p_k} h(k + 1) + \left(1 - p \frac{p_{k+1}}{p_k}\right) h(k - 1), \quad (\text{S.20})$$

$k = 1, 2, \dots, S - 1$ , or, due to the first step equation  $p_k = pp_{k+1} + qp_{k-1}$ ,

$$p_k h(k) = p_k + pp_{k+1} h(k + 1) + qp_{k-1} h(k - 1), \quad k = 1, 2, \dots, S - 1,$$

with the boundary condition  $h(S) = 0$ . When  $p = q = 1/2$  we have  $p_k = k/S$  by Question (e), hence (S.20) becomes

$$h(k) = 1 + \frac{k + 1}{2k} h(k + 1) + \frac{k - 1}{2k} h(k - 1),$$

$k = 1, 2, \dots, S-1$ .

j) We have to solve the equation

$$kh(k) = k + \frac{1}{2}(k+1)h(k+1) + \frac{1}{2}(k-1)h(k-1), \quad k = 1, 2, \dots, S-1,$$

with the boundary condition  $h(S) = 0$ . Letting  $g(k) := kh(k)$  we check that  $g(k)$  satisfies

$$g(k) = k + \frac{1}{2}g(k+1) + \frac{1}{2}g(k-1), \quad k = 1, 2, \dots, S-1, \quad (\text{S.21})$$

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 (S.21) has the form

$$g(k) = -\frac{1}{3}k^3 + C_1 + C_2k,$$

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

$$0 = g(S) = -\frac{1}{3}S^3 + C_1 + C_2S,$$

*i.e.*  $C_1 = 0$  and  $C_2 = S^2/3$ . Consequently, we have

$$g(k) = \frac{k}{3}(S^2 - k^2), \quad k = 0, 1, \dots, S,$$

hence we have

$$h(k) = \mathbb{E}[T_S \mid X_0 = k, T_S < T_0] = \frac{S^2 - k^2}{3}, \quad k = 1, 2, \dots, S.$$

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

$$g(k) = pg(k+2) + rg(k) + 2pg(k-1), \quad 1 \leq k \leq S-1,$$

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  $\lambda$  must satisfy

$$\lambda^k = p\lambda^{k+2} + r\lambda^k + 2p\lambda^{k-1}, \quad 1 \leq k \leq S-1,$$

*i.e.*

$$\lambda = p\lambda^3 + r\lambda + 2p, \quad 1 \leq k \leq S-1,$$

hence, using the relation  $3p + r = 1$  we find

$$\lambda^3 - 3\lambda + 2 = (\lambda^2 - 2\lambda + 1)(\lambda + 2) = (\lambda - 1)^2(\lambda + 2) = 0, \quad k = 1, 2, \dots, S-1,$$

with solutions  $\lambda \in \{1, -2\}$ . Hence we can now try a solution of the form

$$g(k) = C_1 + C_2k + C_3(-2)^k$$

with the boundary conditions

$$\begin{cases} g(0) = 0 = C_1 + C_3, \\ g(S) = 1 = C_1 + C_2S + (-2)^S C_3, \\ g(S+1) = 1 = C_1 + C_2(S+1) - 2 \times (-2)^S C_3, \end{cases}$$

*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

$$\begin{aligned} g(k) &= \frac{1}{1 - (3S+1)(-2)^S} - \frac{3k(-2)^S}{1 - (3S+1)(-2)^S} - \frac{(-2)^k}{1 - (3S+1)(-2)^S} \\ &= \frac{1 - 3k(-2)^S - (-2)^k}{1 - (3S+1)(-2)^S}, \quad k = 0, 1, \dots, S+1. \end{aligned} \quad (\text{S.22})$$

In the graph of Figure [S.1](#) the ruin probability ([S.22](#)) is plotted as a function of  $k$  for  $p = 1/3$  and  $r = 0$ .

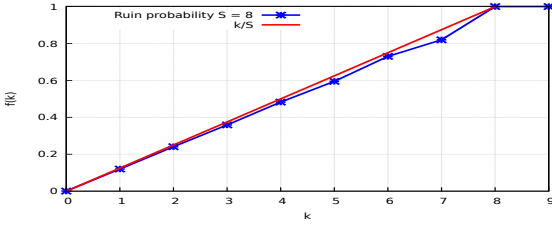


Fig. S.1: Ruin probability  $f_8(k)$  as a function of  $X_0 = k \in [0, 9]$ .

- 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  $n \geq 1$ , 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 (S.22).
- e) We have  $h(0) = h(S) = h(S+1) = 0$ .
- f) We find

$$h(k) = 1 + ph(k+2) + rh(k) + 2ph(k-1), \quad 1 \leq k \leq S-1.$$

- 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

$$Ck^2 = 1 + pC(k+2)^2 + rCk^2 + 2pC(k-1)^2, \quad 1 \leq k \leq S-1,$$

i.e.

$$0 = 1 + pC(2k+4) + 2pC(-2k+1), \quad 1 \leq k \leq S-1,$$

or  $C = -1/(6p)$ .

- h) The general solution takes the form

$$h(k) = -\frac{k^2}{6p} + C_1 + C_2k + C_3(-2)^k$$

with the boundary conditions

$$\begin{cases} h(0) = 0 = C_1 + C_3, \\ h(S) = 0 = -\frac{S^2}{6p} + C_1 + C_2S + (-2)^S C_3, \\ h(S+1) = 0 = -\frac{(S+1)^2}{6p} + C_1 + C_2(S+1) - 2 \times (-2)^S C_3, \end{cases}$$

---

\* 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.

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

$$\begin{aligned} h(k) &= -\frac{k^2}{6p} - \frac{S(S+1)}{6p(1-(3S+1)(-2)^S)} + k \left( \frac{S}{6p} + \frac{(S+1)(1-(-2)^S)}{6p(1-(3S+1)(-2)^S)} \right) \\ &\quad + \frac{S(S+1)(-2)^k}{6p(1-(3S+1)(-2)^S)} \\ &= \frac{k(S-k)}{6p} + \frac{-S(S+1) + k(S+1)(1-(-2)^S) + S(S+1)(-2)^k}{6p(1-(3S+1)(-2)^S)}, \\ &= \frac{k(S-k)}{6p} + \frac{k(S+1)(1-(-2)^S) - S(S+1)(1-(-2)^k)}{6p(1-(3S+1)(-2)^S)} \\ &= \frac{k(S-k)}{6p} + \frac{k(S+1) - S(S+1)}{6p(1-(3S+1)(-2)^S)} + \frac{-k(S+1)(-2)^S + S(S+1)(-2)^k}{6p(1-(3S+1)(-2)^S)} \\ &= \frac{k(S-k)}{6p} + (S+1) \frac{S(-2)^k - k(-2)^S - (S-k)}{6p(1-(3S+1)(-2)^S)}, \end{aligned} \quad (\text{S.23})$$

$k = 0, 1, \dots, S+1$ . In the graph of Figure S.2 the mean game duration (S.23) is plotted as a function of  $k$  with  $p = 1/3$ .

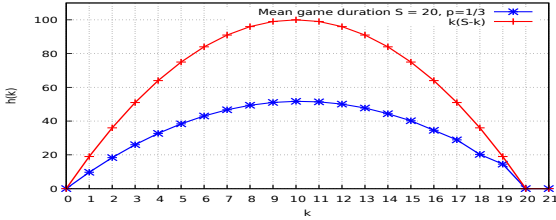


Fig. S.2: Mean game duration  $h_{20}(k)$  as a function of  $X_0 = k \in [0, 20]$ .

- i) Starting from a state  $k \in \{1, 2, \dots, S-1\}$ , 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 ruin probability starting from  $k \geq 1$  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  $\tau$  by one unit in the first step analysis argument, which yields

$$\begin{aligned} h(k) &= p \mathbb{E}[(1 + \tau)^2 \mid X_0 = k + 1] + q \mathbb{E}[(1 + \tau)^2 \mid X_0 = k - 1] \\ &= p \mathbb{E}[1 + 2\tau + \tau^2 \mid X_0 = k + 1] + q \mathbb{E}[1 + 2\tau + \tau^2 \mid X_0 = k - 1] \\ &= 1 + p \mathbb{E}[2\tau + \tau^2 \mid X_0 = k + 1] + q \mathbb{E}[2\tau + \tau^2 \mid X_0 = k - 1] \\ &= 1 + 2p \mathbb{E}[\tau \mid X_0 = k + 1] + 2q \mathbb{E}[\tau \mid X_0 = k - 1] \\ &\quad + p \mathbb{E}[\tau^2 \mid X_0 = k + 1] + q \mathbb{E}[\tau^2 \mid X_0 = k - 1] \\ &= 1 + 2p \mathbb{E}[\tau \mid X_0 = k + 1] + 2q \mathbb{E}[\tau \mid X_0 = k - 1] + ph(k + 1) + qh(k - 1), \end{aligned}$$

$1 \leq k \leq S - 1$ , where we used the relation  $p + q = 1$ .

- c) We have

$$\begin{aligned} h(k) &= 1 + \mathbb{E}[\tau \mid X_0 = k + 1] + \mathbb{E}[\tau \mid X_0 = k - 1] + \frac{1}{2}h(k + 1) + \frac{1}{2}h(k - 1) \\ &= 1 + (S - k - 1)(k + 1) + (S - k + 1)(k - 1) + \frac{1}{2}h(k + 1) + \frac{1}{2}h(k - 1) \\ &= 1 + (S - k - 1)k + (S - k - 1) + (S - k + 1)k - (S - k + 1) \\ &\quad + \frac{1}{2}h(k + 1) + \frac{1}{2}h(k - 1) \\ &= -1 + 2(S - k)k + \frac{1}{2}h(k + 1) + \frac{1}{2}h(k - 1), \end{aligned}$$

$1 \leq k \leq S - 1$ .

- d) We have

$$h(k) = C_1 + C_2k + \frac{2k^2}{3} - \frac{k^3}{3}(2S - k)$$

with

$$0 = h(0) = C_1$$

and

$$0 = h(S) = C_2S + \frac{2S^2}{3} - \frac{S^4}{3}$$

hence

$$C_2 = -\frac{2S}{3} + \frac{S^3}{3} = \frac{S}{3}(S^2 - 2)$$

and

$$h(k) = k \frac{S}{3}(S^2 - 2) + \frac{2k^2}{3} - \frac{k^3}{3}(2S - k).$$

e) We have

$$\begin{aligned}
v(k) &= h(k) - k^2(S - k)^2 \\
&= k \frac{S}{3} (S^2 - 2) + \frac{2k^2}{3} - \frac{k^3}{3} (2S - k) - k^2(S - k)^2 \\
&= k \left( \frac{S}{3} (S^2 - 2) + \frac{2k}{3} - S \frac{k^2}{3} - \frac{k^2}{3} (S - k) - k(S - k)^2 \right) \\
&= \frac{k}{3} (S(S^2 - 2) + 2k - Sk^2 - k^2(S - k) - 3k(S - k)^2) \\
&= \frac{k}{3} ((S - k)(S^2 - 2) + Sk(S - k) - k^2(S - k) - 3k(S - k)^2) \\
&= \frac{k}{3} ((S - k)(S^2 - 2 + Sk) - k^2(S - k) - 3k(S - k)^2) \\
&= \frac{k(S - k)}{3} (k^2 + (S - k)^2 - 2).
\end{aligned}$$

**Remark.** When  $k = S/2$  we find

$$v(k) = \frac{S^2}{12} (S^2/2 - 2)$$

and the standard deviation

$$\sigma(k) \simeq \frac{S^2}{\sqrt{24}} \simeq \frac{S^2}{5}.$$

if  $S$  is large.

In the graph of Figure S.3 the standard deviation of the game duration  $\tau$  is plotted as a function of the initial state  $k = 0, 1, \dots, S$ .

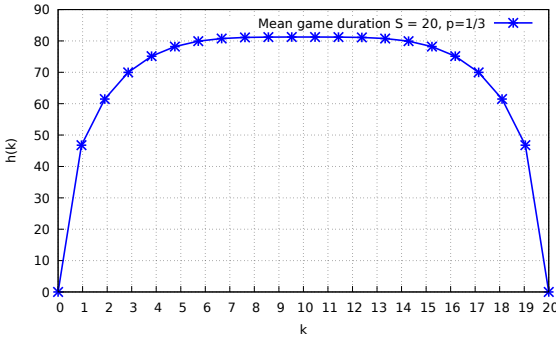


Fig. S.3: Standard deviation of  $\tau$  as a function of  $X_0 = k \in [0, 20]$ .

- 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  $\binom{4}{3} = \binom{4}{1} = 4$  paths, as follows.

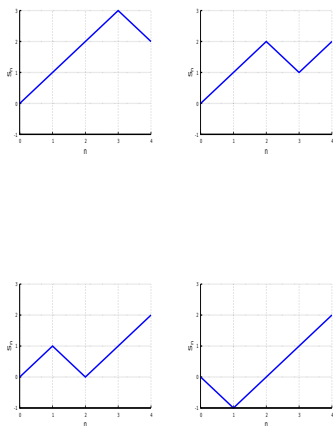


Fig. S.4: Four paths leading from 0 to 2 in four time steps.

- 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  $n$  and  $k$ .

- i) In case  $n$  and  $k$  are even, or written as  $n = 2n'$  and  $k = 2k'$ , Relation (3.3.3) shows that

$$\mathbb{P}(S_n = k \mid S_0 = 0) = \mathbb{P}(S_{2n'} = 2k' \mid S_0 = 0)$$

$$\begin{aligned}
&= \binom{2n'}{n' + k'} p^{n' + k'} q^{n' - k'} \\
&= \binom{n}{(n + k)/2} p^{(n+k)/2} q^{(n-k)/2}, \quad -n \leq k \leq n.
\end{aligned}$$

ii) In case  $n$  and  $k$  are odd, or written as  $n = 2n' + 1$  and  $k = 2k' + 1$ , (3.3.4) shows that

$$\begin{aligned}
\mathbb{P}(S_n = k \mid S_0 = 0) &= \mathbb{P}(S_{2n'+1} = 2k' + 1 \mid S_0 = 0) \\
&= \binom{2n' + 1}{n' + k' + 1} p^{n' + k' + 1} q^{n' - k'} \\
&= \binom{n}{(n + k)/2} p^{(n+k)/2} q^{(n-k)/2}, \quad -n \leq k \leq n.
\end{aligned}$$

d) By a first step analysis started at state  $\textcircled{0}$  we have, letting  $p_{n,k} := \mathbb{P}(S_n = k)$ ,

$$\begin{aligned}
p_{n+1,k} &= \mathbb{P}(S_{n+1} = k) \\
&= \mathbb{P}(S_{n+1} = k \mid S_0 = 0) \\
&= \mathbb{P}(S_{n+1} = k \mid S_1 = 1) \mathbb{P}(S_1 = 1 \mid S_0 = 0) \\
&\quad + \mathbb{P}(S_{n+1} = k \mid S_1 = -1) \mathbb{P}(S_1 = -1 \mid S_0 = 0) \\
&= p \mathbb{P}(S_{n+1} = k \mid S_1 = 1) + q \mathbb{P}(S_{n+1} = k \mid S_1 = -1) \\
&= p \mathbb{P}(S_{n+1} = k - 1 \mid S_1 = 0) + q \mathbb{P}(S_{n+1} = k + 1 \mid S_1 = 0) \\
&= p \mathbb{P}(S_n = k - 1 \mid S_0 = 0) + q \mathbb{P}(S_n = k + 1 \mid S_0 = 0) \\
&= p p_{n,k-1} + q p_{n,k+1},
\end{aligned}$$

which yields

$$p_{n+1,k} = p p_{n,k-1} + q p_{n,k+1},$$

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

$$\begin{aligned}
p p_{n,k-1} + q p_{n,k+1} &= p \binom{n}{(n - 1 + k)/2} p^{(n-1+k)/2} (1 - p)^{(n+1-k)/2} \\
&\quad + q \binom{n}{(n + 1 + k)/2} p^{(n+1+k)/2} (1 - p)^{(n-1-k)/2}
\end{aligned}$$

$$\begin{aligned}
&= \binom{n}{(n-1+k)/2} p^{(n+1+k)/2} q^{(n+1-k)/2} \\
&\quad + \binom{n}{(n+1+k)/2} p^{(n+1+k)/2} q^{(n+1-k)/2} \\
&= p^{(n+1+k)/2} q^{(n+1-k)/2} \left( \binom{n}{(n-1+k)/2} + \binom{n}{(n+1+k)/2} \right) \\
&= p^{(n+1+k)/2} q^{(n+1-k)/2} \left( \frac{n!}{((n+1-k)/2)!((n-1+k)/2)!} \right. \\
&\quad \left. + \frac{n!}{((n-1-k)/2)!((n+1+k)/2)!} \right) \\
&= p^{(n+1+k)/2} q^{(n+1-k)/2} \left( \frac{n!(n+1+k)/2}{((n+1-k)/2)!((n+1+k)/2)!} \right. \\
&\quad \left. + \frac{n!(n+1-k)/2}{((n-1-k)/2)!((n+1+k)/2)!} \right) \\
&= p^{(n+1+k)/2} q^{(n+1-k)/2} \frac{n!(n+1)}{((n+1-k)/2)!((n+1+k)/2)!} \\
&= p^{(n+1+k)/2} q^{(n+1-k)/2} \binom{n+1}{(n+1+k)/2} \\
&= p_{n+1,k},
\end{aligned}$$

which shows that  $p_{n,k}$  satisfies Equation (3.4.25). In addition we clearly have

$$p_{0,0} = \mathbb{P}(S_0 = 0) = 1 \quad \text{and} \quad p_{0,k} = \mathbb{P}(S_0 = k) = 0, \quad k \neq 0.$$

### Exercise 3.2

a) By the spatial 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  $\textcircled{i}$  to  $\boxed{i-1}$  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

$$\begin{aligned}
\mathbb{E}[T_0 \mid S_0 = 1] &= q \times 1 + p \times (1 + \mathbb{E}[T_0 \mid S_0 = 2]) \\
&= q \times 1 + p \times (1 + 2 \mathbb{E}[T_0 \mid S_0 = 1]) \\
&= 1 + 2p \mathbb{E}[T_0 \mid S_0 = 1],
\end{aligned}$$

hence

$$\mathbb{E}[T_0 \mid S_0 = 1] = \frac{1}{1-2p} = \frac{1}{q-p}, \quad q > p.$$

When  $p \geq q$  we find  $\mathbb{E}[T_0 \mid S_0 = 1] = +\infty$ .

b) By the spatial 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  $\boxed{i-1}$  starting from  $\textcircled{i}$  we find

$$\mathbb{P}(T_0 < \infty \mid S_0 = k) = \prod_{i=1}^k \mathbb{P}(T_{i-1} < \infty \mid S_0 = i) = (\mathbb{P}(T_0 < \infty \mid S_0 = 1))^k.$$

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

$$\mathbb{P}(T_0 < \infty \mid S_0 = 1) = \frac{q}{p},$$

when  $p \leq q$  we have  $\mathbb{P}(T_0 < \infty \mid S_0 = 1) = 1$ . We conclude that in general,

$$\mathbb{P}(T_0 < \infty \mid S_0 = k) = \min \left( 1, \left( \frac{q}{p} \right)^k \right), \quad k \in \mathbb{N},$$

which coincides with (2.2.14) 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

$$\mathbb{P}(T_0^r < \infty) = G_{T_0^r}(1) = 1 - \sqrt{1 - 4pq} = 2 \min(p, q),$$

by (3.4.14).

b) In general we have  $\mathbb{P}(T_0^r = n) = \frac{1}{n!} G_{T_0^r}^{(n)}(0)$ ,  $n \geq 0$ , with

$$\begin{cases} G'_{T_0^r}(s) = 4pqs(1 - 4pqs^2)^{-1/2}, \\ G''_{T_0^r}(s) = 4pq(1 - 4pqs^2)^{-3/2}, \\ G'''_{T_0^r}(s) = 48p^2q^2s(1 - 4pqs^2)^{-5/2}, \\ G^{(4)}_{T_0^r}(s) = 48p^2q^2(1 + 16pqs^2)(1 - 4pqs^2)^{-7/2}, \end{cases}$$

hence

$$\begin{cases} \mathbb{P}(T_0^r = 1) = G'_{T_0^r}(0) = 0, \\ \mathbb{P}(T_0^r = 2) = G''_{T_0^r}(0) = 2pq, \\ \mathbb{P}(T_0^r = 3) = \frac{1}{3!} G'''_{T_0^r}(0) = 0, \\ \mathbb{P}(T_0^r = 4) = \frac{1}{4!} G^{(4)}_{T_0^r}(0) = 2p^2q^2. \end{cases}$$

c) We have

$$\mathbb{E}[T_0^r \mathbb{1}_{\{T_0^r < \infty\}}] = G'_{T_0^r}(1) = \frac{4pq}{\sqrt{1 - 4pq}} = \frac{4pq}{|p - q|},$$

hence by (1.6.6) in Lemma 1.10 we find

$$\mathbb{E}[T_0^r \mid T_0^r < \infty] = \frac{1}{\mathbb{P}(T_0^r < \infty)} \mathbb{E}[T_0^r \mathbb{1}_{\{T_0^r < \infty\}}] = 2 \frac{\max(p, q)}{|p - q|},$$

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

$$\begin{aligned} G_i(s) &= \mathbb{E}[s^{T_0} \mid X_0 = i] \\ &= p \mathbb{E}[s^{1+T_0} \mid X_0 = i + 1] + q \mathbb{E}[s^{1+T_0} \mid X_0 = i - 1] \\ &= ps \mathbb{E}[s^{T_0} \mid X_0 = i + 1] + qs \mathbb{E}[s^{T_0} \mid X_0 = i - 1], \quad i \geq 1, \end{aligned}$$

hence we have

$$G_i(s) = psG_{i+1}(s) + qsG_{i-1}(s), \quad i \in \mathbb{Z} \setminus \{0\}, \quad s \in [0, 1). \quad (\text{S.24})$$

b) We note that (S.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 (S.24) we find

$$(\alpha(s))^i = ps(\alpha(s))^{i+1} + qs(\alpha(s))^{i-1},$$

which yields the characteristic equation

$$ps(\alpha(s))^2 - \alpha(s) + qs = 0, \quad s \in [0, 1),$$

with unknown  $\alpha(s)$  and solutions

$$\begin{aligned} \alpha_+(s) &= \frac{1 + \sqrt{1 - 4pqs^2}}{2ps} \\ &= \frac{2qs}{1 - \sqrt{1 - 4pqs^2}} > \frac{1}{2ps}, \quad |s| < 1/(2\sqrt{pq}), \end{aligned}$$

and

$$\begin{aligned} \alpha_-(s) &= \frac{1 - \sqrt{1 - 4pqs^2}}{2ps} \\ &= \frac{2qs}{1 + \sqrt{1 - 4pqs^2}} < 2qs, \quad |s| < 1/(2\sqrt{pq}), \end{aligned}$$

hence the general solution of (S.24) is given by

$$\begin{aligned} G_i(s) &= C_+(\alpha_+(s))^i + C_-(\alpha_-(s))^i \\ &= C_+ \left( \frac{1 + \sqrt{1 - 4pqs^2}}{2ps} \right)^i + C_- \left( \frac{1 - \sqrt{1 - 4pqs^2}}{2ps} \right)^i, \quad i \in \mathbb{Z}. \end{aligned}$$

Hence the boundary conditions  $G_{\pm\infty}(s) = 0$ , or the boundedness condition  $|G_i(s)| \leq 1$  for sufficiently small  $s \in (-1, 1)$ , show that  $C_+ = 0$  and  $C_- = 1$  when  $i \geq 0$ , *i.e.*

$$G_i(s) = \left( \frac{1 - \sqrt{1 - 4pqs^2}}{2ps} \right)^i = \left( \frac{2qs}{1 + \sqrt{1 - 4pqs^2}} \right)^i, \quad i \geq 0, \quad (\text{S.25})$$

and  $C_+ = 1$  and  $C_- = 0$  when  $i \leq 0$ , *i.e.*

$$G_i(s) = \left( \frac{1 + \sqrt{1 - 4pqs^2}}{2ps} \right)^i = \left( \frac{2qs}{1 - \sqrt{1 - 4pqs^2}} \right)^i, \quad i \leq 0,$$

while by symmetry we also have

$$G_i(s) = \left( \frac{2qs}{1 + \sqrt{1 - 4pqs^2}} \right)^i = \left( \frac{1 - \sqrt{1 - 4pqs^2}}{2ps} \right)^i, \quad i \geq 0, \quad (\text{S.26})$$

and

$$G_i(s) = \left( \frac{2qs}{1 - \sqrt{1 - 4pqs^2}} \right)^i = \left( \frac{1 + \sqrt{1 - 4pqs^2}}{2ps} \right)^i, \quad i \leq 0. \quad (\text{S.27})$$

c) Relations (S.25) and (S.26) show from (3.4.14) that

$$\begin{aligned} \mathbb{P}(T_0 < \infty \mid X_0 = i) &= G_i(1) \\ &= \left( \frac{2q}{1 + \sqrt{1 - 4pq}} \right)^i \\ &= \left( \frac{2q}{1 + |p - q|} \right)^i \\ &= \min \left( 1, \left( \frac{q}{p} \right)^i \right), \quad i \geq 0, \end{aligned}$$

which recovers (2.2.14) and (3.4.16).

d) When  $p \geq q$  and  $i \leq 0$ , we have

$$G'_i(s) = \frac{is^{i-1}}{(2q)^{-i}} (1 - \sqrt{1 - 4pqs^2})^{-i} - \frac{4pqi}{(2q)^{-i}} \frac{(1 - \sqrt{1 - 4pqs^2})^{-i-1}}{\sqrt{1 - 4pqs^2}},$$

and when  $s = 1$  this yields

$$\begin{aligned} G'_i(1) &= \frac{i}{(2q)^{-i}} (1 - \sqrt{1 - 4pq})^{-i} - \frac{4pqi}{(2q)^{-i}} \frac{(1 - \sqrt{1 - 4pq})^{-i-1}}{\sqrt{1 - 4pq}} \\ &= \frac{i}{(2q)^{-i}} (1 - (p - q))^{-i} - \frac{4pqi}{(2q)^{-i}} \frac{(1 - (p - q))^{-i-1}}{p - q} \\ &= i - \frac{2pi}{p - q} \\ &= -\frac{i}{p - q}, \quad i \leq 0. \end{aligned}$$

e) By first step analysis starting from state  $\textcircled{0}$  we have, when  $q \geq p$ ,

$$\begin{aligned} \mathbb{E}[s^{T_0} \mid X_0 = 0] &= ps \mathbb{E}[s^{T_0} \mid X_0 = 1] + qs \mathbb{E}[s^{T_0} \mid X_0 = -1] \\ &= psG_1(s) + qsG_{-1}(s) \\ &= \frac{1 - \sqrt{1 - 4pqs^2}}{2} + \frac{2pqs^2}{1 + \sqrt{1 - 4pqs^2}} \\ &= \frac{(1 - \sqrt{1 - 4pqs^2})(1 + \sqrt{1 - 4pqs^2}) + 4pqs^2}{2(1 + \sqrt{1 - 4pqs^2})} \\ &= \frac{4pqs^2}{1 + \sqrt{1 - 4pqs^2}} \end{aligned}$$

$$= 1 - \sqrt{1 - 4pqs^2}, \quad s \in [0, 1),$$

which recovers (3.4.9), while when  $q \leq p$  we find similarly

$$\begin{aligned} \mathbb{E} [s^{T_0^r} \mid X_0 = 0] &= psG_1(s) + qsG_{-1}(s) \\ &= \frac{2qps^2}{1 + \sqrt{1 - 4pqs^2}} + \frac{1 - \sqrt{1 - 4pqs^2}}{2} \\ &= 1 - \sqrt{1 - 4pqs^2}, \quad s \in [0, 1). \end{aligned}$$

Exercise 3.5 By (3.4.21) we have

$$\begin{aligned} \mathbb{E} [T_0^r \mathbf{1}_{\{T_0^r < \infty\}} \mid S_0 = 0] &= \sum_{n \geq 1} n \mathbb{P}(T_0^r = n \mid S_0 = 0) \\ &= \sum_{k \geq 1} 2k \mathbb{P}(T_0^r = 2k \mid S_0 = 0) \\ &= \sum_{k \geq 1} \frac{2k}{2k-1} \binom{2k}{k} (1/4)^k \quad (\text{S.28}) \\ &= \infty, \end{aligned}$$

where the general term of the above series satisfies the equivalence

$$\frac{2k}{2k-1} \binom{2k}{k} (1/4)^k \simeq_{k \rightarrow \infty} \frac{(2k)!}{4^k (k!)^2} \simeq_{k \rightarrow \infty} \sqrt{\frac{1}{2\pi k}}$$

by [Stirling's approximation](#)  $k! \simeq (k/e)^k \sqrt{2\pi k}$  as  $k$  tends to  $\infty$ , from which we conclude to the divergence of the series (S.28).

Exercise 3.6

a) We have

$$\begin{aligned} \mathbb{E} [M_n] &= \mathbb{E} \left[ \sum_{k=1}^n 2^{k-1} X_k \right] \\ &= \sum_{k=1}^n 2^{k-1} \mathbb{E} [X_k] \\ &= (p-q) \sum_{k=0}^{n-1} 2^k \\ &= (p-q) \frac{1-2^n}{1-2} \\ &= (p-q)(2^n - 1), \quad n \geq 0. \end{aligned}$$

- b) The two possible values of  $M_{\mathcal{I} \wedge n}$  are 1 and

$$-\sum_{k=1}^n 2^{k-1} = -\frac{1-2^n}{1-2} = 1-2^n, \quad n \geq 1.$$

We have

$$\mathbb{P}(M_{n \wedge \tau} = 1 - 2^n) = q^n \quad \text{and} \quad \mathbb{P}(M_{n \wedge \tau} = 1) = 1 - q^n, \quad n \geq 1.$$

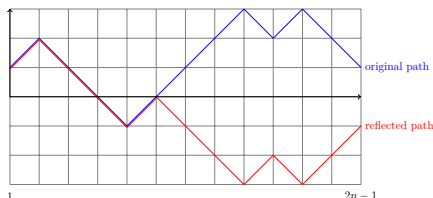
- c) The stopped process  $(M_{\min(n, \tau)})_{n \in \mathbb{N}}$  represents the wealth of a gambler whose strategy of the gambler is to double the stakes each time he loses, and to quit the game as soon as his gains reach \$1.
- d) We have

$$\begin{aligned}\mathbb{E}[M_{n\wedge\tau}] &= (1-2^n)\mathbb{P}(M_{n\wedge\tau} = 1-2^n) + \mathbb{P}(M_{n\wedge\tau} = 1) \\ &= (1-2^n)q^n + 1 - q^n \\ &= 1 - (2q)^n, \quad n \geq 1.\end{aligned}$$

In particular, we find  $\mathbb{E}[M_{n \wedge \tau}] = 0$  in the fair game with  $p = q = 1/2$ .

### Exercise 3.7

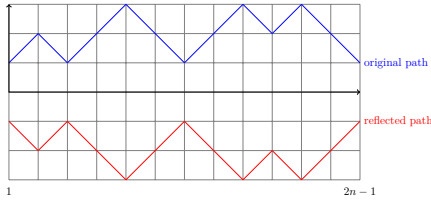
- Applying (3.4.28) with  $m = n - 1$  and  $k = 0$ , we find that this number is  $\binom{2n-2}{n-1}$ .
- Applying (3.4.28) with  $m = n - 1$  and  $k = -1$ , we find that this number is  $\binom{2n-2}{n-2} = \binom{2n-2}{n}$ .
- On the graph below, the blue path joining  $S_1 = 1$  to  $S_{2n-1} = 1$  by crossing  $\textcircled{0}$  is associated to a unique red path joining  $S_1 = 1$  to  $S_{2n-1} = -1$ .



- d) This number is  $\binom{2n-2}{n-2} = \binom{2n-2}{n}$ .
- e) This number is

$$\begin{aligned}
\binom{2n-2}{n-1} - \binom{2n-2}{n-2} &= \frac{(2n-2)!}{(n-1)!(n-1)!} - \frac{(2n-2)!}{(n-2)!n!} \\
&= \frac{(n^2 - n(n-1))(2n-2)!}{n!n!} \\
&= \frac{(2n-2)!}{(n-1)!n!}.
\end{aligned}$$

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  $\textcircled{0}$  between time 1 and time  $2n-1$  we can associate a red path joining  $S_1 = -1$  to  $S_{2n-2} = -1$  *without* crossing  $\textcircled{0}$ .



Adding the number of paths joining  $S_1 = 1$  to  $S_{2n-2} = 1$  *without* crossing  $\textcircled{0}$  from time 1 to time  $2n-1$  to the number of paths joining  $S_1 = -1$  to  $S_{2n-2} = -1$  *without* crossing  $\textcircled{0}$  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  $\textcircled{0}$  from time 0 to time  $2n$ , as follows:

$$2 \frac{(2n-2)!}{(n-1)!n!} = \frac{2n(2n-2)!}{n!n!} = \frac{1}{2n-1} \binom{2n}{n}.$$

We note that the convolution equation (3.4.3) is also satisfied in the following form:

$$\begin{aligned}
\frac{1}{2n-1} \binom{2n}{n} &= \sum_{k=0}^{n-1} \frac{1}{2n-2k-1} \binom{2n-2k}{n-k} \binom{2k}{k} \\
&= \sum_{k=0}^{n-1} \frac{1}{2k-1} \binom{2k}{k} \binom{2n-2k}{n-k}, \quad n \geq 1,
\end{aligned}$$

see *e.g.* [here](#). We note that the number of paths joining  $S_0 = 0$  to  $S_{2n+2} = 0$  *without* crossing  $\textcircled{0}$ , from time 0 to time  $2n+2$  satisfies

$$\frac{1}{2n+1} \binom{2n+2}{n+1} = \frac{2}{n+1} \binom{2n}{n},$$

where  $\frac{1}{n+1} \binom{2n}{n}$  is the *Catalan number* counting the number of Dyck lattice paths joining  $S_0 = 0$  to  $S_n = n$  without passing above the diagonal between time 0 and time  $n$ .

**Problem 3.8**

- a) Since the increment  $X_k$  takes its values in  $\{-1, 1\}$ , the set of distinct values in  $\{S_0, S_1, \dots, S_n\}$  is the integer interval

$$\left[ \inf_{k=0,1,\dots,n} S_k, \sup_{k=0,1,\dots,n} S_k \right],$$

which has

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

elements. In addition we have  $R_0 = 1$  and  $R_1 = 2$ .

- b) At each time step  $k \geq 1$  the range can only either increase by one unit or remain constant, hence  $R_k - R_{k-1} \in \{0, 1\}$  is a Bernoulli random variable. In addition we have the identity

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

hence, applying the probability  $\mathbb{P}$  to both sides, we get

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

- c) By the change of index

$$(X_1, X_2, \dots, X_{k-1}, X_k) \mapsto (X_k, X_{k-1}, \dots, X_2, X_1)$$

under which  $X_1 + X_2 + \dots + X_l$  becomes  $X_k + \dots + X_{k-l+1}$ ,  $l = 1, 2, \dots, k$ , we have

$$\begin{aligned} \mathbb{P}(R_k - R_{k-1} = 1) &= \mathbb{P}(S_k - S_0 \neq 0, S_k - S_1 \neq 0, \dots, S_k - S_{k-1} \neq 0) \\ &= \mathbb{P}(X_1 + \dots + X_k \neq 0, X_2 + \dots + X_k \neq 0, \dots, X_k \neq 0) \\ &= \mathbb{P}(X_1 \neq 0, X_1 + X_2 \neq 0, \dots, X_1 + \dots + X_k \neq 0), \end{aligned}$$

for all  $k \geq 1$ , since the sequence  $(X_k)_{k \geq 1}$  is made of independent and identically distributed random variables.

- d) We have the telescoping sum

$$\begin{aligned} R_n &= R_0 + (R_1 - R_0) + \dots + (R_n - R_{n-1}) \\ &= R_0 + \sum_{k=1}^n (R_k - R_{k-1}), \quad n \geq 0. \end{aligned}$$

e) By (1.2.6) we have

$$\mathbb{P}(T_0 = \infty) = \mathbb{P}\left(\bigcap_{k \geq 1} \{T_0 > k\}\right) = \lim_{k \rightarrow \infty} \mathbb{P}(T_0 > k),$$

since

$$\{T_0 > k+1\} \implies \{T_0 > k\}, \quad k \geq 1,$$

i.e.  $(\{T_0 > k\})_{k \geq 1}$  is a decreasing sequence of events.

f) Noting that  $R_k - R_{k-1} \in \{0, 1\}$  is a Bernoulli random variable with

$$\mathbb{E}[R_k - R_{k-1}] = \mathbb{P}(R_k - R_{k-1} = 1),$$

we find that

$$\begin{aligned} \mathbb{E}[R_n] &= \mathbb{E}\left[R_0 + \sum_{k=1}^n (R_k - R_{k-1})\right] = R_0 + \sum_{k=1}^n \mathbb{E}[R_k - R_{k-1}] \\ &= R_0 + \sum_{k=1}^n \mathbb{P}(R_k - R_{k-1} = 1) \\ &= R_0 + \sum_{k=1}^n \mathbb{P}(X_1 \neq 0, X_1 + X_2 \neq 0, \dots, X_1 + \dots + X_k \neq 0) \\ &= R_0 + \sum_{k=1}^n \mathbb{P}(S_1 \neq 0, S_2 \neq 0, \dots, S_k \neq 0) = R_0 + \sum_{k=1}^n \mathbb{P}(T_0 > k) \\ &= 1 + \sum_{k=1}^n \mathbb{P}(T_0 > k) = \mathbb{P}(T_0 > 0) + \sum_{k=1}^n \mathbb{P}(T_0 > k) = \sum_{k=0}^n \mathbb{P}(T_0 > k). \end{aligned}$$

g) Let  $\varepsilon > 0$ . Since by Question (e) we have  $\mathbb{P}(T_0 = \infty) = \lim_{k \rightarrow \infty} \mathbb{P}(T_0 > k)$ , there exists  $N \geq 1$  such that

$$|\mathbb{P}(T_0 = \infty) - \mathbb{P}(T_0 > k)| < \varepsilon, \quad k \geq N.$$

Hence for  $n \geq N$  we have

$$\begin{aligned} \left| \mathbb{P}(T_0 = \infty) - \frac{1}{n} \sum_{k=1}^n \mathbb{P}(T_0 > k) \right| &= \left| \frac{1}{n} \sum_{k=1}^n (\mathbb{P}(T_0 = \infty) - \mathbb{P}(T_0 > k)) \right| \\ &\leq \frac{1}{n} \sum_{k=1}^n |\mathbb{P}(T_0 = \infty) - \mathbb{P}(T_0 > k)| \\ &\leq \frac{1}{n} \sum_{k=1}^N |\mathbb{P}(T_0 = \infty) - \mathbb{P}(T_0 > k)| + \frac{1}{n} \sum_{k=N+1}^n |\mathbb{P}(T_0 = \infty) - \mathbb{P}(T_0 > k)| \end{aligned}$$

$$\begin{aligned} &\leq \frac{N}{n} + \frac{n-N}{n}\varepsilon \\ &\leq \frac{N}{n} + \varepsilon. \end{aligned}$$

Then, choosing  $N_0 \geq 1$  such that  $(N+1)/n \leq \varepsilon$  for  $n \geq N_0$ , we get

$$\left| \mathbb{P}(T_0 = \infty) - \frac{1}{n} \mathbb{E}[R_n] \right| \leq \frac{1}{n} + \left| \mathbb{P}(T_0 = \infty) - \frac{1}{n} \sum_{k=1}^n \mathbb{P}(T_0 > k) \right| \leq 2\varepsilon,$$

$n \geq N_0$ , which concludes the proof.

Alternatively, the answer to that question can be derived by applying the Cesàro theorem, which states that in general we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n a_k = a$$

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  $\lim_{k \rightarrow \infty} \mathbb{P}(T_0 > k) = \mathbb{P}(T_0 = \infty)$ .

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

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[R_n] = |p - q|,$$

when  $p \neq q$ , and

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[R_n] = 0,$$

when  $p = q = 1/2$ .

**Problem 3.9** The question of recurrence of the  $d$ -dimensional random walk has been first solved in [Pólya \(1921\)](#), cf. [Champion et al. \(2007\)](#) for the solution proposed in this problem.

a) We partition the event  $\{S_n = \vec{0}\}$  into

$$\{S_n = \vec{0}\} = \bigcup_{k=2}^n \{S_{n-k} = \vec{0}, S_{n-k+1} \neq \vec{0}, \dots, S_{n-1} \neq \vec{0}, S_n = \vec{0}\}, \quad n \geq 1,$$

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  $S_0 = \vec{0}$ . Then we have

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

$$\begin{aligned}
&= \sum_{k=2}^n \mathbb{P}(S_{n-k} = \vec{0}, S_{n-k+1} \neq \vec{0}, \dots, S_{n-1} \neq \vec{0}, S_n = \vec{0} \mid S_0 = \vec{0}) \\
&= \sum_{k=2}^n \mathbb{P}(S_{n-k+1} \neq \vec{0}, \dots, S_{n-1} \neq \vec{0}, S_n = \vec{0} \mid S_{n-k} = \vec{0}, S_0 = \vec{0}) \\
&\quad \times \mathbb{P}(S_{n-k} = \vec{0} \mid S_0 = \vec{0}) \\
&= \sum_{k=2}^n \mathbb{P}(S_1 \neq \vec{0}, \dots, S_{k-1} \neq \vec{0}, S_k = \vec{0} \mid S_0 = \vec{0}) \mathbb{P}(S_{n-k} = \vec{0} \mid S_0 = \vec{0}) \\
&= \sum_{k=2}^n \mathbb{P}(T_0^r = k \mid S_0 = \vec{0}) \mathbb{P}(S_{n-k} = \vec{0} \mid S_0 = \vec{0}) \\
&= \sum_{k=2}^n \mathbb{P}(S_{n-k} = \vec{0}) \mathbb{P}(T_0^r = k), \quad n \geq 1.
\end{aligned}$$

b) We have

$$\begin{aligned}
\sum_{n=1}^m \mathbb{P}(S_n = \vec{0}) &= \sum_{n=1}^m \sum_{k=2}^n \mathbb{P}(T_0^r = k) \mathbb{P}(S_{n-k} = \vec{0}) \\
&= \sum_{k=2}^m \sum_{n=k}^m \mathbb{P}(T_0^r = k) \mathbb{P}(S_{n-k} = \vec{0}) \\
&= \sum_{k=2}^m \mathbb{P}(T_0^r = k) \sum_{l=0}^{m-k} \mathbb{P}(S_l = \vec{0}) \\
&\leq \sum_{k=2}^m \mathbb{P}(T_0^r = k) \sum_{l=0}^m \mathbb{P}(S_l = \vec{0}) \\
&= \left( \sum_{n=0}^m \mathbb{P}(S_n = \vec{0}) \right) \left( \sum_{n=2}^m \mathbb{P}(T_0^r = n) \right).
\end{aligned}$$

c) By letting  $m$  tend to  $\infty$  in (3.4.29) we obtain

$$\mathbb{P}(T_0^r < \infty) = \sum_{n \geq 2} \mathbb{P}(T_0^r = n) \geq 1,$$

which allows us to conclude. Note that the sum of the series  $\sum_{n \geq 0} \mathbb{P}(S_n = \vec{0})$  actually represents the average number of visits to state  $\textcircled{0}$ .

d) We have

$$\begin{aligned}
\sum_{n=1}^{2m} \mathbb{P}(S_n = \vec{0}) &= \sum_{n=2}^{2m} \mathbb{P}(T_0^r = n) \sum_{l=0}^{2m-n} \mathbb{P}(S_l = \vec{0}) \\
&\geq \sum_{n=2}^m \mathbb{P}(T_0^r = n) \sum_{l=0}^{2m-n} \mathbb{P}(S_l = \vec{0}) \\
&\geq \sum_{n=2}^m \mathbb{P}(T_0^r = n) \sum_{l=0}^m \mathbb{P}(S_l = \vec{0}).
\end{aligned}$$

e) Letting  $m$  tend to  $+\infty$  in (3.4.30) we find

$$\mathbb{P}(T_0^r < \infty) = \sum_{n \geq 2} \mathbb{P}(T_0^r = n) \leq \frac{\sum_{n \geq 2} \mathbb{P}(S_n = \vec{0})}{\sum_{n \geq 0} \mathbb{P}(S_n = \vec{0})} = 1 - \frac{1}{\sum_{n \geq 0} \mathbb{P}(S_n = \vec{0})} < 1.$$

f) When  $d = 1$  we have

$$\mathbb{P}(S_{2n} = 0) = \frac{1}{2^{2n}} \binom{2n}{n} = \frac{(2n)!}{2^{2n}(n!)^2} \simeq_{n \rightarrow \infty} \frac{1}{\sqrt{\pi n}},$$

by Stirling's approximation, hence

$$\sum_{n \geq 0} \mathbb{P}(S_n = 0) = \infty.$$

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  $k = 0, 1, \dots, n$  the number of ways to arrange those four types of moves among  $2n$  time steps is the multinomial coefficient

$$\binom{2n}{k, k, n-k, n-k} = \frac{(2n)!}{k!k!(n-k)!(n-k)!},$$

hence, since every sequence of  $2n$  moves occur with the same probability  $(1/4)^{2n}$ , by summation over  $k = 0, 1, \dots, n$  we find

$$\mathbb{P}(S_{2n} = \vec{0}) = \left(\frac{1}{4}\right)^{2n} \sum_{k=0}^n \frac{(2n)!}{(k!)^2((n-k)!)^2}$$

$$\begin{aligned}
&= \frac{(2n)!}{4^{2n}(n!)^2} \sum_{k=0}^n \binom{n}{k}^2 \\
&= \frac{(2n)!}{4^{2n}(n!)^2} \binom{2n}{n} \\
&= \frac{((2n)!)^2}{4^{2n}(n!)^4} \simeq_{n \rightarrow \infty} \frac{1}{2\pi n},
\end{aligned}$$

where we used [Stirling's approximation](#), and this yields

$$\sum_{n \geq 0} \mathbb{P}(S_n = \vec{0}) = \infty.$$

To conclude, the result of Question (c) shows again that  $\mathbb{P}(T_0^T < \infty) = 1$ .

- 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  $i_2, \dots, i_d$ . The number of ways to arrange such paths is given by the multinomial coefficient

$$\binom{2n}{i_1, i_1, i_2, i_2, \dots, i_d, i_d} = \frac{(2n)!}{(i_1!)^2 \dots (i_d!)^2},$$

and by summation over all possible indices  $i_1, i_2, \dots, i_d \geq 0$  satisfying  $i_1 + \dots + i_d = n$  and multiplying by the probability  $(1/(2d))^{2n}$  of each path we find

$$\begin{aligned}
\mathbb{P}(S_{2n} = \vec{0}) &= \frac{1}{(2d)^{2n}} \sum_{\substack{i_1 + \dots + i_d = n \\ i_1, i_2, \dots, i_d \geq 0}} \binom{2n}{i_1, i_1, i_2, i_2, \dots, i_d, i_d} \\
&= \frac{1}{(2d)^{2n}} \sum_{\substack{i_1 + \dots + i_d = n \\ i_1, i_2, \dots, i_d \geq 0}} \frac{(2n)!}{(i_1!)^2 \dots (i_d!)^2}.
\end{aligned}$$

- i) By the hint provided that we have

$$\begin{aligned}
\sum_{n \geq 1} \mathbb{P}(S_{2n} = \vec{0}) &= \sum_{n \geq 1} \frac{1}{(2d)^{2n}} \binom{2n}{n!} \sum_{\substack{i_1 + \dots + i_d = n \\ i_1, i_2, \dots, i_d \geq 0}} \frac{(n!)^2}{(i_1!)^2 \dots (i_d!)^2} \\
&\leq \sum_{n \geq 1} \frac{1}{(2d)^{2n}} \binom{2n}{n!} \frac{n!}{(a_n!)^d (a_n + 1)^{b_n}} \sum_{\substack{i_1 + \dots + i_d = n \\ i_1, i_2, \dots, i_d \geq 0}} \frac{n!}{i_1! \dots i_d!} \\
&\leq \sum_{n \geq 1} \frac{1}{(2d)^{2n}} \binom{2n}{n!} \frac{n! d^n}{(a_n!)^d a_n^{b_n}} = \sum_{n \geq 1} \frac{(2n)!}{2^{2n} d^n n! (a_n!)^d a_n^{b_n}}.
\end{aligned}$$

j) For sufficiently large  $n$ , by the Stirling approximation we have

$$\begin{aligned}
\frac{(2n)!}{2^{2n} d^n n! (a_n!)^d a_n^{b_n}} &\simeq \frac{(2n/e)^{2n} \sqrt{4\pi n}}{2^{2n} d^n (n/e)^n \sqrt{2\pi n} ((a_n/e)^{a_n} \sqrt{2\pi a_n})^d a_n^{b_n}} \\
&= \frac{\sqrt{2}}{(2\pi)^{d/2}} \frac{n^n}{e^{b_n} (a_n d)^n a_n^{d/2}} \\
&= \frac{\sqrt{2}}{(2\pi)^{d/2}} \frac{(1 - b_n/n)^{-n}}{e^{b_n} a_n^{d/2}} \\
&\leq \frac{\sqrt{2} d^{d/2}}{(2\pi)^{d/2}} \frac{(1 - (d-1)/n)^{-n}}{(a_n d)^{d/2}} \\
&\simeq \frac{\sqrt{2} d^{d/2} e^{d-1}}{(2\pi)^{d/2}} \frac{1}{n^{d/2}},
\end{aligned}$$

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

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

hence  $\mathbb{P}(T_0^r = \infty) > 0$  by the result of Question (e), and the random walk is *not* recurrent when  $\underline{d \geq 3}$ .

Problem 3.10

- 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  $\boxed{x+1}$  before hitting state  $\textcircled{0}$  starting from  $\textcircled{k}$  is  $k/(x+1)$ , by first step analysis we find the probability

$$p + (1-p) \frac{x-1}{x+1} = 1 - \frac{2q}{x+1}, \quad x \geq 1.$$

- 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  $\textcircled{0}$  to state  $\textcircled{1}$ , then from state  $\textcircled{2}$  to state  $\textcircled{2}$ , etc, up to state  $\textcircled{x}$ , we find

$$\mathbb{P}(\tau_x < \tau_0 \mid S_0 = 0) = \prod_{l=0}^{x-1} \mathbb{P}(\tau_{l+1} < \tau_0 \mid S_0 = l)$$

$$\begin{aligned}
&= \mathbb{P}(\tau_1 < \tau_0 \mid S_0 = 0) \prod_{l=1}^{x-1} \mathbb{P}(\tau_{l+1} < \tau_0 \mid S_0 = l) \\
&= \frac{1}{2} \prod_{l=1}^{x-1} \left(1 - \frac{2q}{l+1}\right) \\
&= \frac{1}{2} \prod_{l=2}^x \left(1 - \frac{2q}{l}\right), \quad x \geq 1.
\end{aligned}$$

d) Using the inequality  $\log(1+z) \leq z$  for  $z > -1$ , we have

$$\begin{aligned}
\sum_{l=2}^x \log \left(1 - \frac{2q}{l}\right) &\leq \int_2^x \log \left(1 - \frac{2q}{y}\right) dy \\
&\leq -2q \int_2^x \frac{1}{y} dy \\
&= -2q \log \frac{x}{2},
\end{aligned}$$

hence

$$\begin{aligned}
\mathbb{P}(\tau_x < \tau_0 \mid S_0 = 0) &= \frac{1}{2} \exp \left( \sum_{l=2}^x \left(1 - \frac{2q}{l}\right) \right) \\
&\leq \exp \left( -2q \log \frac{x}{2} \right) \leq \left( \frac{x}{2} \right)^{-2q},
\end{aligned}$$

$$x \geq 2.$$

e) We have

$$\begin{aligned}
\mathbb{P}(\tau_0 < \infty \mid S_0 = 0) &\geq \mathbb{P} \left( \bigcup_{x \geq 1} \tau_0 < \tau_x \mid S_0 = 0 \right) \\
&= \lim_{x \rightarrow \infty} \mathbb{P}(\tau_0 < \tau_x \mid S_0 = 0) \\
&= \lim_{x \rightarrow \infty} (1 - \mathbb{P}(\tau_x \leq \tau_0 \mid S_0 = 0)) \\
&= \lim_{x \rightarrow \infty} (1 - \mathbb{P}(\tau_x < \tau_0 \mid S_0 = 0)) \\
&\geq 1 - \lim_{x \rightarrow \infty} (x/2)^{-2q} \\
&= 1.
\end{aligned}$$

**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

$$\mathbb{E}[\tau_1 \mid S_0 = 0] = \frac{1}{2} \times 1 + \frac{1}{2}(1 + \mathbb{E}[\tau_1 \mid S_0 = 0]),$$

hence

$$\mathbb{E}[\tau_1 \mid S_0 = 0] = 2.$$

ii) By pathwise analysis. We have

$$\mathbb{E}[\tau_1 \mid S_0 = 0] = \sum_{k \geq 1} \frac{k}{2^k} = \frac{1}{2} \sum_{k \geq 1} \frac{k}{2^{k-1}} = \frac{1/2}{(1 - 1/2)^2} = 2.$$

iii) By applying the result of Question (d) of Problem 2.7 page 74 with  $S = 1$  and  $k = 0$ , which shows that

$$\mathbb{E}[\tau_1 \mid S_0 = 0] = (S + k + 1)(S - k) = 2.$$

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

$$\begin{aligned} \mathbb{E}[\tau_x \mid S_0 = 0] &= 2 + \sum_{k=1}^{x-1} (1 + q(4k + 2)) \\ &= 2 + (1 + 2q)(x - 1) + 4q \sum_{k=1}^{x-1} k \\ &= 2 + (1 + 2q)(x - 1) + 2qx(x - 1) \\ &= 2 + (1 + 2q)x - (1 + 2q) + 2qx^2 - 2qx \\ &= 1 - 2q + x + 2qx^2, \quad x \geq 1. \end{aligned}$$

i) Using the notation of Problem 2.7 we have

$$\begin{aligned} \mathbb{P}(S_1 = x + 1 \mid S_0 = x \text{ and } \tau_{x+1} < \tau_0) &= p \frac{\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x + 1)}{\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x)} \\ &= \frac{p}{\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x)} \end{aligned}$$

and

$$\begin{aligned} \mathbb{P}(S_1 = x - 1 \mid S_0 = x, \tau_{x+1} < \tau_0) &= q \frac{\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x - 1)}{\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x)} \\ &= \frac{q(x - 1)/(x + 1)}{\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x)}, \end{aligned}$$

as we have  $\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x+1) = 1$  and

$$\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x-1) = \frac{x-1}{x+1},$$

because the random walk becomes symmetric when started from state  $\boxed{x-1}$ . Next, we note that, according to Question (b),  $\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x)$  can be computed as

$$\begin{aligned} \mathbb{P}(\tau_{x+1} < \tau_0 \mid S_0 = x) &= p\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_1 = x+1) \\ &\quad + q\mathbb{P}(\tau_{x+1} < \tau_0 \mid S_1 = x-1) \\ &= p + q\frac{x-1}{x+1}, \end{aligned}$$

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

$$\begin{aligned} \mathbb{E}[\tau_{x+1} \mid S_0 = x, \tau_{x+1} < \tau_0] &= 1 + \frac{((x+1)^2 - (x-1)^2)}{3} \frac{q(x-1)/(x+1)}{1 - 2q/(x+1)} \\ &= 1 + \frac{4qx(x-1)/(x+1)}{3(1 - 2q/(x+1))} \\ &= 1 + \frac{4qx(x-1)}{3(x+1-2q)}, \end{aligned}$$

which yields  $1 + (2x-2)/3$  when  $p = q = 1/2$ .

**Remark.** Using the above result we can also compute the average time to reach state  $\textcircled{x}$  with  $x \geq 2$  after starting from state  $\textcircled{1}$  and given one does not hit  $\textcircled{0}$  by the summation

$$\begin{aligned} \mathbb{E}[\tau_x \mid S_0 = 1, \tau_x < \tau_0] &= \sum_{k=2}^x \mathbb{E}[\tau_k \mid S_0 = k-1, \tau_k < \tau_0] \\ &= x-1 + \frac{4q}{3} \sum_{k=2}^x \frac{(k-1)(k-2)}{k-2q} \\ &= x-1 + \frac{2qx(x-1)}{3} - \frac{8p\tilde{q}}{3} \sum_{k=1}^{x-1} \frac{k}{k+1-2q} \end{aligned}$$

$$\begin{aligned}
&= x - 1 + \frac{2qx(x-1)}{3} - \frac{8pq}{3}(x-1) + \frac{8pq}{3}(p-q) \sum_{k=1}^{x-1} \frac{1}{k+1-2q} \\
&= x - 1 + \frac{4q}{3} \left( \frac{x(x-1)}{2} - 2p(x-1) + 2p(p-q) \sum_{k=1}^{x-1} \frac{1}{k+1-2q} \right).
\end{aligned}$$

When  $p = q = 1/2$  we recover the known expression

$$\begin{aligned}
\mathbb{E}[\tau_x \mid S_0 = 1, \tau_x < \tau_0] &= x - 1 + \frac{2}{3} \sum_{k=1}^{x-2} k \\
&= x - 1 + \frac{(x-1)(x-2)}{3} \\
&= \frac{x^2 - 1}{3}, \quad x \geq 2,
\end{aligned}$$

cf. Question (j) of Problem 2.7.

## 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

$$\begin{aligned}
Z_{n+1} &= (S_{n+1})^2 \\
&= (S_{n+1} - S_n + S_n)^2 \\
&= (S_n)^2 + 2(S_{n+1} - S_n)S_n + (S_{n+1} - S_n)^2 \\
&= (S_n)^2 + 2(S_{n+1} - S_n)S_n + 1 \\
&= Z_n + 2X_{n+1}S_n + 1,
\end{aligned}$$

where  $(X_n)_{n \geq 1} = (S_n - S_{n-1})_{n \geq 1}$  denotes the independent increments of  $(S_n)_{n \geq 0}$ , and the increment of  $(Z_n)_{n \in \mathbb{N}}$  is

$$Z_{n+1} - Z_n = 2X_{n+1}S_n + 1, \quad n \geq 0.$$

In the symmetric case  $\mathbb{P}(X_{n+1} = \pm 1) = 1/2$ , the process  $(Z_n)_{n \geq 0}$  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

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

the process  $(Z_n)_{n \geq 0}$  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  $Z_n = |S_n|^2$  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.

#### Exercise 4.2

a) By the Markov property, we have

$$\begin{aligned} & \mathbb{P}(X_7 = 1 \text{ and } X_5 = 2 \mid X_4 = 1 \text{ and } X_3 = 2) \\ &= \mathbb{P}(X_7 = 1 \text{ and } X_5 = 2 \mid X_4 = 1) \\ &= \frac{\mathbb{P}(X_7 = 1, X_5 = 2, X_4 = 1)}{\mathbb{P}(X_4 = 1)} \\ &= \frac{\mathbb{P}(X_7 = 1, X_5 = 2, X_4 = 1)}{\mathbb{P}(X_5 = 2 \text{ and } X_4 = 1)} \frac{\mathbb{P}(X_5 = 2 \text{ and } X_4 = 1)}{\mathbb{P}(X_4 = 1)} \\ &= \mathbb{P}(X_7 = 1 \mid X_5 = 2 \text{ and } X_4 = 1) \mathbb{P}(X_5 = 2 \mid X_4 = 1) \\ &= 0.6 \times \mathbb{P}(X_7 = 1 \mid X_5 = 2). \end{aligned}$$

Next, we note that

$$P^2 = \begin{bmatrix} 0.4 & 0.6 \\ 0.8 & 0.2 \end{bmatrix} \times \begin{bmatrix} 0.4 & 0.6 \\ 0.8 & 0.2 \end{bmatrix} = \begin{bmatrix} 0.64 & 0.36 \\ 0.48 & 0.52 \end{bmatrix},$$

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

$$\begin{aligned}\mathbb{E}[X_2 \mid X_1 = 1] &= 1 \times \mathbb{P}(X_2 = 1 \mid X_1 = 1) + 2 \times \mathbb{P}(X_2 = 2 \mid X_1 = 1) \\ &= 0.4 + 2 \times 0.6 = 1.6.\end{aligned}$$

### Exercise 4.3

a) We have

$$P^n = P = \begin{bmatrix} \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \pi_0 & \pi_1 & \pi_2 & \pi_3 & \cdots & \pi_N \end{bmatrix},$$

for all  $n \geq 1$ .

b) The vector  $\pi$  is an invariant (or stationary) distribution for  $P$  because it satisfies

$$[\pi_0, \pi_1, \dots, \pi_N] = [\pi_0, \pi_1, \dots, \pi_N] \times \begin{bmatrix} \pi_0 & \pi_1 & \pi_2 & \cdots & \pi_N \\ \pi_0 & \pi_1 & \pi_2 & \cdots & \pi_N \\ \pi_0 & \pi_1 & \pi_2 & \cdots & \pi_N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \pi_0 & \pi_1 & \pi_2 & \cdots & \pi_N \end{bmatrix},$$

as in (4.2.5).

c) Since  $P^n = P$ ,  $n \geq 1$ , we have

$$\mathbb{P}(Z_n = i \mid Z_k = j) = \mathbb{P}(Z_n = i) = \pi_i, \quad i, j = 0, 1, \dots, N, \quad 0 \leq k < n,$$

hence by the Markov property we find

$$\begin{aligned}\mathbb{P}(Z_n = i \text{ and } Z_k = j) &= \mathbb{P}(Z_n = i \mid Z_k = j) \mathbb{P}(Z_k = j) \quad (\text{S.29}) \\ &= \mathbb{P}(Z_n = i) \mathbb{P}(Z_k = j),\end{aligned}$$

$i, j = 0, 1, \dots, N$ , 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

$$\begin{aligned}\mathbb{P}(Z_n = i_n, Z_{n-1} = i_{n-1}, \dots, Z_0 = i_0) \\ &= \mathbb{P}(Z_n = i_n \mid Z_{n-1} = i_{n-1}) \cdots \mathbb{P}(Z_1 = i_1 \mid Z_0 = i_0) \mathbb{P}(Z_0 = i_0) \\ &= \mathbb{P}(Z_n = i_n) \mathbb{P}(Z_{n-1} = i_{n-1}) \cdots \mathbb{P}(Z_0 = i_0) \\ &= \pi_{i_n} \pi_{i_{n-1}} \cdots \pi_{i_0},\end{aligned}$$

$i_0, i_1, \dots, i_n \in \{0, 1, \dots, N\}$ , which shows that  $(Z_n)_{n \in \mathbb{N}}$  is an *i.i.d.* sequence of random variables with distribution  $\pi$  over  $\{0, 1, \dots, N\}$ .

#### Exercise 4.4

a) We find

$$\begin{cases} \mathbb{P}(O_0 = a \mid X_0 = 0)\mathbb{P}(O_1 = b \mid X_1 = 0) = m_{0,a}m_{0,b}, \\ \mathbb{P}(O_0 = a \mid X_0 = 0)\mathbb{P}(O_1 = b \mid X_1 = 1) = m_{0,a}m_{1,b}, \\ \mathbb{P}(O_0 = a \mid X_0 = 1)\mathbb{P}(O_1 = b \mid X_1 = 0) = m_{1,a}m_{0,b}, \\ \mathbb{P}(O_0 = a \mid X_0 = 1)\mathbb{P}(O_1 = b \mid X_1 = 1) = m_{1,a}m_{1,b}, \end{cases}$$

b) We find  $\mathbb{P}(X_0 = 1, X_1 = 1) = \mathbb{P}((X_0, X_1) = (1, 1)) = \pi_1 P_{1,1}$ , and

$$\begin{cases} \mathbb{P}((O_0, O_1) = (a, b) \text{ and } (X_0, X_1) = (0, 0)) = \pi_0 P_{0,0} m_{1,a} m_{0,b}, \\ \mathbb{P}((O_0, O_1) = (a, b) \text{ and } (X_0, X_1) = (0, 1)) = \pi_0 P_{0,1} m_{0,a} m_{1,b}, \\ \mathbb{P}((O_0, O_1) = (a, b) \text{ and } (X_0, X_1) = (1, 0)) = \pi_1 P_{1,0} m_{1,a} m_{0,b}, \\ \mathbb{P}((O_0, O_1) = (a, b) \text{ and } (X_0, X_1) = (1, 1)) = \pi_1 P_{1,1} m_{1,a} m_{1,b}, \end{cases}$$

c) We have

$$\begin{aligned} & \mathbb{P}((O_0, O_1) = (a, b)) \\ &= \sum_{x,y \in \{0,1\}} \mathbb{P}((O_0, O_1) = (a, b) \text{ and } (X_0, X_1) = (x, y)) \\ &= \sum_{x,y \in \{0,1\}} \pi_x P_{x,y} m_{x,a} m_{y,b} \\ &= \pi_0 P_{0,0} m_{0,a} m_{0,b} + \pi_0 P_{0,1} m_{0,a} m_{1,b} + \pi_1 P_{1,0} m_{1,a} m_{0,b} + \pi_1 P_{1,1} m_{1,a} m_{1,b}. \end{aligned}$$

d) We have

$$\begin{cases} \mathbb{P}((X_0, X_1) = (0, 1) \mid (O_0, O_1) = (a, b)) = \pi_0 P_{0,1} m_{0,a} m_{1,b} / \mathbb{P}((O_0, O_1) = (a, b)), \\ \mathbb{P}((X_0, X_1) = (1, 0) \mid (O_0, O_1) = (a, b)) = \pi_1 P_{1,0} m_{1,a} m_{0,b} / \mathbb{P}((O_0, O_1) = (a, b)). \end{cases}$$

On the other hand, we have

$$\begin{aligned} \{X_1 = 1\} &= \{(X_0, X_1) = (0, 1)\} \cup \{(X_0, X_1) = (1, 1)\} \\ &= \bigcup_{x \in \{0,1\}} \{(X_0, X_1) = (x, 1)\} \end{aligned}$$

$$= \{(X_0, X_1) = (0, 1)\} \cup \{(X_0, X_1) = (1, 1)\},$$

where the above union is a partition, hence

$$\begin{aligned} & \mathbb{P}(X_1 = 1 \mid (O_0, O_1) = (a, b)) \\ &= \sum_{x \in \{0,1\}} \mathbb{P}((X_0, X_1) = (x, 1) \mid (O_0, O_1) = (a, b)) \\ &= \frac{1}{\mathbb{P}((O_0, O_1) = (a, b))} \sum_{x \in \{0,1\}} \mathbb{P}((X_0, X_1) = (x, 1) \text{ and } (O_0, O_1) = (a, b)) \\ &= \frac{1}{\mathbb{P}((O_0, O_1) = (a, b))} \sum_{x \in \{0,1\}} \pi_x P_{x,1} m_{x,a} m_{1,b} \\ &= \frac{1}{\mathbb{P}((O_0, O_1) = (a, b))} (\pi_0 P_{0,1} m_{0,a} m_{1,b} + \pi_1 P_{1,1} m_{1,a} m_{1,b}). \end{aligned}$$

#### Exercise 4.5

- a) 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 depend on past chain values from time 0. Other examples include  $(99, 101)$  and  $(11, 141)$  whose sums of squares are both equal to  $99^2 + 101^2 = 11^2 + 141^2 = 20002$ .
- b) The random walk  $(S_n)_{n \in \mathbb{N}}$  cannot have the Markov property 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  $n$ .

**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  $(X_1, \dots, X_{n-1})$ .

**Exercise 4.7** The state space of  $(Z_n)_{n \geq 1}$  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  $(Z_n)_{n \geq 1}$  is a  $(0, 1)$ -valued Markov chain and its transition matrix is given by

$$\begin{array}{cc} & \begin{array}{cccc} 00 & 01 & 10 & 11 \end{array} \\ \begin{array}{c} 00 \\ 01 \\ 10 \\ 11 \end{array} & \begin{bmatrix} 1-a & a & 0 & 0 \\ 0 & 0 & b & 1-b \\ 1-a & a & 0 & 0 \\ 0 & 0 & b & 1-b \end{bmatrix}, \end{array}$$

### Exercise 4.8

- a) The probability distribution of  $T_2$  is negative binomial with parameters  $(p, 2)$ , *i.e.*

$$\mathbb{P}(T_2 = k \mid X_0 = 0) = \binom{k-1}{k-2} (1-p)^2 p^{k-2}, \quad k \geq 2.$$

- b) We have

$$\begin{aligned} \mathbb{E}[T_2 \mid X_0 = 0] &= \sum_{k \geq 2} k \mathbb{P}(T_2 = k \mid X_0 = 0) \\ &= (1-p)^2 \sum_{k \geq 2} k \binom{k-1}{k-2} p^{k-2} \\ &= (1-p)^2 \sum_{k \geq 2} k(k-1) p^{k-2} \\ &= (1-p)^2 \frac{\partial^2}{\partial p^2} \sum_{k \geq 0} p^k \\ &= (1-p)^2 \frac{\partial^2}{\partial p^2} \frac{1}{1-p} \\ &= \frac{2}{1-p} \\ &= \frac{2}{q}. \end{aligned}$$

**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

$$\mathbb{E}[T_2 \mid X_0 = 1] = p(1 + \mathbb{E}[T_2 \mid X_0 = 1]) + q \times 1.$$

and

$$\mathbb{E}[T_2 \mid X_0 = 0] = p(1 + \mathbb{E}[T_2 \mid X_0 = 0]) + q \times (1 + \mathbb{E}[T_2 \mid X_0 = 1]),$$

which yield

$$\mathbb{E}[T_2 \mid X_0 = 1] = \frac{1}{q} \quad \text{and} \quad \mathbb{E}[T_2 \mid X_0 = 0] = \frac{2}{q}.$$

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  $N \geq 1$ , the transition matrix reads

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 1/N^2 & 2(N-1)/N^2 & (N-1)^2/N^2 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 2^2/N^2 & 4(N-2)/N^2 & (N-2)^2/N^2 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 3^2/N^2 & 0 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \vdots & 0 & 3^2/N^2 & 0 & 0 \\ 0 & 0 & \cdots & 0 & (N-2)^2/N^2 & 4(N-2)/N^2 & 2^2/N^2 & 0 \\ 0 & 0 & \cdots & 0 & 0 & (N-1)^2/N^2 & 2(N-1)/N^2 & 1/N^2 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 \end{bmatrix},$$

and the matrix entries are given in general as

$$\begin{cases} P_{k,k-1} = \mathbb{P}(X_1 = k-1 \mid X_0 = k) = \frac{k^2}{N^2}, \\ P_{k,k} = \mathbb{P}(X_1 = k \mid X_0 = k) = \frac{2k(N-k)^2}{N^2}, \\ P_{k,k+1} = \mathbb{P}(X_1 = k+1 \mid X_0 = k) = \frac{(N-k)^2}{N^2}, \end{cases}$$

$k = 1, 2, \dots, N-1$ , with  $P_{0,1} = P_{N,N-1} = 1$ . We check that

$$P_{k,k-1} + P_{k,k} + P_{k,k+1} = \frac{k^2 + 2k(N-k)^2 + (N-k)^2}{N^2} = 1, \quad k = 1, 2, \dots, N-1.$$

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

$$P = [P_{i,j}]_{i,j \in \mathbb{N}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ q & 0 & p & 0 & 0 & 0 & \cdots \\ q & 0 & 0 & p & 0 & 0 & \cdots \\ q & 0 & 0 & 0 & p & 0 & \cdots \\ q & 0 & 0 & 0 & 0 & p & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

After  $n$  time steps we have

$$\mathbb{P}(S_n = n + 1 \mid S_0 = 1) = p^n, \quad n \geq 1,$$

and, by decomposing over  $n$  possible paths made of  $l$  steps up and one step down,  $l = 0, 1, \dots, n - 1$ , we find

$$\begin{aligned} \mathbb{P}(S_n = 0 \mid S_0 = 1) &= \sum_{l=0}^{n-1} qp^l \\ &= q \frac{1 - p^n}{1 - p} \\ &= 1 - p^n \\ &= 1 - \mathbb{P}(S_n = n + 1 \mid S_0 = 1), \quad n \geq 1. \end{aligned}$$

More generally, we have

$$\mathbb{P}(S_n = n + k \mid S_0 = k) = p^n, \quad k, n \geq 1,$$

and

$$\begin{aligned} \mathbb{P}(S_n = 0 \mid S_0 = k) &= q \sum_{l=0}^{n-1} p^l \\ &= q \frac{1 - p^n}{1 - p} \\ &= 1 - p^n \\ &= 1 - \mathbb{P}(S_n = n + k \mid S_0 = k), \quad k, n \geq 1, \end{aligned}$$

by (A.2), hence  $P^n$  is given by

$$P^n = \left[ [P^n]_{i,j} \right]_{i,j \in \mathbb{N}} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & p^n & 0 & 0 & 0 & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & 0 & p^n & 0 & 0 & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & 0 & 0 & p^n & 0 & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & 0 & 0 & 0 & p^n & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & p^n & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad (\text{S.30})$$

in which the  $n$  columns  $n^\circ 2$  to  $n+1$  are identically 0. We can also check by matrix multiplication that this relation is consistent with  $P^{n+1} = P \times P^n$ ,  $n \geq 1$ , *i.e.*

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1-p^{n+1} & 0 & \cdots & 0 & p^{n+1} & 0 & 0 & 0 & 0 & \cdots \\ 1-p^{n+1} & 0 & \cdots & 0 & 0 & p^{n+1} & 0 & 0 & 0 & \cdots \\ 1-p^{n+1} & 0 & \cdots & 0 & 0 & 0 & p^{n+1} & 0 & 0 & \cdots \\ 1-p^{n+1} & 0 & \cdots & 0 & 0 & 0 & 0 & p^{n+1} & 0 & \cdots \\ 1-p^{n+1} & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & p^{n+1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \\ = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1-p & 0 & p & 0 & 0 & 0 & 0 & \cdots \\ 1-p & 0 & 0 & p & 0 & 0 & 0 & \cdots \\ 1-p & 0 & 0 & 0 & p & 0 & 0 & \cdots \\ 1-p & 0 & 0 & 0 & 0 & p & 0 & \cdots \\ 1-p & 0 & 0 & 0 & 0 & 0 & p & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \times \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & p^n & 0 & 0 & 0 & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & 0 & p^n & 0 & 0 & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & 0 & 0 & p^n & 0 & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & 0 & 0 & 0 & p^n & 0 & \cdots \\ 1-p^n & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & p^n & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

from the relation  $1 - p + p(1 - p^n) = 1 - p^{n+1}$ ,  $n \in \mathbb{N}$ .

b) In this case, the transition matrix  $P$  becomes

$$P = \left[ P_{i,j} \right]_{i,j \in \mathbb{N}} = \begin{bmatrix} q & p & 0 & 0 & 0 & 0 & \cdots \\ q & 0 & p & 0 & 0 & 0 & \cdots \\ q & 0 & 0 & p & 0 & 0 & \cdots \\ q & 0 & 0 & 0 & p & 0 & \cdots \\ q & 0 & 0 & 0 & 0 & p & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

with

$$P^2 = \begin{bmatrix} q & qp & p^2 & 0 & 0 & 0 & \dots \\ q & qp & 0 & p^2 & 0 & 0 & \dots \\ q & qp & 0 & 0 & p^2 & 0 & \dots \\ q & qp & 0 & 0 & 0 & p^2 & \dots \\ q & qp & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

and more generally, by induction on  $n \geq 2$  we find

$$P^n = \begin{bmatrix} q & qp & qp^2 & \dots & qp^{n-1} & p^n & 0 & 0 & 0 & \dots \\ q & qp & qp^2 & \dots & qp^{n-1} & 0 & p^n & 0 & 0 & \dots \\ q & qp & qp^2 & \dots & qp^{n-1} & 0 & 0 & p^n & 0 & \dots \\ q & qp & qp^2 & \dots & qp^{n-1} & 0 & 0 & 0 & p^n & \dots \\ q & qp & qp^2 & \dots & qp^{n-1} & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (\text{S.31})$$

Exercise 4.11

a) We have

$$\begin{aligned} \mathbb{P}(Z_{n+1} = 2 \mid Z_n = 0 \text{ and } Z_{n-1} = 2) &= \frac{\mathbb{P}(Z_{n+1} = 2, Z_n = 0, Z_{n-1} = 2)}{\mathbb{P}(Z_n = 0 \text{ and } Z_{n-1} = 2)} \\ &= \frac{\mathbb{P}(X_{n+1} = 2, X_n \in \{0, 1\}, X_{n-1} = 2)}{\mathbb{P}(X_n \in \{0, 1\} \text{ and } X_{n-1} = 2)} \\ &= \frac{\mathbb{P}(X_{n+1} = 2, X_n = 0, X_{n-1} = 2)}{\mathbb{P}(X_n \in \{0, 1\} \text{ and } X_{n-1} = 2)} + \frac{\mathbb{P}(X_{n+1} = 2, X_n = 1, X_{n-1} = 2)}{\mathbb{P}(X_n \in \{0, 1\} \text{ and } X_{n-1} = 2)} \\ &= 0, \end{aligned}$$

$n \geq 1$ . Next, we have

$$\begin{aligned} \mathbb{P}(Z_{n+1} = 2 \mid Z_n = 0 \text{ and } Z_{n-1} = 3) &= \frac{\mathbb{P}(Z_{n+1} = 2, Z_n = 0, Z_{n-1} = 3)}{\mathbb{P}(Z_n = 0 \text{ and } Z_{n-1} = 3)} \\ &= \frac{\mathbb{P}(X_{n+1} = 2, X_n \in \{0, 1\}, X_{n-1} = 3)}{\mathbb{P}(X_n \in \{0, 1\} \text{ and } X_{n-1} = 3)} \\ &= \frac{\mathbb{P}(X_{n+1} = 2, X_n = 0, X_{n-1} = 3)}{\mathbb{P}(X_n \in \{0, 1\} \text{ and } X_{n-1} = 3)} + \frac{\mathbb{P}(X_{n+1} = 2, X_n = 1, X_{n-1} = 3)}{\mathbb{P}(X_n \in \{0, 1\} \text{ and } X_{n-1} = 3)} \\ &= \frac{\mathbb{P}(X_{n+1} = 2, X_n = 0, X_{n-1} = 3)}{\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3) + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)} \\ &\quad + \frac{\mathbb{P}(X_{n+1} = 2, X_n = 1, X_{n-1} = 3)}{\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3) + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)} \\ &= \frac{\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3) \mathbb{P}(X_{n+1} = 2 \mid X_n = 0 \text{ and } X_{n-1} = 3)}{\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3) + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)} \end{aligned}$$

$$\begin{aligned}
& + \frac{\mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)\mathbb{P}(X_{n+1} = 2 \mid X_n = 1 \text{ and } X_{n-1} = 3)}{\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3) + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)} \\
& = \frac{\mathbb{P}(X_{n+1} = 2 \mid X_n = 0)}{1 + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)/\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3)} \\
& \quad + \frac{\mathbb{P}(X_{n+1} = 2 \mid X_n = 1)}{1 + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)/\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3)} \\
& = \frac{\mathbb{P}(X_{n+1} = 2 \mid X_n = 0)}{1 + \mathbb{P}(X_n = 1 \text{ and } X_{n-1} = 3)/\mathbb{P}(X_n = 0 \text{ and } X_{n-1} = 3)} \\
& = \frac{1/2}{1 + (1/3)/(1/3)} \\
& = \frac{1}{4},
\end{aligned}$$

$n \geq 1$ .

- b) The process  $(Z_n)_{n \in \mathbb{N}}$  is *not* a Markov chain because by Question (a) we have

$$\begin{aligned}
0 &= \mathbb{P}(Z_{n+1} = 2 \mid Z_n = 0 \text{ and } Z_{n-1} = 2) \\
&\neq \mathbb{P}(Z_{n+1} = 2 \mid Z_n = 0 \text{ and } Z_{n-1} = 3) \\
&= \frac{1}{4}.
\end{aligned}$$

#### 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

$$\begin{aligned}
\eta P &= [50\%, 40\%, 10\%] \times \begin{bmatrix} 2/3 & 1/6 & 1/6 \\ 1/3 & 1/2 & 1/6 \\ 1/6 & 2/3 & 1/6 \end{bmatrix} \\
&= \left[ \frac{2 \times 0.5}{3} + \frac{0.4}{3} + \frac{0.1}{6}, \frac{0.5}{6} + \frac{0.4}{2} + \frac{2 \times 0.1}{3}, \frac{0.5}{6} + \frac{0.4}{6} + \frac{0.1}{6} \right] \\
&= \left[ \frac{29}{60}, \frac{21}{60}, \frac{10}{60} \right] \\
&\simeq [48.3\%, 35\%, 16.7\%].
\end{aligned}$$

- 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

$$[\pi_L, \pi_M, \pi_H] = [\pi_L, \pi_M, \pi_H] \times P = [\pi_L, \pi_M, \pi_H] \times \begin{bmatrix} 8/12 & 2/12 & 2/12 \\ 4/12 & 6/12 & 2/12 \\ 2/12 & 8/12 & 2/12 \end{bmatrix},$$

and it can be rewritten as

$$\pi(P - I) = \pi P - \pi = 0,$$

*i.e.*,

$$\begin{aligned} \pi(P - I) &= [\pi_L, \pi_M, \pi_H](P - I) \\ &= [\pi_L, \pi_M, \pi_H] \times \begin{bmatrix} -4/12 & 2/12 & 2/12 \\ 4/12 & -6/12 & 2/12 \\ 2/12 & 8/12 & -10/12 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \end{aligned}$$

which can be rewritten as

$$\begin{cases} \pi_L - 3\pi_M + 4\pi_H = 0 \\ \pi_L + \pi_M - 5\pi_H = 0. \end{cases}$$

Combining the above with the additional condition

$$\pi_L + \pi_M + \pi_H = 1,$$

we arrive at

$$[\pi_L, \pi_M, \pi_H] = \left[ \frac{11}{24}, \frac{9}{24}, \frac{4}{24} \right] \simeq [45.8\%, 37.5\%, 16.7\%].$$

Problem 4.13

a) We have

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

with  $q := 1 - p$ .

b) We have

$$P = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ q & p & 0 & \cdots & 0 & 0 \\ 0 & q & p & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & q & p & 0 \\ 0 & 0 & \cdots & 0 & q & p \end{bmatrix},$$

with  $q := 1 - p$ .

- c) Clearly, the first row of  $P$  has to be  $[1, 0, \dots, 0]$  because state ① 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.*

$$\alpha_k + \sum_{l=1}^d Q_{k,l} = 1, \quad k = 1, 2, \dots, d,$$

which can be rewritten as

$$\begin{aligned} (I - Q)e &= \begin{bmatrix} 1 - Q_{1,1} & -Q_{1,2} & \cdots & -Q_{1,d} \\ -Q_{1,1} & 1 - Q_{1,2} & \cdots & -Q_{1,d} \\ \vdots & \vdots & \ddots & \vdots \\ -Q_{d,1} & \cdots & Q_{d,d-1} & 1 - Q_{d,d} \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 - Q_{1,1} - \cdots - Q_{1,d} \\ 1 - Q_{2,1} - \cdots - Q_{2,d} \\ \vdots \\ 1 - Q_{d,1} - \cdots - Q_{d,d} \end{bmatrix} \\ &= \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_d \end{bmatrix}, \end{aligned}$$

*i.e.*

$$\alpha = (I - Q)e.$$

- d) Clearly, the conclusion holds for  $n = 0$ , and also at the rank  $n = 1$  since  $\alpha = (I - Q)e$ . Next, we assume that the relation (4.5.9) holds at the rank  $n \geq 0$ . In this case, we have

$$\begin{aligned} P^{n+1} &= P \times P^n \\ &= \begin{bmatrix} 1 & 0 \\ \alpha & Q \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ (I - Q^n)e & Q^n \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ \alpha + Q(I - Q^n)e & Q^{n+1} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ (I - Q^{n+1})e & Q^{n+1} \end{bmatrix}, \end{aligned}$$

since

$$\alpha + Q(I - Q^n)e = (I - Q)e + (Q - Q^{n+1})e = (I - Q^{n+1})e.$$

e) Since  $[P^{n-1}]_{i,k} = [Q^{n-1}]_{i,k}$  and  $\alpha_k = P_{k,0}$ ,  $k = 1, 2, \dots, d$ , we find that

$$\begin{aligned}
 \mathbb{P}(T_0 = n \mid X_0 = i) &= \mathbb{P}\left(\bigcup_{k=1}^d \{X_{n-1} = k, T_0 = n\} \mid X_0 = i\right) \\
 &= \sum_{k=1}^d \mathbb{P}(X_{n-1} = k \mid X_0 = i) \mathbb{P}(X_n = 0 \mid X_{n-1} = k) \\
 &= \sum_{k=1}^d [P^{n-1}]_{i,k} P_{k,0} \\
 &= \sum_{k=1}^d \alpha_k [Q^{n-1}]_{i,k} \\
 &= [Q^{n-1} \alpha]_i, \quad n \geq 1.
 \end{aligned}$$

f) We have

$$\begin{aligned}
 \mathbb{P}(T_0 = n) &= \sum_{i=1}^d \mathbb{P}(X_0 = i) \mathbb{P}(T_0 = n \mid X_0 = i) \\
 &= \sum_{i=1}^d \sum_{k=1}^d \beta_i \alpha_k Q_{i,k}^{n-1} \\
 &= \beta^\top Q^{n-1} \alpha, \quad n \geq 1.
 \end{aligned}$$

g) We have

$$\begin{aligned}
 \mathbb{P}(T_0 \leq n) &= \sum_{k=1}^n \mathbb{P}(T_0 = k) \\
 &= \sum_{k=1}^n \beta^\top Q^{k-1} \alpha \\
 &= \beta^\top (I - Q^n)(I - Q)^{-1} \alpha \\
 &= \beta^\top (I - Q^n) e \\
 &= 1 - \beta^\top Q^n e, \quad n \geq 1.
 \end{aligned}$$

Alternatively, using the relation  $\alpha = (I - Q)e$  and a telescopic sum, we recover this result as

$$\begin{aligned}
 \mathbb{P}(T_0 \leq n) &= \sum_{k=1}^n \mathbb{P}(T_0 = k) \\
 &= \sum_{k=1}^n \beta^\top Q^{k-1} \alpha
 \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^n \beta^\top Q^{k-1} (I - Q) e \\
&= \sum_{k=0}^{n-1} \beta^\top Q^k e - \sum_{k=1}^n \beta^\top Q^k e \\
&= \beta^\top e - \beta^\top Q^n e \\
&= 1 - \beta^\top Q^n e, \quad n \geq 1.
\end{aligned}$$

h) We have

$$\begin{aligned}
\mathbb{P}(T_0 \leq n) &= 1 - \sum_{k=1}^d \mathbb{P}(X_n = k) \\
&= 1 - \sum_{k=1}^d \sum_{i=1}^d \beta_i \mathbb{P}(X_n = k \mid X_0 = i) \\
&= 1 - \sum_{k=1}^d \sum_{i=1}^d \beta_i Q_{i,k}^n \\
&= 1 - \beta^\top Q^n e, \quad n \geq 1.
\end{aligned}$$

Alternatively we could also write

$$\begin{aligned}
\mathbb{P}(T_0 \leq n) &= \mathbb{P}(X_n = 0) \\
&= \sum_{i=1}^d \beta_i \mathbb{P}(X_n = 0 \mid X_0 = i) \\
&= \sum_{i=1}^d \beta_i [P^n]_{i,0} \\
&= \sum_{i=1}^d \beta_i [(I - Q^n)e]_i \\
&= \sum_{i=1}^d \beta_i - \sum_{i=1}^d \beta_i [Q^n e]_i \\
&= 1 - \beta^\top Q^n e, \quad n \geq 1.
\end{aligned}$$

i) We note that  $T_0$  is finite with probability one, since

$$\begin{aligned}
\mathbb{P}(T_0 < \infty) &= \sum_{n=0}^{\infty} \mathbb{P}(T_0 = n) \\
&= \sum_{n=1}^{\infty} \beta^\top Q^{n-1} \alpha
\end{aligned}$$

$$\begin{aligned}
&= \beta^\top \sum_{n=0}^{\infty} Q^n \alpha \\
&= \beta^\top (I - Q)^{-1} \alpha \\
&= \beta^\top (I - Q)^{-1} (I - Q) e \\
&= \sum_{i=1}^d \beta_i \\
&= \sum_{i=1}^d \mathbb{P}(X_0 = i) \\
&= 1.
\end{aligned}$$

Since the states  $\{1, 2, \dots, d\}$  are transient, Corollary 6.9 shows that the matrix inverse  $(I - sQ)^{-1}$  exists and is given by the series

$$(I - sQ)^{-1} = \sum_{k \geq 0} s^k Q^k, \quad s \in (-1, 1],$$

hence we have

$$\begin{aligned}
G_{T_0}(s) &= \mathbb{E}[s^{T_0}] \\
&= \sum_{k \geq 0} s^k \mathbb{P}(T_0 = k) \\
&= \mathbb{P}(X_0 = 0) + \sum_{k \geq 1} s^k \beta^\top Q^{k-1} \alpha \\
&= s \sum_{k \geq 0} s^k \beta^\top Q^k \alpha \\
&= s \beta^\top (I - sQ)^{-1} \alpha \\
&= s \beta^\top (I - sQ)^{-1} (I - Q) e.
\end{aligned}$$

We note that

$$\mathbb{P}(T_0 < \infty) = G_{T_0}(1) = \beta^\top (I - Q)^{-1} (I - Q) e = \beta^\top e = 1,$$

hence state ① is reached in finite time with probability one.

j) We have

$$G'_{T_0}(s) = \beta^\top (I - sQ)^{-1} \alpha + s \beta^\top Q (I - sQ)^{-2} \alpha,$$

hence

$$\begin{aligned}
\mathbb{E}[T_0] &= G'_{T_0}(1) \\
&= \beta^\top (I - Q)^{-1} \alpha + \beta^\top Q (I - Q)^{-2} \alpha
\end{aligned}$$

$$\begin{aligned}
&= \beta^\top (I - Q)(I - Q)^{-2} \alpha + \beta^\top Q(I - Q)^{-2} \alpha \\
&= \beta^\top (I - Q)^{-2} \alpha \\
&= \beta^\top (I - Q)^{-1} e.
\end{aligned}$$

We also have

$$G''_{T_0}(s) = \beta^\top Q(I - sQ)^{-2} \alpha + \beta^\top Q(I - sQ)^{-2} \alpha + 2s\beta^\top Q^2(I - sQ)^{-3} \alpha,$$

hence

$$\begin{aligned}
\mathbb{E}[T_0(T_0 - 1)] &= G''_{T_0}(1) \\
&= 2\beta^\top Q(I - Q)^{-2} \alpha + 2\beta^\top Q^2(I - Q)^{-3} \alpha \\
&= 2\beta^\top Q(I - Q)^{-3} \alpha, \\
&= 2\beta^\top Q(I - Q)^{-2} e,
\end{aligned}$$

hence

$$\begin{aligned}
\mathbb{E}[T_0^2] &= \mathbb{E}[T_0(T_0 - 1)] + \mathbb{E}[T_0] \\
&= 2\beta^\top Q(I - Q)^{-2} e + \beta^\top (I - Q)^{-1} e \\
&= 2\beta^\top Q(I - Q)^{-2} e + \beta^\top (I - Q)(I - Q)^{-2} e \\
&= \beta^\top (I + Q)(I - Q)^{-2} e.
\end{aligned}$$

More generally, by (1.7.7) we could also compute the *factorial moment*

$$\mathbb{E}[T_0(T_0 - 1) \cdots (T_0 - k + 1)] = G_{T_0}^{(k)}(1) = k! \beta^\top Q^{k-1} (I - Q)^{-k} e,$$

for all  $k \geq 1$ .

#### Problem 4.14

a) We have

$$\begin{aligned}
&\mathbb{P}((O_0, O_1, O_2) = (c, a, b) \text{ and } (X_0, X_1, X_2) = (1, 1, 0)) \\
&= \mathbb{P}((O_0, O_1, O_2) = (c, a, b) \mid (X_0, X_1, X_2) = (1, 1, 0)) \\
&\quad \times \mathbb{P}((X_0, X_1, X_2) = (1, 1, 0)) \\
&= \mathbb{P}(O_0 = c \mid X_0 = 1) \mathbb{P}(O_1 = a \mid X_1 = 1) \\
&\quad \times \mathbb{P}(O_2 = b \mid X_2 = 0) \mathbb{P}((X_0, X_1, X_2) = (1, 1, 0)) \\
&= \pi_1 P_{1,1} P_{1,0} m_{1,c} m_{1,a} m_{0,b},
\end{aligned}$$

since

$$\mathbb{P}((X_0, X_1, X_2) = (1, 1, 0)) = \pi_1 P_{1,1} P_{1,0}.$$

b) We have

$$\begin{aligned}
& \mathbb{P}((O_0, O_1, O_2) = (c, a, b)) \\
&= \sum_{x, y, z \in \{0,1\}} \mathbb{P}((O_0, O_1, O_2) = (c, a, b) \text{ and } (X_0, X_1, X_2) = (x, y, z)) \\
&= \sum_{x, y, z \in \{0,1\}} \pi_x P_{x,y} P_{y,z} m_{x,c} m_{y,a} m_{z,b}.
\end{aligned}$$

c) We have

$$\begin{aligned}
\{X_1 = 1\} &= \{(X_0, X_1, X_2) = (0, 1, 0)\} \cup \{(X_0, X_1, X_2) = (0, 1, 1)\} \\
&\quad \cup \{(X_0, X_1, X_2) = (1, 1, 0)\} \cup \{(X_0, X_1, X_2) = (1, 1, 1)\} \\
&= \bigcup_{x, z \in \{0,1\}} \{(X_0, X_1, X_2) = (x, 1, z)\},
\end{aligned}$$

where the above union is a partition, hence

$$\begin{aligned}
& \mathbb{P}(X_1 = 1 \mid (O_0, O_1, O_2) = (c, a, b)) \\
&= \sum_{x, z \in \{0,1\}} \mathbb{P}((X_0, X_1, X_2) = (x, 1, z) \mid (O_0, O_1, O_2) = (c, a, b)) \\
&= \frac{1}{\mathbb{P}((O_0, O_1, O_2) = (c, a, b))} \\
&\times \sum_{x, z \in \{0,1\}} \mathbb{P}((X_0, X_1, X_2) = (x, 1, z) \text{ and } (O_0, O_1, O_2) = (c, a, b)) \\
&= \frac{1}{\mathbb{P}((O_0, O_1, O_2) = (c, a, b))} \sum_{x, z \in \{0,1\}} \pi_x P_{x,1} P_{1,z} m_{x,c} m_{1,b} m_{z,c}.
\end{aligned}$$

d) By the result of Question (a), we find

$$\begin{cases}
\mathbb{P}((X_0, X_1, X_2) = (0, 0, 0) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.00588, \\
\mathbb{P}((X_0, X_1, X_2) = (0, 0, 1) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.00126, \\
\mathbb{P}((X_0, X_1, X_2) = (0, 1, 0) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.0101, \\
\mathbb{P}((X_0, X_1, X_2) = (0, 1, 1) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.00756, \\
\mathbb{P}((X_0, X_1, X_2) = (1, 0, 0) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.000448, \\
\mathbb{P}((X_0, X_1, X_2) = (1, 0, 1) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.0000960, \\
\mathbb{P}((X_0, X_1, X_2) = (1, 1, 0) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.00269, \\
\mathbb{P}((X_0, X_1, X_2) = (1, 1, 1) \text{ and } (O_0, O_1, O_2) = (c, a, b)) = 0.00202.
\end{cases}$$

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) Proceeding as in the answer to Question (c), we obtain

$$\begin{cases} \mathbb{P}(X_0 = 0 \mid (O_0, O_1, O_2) = (c, a, b)) = 0.825, \\ \mathbb{P}(X_0 = 1 \mid (O_0, O_1, O_2) = (c, a, b)) = 0.175, \\ \mathbb{P}(X_1 = 0 \mid (O_0, O_1, O_2) = (c, a, b)) = 0.256, \\ \mathbb{P}(X_1 = 1 \mid (O_0, O_1, O_2) = (c, a, b)) = 0.744, \\ \mathbb{P}(X_2 = 0 \mid (O_0, O_1, O_2) = (c, a, b)) = 0.636, \\ \mathbb{P}(X_2 = 1 \mid (O_0, O_1, O_2) = (c, a, b)) = 0.364. \end{cases}$$

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

$$\hat{P} = \begin{bmatrix} 0.415 & 0.585 \\ 0.482 & 0.518 \end{bmatrix}.$$

i) We find

$$\hat{M} = \begin{bmatrix} 0.149 & 0.370 & 0.481 \\ 0.580 & 0.284 & 0.136 \end{bmatrix}.$$

j) The estimates of the matrix  $M$  obtained from the  code can be plotted as follows:

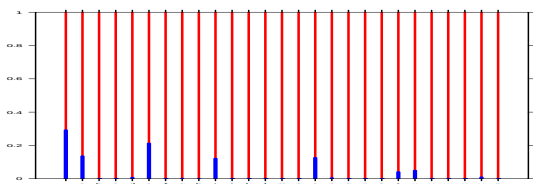


Fig. S.5: Plot of `estimate$hmm$emissionProbs[1,]`.

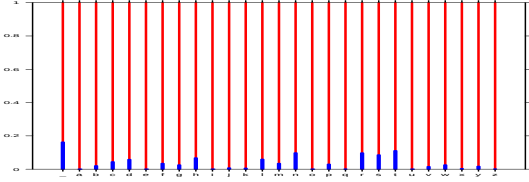


Fig. S.6: Plot of `estimate$hmm$emissionProbs[2,]`.

From Figures S.5 and S.6 we can infer that the vowels  $\{a, e, i, u, o\}$  are more frequently associated to the state ① 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 ①, and the combined probabilities of vowels from state ② is only  $6.2 \times 10^{-9} \%$ .

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 Figure S.7 that plots  $\eta \mapsto (M_{0,\eta}/M_{0,\text{"_"}})((M_{1,\text{"_"}} - M_{1,\eta})/M_{1,\text{"_"}})^2$  by combining the information available in the two rows of the emission matrix  $M$ , showing that “y” recovers its “semi-vowel” status.

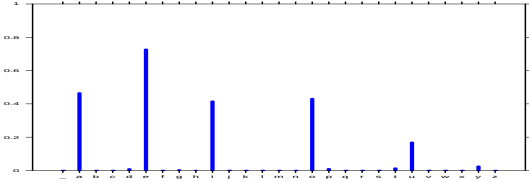


Fig. S.7: Plot of  $\eta \mapsto (M_{0,\eta}/M_{0,\text{"_"}})((M_{1,\text{"_"}} - M_{1,\eta})/M_{1,\text{"_"}})^2$ .

Note that the graphs of Figures S.5 and S.6 do *not* represent a frequency analysis. A frequency analysis of letters can be represented as the histogram of Figure S.8 using the command

```
1 text = readChar("brown.txt",nchars=10000)
2 data <- unlist(strsplit(gsub("[^a-z]", "_", tolower(text)), ""))
3 barplot(col = rainbow(30), table(data), cex.names=0.7)
```

with the following output:

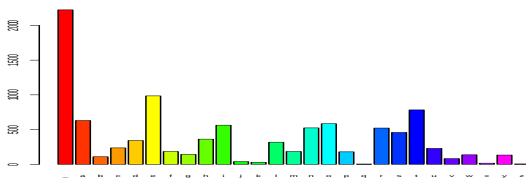


Fig. S.8: Frequency analysis of alphabet letters.

The command

```
1 estimate$hmm$transProbs[1,]
```

yields the transition probability estimate

$$\hat{P} = \begin{bmatrix} 0 & 1 \\ 0.1906356 & 0.8093644 \end{bmatrix}$$

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

$$\begin{bmatrix} 5/6 & 1/6 \\ 4/5 & 1/5 \end{bmatrix},$$

assuming the alphabet has *already* been partitioned. Such a matrix can be estimated on the whole text, from the following code:

```
1 x <- unlist(strsplit(gsub("[^a-z]", "", tolower(text)), ""))
2 y <- unlist(strsplit(gsub("[^a,e,i,o,u]", "2", tolower(x)), ""))
3 z <- as.numeric(noquote(unlist(strsplit(gsub("[a,e,i,o,u]", "1", y), ""))))
4 p <- matrix(nrow = 2, ncol = 2, 0)
5 for (t in 1:(length(z) - 1)) p[z[t], z[t + 1]] <- p[z[t], z[t + 1]] + 1
6 for (i in 1:2) p[i, ] <- p[i, ] / sum(p[i, ])
```

as

$$\begin{bmatrix} 0.1424749 & 0.8575251 \\ 0.5360502 & 0.4639498 \end{bmatrix},$$

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.9, and a special weight given to the letters  $h$  and  $t$  from state ① in Figure S.10.

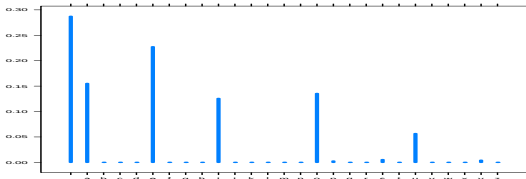


Fig. S.9: Plot of `estimate$hmm$emissionProbs[3,]`.

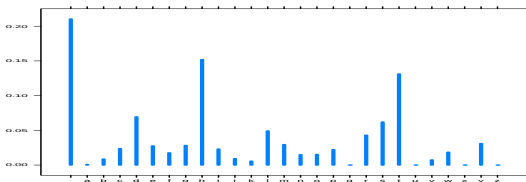


Fig. S.10: Plot of `estimate$hmm$emissionProbs[1,]`.

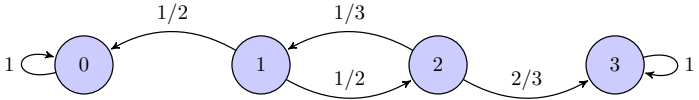
## Chapter 5 - First Step Analysis

### Exercise 5.1

a) The chain has the transition matrix

$$[P_{i,j}]_{i,j \in \mathbb{N}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/3 & 0 & 2/3 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

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

$$T_3 := \inf\{n \geq 0 : X_n = 3\}.$$

By first step analysis, we have

$$\begin{cases} \alpha = \mathbb{P}(T_3 < \infty \mid X_0 = 1) = \frac{1}{2} \times 0 + \frac{1}{2} \times \beta \\ \beta = \mathbb{P}(T_3 < \infty \mid X_0 = 2) = \frac{1}{3} \times \alpha + \frac{2}{3} \times 1, \end{cases}$$

hence

$$\alpha = \mathbb{P}(T_3 < \infty \mid X_0 = 1) = \frac{2}{5} \quad \text{and} \quad \beta = \mathbb{P}(T_3 < \infty \mid X_0 = 2) = \frac{4}{5}.$$

- c) By first step analysis, we have

$$\begin{cases} \mathbb{E}[T_{0,3} \mid X_0 = 1] = \frac{1}{2} (1 + \mathbb{E}[T_{0,3} \mid X_0 = 0]) + \frac{1}{2} (1 + \mathbb{E}[T_{0,3} \mid X_0 = 2]) \\ \quad = 1 + \frac{1}{2} \mathbb{E}[T_{0,3} \mid X_0 = 2], \\ \mathbb{E}[T_{0,3} \mid X_0 = 2] = 1 + \frac{1}{3} \mathbb{E}[T_{0,3} \mid X_0 = 1] + \frac{2}{3} (1 + \mathbb{E}[T_{0,3} \mid X_0 = 3]) \\ \quad = 1 + \frac{1}{3} \mathbb{E}[T_{0,3} \mid X_0 = 1], \end{cases}$$

*i.e.*

$$\begin{cases} \mathbb{E}[T_{0,3} \mid X_0 = 1] = \frac{3}{2} + \frac{1}{6} \mathbb{E}[T_{0,3} \mid X_0 = 1] \\ \mathbb{E}[T_{0,3} \mid X_0 = 2] = \frac{4}{3} + \frac{1}{6} \mathbb{E}[T_{0,3} \mid X_0 = 2], \end{cases}$$

hence

$$\begin{cases} \mathbb{E}[T_{0,3} \mid X_0 = 1] = \frac{9}{5}, \\ \mathbb{E}[T_{0,3} \mid X_0 = 2] = \frac{8}{5}. \end{cases}$$

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 state ①, we find

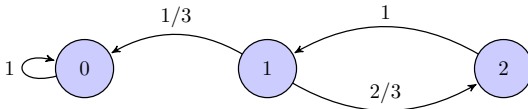
$$\mu_1(1) = 1 + \frac{b}{a} = 1 + \frac{0.8}{0.6} = \frac{7}{3}.$$

from *e.g.* (5.3.6). This mean time can also be recovered by path-wise analysis using the mean  $1/a$  of the geometric distribution on  $\{1, 2, 3, \dots\}$  with parameter  $a \in (0, 1]$ , as

$$(1 - b) \times 1 + b \times \left(1 + \frac{1}{a}\right) = 1 + \frac{b}{a} = \frac{7}{3}.$$

Exercise 5.3

a) The chain has the following graph:



Noting that state ① is absorbing, by first step analysis we have

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = \frac{1}{3}g_0(0) + \frac{2}{3}g_0(2) \\ g_0(2) = g_0(1), \end{cases}$$

which has for solution

$$g_0(0) = g_0(1) = g_0(2) = 1.$$

```

1  install.packages("igraph");install.packages("markovchain")
2  library("igraph");library(markovchain)
3  P<-matrix(c(1,0,0,1/3,0,2/3,0,1,0),nrow=3,byrow=TRUE);
   MC<-new("markovchain",transitionMatrix=P)
4  graph <- as(MC, "igraph")
5  plot(graph,vertex.size=50,edge.label=cex=2,edge.label=E(graph)$prob,edge.color='black',
   vertex.color='dodgerblue',vertex.label.cex=3)
6  hittingProbabilities(object = MC)
   1 2 3
8  1 1 0.0000000 0.0000000
   2 1 0.6666667 0.6666667
10 3 1 1.0000000 0.6666667
  
```

b) By first step analysis, we have

$$\begin{cases} h_0(0) = 0 \\ h_0(1) = 1 + \frac{1}{3}h_0(0) + \frac{2}{3}h_0(2) \\ h_0(2) = 1 + h_0(1), \end{cases}$$

which has for solution

$$h_0(0) = 0, \quad h_0(1) = 5, \quad h_0(2) = 6.$$

```
1 meanAbsorptionTime(object = MC)
2 5 6
```

#### Exercise 5.4

a) By first step analysis, we have

$$\begin{cases} p_2(0) = 0.2 \times p_2(1) + 0.8 \\ p_2(1) = 0.1 \times p_2(1) + 0.3 \times p_2(0) + 0.6 \times p_2(3) \\ p_2(2) = 0.5 \times p_2(0) + 0.5 \times p_2(3) \\ p_2(3) = 0, \end{cases}$$

hence

$$\begin{cases} p_2(0) = 0.2 \times p_2(1) + 0.8 \\ p_2(1) = \frac{p_2(0)}{3} \\ p_2(2) = \frac{p_2(0)}{2} \\ p_2(3) = 0, \end{cases}$$

and

$$\begin{cases} p_2(0) = \frac{6}{7} \\ p_2(1) = \frac{2}{7} \\ p_2(2) = \frac{3}{7} \\ p_2(3) = 0. \end{cases}$$

We note that

$$\mathbb{P}(T_2^r = \infty \mid X_0 = 2) = 1 - p_2(2) = \frac{4}{7} \geq \mathbb{P}(X_1 = 3 \mid X_0 = 2) = 0.5,$$

as seen in (6.3.6).

b) By first step analysis, we have

$$\begin{cases} p_1(0) = 0.2 + 0.8 \times p_1(2) \\ p_1(1) = 0.1 + 0.3 \times p_1(0) + 0.6 \times p_1(3) \\ p_1(2) = 0.5 \times p_1(0) + 0.5 \times p_1(3) \\ p_1(3) = 0, \end{cases}$$

hence

$$\begin{cases} p_1(0) = 0.2 + 0.8 \times p_1(2) = 0.2 + 0.4 \times p_1(0) \\ p_1(1) = 0.1 + 0.3 \times p_1(0) \\ p_1(2) = 0.5 \times p_1(0) \\ p_1(3) = 0, \end{cases}$$

and

$$\begin{cases} p_1(0) = \frac{1}{3} \\ p_1(1) = 0.2 \\ p_1(2) = \frac{1}{6} \\ p_1(3) = 0. \end{cases}$$

We note that

$$\mathbb{P}(T_1^r = \infty \mid X_0 = 1) = 1 - p_1(1) = 0.8 \geq \mathbb{P}(X_1 = 3 \mid X_0 = 1) = 0.6,$$

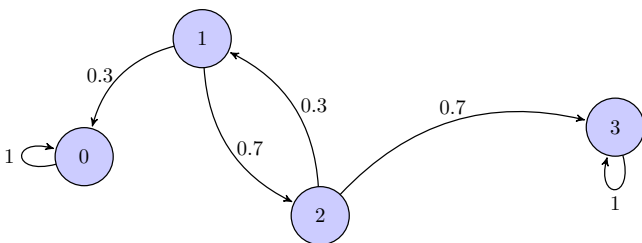
as seen in (6.3.7).

c) We have

$$\begin{cases} \mu_2(0) = 0.2 \times (1 + \mu_2(1)) + 0.8 \\ \mu_2(1) = 0.3 \times (1 + \mu_2(0)) + 0.1 \times (1 + \mu_2(1)) + 0.6 \times (1 + \mu_2(3)) \\ \mu_2(2) = 0.5 \times (1 + \mu_2(0)) + 0.5 \times (1 + \mu_2(3)) \\ \mu_2(3) = +\infty, \end{cases}$$

hence  $\mu_0(1) = \mu_2(1) = \mu_2(2) = \mu_3(2) = +\infty$ , as can be checked from  $\mathbb{P}(T_2^r = \infty \mid X_0 = k) = 1 - p_2(k) > 0$ ,  $k = 0, 1, 2, 3$ .

**Exercise 5.5** 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 ① and ③ are absorbing.  
 b) By first step analysis we find the equations

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = 0.3 + 0.7g_0(2) \\ g_0(2) = 0.3g_0(1) \\ g_0(3) = 0, \end{cases}$$

hence

$$g_0(1) = \frac{0.3}{1 - 0.7 \times 0.3} \quad \text{and} \quad g_0(2) = \frac{0.3 \times 0.3}{1 - 0.7 \times 0.3},$$

which is consistent with the answer

$$\begin{cases} g_0(1) = \frac{a}{1 - \beta c} \\ g_0(2) = \frac{a\beta}{1 - \beta c} \end{cases}$$

obtained from (5.1.11). Note that this chain is actually a particular case of the gambling process of Chapter 2 with  $S = 3$ , hence by (2.2.12) we can also write

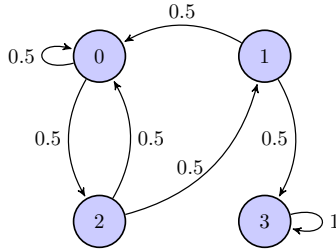
$$g_0(1) = \mathbb{P}(R_A \mid X_0 = 1) = \frac{(3/7) - (3/7)^3}{1 - (3/7)^3} = \frac{(7/3)^{3-1} - 1}{(7/3)^3 - 1} \simeq 0.3797.$$

- c) By first step analysis we find the equations

$$\begin{cases} h_1(0) = 1 + h_1(0) \\ h_1(1) = 0 \\ h_1(2) = 0.3(1 + h_1(1)) + 0.7(1 + h_1(3)) = 1 + 0.7h_1(3) \\ h_1(3) = 1 + h_1(3), \end{cases}$$

which admit the only solution  $h_1(1) = 0$  and  $h_1(0) = h_1(2) = h_1(2) = +\infty$ .

Exercise 5.6 We observe that state ③ is absorbing:



Let

$$h_3(k) := \mathbb{E}[T_3 \mid X_0 = k]$$

denote the mean (hitting) time needed to reach state ③ starting from state  $k = 0, 1, 2, 3$ . We have

$$\begin{cases} h_3(0) = 1 + \frac{1}{2}h_3(0) + \frac{1}{2}h_3(2) \\ h_3(1) = 1 + \frac{1}{2}h_3(0) \\ h_3(2) = 1 + \frac{1}{2}h_3(0) + \frac{1}{2}h_3(1) \\ h_3(3) = 0, \end{cases}$$

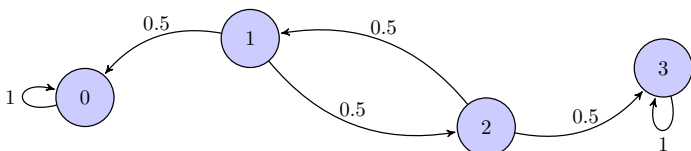
which yields

$$h_3(3) = 0, \quad h_3(1) = 8, \quad h_3(2) = 12, \quad h_3(0) = 14.$$

We check that  $h_3(3) < h_3(1) < h_3(2) < h_3(0)$ , as can be expected from the graph.

Exercise 5.7 This exercise has some similarities with the gambling problem of Chapter 2, with the maze problems of Section 5.3, and with Exercise 5.11 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 ① and ③ are absorbing, by first step analysis we have

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = \frac{1}{2} + \frac{1}{2}g_0(2) \\ g_0(2) = \frac{1}{2}g_0(1) \\ g_0(3) = 0, \end{cases}$$

which has for solution

$$g_0(0) = 1, \quad g_0(1) = \frac{2}{3}, \quad g_0(2) = \frac{1}{3}, \quad g_0(3) = 0.$$

- c) By first step analysis, we have

$$\begin{cases} h_{0,3}(0) = 0 \\ h_{0,3}(1) = 1 + \frac{1}{2}h_{0,3}(2) \\ h_{0,3}(2) = 1 + \frac{1}{2}h_{0,3}(1) \\ h_{0,3}(3) = 0, \end{cases}$$

which has for solution

$$h_{0,3}(0) = 0, \quad h_{0,3}(1) = 2, \quad h_{0,3}(2) = 2, \quad h_{0,3}(3) = 0.$$

Exercise 5.8 Using the law of total probability with respect to the partition  $\Omega = \bigcup_{m \geq 0} \{T_A = m\}$ , we have, provided that  $i_0 \in A$ ,

$$\begin{aligned} & \mathbb{P}(X_{T_A+n} = j \mid X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty) \\ &= \sum_{m \geq 0} \mathbb{P}(X_{T_A+n} = j \text{ and } T_A = m \mid X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty) \\ &= \sum_{m \geq 0} \frac{\mathbb{P}(X_{T_A+n} = j, T_A = m, X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty)}{\mathbb{P}(X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty)} \end{aligned}$$

$$\begin{aligned}
&= \sum_{m \geq 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_{T_A} \text{ and } T_A < +\infty)} \\
&= \sum_{m \geq 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_{T_A} \text{ and } T_A < +\infty)} \\
&= \sum_{m \geq 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)} \\
&\quad \times \frac{\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)}{\mathbb{P}(X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty)} \\
&= \sum_{m \geq 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) \\
&\quad \times \frac{\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)}{\mathbb{P}(X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty)} \\
&= \sum_{m \geq 0} \mathbb{P}(X_{n+m} = j \mid X_m = i_0 \text{ and } X_m \in A) \\
&\quad \times \frac{\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)}{\mathbb{P}(X_{T_A} = i_0, \dots, X_0 = i_{T_A} \text{ and } T_A < +\infty)} \\
&= \sum_{m \geq 0} \mathbb{P}(X_{n+m} = j \mid X_m = i_0) \\
&\quad \times \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)} \\
&= \mathbb{P}(X_n = j \mid X_0 = i_0) \sum_{m \geq 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)} \\
&= \mathbb{P}(X_n = j \mid X_0 = i_0) \sum_{m \geq 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 \geq 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),
\end{aligned}$$

for all  $n, m \geq 0$ ,  $j \in \mathbb{S}$ , and  $(i_k)_{k \in \mathbb{N}} \subset \mathbb{S}$ .

### Exercise 5.9

a) Letting  $f(k) := \mathbb{P}(T_0 < \infty \mid S_0 = k)$  we have the boundary condition  $f(0) = 1$  and by first step analysis we find that  $f(k)$  satisfies

$$f(k) = pf(k+1) + qf(k-1), \quad k \geq 1,$$

which is (2.2.6), and has the general solution

$$f(k) = C_1 + C_2 r^k, \quad k \in \mathbb{N}, \quad (\text{S.32})$$

where  $r = q/p$ , by (2.2.17), see also the command

`RSolve[f[k]=pf[k+1]+(1-p)f[k],f[k],k].`

- i) In case  $q \geq p$ ,  $f(k)$  would tend to (positive or negative) infinity if  $C_2 \neq 0$ , hence we should have  $C_2 = 0$ , and  $C_1 = f(0) = 1$ , showing that  $f(k) = 1$  for all  $k \in \mathbb{N}$ .
- ii) In case  $q < p$ , the probability of hitting  $\textcircled{0}$  in finite time after starting from state  $\textcircled{k}$  becomes 0 in the limit as  $k$  tends to infinity, *i.e.* we have

$$\lim_{k \rightarrow \infty} f(k) = \lim_{k \rightarrow \infty} \mathbb{P}(T_0 < \infty \mid S_0 = k) = 0, \quad k \in \mathbb{N},$$

which shows that  $C_1 = 0$ .

On the other hand, the condition  $f(0) = 1$  yields  $C_2 = 1$ , hence we find  $f(k) = (q/p)^k$  for all  $k \geq 0$ .

Alternatively, denoting by  $T_k$  the hitting time of state  $\textcircled{k}$  we could show the decomposition

$$\begin{aligned} f(k+1) &= \mathbb{P}(T_0 < \infty \mid S_0 = k+1) \\ &= \mathbb{P}(T_k < \infty \mid S_0 = k+1) \times \mathbb{P}(T_0 < \infty \mid S_0 = k) \\ &= \mathbb{P}(T_0 < \infty \mid S_0 = 1) \times \mathbb{P}(T_0 < \infty \mid S_0 = k) \\ &= f(1) \times f(k), \quad k \in \mathbb{N}, \end{aligned}$$

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

$$f(k) = f(0)(f(1))^k = (f(1))^k, \quad k \geq 0,$$

hence we have  $C_1 = 0$ ,  $C_2 = 1$ , and  $f(1) = r = q/p$ , which yields

$$f(k) = \left(\frac{q}{p}\right)^k, \quad k \in \mathbb{N},$$

and recovers (2.2.14) and (3.4.16), cf. also (S.25) and (S.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

$$\begin{aligned} h(k) &= p(1 + h(k+1)) + q(1 + h(k-1)) \\ &= 1 + ph(k+1) + qh(k-1), \quad k \geq 1, \end{aligned}$$

which is (2.3.7a) and has a general solution of the form

$$h(k) = C_1 + C_2 r^k + \frac{1}{q-p} k, \quad k \in \mathbb{N}, \quad (\text{S.33})$$

by (2.3.10). Next, we note that by the Markov property we should have the decomposition

$$\begin{aligned} h(k+1) &= \mathbb{E}[T_0 \mid S_0 = k+1] \\ &= \mathbb{E}[T_0 \mid S_0 = 1] + \mathbb{E}[T_0 \mid S_0 = k] \\ &= h(1) + h(k), \quad k \in \mathbb{N}, \end{aligned}$$

*i.e.* the mean time to go down from  $k+1$  to 0 should be the sum of the mean time needed to go down one step plus the mean time needed to go down  $k$  steps. This shows that

$$h(k) = h(0) + kh(1) = kh(1),$$

hence by (S.33) we have  $C_1 = C_2 = 0$ ,  $h(1) = 1/(q-p)$ , and

$$h(k) = \frac{k}{q-p}, \quad k \in \mathbb{N}.$$

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 5.10** First, we take a look at the complexity of the problem. Starting from (10) there are multiple ways to reach state (13) without reaching (11) or (12). For example:

$$13 = 3+4+1+5, \quad \text{or} \quad 13 = 1+6+3+3, \quad \text{or} \quad 13 = 1+1+2+1+3+1+4,$$

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

$$X_n = \sum_{k=1}^n \xi_k, \quad n \geq 0,$$

where  $(\xi_k)_{k \geq 1}$  is a family of independent random variables uniformly distributed over  $\{1, 2, 3, 4, 5, 6\}$ . The process  $(X_n)_{n \geq 0}$  is a Markov chain since given the history of  $(X_k)_{k=0,1,\dots,n}$  up to time  $n$ , the value

$$X_{n+1} = X_n + \xi_{n+1}$$

depends only on  $X_n$  and on  $\xi_{n+1}$  which is independent of  $X_0, X_1, \dots, X_n$ . The process  $(X_n)_{n \geq 0}$  is actually a random walk with independent increments  $\xi_1, \xi_2, \dots$

The chain  $(X_n)_{n \geq 0}$  has the transition matrix

$$[P_{i,j}]_{i,j \in \mathbb{N}} = \begin{bmatrix} 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Letting  $A := \{11, 12, 13, 14, 15, 16\}$ , we are looking at the probability

$$g_0 := \mathbb{P}(X_{T_A} = 13 \mid X_0 = 0)$$

of hitting the set  $A$  through state 13 after starting from state ①. More generally, letting

$$g_k := \mathbb{P}(X_{T_A} = 13 \mid X_0 = k)$$

denote the probability of hitting the set  $A$  *through state* ⑬ after starting from state  $k \geq 0$ , we have  $g_k = 0$  for all  $k \geq 14$ , and for  $k = 8, 9, 10, 11, 12, 13$  the probabilities  $g_k$  are easily computed as

$$g_{13} = \mathbb{P}(X_{T_A} = 13 \mid X_0 = 13) = 1,$$

$$g_{12} = \mathbb{P}(X_{T_A} = 13 \mid X_0 = 12) = 0,$$

$$g_{11} = \mathbb{P}(X_{T_A} = 13 \mid X_0 = 11) = 0,$$

$$g_{10} = \mathbb{P}(X_{T_A} = 13 \mid X_0 = 10) = \frac{1}{6},$$

$$g_9 = \mathbb{P}(X_{T_A} = 13 \mid X_0 = 9) = \frac{1}{6} + \frac{1}{6} \times \frac{1}{6} = \frac{7}{36},$$

$$g_8 = \mathbb{P}(X_{T_A} = 13 \mid X_0 = 8) = \frac{1}{6} + 2 \times \frac{1}{6} \times \frac{1}{6} + \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \frac{49}{216},$$

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

$$13 = 8+5, \quad 13 = 8+1+4, \quad 13 = 8+2+3 \quad 13 = 8+1+1+3.$$

Clearly, things become easily complicated for  $k \leq 7$ , and we will rely on first step analysis to continue the calculation. We have

$$g_k = \frac{1}{6} \sum_{i=1}^6 g_{k+i}, \quad k \in \mathbb{N},$$

*i.e.*

$$\begin{aligned} g_{10} &= \frac{1}{6} g_{13} \\ g_9 &= \frac{1}{6} (g_{10} + g_{13}) \\ g_8 &= \frac{1}{6} (g_9 + g_{10} + g_{13}) \\ g_7 &= \frac{1}{6} (g_8 + g_9 + g_{10} + g_{13}) \\ g_6 &= \frac{1}{6} (g_7 + g_8 + g_9 + g_{10}) \\ g_5 &= \frac{1}{6} (g_6 + g_7 + g_8 + g_9 + g_{10}) \\ g_4 &= \frac{1}{6} (g_5 + g_6 + g_7 + g_8 + g_9 + g_{10}) \\ g_3 &= \frac{1}{6} (g_4 + g_5 + g_6 + g_7 + g_8 + g_9) \\ g_2 &= \frac{1}{6} (g_3 + g_4 + g_5 + g_6 + g_7 + g_8) \\ g_1 &= \frac{1}{6} (g_2 + g_3 + g_4 + g_5 + g_6 + g_7) \\ g_0 &= \frac{1}{6} (g_1 + g_2 + g_3 + g_4 + g_5 + g_6). \end{aligned}$$

In order to solve this system of equations we rewrite it as

$$\left\{ \begin{array}{l} g_0 - g_1 = \frac{1}{6}(g_1 - g_7) \\ g_1 - g_2 = \frac{1}{6}(g_2 - g_8) \\ g_2 - g_3 = \frac{1}{6}(g_3 - g_9) \\ g_3 - g_4 = \frac{1}{6}(g_4 - g_{10}) \\ g_4 - g_5 = \frac{1}{6}g_5 \\ g_5 - g_6 = \frac{1}{6}g_6 \\ g_6 - g_7 = \frac{1}{6}(g_7 - 1) \\ g_7 - g_8 = \frac{1}{6}g_8 \\ g_8 - g_9 = \frac{1}{6}g_9 \\ g_9 - g_{10} = \frac{1}{6}g_{10} \\ g_{10} = \frac{1}{6}g_{13}, \end{array} \right. \quad \text{or} \quad \left\{ \begin{array}{l} g_0 = \frac{7^{10} - 7^6 \times 6^4 - 4 \times 7^3 \times 6^6}{6^{11}} \\ g_1 = \frac{7^9 - 7^5 \times 6^4 - 3 \times 7^2 \times 6^6}{6^{10}} \\ g_2 = \frac{7^8 - 7^4 \times 6^4 - 2 \times 7 \times 6^6}{6^9} \\ g_3 = \frac{7^7 - 7^3 \times 6^4 - 6^6}{6^8} \\ g_4 = \frac{7^6 - 7^2 \times 6^4}{6^7} \\ g_5 = \frac{7^5 - 7 \times 6^4}{6^6} \\ g_6 = \frac{7^4}{6^5} - \frac{1}{6} = \frac{7^4 - 6^4}{6^5} \\ g_7 = \frac{7^3}{6^4} \\ g_8 = \frac{7^2}{6^3} \\ g_9 = \frac{7}{6^2} \\ g_{10} = \frac{1}{6}, \end{array} \right.$$

*i.e.*

$$g_0 = \frac{7^{10} - 7^6 \times 6^4 - 4 \times 7^3 \times 6^6}{6^{11}} \simeq 0.181892636.$$

The total number of paths leading to 13 in the above fashion is 928. More generally, the numbers of such paths starting from  $k = 0, 1, \dots, 10$  can be computed by induction according to the following table:

| 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 2.1: Number of paths starting from  $k = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$ .

In Figure S.11 we enumerate the corresponding possible paths.

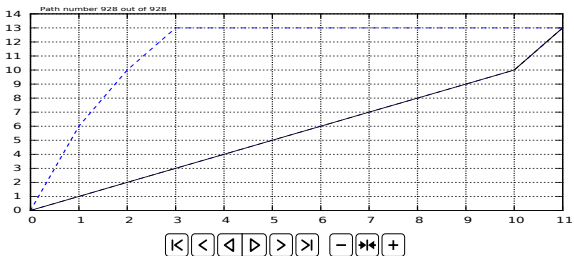


Fig. S.11: Graph of 928 paths leading from 0 to state 13.\*

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 2.2: Number of paths starting taking  $k$  rounds,  $k = 3, 4, 5, 6, 7, 8, 9, 10, 11$ .

This would also allow us to recover the value of  $g_0$  as

$$\begin{aligned}
 g_0 &= \frac{18}{6^3} + \frac{88}{6^4} + \frac{191}{6^5} + \frac{246}{6^6} + \frac{209}{6^7} + \frac{120}{6^8} + \frac{45}{6^9} + \frac{10}{6^{10}} + \frac{1}{6^{11}} \\
 &= \frac{65990113}{362797056} \simeq 0.181892636,
 \end{aligned}$$

see *e.g.* [here](#).

### Exercise 5.11

a) The transition matrix is given by

$$\begin{bmatrix}
 \times & \times & \times & \times & \times & \times \\
 q & 0 & p & 0 & 0 & 0 \\
 0 & q & 0 & p & 0 & 0 \\
 0 & 0 & q & 0 & p & 0 \\
 0 & 0 & 0 & q & 0 & p \\
 \times & \times & \times & \times & \times & \times
 \end{bmatrix}.$$

\* Animated figure (works in Acrobat Reader).

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

$$g(k) = q \times g(k-1) + p \times g(k+1), \quad 1 \leq k \leq 4. \quad (\text{S.34})$$

- b) When  $p = q = 1/2$  the probability that starting from state  $\textcircled{k}$  the fish finds the food before getting shocked is obtained by solving Equation (S.34) rewritten as

$$g(k) = \frac{1}{2} \times g(k-1) + \frac{1}{2} \times g(k+1), \quad 1 \leq k \leq 4.$$

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

$$g(k) = \frac{k}{5}, \quad k = 0, 1, \dots, 5.$$

#### Exercise 5.12

- a) The transition matrix is given by

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 & \dots \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 & \dots \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 & \dots \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Note that an arbitrary choice has been made for the first line (*i.e.* state  $\textcircled{0}$  is absorbing), however, other choices would not change the answer to the question.

- b) We have

$$h_0(m) = \sum_{k=0}^{m-1} \frac{1}{m} (1 + h_0(k)) = 1 + \frac{1}{m} \sum_{k=0}^{m-1} h_0(k), \quad m \geq 1,$$

and  $h_0(0) = 0$ ,  $h_0(1) = 1$ .

- c) We have

$$h_0(m) = 1 + \frac{1}{m} \sum_{k=0}^{m-1} h_0(k)$$

$$\begin{aligned}
&= 1 + \frac{1}{m}h_0(m-1) + \frac{m-1}{m} \sum_{k=0}^{m-2} \frac{h_0(k)}{m-1} \\
&= 1 + \frac{1}{m}h_0(m-1) - \frac{m-1}{m} + \frac{m-1}{m} \left( 1 + \sum_{k=0}^{m-2} \frac{h_0(k)}{m-1} \right) \\
&= 1 + \frac{1}{m}h_0(m-1) - \frac{m-1}{m} + \frac{m-1}{m}h_0(m-1) \\
&= 1 - \frac{m-1}{m} + \frac{1}{m}h_0(m-1) + \frac{m-1}{m}h_0(m-1) \\
&= h_0(m-1) + \frac{1}{m}, \quad m \geq 1,
\end{aligned}$$

hence

$$\begin{aligned}
h_0(m) &= h_0(m-1) + \frac{1}{m} \\
&= h_0(m-2) + \frac{1}{m-1} + \frac{1}{m} = \sum_{k=1}^m \frac{1}{k}, \quad m \geq 1.
\end{aligned}$$

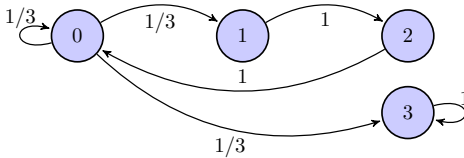
We can also note that from the Markov property, we have

$$h_0(m) = h_{m-1}(m) + h_0(m-1),$$

where the mean time  $h_{m-1}(m)$  from state  $\textcircled{m}$  to state  $\boxed{m-1}$  is equal to  $1/m$  since state  $\boxed{m-1}$  is always reached in one step from state  $\textcircled{m}$ , with probability  $1/m$ .

### Exercise 5.13

- a) Assuming that it takes one day per state transition, the graph of the chain can be drawn as follows:



In this model, state  $\textcircled{0}$  represents the tower, states  $\textcircled{1}$  and  $\textcircled{2}$  correspond to the path returning to the tower in three steps through exit A, the returning loop to state  $\textcircled{0}$  accounts for the time to return to the tower through exit B, and state  $\textcircled{3}$  represents the outside.

- b) We have

$$P = \begin{bmatrix} 1/3 & 1/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

- c) By first step analysis, the average time  $h_3(i)$  from state  $\textcircled{i}$  to state  $\textcircled{3}$  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

$$h_3(0) = \frac{1}{3}(1 + h_3(0)) + \frac{1}{3}(3 + h_3(0)) + \frac{1}{3},$$

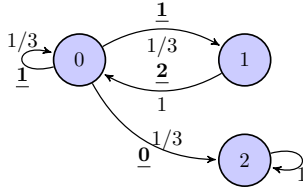
*i.e.*  $h_3(0) = 5$ , *i.e.* 4 times steps on average to reach the exit, plus one time step from the exit to the outside. The result  $h_3(0) = 4$  can also be obtained directly by writing the equations

$$\begin{cases} h_3(0) = \frac{1}{3}(1 + h_3(0)) + \frac{1}{3}(1 + h_3(1)) + \frac{1}{3} \times 0 \\ h_3(1) = 1 + h_3(2) \\ h_3(2) = 1 + h_3(0). \end{cases}$$

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  $h = \mathbf{1} + Ph$  *cannot* be directly used when the links have different weights starting from state  $\textcircled{1}$ .

#### 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  $\textcircled{i}$  to state  $\textcircled{2}$  solves

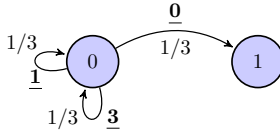
$$\begin{cases} h_2(0) = \frac{1}{3}(\underline{1} + h_2(0)) + \frac{1}{3}(\underline{1} + h_2(1)) + \frac{1}{3} \times \underline{0} \\ h_2(1) = \underline{2} + h_2(0) \\ h_2(2) = 0, \end{cases}$$

which yields

$$h_2(0) = \frac{1}{3}(\underline{1} + h_2(0)) + \frac{1}{3}(\underline{3} + h_2(0)) + \frac{1}{3} \times \underline{0},$$

*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:

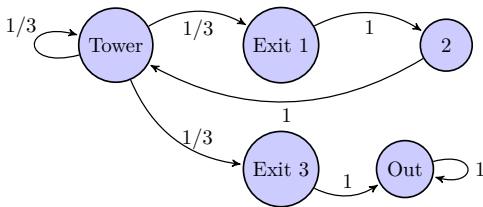


By first step analysis, the average time  $t_1(0)$  to travel from state  $\textcircled{0}$  to state  $\textcircled{1}$  is directly given by

$$t_1(0) = \frac{1}{3}(\underline{1} + t_1(0)) + \frac{1}{3}(\underline{3} + t_1(0)) + \frac{1}{3} \times \underline{0},$$

*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:

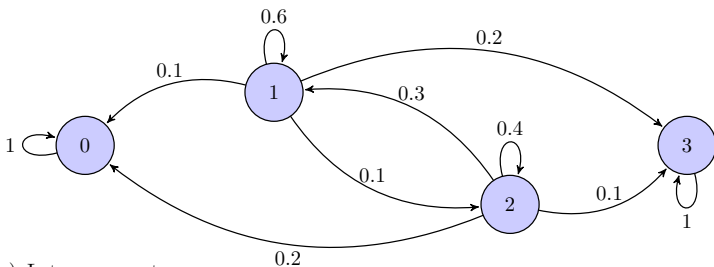


Exercise 5.14 The average time  $t$  spent inside the maze can be quickly computed by the following first step analysis using weighted links:

$$t = \frac{1}{2} \times (t + 3) + \frac{1}{6} \times 2 + \frac{2}{6} \times (t + 5),$$

which yields  $t = 21$ . We refer to Exercise 5.13 and its [solution](#) for a more detailed analysis of a related problem.

Exercise 5.15 The chain has the following graph:



a) Let us compute

$$g_0(k) = \mathbb{P}(T_0 < \infty \mid X_0 = k) = \mathbb{P}(X_{T_{\{0,3\}}} = 0 \mid X_0 = k), \quad k = 0, 1, 2, 3.$$

Since states ① and ③ are absorbing, by first step analysis we have

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = 0.1 \times g_0(0) + 0.6 \times g_0(1) + 0.1 \times g_0(2) + 0.2 \times g_0(3) \\ g_0(2) = 0.2 \times g_0(0) + 0.3 \times g_0(1) + 0.4 \times g_0(2) + 0.1 \times g_0(3) \\ g_0(3) = 0, \end{cases}$$

*i.e.*

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = 0.1 + 0.6 \times g_0(1) + 0.1 \times g_0(2) \\ g_0(2) = 0.2 + 0.3 \times g_0(1) + 0.4 \times g_0(2) \\ g_0(3) = 0, \end{cases}$$

which has for solution

$$g_0(0) = 1, \quad g_0(1) = \frac{8}{21}, \quad g_0(2) = \frac{11}{21}, \quad g_0(3) = 0,$$

cf. also (5.1.11).

b) Let

$$h_{0,3}(k) = \mathbb{E}[T_{\{0,3\}} \mid X_0 = k]$$

denote the mean time to reach the set  $A = \{0, 3\}$  starting from  $k = 0, 1, 2, 3$ . By first step analysis, we have

$$\begin{cases} h_{0,3}(0) = 0 \\ h_{0,3}(1) = 0.1 \times 1 + 0.6 \times (1 + h_{0,3}(1)) + 0.1 \times (1 + h_{0,3}(2)) + 0.2 \times (1 + h_{0,3}(3)) \\ h_{0,3}(2) = 0.2 \times 1 + 0.4 \times (1 + h_{0,3}(1)) + 0.3 \times (1 + h_{0,3}(2)) + 0.1 \times (1 + h_{0,3}(3)) \\ h_{0,3}(3) = 0, \end{cases}$$

i.e.

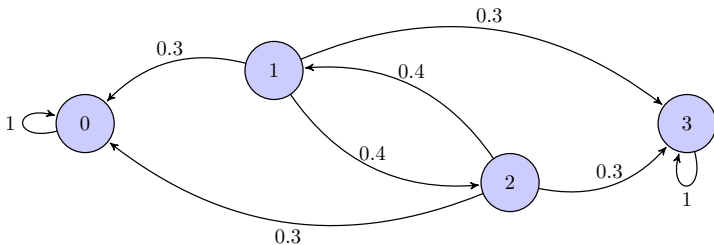
$$\begin{cases} h_{0,3}(0) = 0 \\ h_{0,3}(1) = 1 + 0.6 \times h_{0,3}(1) + 0.1 \times h_{0,3}(2) \\ h_{0,3}(2) = 1 + 0.4 \times h_{0,3}(1) + 0.3 \times h_{0,3}(2) \\ h_{0,3}(3) = 0, \end{cases}$$

which has for solution

$$h_{0,3}(0) = 0, \quad h_{0,3}(1) = \frac{10}{3}, \quad h_{0,3}(2) = \frac{10}{3}, \quad h_{0,3}(3) = 0.$$

Note that the relation  $h_{0,3}(1) = h_{0,3}(2)$  can be guessed from the symmetry of the problem.

Exercise 5.16 The chain has the following graph:



- a) The absorbing states are ① and ③.  
 b) By (5.1.12) we have  $g_0(1) = g_3(1) = 1/2$ . On the other hand, we clearly have  $g_1(0) = g_1(3) = 0$  and  $g_1(1) = 1$ , hence

$$g_1(2) = 0.3 \times g_1(0) + 0.4 \times g_1(1) + 0.3 \times g_1(3) = 0.4.$$

- c) We clearly have  $p_1(0) = p_1(3) = 0$ , and

$$\begin{cases} p_1(1) = 0.3 \times p_1(0) + 0.4 \times p_1(2) + 0.3 \times p_1(3) = 0.4 \times p_1(2) \\ p_1(2) = 0.3 \times p_1(0) + 0.4 + 0.3 \times p_1(3) = 0.4, \end{cases}$$

hence  $p_1(1) = 0.16$ .

- d) We have  $h_1(1) = 0$  by construction and  $h_1(0) = h_1(3) = +\infty$  because states ① and ③ are absorbing, and  $h_1(2) = +\infty$  because  $g_0(2) \geq 0.3 > 0$ . Regarding mean return times, we have  $\mu_1(0) = \mu_1(1) = \mu_1(2) = \mu_1(3) = +\infty$  because states ① and ② communicate while states ① and ③ are absorbing.

#### Exercise 5.17

- a) Let  $g_2(k) = \mathbf{P}(T_2 < \infty \mid X_0 = k)$ ,  $k = 0, 1, 2$ . We have  $g_2(2) = 1$  and by first step analysis,

$$\begin{cases} g_2(0) = \frac{1}{3}g_2(0) + \frac{1}{3}g_2(1) + \frac{1}{3} \\ g_2(1) = \frac{1}{4}g_2(0) + \frac{3}{4}g_2(1), \end{cases}$$

with  $g_2(2) = 1$  since state ② is absorbing. This shows that

$$\begin{cases} \frac{2}{3}g_2(0) = \frac{1}{3}g_2(1) + \frac{1}{3} \\ g_2(1) = g_2(0), \end{cases}$$

hence  $\mathbf{P}(T_2 < \infty \mid X_0 = 1) = g_2(0) = g_2(1) = g_2(2) = 1$ .

Next, let  $g_1(k) = \mathbf{P}(T_1^r < \infty \mid X_0 = k)$ ,  $k = 0, 1, 2$ . We have  $g_1(2) = 0$  and by first step analysis we have

$$\begin{cases} g_1(0) = \frac{1}{3}g_1(0) + \frac{1}{3} \\ g_1(1) = \frac{1}{4}g_1(0) + \frac{3}{4} \\ g_1(2) = 0, \end{cases}$$

hence

$$\begin{cases} g_1(0) = \frac{1}{2} \\ g_1(1) = \frac{7}{8} \\ g_1(2) = 0, \end{cases}$$

and  $\mathbf{P}(T_1^r < \infty \mid X_0 = 1) = g_1(1) = 7/8$ .

- 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

$$\begin{cases} h_2(0) = 1 + \frac{1}{3}h_2(0) + \frac{1}{3}h_2(1) \\ h_2(1) = 1 + \frac{1}{4}h_2(0) + \frac{3}{4}h_2(1) \\ h_2(2) = 0, \end{cases}$$

hence

$$\begin{cases} 2h_2(0) = 3 + h_2(1) \\ h_2(1) = 4 + h_2(0) \\ h_2(2) = 0, \end{cases}$$

hence

$$\begin{cases} h_2(0) = 7 \\ h_2(1) = 11 \\ h_2(2) = 0, \end{cases}$$

and  $\mathbb{E}[T_2^r < \infty \mid X_0 = 1] = h_2(1) = 11$ .

Exercise 5.18

- a) We have

$$\mathbb{E}[T^{(m)}] = mp^m + \sum_{k=0}^{m-1} p^k q(k+1 + \mathbb{E}[T^{(m)}]).$$

b) We find

$$\begin{aligned} \mathbb{E}[T^{(m)}] &= \frac{mp^m + q \sum_{k=0}^{m-1} p^k (k+1)}{1 - q \sum_{k=0}^{m-1} p^k} \\ &= \frac{mp^m + \frac{1 - (m+1)p^m + mp^{m+1}}{1-p}}{p^m} \\ &= \frac{m(1-p)p^m + 1 - (m+1)p^m + mp^{m+1}}{(1-p)p^m} \\ &= \frac{1/p^m - 1}{1-p} \\ &= \sum_{k=1}^m \frac{1}{p^k}. \end{aligned} \tag{S.35}$$

Alternative solution: We note the recurrence relation

$$\mathbb{E}[T^{(m)}] = \mathbb{E}[T^{(m-1)}] + p \times 1 + (1-p) \mathbb{E}[T^{(m)}], \quad m \geq 2,$$

or

$$\mathbb{E}[T^{(m)}] = \frac{\mathbb{E}[T^{(m-1)}] + p}{p}, \quad m \geq 2,$$

with  $\mathbb{E}[T^{(1)}] = 1/p$ , which also recovers (S.35).

Exercise 5.19 We have

$$\begin{aligned} V^*(k) &= \max_{\pi} \mathbb{E} \left[ \sum_{n \geq 0} \gamma^n R(X_n) \mid X_0 = k \right] \\ &= \max_{\pi} \mathbb{E} \left[ R(X_0) + \sum_{n \geq 1} \gamma^n R(X_n) \mid X_0 = k \right] \\ &= \max_{\pi} \mathbb{E} \left[ R(k) + \gamma \sum_{n \geq 1} \gamma^{n-1} R(X_n) \mid X_0 = k \right] \\ &= \max_{\pi} \left( R(k) + \mathbb{E} \left[ \gamma \sum_{n \geq 1} \gamma^{n-1} R(X_n) \mid X_0 = k \right] \right) \end{aligned}$$

$$\begin{aligned}
&= R(k) + \gamma \max_{\pi} \mathbb{E} \left[ \sum_{n \geq 1} \gamma^{n-1} R(X_n) \mid X_0 = k \right] \\
&= R(k) + \gamma \max_{a \in \mathcal{A}} \sum_{l \in \mathbb{S}} P_{k,l}^a \max_{\pi} \mathbb{E} \left[ \sum_{n \geq 1} \gamma^{n-1} R(X_n) \mid X_1 = l \right] \\
&= R(k) + \gamma \max_{a \in \mathcal{A}} \sum_{l \in \mathbb{S}} P_{k,l}^a \max_{\pi} \mathbb{E} \left[ \sum_{n \geq 0} \gamma^n R(X_{n+1}) \mid X_1 = l \right] \\
&= R(k) + \gamma \max_{a \in \mathcal{A}} \sum_{l \in \mathbb{S}} P_{k,l}^a \mathbb{E} \left[ \sum_{n \geq 0} \gamma^n R(X_n) \mid X_0 = l \right] \\
&= R(k) + \gamma \max_{a \in \mathcal{A}} \sum_{l \in \mathbb{S}} P_{k,l}^a V^*(l), \quad k \in \mathbb{S}.
\end{aligned}$$

The optimal policy  $\pi_k^* \in \mathbb{A}$  starting from state  $k \in \mathbb{S}$  is given by

$$\pi_k^* = \operatorname{argmax}_{a \in \mathbb{A}} \sum_{l \in \mathbb{S}} P_{k,l}^a V^*(l).$$

Problem 5.20

- a) The boundary conditions  $g(0)$  and  $g(N)$  are given by  $g(0) = 1$  and  $g(N) = 0$ .
- b) We have

$$\begin{aligned}
g(k) &= \mathbb{P}(T_0 < T_N \mid X_0 = k) \\
&= \sum_{l=0}^N \mathbb{P}(T_0 < T_N \mid X_1 = l) \mathbb{P}(X_1 = l \mid X_0 = k) \\
&= \sum_{l=0}^N \mathbb{P}(T_0 < T_N \mid X_1 = l) P_{k,l} \\
&= \sum_{l=0}^N \mathbb{P}(T_0 < T_N \mid X_0 = l) P_{k,l} \\
&= \sum_{l=0}^N g(l) P_{k,l}, \quad k = 0, 1, \dots, N.
\end{aligned}$$

- c) We find

$$[P_{i,j}]_{0 \leq i,j \leq 3} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ (2/3)^3 & (2/3)^2 & 2/3^2 & 1/3^3 \\ 1/3^3 & 2/3^2 & (2/3)^2 & (2/3)^3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

- d) Letting  $g(k) = 1 - k/N$ , we check that  $g(k)$  satisfies the boundary conditions  $g(0) = 1$  and  $g(N) = 0$ , and in addition we have

$$\begin{aligned} \sum_{l=0}^N g(l)P_{k,l} &= \sum_{l=0}^N \frac{N!}{(N-l)!l!} \left(\frac{k}{N}\right)^l \left(1 - \frac{k}{N}\right)^{N-l} \frac{N-l}{N} \\ &= \sum_{l=0}^{N-1} \frac{N!}{(N-l)!l!} \left(\frac{k}{N}\right)^l \left(1 - \frac{k}{N}\right)^{N-l} \frac{N-l}{N} \\ &= \sum_{l=0}^{N-1} \frac{(N-1)!}{(N-l-1)!l!} \left(\frac{k}{N}\right)^l \left(1 - \frac{k}{N}\right)^{N-l} \\ &= \left(1 - \frac{k}{N}\right) \sum_{l=0}^{N-1} \binom{N-1}{l} \left(\frac{k}{N}\right)^l \left(1 - \frac{k}{N}\right)^{N-1-l} \\ &= \left(1 - \frac{k}{N}\right) \left(\frac{k}{N} + 1 - \frac{k}{N}\right)^{N-1} \\ &= \frac{N-k}{N} \\ &= g(k), \quad k = 0, 1, \dots, N, \end{aligned}$$

which allows us to conclude by uniqueness of the solution given two boundary conditions, cf. Exercise 5.9 for cases of non-uniqueness under a single boundary condition.

- e) The boundary conditions  $h(0)$  and  $h(N)$  are given by  $h(0) = 0$  and  $h(N) = 0$  since the states  $\textcircled{0}$  and  $\textcircled{N}$  are absorbing.  
f) We have

$$\begin{aligned} h(k) &= \mathbb{E}[T_{0,N} \mid X_0 = k] \\ &= \sum_{l=0}^N (1 + \mathbb{E}[T_{0,N} \mid X_1 = l]) \mathbb{P}(X_1 = l \mid X_0 = k) \\ &= \sum_{l=0}^N (1 + \mathbb{E}[T_{0,N} \mid X_1 = l]) P_{k,l} = \sum_{l=0}^N P_{k,l} + \sum_{l=0}^N \mathbb{E}[T_{0,N} \mid X_1 = l] P_{k,l} \\ &= 1 + \sum_{l=1}^{N-1} \mathbb{E}[T_{0,N} \mid X_1 = l] P_{k,l} = 1 + \sum_{l=1}^{N-1} h(l) P_{k,l}, \end{aligned} \tag{S.36}$$

$k = 1, 2, \dots, N - 1$ .

g) In this case, the Equation (S.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

$$h(0) = 0, \quad h(1) = 3, \quad h(2) = 3, \quad h(3) = 0.$$

Problem 5.21

- a) Since we consider the time until we hit either 0 or  $N$ , we have  $h(0) = 0$  as well as  $h(N) = 0$ .  
b) We have

$$\begin{aligned} h(k) &= \mathbb{E} \left[ \sum_{i=0}^{\tau-1} X_i \mid X_0 = k \right] \\ &= \mathbb{E} [X_0 \mid X_0 = k] + \mathbb{E} \left[ \sum_{i=1}^{\tau-1} X_i \mid X_0 = k \right] \\ &= k + p \mathbb{E} \left[ \sum_{i=1}^{\tau-1} X_i \mid X_1 = k + 1 \right] + q \mathbb{E} \left[ \sum_{i=1}^{\tau-1} X_i \mid X_1 = k - 1 \right] \\ &= k + p \mathbb{E} \left[ \sum_{i=0}^{\tau-2} X_{i+1} \mid X_1 = k + 1 \right] + q \mathbb{E} \left[ \sum_{i=0}^{\tau-2} X_{i+1} \mid X_1 = k - 1 \right] \end{aligned} \quad (\text{S.37})$$

$$\begin{aligned} &= k + p \mathbb{E} \left[ \sum_{i=0}^{\tau-1} X_i \mid X_0 = k + 1 \right] + q \mathbb{E} \left[ \sum_{i=0}^{\tau-1} X_i \mid X_0 = k - 1 \right] \quad (\text{S.38}) \\ &= k + ph(k + 1) + qh(k - 1), \quad 1 \leq k \leq N - 1, \end{aligned}$$

where we used the fact that  $\tau - 1$  in (S.37) becomes  $\tau$  in (S.38).

When passing from

$$p \mathbb{E} \left[ \sum_{i=1}^{\tau-1} X_i \mid X_1 = k + 1 \right] + q \mathbb{E} \left[ \sum_{i=1}^{\tau-1} X_i \mid X_1 = k - 1 \right]$$

to

$$p \mathbb{E} \left[ \sum_{i=0}^{\tau-1} X_i \mid X_0 = k+1 \right] + q \mathbb{E} \left[ \sum_{i=0}^{\tau-1} X_i \mid X_0 = k-1 \right],$$

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,  $\tau$  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  $\tau$  (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

$$h(k) = -\frac{k^3}{3} + C_1 + C_2 k,$$

and the boundary conditions show that

$$\begin{cases} 0 = h(0) = C_1, \\ 0 = h(N) = -\frac{N^3}{3} + C_1 + C_2 N, \end{cases}$$

hence  $C_1 = 0$ ,  $C_2 = N^2/3$ , and

$$h(k) = -\frac{k^3}{3} + N^2 \frac{k}{3} = \frac{k}{3} (N^2 - k^2) = k(N-k) \frac{N+k}{3}, \quad k = 0, 1, \dots, N. \quad (\text{S.39})$$

- e) When  $N = 2$  we find  $h(1) = 1$  since starting from  $k = 1$  we can only move to state ① 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

$$\mathbb{E}[T_{0,N} \mid X_0 = 4] = 4(70 - 4) = 264 \text{ months} = 22 \text{ years}.$$

ii) By (S.39), we find

$$h(4) = \frac{4}{3}(70^2 - 4^2) = \$6,512\text{K} = \$6.512\text{M}.$$

iii) In that case we find

$$\$4\text{K} \times 264 = \$1056\text{K} = \$1.056\text{M}.$$

iv) It appears that starting a (potentially risky) business is more profitable on the average than keeping the same fixed initial income over an equivalent (average) period of time.

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

$$h(k) = \mathbb{E} \left[ \sum_{i=0}^{\tau-1} \beta^i X_i \mid X_0 = k \right]$$

with discount factor  $\beta \in (0, 1)$  in the case where  $p < q$ , with additional computational difficulties.

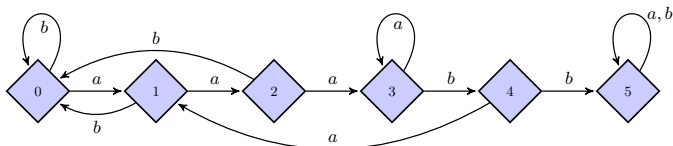
Note that the probability of hitting 0 first (*i.e.* ending with a bankruptcy) is given by

$$\mathbb{P}(X_{T_{0,N}} = 0 \mid X_0 = k) = 1 - \frac{k}{N} = \%94$$

when  $k = 4$  and  $N = 70$ .

**Problem 5.22** This problem is based on a simplified version of questions considered in [Gusev \(2014\)](#).

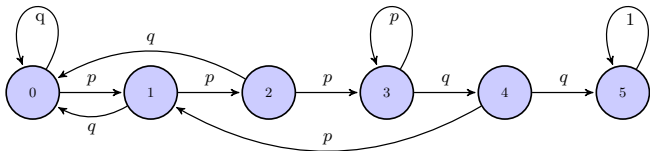
a) We find



The smallest integers are  $l = 3$ ,  $m = 2$  and the word is  $a^3b^2$ .

b) This language can be denoted by  $\Sigma^*a^3b^2\Sigma^*$ . 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  $Y_0, \dots, Y_{k-1}$ . 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

$$[P_{i,j}]_{0 \leq i,j \leq 5} = \begin{bmatrix} q & p & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & 0 \\ q & 0 & 0 & p & 0 & 0 \\ 0 & 0 & 0 & p & q & 0 \\ 0 & p & 0 & 0 & 0 & q \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

- d) Denoting by  $h_5(k)$  the time it takes to reach state  $\textcircled{5}$  starting from state  $k = 0, 1, 2, 3, 4, 5$ , we have the equations

$$\begin{cases} h_5(0) = 1 + qh_5(0) + ph_5(1) \\ h_5(1) = 1 + qh_5(0) + ph_5(2) \\ h_5(2) = 1 + qh_5(0) + ph_5(3) \\ h_5(3) = 1 + ph_5(3) + qh_5(4) \\ h_5(4) = 1 + ph_5(1) + qh_5(5) \\ h_5(5) = 0, \end{cases}$$

*i.e.*

$$\begin{cases} ph_5(0) = 1 + ph_5(1) \\ h_5(1) = 1 + qh_5(0) + ph_5(2) \\ h_5(2) = 1 + qh_5(0) + ph_5(3) \\ qh_5(3) = 1 + qh_5(4) = 1 + qph_5(0) \\ h_5(4) = 1 + ph_5(1) = ph_5(0) \\ h_5(5) = 0, \end{cases}$$

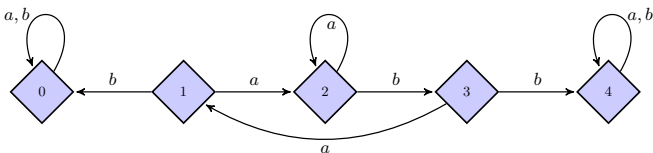
*i.e.*

$$\begin{cases} h_5(0) = \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + h_5(3) \\ h_5(1) = \frac{1}{p^2} + \frac{1}{p^3} + h_5(3) \\ h_5(2) = \frac{1}{p^3} + h_5(3) \\ h_5(3) = \frac{1}{q} + ph_5(0) \\ h_5(4) = ph_5(0) \\ h_5(5) = 0, \end{cases}$$

i.e.

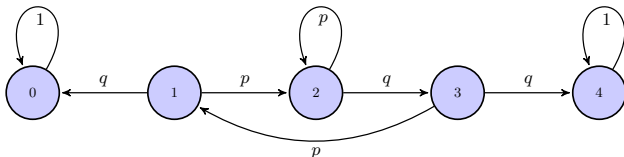
$$\begin{cases} h_5(0) = \frac{q(p^2 + p + 1) + p^3}{p^3q^2} = \frac{1}{p^3q^2} \\ h_5(1) = \frac{1}{p^3q^2} - \frac{1}{p} \\ h_5(2) = \frac{1}{p^3q^2} - \frac{1}{p} - \frac{1}{p^2} \\ h_5(3) = \frac{1}{q} + \frac{1}{p^2q^2} \\ h_5(4) = \frac{1}{p^2q^2} \\ h_5(5) = 0. \end{cases}$$

e) We find



f) The synchronizing 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

$$[P_{i,j}]_{0 \leq i,j \leq 4} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & 0 & p & q & 0 \\ 0 & p & 0 & 0 & q \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

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

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = qg_0(0) + pg_0(2) = q + pg_0(2) \\ g_0(2) = pg_0(2) + qg_0(3) \\ g_0(3) = pg_0(1) + qg_0(4) = pg_0(1) \\ g_0(4) = 0, \end{cases}$$

*i.e.*

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = q + pg_0(2) \\ g_0(2) = pg_0(2) + qg_0(3) = pg_0(2) + qp g_0(1) \\ g_0(3) = pg_0(1) \\ g_0(4) = 0, \end{cases}$$

*i.e.*

$$\begin{cases} g_0(0) = 1 \\ g_0(1) = q + p^2 g_0(1) \\ g_0(2) = pg_0(1) \\ g_0(3) = pg_0(1) \\ g_0(4) = 0, \end{cases}$$

*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

$$g_0(1) = 1 - pq \sum_{k \geq 0} p^{2k} = 1 - \frac{pq}{1-p^2} = 1 - \frac{p}{1+p} = \frac{1}{1+p}.$$

Now, starting from state ① one may move directly to state ① with probability  $q$ , in which case the first synchronized word is “b”, 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.

Problem 5.23

a) We have

$$\begin{aligned}
 V(k) &= \mathbb{E} \left[ \sum_{n \geq 0} R(X_n) \mid X_0 = k \right] \\
 &= \sum_{m \in \mathbb{S}} P_{k,m} \left( R(k) + \mathbb{E} \left[ \sum_{n \geq 1} R(X_n) \mid X_1 = m \right] \right) \\
 &= \sum_{m \in \mathbb{S}} P_{k,m} R(k) + \sum_{m \in \mathbb{S}} P_{k,m} \mathbb{E} \left[ \sum_{n \geq 1} R(X_n) \mid X_1 = m \right] \\
 &= R(k) \sum_{m \in \mathbb{S}} P_{k,m} + \sum_{m \in \mathbb{S}} P_{k,m} \mathbb{E} \left[ \sum_{n \geq 0} R(X_n) \mid X_0 = m \right] \\
 &= R(k) + \sum_{m \in \mathbb{S}} P_{k,m} V(m), \quad k \in \mathbb{S}.
 \end{aligned}$$

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

$$Q^*(3, \rightarrow) = -1 + \max(Q^*(3, \downarrow), Q^*(3, \rightarrow)),$$

which implies

$$Q^*(3, \rightarrow) < Q^*(3, \downarrow),$$

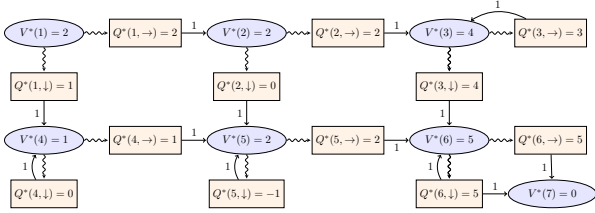
hence

$$Q^*(3, \rightarrow) = -1 + Q^*(3, \downarrow) = 3.$$

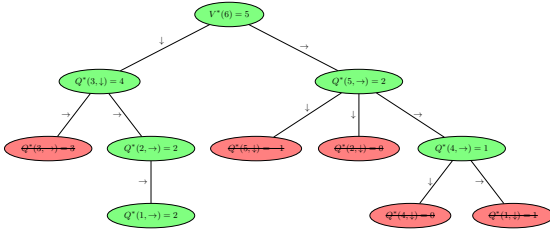
Similarly, we find

$$\begin{cases} Q^*(5, \downarrow) = -1, & Q^*(5, \rightarrow) = 2, \\ Q^*(2, \downarrow) = 0, & Q^*(2, \rightarrow) = 2, \\ Q^*(4, \downarrow) = 0, & Q^*(4, \rightarrow) = 1, \\ Q^*(1, \downarrow) = 1, & Q^*(1, \rightarrow) = 2. \end{cases}$$

The optimal action-value functional  $Q^*(k, a)$  can be summarized as follows:



This solution can also be obtained by backward optimization (or dynamic programming):



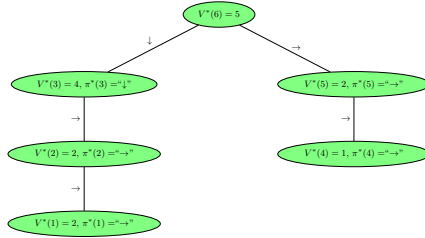
b2) At every state  $(k)$  we have

$$V^*(k) = \max(Q^*(k, \downarrow), Q^*(k, \rightarrow)),$$

hence

$$\begin{cases} V^*(7) = 0, \\ V^*(6) = 5, \\ V^*(3) = 4, \\ V^*(5) = 2, \\ V^*(2) = 2, \\ V^*(4) = 1, \\ V^*(1) = 2. \end{cases}$$

We obtain the following backward optimization tree:



b3) We find

$$\pi^* = (\pi^*(1), \pi^*(2), \pi^*(3), \pi^*(4), \pi^*(5), \pi^*(6), \pi^*(7)) = (\rightarrow, \rightarrow, \downarrow, \rightarrow, \rightarrow, \rightarrow, \uparrow),$$

which is consistent with the following MDPtoolbox output.

```

1  install.packages("MDPtoolbox")
2  library(MDPtoolbox)
3  P <- array(0, c(7, 7, 2))
4  P[,1] <- matrix(c(0,0,0,1,0,0,0,
5                    0,0,0,0,1,0,0,
6                    0,0,0,0,0,1,0,
7                    0,0,0,1,0,0,0,
8                    0,0,0,0,1,0,0,
9                    0,0,0,0,0,0,1,
10                   0,0,0,0,0,0,1), nrow=7, ncol=7, byrow=TRUE)
11
12 P[,2] <- matrix(c(0,1,0,0,0,0,0,
13                   0,0,1,0,0,0,0,
14                   0,0,1,0,0,0,0,
15                   0,0,0,0,1,0,0,
16                   0,0,0,0,0,1,0,
17                   0,0,0,0,0,0,1,
18                   0,0,0,0,0,0,1), nrow=7, ncol=7, byrow=TRUE)
19
20 R <- array(0, c(7, 2))
21 R[,1] <- matrix(c(0, -2, -1, -1, -3, 5, 0), nrow=1, ncol=7, byrow=TRUE)
22 R[,2] <- R[,1]
23 mdp_check(P, R)
24 mdp_value_iteration(P,R,discount=1)
25 $V
26 [1] 2 2 4 1 2 5 0
27 $policy
28 [1] 2 2 1 2 2 1 1

```

c) c1) We have

$$Q^*(k, a) := \max_{\pi : \pi(k)=a} \mathbb{E} \left[ \sum_{n \geq 0} R(X_n) \mid X_0 = k \right]$$

$$\begin{aligned}
&= \max_{\pi : \pi(k)=a} \mathbb{E} \left[ R(X_0) + \sum_{n \geq 1} R(X_n) \mid X_0 = k \right] \\
&= \max_{\pi : \pi(k)=a} \mathbb{E} \left[ R(k) + \sum_{n \geq 1} R(X_n) \mid X_0 = k \right] \\
&= \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} \left( R(k) + \max_{\pi} \mathbb{E} \left[ \sum_{n \geq 1} R(X_n) \mid X_1 = l \right] \right) \\
&= R(k) + \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} \max_{\pi} \mathbb{E} \left[ \sum_{n \geq 0} R(X_{n+1}) \mid X_1 = l \right] \\
&= R(k) + \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} \max_{\pi} \mathbb{E} \left[ \sum_{n \geq 0} R(X_n) \mid X_0 = l \right] \\
&= R(k) + \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} V^*(l), \quad k \in \mathbb{S}.
\end{aligned}$$

Similarly, the value function  $Q^\pi(k, a)$  for a given policy  $\pi$  by setting the first action at state  $(k)$  to “ $a$ ” satisfies the equation

$$Q^\pi(k, a) = R(k) + \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} V^\pi(l), \quad k \in \mathbb{S}. \quad (\text{S.40})$$

c2) The optimal policy  $\pi_k^* \in \mathbb{A}$  at state  $k \in \mathbb{S}$  is given by

$$\pi_k^* = \operatorname{argmax}_{a \in \mathbb{A}} Q^*(k, a) = \operatorname{argmax}_{a \in \mathbb{A}} \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} V^*(l).$$

c3) We have

$$\begin{aligned}
V^*(k) &= \max_{\pi} \mathbb{E} \left[ \sum_{n \geq 0} R(X_n) \mid X_0 = k \right] \\
&= \max_{a \in \mathcal{A}} \max_{\pi : \pi(k)=a} \mathbb{E} \left[ \sum_{n \geq 0} R(X_n) \mid X_0 = k \right] \\
&= \max_{a \in \mathcal{A}} Q^*(k, a) \\
&= \max_{a \in \mathcal{A}} \left( R(k) + \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} V^*(l) \right) \\
&= R(k) + \max_{a \in \mathcal{A}} \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} V^*(l), \quad k \in \mathbb{S}.
\end{aligned}$$

This also implies

$$\begin{aligned}
V^*(k) &= R(k) + \max_{a \in \mathcal{A}} \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} V^*(l) \\
&= \max_{a \in \mathcal{A}} \left( R(k) + \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} V^*(l) \right) \\
&= \max_{a \in \mathcal{A}} Q^*(k, a), \quad k \in \mathbb{S}.
\end{aligned} \tag{S.41}$$

Similarly, the value function  $V^\pi(k)$  for a given policy satisfies the equation

$$V^\pi(k) = R(k) + \sum_{l \in \mathbb{S}} P_{k,l}^{(\pi(k))} V^\pi(l) \quad k \in \mathbb{S}, \tag{S.42}$$

and the Bellman equation for  $Q^*(k, a)$  can also be rewritten as

$$Q^*(k, a) = R(k) + \sum_{l \in \mathbb{S}} P_{k,l}^{(a)} \max_{b \in \mathcal{A}} Q^*(k, b), \quad k \in \mathbb{S}.$$

In order to solve the MDP problem, starting from an arbitrary initial policy choice  $\pi$ , one can:

- i) Compute  $V^\pi$  by applying (S.42) iteratively.
  - ii) Deduce the value of  $Q^\pi(k, a)$  for every state  $\widehat{k}$  and action “ $a$ ” from (S.40).
  - iii) For every state  $\widehat{k}$  and action “ $a$ ”, based on (S.41), choose to update  $\pi$  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

$$\begin{cases} Q^*(7, \downarrow) = 0, & Q^*(7, \rightarrow) = 0, \\ Q^*(6, \downarrow) = 5, & Q^*(6, \rightarrow) = 5, \\ Q^*(3, \downarrow) = 4, & Q^*(3, \rightarrow) = 3, \\ Q^*(5, \downarrow) = -1, & Q^*(5, \rightarrow) = 2. \end{cases}$$

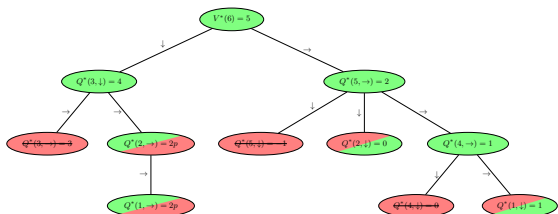
We also have  $Q^*(2, \downarrow) = 0$  and

$$\begin{aligned}
Q^*(2, \rightarrow) &= -2 + p \max(Q^*(3, \downarrow), Q^*(3, \rightarrow)) + q \max(Q^*(5, \rightarrow), Q^*(5, \rightarrow)) \\
&= -2 + pQ^*(3, \downarrow) + qQ^*(5, \rightarrow) \\
&= -2 + 4p + 2q = 2p,
\end{aligned}$$

and

$$Q^*(4, \downarrow) = 0, \quad Q^*(4, \rightarrow) = 1, \quad Q^*(1, \downarrow) = 1, \quad Q^*(1, \rightarrow) = Q^*(2, \rightarrow) = 2p.$$

In other words, we have the following backward optimization (or dynamic programming) graph:



d2) At every state  $\textcircled{k}$  we have

$$V^*(k) = \max(Q^*(k, \downarrow), Q^*(k, \rightarrow)),$$

hence

$$\begin{cases} V^*(7) = 0, \\ V^*(6) = 5, \\ V^*(5) = 2, \\ V^*(4) = 1, \\ V^*(3) = 4, \\ V^*(2) = 2p, \\ V^*(1) = \max(2p, 1). \end{cases}$$

d3) When  $p = 0$ , we find

$$\pi^* = (\pi^*(1), \pi^*(2), \pi^*(3), \pi^*(4), \pi^*(5), \pi^*(6), \pi^*(7)) = (\downarrow, \uparrow, \downarrow, \rightarrow, \rightarrow, \uparrow, \uparrow)$$

which is consistent with the MDPtoolbox output:

```

1 $V
2 [1] 1 0 4 1 2 5 0
3 $policy
4 [1] 1 1 1 2 2 1 1

```

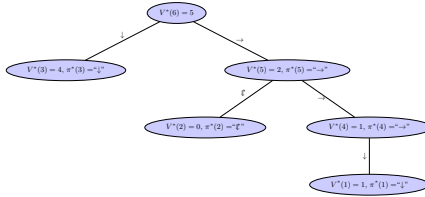


Fig. S.12: Optimal value function with  $p = 0$ .

When  $0 < p < 1/2$  we obtain

$$\pi^* = (\pi^*(1), \pi^*(2), \pi^*(3), \pi^*(4), \pi^*(5), \pi^*(6), \pi^*(7)) = (\downarrow, \rightarrow, \downarrow, \rightarrow, \rightarrow, \uparrow, \uparrow)$$

which is consistent with the MDPtoolbox output, here with  $p = 0.25$ :

```

1 $V
2 [1] 1.0 0.5 4.0 1.0 2.0 5.0 0.0
3 $policy
4 [1] 1 2 1 2 2 1 1

```

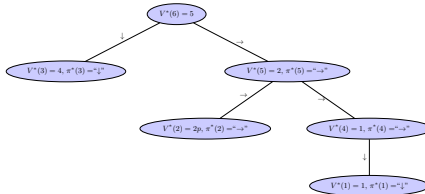


Fig. S.13: Optimal value function with  $0 < p < 1/2$ .

When  $p = 1/2$  we find

$$\pi^* = (\pi^*(1), \pi^*(2), \pi^*(3), \pi^*(4), \pi^*(5), \pi^*(6), \pi^*(7)) = (\uparrow, \rightarrow, \downarrow, \rightarrow, \rightarrow, \uparrow, \uparrow)$$

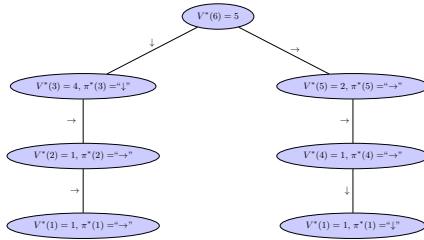


Fig. S.14: Optimal value function with  $p = 1/2$ .

which is consistent with the MDPtoolbox output, with  $p = 0.5$ :

|   |                   |
|---|-------------------|
| 1 | \$V               |
| 2 | [1] 1 1 4 1 2 5 0 |
|   | \$policy          |
| 4 | [1] 1 2 1 2 2 1 1 |

When  $1/2 < p \leq 1$ , we obtain

$$\pi^* = (\pi^*(1), \pi^*(1), \pi_3^*, \pi_4^*, \pi_5^*, \pi_6^*, \pi_7^*) = (\rightarrow, \rightarrow, \downarrow, \rightarrow, \rightarrow, \uparrow, \uparrow).$$

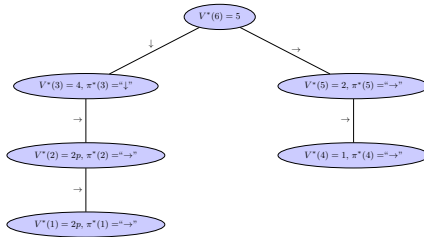


Fig. S.15: Optimal value function with  $1/2 < p \leq 1$ .

which is also consistent with the following MDPtoolbox output, here with  $p = 0.75$ :

```

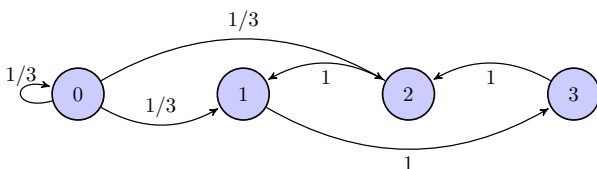
1 library(MDPtoolbox)
2 P <- array(0, c(7, 7, 2));p=0.75;q=1-p
3 P[,1] <- matrix(c(0,0,0,1,0,0,0,
4                   0,0,0,0,1,0,0,
5                   0,0,0,0,0,1,0,
6                   0,0,0,1,0,0,0,
7                   0,0,0,0,1,0,0,
8                   0,0,0,0,0,0,1,
9                   0,0,0,0,0,0,1), nrow=7, ncol=7, byrow=TRUE)
10 P[,2] <- matrix(c(0,1,0,0,0,0,0,
11                  0,0,p,q,0,0,0,
12                  0,0,1,0,0,0,0,
13                  0,0,0,0,1,0,0,
14                  0,0,0,0,0,1,0,
15                  0,0,0,0,0,0,1,
16                  0,0,0,0,0,0,1), nrow=7, ncol=7, byrow=TRUE)
17 R <- array(0, c(7, 2))
18 R[,1] <- matrix(c(0, -2, -1, -1, -3, 5, 0), nrow=1, ncol=7, byrow=TRUE);R[,2]
19   <- R[,1]
20 mdp_check(P, R);mdp_value_iteration(P,R,discount=1)
21 $V
22 [1] 1.5 1.5 4.0 1.0 2.0 5.0 0.0
23 $policy
24 [1] 2 2 1 2 2 1 1

```

## Chapter 6 - Classification of States

### Exercise 6.1

a) The graph of the chain is



This chain is reducible because its state space can be partitioned into two communicating classes as  $\mathbb{S} = \{0\} \cup \{1, 2, 3\}$ .

- b) State ① has period 1 and states ①, ②, ③ have period 3.  
c) We have

$$p_{0,0} = \mathbb{P}(T_0 < \infty \mid X_0 = 0) = \mathbb{P}(T_0 = 1 \mid X_0 = 0) = \frac{1}{3},$$

and

$$\mathbb{P}(T_0 = \infty \mid X_0 = 0) = 1 - \mathbb{P}(T_0 < \infty \mid X_0 = 0) = \frac{2}{3}.$$

We also have

$$\begin{aligned} \mathbb{P}(R_0 < \infty \mid X_0 = 0) &= \mathbb{P}(T_0 = \infty \mid X_0 = 0) \sum_{n \geq 1} (\mathbb{P}(T_0 < \infty \mid X_0 = 0))^n \\ &= \frac{2}{3} \sum_{n \geq 1} \left(\frac{1}{3}\right)^n = 1. \end{aligned}$$

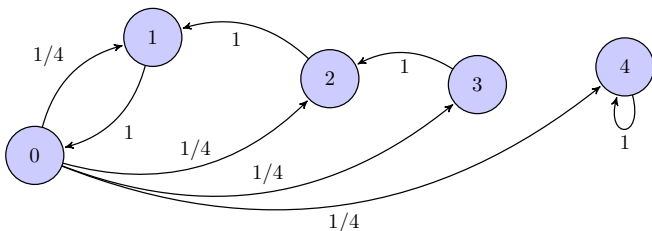
- d) There are no absorbing states, state ① is transient, and states ①, ②, ③ are recurrent by Corollary 6.7. State ① is transient since  $\mathbb{P}(R_0 < \infty \mid X_0 = 0) = 1$ , as expected from (5.4.4).

### Exercise 6.2

- The chain is reducible and its communicating classes are  $\{0, 1\}$  and  $\{2\}$ .
- States ① and ② are transient and the state ③ is (positive) recurrent and absorbing.
- All states have period 1.

### Exercise 6.3

- a) The chain has the following graph



- All states ①, ②, ③ and ④ have period 1, which can be obtained as the greatest common divisor (GCD) of  $\{2, 3\}$  for states ①, ②, ③ and  $\{4, 6, 7\}$  for state ④. The chain is aperiodic.
- State ④ is absorbing (and therefore recurrent), state ① is transient because

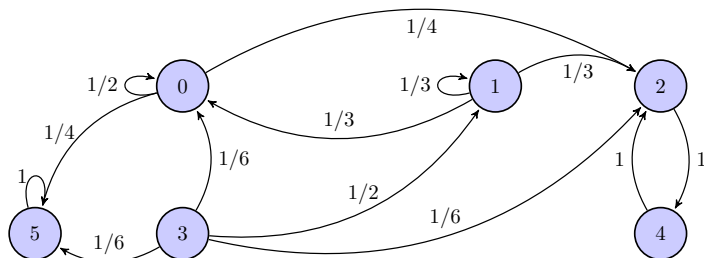
$$\mathbb{P}(T_0^r = \infty \mid X_0 = 0) \geq \frac{1}{4} > 0,$$

and the remaining states ①, ②, ③ are also transient because they communicate with the transient state ①, cf. Corollary 6.7. By pathwise or first step analysis, we can actually check that

$$\begin{aligned}
& \mathbb{P}(T_0^r < \infty | X_0 = 0) \\
&= \frac{1}{4} (\mathbb{P}(T_0^r < \infty | X_0 = 1) + \mathbb{P}(T_0^r < \infty | X_0 = 2) + \mathbb{P}(T_0^r < \infty | X_0 = 3)) \\
&= \frac{3}{4}.
\end{aligned}$$

- d) The Markov chain is reducible because its state space  $\mathbb{S} = \{0, 1, 2, 3, 4\}$  can be partitioned into two communicating classes  $\{0, 1, 2, 3\}$  and  $\{4\}$ .

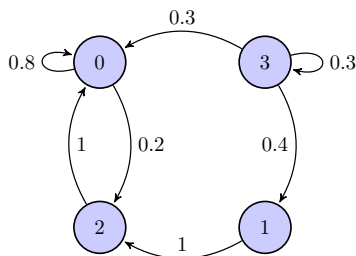
Exercise 6.4 The graph of the chain is



- a) The chain is reducible and its communicating classes are  $\{0\}$ ,  $\{1\}$ ,  $\{3\}$ ,  $\{5\}$ , and  $\{2, 4\}$ .  
b) States  $\textcircled{0}$ ,  $\textcircled{1}$ ,  $\textcircled{3}$  are transient and states  $\textcircled{2}$ ,  $\textcircled{4}$ ,  $\textcircled{5}$  are recurrent.  
c) State  $\textcircled{3}$  has period 0, states  $\textcircled{2}$  and  $\textcircled{4}$  have period 2, and states  $\textcircled{0}$ ,  $\textcircled{1}$ ,  $\textcircled{5}$  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 ①, ②, ③ have period 1, and state ④ has period 0. States ① and ③ are transient, states ② and ④ are recurrent by Theorem 6.11 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) States ①, ②, ③, ④ and ⑤ are transient.  
 c) State ⑥ is recurrent.  
 d) State ⑥ is positive recurrent.  
 e) State ⑥ has period 1, states ① and ② have period 2, and states ③, ④ and ⑤ have period 3.

#### Exercise 6.7

- a) This question refers to the Ehrenfest chain, cf. Example (iv) page 112. By (4.3.2) and (4.3.3) the transition matrix of this chain takes the form

$$\begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 1/N & 0 & (N-1)/N & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 2/N & 0 & (N-2)/N & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 3/N & 0 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \vdots & 0 & 3/N & 0 & 0 \\ 0 & 0 & \cdots & 0 & (N-2)/N & 0 & 2/N & 0 \\ 0 & 0 & \cdots & 0 & 0 & (N-1)/N & 0 & 1/N \\ 0 & 0 & \cdots & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

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 ① is (positive) recurrent as it is absorbing, while every state ② is transient when  $k \geq 1$  since  $\mathbb{P}(T_k^r < \infty \mid X_0 = k) = 0$ ,  $k \geq 1$ . State ① is aperiodic, and all other states have period 0.

b) For the success runs chain of Exercise 4.10-(b), we have

$$\mathbb{P}(T_0^r < \infty \mid X_0) = q \sum_{n \geq 0} p^n = \frac{q}{1-p} = 1,$$

hence state ① is recurrent. Since the chain is irreducible, it is recurrent by Corollary 6.7. State ① is also positive recurrent as we have

$$\mathbb{E}[T_0^r \mid X_0] = q \sum_{n \geq 1} np^{n-1} = \frac{q}{(1-p)^2} = \frac{1}{q}.$$

Next, by first step analysis we have

$$\begin{aligned} \mathbb{E}[T_k^r \mid X_0 = 0] &= \sum_{i=1}^k \mathbb{E}[T_i^r \mid X_0 = i-1] \\ &= \sum_{i=1}^k (q(1 + \mathbb{E}[T_{i-1}^r \mid X_0 = 0]) + p) \\ &= \sum_{i=1}^k (1 + q \mathbb{E}[T_{i-1}^r \mid X_0 = 0]) \\ &= k + q \sum_{i=1}^k \mathbb{E}[T_{i-1}^r \mid X_0 = 0], \quad k \geq 1, \end{aligned}$$

which shows by induction that  $\mathbb{E}[T_k^r \mid X_0 = 0] < \infty$ ,  $k \geq 0$ . On the other hand, we have

$$\begin{aligned} \mathbb{E}[T_k^r \mid X_0 = k] &= 1 + q \mathbb{E}[T_k^r \mid X_0 = 0] + p \mathbb{E}[T_k^r \mid X_0 = k+1] \\ &= 1 + q \mathbb{E}[T_k^r \mid X_0 = 0] + p \mathbb{E}[T_k^r \mid X_0 = k], \quad k \geq 0, \end{aligned}$$

since the starting point does not matter as long as we start from  $k \geq 1$ , hence

$$\mathbb{E}[T_k^r \mid X_0 = k] = \frac{1}{q} + \mathbb{E}[T_k^r \mid X_0 = 0], \quad k \geq 1,$$

which shows that  $\mathbb{E}[T_k^r \mid X_k = 0] < \infty$ ,  $k \geq 0$ , therefore the success runs Markov chain is also positive recurrent.

c) The chain is aperiodic because it is recurrent and state ① has a returning loop.

**Problem 6.9** By the same argument as in Exercise 3.2, by first step analysis we have

$$\begin{aligned}
\mathbb{E}[T_0^r \mid X_0 = 1] &= \frac{1}{\alpha + 1} \times 1 + \frac{\alpha}{\alpha + 1} \times (1 + \mathbb{E}[T_0^r \mid X_0 = 2]) \\
&= \frac{1}{\alpha + 1} \times 1 + \frac{\alpha}{\alpha + 1} \times (1 + 2 \mathbb{E}[T_0^r \mid X_0 = 1]) \\
&= 1 + \frac{2\alpha}{\alpha + 1} \mathbb{E}[T_0^r \mid X_0 = 1],
\end{aligned}$$

hence

$$\mathbb{E}[T_0^r \mid X_0 = 1] = \frac{1 + \alpha}{1 - \alpha}, \quad \alpha < 1.$$

From the spatial homogeneity of the chain, we deduce that

$$\begin{aligned}
\mathbb{E}[T_0^r \mid X_0 = k] &= \sum_{i=1}^k \mathbb{E}[T_{i-1}^r \mid X_0 = i] \\
&= k \mathbb{E}[T_0^r \mid X_0 = 1] \\
&= k \frac{1 + \alpha}{1 - \alpha}, \quad k \geq 1, \quad \alpha < 1,
\end{aligned}$$

and

$$\mathbb{E}[T_0^r \mid X_0 = 0] = 1 + \mathbb{E}[T_0^r \mid X_0 = 1] = 1 + \frac{1 + \alpha}{1 - \alpha} = \frac{2}{1 - \alpha},$$

hence state ① is positive recurrent if  $\alpha < 1$ . On the other hand, we have

$$\begin{aligned}
\mathbb{E}[T_1^r \mid X_0 = 1] &= \frac{1}{\alpha + 1} \times 2 + \frac{\alpha}{\alpha + 1} \times (1 + \mathbb{E}[T_1^r \mid X_0 = 2]) \\
&= 1 + \frac{1}{\alpha + 1} + \frac{\alpha}{\alpha + 1} \mathbb{E}[T_0^r \mid X_0 = 1] \\
&= 1 + \frac{1}{\alpha + 1} + \frac{\alpha}{1 - \alpha} \\
&= \frac{2}{1 - \alpha^2}.
\end{aligned}$$

Similarly to the above, we have  $\mathbb{E}[T_1^r \mid X_0 = 0] = 1$  and, for  $k \geq 1$ ,

$$\begin{aligned}
\mathbb{E}[T_{k+1}^r \mid X_0 = k] &= \frac{\alpha}{\alpha + 1} \times 1 + \frac{1}{\alpha + 1} \times (1 + \mathbb{E}[T_{k+1}^r \mid X_0 = k - 1]) \\
&= 1 + \frac{1}{\alpha + 1} \times (\mathbb{E}[T_k^r \mid X_0 = k - 1] + \mathbb{E}[T_{k+1}^r \mid X_0 = k]),
\end{aligned}$$

hence

$$\mathbb{E}[T_{k+1}^r \mid X_0 = k] = \frac{\alpha + 1 + \mathbb{E}[T_k^r \mid X_0 = k - 1]}{\alpha}, \quad k \geq 1,$$

which can be solved as

$$\mathbb{E}[T_{k+1}^r \mid X_0 = k] = \frac{2\alpha^{-k} - \alpha - 1}{1 - \alpha}, \quad k \geq 0,$$

see *e.g.* the command

$$\text{RSolve}[f[k+1]==1+1/a+f[k]/a, f[0]==1, f[k], k],$$

which shows that

$$\begin{aligned} \mathbb{E}[T_k^r \mid X_0 = k] &= \frac{1}{\alpha + 1} (1 + \mathbb{E}[T_k^r \mid X_0 = k - 1]) + \frac{\alpha}{\alpha + 1} (1 + \mathbb{E}[T_k^r \mid X_0 = k + 1]) \\ &= 1 + \frac{1}{\alpha + 1} \mathbb{E}[T_k^r \mid X_0 = k - 1] + \frac{\alpha}{\alpha + 1} \mathbb{E}[T_k^r \mid X_0 = k + 1] \\ &= 1 + \frac{1}{\alpha + 1} \mathbb{E}[T_k^r \mid X_0 = k - 1] + \frac{\alpha}{1 - \alpha} \\ &= \frac{1}{1 - \alpha} + \frac{1}{\alpha + 1} \mathbb{E}[T_k^r \mid X_0 = k - 1] \\ &= \frac{2\alpha^{-k}}{1 - \alpha^2}, \end{aligned}$$

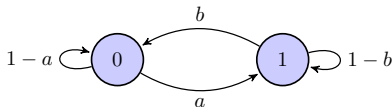
therefore the chain is positive recurrent if and only if  $\alpha < 1$ .

## Chapter 7 - Long-Run Behavior of Markov Chains

Exercise 7.1 Writing the condition  $\pi P = \pi$  leads to the equations

$$\begin{cases} \frac{\pi_0}{3} + 2\frac{\pi_1}{3} = \pi_0 \\ 2\frac{\pi_0}{3} + \frac{\pi_1}{3} = \pi_1 \end{cases}$$

*i.e.*  $\pi_0 = \pi_1$ . Combining this relation with the condition  $\pi_0 + \pi_1 = 1$  shows that  $\pi_0 = \pi_1 = 1/2$ .



Using the general relation

$$[\pi_0, \pi_1] = \left[ \frac{b}{a+b}, \frac{a}{a+b} \right],$$

with  $(a, b) \neq (0, 0)$  and  $(a, b) \neq (1, 1)$  for the two-state chain with transition matrix

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

yields the same answer

$$[\pi_0, \pi_1] = \left[ \frac{1}{2}, \frac{1}{2} \right]$$

when  $a = b$ , in which case the matrix  $P$  is also *column-stochastic*.

```

1 install.packages("igraph");install.packages("markovchain")
2 library("igraph");library(markovchain)
3 P<-matrix(c(1/3,2/3,2/3,1/3),nrow=2,byrow=TRUE);MC
4   <-new("markovchain",transitionMatrix=P)
5 graph <- as(MC, "igraph")
6 plot(graph,vertex.size=50,edge.label.cex=2,edge.label=E(graph)$prob,edge.color='black',
7       vertex.color='dodgerblue',vertex.label.cex=3)
8 steadyStates(object = MC)
9      1 2
10 [1,] 0.5 0.5

```

## Exercise 7.2

- 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  $\pi = \pi P$ , i.e.

$$[\pi_0, \pi_1, \pi_2, \pi_3, \dots] = [\pi_0, \pi_1, \pi_2, \pi_3, \dots] \times \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ q & 0 & p & 0 & 0 & \dots \\ q & 0 & 0 & p & 0 & \dots \\ q & 0 & 0 & 0 & p & \dots \\ q & 0 & 0 & 0 & 0 & \dots \\ \vdots & 0 & 0 & 0 & 0 & \ddots \end{bmatrix},$$

shows that

$$\begin{cases} \pi_0 = \pi_0 + q(\pi_1 + \pi_2 + \dots) = \pi_0 + q(1 - \pi_0) = q + p\pi_0, \\ \pi_1 = 0, \\ \pi_2 = p\pi_1, \\ \pi_3 = p\pi_2, \\ \vdots \end{cases}$$

hence  $\pi_k = 0$ ,  $k \geq 1$ , and the stationary distribution is given by

$$\pi = [1, 0, 0, 0, \dots].$$

c) Taking the limit as  $n$  tends to infinity in (S.30), we find

$$\lim_{n \rightarrow \infty} P^n = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

hence the limiting distribution coincides with the stationary distribution  $\pi$  obtained in (b).

### Exercise 7.3

- 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  $\pi = \pi P$ , *i.e.*

$$[\pi_0, \pi_1, \pi_2, \pi_3, \dots] = [\pi_0, \pi_1, \pi_2, \pi_3, \dots] \times \begin{bmatrix} q & p & 0 & 0 & 0 & \cdots \\ q & 0 & p & 0 & 0 & \cdots \\ q & 0 & 0 & p & 0 & \cdots \\ q & 0 & 0 & 0 & p & \cdots \\ q & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

shows that

$$\begin{cases} \pi_1 = p\pi_0, \\ \pi_2 = p\pi_1, \\ \pi_3 = p\pi_2, \\ \vdots \end{cases}$$

hence by induction we get  $\pi_k = p^k \pi_0$ ,  $k \geq 1$ . By the condition  $\sum_{k \geq 0} \pi_k = 1$  we find

$$1 = \sum_{k \geq 0} \pi_k = \pi_0 \sum_{k \geq 0} p^k = \frac{\pi_0}{1 - p},$$

which shows that  $\pi_0 = q$ . We conclude that  $\pi$  is the geometric distribution

$$\pi_k = (1 - p)p^k, \quad k \geq 0.$$

- c) Taking the limit as  $n$  tends to infinity in (S.31), we find, when  $p < 1$ ,

$$\lim_{n \rightarrow \infty} P^n = \begin{bmatrix} q & qp & qp^2 & qp^3 & qp^4 & \dots \\ q & qp & qp^2 & qp^3 & qp^4 & \dots \\ q & qp & qp^2 & qp^3 & qp^4 & \dots \\ q & qp & qp^2 & qp^3 & qp^4 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

hence the chain admits the geometric distribution with parameter  $p$  as a limiting distribution, which coincides with the stationary distribution  $\pi$  obtained in (b).

#### Exercise 7.4

- 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.
- The stationary distribution of the Ehrenfest chain satisfies the equation  $\pi P = \pi$ , which reads

$$\begin{aligned} [\pi_0, \pi_1, \pi_2, \dots] \times \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & \dots & 0 & 0 \\ 1/N & 0 & (N-1)/N & 0 & \dots & \dots & 0 & 0 \\ 0 & 2/N & 0 & (N-2)/N & \dots & \dots & 0 & 0 \\ 0 & 0 & 3/N & 0 & \dots & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \vdots & 0 & 3/N & 0 & 0 \\ 0 & 0 & \dots & 0 & (N-2)/N & 0 & 2/N & 0 \\ 0 & 0 & \dots & 0 & 0 & (N-1)/N & 0 & 1/N \\ 0 & 0 & \dots & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \\ = [\pi_0, \pi_1, \pi_2, \dots] = \pi, \end{aligned}$$

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

$$\begin{cases} \pi_0 = \frac{\pi_1}{N} \\ \pi_1 = \frac{2}{N-1}\pi_2, \\ \pi_2 = \frac{3}{N-2}\pi_3, \\ \pi_3 = \frac{4}{N-3}\pi_4, \\ \vdots \end{cases}$$

hence

$$\pi_0 = \frac{1}{N} \times \frac{2}{N-1} \times \frac{3}{N-2} \times \cdots \times \frac{k}{N-k+1} \pi_k = \frac{k!(N-k)!}{N!} \pi_k,$$

or

$$\pi_k = \binom{N}{k} \pi_0, \quad k = 0, \dots, N.$$

In addition we have

$$1 = \sum_{k=0}^N \pi_k = \pi_0 \sum_{k=0}^N \binom{N}{k} = (1+1)^N \pi_0 = 2^N \pi_0,$$

hence  $\pi_0 = 2^{-N}$  and  $(\pi_k)_{k=0,1,\dots,N}$  is the binomial distribution with parameter  $(N, 1/2)$ , *i.e.*

$$\pi_k = \binom{N}{k} \left(\frac{1}{2}\right)^N, \quad k = 0, 1, \dots, N. \quad (\text{S.43})$$

Conversely, from (S.43) we can recover

$$\begin{aligned} \sum_{l=0}^N \pi_l P_{l,k} &= \pi_{k-1} \frac{N-(k-1)}{N} + \pi_{k+1} \frac{k+1}{N} \\ &= \binom{N}{k-1} \left(\frac{1}{2}\right)^N \frac{N-k+1}{N} + \binom{N}{k+1} \left(\frac{1}{2}\right)^N \frac{k+1}{N} \\ &= \left(\frac{1}{2}\right)^N \frac{(N-1)!k}{(N-k)!k!} + \left(\frac{1}{2}\right)^N \frac{(N-1)!(N-k)}{(N-k)!k!} \\ &= \left(\frac{1}{2}\right)^N \binom{N}{k} \\ &= \pi_k, \quad k = 1, 2, \dots, N-1, \end{aligned}$$

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  code which shows that  $(P^n)_{n \geq 1}$  admits two subsequences converging to two distinct limits, therefore  $\lim_{n \rightarrow \infty} P^n$  does not exist and this chain does not admit a limiting distribution. Note that the conditions of Theorems 7.8 and 7.10 are not satisfied here, as this chain has period 2.

```

1  install.packages("expm");library(expm)
2  Ehrenfest <- function(n) {States <- c(0, seq(1,n))
   TPM <- matrix(0,nrow=length(States),ncol=length(States),dimnames=
4     list(seq(0,n),seq(0,n)))
   tran_prob <- function(i,n) {
6     tranRow <- rep(0,n+1)
     if(i==0) tranRow[2] <- 1
8     if(i==n) tranRow[(n+1)-1] <- 1
     if(i!=0 & i!=n) {
10      j=i+1
      tranRow[j-1] <- i/(n)
12      tranRow[j+1] <- 1-i/(n)}
     return(tranRow)}
14   for(j in 0:(n))TPM[j+1,] <- tran_prob(j,n)
   return(TPM)}
16  P=Ehrenfest(4)
   P%~%1000
18  P%~%1001
   P%~%1002

```

### Exercise 7.5

- 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

$$\left\{ \begin{array}{l} \pi_0 = \frac{\pi_1}{N^2} \\ \pi_1 = \pi_0 + \frac{2(N-1)}{N^2} \pi_1 + \frac{2^2}{N^2} \pi_2, \\ \pi_2 = \frac{(N-1)^2}{N^2} \pi_1 + \frac{4(N-2)}{N^2} \pi_2 + \frac{3^2}{N^2} \pi_3, \\ \pi_3 = \frac{(N-2)^2}{N^2} \pi_1 + \frac{6(N-3)}{N^2} \pi_2 + \frac{3^2}{N^2} \pi_3, \\ \vdots \end{array} \right.$$

from which we deduce

$$\left\{ \begin{array}{l} \pi_0 = \frac{\pi_1}{N^2} \\ \pi_1 = \frac{2^2}{(N-1)^2} \pi_2, \\ \pi_2 = \frac{3^2}{(N-2)^2} \pi_3, \\ \pi_3 = \frac{4^2}{(N-3)^2} \pi_4, \\ \vdots \end{array} \right.$$

hence

$$\pi_0 = \frac{1}{N^2} \times \frac{2^2}{(N-1)^2} \times \frac{3^2}{(N-2)^2} \times \cdots \times \frac{k^2}{(N-k+1)^2} \pi_k = \frac{k!^2 (N-k)!^2}{N!^2} \pi_k,$$

or

$$\pi_k = \binom{N}{k}^2 \pi_0, \quad k = 0, \dots, N.$$

In addition we have

$$1 = \sum_{k=0}^N \pi_k = \pi_0 \sum_{k=0}^N \binom{N}{k}^2 = \binom{2N}{N} \pi_0,$$

hence  $\pi_0 = 1/\binom{2N}{N}$  and  $(\pi_k)_{k=0,1,\dots,N}$  is the probability distribution given by

$$\pi_k = \frac{\binom{N}{k}^2}{\binom{2N}{N}}, \quad k = 0, 1, \dots, N.$$

- c) The Bernoulli-Laplace chain admits a limiting distribution which coincides with its stationary distribution by Theorem 7.8.

Exercise 7.6

- a) We have

$$\begin{bmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.5 & 0 \end{bmatrix}.$$

- b) By first step analysis, we have

$$\begin{cases} \mu_0(0) = 1 + \frac{1}{2}\mu_0(1) + \frac{1}{2}\mu_0(3) \\ \mu_0(1) = 1 + \frac{1}{2}\mu_0(2) \\ \mu_0(2) = 1 + \frac{1}{2}\mu_0(1) + \frac{1}{2}\mu_0(3) \\ \mu_0(3) = 1 + \frac{1}{2}\mu_0(2). \end{cases}$$

By symmetry of the maze we have  $\mu_0(1) = \mu_0(3)$ , hence

$$\begin{cases} \mu_0(0) = \mu_0(2) = 1 + \mu_0(1) \\ \mu_0(1) = \mu_0(3) = 1 + \frac{1}{2}\mu_0(2). \end{cases}$$

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

$$(\pi_0, \pi_1, \pi_2, \pi_3) = \left( \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \right)$$

is invariant and satisfies the condition  $\pi = \pi P$ , see also Exercise 7.15.

Recall that the family  $(\pi_0, \pi_1, \pi_2, \pi_3)$  is a *(probability) distribution if and only if*

$$(i) \quad \pi_0 \geq 0, \quad \pi_1 \geq 0, \quad \pi_2 \geq 0, \quad \pi_3 \geq 0,$$

and

(ii)

$$\pi_0 + \pi_1 + \pi_2 + \pi_3 = \mathbf{1}.$$

In the present situation we have

$$\pi_0 = \pi_1 = \pi_2 = \pi_3$$

and

$$\pi_0 + \pi_1 + \pi_2 + \pi_3 = \mathbf{1} = 4\pi_0,$$

which necessarily yields

$$\pi_0 = \pi_1 = \pi_2 = \pi_3 = \frac{1}{4} = \frac{1}{\mu_0(0)}.$$

### Exercise 7.7

- a) Clearly, the transition from the current state to the next state depends only on the current state on the chain, hence the process is Markov. The transition matrix of the chain on the state space  $\mathbb{S} = (D, N)$  is

$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = \begin{bmatrix} 1/4 & 3/4 \\ 1/4 & 3/4 \end{bmatrix}.$$

- b) The stationary distribution  $\pi = (\pi_D, \pi_N)$  is solution of  $\pi = \pi P$ , *i.e.*

$$\begin{cases} \pi_D = (1-a)\pi_D + b\pi_N = \frac{1}{4}\pi_D + \frac{1}{4}\pi_N \\ \pi_N = a\pi_D + (1-b)\pi_N = \frac{3}{4}\pi_D + \frac{3}{4}\pi_N \end{cases}$$

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  $\textcircled{N}$  starting from state  $\textcircled{D}$  satisfies

$$\mu_N(D) = (1-a)(1 + \mu_N(D)) + a \quad (\text{S.44})$$

hence  $\mu_N(D) = 1/a = 4/3$ .

### Additional comments:

- (i) The stationary distribution  $[\pi_D, \pi_N]$  satisfies

$$[\pi_D, \pi_N] = \left[ \frac{\mu_N(D)}{\mu_D(N) + \mu_N(D)}, \frac{\mu_D(N)}{\mu_D(N) + \mu_N(D)} \right],$$

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

$$\mu_N(D) = \frac{3}{4} \sum_{k \geq 1} k \left( \frac{1}{4} \right)^{k-1} = \frac{3}{4} \frac{1}{(1-1/4)^2} = \frac{4}{3},$$

cf. (A.4).

- e) The average time  $\mu_D(N)$  to reach state  $\textcircled{D}$  starting from state  $\textcircled{N}$  satisfies

$$\mu_D(N) = (1 - b)(1 + \mu_D(N)) + b \quad (\text{S.45})$$

hence  $\mu_N(D) = 1/b = 4$ .

**Additional comments:**

- (i) The value of  $\mu_D(N)$  may also be recovered as

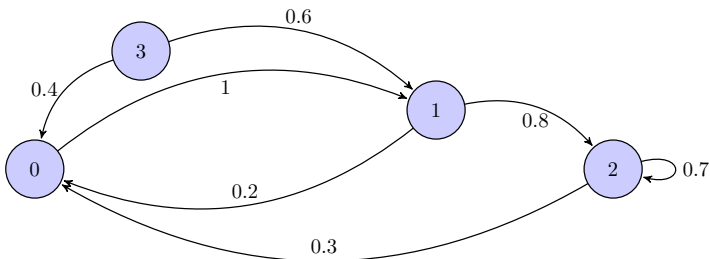
$$\mu_D(N) = \frac{1}{4} \sum_{k \geq 1} k \left( \frac{3}{4} \right)^{k-1} = \frac{1}{4} \frac{1}{(1 - 3/4)^2} = 4,$$

cf. (A.4).

- (ii) The values of  $\mu_D^D$  and  $\mu_N^N$  can also be computed from  $\mu_D^D = 1/\pi_D$  and  $\mu_N^N = 1/\pi_N$ .

**Exercise 7.8**

- a) The chain has the following graph:



The chain is reducible and its communicating classes are  $\{0, 1, 2\}$  and  $\{3\}$ .

- b) State  $\textcircled{3}$  is transient because  $\mathbb{P}(T_3 = \infty \mid X_0 = 3) = 0.4 + 0.6 = 1$ , cf. (6.3.1), and states  $\textcircled{0}, \textcircled{1}, \textcircled{2}$  are recurrent by Theorem 6.11 and Corollary 6.7.
- c) It suffices to consider the subchain on  $\{0, 1, 2\}$  with transition matrix

$$\tilde{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0.2 & 0 & 0.8 \\ 0.3 & 0 & 0.7 \end{bmatrix},$$

and to solve  $\pi = \pi \tilde{P}$ , i.e.

$$\begin{cases} \pi_0 = 0.2\pi_1 + 0.3\pi_2 \\ \pi_1 = \pi_0 \\ \pi_2 = 0.8\pi_1 + 0.7\pi_2 \end{cases}$$

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.*

$$\pi_0 = \frac{3}{14}, \quad \pi_1 = \frac{3}{14}, \quad \pi_2 = \frac{4}{7},$$

and the fraction of time spent at state ① 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 ①, we have

$$\begin{cases} h_0(0) = 0 \\ h_0(1) = 0.2(1 + h_0(0)) + 0.8(1 + h_0(2)) \\ h_0(2) = 0.3(1 + h_0(0)) + 0.7(1 + h_0(2)) \\ h_0(3) = 0.4(1 + h_0(0)) + 0.6(1 + h_0(1)), \end{cases}$$

*i.e.*

$$\begin{cases} h_0(0) = 0, \\ h_0(1) = 1 + 0.8h_0(2), \\ h_0(2) = 1 + 0.7h_0(2), \\ h_0(3) = 1 + 0.6h_0(1), \end{cases} \quad \text{or} \quad \begin{cases} h_0(0) = 0, \\ h_0(1) = 1 + 0.8h_0(2), \\ 0.3h_0(2) = 1, \\ h_0(3) = 1 + 0.6h_0(1), \end{cases}$$

hence

$$h_0(0) = 0, \quad h_0(1) = \frac{11}{3}, \quad h_0(2) = \frac{10}{3}, \quad h_0(3) = \frac{16}{5},$$

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

$$h_0(2) = 0.3 \sum_{k \geq 1} k(0.7)^{k-1} = \frac{0.3}{(1 - 0.7)^2} = \frac{10}{3}.$$

Note that the value of  $h_0(2)$  could also be computed by restriction to the sub-chain  $\{0, 1, 2\}$ , by solving

$$\begin{cases} h_0(0) = 0 \\ h_0(1) = 0.2(1 + h_0(0)) + 0.8(1 + h_0(2)) \\ h_0(2) = 0.3(1 + h_0(0)) + 0.7(1 + h_0(2)). \end{cases}$$

### Exercise 7.9

- a) First, we note that the chain has finite state space and it is irreducible, positive recurrent and aperiodic, hence by Theorem 7.8 its limiting distribution coincides with its stationary distribution which is the unique solution of  $\pi = \pi P$ . After calculations, this equation can be solved as

$$\pi_0 = c \times 161, \quad \pi_1 = c \times 460, \quad \pi_2 = c \times 320, \quad \pi_3 = c \times 170,$$

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

$$\pi_0 = \frac{161}{1111}, \quad \pi_1 = \frac{460}{1111}, \quad \pi_2 = \frac{320}{1111}, \quad \pi_3 = \frac{170}{1111}. \quad (\text{S.46})$$

- 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

$$\begin{cases} \mu_0(0) = 1 + \mu_0(1), \\ \mu_0(1) = 0.1 + 0.4(1 + \mu_0(1)) + 0.2(1 + \mu_0(2)) + 0.3(1 + \mu_0(3)) \\ \quad = 1 + 0.4\mu_0(1) + 0.2\mu_0(2) + 0.3\mu_0(3), \\ \mu_0(2) = 0.2 + 0.2(1 + \mu_0(1)) + 0.5(1 + \mu_0(2)) + 0.1(1 + \mu_0(3)) \\ \quad = 1 + 0.2\mu_0(1) + 0.5\mu_0(2) + 0.1\mu_0(3), \\ \mu_0(3) = 1 + 0.3\mu_0(1) + 0.4\mu_0(2), \end{cases}$$

hence

$$\mu_0(1) = \frac{950}{161}, \quad \mu_0(2) = \frac{860}{161}, \quad \mu_0(3) = \frac{790}{161},$$

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

$$\mu_0(0) = 1 + \mu_0(1) = 1 + \frac{950}{161} = \frac{161 + 950}{161} = \frac{1111}{161},$$

hence the relation  $\pi_0 = 1/\mu_0(0)$  is satisfied from (S.46).

### Exercise 7.10



- 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 6.13. By Proposition 6.16, 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

$$\begin{cases} \pi_0 = \frac{1}{4}\pi_1 + \frac{1}{2}\pi_3 \\ \pi_1 = \frac{1}{2}\pi_0 + \frac{1}{3}\pi_2 \\ \pi_2 = \frac{3}{4}\pi_1 + \frac{1}{2}\pi_3 \\ \pi_3 = \frac{1}{2}\pi_0 + \frac{2}{3}\pi_2, \end{cases}$$

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

$$\pi_0 = 2c, \quad \pi_1 = 2c, \quad \pi_2 = 3c, \quad \pi_3 = 3c,$$

and  $c = 1/10$ , *i.e.*

$$\pi_0 = 20\%, \quad \pi_1 = 20\%, \quad \pi_2 = 30\%, \quad \pi_3 = 30\%.$$

**Exercise 7.11** We choose to model the problem on the state space  $\{1, 2, 3, 4\}$ , meaning that the replacement of a component is immediate upon failure. Let  $X_n$  denote the remaining active time of the component at time  $n$ . Given that at time  $n$  there remains  $X_n = k \geq 2$  units of time until failure, we know with certainty that at the next time step  $n+1$  there will remain  $X_{n+1} = k-1 \geq 1$  units of time until failure. Hence at any time  $n \geq 1$  we have

$$X_n = 4 \implies X_{n+1} = 3 \implies X_{n+2} = 2 \implies X_{n+3} = 1,$$

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  $T \in \{1, 2, 3, 4\}$ . Hence we have

$$\mathbb{P}(X_{n+1} = k \mid X_n = 1) = \mathbb{P}(T = k), \quad k = 1, 2, 3, 4,$$

and the process  $(X_n)_{n \in \mathbb{N}}$  is a Markov chain on  $\mathbb{S} = \{1, 2, 3, 4\}$ , with transition matrix

$$P = \begin{bmatrix} \mathbb{P}(Y=1) & \mathbb{P}(Y=2) & \mathbb{P}(Y=3) & \mathbb{P}(Y=4) \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

We now look for the limit

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = 1).$$

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

$$\pi_1 = \lim_{n \rightarrow \infty} \mathbb{P}(X_n = 1),$$

where  $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$  is the stationary distribution  $\pi$  uniquely determined from the equation  $\pi = \pi P$ , as follows:

$$\begin{cases} \pi_1 = 0.1\pi_1 + \pi_2 \\ \pi_2 = 0.2\pi_1 + \pi_3 \\ \pi_3 = 0.3\pi_1 + \pi_4 \\ \pi_4 = 0.4\pi_1. \end{cases}$$

hence

$$\pi_2 = 0.9\pi_1, \quad \pi_3 = 0.7\pi_1, \quad \pi_4 = 0.4\pi_1,$$

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

$$\pi_1 = \frac{1}{3}, \quad \pi_2 = \frac{9}{30}, \quad \pi_3 = \frac{7}{30}, \quad \pi_4 = \frac{4}{30}.$$

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:

```

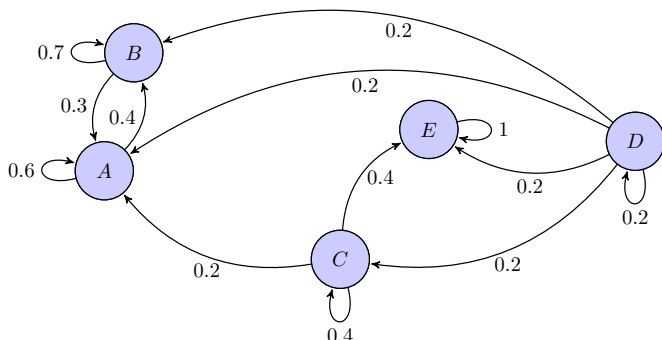
1 P = [0.1,0.2,0.3,0.4;
2   1,0,0,0;
3   0,1,0,0;
4   0,0,1,0;
5   mpower(P,1000)

```

showing that

$$\lim_{n \rightarrow \infty} P^n = \begin{bmatrix} 0.33333 & 0.30000 & 0.23333 & 0.13333 \\ 0.33333 & 0.30000 & 0.23333 & 0.13333 \\ 0.33333 & 0.30000 & 0.23333 & 0.13333 \\ 0.33333 & 0.30000 & 0.23333 & 0.13333 \end{bmatrix}.$$

Exercise 7.12 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  $\textcircled{C}$ , one can only return to  $\textcircled{C}$  or end up in one of the absorbing classes  $\{A, B\}$  or  $\{E\}$ . Let us denote by

$$T_{\{A,B\}} = \inf\{n \geq 0 : X_n \in \{A, B\}\}$$

the hitting time of  $\{A, B\}$ . We start by computing  $\mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C)$ . By first step analysis, we find that this probability satisfies

$$\mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C) = 0.2 + 0.4 \times \mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C) + 0.4 \times 0,$$

hence

$$\mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C) = \frac{1}{3},$$

which can also be recovered using a geometric series as

$$\mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C) = 0.2 \sum_{n \geq 0} (0.4)^n = \frac{0.2}{1 - 0.4} = \frac{1}{3}.$$

On the other hand,  $\{A, B\}$  is a closed two-state chain with transition matrix

$$\begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix},$$

hence, starting from any state within  $\{A, B\}$ , the long run probability of being in  $A$  is given by

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{P}(X_n = A \mid X_0 \in \{A, B\}) \\ &= \lim_{n \rightarrow \infty} \mathbb{P}(X_n = A \mid X_0 = A) \frac{\mathbb{P}(X_0 = A)}{\mathbb{P}(X_0 \in \{A, B\})} \\ & \quad + \lim_{n \rightarrow \infty} \mathbb{P}(X_n = A \mid X_0 = B) \frac{\mathbb{P}(X_0 = B)}{\mathbb{P}(X_0 \in \{A, B\})} \\ &= \frac{b}{a+b} \left( \frac{\mathbb{P}(X_0 = A)}{\mathbb{P}(X_0 \in \{A, B\})} + \frac{\mathbb{P}(X_0 = B)}{\mathbb{P}(X_0 \in \{A, B\})} \right) \\ &= \frac{0.3}{0.3+0.4} = \frac{3}{7}. \end{aligned}$$

Since

$$\{X_n = A\} \subset \{T_{\{A,B\}} \leq n\} \subset \{T_{\{A,B\}} < \infty\}, \quad n \geq 0,$$

we conclude that

$$\begin{aligned} \alpha &:= \lim_{n \rightarrow \infty} \mathbb{P}(X_n = A \mid X_0 = C) \\ &= \lim_{n \rightarrow \infty} \mathbb{P}(T_{\{A,B\}} < \infty \text{ and } X_n = A \mid X_0 = C) \\ &= \lim_{n \rightarrow \infty} \mathbb{P}(T_{\{A,B\}} < \infty \text{ and } X_{n+T_{\{A,B\}}} = A \mid X_0 = C) \\ &= \mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C) \lim_{n \rightarrow \infty} \mathbb{P}(X_{n+T_{\{A,B\}}} = A \mid T_{\{A,B\}} < \infty \text{ and } X_0 = C) \\ &= \mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C) \\ & \quad \times \lim_{n \rightarrow \infty} \mathbb{P}(X_{n+T_{\{A,B\}}} = A \mid T_{\{A,B\}} < \infty, X_{T_{\{A,B\}}} \in \{A, B\}, X_0 = C) \\ &= \mathbb{P}(T_{\{A,B\}} < \infty \mid X_0 = C) \lim_{n \rightarrow \infty} \mathbb{P}(X_n = A \mid X_0 \in \{A, B\}) \\ &= \frac{1}{3} \times \frac{3}{7} \\ &= \frac{1}{7}, \end{aligned}$$

where we used the strong Markov property, cf. Exercise 5.8.

The value  $\alpha$  of this probability could also be obtained in a faster way by the following more intuitive first step analysis:

$$\alpha = 0.4 \times \alpha + 0.2 \times \frac{0.3}{0.7},$$

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

$$P^{20} = \begin{bmatrix} 0.4285 & 0.5714 & 0 & 0 & 0 \\ 0.4285 & 0.5714 & 0 & 0 & 0 \\ 0.1428 & 0.1904 & 1.099510^{-8} & 0 & 0.6666 \\ 0.2500 & 0.3333 & 1.099510^{-8} & 1.048510^{-14} & 0.4166 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

which recovers

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = A \mid X_0 = C) = \frac{1}{7} \simeq 0.142857$$

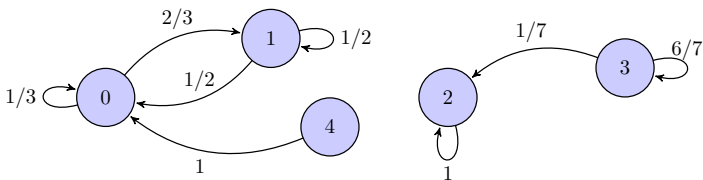
up to 6 decimals, and one can reasonably conjecture that the limiting distribution of the chain is given by

$$\lim_{n \rightarrow \infty} P^n = \begin{bmatrix} 3/7 & 4/7 & 0 & 0 & 0 \\ 3/7 & 4/7 & 0 & 0 & 0 \\ 1/7 & 4/21 & 0 & 0 & 2/3 \\ 1/4 & 1/3 & 0 & 0 & 5/12 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (\text{S.47})$$

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 (S.47) we also note that  $\lim_{n \rightarrow \infty} \mathbb{P}(X_n = i \mid X_0 = j)$  is dependent on the initial state  $\textcircled{j}$ , 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 (S.47) is a stationary distribution solution of  $\pi = \pi P$ , cf. Proposition 7.7.

### Exercise 7.13

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 ① and ② are recurrent, and state ② is absorbing (hence it is recurrent).  
 d) By (4.5.7) we have

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0 \mid X_0 = 4) = \lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0 \mid X_0 = 0) = \frac{1/2}{2/3 + 1/2} = \frac{3}{7},$$

cf. also the Table 8.1.

#### Exercise 7.14

- a) The transition matrix  $P$  of the chain on the state space  $\mathbb{S} = (C, T)$  is given by

$$\begin{bmatrix} 4/5 & 1/5 \\ 3/4 & 1/4 \end{bmatrix}.$$

- b) The stationary distribution  $\pi = (\pi_C, \pi_T)$  is solution of  $\pi = \pi P$ , i.e.

$$\begin{cases} \pi_C = \frac{4}{5}\pi_C + \frac{3}{4}\pi_T \\ \pi_T = \frac{1}{5}\pi_C + \frac{1}{4}\pi_T \end{cases}$$

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  $\textcircled{T}$  starting from  $\textcircled{C}$  and  $\textcircled{T}$ , respectively. By first step analysis, we have

$$\begin{cases} \mu_T(C) = 1 + \frac{4}{5}\mu_T(C) \\ \mu_T(T) = 1 + \frac{3}{4}\mu_T(C) \end{cases}$$

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 7.15

- a) We note that, due to the symmetry of  $P$ , the vector

$$\pi = [\pi_1, \pi_2, \dots, \pi_N] := [1/N, \dots, 1/N]$$

is an invariant probability distribution for the chain, since we can check the balance condition

$$\begin{aligned} [\pi P]_j &= \sum_{i=1}^N \pi_i P_{i,j} \\ &= \frac{1}{N} \sum_{i=1}^N P_{j,i} \\ &= \frac{1}{N} = \pi_j, \quad j = 1, 2, \dots, N. \end{aligned}$$

- b) When  $N = 2$  the chain can only have the transition matrix

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

which clearly has period 2. When  $N \geq 3$  the chain is aperiodic because it is irreducible and one can travel from  $\textcircled{0}$  to  $\textcircled{0}$  in 2 steps via the path  $\textcircled{0} \rightarrow \textcircled{1} \rightarrow \textcircled{0}$ , and in 3 steps via the path  $\textcircled{0} \rightarrow \textcircled{1} \rightarrow \textcircled{2} \rightarrow \textcircled{0}$ .

### Exercise 7.16 (Problem 5.22-(d) 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  $\textcircled{5}$  is absorbing. This means that

$$\lim_{n \rightarrow \infty} P^n = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

which would be difficult to recover by a direct computation of  $P^n$ .

For the stationary distribution, the equation  $\pi = \pi P$  reads

$$\begin{cases} \pi_0 = q\pi_0 + q\pi_1 + q\pi_2 \\ \pi_1 = p\pi_0 + p\pi_4 \\ \pi_2 = p\pi_1 \\ \pi_3 = p\pi_2 + p\pi_3 \\ \pi_4 = p\pi_3 \\ \pi_5 = q\pi_4 + \pi_5, \end{cases}$$

*i.e.*

$$\begin{cases} p\pi_0 = q\pi_1 + q\pi_2 \\ \pi_1 = p\pi_0 + p\pi_4 \\ \pi_2 = p\pi_1 \\ q\pi_3 = p\pi_2 \\ \pi_4 = p\pi_3 \\ \pi_4 = 0, \end{cases}$$

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 7.17

a) We solve the system of equations

$$\begin{cases} \pi_0 = q(\pi_0 + \pi_1 + \pi_2 + \pi_3) + \pi_4 = q + p\pi_4 \\ \pi_1 = p\pi_0 \\ \pi_2 = p\pi_1 = p^2\pi_0 \\ \pi_3 = p\pi_2 = p^3\pi_0 \\ \pi_4 = p\pi_3 = p^4\pi_0, \end{cases}$$

which yields

$$1 = \pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 = \pi_0(1 + p + p^2 + p^3 + p^4),$$

and

$$\begin{cases} \pi_0 = \frac{1}{1 + p + p^2 + p^3 + p^4} \\ \pi_1 = \frac{p}{1 + p + p^2 + p^3 + p^4} \\ \pi_2 = \frac{p^2}{1 + p + p^2 + p^3 + p^4} \\ \pi_3 = \frac{p^3}{1 + p + p^2 + p^3 + p^4} \\ \pi_4 = \frac{p^4}{1 + p + p^2 + p^3 + p^4}. \end{cases}$$

- b) Since the chain is irreducible and aperiodic with finite state space, its limiting distribution coincides with its stationary distribution.

### Exercise 7.18

- a) The transition matrix  $P$  is given by

$$P = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/3 & 0 & 1/3 & 1/3 \\ 0 & 1 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix}.$$

- b) The chain is aperiodic, irreducible, and has finite state space hence we can apply Theorem 7.8 or Theorem 7.10. The equation  $\pi P = \pi$  reads

$$\begin{aligned} \pi P = [\pi_A, \pi_B, \pi_C, \pi_D] \times \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/3 & 0 & 1/3 & 1/3 \\ 0 & 1 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix} &= \begin{bmatrix} \frac{1}{3}\pi_B + \frac{1}{2}\pi_D \\ \frac{1}{2}\pi_A + \pi_C + \frac{1}{2}\pi_D \\ \frac{1}{3}\pi_B \\ \frac{1}{2}\pi_A + \frac{1}{3}\pi_B \end{bmatrix} \\ &= [\pi_A, \pi_B, \pi_C, \pi_D], \end{aligned}$$

*i.e.*  $\pi_A = \pi_D = 2\pi_C$  and  $\pi_B = 3\pi_C$ , which, under the condition  $\pi_A + \pi_B + \pi_C + \pi_D = 1$ , gives  $\pi_A = 1/4$ ,  $\pi_B = 3/8$ ,  $\pi_C = 1/8$ ,  $\pi_D = 1/4$ .

- c) We solve the system

$$\begin{cases} \mu_D(A) = \frac{1}{2} + \frac{1}{2}(1 + \mu_D(B)) = 1 + \frac{1}{2}\mu_D(B) \\ \mu_D(B) = \frac{1}{3} + \frac{1}{3}(1 + \mu_D(A)) + \frac{1}{3}(1 + \mu_D(C)) = 1 + \frac{1}{3}(\mu_D(A) + \mu_D(C)) \\ \mu_D(C) = 1 + \mu_D(B) \\ \mu_D(D) = \frac{1}{2}(1 + \mu_D(A)) + \frac{1}{2}(1 + \mu_D(B)) = 1 + \frac{1}{2}(\mu_D(A) + \mu_D(B)), \end{cases}$$

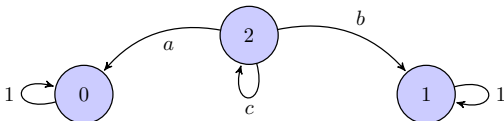
which has for solution

$$\mu_D(A) = \frac{8}{3}, \quad \mu_D(B) = \frac{10}{3}, \quad \mu_D(C) = \frac{13}{3}, \quad \mu_D(D) = 4.$$

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 7.19 Clearly we may assume that  $c < 1$ , as the case  $c = 1$  corresponds to the identity matrix, or to constant a chain. On the other hand, we cannot directly apply Theorem 7.8 since the chain is reducible. The chain has the following graph:



a) By observation of

$$P^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a(1+c) & b(1+c) & c^2 \end{bmatrix}$$

and

$$P^3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a(1+c+c^2) & b(1+c+c^2) & c^3 \end{bmatrix}$$

we infer that  $P^n$  takes the general form

$$P^n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_n & b_n & c_n \end{bmatrix},$$

where  $a_n$ ,  $b_n$  and  $c_n$  are coefficients to be determined by the following induction argument. Writing down the relation  $P^{n+1} = P \times P^n$  as

$$P^{n+1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_{n+1} & b_{n+1} & c_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & b & c \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_n & b_n & c_n \end{bmatrix}$$

shows that we have the recurrence relations

$$\begin{cases} a_{n+1} = a + ca_n, \\ b_{n+1} = b + cb_n, \\ c_{n+1} = c \times c_n, \end{cases} \quad \text{which yield} \quad \begin{cases} a_n = a + ac + \dots + ac^{n-1} = a \frac{1-c^n}{1-c}, \\ b_n = b + bc + \dots + bc^{n-1} = b \frac{1-c^n}{1-c}, \\ c_{n+1} = c^n, \end{cases}$$

hence

$$P^n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a \frac{1-c^n}{1-c} & b \frac{1-c^n}{1-c} & c^n \end{bmatrix}.$$

- b) From the structure of  $P^n$  it follows that the chain admits a limiting distribution

$$\lim_{n \rightarrow \infty} P^n = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{a}{1-c} & \frac{b}{1-c} & 0 \end{bmatrix}.$$

which is dependent of the initial state, provided that  $c < 1$ . The limiting probabilities

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0 \mid X_0 = 2) = \frac{a}{1-c},$$

resp.

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = 1 \mid X_0 = 2) = \frac{b}{1-c},$$

correspond to the probability of moving to state ②, 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 7.20

- a) The process  $(X_n)_{n \geq 0}$  is a two-state Markov chain on  $\{0, 1\}$  with transition matrix

$$\begin{bmatrix} \alpha & \beta \\ p & q \end{bmatrix}$$

and  $\alpha = 1 - \beta$ . The entries on the second line are easily obtained. Concerning the first line we note that  $\mathbb{P}(N = 1) = \beta$  is the probability of switching from 0 to 1 in one time step, while the equality  $\mathbb{P}(N \geq 2) = 1 - \beta$  shows that the probability of remaining at 0 for one time step is  $1 - \beta$ .

- b) This probability is given from the stationary distribution  $(\pi_0, \pi_1)$  as  $\pi_1 = \frac{\beta}{p + \beta}$ .

### Exercise 7.21

- 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

$$\pi_i = \frac{1}{\mu_i(i)}, \quad i = 1, 2, \dots, N,$$

by Corollary 7.9. Since  $\pi = (\pi_1, \pi_2, \dots, \pi_N)$  is a probability distribution we have  $\pi_1 + \pi_2 + \dots + \pi_N = 1$  hence there exists  $i \in \{1, 2, \dots, N\}$  with  $\pi_i = 1/\mu_i(i) \geq 1/N$ , i.e.  $\mu_i(i) \leq N$ , otherwise the sum  $\pi_1 + \pi_2 + \dots + \pi_N$  would be lower than one.

- b) Similarly, the condition  $\pi_1 + \pi_2 + \dots + \pi_N = 1$  shows that there must exist  $i \in \{1, 2, \dots, N\}$  with  $\pi_i = 1/\mu_i(i) \leq 1/N$ , i.e.  $\mu_i(i) \geq N$ , otherwise the sum  $\pi_1 + \pi_2 + \dots + \pi_N$  would be larger than one.

### Exercise 7.22

- a) The  $N \times N$  transition matrix of the chain is

$$P = \begin{bmatrix} q & p & 0 & \cdots & \cdots & 0 & 0 & 0 \\ q & 0 & p & \cdots & \cdots & 0 & 0 & 0 \\ 0 & q & 0 & \ddots & \ddots & 0 & 0 & 0 \\ 0 & 0 & q & \ddots & \ddots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & \ddots & 0 & p & 0 \\ 0 & 0 & 0 & \cdots & \cdots & q & 0 & p \\ 0 & 0 & 0 & \cdots & \cdots & 0 & q & p \end{bmatrix}.$$

- b) The chain is irreducible if  $p \in (0, 1)$ , and reducible if  $p = 0$  or  $p = 1$ .  
c) If  $p \in (0, 1)$  there are no absorbing states and all states are positive recurrent. If  $p = 0$ , state ① is absorbing and all other states are transient. If  $p = 1$ , state  $N$  is absorbing and all other states are transient.  
d) The equation  $\pi = \pi P$  yields

$$\pi_2 = \frac{p}{q}\pi_1 \quad \text{and} \quad \pi_N = \frac{p}{q}\pi_{N-1}, \quad k = 1, 2, \dots, N-1,$$

and

$$p(\pi_k - \pi_{k-1}) = q(\pi_{k+1} - \pi_k) \quad k = 2, 3, \dots, N-1.$$

We check that  $\pi_k$  given by

$$\pi_k = \frac{p^{k-1}}{q^{k-1}}\pi_1, \quad k = 1, 2, \dots, N,$$

satisfies the above conditions. The normalization condition

$$1 = \sum_{k=1}^N \pi_k = \pi_1 \sum_{k=1}^N \frac{p^{k-1}}{q^{k-1}} = \pi_1 \sum_{k=0}^{N-1} \left(\frac{p}{q}\right)^k = \pi_1 \frac{1 - (p/q)^N}{1 - p/q}$$

shows that

$$\pi_k = \frac{1 - p/q}{1 - (p/q)^N} \frac{p^{k-1}}{q^{k-1}}, \quad k = 1, 2, \dots, N,$$

provided that  $0 < p \neq q < 1$ . When  $p = q = 1/2$  we find that the uniform distribution

$$\pi_k = \frac{1}{N}, \quad k = 1, 2, \dots, N,$$

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 7.23** (Problem 6.9 continued). We note that the chain is irreducible and has period 2.

- a) By Problem 6.9 the chain is positive recurrent if and only if  $\alpha < 1$ , hence by Theorem 7.10 it admits a stationary distribution  $\pi$  in this case. The relation  $\pi = \pi P$  reads

$$\pi = [\pi_0, \pi_1, \pi_2, \dots] = [\pi_0, \pi_1, \pi_2, \dots] \times \begin{bmatrix} 0 & 1 & 0 & 0 & \dots \\ \frac{1}{\alpha+1} & 0 & \frac{\alpha}{\alpha+1} & 0 & \dots \\ 0 & \frac{1}{\alpha+1} & 0 & \frac{\alpha}{\alpha+1} & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix},$$

hence

$$\pi_0 = \frac{\pi_1}{\alpha+1}, \quad \pi_1 = \pi_0 + \frac{\pi_2}{\alpha+1}, \quad \pi_2 = \frac{\alpha\pi_1}{\alpha+1} + \frac{\pi_3}{\alpha+1}, \dots$$

i.e.

$$\pi_0 = \frac{\pi_2}{\alpha(\alpha+1)}, \quad \pi_1 = \frac{\pi_2}{\alpha}, \quad \pi_2 = \frac{\pi_3}{\alpha}, \dots$$

from which we can deduce that

$$\pi_k = \pi_0 \alpha^{k-1} (1 + \alpha), \quad k \geq 1.$$

From the condition  $\sum_{k \geq 0} \pi_k = 1$  we find

$$1 = \pi_0 + \pi_0(1 + \alpha) \sum_{k \geq 1} \alpha^{k-1} = \pi_0 + \pi_0 \frac{1 + \alpha}{1 - \alpha},$$

hence

$$\pi_0 = \frac{1 - \alpha}{2} \quad \text{and} \quad \pi_k = \alpha^{k-1} \frac{1 - \alpha^2}{2}, \quad k \geq 1.$$

More generally, from the relation  $\pi = \pi P$  we can also write

$$\pi_k = \frac{\alpha \pi_{k-1}}{\alpha + 1} + \frac{\pi_{k+1}}{\alpha + 1}, \quad k \geq 2,$$

which can be written as

$$\pi_k = q\pi_{k-1} + p\pi_{k+1}, \quad k \geq 2, \quad (\text{S.48})$$

where  $p = 1/(\alpha + 1)$  and  $q = \alpha/(\alpha + 1)$ . We look for a solution of (S.48) of the form  $\pi_k = C(q/p)^k = C\alpha^k$  as in (2.2.15),  $k \geq 1$ , which clearly satisfies (S.48) written as

$$\frac{\alpha \alpha^{k-1}}{\alpha + 1} + \frac{\alpha^{k+1}}{\alpha + 1} = \alpha^k, \quad k \geq 2.$$

When  $\alpha < 1$  the value of  $C$  can be found from the condition

$$\begin{aligned} 1 &= \sum_{k \geq 0} \pi_k \\ &= \pi_0 + \sum_{k \geq 2} \pi_k \\ &= \frac{\pi_2}{\alpha(\alpha + 1)} + \sum_{k \geq 1} \pi_k \\ &= C \frac{\alpha}{\alpha + 1} + C \frac{\alpha}{1 - \alpha} \\ &= C \frac{2\alpha}{1 - \alpha^2} \end{aligned}$$

hence

$$C = \frac{1 - \alpha^2}{2\alpha}$$

and

$$\pi_0 = \frac{1 - \alpha}{2}, \quad \pi_k = \alpha^{k-1} \frac{1 - \alpha^2}{2}, \quad k \geq 1.$$

- b) When  $\alpha > 1$  the series  $\sum_{k \geq 1} \pi_k$  does not converge so there cannot be a stationary distribution. In addition,  $\pi_0$  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

$$\mathbb{E}[T_0 \mid X_0 = 0] = \frac{1}{1 - \alpha} = \frac{1}{\pi_0},$$

and

$$\mathbb{E}[T_k \mid X_0 = k] = \frac{2\alpha^{k-1}}{1 - \alpha^2} = \frac{1}{\pi_k} < \infty, \quad k \geq 1,$$

as expected from Theorem 7.10.

### Exercise 7.24

- a) We have

$$\begin{aligned} \mathbb{P}(X_n = x) &= \mathbb{P}(\{X_n = x\} \cap \{\tau \leq n\}) + \mathbb{P}(\{X_n = x\} \cap \{\tau > n\}) \\ &= \mathbb{P}(\{Y_n = x\} \cap \{\tau \leq n\}) + \mathbb{P}(\{X_n = x\} \cap \{\tau > n\}) \\ &\leq \mathbb{P}(Y_n = x) + \mathbb{P}(\tau > n). \end{aligned}$$

- b) Similarly to Question (a), we have

$$\begin{aligned} \mathbb{P}(Y_n = x) &= \mathbb{P}(\{Y_n = x\} \cap \{\tau \leq n\}) + \mathbb{P}(\{Y_n = x\} \cap \{\tau > n\}) \\ &= \mathbb{P}(\{X_n = x\} \cap \{\tau \leq n\}) + \mathbb{P}(\{Y_n = x\} \cap \{\tau > n\}) \\ &\leq \mathbb{P}(X_n = x) + \mathbb{P}(\tau > n), \end{aligned}$$

hence

$$-\mathbb{P}(\tau > n) \leq \mathbb{P}(X_n = x) - \mathbb{P}(Y_n = x) \leq \mathbb{P}(\tau > n),$$

which leads to the conclusion.

### Problem 7.25

- a) We have

$$\begin{aligned} \mathbb{P}(Y_{n+1} = j \mid Y_n = i) &= \frac{\mathbb{P}(Y_{n+1} = j \text{ and } Y_n = i)}{\mathbb{P}(Y_n = i)} \\ &= \frac{\mathbb{P}(Y_{n+1} = j)}{\mathbb{P}(Y_n = i)} \frac{\mathbb{P}(Y_n = i \text{ and } Y_{n+1} = j)}{\mathbb{P}(Y_{n+1} = j)} \\ &= \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i)} \frac{\mathbb{P}(X_{N-n} = i \text{ and } X_{N-n-1} = j)}{\mathbb{P}(X_{N-n-1} = j)} \\ &= \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i)} \mathbb{P}(X_{N-n} = i \mid X_{N-n-1} = j) = \frac{\pi_j}{\pi_i} P_{j,i}. \end{aligned}$$

On the other hand, we have

$$\begin{aligned}
\mathbb{P}(Y_{n+1} = j \mid Y_n = i_n, \dots, Y_0 = i_0) &= \frac{\mathbb{P}(Y_{n+1} = j, Y_n = i_n, \dots, Y_0 = i_0)}{\mathbb{P}(Y_n = i_n, \dots, Y_0 = i_0)} \\
&= \frac{\mathbb{P}(X_{N-n-1} = j, X_{N-n} = i_n, \dots, X_N = i_0)}{\mathbb{P}(X_{N-n} = i_n, \dots, X_N = i_0)} \\
&= \mathbb{P}(X_{N-n-1} = j \text{ and } X_{N-n} = i_n) \\
&\quad \times \frac{\mathbb{P}(X_{N-n+1} = i_{n-1}, \dots, X_N = i_0 \mid X_{N-n-1} = j, X_{N-n} = i_n)}{\mathbb{P}(X_{N-n+1} = i_{n-1}, \dots, X_N = i_0)} \\
&= \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i_n)} \frac{\mathbb{P}(X_{N-n-1} = j \text{ and } X_{N-n} = i_n)}{\mathbb{P}(X_{N-n-1} = j)} \\
&\quad \times \frac{\mathbb{P}(X_{N-n+1} = i_{n-1}, \dots, X_N = i_0 \mid X_{N-n-1} = j, X_{N-n} = i_n)}{\mathbb{P}(X_{N-n+1} = i_{n-1}, \dots, X_N = i_0 \mid X_{N-n} = i_n)} \\
&= \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i_n)} \mathbb{P}(X_{N-n} = i_n \mid X_{N-n-1} = j, ) \\
&\quad \times \frac{\mathbb{P}(X_{N-n+1} = i_{n-1}, \dots, X_N = i_0 \mid X_{N-n-1} = j)}{\mathbb{P}(X_{N-n+1} = i_{n-1}, \dots, X_N = i_0 \mid X_{N-n} = i_n)} \\
&= \frac{\mathbb{P}(X_{N-n-1} = j)}{\mathbb{P}(X_{N-n} = i_n)} \mathbb{P}(X_{N-n} = i_n \mid X_{N-n-1} = j) = \frac{\pi_j}{\pi_{i_n}} P_{j, i_n},
\end{aligned}$$

and this shows that

$$\mathbb{P}(Y_{n+1} = j \mid Y_n = i_n, \dots, Y_0 = i_0) = \mathbb{P}(Y_{n+1} = j \mid Y_n = i_n) = \frac{\pi_j}{\pi_{i_n}} P_{j, i_n},$$

*i.e.* the time-reversed process  $(Y_n)_{n=0,1,\dots,N}$  has the Markov property.

b) We find

$$P_{i,j} = \frac{\pi_i}{\pi_j} P_{j,i}, \quad (\text{S.49})$$

*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

$$\pi_j = \sum_i \pi_j P_{j,i} = \sum_i \pi_i P_{i,j} = [\pi P]_j.$$

d) According to the detailed balance condition (S.49), we have

$$P_{k_1, k_2} P_{k_2, k_3} \cdots P_{k_n, k_1} = P_{k_n, k_1} \prod_{i=1}^{n-1} P_{k_i, k_{i+1}} = P_{k_n, k_1} \prod_{i=1}^{n-1} \frac{\pi_{k_{i+1}}}{\pi_{k_i}} P_{k_{i+1}, k_i}$$

$$= \frac{\pi_{k_n}}{\pi_{k_1}} P_{k_n, k_1} \prod_{i=1}^{n-1} P_{k_{i+1}, k_i} = P_{k_1, k_n} \prod_{i=1}^{n-1} P_{k_{i+1}, k_i},$$

holds for all sequences  $\{k_1, k_2, \dots, k_n\}$  of states and  $n \geq 2$ .

e) If the Markov chain satisfies

$$P_{i, k_2} P_{k_2, k_3} \cdots P_{k_{n-1}, j} P_{j, i} = P_{i, j} P_{j, k_{n-1}} \cdots P_{k_3, k_2} P_{k_2, i}$$

then by summation over the indices  $k_2, k_3, \dots, k_{n-1}$ , using the matrix power relation

$$[P^{n-1}]_{i, j} = \sum_{k_2, \dots, k_{n-1}} P_{i, k_2} P_{k_2, k_3} \cdots P_{k_{n-1}, j},$$

we get

$$\begin{aligned} [P^{n-1}]_{i, j} P_{j, i} &= P_{j, i} \sum_{k_2, \dots, k_{n-1}} P_{i, k_2} P_{k_2, k_3} \cdots P_{k_{n-1}, j} \\ &= P_{i, j} \sum_{k_2, \dots, k_{n-1}} P_{j, k_{n-1}} \cdots P_{k_3, k_2} P_{k_2, i} \\ &= P_{i, j} [P^{n-1}]_{j, i}. \end{aligned}$$

On the other hand, taking the limit as  $n$  goes to infinity, Theorem 7.8 shows that

$$\lim_{n \rightarrow \infty} [P^{n-1}]_{j, i} = \lim_{n \rightarrow \infty} [P^n]_{j, i} = \pi_i, \quad i, j \in \mathbb{S},$$

since the limiting and stationary distributions coincide, and we get

$$\pi_j P_{j, i} = P_{i, j} \pi_i, \quad i, j \in \mathbb{S},$$

which is the detailed balance condition.

f) The detailed balance condition reads

$$\pi_i P_{i, i+1} = \pi_i \left( \frac{1}{2} - \frac{i}{2M} \right) = \pi_{i+1} P_{i+1, i} = \pi_{i+1} \frac{i+1}{2M},$$

hence

$$\frac{\pi_{i+1}}{\pi_i} = \frac{1 - i/M}{(i+1)/M} = \frac{M-i}{i+1},$$

which shows that

$$\pi_i = \frac{(M-i+1)}{i} \frac{(M-i+2)}{(i-1)} \cdots \frac{(M-1)}{2} \frac{M}{1} \pi_0 = \frac{M!}{i!(M-i)!} \pi_0 = \pi_0 \binom{M}{i},$$

$i = 0, 1, \dots, M$ , where the constant  $\pi_0 > 0$  is given by

$$1 = \sum_i \pi_i = \pi_0 \sum_{i=0}^M \binom{M}{i} = \pi_0 2^M,$$

hence  $\pi_0 = 2^{-M}$  and

$$\pi_i = \frac{1}{2^M} \binom{M}{i}, \quad i = 0, 1, \dots, M.$$

g) We have

$$\begin{aligned} [\pi P]_i &= P_{i+1,i} \pi_{i+1} + P_{i,i} \pi_i + P_{i-1,i} \pi_{i-1} \\ &= \frac{1}{2^M} \frac{i+1}{2M} \binom{M}{i+1} + \frac{1}{2^M} \left( \frac{1}{2} - \frac{i-1}{2M} \right) \binom{M}{i-1} + \frac{1}{2^M} \times \frac{1}{2} \binom{M}{i} \\ &= \frac{1}{2^M} \frac{i+1}{2M} \frac{M!}{(i+1)!(M-i-1)!} + \frac{1}{2^M} \frac{M-i+1}{2M} \frac{M!}{(i-1)!(M-i+1)!} \\ &\quad + \frac{1}{2^M \times 2} \frac{M!}{i!(M-i)!} \\ &= \frac{1}{2} \left( \frac{1}{2^M} \frac{(M-1)!}{i!(M-i-1)!} + \frac{1}{2^M} \frac{(M-1)!}{(i-1)!(M-i)!} \right) + \frac{1}{2^{M+1}} \binom{M}{i} \\ &= \frac{1}{2^{M+1}} \binom{M-1}{i} + \frac{1}{2^{M+1}} \binom{M-1}{i-1} + \frac{1}{2^{M+1}} \binom{M}{i} \\ &= \frac{1}{2^M} \binom{M}{i}, \end{aligned}$$

which is also known as [Pascal's triangle](#).

- h) The chain is positive recurrent, irreducible and aperiodic, therefore by Theorem 7.8 it admits a limiting distribution equal to  $\pi$ .  
i) We have

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}[X_n \mid X_0 = i] &= \lim_{n \rightarrow \infty} \sum_{j=0}^M j \mathbb{P}(X_n = j \mid X_0 = i) \\ &= \sum_{j=0}^M j \lim_{n \rightarrow \infty} \mathbb{P}(X_n = j \mid X_0 = i) = \sum_{j=0}^M j \pi_j \\ &= \frac{1}{2^M} \sum_{j=0}^M j \binom{M}{j} = \frac{1}{2^M} \sum_{j=0}^M j \frac{M!}{j!(M-j)!} \\ &= \frac{M}{2^M} \sum_{j=1}^M \frac{(M-1)!}{(j-1)!(M-j)!} = \frac{M}{2^M} \sum_{j=0}^{M-1} \frac{(M-1)!}{j!(M-1-j)!} = \frac{M}{2}, \end{aligned}$$

independently of  $i = 0, 1, \dots, M$ .

- j) Clearly, the relation

$$\mathbb{E} \left[ X_0 - \frac{M}{2} \mid X_0 = i \right] = \left( i - \frac{M}{2} \right)$$

holds when  $n = 0$ . Next, assuming that the relation holds at the rank  $n \geq 0$ , we have

$$\begin{aligned} h_{n+1}(i) &= \mathbb{E} \left[ X_{n+1} - \frac{M}{2} \mid X_0 = i \right] \\ &= P_{i,i+1} \mathbb{E} \left[ X_{n+1} - \frac{M}{2} \mid X_1 = i+1 \right] + P_{i,i} \mathbb{E} \left[ X_{n+1} - \frac{M}{2} \mid X_1 = i \right] \\ &\quad + P_{i,i-1} \mathbb{E} \left[ X_{n+1} - \frac{M}{2} \mid X_1 = i-1 \right] \\ &= \left( \frac{1}{2} - \frac{i}{2M} \right) \mathbb{E} \left[ X_{n+1} - \frac{M}{2} \mid X_1 = i+1 \right] + \frac{1}{2} \mathbb{E} \left[ X_{n+1} - \frac{M}{2} \mid X_1 = i \right] \\ &\quad + \frac{i}{2M} \mathbb{E} \left[ X_{n+1} - \frac{M}{2} \mid X_1 = i-1 \right] \\ &= \left( \frac{1}{2} - \frac{i}{2M} \right) \mathbb{E} \left[ X_n - \frac{M}{2} \mid X_0 = i+1 \right] + \frac{1}{2} \mathbb{E} \left[ X_n - \frac{M}{2} \mid X_0 = i \right] \\ &\quad + \frac{i}{2M} \mathbb{E} \left[ X_n - \frac{M}{2} \mid X_0 = i-1 \right] \\ &= \left( \frac{1}{2} - \frac{i}{2M} \right) \left( i+1 - \frac{M}{2} \right) \left( 1 - \frac{1}{M} \right)^n + \frac{1}{2} \left( i - \frac{M}{2} \right) \left( 1 - \frac{1}{M} \right)^n \\ &\quad + \frac{i}{2M} \left( i-1 - \frac{M}{2} \right) \left( 1 - \frac{1}{M} \right)^n = \left( i - \frac{M}{2} \right) \left( 1 - \frac{1}{M} \right)^{n+1}, \quad n \geq 0, \end{aligned}$$

for all  $i = 0, 1, \dots, M$ .

Taking the limit as  $n$  goes to infinity, we get

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ X_n - \frac{M}{2} \mid X_0 = i \right] = \lim_{n \rightarrow \infty} \left( i - \frac{M}{2} \right) \left( 1 - \frac{1}{M} \right)^n = 0,$$

hence

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n \mid X_0 = i] = \frac{M}{2},$$

for all  $i = 0, 1, \dots, M$ , which recovers the result of Question (i).

#### Problem 7.26

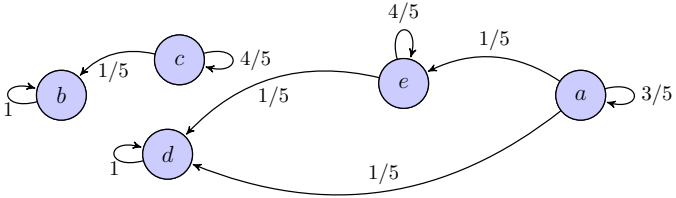
a) The ranking table is completed as follows:

| $\preceq$ | a             | b             | c             | d             | e             |
|-----------|---------------|---------------|---------------|---------------|---------------|
| a         | =             | $\not\preceq$ | $\not\preceq$ | $\preceq$     | $\preceq$     |
| b         | $\not\preceq$ | =             | $\succ$       | $\not\preceq$ | $\not\preceq$ |
| c         | $\not\preceq$ | $\preceq$     | =             | $\not\preceq$ | $\not\preceq$ |
| d         | $\succ$       | $\not\preceq$ | $\not\preceq$ | =             | $\succ$       |
| e         | $\succ$       | $\not\preceq$ | $\not\preceq$ | $\preceq$     | =             |

- 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  $(X_n)_{n \in \mathbb{N}}$  admits the following graph, and is clearly reducible:



- d) Starting from state ①a, ①d or ①e, the limiting distribution is  $(0, 0, 0, 1, 0)$ , starting from state ①b or ①c, the limiting distribution is  $(0, 1, 0, 0, 0)$ , so that although the chain admits limiting distributions, it does *not* admit a limiting distribution independent of the initial state.

More precisely, it can be checked that the power  $P^n$  of order  $n \geq 1$  of the transition matrix  $P$  takes the form

$$P^n = \begin{bmatrix} (3/5)^n & 0 & 0 & 1 - (4/5)^n & a_n \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 - (4/5)^n & (4/5)^n & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 - (4/5)^n & (4/5)^n \end{bmatrix}$$

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

$$\lim_{n \rightarrow \infty} P^n = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

e) The equation  $\pi = \pi P$  reads

$$\begin{aligned} \pi &= [\pi_a, \pi_b, \pi_c, \pi_d, \pi_e] \\ &= \pi \begin{bmatrix} 3/5 & 0 & 0 & 1/5 & 1/5 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1/5 & 4/5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/5 & 4/5 \end{bmatrix} \\ &= [3\pi_a/5, \pi_b + \pi_c/5, 4\pi_c/5, \pi_a/5 + \pi_d + \pi_e/5, \pi_a/5 + 4\pi_e/5], \end{aligned}$$

*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

$$\begin{cases} \pi_a = 0 \\ \pi_c = 0 \\ \pi_e = 0. \end{cases}$$

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.

- 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

$$\begin{aligned} \tilde{\pi} &= \tilde{\pi} \tilde{P} \\ &= \frac{\varepsilon}{n} \tilde{\pi} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} + (1 - \varepsilon) \tilde{\pi} P \end{aligned}$$

$$= \left[ \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5} \right] + (1 - \varepsilon)\tilde{\pi}P.$$

h) The equation

$$\tilde{\pi} = \left[ \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5} \right] + (1 - \varepsilon)\tilde{\pi}P$$

reads

$$\begin{aligned} [\pi_a, \pi_b, \pi_c, \pi_d, \pi_e] &= \left[ \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5} \right] + (1 - \varepsilon)\tilde{\pi} \begin{bmatrix} 3/5 & 0 & 0 & 1/5 & 1/5 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1/5 & 4/5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/5 & 4/5 \end{bmatrix} \\ &= \left[ \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5}, \frac{\varepsilon}{5} \right] \\ &\quad + (1 - \varepsilon) [3\pi_a/5, \pi_b + \pi_c/5, 4\pi_c/5, \pi_a/5 + \pi_d + \pi_e/5, \pi_a/5 + 4\pi_e/5], \end{aligned}$$

*i.e.*

$$\begin{aligned} [0, 0, 0, 0, 0] \\ = [\varepsilon + 3(1 - \varepsilon)\pi_a - 5\pi_a, \varepsilon + 5(1 - \varepsilon)\pi_b - 5\pi_b + (1 - \varepsilon)\pi_c, \varepsilon + 4(1 - \varepsilon)\pi_c - 5\pi_c, \\ \varepsilon + (1 - \varepsilon)\pi_a + 5(1 - \varepsilon)\pi_d - 5\pi_d + (1 - \varepsilon)\pi_e, \varepsilon + (1 - \varepsilon)\pi_a + 4(1 - \varepsilon)\pi_e - 5\pi_e], \end{aligned}$$

*i.e.*

$$\begin{cases} \varepsilon - (2 + 3\varepsilon)\pi_a = 0 \\ \varepsilon - 5\varepsilon\pi_b + (1 - \varepsilon)\pi_c = 0 \\ \varepsilon - \pi_c(1 + 4\varepsilon) = 0 \\ \varepsilon + (1 - \varepsilon)\pi_a - 5\varepsilon\pi_d + (1 - \varepsilon)\pi_e = 0 \\ \varepsilon + (1 - \varepsilon)\pi_a - (1 + 4\varepsilon)\pi_e = 0, \end{cases}$$

*i.e.*

$$\begin{cases} \pi_a = \frac{\varepsilon}{2 + 3\varepsilon}, \\ \pi_b = \frac{2 + 3\varepsilon}{5(1 + 4\varepsilon)}, \\ \pi_c = \frac{\varepsilon}{1 + 4\varepsilon}, \\ \pi_d = \frac{3 + 2\varepsilon}{5(1 + 4\varepsilon)}, \\ \pi_e = \frac{\varepsilon(3 + 2\varepsilon)}{(1 + 4\varepsilon)(2 + 3\varepsilon)}. \end{cases}$$

We also check that

$$\begin{aligned}
& \pi_a + \pi_b + \pi_c + \pi_d + \pi_e \\
&= \frac{\varepsilon}{2+3\varepsilon} + \frac{2+3\varepsilon}{5(1+4\varepsilon)} + \frac{\varepsilon}{1+4\varepsilon} + \frac{3+2\varepsilon}{5(1+4\varepsilon)} + \frac{\varepsilon(3+2\varepsilon)}{(1+4\varepsilon)(2+3\varepsilon)} \\
&= \frac{5\varepsilon(1+4\varepsilon)}{5(2+3\varepsilon)(1+4\varepsilon)} + \frac{(2+3\varepsilon)^2}{5(1+4\varepsilon)(2+3\varepsilon)} + \frac{5\varepsilon(2+3\varepsilon)}{5(1+4\varepsilon)(2+3\varepsilon)} \\
&\quad + \frac{(3+2\varepsilon)(2+3\varepsilon)}{5(2+3\varepsilon)(1+4\varepsilon)} + \frac{5\varepsilon(3+2\varepsilon)}{5(1+4\varepsilon)(2+3\varepsilon)} \\
&= \frac{5\varepsilon(1+4\varepsilon) + (2+3\varepsilon)(5+10\varepsilon) + 5\varepsilon(3+2\varepsilon)}{5(1+4\varepsilon)(2+3\varepsilon)} \\
&= 1.
\end{aligned}$$

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    | c     |
| 5    | a     |

based on the idea that the most visited states should rank higher. In the graph of Figure S.16, the stationary distribution is plotted as a function of  $\varepsilon \in [0, 1]$ .

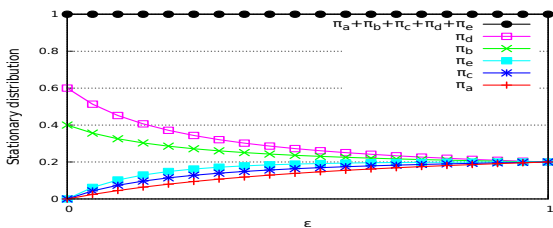


Fig. S.16: Stationary distribution as a function of  $\varepsilon \in [0, 1]$ .

j) By Corollary 7.9, we find

$$\begin{cases} \mu_a(a) = 3 + \frac{2}{\varepsilon} \\ \mu_b(b) = \frac{5(1+4\varepsilon)}{2+3\varepsilon} \\ \mu_c(c) = 4 + \frac{1}{\varepsilon} \\ \mu_d(d) = \frac{5(1+4\varepsilon)}{3+2\varepsilon} \\ \mu_e(e) = \frac{(1+4\varepsilon)(2+3\varepsilon)}{\varepsilon(3+2\varepsilon)} \end{cases}$$

In the graph of Figure S.17 the mean return times are plotted as a function of  $\varepsilon \in [0, 1]$ . A commonly used value in the literature is  $\varepsilon = 1/7$ .

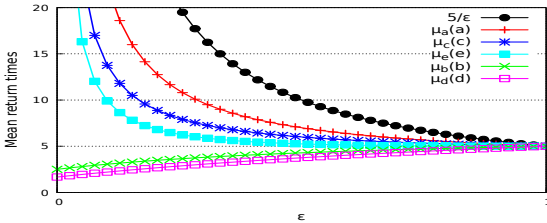


Fig. S.17: Mean return times as functions of  $\varepsilon \in [0, 1]$ .

We note that the ranking of states is clearer for smaller values of  $\varepsilon$ . In particular  $\varepsilon$  cannot be chosen too large, for example taking  $\varepsilon = 1$  makes all mean return times equal and corresponds to a uniform stationary distribution. However, the mean return times can be higher and hence the simulations can take a longer time for small values of  $\varepsilon$ .

#### Problem 7.27

- The cardinality of  $\mathbb{S}$  is  $2^N$ .
- 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.
- We have

$$\mathbb{P}(Z_1 = z \mid Z_0 = z) = 1 - \sum_{k=1}^N \mathbb{P}(Z_1 = \bar{z}^k \mid Z_0 = z)$$

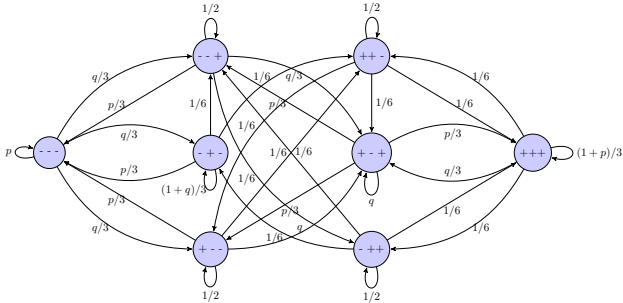
$$= 1 - \frac{1}{N} \sum_{k=1}^N \frac{1}{1 + (p/q)^{z_k(z_{k-1} + z_{k+1})/2}}, \quad z \in \mathbb{S}.$$

d) The transition probability matrix  $P$  of  $(Z_n)_{n \in \mathbb{N}}$  is given by

$$P = \begin{array}{c} \begin{array}{cccccccc} & --- & --+ & -+- & -++ & +-- & +-+ & ++- & +++ \\ \begin{array}{l} --- \\ --+ \\ -+- \\ -++ \\ +-- \\ +-+ \\ ++- \\ +++ \end{array} & \left[ \begin{array}{cccccccc} p & q/3 & q/3 & 0 & q/3 & 0 & 0 & 0 \\ p/3 & 1/2 & 0 & 1/6 & 0 & q/3 & 0 & 0 \\ p/3 & 0 & (1+q)/3 & 1/6 & 0 & 0 & 1/6 & 0 \\ 0 & 1/6 & 1/6 & 1/2 & 0 & 0 & 0 & 1/6 \\ p/3 & 0 & 0 & 0 & 1/2 & q/3 & 1/6 & 0 \\ 0 & p/3 & 0 & 0 & p/3 & q & 0 & p/3 \\ 0 & 0 & 1/6 & 0 & 1/6 & 0 & 1/2 & 1/6 \\ 0 & 0 & 0 & 1/6 & 0 & q/3 & 1/6 & (1+p)/3 \end{array} \right] \end{array} \end{array}.$$

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  $z = (+1, +1, \dots, +1)$  until we reach  $z = (-1, -1, \dots, -1)$ . When  $N = 3$  the chain has the following graph:



From (7.3.5) we note the property

$$\mathbb{P}(Z_1 = \bar{z}^k \mid Z_0 = z) + \mathbb{P}(Z_1 = z \mid Z_0 = \bar{z}^k)$$

$$\begin{aligned}
&= \frac{1}{N(1 + (p/q)^{z_k(z_{k-1} + z_{k+1})/2})} + \frac{1}{N(1 + (p/q)^{-z_k(z_{k-1} + z_{k+1})/2})}, \\
&= \frac{1}{N}, \quad k = 1, 2, \dots, N.
\end{aligned}$$

- 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 6.14. In addition it is aperiodic since every state has a returning loop because

$$\mathbb{P}(Z_1 = z \mid Z_0 = z) \geq \min(p, q) > 0, \quad z \in \mathbb{S}.$$

Hence by *e.g.* Theorem 7.2 it admits a unique stationary distribution which coincides with its limiting distribution.

- f) Under (7.3.7) we have

$$\begin{aligned}
\mathbb{P}(Z_1 = z) &= \mathbb{P}(Z_1 = z \mid Z_0 = z)\pi_z + \sum_{k=1}^N \mathbb{P}(Z_1 = z \mid Z_0 = \bar{z}^k)\pi_{\bar{z}^k} \\
&= \pi_z \mathbb{P}(Z_1 = z \mid Z_0 = z) + \pi_z \sum_{k=1}^N \mathbb{P}(Z_1 = \bar{z}^k \mid Z_0 = z) \\
&= \pi_z \left( \mathbb{P}(Z_1 = z \mid Z_0 = z) + \sum_{k=1}^N \mathbb{P}(Z_1 = \bar{z}^k \mid Z_0 = z) \right) \\
&= \pi_z,
\end{aligned}$$

hence  $(\pi_z)_{z \in \mathbb{S}}$  is a stationary distribution.

- g) By (7.3.5) we have

$$\begin{aligned}
\frac{\mathbb{P}(Z_1 = \bar{z}^k \mid Z_0 = z)}{\mathbb{P}(Z_1 = z \mid Z_0 = \bar{z}^k)} &= \frac{1 + (p/q)^{\bar{z}^k(\bar{z}_{k-1}^k + \bar{z}_{k+1}^k)/2}}{1 + (p/q)^{z_k(z_{k-1} + z_{k+1})/2}} \\
&= \frac{1 + (p/q)^{-z_k(z_{k-1} + z_{k+1})/2}}{1 + (p/q)^{z_k(z_{k-1} + z_{k+1})/2}} \\
&= \frac{q^{z_k(z_{k-1} + z_{k+1})/2} (1 + (q/p)^{z_k(z_{k-1} + z_{k+1})/2})}{q^{z_k(z_{k-1} + z_{k+1})/2} + p^{z_k(z_{k-1} + z_{k+1})/2}} \\
&= \frac{(q/p)^{z_k(z_{k-1} + z_{k+1})/2} (p^{z_k(z_{k-1} + z_{k+1})/2} + q^{z_k(z_{k-1} + z_{k+1})/2})}{q^{z_k(z_{k-1} + z_{k+1})/2} + p^{z_k(z_{k-1} + z_{k+1})/2}} \\
&= \left( \frac{q}{p} \right)^{z_k(z_{k-1} + z_{k+1})/2}, \quad k = 1, 2, \dots, N.
\end{aligned}$$

- h) For all  $z \in \mathbb{S}$  we have

$$\begin{aligned}
\pi_{\bar{z}^k} &= C_\beta \exp \left( \beta \sum_{l=0}^{k-2} z_l z_{l+1} - \beta z_{k-1} z_k - \beta z_k z_{k+1} + \beta \sum_{l=k+1}^N z_l z_{l+1} \right) \\
&= C_\beta \exp \left( -2\beta z_k (z_{k-1} + z_{k+1}) + \beta \sum_{l=0}^N z_l z_{l+1} \right) \\
&= \pi_z e^{-2\beta z_k (z_{k-1} + z_{k+1})} \\
&= \left( \frac{q}{p} \right)^{z_k (z_{k-1} + z_{k+1})/2} \pi_z, \quad k = 1, 2, \dots, N,
\end{aligned}$$

and the *inverse temperature*  $\beta$  is given by

$$\beta = \frac{1}{4} \log \frac{p}{q},$$

*i.e.*

$$p = \frac{1}{1 + e^{-4\beta}}.$$

The constant  $C_\beta$  is chosen so that

$$\sum_{z \in \mathbb{S}} \pi_z = C_\beta \sum_{z \in \mathbb{S}} \exp \left( \beta \sum_{l=0}^N z_l z_{l+1} \right) = 1,$$

*i.e.*

$$C_\beta = \left( \sum_{z \in \mathbb{S}} \exp \left( \beta \sum_{l=0}^N z_l z_{l+1} \right) \right)^{-1}.$$

The stationary distribution  $(\pi_z)_{z \in \mathbb{S}}$  is known as the *Boltzmann distribution*.

i) We have

$$\pi = \begin{bmatrix} \pi_{----} \\ \pi_{---+} \\ \pi_{--+-} \\ \pi_{-++-} \\ \pi_{+--+} \\ \pi_{+-+-} \\ \pi_{++-+} \\ \pi_{++++} \end{bmatrix} = \begin{bmatrix} C_\beta \\ C_\beta \\ C_\beta e^{-4\beta} \\ C_\beta \\ C_\beta \\ C_\beta \\ C_\beta \\ C_\beta e^{4\beta} \end{bmatrix} = C_\beta \begin{bmatrix} 1 \\ 1 \\ q/p \\ 1 \\ 1 \\ 1 \\ 1 \\ p/q \end{bmatrix} = \frac{1}{1 + 4pq} \begin{bmatrix} pq & --- \\ pq & --+ \\ q^2 & -+- \\ pq & -++ \\ pq & +-- \\ pq & +-+ \\ pq & ++- \\ p^2 & +++ \end{bmatrix}$$

and from the relation

$$\pi_{---} + \pi_{--} + \pi_{-+-} + \pi_{-++} + \pi_{+--} + \pi_{+-+} + \pi_{++-} + \pi_{+++} = 1,$$

we find

$$\begin{aligned} C_\beta &= \frac{1}{e^{4\beta} + e^{-4\beta} + 6} \\ &= \frac{1}{4 \cosh^2(2\beta) + 4} \\ &= \frac{pq}{q^2 + p^2 + 6pq} \\ &= \frac{1}{6 + p/q + q/p} \\ &= \frac{pq}{1 + 4pq}. \end{aligned}$$

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

$$\pi_{-++} + \pi_{+-+} + \pi_{++-} + \pi_{+++} = \frac{p(1+2q)}{1+4pq},$$

while the probability of having more “-” than “+” is

$$\pi_{---} + \pi_{--} + \pi_{-+-} + \pi_{+--} = \frac{q(1+2p)}{1+4pq}.$$

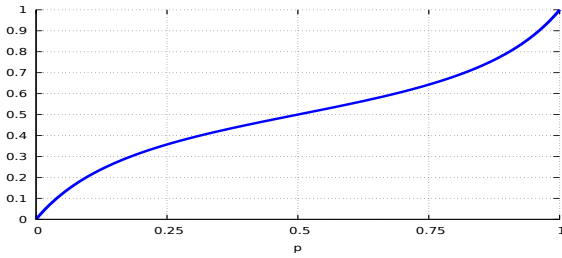


Fig. S.18: Probability of a majority of “+” in the long run as a function of  $p \in [0, 1]$ .

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

$$\pi = \begin{bmatrix} \pi_{---} \\ \pi_{--+} \\ \pi_{-+-} \\ \pi_{-++} \\ \pi_{+-+} \\ \pi_{+--} \\ \pi_{++-} \\ \pi_{+++} \end{bmatrix} = \begin{bmatrix} C_\beta e^{2\beta} \\ C_\beta e^{2\beta} \\ C_\beta e^{-2\beta} \\ C_\beta e^{2\beta} \\ C_\beta e^{-2\beta} \\ C_\beta e^{-2\beta} \\ C_\beta e^{-2\beta} \\ C_\beta e^{2\beta} \end{bmatrix} = C_\beta \begin{bmatrix} \sqrt{p/q} \\ \sqrt{p/q} \\ \sqrt{q/p} \\ \sqrt{p/q} \\ \sqrt{q/p} \\ \sqrt{q/p} \\ \sqrt{q/p} \\ \sqrt{p/q} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} p & --- \\ p & --+ \\ q & -+- \\ p & -++ \\ q & +-+ \\ q & +-- \\ q & ++- \\ p & +++ \end{bmatrix}$$

where

$$C_\beta = \frac{1}{4\sqrt{p/q} + 4\sqrt{q/p}} = \frac{\sqrt{pq}}{4}.$$

The probabilities of having more “+” than “-” in the long run are

$$\pi_{-++} + \pi_{+-+} + \pi_{++-} + \pi_{+++} = \frac{1}{2}$$

while the probability of having more “-” than “+” is also

$$\pi_{---} + \pi_{--+} + \pi_{-+-} + \pi_{-++} = \frac{1}{2}.$$

The study undertaken in [Bhattacharya and Mukherjee \(2018\)](#) 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.*

$$p = \frac{1}{1 + e^{-4\beta}} = 0.968077467.$$

A similar analysis on the 2012 US presidential election shows that  $\beta = 0.96113$ , *i.e.*  $p = 0.979051557$ .

## Chapter 8 - Branching Processes

### Exercise 8.1

a) We have

$$\begin{aligned}\mathbb{P}(X_2 > 0) &= 1 - \mathbb{P}(X_2 = 0) \\ &= 1 - \left( \frac{1}{5} + \frac{3}{5} \times \frac{1}{5} + \frac{1}{5} \times \frac{1}{5} \times \frac{1}{5} \right) \\ &= 1 - \frac{41}{125} = \frac{84}{125}.\end{aligned}$$

b) We have

$$\begin{aligned}\mathbb{P}(X_2 = 1) &= \mathbb{P}(X_2 = 1 \text{ and } X_1 = 0) + \mathbb{P}(X_2 = 1 \text{ and } X_1 = 1) \\ &\quad + \mathbb{P}(X_2 = 1 \text{ and } X_1 = 2) \\ &= \mathbb{P}(X_2 = 1 \mid X_1 = 0)\mathbb{P}(X_1 = 0) \\ &\quad + \mathbb{P}(X_2 = 1 \mid X_1 = 1)\mathbb{P}(X_1 = 1) \\ &\quad + \mathbb{P}(X_2 = 1 \mid X_1 = 2)\mathbb{P}(X_1 = 2) \\ &= 0 \times \frac{1}{5} + \frac{3}{5} \times \frac{3}{5} + 2 \times \frac{1}{5} \times \frac{3}{5} \times \frac{1}{5} \\ &= \frac{51}{125}.\end{aligned}$$

We could also compute

$$\begin{aligned}\mathbb{P}(X_2 = 2) &= \mathbb{P}(X_2 = 2 \text{ and } X_1 = 0) + \mathbb{P}(X_2 = 2 \text{ and } X_1 = 1) \\ &\quad + \mathbb{P}(X_2 = 2 \text{ and } X_1 = 2) \\ &= \mathbb{P}(X_2 = 2 \mid X_1 = 0)\mathbb{P}(X_1 = 0) \\ &\quad + \mathbb{P}(X_2 = 2 \mid X_1 = 1)\mathbb{P}(X_1 = 1) \\ &\quad + \mathbb{P}(X_2 = 2 \mid X_1 = 2)\mathbb{P}(X_1 = 2)\end{aligned}$$

$$\begin{aligned}
&= 0 \times \frac{1}{5} + \frac{1}{5} \times \frac{3}{5} + \frac{1}{5} \times \left(\frac{3}{5}\right)^2 + 2 \times \frac{1}{5} \times \frac{1}{5} \times \frac{1}{5} \\
&= \frac{26}{125},
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{P}(X_2 = 3) &= \mathbb{P}(X_2 = 3 \text{ and } X_1 = 0) + \mathbb{P}(X_2 = 3 \text{ and } X_1 = 1) \\
&\quad + \mathbb{P}(X_2 = 3 \text{ and } X_1 = 2) \\
&= \mathbb{P}(X_2 = 3 \mid X_1 = 0)\mathbb{P}(X_1 = 0) \\
&\quad + \mathbb{P}(X_2 = 3 \mid X_1 = 1)\mathbb{P}(X_1 = 1) \\
&\quad + \mathbb{P}(X_2 = 3 \mid X_1 = 2)\mathbb{P}(X_1 = 2) \\
&= 2 \times \frac{1}{5} \times \frac{1}{5} \times \frac{3}{5} \\
&= \frac{6}{125},
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{P}(X_2 = 4) &= \mathbb{P}(X_2 = 4 \text{ and } X_1 = 0) + \mathbb{P}(X_2 = 4 \text{ and } X_1 = 1) \\
&\quad + \mathbb{P}(X_2 = 4 \text{ and } X_1 = 2) \\
&= \mathbb{P}(X_2 = 4 \mid X_1 = 0)\mathbb{P}(X_1 = 0) \\
&\quad + \mathbb{P}(X_2 = 4 \mid X_1 = 1)\mathbb{P}(X_1 = 1) \\
&\quad + \mathbb{P}(X_2 = 4 \mid X_1 = 2)\mathbb{P}(X_1 = 2) \\
&= \frac{1}{5} \times \frac{1}{5} \times \frac{1}{5} \\
&= \frac{1}{125},
\end{aligned}$$

which recovers

$$\mathbb{P}(X_2 \geq 1) = \frac{51 + 26 + 6 + 1}{125} = \frac{84}{125}.$$

c) We have

$$\begin{aligned}
\mathbb{P}(X_1 = 2 \mid X_2 = 1) &= \frac{\mathbb{P}(X_1 = 2 \text{ and } X_2 = 1)}{\mathbb{P}(X_2 = 1)} \\
&= \mathbb{P}(X_2 = 1 \mid X_1 = 2) \frac{\mathbb{P}(X_1 = 2)}{\mathbb{P}(X_2 = 1)} \\
&= 2 \times \frac{1}{5} \times \frac{3}{5} \times \frac{1}{5} \times \frac{125}{51} = \frac{2}{17}.
\end{aligned}$$

We can also compute

$$\begin{aligned}
\mathbb{P}(X_1 = 1 \mid X_2 = 1) &= \mathbb{P}(X_2 = 1 \mid X_1 = 1) \frac{\mathbb{P}(X_1 = 1)}{\mathbb{P}(X_2 = 1)} \\
&= \frac{3}{5} \times \frac{3/5}{51/125} \\
&= \frac{15}{17},
\end{aligned}$$

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

$$G_1(s) = \mathbb{E}[s^Y] = s^0 \mathbb{P}(Y = 0) + s^1 \mathbb{P}(Y = 1) = \frac{1}{2} + \frac{1}{2}s, \quad s \in \mathbb{R}.$$

b) We prove this statement by induction. Clearly, it holds at the order 1. Next, assuming that (8.3.10) holds at the order  $n \geq 1$  we get

$$\begin{aligned}
G_{n+1}(s) &= G_1(G_n(s)) = G_1\left(1 - \frac{1}{2^n} + \frac{s}{2^n}\right) \\
&= \frac{1}{2} + \frac{1}{2}\left(1 - \frac{1}{2^n} + \frac{s}{2^n}\right) = 1 - \frac{1}{2^{n+1}} + \frac{s}{2^{n+1}}.
\end{aligned}$$

### 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

$$\mathbb{P}(X_n = 0 \mid X_0 = 1) = 1 - \mathbb{P}(X_n = 1 \mid X_0 = 1) = 1 - \frac{1}{2^n},$$

and

$$G_n(s) = \mathbb{P}(X_n = 0 \mid X_0 = 1) + s \mathbb{P}(X_n = 1 \mid X_0 = 1) = 1 - \frac{1}{2^n} + \frac{s}{2^n}.$$

(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

$$G_n(s) = \mathbb{E}[s^{Y_1 + \dots + Y_{X_n}}] = \mathbb{E}[s^{Y_1}] \cdot \dots \cdot \mathbb{E}[s^{Y_{X_n}}].$$

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  $n \geq 2$  we do **not** have  $G_n(s) = (G_1(s))^n$ .

c) We have

$$\mathbb{P}(X_n = 0 \mid X_0 = 1) = G_n(0) = 1 - \frac{1}{2^n}.$$

#### Additional comments:

(i) We may also directly write

$$\begin{aligned} \mathbb{P}(X_n = 0 \mid X_0 = 1) &= 1 - \mathbb{P}(X_n = 1 \mid X_0 = 1) \\ &= 1 - \mathbb{P}(Y_1, Y_2 = 1, \dots, Y_n = 1) \\ &= 1 - (\mathbb{P}(Y_1 = 1))^n \\ &= 1 - \frac{1}{2^n}. \end{aligned}$$

On the other hand, we do **not** have

$$\mathbb{P}(X_n = 0 \mid X_0 = 1) = (\mathbb{P}(Y_1 \neq 0))^n,$$

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

$$\{Y_1 = 0, Y_2 = 0, \dots, Y_n = 0\} \subsetneq \{X_n = 0\},$$

hence

$$\frac{1}{2^n} \leq \mathbb{P}(X_n = 0 \mid X_0 = 1) = 1 - \frac{1}{2^n}, \quad n \geq 1.$$

(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

$$G_n\left(1 - \frac{1}{2^n}\right) = 1 \neq \frac{1}{2^n}$$

does *not* hold for  $n \geq 1$ .

- (iii) In fact,  $\{X_n = 0\}$  means that extinction occurs at time  $n$ , *or* has already occurred *before* time  $n$ .

d) We have

$$\mathbb{E}[X_n \mid X_0 = 1] = G'_n(s)|_{s=1} = (\mathbb{E}[Y_1])^n = \frac{1}{2^n}.$$

**Additional comment:**

In this simple setting we could also write

$$\begin{aligned}\mathbb{E}[X_n \mid X_0 = 1] &= 0 \times \mathbb{P}(X_n = 0 \mid X_0 = 1) + 1 \times \mathbb{P}(X_n = 1 \mid X_0 = 1) \\ &= \mathbb{P}(X_n = 1 \mid X_0 = 1) \\ &= \frac{1}{2^n}.\end{aligned}$$

e) The extinction probability  $\alpha$  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  $(\{X_n = 0\})_{n \geq 1}$  is increasing, we also have

$$\alpha = \mathbb{P}\left(\bigcup_{n \geq 1} \{X_n = 0\}\right) = \lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0) = \lim_{n \rightarrow \infty} 1 - \frac{1}{2^n} = 1.$$

### 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

$$\begin{aligned}G_2(s) &= G_1(G_1(s)) \\ &= G_1(0.2 + 0.5s + 0.3s^2) \\ &= 0.2 + 0.5(0.2 + 0.5s + 0.3s^2) + 0.3(0.2 + 0.5s + 0.3s^2)^2 \\ &= 0.312 + 0.31s + 0.261s^2 + 0.09s^3 + 0.027s^4,\end{aligned}$$

with



$$G_2'(s) = 0.31 + 0.522s + 0.27s^2 + 0.108s^3$$

and

$$G_2''(s) = 0.522 + 0.54s + 0.324s^2,$$

hence

$$G_2'(1) = (G_1'(1))^2 = (1.1)^2 = 1.21 \quad \text{and} \quad G_2''(1) = 1.386,$$

and

$$\mathbb{E}[X_2^2] = G_2''(1) + G_2'(1) = 1.386 + 1.21 = 2.596.$$

By (1.7.8), this yields

$$\text{Var}[X_2] = 2.596 - (1.21)^2.$$

b) We have

$$G_2(s) = 0.312 + 0.31s + 0.261s^2 + 0.09s^3 + 0.027s^4,$$

hence

$$\mathbb{P}(X_2 = 0) = 0.312, \quad \mathbb{P}(X_2 = 1) = 0.31, \quad \mathbb{P}(X_2 = 2) = 0.261,$$

and

$$\mathbb{P}(X_2 = 3) = 0.09, \quad \mathbb{P}(X_2 = 4) = 0.027.$$

c) We have

$$\begin{aligned} \mathbb{P}(X_4 = 0) &= G_4(0) \\ &= G_2(G_2(0)) \\ &= 0.312 + 0.31G_2(0) + 0.261(G_2(0))^2 + 0.09(G_2(0))^3 + 0.027(G_2(0))^4 \\ &= 0.312 + 0.31 \times 0.312 + 0.261 \times 0.312^2 + 0.09 \times 0.312^3 + 0.027 \times 0.312^4 \\ &\simeq 0.44314. \end{aligned}$$

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  $\alpha$  solves the equation

$$\alpha = G_1(\alpha) = 0.2 + 0.5\alpha + 0.3\alpha^2,$$

*i.e.*

$$0.3\alpha^2 - 0.5\alpha + 0.2 = 0.3(\alpha - 1)(\alpha - 2/3) = 0,$$

hence  $\alpha = 2/3$ .

#### Exercise 8.4

a) We have

$$G_1(s) = \mathbb{P}(Y=0) + s\mathbb{P}(Y=1) + s^2\mathbb{P}(Y=2) = as^2 + bs + c, \quad s \in \mathbb{R}.$$

b) Letting  $X_n$  denote the number of individuals in the population at generation  $n \geq 0$ , we have

$$\mathbb{P}(X_2 = 0 \mid X_0 = 1) = G_1(G_1(0)) = G_1(c) = ac^2 + bc + c.$$

This probability can actually be recovered by pathwise analysis, by noting that in order to reach  $\{X_2 = 0\}$  we should have either

- i)  $Y_1 = 0$  with probability  $c$ , or
- ii)  $Y_1 = 1$  with probability  $b$  and then  $Y_1 = 0$  with probability  $c$ ,  
or
- iii)  $Y_1 = 2$  with probability  $a$  and then  $Y_1 = 0$  (two times) with probability  $c$ ,

which yields

$$\mathbb{P}(X_2 = 0 \mid X_0 = 1) = c + bc + ac^2.$$

c) We have

$$\mathbb{P}(X_2 = 0 \mid X_0 = 2) = (\mathbb{P}(X_2 = 0 \mid X_0 = 1))^2 = (ac^2 + bc + c)^2,$$

as in (8.3.1).

d) The extinction probability  $\alpha_1$  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  $\alpha_1$  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  $\alpha_2$  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

$$\min \left( 1, \frac{q^2}{p^2} \right) = \begin{cases} 1 & \text{if } p \leq q, \\ \frac{q^2}{p^2} & \text{if } p \geq q. \end{cases}$$

c) Since  $\{Z_1 = 0\} \subset \{Z_2 = 0\}$ , this probability is

$$\begin{aligned} \mathbb{P}(Z_2 = 0 \text{ and } Z_1 \geq 1) &= \mathbb{P}(Z_2 = 0 \cap \{Z_1 = 0\}^c) \\ &= \mathbb{P}(Z_2 = 0) - \mathbb{P}(Z_1 = 0) \\ &= G_Y(G_Y(0)) - G_Y(0) \\ &= 2pq^3 + q^2p^4. \end{aligned}$$

Alternatively, it can be recovered pathwise as

$$2pq \times q^2 + p^2 \times q^2 \times q^2 = 2pq^3 + p^2q^4.$$

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

$$\begin{aligned} \mathbb{P}(T_0 < \infty) &= \sum_{n \geq 0} \mathbb{P}(T_0 < \infty \text{ and } Z_0 = n) \\ &= \mathbb{P}(Z_0 = n) \sum_{n \geq 0} \mathbb{P}(T_0 < \infty \mid Z_0 = n) \\ &= \sum_{n \geq 0} \left( \frac{q^2}{p^2} \right)^n \mathbb{P}(Z_0 = n) \\ &= e^{-\lambda} \sum_{n \geq 0} \frac{1}{n!} \left( \frac{\lambda q^2}{p^2} \right)^n \\ &= e^{\lambda(q^2/p^2 - 1)} \\ &= e^{\lambda(q^2 - p^2)/p^2} \\ &= e^{\lambda(q-p)/p^2}, \end{aligned}$$

whereas when  $q \geq p$  the extinction probability becomes

$$\sum_{n \geq 0} \mathbb{P}(Z_0 = n) = e^{-\lambda} \sum_{n \geq 0} \frac{\lambda^n}{n!} = 1.$$

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

$$\begin{aligned} \frac{1}{4} \times \left(\frac{1}{4}\right)^2 \times \cdots \times \left(\frac{1}{4}\right)^{2^{n-1}} &= \prod_{k=0}^{n-1} \left(\frac{1}{4}\right)^{2^k} \\ &= \left(\frac{1}{4}\right)^{\sum_{k=0}^{n-1} 2^k} \\ &= \left(\frac{1}{4}\right)^{2^n - 1}, \quad n \geq 0. \end{aligned}$$

- 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

$$\mathbb{P}(Y = 0) = \frac{1}{12}, \quad \mathbb{P}(Y = 1) = \frac{2}{3}, \quad \mathbb{P}(Y = 2) = \frac{1}{4},$$

which has the generating function

$$\begin{aligned} G_1(s) &= \mathbb{P}(Y = 0) + s\mathbb{P}(Y = 1) + s^2\mathbb{P}(Y = 2) \\ &= \frac{1}{12} + \frac{2s}{3} + \frac{s^2}{4} = \frac{1}{12}(1 + 8s + 3s^2), \end{aligned}$$

hence the equation  $G_1(\alpha) = \alpha$  reads

$$3\alpha^2 - 4\alpha + 1 = 3(\alpha - 1)(\alpha - 1/3) = 0,$$

which has  $\alpha = 1/3$  for smallest solution. Consequently, the extinction probability of the culture is equal to  $1/3$ .

- c) The probability that only red cells are generated from time 0 to time  $n$  is

$$\frac{1}{3} \times \left(\frac{1}{3}\right)^2 \times \cdots \times \left(\frac{1}{3}\right)^{2^{n-1}} = \prod_{k=0}^{n-1} \left(\frac{1}{3}\right)^{2^k} = \left(\frac{1}{3}\right)^{2^n - 1},$$

$n \geq 0$ . The probability distribution

$$\mathbf{P}(Y = 0) = \frac{1}{6}, \quad \mathbf{P}(Y = 1) = \frac{2}{2}, \quad \mathbf{P}(Y = 2) = \frac{1}{3},$$

of the number  $Y$  of red cells has the generating function

$$G_1(s) = \mathbf{P}(Y = 0) + s\mathbf{P}(Y = 1) + s^2\mathbf{P}(Y = 2)$$

$$= \frac{1}{6} + \frac{s}{2} + \frac{s^2}{3} = \frac{1}{12}(2 + 6s + 4s^2),$$

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\*

$$\mathbb{E}[T_0 \mid X_0 = k] \leq k \mathbb{E}[T_0 \mid X_0 = 1], \quad k \in \mathbb{N},$$

we have

$$\begin{aligned} \mathbb{E}[T_0 \mid X_0 = 1] &= \sum_{k \geq 0} \mathbb{P}(Y_1 = k) (1 + \mathbb{E}[T_0 \mid X_0 = k]) \\ &= \sum_{k \geq 0} \mathbb{P}(Y_1 = k) + \sum_{k \geq 0} \mathbb{P}(Y_1 = k) \mathbb{E}[T_0 \mid X_0 = k] \\ &= 1 + \sum_{k \geq 0} \mathbb{P}(Y_1 = k) \mathbb{E}[T_0 \mid X_0 = k] \\ &\leq 1 + \mathbb{E}[T_0 \mid X_0 = 1] \sum_{k \geq 0} k \mathbb{P}(Y_1 = k) \\ &= 1 + \mathbb{E}[T_0 \mid X_0 = 1] \mathbb{E}[Y_1], \end{aligned}$$

hence

$$\mathbb{E}[T_0 \mid X_0 = 1] \leq \frac{1}{1 - \mathbb{E}[Y_1]} < \infty.$$

Exercise 8.8 This is a particular case of Example (iv) on page 245.

a) We have

$$G_1(s) = q^2 \sum_{n \geq 1} n(ps)^{n-1} = \frac{q^2}{(1 - ps)^2}.$$

b) The equation  $G_1(\alpha) = \alpha$  reads

$$\alpha = \frac{q^2}{(1 - p\alpha)^2},$$

*i.e.*

$$p^2\alpha^3 - 2p\alpha^2 + \alpha - q^2 = 0,$$

which is known to admit  $\alpha = 1$  for solution, *i.e.*

---

\* This inequality follows from the relation  $\mathbb{E}[\max(X_1, \dots, X_n)] \leq \mathbb{E}[X_1 + \dots + X_n]$ .

$$(\alpha - 1)(p^2\alpha^2 - (1 - q^2)\alpha + q^2) = (\alpha - 1)(p^2\alpha^2 - p(1 + q)\alpha + q^2) = 0,$$

whose smallest solution

$$\alpha = \frac{p(1 + q) - \sqrt{p^2(1 + q)^2 - 4p^2q^2}}{2} = p \frac{2 - p - \sqrt{4p - 3p^2}}{2}$$

is the extinction probability.

#### Exercise 8.9

a) We have  $\mathbb{P}(X = k) = (1/2)^{k+1}$ ,  $k \geq 0$ .

b) The probability generating function of  $X$  is given by

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k \geq 0} s^k \mathbb{P}(X = k) = \frac{1}{2} \sum_{k \geq 0} (s/2)^k = \frac{1}{2 - s},$$

$$-1 < s \leq 1.$$

c) The probability we are looking for is

$$\mathbb{P}(X_3 = 0 \mid X_0 = 0) = G_X(G_X(G_X(0))) = \frac{1}{2 - \frac{1}{2 - 1/2}} = \frac{3}{4}.$$

d) Since giving birth to a girl is equivalent to having at least one child, and this happens to every couple with probability  $3/4$ , the probability we are looking for is equal to

$$\frac{1}{4} + \frac{3}{4} \times \frac{1}{4} + \left(\frac{3}{4}\right)^2 \times \frac{1}{4} = \frac{1}{4} \times \frac{1 - (3/4)^3}{1 - 3/4} = 1 - (3/4)^3 = \frac{37}{64} = 0.578125.$$

It can also be recovered from

$$G_Z^{(3)}(s) = G_Z(G_Z(G_Z(s))) = \frac{37}{64} + \frac{27s}{64}$$

at  $s = 0$ , where  $G_Z$  is the probability generating function

$$G_Z(s) = \frac{1}{4} + \frac{3s}{4}.$$

#### Exercise 8.10

a) We have

$$\begin{aligned} G_{Z_1}(s) &= 1 - \frac{q}{1 - p} + qs \sum_{k \geq 1} (ps)^{k-1} \\ &= 1 - \frac{q}{1 - p} + \frac{qs}{1 - ps} \end{aligned}$$

$$= \frac{1-p-q}{1-p} + \frac{qs}{1-ps}.$$

b) We have

$$G'_{Z_1}(s) = \frac{q}{1-ps} + \frac{pq s}{(1-ps)^2},$$

hence

$$\mathbb{E}[Z_1] = G'_{Z_1}(1) = \frac{q}{1-p} + \frac{pq}{(1-p)^2} = \frac{q(1-p) + pq}{(1-p)^2} = \frac{q}{(1-p)^2}.$$

c) The critical value of  $q$  is  $q = (1-p)^2$ .

d) When  $q = (1-p)^2$  we have

$$G_{Z_1}(s) = p + \frac{(1-p)^2 s}{1-ps}, \quad -1 < s < 1.$$

The relation is clearly satisfied at the rank  $k = 1$ . Next, assuming that the relation

$$G_{Z_k}(s) = \frac{kp - (kp + p - 1)s}{1-p + kp - kps}, \quad -1 < s < 1,$$

holds at the rank  $k \geq 1$ , we have

$$\begin{aligned} G_{Z_{k+1}}(s) &= G_{Z_k}(G_{Z_1}(s)) \\ &= \frac{kp - (kp + p - 1)G_{Z_1}(s)}{1-p + kp - kpG_{Z_1}(s)}, \\ &= \frac{kp - (kp + p - 1)\left(p + \frac{(1-p)^2 s}{1-ps}\right)}{1-p + kp - kp\left(p + \frac{(1-p)^2 s}{1-ps}\right)} \\ &= \frac{kp(1-ps) - (kp + p - 1)(p(1-ps) + (1-p)^2 s)}{(1-p + kp)(1-ps) - kp(p(1-ps) + (1-p)^2 s)} \\ &= \frac{kp(1-ps) - (kp + p - 1)(p + (1-2p)s)}{(1-p + kp)(1-ps) - kp(p + (1-2p)s)} \\ &= \frac{(1-p)(k+1)p - (1-p)((k+2)p - 1)s}{(1-p)(kp+1) - (1-p)(k+1)ps} \\ &= \frac{(k+1)p - ((k+1)p + p - 1)s}{1-p + (k+1)p - (k+1)ps}, \quad -1 < s < 1. \end{aligned}$$

Exercise 8.11

a) We have

$$\begin{aligned}
G(s) &= \mathbb{E}[s^X] \\
&= \mathbb{E}[s^{1+X_1+\dots+X_N}] \\
&= s \mathbb{E}\left[\prod_{l=1}^N s^{X_l}\right] \\
&= s \sum_{k \geq 0} \mathbb{E}\left[\prod_{l=1}^N s^{X_l} \mid N = k\right] \mathbb{P}(N = k) \\
&= s \sum_{k \geq 0} \mathbb{E}\left[\prod_{l=1}^k s^{X_l} \mid N = k\right] \mathbb{P}(N = k) \\
&= s \sum_{k \geq 0} \mathbb{E}\left[\prod_{l=1}^k s^{X_l}\right] \mathbb{P}(N = k) \\
&= s \sum_{k \geq 0} \left(\prod_{l=1}^k \mathbb{E}[s^{X_l}]\right) \mathbb{P}(N = k) \\
&= s e^{-\mu} \sum_{k \geq 0} (\mathbb{E}[s^{X_1}])^k \frac{\mu^k}{k!} \\
&= s G_\mu(\mathbb{E}[s^{X_1}]) \\
&= s G_\mu(G(s)), \quad -1 \leq s \leq 1,
\end{aligned}$$

where  $(X_k)_{k \geq 1}$  denotes a sequence of independent copies of  $X$ , see also the recursion in Proposition 8.1 and Relation (13) in [Haight and Breuer \(1960\)](#). This relation can be solved using Lagrange series, see page 145 of [Pólya and Szegő \(1998\)](#), as

$$G(s) = \sum_{n=1}^{\infty} s^n \mathbb{P}(X = n) = \sum_{n=1}^{\infty} s^n e^{-\mu n} \frac{(\mu n)^{n-1}}{n!},$$

where

$$\mathbb{P}(X = n) = e^{-\mu n} \frac{(\mu n)^{n-1}}{n!}, \quad n \geq 1,$$

is the Borel distribution, see also [Finner et al. \(2015\)](#).

b) We have

$$G'(s) = G_\mu(G(s)) + s G'_\mu(s) G'_\mu(G(s))$$

at  $s = 1$ , which gives

$$G'(1) = G_\mu(1) + G'_\mu(1) G'(1) = 1 + \mu G'(1),$$

and

$$\mathbb{E}[X] = \frac{1}{1 - \mu},$$

which is finite if  $\mu < 1$ .

c) Similarly, knowing that  $G''_\mu(1) = \mu^2$ , the relation

$$G''(s) = 2G'(s)G'_\mu(G(s)) + sG''(s)G'_\mu(G(s)) + s(G'(s))^2G''_\mu(G(s))$$

at  $s = 1$  gives

$$\begin{aligned} G''(1) &= 2G'(1)G'_\mu(1) + G''(1)G'_\mu(1) + (G'(1))^2G''_\mu(1) \\ &= \frac{2\mu - \mu^2}{(1 - \mu)^2} + \mu G''(1), \end{aligned}$$

hence

$$G''(1) = \frac{2\mu - \mu^2}{(1 - \mu)^3}$$

and

$$\begin{aligned} \text{Var}[X] &= G''(1) + G'(1) - (G'(1))^2 \\ &= \frac{2\mu - \mu^2}{(1 - \mu)^3} + \frac{1}{1 - \mu} - \frac{1}{(1 - \mu)^2} \\ &= \frac{\mu}{(1 - \mu)^3}, \end{aligned}$$

see § 7.2.2 of [Johnson et al. \(2005\)](#).

Problem 8.12

a) We have

$$\mathbb{E}[Z_n] = \sum_{k=1}^n \mathbb{E}[X_k] = \sum_{k=1}^n \mu^k = \mu \sum_{k=0}^{n-1} \mu^k = \mu \frac{1 - \mu^n}{1 - \mu}, \quad n \in \mathbb{N}.$$

b) We have

$$\mathbb{E}[Z] = \mathbb{E}\left[\sum_{k \geq 1} X_k\right] = \sum_{k \geq 1} \mathbb{E}[X_k] = \mu \sum_{k \geq 0} \mu^k = \frac{\mu}{1 - \mu}, \quad n \in \mathbb{N},$$

provided that  $\mu < 1$ .

c) We have

$$\begin{aligned} H(s) &= \mathbb{E}[s^Z \mid X_0 = 1] \\ &= s^0 \mathbb{P}(Y_1 = 0) + \sum_{k \geq 1} \mathbb{E}[s^Z \mid X_1 = k] \mathbb{P}(Y_1 = k) \\ &= \mathbb{P}(Y_1 = 0) + \sum_{k \geq 1} (\mathbb{E}[s^Z \mid X_1 = 1])^k \mathbb{P}(Y_1 = k) \end{aligned}$$

$$\begin{aligned}
&= \mathbb{P}(Y_1 = 0) + \sum_{k \geq 1} \left( \mathbb{E} \left[ s^{1 + \sum_{i \geq 2} X_i} \mid X_1 = 1 \right] \right)^k \mathbb{P}(Y_1 = k) \\
&= \mathbb{P}(Y_1 = 0) + \sum_{k \geq 1} \left( \mathbb{E} \left[ s^{1 + \sum_{i \geq 1} X_i} \mid X_0 = 1 \right] \right)^k \mathbb{P}(Y_1 = k) \\
&= \mathbb{P}(Y_1 = 0) + \sum_{k \geq 1} \left( \mathbb{E} [s^{1+Z} \mid X_0 = 1] \right)^k \mathbb{P}(Y_1 = k) \\
&= \sum_{k \geq 0} (s \mathbb{E} [s^Z \mid X_0 = 1])^k \mathbb{P}(Y_1 = k) \\
&= G_1(sH(s)).
\end{aligned}$$

d) We have

$$H(s) = G_1(sH(s)) = \frac{1-p}{1-psH(s)},$$

hence

$$psH^2(s) - H(s) + q = 0,$$

and

$$H(s) = \frac{1 \pm \sqrt{1-4pqs}}{2ps} = \frac{1 - \sqrt{1-4pqs}}{2ps},$$

where we have chosen the minus sign since the plus sign leads to  $H(0) = +\infty$  whereas we should have  $H(0) = \mathbb{P}(Z = 0) \leq 1$ . In addition we have  $\mu = p/q < 1$  hence  $p < 1/2 < q$  and the minus sign gives

$$H(1) = \frac{1 - \sqrt{1-4pq}}{2p} = \frac{1 - |q-p|}{2p} = 1.$$

e) We have

$$\lim_{s \searrow 0^+} H(s) = \lim_{s \searrow 0^+} \frac{1 - (1-2pqs)}{2ps} = q = \mathbb{P}(Z = 0) = \mathbb{P}(Y_1 = 0) = H(0).$$

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

$$\begin{aligned}
H'(1) &= \frac{pq}{p\sqrt{1-4pq}} - \frac{1 - \sqrt{1-4pq}}{2p} = \frac{pq}{p(q-p)} - \frac{1 - (q-p)}{2p} \\
&= \frac{q}{q-p} - 1 = \frac{p}{q-p} = \frac{\mu}{1-\mu},
\end{aligned}$$

with  $\mu = p/q$  for  $p < 1/2$ , which shows that

$$\mathbb{E}[Z] = \frac{\mu}{1 - \mu}$$

and recovers the result of Question (b).

g) We have

$$\begin{aligned} \mathbb{E} \left[ \sum_{k=1}^Z U_k \right] &= \sum_{n \geq 0} \mathbb{E} \left[ \sum_{k=1}^Z U_k \mid Z = n \right] \mathbb{P}(Z = n) \\ &= \sum_{n \geq 0} \mathbb{E} \left[ \sum_{k=1}^n U_k \right] \mathbb{P}(Z = n) \\ &= \sum_{n \geq 0} n \mathbb{E}[U_1] \mathbb{P}(Z = n) \\ &= \mathbb{E}[U_1] \mathbb{E}[Z] = \mathbb{E}[U_1] \frac{\mu}{1 - \mu}. \end{aligned}$$

h) We have

$$\begin{aligned} \mathbb{P}(U_k < x, k = 1, 2, \dots, Z) &= \sum_{n \geq 0} \mathbb{P}(U_k < x, k = 1, 2, \dots, Z \mid Z = n) \mathbb{P}(Z = n) \\ &= \sum_{n \geq 0} \mathbb{P}(U_k < x, k = 1, 2, \dots, n) \mathbb{P}(Z = n) \\ &= \sum_{n \geq 0} \mathbb{P}(U_1 < x) \times \dots \times \mathbb{P}(U_n < x) \mathbb{P}(Z = n) \\ &= \sum_{n \geq 0} (\mathbb{P}(U_1 < x))^n \mathbb{P}(Z = n) = \sum_{n \geq 0} (F(x))^n \mathbb{P}(Z = n) = H(F(x)), \end{aligned}$$

under the convention that the condition is satisfied by default when  $Z = 0$ .

**Remark.** We could also compute the same probability given that  $Z \geq 1$ , and this would give

$$\begin{aligned} &\mathbb{P}(U_k < x, k = 1, 2, \dots, Z \mid Z \geq 1) \\ &= \frac{1}{\mathbb{P}(Z \geq 1)} \sum_{n \geq 1} \mathbb{P}(U_k < x, k = 1, 2, \dots, Z \mid Z = n) \mathbb{P}(Z = n) \\ &= \frac{1}{\mathbb{P}(Z \geq 1)} \sum_{n \geq 1} \mathbb{P}(U_k < x, k = 1, 2, \dots, n) \mathbb{P}(Z = n) \\ &= \frac{1}{\mathbb{P}(Z \geq 1)} \sum_{n \geq 1} \mathbb{P}(U_1 < x)^n \mathbb{P}(Z = n) = \sum_{n \geq 1} F(x)^n \frac{\mathbb{P}(Z = n)}{\mathbb{P}(Z \geq 1)} \end{aligned}$$

$$= \frac{1}{\mathbb{P}(Z \geq 1)}(H(F(x)) - \mathbb{P}(Z = 0)) = \frac{1}{p}(H(F(x)) - q).$$

i) We have

$$\mathbb{E} \left[ \sum_{k=1}^Z U_k \right] = \mathbb{E}[U_1] \frac{\mu}{1-\mu} = \frac{\mu}{1-\mu} = \frac{p}{q-p}.$$

We find

$$\begin{aligned} \mathbb{P}(U_k < x, k = 1, 2, \dots, Z) &= H(F(x)) \\ &= H(1 - e^{-x}) \\ &= \frac{1 - \sqrt{1 - 4pq(1 - e^{-x})}}{2p(1 - e^{-x})}. \end{aligned}$$

## Chapter 9 - Continuous-Time Markov Chains

**Exercise 9.1** We model the number of operating machines as a birth and death process  $(X_t)_{t \in \mathbb{R}_+}$  on the state space  $\{0, 1, 2, 3, 4, 5\}$ .

A new machine can only be added at the rate  $\lambda$  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  $\tau_i$  of machine  $i \in \{1, 2, \dots, k\}$  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  $\min(\tau_1, \tau_2, \dots, \tau_k)$ , and we have

$$\begin{aligned} \mathbb{P}(\min(\tau_1, \tau_2, \dots, \tau_k) > t) &= \mathbb{P}(\tau_1 > t, \tau_2 > t, \dots, \tau_k > t) \\ &= \mathbb{P}(\tau_1 > t) \mathbb{P}(\tau_2 > t) \cdots \mathbb{P}(\tau_k > t) = (e^{-\mu t})^k = e^{-k\mu t}, \end{aligned}$$

$t \geq 0$ , hence the time until the first machine failure is exponentially distributed with parameter  $k\mu$ , *i.e.* the birth rate  $\mu_k$  of  $(X_t)_{t \in \mathbb{R}_+}$  is  $\mu_k = k\mu$ ,  $k = 1, 2, 3, 4, 5$ .

Consequently, the infinitesimal generator  $Q$  of  $(X_t)_{t \in \mathbb{R}_+}$  is given by

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & 0 & 0 \\ \mu & -\mu - \lambda & \lambda & 0 & 0 & 0 \\ 0 & 2\mu & -2\mu - \lambda & \lambda & 0 & 0 \\ 0 & 0 & 3\mu & -3\mu - \lambda & \lambda & 0 \\ 0 & 0 & 0 & 4\mu & -4\mu - \lambda & \lambda \\ 0 & 0 & 0 & 0 & 5\mu & -5\mu \end{bmatrix},$$

with  $\lambda = 0.5$  and  $\mu = 0.2$ . We look for a stationary distribution of the form

$$\pi = (\pi_0, \pi_1, \pi_2, \pi_3, \pi_4, \pi_5)$$

by solving  $\pi Q = 0$ , *i.e.*

$$\begin{cases} 0 = -\lambda\pi_0 + \mu\pi_1 \\ 0 = \lambda\pi_0 - (\lambda + \mu)\pi_1 + 2\mu\pi_2 \\ 0 = \lambda\pi_1 - (\lambda + 2\mu)\pi_2 + 3\mu\pi_3 \\ 0 = \lambda\pi_2 - (\lambda + 3\mu)\pi_3 + 4\mu\pi_4 \\ 0 = \lambda\pi_3 - (\lambda + 4\mu)\pi_4 + 5\mu\pi_5 \\ 0 = \lambda\pi_4 - 5\mu\pi_5, \end{cases} \quad \text{which gives} \quad \begin{cases} 0 = -\lambda\pi_0 + \mu\pi_1 \\ 0 = -\lambda\pi_1 + 2\mu\pi_2 \\ 0 = -\lambda\pi_2 + 3\mu\pi_3 \\ 0 = -\lambda\pi_3 + 4\mu\pi_4 \\ 0 = -\lambda\pi_4 + 5\mu\pi_5, \end{cases}$$

hence

$$\pi_1 = \frac{\lambda}{\mu}\pi_0, \quad \pi_2 = \frac{\lambda}{2\mu}\pi_1, \quad \pi_3 = \frac{\lambda}{3\mu}\pi_2, \quad \pi_4 = \frac{\lambda}{4\mu}\pi_3, \quad \pi_5 = \frac{\lambda}{5\mu}\pi_4,$$

*i.e.*

$$\pi_1 = \frac{\lambda}{\mu}\pi_0, \quad \pi_2 = \frac{\lambda^2}{2\mu^2}\pi_0, \quad \pi_3 = \frac{\lambda^3}{3!\mu^3}\pi_0, \quad \pi_4 = \frac{\lambda^4}{4!\mu^4}\pi_0, \quad \pi_5 = \frac{\lambda^5}{5!\mu^5}\pi_0,$$

which is a truncated Poisson distribution with

$$\pi_0 + \frac{\lambda}{\mu}\pi_0 + \frac{\lambda^2}{2\mu^2}\pi_0 + \frac{\lambda^3}{3!\mu^3}\pi_0 + \frac{\lambda^4}{4!\mu^4}\pi_0 + \frac{\lambda^5}{5!\mu^5}\pi_0 = 1,$$

hence

$$\begin{aligned} \pi_0 &= \frac{1}{1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2\mu^2} + \frac{\lambda^3}{3!\mu^3} + \frac{\lambda^4}{4!\mu^4} + \frac{\lambda^5}{5!\mu^5}} \\ &= \frac{\mu^5}{\mu^5 + \lambda\mu^4 + \lambda^2\mu^3/2 + \lambda^3\mu^2/3! + \lambda^4\mu/4! + \lambda^5/5!}. \end{aligned}$$

Finally, since  $\pi_5$  is the probability that all 5 machines are operating, the fraction of time the repairman is idle in the long run is

$$\pi_5 = \frac{\lambda^5}{120\mu^5 + 120\lambda\mu^4 + 60\lambda^2\mu^3 + 20\lambda^3\mu^2 + 5\lambda^4\mu + \lambda^5}.$$

Note that if at most two machines can be under repair, the infinitesimal generator  $Q$  of  $(X_t)_{t \in \mathbb{R}_+}$  will become

$$Q = \begin{bmatrix} -2\lambda & 2\lambda & 0 & 0 & 0 & 0 \\ \mu & -\mu - 2\lambda & 2\lambda & 0 & 0 & 0 \\ 0 & 2\mu & -2\mu - 2\lambda & 2\lambda & 0 & 0 \\ 0 & 0 & 3\mu & -3\mu - 2\lambda & 2\lambda & 0 \\ 0 & 0 & 0 & 4\mu & -4\mu - \lambda & \lambda \\ 0 & 0 & 0 & 0 & 5\mu & -5\mu \end{bmatrix}.$$

### Exercise 9.2

- a) Since the time  $\tau_k^R$  spent between two Poisson arrivals n<sup>o</sup>  $k$  and  $k+1$  is an exponentially distributed random variable with parameter  $\lambda_R$ , the probability we are looking for is given by

$$\mathbb{P}(\tau_k^R > t) = \mathbb{P}(N_t^R = 0) = e^{-\lambda_R t},$$

where  $N_t^R$  denotes a Poisson process with intensity  $\lambda_R$ .

- b) This probability is given by

$$\begin{aligned} \mathbb{P}(N_t^W \leq 3) &= \mathbb{P}(N_t^W = 0) + \mathbb{P}(N_t^W = 1) + \mathbb{P}(N_t^W = 2) + \mathbb{P}(N_t^W = 3) \\ &= e^{-\lambda_W t} + \lambda_W t e^{-\lambda_W t} + e^{-\lambda_W t} \frac{\lambda_W^2 t^2}{2} + e^{-\lambda_W t} \frac{\lambda_W^3 t^3}{6}, \end{aligned}$$

where  $N_t^W$  denotes a Poisson process with intensity  $\lambda_W$ .

- 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  $\tau^R$  and  $\tau^W$  are independent exponential random variables with parameters  $\lambda_R$  and  $\lambda_W$ , 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  $\lambda_R$  and  $\lambda_W$ .

- d) This distribution is given by  $\mathbb{P}(N_t^R = k \mid N_t^R + N_t^W = n)$  where  $N_t^R$ ,  $N_t^W$  are independent Poisson random variables with parameters  $\lambda_R t$  and  $\lambda_W t$  respectively. We have

$$\begin{aligned} \mathbb{P}(N_t^R = k \mid N_t^R + N_t^W = n) &= \frac{\mathbb{P}(N_t^R = k \text{ and } N_t^R + N_t^W = n)}{\mathbb{P}(N_t^R + N_t^W = n)} \\ &= \frac{\mathbb{P}(N_t^R = k \text{ and } N_t^W = n - k)}{\mathbb{P}(N_t^R + N_t^W = n)} = \frac{\mathbb{P}(N_t^R = k) \mathbb{P}(N_t^W = n - k)}{\mathbb{P}(N_t^R + N_t^W = n)} \\ &= e^{-\lambda_R t} \frac{(\lambda_R t)^k}{k!} e^{-\lambda_W t} \frac{(\lambda_W t)^{n-k}}{(n-k)!} \left( e^{-\lambda_W t - \lambda_R t} \frac{(\lambda_W + \lambda_R)^n t^n}{n!} \right)^{-1} \\ &= \binom{n}{k} \left( \frac{\lambda_R}{\lambda_R + \lambda_W} \right)^k \left( \frac{\lambda_W}{\lambda_R + \lambda_W} \right)^{n-k}, \quad k = 0, 1, \dots, n, \end{aligned}$$

cf. (S.10) in the [solution](#) of Exercise 1.7.

### Exercise 9.3

- a) The number  $X_t$  of machines operating at time  $t$  is a birth and death process on  $\{0, 1, 2\}$  with infinitesimal generator

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 \\ \mu & -(\lambda + \mu) & \lambda \\ 0 & 2\mu & -2\mu \end{bmatrix}.$$

The stationary distribution  $\pi = (\pi_0, \pi_1, \pi_2)$  is solution of  $\pi Q = 0$ , *i.e.*

$$\begin{cases} 0 = -\lambda\pi_0 + \mu\pi_1 \\ 0 = \lambda\pi_0 - (\lambda + \mu)\pi_1 + 2\mu\pi_2 \\ 0 = \lambda\pi_1 - 2\mu\pi_2 \end{cases}$$

under the condition  $\pi_0 + \pi_1 + \pi_2 = 1$ , which yields

$$(\pi_0, \pi_1, \pi_2) = \left( \frac{2\mu^2}{2\mu^2 + 2\lambda\mu + \lambda^2}, \frac{2\lambda\mu}{2\mu^2 + 2\lambda\mu + \lambda^2}, \frac{\lambda^2}{2\mu^2 + 2\lambda\mu + \lambda^2} \right),$$

*i.e.* the probability that no machine is operating is  $\pi_0 = 2/5$  when  $\lambda = \mu = 1$ .

- b) The number  $X_t$  of machines operating at time  $t$  is now a birth and death process on  $\{0, 1\}$ . The time spent in state ① 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 ① has to be computed using first step analysis on the discrete-time embedded chain. We have

$$\begin{aligned} \mathbb{E}[T_0 \mid X_0 = 1] &= \mathbb{P}(\tau_\lambda < \tau_\mu) \times \left( \frac{1}{\mu} + \mathbb{E}[T_0 \mid X_0 = 1] \right) + \mathbb{P}(\tau_\mu < \tau_\lambda) \times \frac{1}{\mu} \\ &= \frac{1}{\mu} + \mathbb{P}(\tau_\lambda < \tau_\mu) \times \mathbb{E}[T_0 \mid X_0 = 1] \\ &= \frac{1}{\mu} + \frac{\lambda}{\lambda + \mu} \mathbb{E}[T_0 \mid X_0 = 1], \end{aligned}$$

where, by (9.6.2) and (9.6.3) or (1.5.9),  $\mathbb{P}(\tau_\lambda < \tau_\mu) = \lambda/(\lambda + \mu)$  is the probability that an exponential random variable  $\tau_\lambda$  with parameter  $\lambda > 0$  is smaller than another independent exponential random variable  $\tau_\mu$  with parameter  $\mu > 0$ . In other words,  $\mathbb{P}(\tau_\lambda < \tau_\mu)$  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 9.18. This yields

$$\mathbb{E}[T_0 \mid X_0 = 1] = \frac{\lambda + \mu}{\mu^2},$$

hence the corresponding rate is  $\mu^2/(\lambda + \mu)$  and the infinitesimal generator of the chain becomes

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \frac{1}{\mathbb{E}[T_0 \mid X_0 = 1]} & -\frac{1}{\mathbb{E}[T_0 \mid X_0 = 1]} \end{bmatrix} = \begin{bmatrix} -\lambda & \lambda \\ \frac{\mu^2}{\lambda + \mu} & -\frac{\mu^2}{\lambda + \mu} \end{bmatrix}.$$

The stationary distribution  $\pi = (\pi_0, \pi_1)$  is solution of  $\pi Q = 0$ , *i.e.*

$$\begin{cases} 0 = -\lambda\pi_0 + \pi_1 \frac{\mu^2}{\lambda + \mu} \\ 0 = \lambda\pi_0 - \pi_1 \frac{\mu^2}{\lambda + \mu} \end{cases}$$

under the condition  $\pi_0 + \pi_1 = 1$ , which yields

$$(\pi_0, \pi_1) = \left( \frac{\mu^2}{\mu^2 + \lambda\mu + \lambda^2}, \frac{\lambda\mu + \lambda^2}{\mu^2 + \lambda\mu + \lambda^2} \right),$$

*i.e.* the probability that no machine is operating when  $\lambda = \mu = 1$  is  $\pi_0 = 1/3$ .

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

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 \\ \mu & -(\lambda + \mu) & \lambda \\ 0 & \mu & -\mu \end{bmatrix},$$

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

$$(\pi_0, \pi_1, \pi_2) = \left( \frac{\mu^2}{\mu^2 + \lambda\mu + \lambda^2}, \frac{\lambda\mu}{\mu^2 + \lambda\mu + \lambda^2}, \frac{\lambda^2}{\mu^2 + \lambda\mu + \lambda^2} \right),$$

hence

$$\pi_0 = \frac{1}{3}, \quad \pi_1 = \frac{1}{3}, \quad \pi_2 = \frac{1}{3},$$

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

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 \\ 0 & -\lambda & \lambda & 0 \\ 0 & 0 & -\lambda & \lambda \\ \lambda & 0 & 0 & -\lambda \end{bmatrix}.$$

- b) As the chain is irreducible, the limiting distribution  $\pi = [\pi_0, \pi_1, \pi_2, \pi_3]$  is obtained from the stationary distribution, by solving the equation  $\pi Q = 0$ , which yields  $\pi = [\pi_0, \pi_1, \pi_2, \pi_3] = [1/4, 1/4, 1/4, 1/4]$
- c) This mean time is the sum

$$\mathbb{E}[\tau_0] + \mathbb{E}[\tau_1] + \mathbb{E}[\tau_2] + \mathbb{E}[\tau_3] = \frac{4}{\lambda}$$

of the means of four exponential random variables with parameter  $\lambda > 0$ , hence it equals  $4/\lambda$ .

#### Exercise 9.5 Muni Toke (2015)

- 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

$$\begin{bmatrix} -5\lambda & 5\lambda & 0 & 0 & 0 & 0 & \cdots \\ \mu + \theta & -\mu - \theta - 5\lambda & 5\lambda & 0 & 0 & 0 & \cdots \\ 0 & 2\mu + \theta & -2\mu - \theta - 5\lambda & 5\lambda & 0 & 0 & \cdots \\ 0 & 0 & 3\mu + \theta & -3\mu - \theta - 5\lambda & 5\lambda & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

- b) Given that the mean number of issued order until time  $t$  is  $\lambda t$  and the mean number of cancelled order until time  $t$  is  $\mu t$ , 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  $\{1, 2, 3, \dots\}$  with state-dependent rate  $\lambda_k = (1+k)^\rho$ ,  $k \geq 1$ . Let us denote by  $\tau_k$  the time spent at state  $k \in \mathbb{N}$  between two increases, which is an exponentially distributed random variable with parameter  $\lambda_k$ . The time it takes for the crack length to grow to infinity is  $\sum_{k \geq 1} \tau_k$ . It is known that  $\sum_{k \geq 1} \tau_k < \infty$  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

$$\mathbb{E} \left[ \sum_{k \geq 1} \tau_k \right] = \sum_{k \geq 1} \mathbb{E}[\tau_k] = \sum_{k \geq 1} \frac{1}{\lambda_k} = \sum_{k \geq 1} \frac{1}{(1+k)^\rho}.$$

By comparison with the integral of the function  $x \mapsto 1/(1+x)^\rho$  we get

$$\begin{aligned} \mathbb{E} \left[ \sum_{k \geq 1} \tau_k \right] &= \sum_{k \geq 1} \frac{1}{(1+k)^\rho} \\ &\leq \sum_{k \geq 1} \int_{k-1}^k \frac{1}{(1+x)^\rho} dx \\ &\leq \int_0^\infty \frac{1}{(1+x)^\rho} dx \\ &= \frac{1}{1-\rho} [(1+x)^{1-\rho}]_0^\infty \\ &= \frac{1}{\rho-1} < \infty, \end{aligned}$$

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

$$\begin{aligned} \mathbb{E} \left[ \sum_{k \geq 1} \tau_k \right] &= \sum_{k \geq 1} \frac{1}{(1+k)^\rho} \\ &\geq \sum_{k \geq 1} \int_k^{k+1} \frac{1}{(1+x)^\rho} dx \\ &\geq \int_1^\infty \frac{1}{(1+x)^\rho} dx \\ &= \infty, \end{aligned}$$

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  $\tau_k$  is the (random) crack length increase at each step. In this case, the relation

$$\mathbb{E} \left[ \sum_{k \geq 1} \tau_k \right] < \infty$$

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.*

$$\mathbb{E}[T_3] = \mathbb{E}[\tau_0] + \mathbb{E}[\tau_1] + \mathbb{E}[\tau_2] = \frac{3}{\lambda} = 30 \text{ minutes.}$$

- b) This probability is

$$\begin{aligned} \mathbb{P}(N_{60} < 3) &= \mathbb{P}(N_{60} = 0) + \mathbb{P}(N_{60} = 1) + \mathbb{P}(N_{60} = 2) \\ &= e^{-60\lambda}(1 + 60\lambda + (60\lambda)^2/2) \\ &= 25 e^{-6} \simeq 0.062. \end{aligned}$$

### Exercise 9.8

- a) The number of users checking the website at time  $t$  consists in the number  $N_t - N_{t-\min(\theta, t)}$  of users who have connected between time  $t - \min(\theta, t)$  and time  $t$ , hence its average is given by

$$\mathbb{E}[N_t - N_{t-\min(\theta, t)}] = \lambda(t - (t - \min(\theta, t))) = \lambda \min(\theta, t).$$

- b) The event that no user is checking the website at time  $t$  can be written as  $\{N_t - N_{\max(\theta, t)} = 0\}$ , and its probability is given by

$$\mathbb{P}(N_t - N_{\max(\theta, t)} = 0) = e^{-\lambda(t - (t - \min(\theta, t)))} = e^{-\lambda \min(\theta, t)}.$$

### Exercise 9.9

- a) By the independence of increments of the Poisson process  $(N_t)_{t \in \mathbb{R}_+}$  we find

$$\begin{aligned} \mathbb{P}(N_{t_3} = 5 \mid N_{t_1} = 1) &= \frac{\mathbb{P}(N_{t_3} = 5 \text{ and } N_{t_1} = 1)}{\mathbb{P}(N_{t_1} = 1)} \\ &= \frac{\mathbb{P}(N_{t_3} - N_{t_1} = 4 \text{ and } N_{t_1} = 1)}{\mathbb{P}(N_{t_1} = 1)} = \frac{\mathbb{P}(N_{t_3} - N_{t_1} = 4) \mathbb{P}(N_{t_1} = 1)}{\mathbb{P}(N_{t_1} = 1)} \\ &= \mathbb{P}(N_{t_3} - N_{t_1} = 4) = \frac{(\lambda(t_3 - t_1))^4}{4!} e^{-\lambda(t_3 - t_1)}. \end{aligned}$$

- b) We expand  $N_{t_4}$  into the telescoping sum

$$N_{t_4} = (N_{t_4} - N_{t_3}) + (N_{t_3} - N_{t_2}) + (N_{t_2} - N_{t_1}) + (N_{t_1} - N_0)$$

of independent increments on disjoint time intervals, to obtain

$$\begin{aligned}
& \mathbb{E}[N_{t_1} N_{t_4} (N_{t_3} - N_{t_2})] \\
&= \mathbb{E}[N_{t_1} (N_{t_4} - N_{t_3}) (N_{t_3} - N_{t_2})] + \mathbb{E}[N_{t_1} (N_{t_3} - N_{t_2}) (N_{t_3} - N_{t_2})] \\
&\quad + \mathbb{E}[N_{t_1} (N_{t_2} - N_{t_1}) (N_{t_3} - N_{t_2})] + \mathbb{E}[N_{t_1} N_{t_1} (N_{t_3} - N_{t_2})] \\
&= \mathbb{E}[N_{t_1}] \mathbb{E}[N_{t_4} - N_{t_3}] \mathbb{E}[N_{t_3} - N_{t_2}] + \mathbb{E}[N_{t_1}] \mathbb{E}[(N_{t_3} - N_{t_2})^2] \\
&\quad + \mathbb{E}[N_{t_1}] \mathbb{E}[N_{t_2} - N_{t_1}] \mathbb{E}[N_{t_3} - N_{t_2}] + \mathbb{E}[N_{t_1}^2] \mathbb{E}[N_{t_3} - N_{t_2}] \\
&= \lambda^3 t_1 (t_4 - t_3) (t_3 - t_2) + \lambda^2 t_1 (t_3 - t_2) (1 + \lambda(t_3 - t_2)) \\
&\quad + \lambda^3 t_1 (t_2 - t_1) (t_3 - t_2) + \lambda^2 t_1 (1 + \lambda t_1) (t_3 - t_2).
\end{aligned}$$

c) We have  $\{T_2 > t\} = \{N_t \leq 1\}$ ,  $t \geq 0$ , hence

$$\mathbb{E}[N_{t_2} \mid T_2 > t_1] = \mathbb{E}[N_{t_2} \mid N_{t_1} \leq 1] = \frac{1}{\mathbb{P}(N_{t_1} \leq 1)} \mathbb{E}[N_{t_2} \mathbb{1}_{\{N_{t_1} \leq 1\}}],$$

by (1.6.6). Now, using the independence of increments between  $N_{t_2} - N_{t_1}$  and  $N_{t_1}$ , we have

$$\begin{aligned}
\mathbb{E}[N_{t_2} \mathbb{1}_{\{N_{t_1} \leq 1\}}] &= \mathbb{E}[(N_{t_2} - N_{t_1}) \mathbb{1}_{\{N_{t_1} \leq 1\}}] + \mathbb{E}[N_{t_1} \mathbb{1}_{\{N_{t_1} \leq 1\}}] \\
&= \mathbb{E}[N_{t_2} - N_{t_1}] \mathbb{E}[\mathbb{1}_{\{N_{t_1} \leq 1\}}] + \mathbb{E}[N_{t_1} \mathbb{1}_{\{N_{t_1} \leq 1\}}] \\
&= \mathbb{E}[N_{t_2} - N_{t_1}] \mathbb{P}(N_{t_1} \leq 1) + \mathbb{E}[N_{t_1} \mathbb{1}_{\{N_{t_1} \leq 1\}}],
\end{aligned}$$

hence

$$\begin{aligned}
\mathbb{E}[N_{t_2} \mid T_2 > t_1] &= \mathbb{E}[N_{t_2} \mid N_{t_1} \leq 1] = \frac{1}{\mathbb{P}(N_{t_1} \leq 1)} \mathbb{E}[N_{t_2} \mathbb{1}_{\{N_{t_1} \leq 1\}}] \\
&= \frac{1}{\mathbb{P}(N_{t_1} \leq 1)} \mathbb{E}[(N_{t_2} - N_{t_1} + N_{t_1}) \mathbb{1}_{\{N_{t_1} \leq 1\}}] \\
&= \frac{1}{\mathbb{P}(N_{t_1} \leq 1)} \mathbb{E}[(N_{t_2} - N_{t_1}) \mathbb{1}_{\{N_{t_1} \leq 1\}}] + \frac{1}{\mathbb{P}(N_{t_1} \leq 1)} \mathbb{E}[N_{t_1} \mathbb{1}_{\{N_{t_1} \leq 1\}}] \\
&= \mathbb{E}[N_{t_2} - N_{t_1}] + \frac{1}{\mathbb{P}(N_{t_1} \leq 1)} \mathbb{E}[N_{t_1} \mathbb{1}_{\{N_{t_1} \leq 1\}}] \\
&= \mathbb{E}[N_{t_2} - N_{t_1}] + \frac{1}{\mathbb{P}(N_{t_1} \leq 1)} (0 \times \mathbb{P}(N_{t_1} = 0) + 1 \times \mathbb{P}(N_{t_1} = 1)) \\
&= \mathbb{E}[N_{t_2} - N_{t_1}] + \frac{\mathbb{P}(N_{t_1} = 1)}{\mathbb{P}(N_{t_1} \leq 1)} = \lambda(t_2 - t_1) + \frac{\lambda t_1 e^{-\lambda t_1}}{e^{-\lambda t_1} + \lambda t_1 e^{-\lambda t_1}}.
\end{aligned}$$

Exercise 9.10 The generator of the process is given by

$$Q = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 \\ \mu_1 & -\lambda_1 - \mu_1 & \lambda_1 \\ 0 & \mu_2 & -\mu_2 \end{bmatrix} = \begin{bmatrix} -2\alpha & 2\alpha & 0 \\ \beta & -\alpha - \beta & \alpha \\ 0 & 2\beta & -2\beta \end{bmatrix}.$$

Writing the equation  $\pi Q = 0$  shows that

$$\pi_1 = \frac{2\alpha}{\beta}\pi_0 \quad \text{and} \quad \pi_2 = \frac{\alpha}{2\beta}\pi_1,$$

and the condition

$$\pi_0 + \pi_1 + \pi_2 = 1$$

shows that

$$\pi_0 = \left(1 + \frac{\alpha}{\beta}\right)^{-2}$$

and

$$\pi_0 = \left(\frac{\beta}{\alpha + \beta}\right)^2, \quad \pi_1 = \frac{2\alpha\beta}{(\alpha + \beta)^2} \quad \text{and} \quad \pi_2 = \left(\frac{\alpha}{\alpha + \beta}\right)^2,$$

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 9.11** This is an extension of Exercise 9.10. 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

$$\pi_{k+1} = \frac{\alpha}{\beta} \frac{N-k}{k+1} \pi_k, \quad k = 0, 1, \dots, N-1,$$

and by induction on  $k = 1, 2, \dots, N$  we find  $\pi_1 = N\pi_0\alpha/\beta$  and

$$\pi_2 = \frac{\alpha}{\beta} \frac{(N-1)}{2} \pi_1 = \left(\frac{\alpha}{\beta}\right)^2 \frac{N(N-1)}{1 \times 2} \pi_0, \quad \pi_3 = \left(\frac{\alpha}{\beta}\right)^3 \frac{N(N-1)(N-2)}{1 \times 2 \times 3} \pi_0,$$

hence

$$\begin{aligned}\pi_k &= \left(\frac{\alpha}{\beta}\right)^k \frac{N(N-1)\cdots(N-k+1)}{k!} \pi_0 \\ &= \left(\frac{\alpha}{\beta}\right)^k \frac{N!}{(N-k)!k!} \pi_0 = \left(\frac{\alpha}{\beta}\right)^k \binom{N}{k} \pi_0,\end{aligned}$$

$k = 0, 1, \dots, N$ . The condition

$$\pi_0 + \pi_1 + \cdots + \pi_N = 1$$

shows that

$$1 = \pi_0 \sum_{k=0}^N \left(\frac{\alpha}{\beta}\right)^k \frac{N!}{(N-k)!k!} = \left(1 + \frac{\alpha}{\beta}\right)^N \pi_0,$$

hence

$$\pi_0 = \left(1 + \frac{\alpha}{\beta}\right)^{-N}$$

and we have

$$\begin{aligned}\pi_k &= \left(1 + \frac{\alpha}{\beta}\right)^{-N} \left(\frac{\alpha}{\beta}\right)^k \frac{N!}{(N-k)!k!} \\ &= \left(\frac{\alpha}{\alpha + \beta}\right)^k \left(\frac{\beta}{\alpha + \beta}\right)^{N-k} \frac{N!}{(N-k)!k!}, \quad k = 0, 1, \dots, N,\end{aligned}$$

hence the stationary distribution  $\pi$  is a binomial distribution with parameter  $(N, p) = (N, \alpha/(\alpha + \beta))$ .

**Exercise 9.12** The generator  $Q$  of this pure birth process is given by

$$Q = [\lambda_{i,j}]_{i,j \in \mathbb{N}} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & \cdots \\ 0 & -3 & 3 & 0 & 0 & \cdots \\ 0 & 0 & -2 & 2 & 0 & \cdots \\ 0 & 0 & 0 & -5 & 5 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix},$$

hence the forward Kolmogorov equation  $P'(t) = P(t)Q$  reads

$$\begin{bmatrix} P'_{0,0}(t) & P'_{0,1}(t) & P'_{0,2}(t) & P'_{0,3}(t) & \cdots \\ P'_{1,0}(t) & P'_{1,1}(t) & P'_{1,2}(t) & P'_{1,3}(t) & \cdots \\ P'_{2,0}(t) & P'_{2,1}(t) & P'_{2,2}(t) & P'_{2,3}(t) & \cdots \\ P'_{3,0}(t) & P'_{3,1}(t) & P'_{3,2}(t) & P'_{3,3}(t) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$= \begin{bmatrix} P_{0,0}(t) & P_{0,1}(t) & P_{0,2}(t) & P_{0,3}(t) & \cdots \\ P_{1,0}(t) & P_{1,1}(t) & P_{1,2}(t) & P_{1,3}(t) & \cdots \\ P_{2,0}(t) & P_{2,1}(t) & P_{2,2}(t) & P_{2,3}(t) & \cdots \\ P_{3,0}(t) & P_{3,1}(t) & P_{3,2}(t) & P_{3,3}(t) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \times \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & \cdots \\ 0 & -3 & 3 & 0 & 0 & \cdots \\ 0 & 0 & -2 & 2 & 0 & \cdots \\ 0 & 0 & 0 & -5 & 5 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix},$$

which yields

$$\begin{cases} P'_{0,0}(t) = -P_{0,0}(t), \\ P'_{0,1}(t) = P_{0,0}(t) - 3P_{0,1}(t), \\ P'_{0,2}(t) = 3P_{0,1}(t) - 2P_{0,2}(t), \\ P'_{0,3}(t) = 2P_{0,2}(t) - 5P_{0,3}(t). \end{cases}$$

The first equation is solved by (A.9) as

$$P_{0,0}(t) = P_{0,0}(0)e^{-t} = e^{-t}, \quad t \geq 0,$$

and this solution can be easily recovered from

$$P_{0,0}(t) = \mathbb{P}(X_t = 0 \mid X_0 = 0) = \mathbb{P}(\tau_0 > t) = e^{-t}, \quad t \geq 0.$$

The second equation becomes

$$P'_{0,1}(t) = e^{-t} - 3P_{0,1}(t), \quad (\text{S.50})$$

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 (S.50), we get

$$-C_2 e^{-t} = e^{-t} - 3C_2 e^{-t},$$

*i.e.*  $C_2 = 1/2$ , hence the general solution

$$P_{0,1}(t) = \frac{1}{2}e^{-t} + C_1 e^{-3t},$$

and from the initial condition  $P_{0,1}(0) = 0$  we get  $C_1 = -1/2$ , *i.e.*

$$P_{0,1}(t) = \frac{1}{2}(e^{-t} - e^{-3t}), \quad t \geq 0.$$

The remaining equations can be solved similarly by searching for a suitable particular solution. For

$$P'_{0,2}(t) = \frac{3}{2}(e^{-t} - e^{-3t}) - 2P_{0,2}(t),$$

we find, searching for a particular solution of the form  $t \mapsto ae^{-t} + be^{-3t}$ ,

$$P_{0,2}(t) = \frac{3}{2}e^{-3t}(1 - e^t)^2, \quad t \geq 0, \quad (\text{S.51})$$

see [here](#), and for

$$P'_{0,3}(t) = 3e^{-3t}(1 - e^t)^2 - 5P_{0,3}(t),$$

we find

$$P_{0,3}(t) = \frac{1}{4}e^{-5t}(e^t - 1)^3(1 + 3e^t), \quad t \geq 0, \quad (\text{S.52})$$

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

$$\begin{aligned} P_{0,1}(t) &= \mathbb{P}(X_t = 1 \mid X_0 = 0) \\ &= \mathbb{P}(\tau_0 < t \text{ and } \tau_0 + \tau_1 > t) \\ &= \mathbb{P}(\tau_0 < t < \tau_0 + \tau_1) \\ &= \lambda_0 \lambda_1 \int_{\{(x,y) : 0 < x < t, x+y > t\}} e^{-\lambda_0 x} e^{-\lambda_1 y} dx dy \\ &= \lambda_0 \lambda_1 \int_0^t \int_t^\infty e^{-\lambda_0 x} e^{-\lambda_1(z-x)} dx dz \\ &= \lambda_0 \lambda_1 \int_0^t \int_t^\infty e^{-(\lambda_0 - \lambda_1)x} e^{-\lambda_1 z} dx dz \\ &= e^{-\lambda_1 t} \lambda_0 \int_0^t e^{-(\lambda_0 - \lambda_1)x} dx \\ &= \frac{\lambda_0}{\lambda_0 - \lambda_1} (e^{-\lambda_1 t} - e^{-\lambda_0 t}) \\ &= \frac{1}{2} (e^{-t} - e^{-3t}), \quad t \geq 0. \end{aligned} \quad (\text{S.53})$$

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 (S.51) can be recovered via a triple integral. Similarly, (S.52) can be approached by a quadruple integral, etc.

**Exercise 9.13** Noting that the two events

$$\{T_1 > t, T_2 > t + s\} = \{X_t = 0, 0 \leq X_{t+s} \leq 1\}$$

coincides for all  $s, t \geq 0$ , we find that

$$\begin{aligned}
\mathbb{P}(T_1 > t \text{ and } T_2 > t + s) &= \mathbb{P}(X_t = 0 \text{ and } X_{t+s} \in \{0, 1\} \mid X_0 = 0) \\
&= \mathbb{P}(X_{t+s} \in \{0, 1\} \mid X_t = 0) \mathbb{P}(X_t = 0 \mid X_0 = 0) \\
&= \mathbb{P}(X_s \in \{0, 1\} \mid X_0 = 0) \mathbb{P}(X_t = 0 \mid X_0 = 0) \\
&= (\mathbb{P}(X_s = 0 \mid X_0 = 0) + \mathbb{P}(X_s = 1 \mid X_0 = 0)) \mathbb{P}(X_t = 0 \mid X_0 = 0) \\
&= P_{0,0}(t)(P_{0,0}(s) + P_{0,1}(s)).
\end{aligned}$$

Next, we note that we have

$$P_{0,0}(t) = e^{-\lambda_0 t}, \quad t \geq 0,$$

and, from Relation (S.53) above,

$$P_{0,1}(t) = \frac{\lambda_0}{\lambda_1 - \lambda_0} (e^{-\lambda_0 t} - e^{-\lambda_1 t}), \quad t \geq 0,$$

hence

$$\begin{aligned}
\mathbb{P}(T_1 > t \text{ and } T_2 > t + s) &= e^{-\lambda_0(t+s)} + \frac{\lambda_0}{\lambda_1 - \lambda_0} (e^{-\lambda_0(t+s)} - e^{-\lambda_0 t - \lambda_1 s}) \\
&= \frac{\lambda_1}{\lambda_1 - \lambda_0} e^{-\lambda_0(t+s)} - \frac{\lambda_0}{\lambda_1 - \lambda_0} e^{-\lambda_0 t - \lambda_1 s}.
\end{aligned}$$

Then, since

$$\mathbb{P}(T_1 > x \text{ and } T_2 > y) = \int_x^\infty \int_y^\infty f_{(T_1, T_2)}(u, v) du dv,$$

by (1.5.4) we get

$$\begin{aligned}
f_{(T_1, T_2)}(x, y) &= \frac{\partial^2}{\partial y \partial x} \mathbb{P}(T_1 > x \text{ and } T_2 > y) \\
&= \frac{\partial^2}{\partial y \partial x} \left( \frac{\lambda_1}{\lambda_1 - \lambda_0} e^{-\lambda_0 y} - \frac{\lambda_0}{\lambda_1 - \lambda_0} e^{-\lambda_0 x - \lambda_1(y-x)} \right) \\
&= -\lambda_0 \frac{\partial}{\partial y} e^{-\lambda_0 x - \lambda_1(y-x)} = \lambda_0 \lambda_1 e^{-\lambda_0 x - \lambda_1(y-x)},
\end{aligned}$$

provided that  $y \geq x \geq 0$ . When  $x > y \geq 0$  we have

$$f_{(T_1, T_2)}(x, y) = 0.$$

The density of  $(\tau_0, \tau_1)$  is given under the change of variable  $T_1 = \tau_0$ ,  $T_2 = \tau_0 + \tau_1$ , by

$$f_{(\tau_0, \tau_1)}(s, t) = f_{(T_1, T_2)}(s, s + t) = \lambda_0 \lambda_1 e^{-\lambda_0 s - \lambda_1 t}, \quad s, t \geq 0,$$

which shows that  $\tau_0, \tau_1$  are two independent exponentially distributed random variables with parameters  $\lambda_0$  and  $\lambda_1$ , respectively.

**Exercise 9.14** Let  $(N_t)_{t \in \mathbb{R}_+}$  denote a Poisson process with intensity  $\lambda > 0$ .

a) This probability is equal to

$$\mathbb{P}(N_T = 0) = \mathbb{P}(\tau_0 > T) = e^{-\lambda T}.$$

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 1.10 and (1.6.12))

$$\begin{aligned} \mathbb{E}[\tau_0 \mid \tau_0 < T] &= \frac{1}{\mathbb{P}(\tau_0 < T)} \mathbb{E}[\tau_0 \mathbb{1}_{\{\tau_0 < T\}}] \\ &= \frac{\lambda}{1 - e^{-\lambda T}} \int_0^T x e^{-\lambda x} dx \\ &= \frac{1 - (1 + \lambda T)e^{-\lambda T}}{\lambda(1 - e^{-\lambda T})} \end{aligned}$$

for the first car to pass, after which the process is reinitialized and the average waiting time is again  $t$ . Hence by first step analysis in continuous time we find the equation

$$\begin{aligned} t &= 0 \times \mathbb{P}(N_T = 0) + (\mathbb{E}[\tau_0 \mid \tau_0 < T] + t) \times \mathbb{P}(\tau_0 \leq T) \\ &= \frac{1 - (1 + \lambda T)e^{-\lambda T}}{\lambda} + t(1 - e^{-\lambda T}) \end{aligned} \quad (\text{S.54})$$

with unknown  $t$ , and solution

$$t = \frac{e^{\lambda T} - 1 - \lambda T}{\lambda}.$$

Alternatively, we could also rewrite the above equation (S.54) as

$$t = \mathbb{E}[0 \times \mathbb{1}_{\{\tau_0 \geq T\}} + (\tau_0 + t) \mathbb{1}_{\{\tau_0 < T\}}] = \frac{1 - (1 + \lambda T)e^{-\lambda T}}{\lambda} + t(1 - e^{-\lambda T}).$$

c) Denoting by  $t$  the mean time until she finishes crossing the street we have, by first step analysis in continuous time,

$$\begin{aligned} t &= \mathbb{E}[T \mathbb{1}_{\{T < \tau_0\}} + (\tau_0 + t) \mathbb{1}_{\{T \geq \tau_0\}}] \\ &= \lambda T \int_T^\infty e^{-\lambda s} ds + \lambda \int_0^T (s + t) e^{-\lambda s} ds \end{aligned}$$

$$\begin{aligned}
&= T\mathbb{P}(\tau_0 > T) + \lambda \int_0^T se^{-\lambda s} ds + t\mathbb{P}(\tau_0 < T) \\
&= Te^{-\lambda T} + \frac{1 - e^{-\lambda T}(1 + \lambda T)}{\lambda} + t(1 - e^{-\lambda T}),
\end{aligned}$$

which yields

$$t = \frac{e^{\lambda T} - 1}{\lambda}.$$

Alternatively, we can write

$$\begin{aligned}
t &= T\mathbb{P}(\tau_0 > T) + (\mathbb{E}[\tau_0 \mid \tau_0 > T] + t)\mathbb{P}(\tau_0 > T) \\
&= T\mathbb{P}(\tau_0 > T) + \mathbb{E}[\tau_0 \mathbb{1}_{\{\tau_0 > T\}}] + t\mathbb{P}(\tau_0 > T) \\
&= T\mathbb{P}(\tau_0 > T) + \lambda \int_0^T se^{-\lambda s} ds + t\mathbb{P}(\tau_0 < T),
\end{aligned}$$

which yields the same result. The same conclusion follows by adding the extra crossing time  $T$  to the answer of Question (b).

- d) In this case,  $T$  becomes an independent exponentially distributed random variable with parameter  $\mu > 0$ , hence we can write

$$\begin{aligned}
t &= \mathbb{E}\left[\frac{e^{\lambda T} - 1}{\lambda}\right] \\
&= \mu \int_0^\infty \frac{e^{\lambda u} - 1}{\lambda} e^{-\mu u} du \\
&= \frac{\mu}{\lambda} \int_0^\infty e^{(\lambda - \mu)u} du - \frac{\mu}{\lambda} \int_0^\infty e^{-\mu u} du \\
&= \frac{\mu}{\lambda(\mu - \lambda)} - \frac{1}{\lambda} \\
&= \frac{1}{\mu - \lambda}
\end{aligned}$$

if  $\mu > \lambda$ , with  $t = +\infty$  if  $\mu \leq \lambda$ .

- e) Crossing two lanes at once will require

$$t_1(\lambda) = \frac{e^{2\lambda \times 2T/2} - 1}{2\lambda}$$

units of time on average, whereas making a stop in between will take on average

$$t_2(\lambda) = 2 \times \frac{e^{\lambda T/2} - 1}{\lambda}$$

with  $t_2(\lambda) < t_1(\lambda)$  for all  $\lambda > 0$ , hence it is better to make a stop in between two lanes.

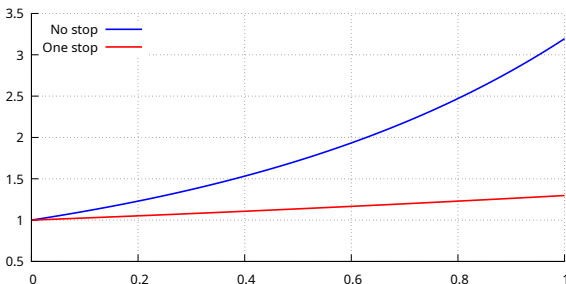


Fig. S.19: Comparison between average waiting times with  $T = 1$ .

### Exercise 9.15

- This is the probability that an exponential random variable with parameter  $\mu$  is less than  $T$ , *i.e.*  $\mathbb{P}(\tau_0 > T) = e^{-\mu T}$ .
- Denoting by  $t = \mathbb{E}[W]$  the mean time until the machine breaks down we have, by first step analysis in continuous time,

$$\begin{aligned}
 t &= \mathbb{E} \left[ T \mathbb{1}_{\{T < \tau_0\}} + (\tau_0 + t) \mathbb{1}_{\{T \geq \tau_0\}} \right] \\
 &= \mu T \int_0^\infty e^{-\mu s} ds + \mu \int_0^T (s + t) e^{-\mu s} ds \\
 &= T \mathbb{P}(\tau_0 > T) + \mu \int_0^T s e^{-\mu s} ds + t \mathbb{P}(\tau_0 < T) \\
 &= T e^{-\mu T} + \frac{1 - e^{-\mu T} (1 + \mu T)}{\mu} + t (1 - e^{-\mu T}),
 \end{aligned}$$

which yields

$$t = \mathbb{E}[W] = \frac{e^{\mu T} - 1}{\mu}.$$

Alternatively, we can write

$$\begin{aligned}
 t &= T \mathbb{P}(\tau_0 > T) + (\mathbb{E}[\tau_0 \mid \tau_0 > T] + t) \mathbb{P}(\tau_0 > T) \\
 &= T \mathbb{P}(\tau_0 > T) + \mathbb{E}[\tau_0 \mathbb{1}_{\{\tau_0 > T\}}] + t \mathbb{P}(\tau_0 > T) \\
 &= T \mathbb{P}(\tau_0 > T) + \mu \int_0^T s e^{-\mu s} ds + t \mathbb{P}(\tau_0 < T),
 \end{aligned}$$

which yields the same result.

- This proportion is

$$\frac{\mathbb{E}[W]}{\mathbb{E}[W] + 1/\lambda} = \frac{e^{\mu T} - 1}{e^{\mu T} - 1 + \mu/\lambda}.$$

The state of the machine can be modeled using a two-state Markov chain with infinitesimal generator

$$Q = \begin{bmatrix} \frac{-\lambda}{\mathbb{E}[W]} & \frac{\lambda}{\mathbb{E}[W]} \\ \frac{\mu}{e^{\mu T} - 1} & \frac{\mu}{1 - e^{\mu T}} \end{bmatrix},$$

and the above proportion can be rewritten using the stationary distribution

$$\begin{aligned} [\pi_0, \pi_1] &= \left[ \frac{1/\lambda}{1/\lambda + (e^{\mu T} - 1)/\mu}, \frac{(e^{\mu T} - 1)/\mu}{1/\lambda + (e^{\mu T} - 1)/\mu} \right] \\ &= \left[ \frac{\mu}{\mu + \lambda(e^{\mu T} - 1)}, \frac{\lambda(e^{\mu T} - 1)}{\mu + \lambda(e^{\mu T} - 1)} \right] \\ &= \left[ \frac{1/\lambda}{\mathbb{E}[W] + 1/\lambda}, \frac{\mathbb{E}[W]}{\mathbb{E}[W] + 1/\lambda} \right]. \end{aligned}$$

#### Exercise 9.16

a) The generator  $Q$  of  $(X_t)_{t \in \mathbb{R}_+}$  is given by

$$Q = \begin{bmatrix} -0.5 & 0.5 & 0 \\ 0.2 & -0.45 & 0.25 \\ 0 & 0.4 & -0.4 \end{bmatrix}.$$

b) Solving for  $\pi Q = 0$  we have

$$\begin{aligned} \pi Q &= [\pi_0, \pi_1, \pi_2] \times \begin{bmatrix} -0.5 & 0.5 & 0 \\ 0.2 & -0.45 & 0.25 \\ 0 & 0.4 & -0.4 \end{bmatrix} \\ &= \begin{bmatrix} -0.5 \times \pi_0 + 0.2 \times \pi_1 \\ 0.5 \times \pi_0 - 0.45 \times \pi_1 + 0.4 \times \pi_2 \\ 0.25 \times \pi_1 - 0.4 \times \pi_2 \end{bmatrix}^T = [0, 0, 0], \end{aligned}$$

*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

$$100 \times \frac{90}{81} = \frac{1000}{9}.$$

e) We have

$$Q = \begin{bmatrix} -0.5 & 0.5 & 0 \\ 0.2 & -0.7 & 0.5 \\ 0 & 0.4 & -0.4 \end{bmatrix},$$

and solving  $\pi Q = 0$  shows that

$$[\pi_0, \pi_1, \pi_2] = \left[ \frac{0.32}{2.12}, \frac{0.8}{2.12}, \frac{1}{2.12} \right] = [0.15094, 0.37736, 0.47170].$$

**Exercise 9.17** Both chains  $(X_1(t))_{t \in \mathbb{R}}$  and  $(X_2(t))_{t \in \mathbb{R}}$  have the same infinitesimal generator

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}.$$

The infinitesimal generator of  $Z(t) := X_1(t) + X_2(t)$  is given by

$$\begin{bmatrix} -2\lambda & 2\lambda & 0 \\ \mu & -\lambda - \mu & \lambda \\ 0 & 2\mu & -2\mu \end{bmatrix},$$

as the birth rate  $\lambda$  is doubled when both chains are in state  $\textcircled{0}$ , and the death rate  $\mu$  is also doubled when both chains are in state  $\textcircled{1}$ .

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

$$\begin{aligned} & \begin{bmatrix} \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0) & \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0) \\ \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 1) & \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 1) \end{bmatrix} \\ &= \begin{bmatrix} \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 0) & \mathbb{P}(X_2(t) = 1 \mid X_2(0) = 0) \\ \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 1) & \mathbb{P}(X_2(t) = 1 \mid X_2(0) = 1) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-t(\lambda + \mu)} & \frac{\lambda}{\lambda + \mu} - \frac{\lambda}{\lambda + \mu} e^{-t(\lambda + \mu)} \\ \frac{\mu}{\lambda + \mu} - \frac{\mu}{\lambda + \mu} e^{-t(\lambda + \mu)} & \frac{\lambda}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} e^{-t(\lambda + \mu)} \end{bmatrix}. \end{aligned}$$

As for the transition semi-group of  $Z(t)$ , we have

$$\begin{aligned} P_{0,0}(t) &= \mathbb{P}(Z(t) = 0 \mid Z(0) = 0) \\ &= \mathbb{P}(X_1(t) = 0 \text{ and } X_2(t) = 0 \mid X_1(0) = 0 \text{ and } X_2(0) = 0) \\ &= \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 0) \\ &= (\mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0))^2 \end{aligned}$$

$$= \left( \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-t(\lambda + \mu)} \right)^2.$$

For  $P_{0,1}(t)$  we have

$$\begin{aligned} P_{0,1}(t) &= \mathbb{P}(Z(t) = 1 \mid Z(0) = 0) \\ &= \mathbb{P}(X_1(t) = 0 \text{ and } X_2(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 0) \\ &\quad + \mathbb{P}(X_1(t) = 1 \text{ and } X_2(t) = 0 \mid X_1(0) = 0 \text{ and } X_2(0) = 0) \\ &= \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 0) \\ &\quad + \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 1 \mid X_2(0) = 0) \\ &= 2\mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 0). \end{aligned}$$

Starting from  $Z(0) = 1$  and ending at  $Z(t) = 1$  we have two possibilities  $(0, 1)$  or  $(1, 0)$  for the terminal condition. As for the initial condition  $Z(0) = 1$  the two possibilities  $(0, 1)$  and  $(1, 0)$  count for one only since they both give  $Z(0) = 1$ . We have

$$\begin{aligned} &\mathbb{P}(Z(t) = 1 \mid Z(0) = 1) \\ &= \mathbb{P}(Z(t) = 1 \mid \{X_1(0) = 0 \text{ and } X_2(0) = 1\} \cup \{X_1(0) = 1 \text{ and } X_2(0) = 0\}) \\ &= \frac{\mathbb{P}(\{Z(t) = 1\} \cap (\{X_1(0) = 0 \text{ and } X_2(0) = 1\} \cup \{X_1(0) = 1 \text{ and } X_2(0) = 0\}))}{\mathbb{P}(\{X_1(0) = 0 \text{ and } X_2(0) = 1\} \cup \{X_1(0) = 1 \text{ and } X_2(0) = 0\})} \\ &= \frac{\mathbb{P}(\{Z(t) = 1\} \cap \{X_1(0) = 0 \text{ and } X_2(0) = 1\})}{\mathbb{P}(Z(0) = 1)} \\ &\quad + \frac{\mathbb{P}(\{Z(t) = 1\} \cap \{X_1(0) = 1 \text{ and } X_2(0) = 0\})}{\mathbb{P}(Z(0) = 1)} \\ &= \mathbb{P}(Z(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1) \frac{\mathbb{P}(X_1(0) = 0 \text{ and } X_2(0) = 1)}{\mathbb{P}(Z(0) = 1)} \\ &\quad + \mathbb{P}(Z(t) = 1 \mid X_1(0) = 1 \text{ and } X_2(0) = 0) \frac{\mathbb{P}(X_1(0) = 1 \text{ and } X_2(0) = 0)}{\mathbb{P}(Z(0) = 1)} \\ &= \mathbb{P}(Z(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1) \frac{\mathbb{P}(X_1(0) = 0 \text{ and } X_2(0) = 1)}{\mathbb{P}(Z(0) = 1)} \\ &\quad + \mathbb{P}(Z(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1) \frac{\mathbb{P}(X_1(0) = 1 \text{ and } X_2(0) = 0)}{\mathbb{P}(Z(0) = 1)} \\ &= \mathbb{P}(Z(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1). \end{aligned}$$

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

$$\begin{aligned} P_{1,1}(t) &= \mathbb{P}(Z(t) = 1 \mid Z(0) = 1) \\ &= \mathbb{P}(X_1(t) = 0 \text{ and } X_2(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1) \\ &\quad + \mathbb{P}(X_1(t) = 1 \text{ and } X_2(t) = 0 \mid X_1(0) = 0 \text{ and } X_2(0) = 1) \\ &= \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 1 \mid X_2(0) = 1) \\ &\quad + \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 1). \end{aligned}$$

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.

Concerning  $P_{1,0}(t)$  we have

$$\begin{aligned} P_{1,0}(t) &= \mathbb{P}(Z(t) = 0 \mid Z(0) = 1) \\ &= \mathbb{P}(X_1(t) = 0 \text{ and } X_2(t) = 0 \mid X_1(0) = 0 \text{ and } X_2(0) = 1). \end{aligned}$$

On the other hand, we have

$$\begin{aligned} P_{1,2}(t) &= \mathbb{P}(Z(t) = 2 \mid Z(0) = 1) \\ &= \mathbb{P}(X_1(t) = 1 \text{ and } X_2(t) = 1 \mid X_1(0) = 0 \text{ and } X_2(0) = 1) \\ &= \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 1 \mid X_2(0) = 1). \end{aligned}$$

We check that

$$\begin{aligned} P_{1,0}(t) + P_{1,1}(t) + P_{1,2}(t) &= \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 1) \\ &\quad + \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 1 \mid X_2(0) = 1) \\ &\quad + \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 1) \\ &\quad + \mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0) \mathbb{P}(X_2(t) = 1 \mid X_2(0) = 1) \\ &= (\mathbb{P}(X_1(t) = 1 \mid X_1(0) = 0) + \mathbb{P}(X_1(t) = 0 \mid X_1(0) = 0)) \\ &\quad \times (\mathbb{P}(X_2(t) = 1 \mid X_2(0) = 1) + \mathbb{P}(X_2(t) = 0 \mid X_2(0) = 1)) \\ &= 1. \end{aligned}$$

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 9.18** Starting from state ①, the process  $X_t = \xi_{N_t}$  stays at state ① during an exponentially distributed Poisson interjump time with parameter  $\lambda$ , 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 (9.7.4), hence the birth rate of  $X_t$  from state ① to state ② is  $\lambda$ .

Then, starting from state ②, the process  $X_t$  stays at ② during an exponentially distributed time with parameter  $\lambda$ . 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  $\alpha$ .

In fact, due to the Markov property,  $X_t$  will remain at 1 during an exponentially distributed time whose expectation may be higher than  $1/\lambda$  when  $\alpha > 0$ . We will compute the expectation of this random time.

a) We have

$$\begin{aligned}\mathbb{E}[T_0^r | X_0 = 1] &= \alpha \left( \frac{1}{\lambda} + \mathbb{E}[T_0^r | X_0 = 1] \right) + (1 - \alpha) \times \left( \frac{1}{\lambda} + 0 \right) \\ &= \frac{1}{\lambda} + \alpha \mathbb{E}[T_0^r | X_0 = 1],\end{aligned}$$

hence

$$\mathbb{E}[T_0^r | X_0 = 1] = \frac{1}{\lambda(1 - \alpha)} \quad (\text{S.55})$$

and

$$\mathbb{E}[T_0^r | X_0 = 0] = \frac{1}{\lambda} + 1 \times \mathbb{E}[T_0^r | X_0 = 1] = \frac{2 - \alpha}{\lambda(1 - \alpha)}.$$

Note that (S.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

$$\mathbb{E}[T_1^r | X_0 = 1] = \frac{1}{\lambda} + (1 - \alpha) \mathbb{E}[T_1^r | X_0 = 0] = \frac{2 - \alpha}{\lambda},$$

since  $\mathbb{E}[T_1^r | X_0 = 0] = 1/\lambda$ .

c) This continuous-time first step analysis argument is similar to the one used in the [solution](#) of Exercise 9.3. Since

$$\mathbb{E}[T_0^r | X_0 = 1] = \frac{1}{\lambda(1 - \alpha)},$$

it takes an exponential random time with parameter  $\lambda(1 - \alpha)$  for the process  $(X_t)_{t \in \mathbb{R}_+}$  to switch from state ① to state ②. Hence the death rate is

$$\frac{1}{\mathbb{E}[T_0^r \mid X_0 = 1]} = \lambda(1 - \alpha),$$

and the infinitesimal generator  $Q$  of  $X_t$  is

$$\begin{bmatrix} \frac{-\lambda}{\mathbb{E}[T_0^r \mid X_0 = 1]} & \frac{\lambda}{\mathbb{E}[T_0^r \mid X_0 = 1]} \\ \frac{\lambda}{\mathbb{E}[T_0^r \mid X_0 = 1]} & \frac{-\lambda}{\mathbb{E}[T_0^r \mid X_0 = 1]} \end{bmatrix} = \begin{bmatrix} -\lambda & \lambda \\ (1 - \alpha)\lambda & -(1 - \alpha)\lambda \end{bmatrix}.$$

#### Problem 9.19

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  $0 < t_1 < t_2 < \dots < t_n$ , the family

$$\begin{aligned} & (N_{t_n}^1 + N_{t_n}^2 - (N_{t_{n-1}}^1 + N_{t_{n-1}}^2), \dots, N_{t_2}^1 + N_{t_2}^2 - (N_{t_1}^1 + N_{t_1}^2)) \\ &= (N_{t_n}^1 - N_{t_{n-1}}^1 + N_{t_n}^2 - N_{t_{n-1}}^2, \dots, N_{t_2}^1 - N_{t_1}^1 + N_{t_2}^2 - N_{t_1}^2) \end{aligned}$$

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

$$N_{t_n}^2 - N_{t_{n-1}}^2, \dots, N_{t_2}^2 - N_{t_1}^2,$$

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  $n \geq 1$ .

(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

$$\begin{aligned} \mathbb{P}(N_{t+h}^1 + N_{t+h}^2 - (N_{s+h}^1 + N_{s+h}^2) = n) \\ = \sum_{k=0}^n \mathbb{P}(N_{t+h}^1 - N_{s+h}^1 = k) \mathbb{P}(N_{t+h}^2 - N_{s+h}^2 = n - k) \end{aligned}$$

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

$$\begin{aligned} \mathbb{P}(M_{t+h} - M_t = n) &= \mathbb{P}(N_{t+h}^1 - N_{t+h}^2 - (N_{s+h}^1 - N_{s+h}^2) = n) \\ &= \mathbb{P}(N_{t+h}^1 - N_{s+h}^1 - (N_{t+h}^2 - N_{s+h}^2) = n) \\ &= \sum_{k \geq 0} \mathbb{P}(N_{t+h}^1 - N_{s+h}^1 = n + k) \mathbb{P}(N_{t+h}^2 - N_{s+h}^2 = k) \end{aligned}$$

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  $h \in \mathbb{R}_+$ .

c) For  $n \in \mathbb{N}$  we have

$$\begin{aligned} \mathbb{P}(M_t = n) &= \mathbb{P}(N_t^1 - N_t^2 = n) \\ &= \sum_{k \geq \max(0, -n)} \mathbb{P}(N_t^1 = n + k) \mathbb{P}(N_t^2 = k) \\ &= e^{-(\lambda_1 + \lambda_2)t} \sum_{k \geq \max(0, -n)} \frac{\lambda_1^{n+k} \lambda_2^k t^{n+2k}}{k!(n+k)!} \\ &= \left(\frac{\lambda_1}{\lambda_2}\right)^{n/2} e^{-(\lambda_1 + \lambda_2)t} \sum_{k \geq \max(0, -n)} \frac{(t\sqrt{\lambda_1 \lambda_2})^{n+2k}}{(n+k)!k!} \\ &= \left(\frac{\lambda_1}{\lambda_2}\right)^{n/2} e^{-(\lambda_1 + \lambda_2)t} I_{|n|}(2t\sqrt{\lambda_1 \lambda_2}), \end{aligned}$$

where

$$I_n(x) = \sum_{k \geq 0} \frac{(x/2)^{n+2k}}{k!(n+k)!}, \quad x > 0,$$

is the modified Bessel function with parameter  $n \geq 0$ . When  $n \leq 0$ , by exchanging  $\lambda_1$  and  $\lambda_2$  we get

$$\begin{aligned}\mathbb{P}(M_t = n) &= \mathbb{P}(-M_t = -n) \\ &= \left(\frac{\lambda_2}{\lambda_1}\right)^{-n/2} e^{-(\lambda_1 + \lambda_2)t} I_{-n}(2t\sqrt{\lambda_1\lambda_2}) \\ &= \left(\frac{\lambda_1}{\lambda_2}\right)^{n/2} e^{-(\lambda_1 + \lambda_2)t} I_{-n}(2t\sqrt{\lambda_1\lambda_2}),\end{aligned}$$

hence in the general case we have

$$\mathbb{P}(M_t = n) = \left(\frac{\lambda_1}{\lambda_2}\right)^{n/2} e^{-(\lambda_1 + \lambda_2)t} I_{|n|}(2t\sqrt{\lambda_1\lambda_2}), \quad n \in \mathbb{Z},$$

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  $\lambda_1$  and  $\lambda_2$  satisfies

$$\begin{aligned}P_{i,j}(t) &= \mathbb{P}(M_t = j \mid M_0 = i) \\ &= \mathbb{P}(M_t = j - i) \\ &= \left(\frac{\lambda_1}{\lambda_2}\right)^{(j-i)/2} e^{-(\lambda_1 + \lambda_2)t} I_{|j-i|}(2t\sqrt{\lambda_1\lambda_2}),\end{aligned}$$

$i, j \in \mathbb{Z}$ ,  $t \geq 0$ .

When  $\lambda_1 = \lambda_2 = \lambda$  we find

$$\mathbb{P}(M_t = n) = e^{-2\lambda t} I_{|n|}(2\lambda t).$$

d) From the bound\*

$$I_{|n|}(y) < C_n y^{|n|} e^y, \quad y > 1,$$

we get

$$\begin{aligned}\mathbb{P}(M_t = n) &\leq C_n e^{-(\lambda_1 + \lambda_2)t} \left(\frac{\lambda_1}{\lambda_2}\right)^{n/2} (2t\sqrt{\lambda_1\lambda_2})^{|n|} e^{2t\sqrt{\lambda_1\lambda_2}} \\ &= C_n e^{-t(\sqrt{\lambda_1} - \sqrt{\lambda_2})^2} \left(\frac{\lambda_1}{\lambda_2}\right)^{n/2} (2t\sqrt{\lambda_1\lambda_2})^{|n|},\end{aligned}$$

which tends to 0 as  $t$  goes to infinity when  $\lambda_1 \neq \lambda_2$ .

---

\* See e.g. Theorem 2.1 of [Laforgia \(1991\)](#) for a proof of this inequality.

Hence we have\*

$$\lim_{t \rightarrow \infty} \mathbb{P}(|M_t| < c) = \sum_{-c < k < c} \lim_{t \rightarrow \infty} \mathbb{P}(|M_t| = k) = 0, \quad c > 0. \quad (\text{S.56})$$

**Remark.** There exists a shorter proof by a probabilistic argument (not required here), which is displayed below, cf. [Bosq and Nguyen \(1996\)](#).

Noting that since  $\mathbb{E}[M_t] = t(\lambda_1 - \lambda_2)$ , for  $t$  large enough we have, depending on the sign of  $\lambda_1 - \lambda_2$ ,

$$c - \mathbb{E}[M_t] \leq -\frac{1}{2} \mathbb{E}[M_t] \quad \text{or} \quad -c - \mathbb{E}[M_t] \geq \frac{1}{2} \mathbb{E}[M_t],$$

we get

$$-c \leq M_t \leq c \implies M_t - \mathbb{E}[M_t] \leq -\frac{1}{2} \mathbb{E}[M_t] \quad \text{or} \quad M_t - \mathbb{E}[M_t] \geq \frac{1}{2} \mathbb{E}[M_t],$$

hence by Chebyshev's inequality, we have

$$\begin{aligned} \mathbb{P}(|M_t| \leq c) &\leq \mathbb{P}\left(|M_t - \mathbb{E}[M_t]| \geq \frac{1}{2} \mathbb{E}[M_t]\right) \\ &\leq 4 \frac{\text{Var}[M_t]}{(\mathbb{E}[M_t])^2} = \frac{4}{t} \frac{\lambda_1 + \lambda_2}{(\lambda_1 - \lambda_2)^2}, \end{aligned}$$

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  $M_t \geq 0$ ,  $M_t$  represents the number of waiting customers. When  $M_t \leq 0$ ,  $-M_t$  represents the number of waiting drivers.

Relation (S.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.

## Problem 9.20

- a) We have

$$Q = \begin{bmatrix} -N\lambda & N\lambda & 0 & \cdots & \cdots & 0 & 0 & 0 \\ \mu & -\mu - (N-1)\lambda & (N-1)\lambda & \cdots & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & (N-1)\mu & -(N-1)\mu - \lambda & \lambda \\ 0 & 0 & 0 & \cdots & \cdots & 0 & N\mu & -N\mu \end{bmatrix}.$$

\* Treating the case  $\lambda_1 = \lambda_2$  is more complicated and not required.

- 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

$$\frac{\partial G_k}{\partial t}(s, t) = \sum_{n=0}^N s^n P'_{k,n}(t),$$

and use the expression

$$\frac{\partial G_k}{\partial s}(s, t) = \sum_{n=1}^N n s^{n-1} P'_{k,n}(t).$$

- d) We have

$$\begin{aligned} & \lambda N(s-1)G_k(s, t) + (\mu + (\lambda - \mu)s - \lambda s^2) \frac{\partial G_k}{\partial s}(s, t) - \frac{\partial G_k}{\partial t}(s, t) \\ &= -(s-1)(\lambda + \mu)(N-k) \left( (s-1)\mu e^{-(\lambda+\mu)t} + \lambda s + \mu \right)^k \lambda \left( -(s-1)\lambda e^{-(\lambda+\mu)t} + \lambda s + \mu \right)^{N-k-1} e^{-(\lambda+\mu)t} \\ &+ (s-1)(\lambda + \mu)k\mu \left( (s-1)\mu e^{-(\lambda+\mu)t} + \lambda s + \mu \right)^{k-1} \left( -(s-1)\lambda e^{-(\lambda+\mu)t} + \lambda s + \mu \right)^{N-k} e^{-(\lambda+\mu)t} \\ &+ (s-1) \left( (s-1)\mu e^{-(\lambda+\mu)t} + \lambda s + \mu \right)^k N\lambda \left( -(s-1)\lambda e^{-(\lambda+\mu)t} + \lambda s + \mu \right)^{N-k} \\ &+ \left( \lambda s^2 - (\lambda - \mu)s - \mu \right) (N-k) \left( \lambda e^{-(\lambda+\mu)t} - \lambda \right) \left( (s-1)\mu e^{-(\lambda+\mu)t} + \lambda s + \mu \right)^k \left( -(s-1)\lambda e^{-(\lambda+\mu)t} + \lambda s + \mu \right)^{N-k} \\ &- \left( \mu e^{-(\lambda+\mu)t} + \lambda \right) k \left( (s-1)\mu e^{-(\lambda+\mu)t} + \lambda s + \mu \right)^{k-1} \left( -(s-1)\lambda e^{-(\lambda+\mu)t} + \lambda s + \mu \right)^{N-k} \left( \lambda s^2 - (\lambda - \mu)s - \mu \right) \\ &= 0. \end{aligned}$$

- e) This expression follows from the relation

$$\mathbb{E}[X_t \mid X_0 = k] = \frac{\partial G_k}{\partial s}(s, t)|_{s=1}$$

and the result of Question (d).

- f) We have

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbb{E}[X_t \mid X_0 = k] &= k \frac{\lambda(\lambda + \mu)^{k-1}}{(\lambda + \mu)^N} (\mu + \lambda)^{N-k} \\ &+ (N-k) \frac{(\mu + \lambda)^k \lambda(\mu + \lambda)^{N-k-1}}{(\lambda + \mu)^N} = \frac{N\lambda}{\lambda + \mu}. \end{aligned}$$

Problem 9.21 First, we note that in case

$$f(x) = \mathbb{1}_{[a,b]}(x), \quad 0 \leq a \leq b \leq t,$$

we have, by direct counting,

$$\sum_{k=1}^{N_t} f(T_k) = \sum_{k=1}^{N_t} \mathbb{1}_{[a,b]}(T_k) = N_b - N_a.$$

Hence

$$\mathbb{E} \left[ \sum_{k=1}^{N_t} f(T_k) \right] = \mathbb{E}[N_b - N_a] = \lambda(b-a) = \lambda \int_0^t \mathbb{1}_{[a,b]}(s) ds = \lambda \int_0^t f(s) ds,$$

hence (9.7.6) is proved for  $f(x) = \mathbb{1}_{[a,b]}(x)$ .

Next, we check by linearity that the (9.7.6) extends to a linear combination of [indicator functions](#) of the form

$$f(x) = \sum_{k=1}^n \alpha_k \mathbb{1}_{[a_{k-1}, a_k]}(x), \quad 0 \leq a_0 < a_1 < \cdots < a_n < t.$$

The difficult part is to do the extension from the linear combinations of indicator functions to “any” integrable function  $f : [0, t] \rightarrow \mathbb{R}$ . 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 [Privault \(2009\)](#).

Let  $g_n$ ,  $n \geq 1$ , be defined as

$$\begin{aligned} g_n(t_1, t_2, \dots, t_n) &= \sum_{k=1}^n \mathbb{1}_{\{t_{k-1} < t < t_k\}} (f(t_1) + \cdots + f(t_{k-1})) \\ &\quad + \mathbb{1}_{\{t_n < t\}} (f(t_1) + \cdots + f(t_n)), \end{aligned}$$

with  $t_0 := 0$ , so that

$$\sum_{k=1}^{\min(n, N_t)} f(T_k) = g_n(T_1, T_2, \dots, T_n).$$

Then

$$\begin{aligned} \mathbb{E} \left[ \sum_{k=1}^{\min(n, N_t)} f(T_k) \right] &= \mathbb{E}[g(T_1, T_2, \dots, T_n)] \\ &= \lambda^n \sum_{k=1}^n \int_0^\infty e^{-\lambda t_n} \int_0^{t_n} \cdots \int_0^{t_2} \mathbb{1}_{\{t_{k-1} < t < t_k\}} (f(t_1) + \cdots + f(t_{k-1})) dt_1 \cdots dt_n \\ &\quad + \lambda^n \int_0^\infty e^{-\lambda t_n} \int_0^{t_n} \cdots \int_0^{t_2} (f(t_1) + \cdots + f(t_n)) dt_1 \cdots dt_n \\ &= \lambda^n \sum_{k=1}^n \int_t^\infty e^{-\lambda t_n} \frac{(t_n - t)^{n-k}}{(n-k-1)!} dt_n \\ &\quad \times \int_0^t \int_0^{t_{k-1}} \cdots \int_0^{t_2} (f(t_1) + \cdots + f(t_{k-1})) dt_1 \cdots dt_{k-1} \\ &\quad + \lambda^n \int_0^t e^{-\lambda t_n} \int_0^{t_n} \cdots \int_0^{t_2} (f(t_1) + \cdots + f(t_n)) dt_1 \cdots dt_n \end{aligned}$$

$$\begin{aligned}
&= e^{-\lambda t} \sum_{k=1}^n \lambda^k \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^{n-k}}{(n-k-1)!} dt \\
&\quad \times \int_0^t \int_0^{t_{k-1}} \cdots \int_0^{t_2} (f(t_1) + \cdots + f(t_{k-1})) dt_1 \cdots dt_{k-1} \\
&\quad + \lambda^n \int_0^t e^{-\lambda t_n} \int_0^{t_n} \cdots \int_0^{t_2} (f(t_1) + \cdots + f(t_n)) dt_1 \cdots dt_n \\
&= e^{-\lambda t} \sum_{k=1}^n \lambda^k \int_0^t \int_0^{t_{k-1}} \cdots \int_0^{t_2} (f(t_1) + \cdots + f(t_{k-1})) dt_1 \cdots dt_{k-1} \\
&\quad + e^{-\lambda t} \sum_{k \geq n} \lambda^k \int_0^t \frac{(t-t_n)^{k-n}}{(k-n)!} \int_0^{t_n} \cdots \int_0^{t_2} (f(t_1) + \cdots + f(t_n)) dt_1 \cdots dt_n \\
&= e^{-\lambda t} \sum_{k=1}^n \lambda^k \int_0^t \int_0^{t_{k-1}} \cdots \int_0^{t_2} (f(t_1) + \cdots + f(t_{k-1})) dt_1 \cdots dt_{k-1} \\
&\quad + e^{-\lambda t} \sum_{k \geq n} \lambda^k \int_0^t \int_{t_n}^t \int_{t_n}^{t_k} \cdots \int_{t_n}^{t_{n+2}} dt_{n+1} \cdots dt_k \\
&\quad \times \int_0^{t_{n+1}} \cdots \int_0^{t_2} (f(t_1) + \cdots + f(t_n)) dt_1 \cdots dt_n \\
&= e^{-\lambda t} \sum_{k \geq 0} \lambda^k \int_0^t \int_0^{t_k} \cdots \int_0^{t_2} (f(t_1) + \cdots + f(t_{\min(k,n)})) dt_1 \cdots dt_k \\
&= e^{-\lambda t} \sum_{k \geq 0} \frac{\lambda^k}{k!} \int_0^t \cdots \int_0^t (f(t_1) + \cdots + f(t_{\min(k,n)})) dt_1 \cdots dt_k.
\end{aligned}$$

Hence as  $n$  goes to infinity,

$$\begin{aligned}
\mathbb{E} \left[ \sum_{k=1}^{N_t} f(T_k) \right] &= \lim_{n \rightarrow \infty} \mathbb{E} \left[ \sum_{k=1}^{\min(n, N_t)} f(T_k) \right] \\
&= \lim_{n \rightarrow \infty} e^{-\lambda t} \sum_{k \geq 0} \frac{\lambda^k}{k!} \int_0^t \cdots \int_0^t (f(t_1) + \cdots + f(t_{\min(k,n)})) dt_1 \cdots dt_k \\
&= e^{-\lambda t} \sum_{k \geq 0} \frac{\lambda^k}{k!} \int_0^t \cdots \int_0^t (f(t_1) + \cdots + f(t_k)) dt_1 \cdots dt_k \\
&= \lambda e^{-\lambda t} \int_0^t f(s) ds \sum_{k \geq 1} \frac{(\lambda t)^{k-1}}{(k-1)!} = \lambda \int_0^t f(s) ds.
\end{aligned}$$

## Chapter 10 - Discrete-Time Martingales

Exercise 10.1 For all  $k = 0, 1, \dots, B$  and  $n \geq 1$  we have

$$\begin{aligned}\mathbb{P}(\tau_{0,B} = \infty \mid S_0 = k) &\leq \mathbb{P}(\tau_{0,B} > nB \mid S_0 = k) \\ &\leq \mathbb{P}\left(\bigcap_{k=0}^{n-1} \{X_{kB+1} = 1, \dots, X_{kB+B} = 1\}^c\right) \\ &= (1 - p^k)^n,\end{aligned}$$

from which we obtain  $\mathbb{P}(\tau_{0,B} = \infty \mid S_0 = k) = 0$  after letting  $n$  tend to infinity when  $p \in [0, 1)$ , hence  $\mathbb{P}(\tau_{0,B} < \infty \mid S_0 = k) = 1$ . In case  $p = 1$ , we clearly have  $\mathbb{P}(\tau_{0,B} < \infty \mid S_0 = k) = 1$ .

Exercise 10.2

a) We have

$$\begin{aligned}\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}\left[\sum_{k=1}^{n+1} 2^{k-1} X_k \mid \mathcal{F}_n\right] \\ &= \mathbb{E}\left[\sum_{k=1}^n 2^{k-1} X_k \mid \mathcal{F}_n\right] + \mathbb{E}[2^n X_{n+1} \mid \mathcal{F}_n] \\ &= \sum_{k=1}^n 2^{k-1} X_k + 2^n \mathbb{E}[X_{n+1}] \\ &= M_n, \quad n \geq 0.\end{aligned}$$

- 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

$$-\sum_{k=1}^n 2^{k-1} = -\frac{1-2^n}{1-2} = 1-2^n, \quad n \geq 1.$$

We have

$$\mathbb{P}(M_{n \wedge \tau} = 1-2^n) = 2^{-n} \quad \text{and} \quad \mathbb{P}(M_{n \wedge \tau} = 1) = 1-2^{-n}, \quad n \geq 1.$$

e) We have

$$\begin{aligned}\mathbb{E}[M_{n \wedge \tau}] &= (1-2^n)\mathbb{P}(M_{n \wedge \tau} = 1-2^n) + \mathbb{P}(M_{n \wedge \tau} = 1) \\ &= (1-2^n)2^{-n} + (1-2^{-n})\end{aligned}$$

$$= 0, \quad n \geq 1.$$

f) The Stopping Time Theorem 10.9 directly states that

$$\mathbb{E}[M_{n \wedge \tau}] = \mathbb{E}[M_0] = 0.$$

g) We have

$$\mathbb{P}(\tau < \infty) = \sum_{n \geq 1} \mathbb{P}(\tau = n) = \sum_{n \geq 1} \frac{1}{2^n} = \frac{1}{2} \frac{1}{1 - 1/2} = 1.$$

h) We have  $\mathbb{E}[M_\tau] = 1$  and  $\lim_{n \rightarrow \infty} \mathbb{E}[M_{\tau \wedge n}] = 0$ .

### Exercise 10.3

- a) The random time  $\tau$  is a stopping time because for every  $n \geq 0$ , the validity of the event  $\{\tau > n\}$  can be decided by studying the path of the process  $(S_k)_{k \geq 0}$  until time  $n$ .
- b) The range of possible values of  $S_\tau$  is  $\{-1, 0, 1, 2, 3, 4, \dots\}$ .

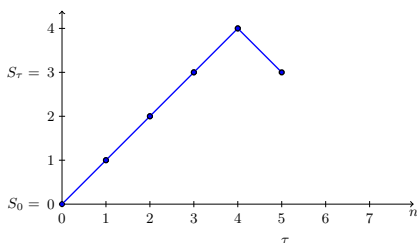


Fig. S.21: Sample path of the random walk  $(S_n)_{n \geq 0}$ .

- c) According to the stopping time theorem, the stopped process  $(S_{n \wedge \tau})_{n \geq 0}$  is a martingale, and therefore we have

$$\mathbb{E}[S_\tau] = \mathbb{E} \left[ \lim_{n \rightarrow \infty} S_{n \wedge \tau} \right] = \lim_{n \rightarrow \infty} \mathbb{E}[S_{n \wedge \tau}] = \lim_{n \rightarrow \infty} \mathbb{E}[S_0] = 0.$$

The exchange between limit and expected value can be justified from the dominated convergence theorem, see the next question.

- d) We have  $\mathbb{P}(S_\tau = -1) = 1/2$  and  $\mathbb{P}(S_\tau = k) = 1/2^{k+2}$ ,  $k \geq 0$ . In particular, we have  $S_{n \wedge \tau} \leq 1 + S_\tau$ , and the non-negative random variable  $1 + S_\tau$  is integrable according to the next question.
- e) We have

$$\mathbb{E}[1 + S_\tau] = 1 + \sum_{k \in \mathbb{Z}} k \mathbb{P}(S_\tau = k)$$

$$\begin{aligned}
&= \frac{1}{2} + \sum_{k \geq 0} k \mathbb{P}(S_\tau = k) \\
&= \frac{1}{2} + \sum_{k \geq 0} \frac{k}{2^{k+2}} \\
&= \frac{1}{2} + \frac{1}{8} \sum_{k \geq 0} \frac{k}{2^{k-1}} \\
&= \frac{1}{2} + \frac{1}{8} \frac{1}{(1 - 1/2)^2} \\
&= 1,
\end{aligned}$$

hence  $\mathbb{E}[S_\tau] = 0$ .

Exercise 10.4 We note that for all  $n \geq 1$  we have

$$\{\tau = n\} = \{\tau \geq n\} \setminus \{\tau > n\} = \{\tau > n-1\} \cap \{\tau > n\}^c \in \mathcal{F}_n,$$

since  $\{\tau > n\} \in \mathcal{F}_n$  and  $\{\tau > n-1\} \in \mathcal{F}_{n-1} \subset \mathcal{F}_n$ . In case  $n = 0$ , we have

$$\{\tau = 0\} = \{\tau > 0\}^c \in \mathcal{F}_0.$$

Conversely, assuming that (10.5.1) holds, we have

$$\{\tau > n\} = \{\tau \leq n\}^c = \left( \bigcup_{k=0}^n \{\tau = k\} \right)^c = \bigcap_{k=0}^n \{\tau = k\}^c \in \mathcal{F}_n,$$

since

$$\{\tau = k\} \in \mathcal{F}_k \subset \mathcal{F}_n, \quad k = 0, 1, \dots, n.$$

Exercise 10.5 We consider three examples.

- i) Consider the stopped martingale  $(X_n)_{n \geq 0} := (M_{\tau \wedge n})_{n \geq 0}$  of Exercise 10.2, for which by Question (e) therein we have  $\mathbb{E}[X_n] = \mathbb{E}[M_{\tau \wedge n}] = 0$ ,  $n \geq 0$ , hence

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n] = \lim_{n \rightarrow \infty} \mathbb{E}[M_{\tau \wedge n}] = 0,$$

while

$$\lim_{n \rightarrow \infty} X_n = \lim_{n \rightarrow \infty} M_{\tau \wedge n} = 1,$$

*a.s.* Indeed, by (1.2.5), the *non-decreasing* sequence of events  $(\{X_n = 1\})_{n \geq 1}$  satisfies

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} X_n = 1\right) = \mathbb{P}\left(\bigcup_{n \geq 1} \{X_n = 1\}\right)$$

$$\begin{aligned}
&= \lim_{n \rightarrow \infty} \mathbb{P}(X_n = 1) \\
&= \lim_{n \rightarrow \infty} \left(1 - \frac{1}{2^n}\right) \\
&= 1.
\end{aligned}$$

Note that here the sequence  $(M_n)_{n \geq 0}$  is unbounded, as

$$\inf_{n \geq 0} M_n = -\infty \quad \text{and} \quad \sup_{n \geq 0} |M_n| = +\infty.$$

- ii) Consider the random sequence  $X_n := n \mathbb{1}_{\{U < 1/n\}}$ ,  $n \geq 1$ , where  $U \simeq U(0, 1]$  is a uniformly distributed random variable on  $(0, 1]$ . More formally,  $(X_n)_{n \geq 1}$  can be constructed by *e.g.* taking  $\Omega = (0, 1)$  with the probability measure  $\mathbb{P}$  defined by  $\mathbb{P}([a, b]) = b - a$ ,  $0 < a \leq b < 1$ , and defining the random variable  $X_n(\omega) = n \mathbb{1}_{\{0 < \omega < 1/n\}}$ ,  $\omega \in (0, 1)$ , in which case we have  $\sup_{n \geq 1} |X_n| = +\infty$ ,  $\mathbb{E}[X_n] = n \times (1/n - 0) = 1$ ,  $n \geq 1$ , and  $\lim_{n \rightarrow \infty} X_n(\omega) = 0$  for all  $\omega \in (0, 1)$ .
- iii) Consider the random sequence  $X_n := n^2 \mathbb{1}_{\{U_n < 1/n^2\}}$ ,  $n \geq 1$ , where  $(U_n)_{n \geq 1} \simeq U(0, 1]$  is a sequence of uniformly distributed random variables on  $(0, 1]$ . On the one hand, we have

$$\mathbb{E}[X_n] = n^2 \mathbb{E}[\mathbb{1}_{\{U_n < 1/n^2\}}] = n^2 \mathbb{P}(\mathbb{1}_{\{U_n < 1/n^2\}}) = 1, \quad n \geq 1.$$

On the other hand, we have

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n > 0) = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty,$$

hence by the [Borel–Cantelli Lemma](#), the probability that  $X_n > 0$  infinitely many times is zero, which yields  $\lim_{n \rightarrow \infty} X_n = 0$  almost surely, *i.e.* with probability one.

#### Exercise 10.6

- a) From the tower property of conditional expectations (1.6.8), we have:

$$\mathbb{E}[M_{n+1}] = \mathbb{E}[\mathbb{E}[M_{n+1} \mid \mathcal{F}_n]] \geq \mathbb{E}[M_n], \quad n \geq 0.$$

- b) If  $(Z_n)_{n \geq 0}$  is a stochastic process with independent increments having nonnegative expectations, we have

$$\begin{aligned}
\mathbb{E}[Z_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}[Z_n \mid \mathcal{F}_n] + \mathbb{E}[Z_{n+1} - Z_n \mid \mathcal{F}_n] \\
&= \mathbb{E}[Z_n \mid \mathcal{F}_n] + \mathbb{E}[Z_{n+1} - Z_n] \\
&\geq \mathbb{E}[Z_n \mid \mathcal{F}_n] = Z_n, \quad n \geq 0.
\end{aligned}$$

c) We let  $A_0 := 0$  and

$$A_{n+1} := A_n + \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n], \quad n \geq 0,$$

which defines a predictable process by induction, as  $A_{n+1}$  is  $\mathcal{F}_n$ -measurable,  $n \geq 0$ , and

$$N_n := M_n - A_n, \quad n \geq 0. \quad (\text{S.57})$$

(i) For all  $n \geq 0$ , we have

$$\begin{aligned} \mathbb{E}[N_{n+1} \mid \mathcal{F}_n] &= \mathbb{E}[M_{n+1} - A_{n+1} \mid \mathcal{F}_n] \\ &= \mathbb{E}[M_{n+1} - A_n - \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] \mid \mathcal{F}_n] \\ &= \mathbb{E}[M_{n+1} - A_n \mid \mathcal{F}_n] - \mathbb{E}[\mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] \mid \mathcal{F}_n] \\ &= \mathbb{E}[M_{n+1} - A_n \mid \mathcal{F}_n] - \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] \\ &= -\mathbb{E}[A_n \mid \mathcal{F}_n] + \mathbb{E}[M_n \mid \mathcal{F}_n] \\ &= M_n - A_n \\ &= N_n, \end{aligned}$$

hence  $(N_n)_{n \geq 0}$  is a martingale with respect to  $(\mathcal{F}_n)_{n \geq 0}$ .

(ii) For all  $n \geq 0$ , we have

$$\begin{aligned} A_{n+1} - A_n &= \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] \\ &= \mathbb{E}[M_{n+1} \mid \mathcal{F}_n] - \mathbb{E}[M_n \mid \mathcal{F}_n] \\ &= \mathbb{E}[M_{n+1} \mid \mathcal{F}_n] - M_n \geq 0, \end{aligned}$$

since  $(M_n)_{n \geq 0}$  is a submartingale.

(iii) By induction we have  $A_{n+1} = A_n + \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n]$ ,  $n \geq 0$ , which is  $\mathcal{F}_n$ -measurable if  $A_n$  is  $\mathcal{F}_{n-1}$ -measurable,  $n \geq 1$ .

(iv) This property is obtained by construction in (S.57).

d) For all bounded stopping times  $\sigma$  and  $\tau$  such that  $\sigma \leq \tau$  a.s., we have  $\mathbb{E}[A_\sigma] \leq \mathbb{E}[A_\tau]$ , and  $\mathbb{E}[N_\sigma] = \mathbb{E}[N_\tau]$  by (10.3.3), hence

$$\begin{aligned} \mathbb{E}[M_\sigma] &= \mathbb{E}[N_\sigma] + \mathbb{E}[A_\sigma] \\ &= \mathbb{E}[N_\tau] + \mathbb{E}[A_\sigma] \\ &\leq \mathbb{E}[N_\tau] + \mathbb{E}[A_\tau] \\ &= \mathbb{E}[M_\tau], \end{aligned}$$

by (10.3.3), since  $(N_n)_{n \geq 0}$  is a martingale and  $(A_n)_{n \geq 0}$  is non-decreasing.

## Exercise 10.7

a) We will show more generally that

$$\phi(p_1x_1 + p_2x_2 + \cdots + p_nx_n) \leq p_1\phi(x_1) + p_2\phi(x_2) + \cdots + p_n\phi(x_n), \quad (\text{S.58})$$

$x_1, \dots, x_n \in \mathbb{R}$ , for any sequence of coefficients  $p_1, p_2, \dots, p_n \geq 0$  such that  $p_1 + p_2 + \cdots + p_n = 1$ . The inequality (S.58) clearly holds for  $n = 1$ , and for  $n = 2$  it coincides with the convexity property of  $\phi$ , *i.e.*

$$\phi(p_1x_1 + p_2x_2) \leq p_1\phi(x_1) + p_2\phi(x_2), \quad x_1, x_2 \in \mathbb{R}.$$

Assuming that (S.58) holds for some  $n \geq 1$  and taking  $p_1, p_2, \dots, p_{n+1} \geq 0$  such that  $p_1 + p_2 + \cdots + p_{n+1} = 1$  and  $0 < p_{n+1} < 1$  and applying (10.5.2) at the second order, we have

$$\begin{aligned} & \phi(p_1x_1 + p_2x_2 + \cdots + p_{n+1}x_{n+1}) \\ &= \phi\left((1 - p_{n+1})\frac{p_1x_1 + p_2x_2 + \cdots + p_nx_n}{1 - p_{n+1}} + p_{n+1}x_{n+1}\right) \\ &\leq (1 - p_{n+1})\phi\left(\frac{p_1x_1 + p_2x_2 + \cdots + p_nx_n}{1 - p_{n+1}}\right) + p_{n+1}\phi(x_{n+1}) \\ &\leq (1 - p_{n+1})\left(\frac{p_1\phi(x_1) + p_2\phi(x_2) + \cdots + p_n\phi(x_n)}{1 - p_{n+1}}\right) + p_{n+1}\phi(x_{n+1}) \\ &= p_1\phi(x_1) + p_2\phi(x_2) + \cdots + p_{n+1}\phi(x_{n+1}), \end{aligned}$$

and we conclude by induction.

b) Taking  $p_1 = p_1 = \cdots = p_N = 1/N$ , we have

$$\begin{aligned} \mathbb{E}^* \left[ \phi \left( \frac{S_1 + \cdots + S_N}{N} \right) \right] &\leq \mathbb{E}^* \left[ \frac{\phi(S_1) + \cdots + \phi(S_N)}{N} \right] && \text{since } \phi \text{ is convex,} \\ &= \frac{\mathbb{E}^*[\phi(S_1)] + \cdots + \mathbb{E}^*[\phi(S_N)]}{N} \\ &= \frac{\mathbb{E}^*[\phi(\mathbb{E}^*[S_N \mid \mathcal{F}_1])] + \cdots + \mathbb{E}^*[\phi(\mathbb{E}^*[S_N \mid \mathcal{F}_N])]}{N} && \text{because } (S_n)_{n \geq 0} \text{ is a martingale,} \\ &\leq \frac{\mathbb{E}^*[\mathbb{E}^*[\phi(S_N) \mid \mathcal{F}_1]] + \cdots + \mathbb{E}^*[\mathbb{E}^*[\phi(S_N) \mid \mathcal{F}_N]]}{N} && \text{by Jensen's inequality,} \\ &= \frac{\mathbb{E}^*[\phi(S_N)] + \cdots + \mathbb{E}^*[\phi(S_N)]}{N} && \text{by the tower property,} \\ &= \mathbb{E}^*[\phi(S_N)]. \end{aligned}$$

c) This is an application of the above bound to the convex payoff function  $x \mapsto (x - K)^+$ , see Figure S.22 and the  code below for an illustration.

```

1  nSim=99999;p=0.4;q=1-p;n=7;a=q/p;r=1;european=0;asian=0;K=1.5
   dev.new(width=16,height=7)
3  for (j in 1:nSim){S<-a~cumsum(2*rbinom(n,1,p)-1);color="blue"
   A<-sum(c(1,S))/(n+1);if (S[n]>=K){european=european+S[n]-K}
5  if (A>=K){asian=asian+A-K};if (S[n]>A){color="darkred"} else {color="darkgreen"}
   plot(seq(0,n),c(1,S), xlab="Time", xlim=c(0,n), type='o', ylim=c(0,a^(n-2)), lwd=3,
        ylab="", col=color,main=paste("Asian Price=",format(round(asian,2)),"/",
        j,"j,"=,format(round(european/j,2))), xaxs='i',xast='n', yaxt='n', yaxs='i', yaxp
        =c(0,10,10))
7  text(3,6,paste("A-Payoff=",format(round(max(A-K,0,2)), " E-Payoff=",
        format(round(max(S[n]-K,0,2))), col=color,cex=2)
   axis(1, at=seq(0,n), labels=seq(0,n), las=1)
9  axis(2, at=c(0,K,A,1,2,3,4,5,6,7,8,9,10), labels=c(0,"K", "Average",1,2,3,4,5,6,7,8,9,10),
        las=2)
   lines(seq(0,n),rep(K,n+1),col="red",lty=1,lwd=4);
11  lines(seq(0,n),rep(A,n+1),col="darkgreen",lty=2,lwd=4); Sys.sleep(0.1)
   if (S[n]>K || A>K){readline(prompt="Pause. Press <Enter> to continue...")}

```

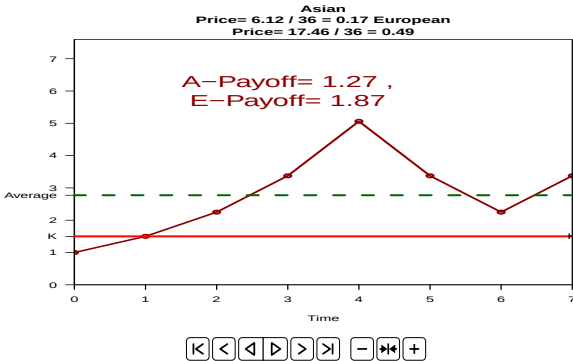


Fig. S.22: Asian option price vs. European option price.\*

## Exercise 10.8

a) We have

$$\mathbb{E}[M_n] \leq \mathbb{E}[\mathbb{E}[M_{n+1} \mid \mathcal{F}_n]] = \mathbb{E}[M_{n+1}], \quad n \geq 0.$$

b) We write

$$\begin{aligned}
 S_n - \alpha n &= \sum_{k=1}^n (X_k - p) + n(p - \alpha) \\
 &= \underbrace{(X_1 + X_2 + \cdots + X_n - np)}_{\text{martingale}} + n(p - \alpha)
 \end{aligned}$$

\* The animation works in Acrobat Reader.

$$= S_n - np + (p - \alpha)n$$

as the sum of a martingale (a stochastic process with centered independent increments) and  $(p - \alpha)n$ . As in the Doob-Meyer decomposition of Exercise 10.6-(c), we conclude that  $(S_n)_{n \geq 0}$  is a submartingale if and only if  $p \geq \alpha$ . Indeed, we have

$$\begin{aligned} \mathbb{E}[S_n - \alpha n \mid \mathcal{F}_k] &= \mathbb{E}[X_1 + X_2 + \cdots + X_n - np \mid \mathcal{F}_k] + (p - \alpha)n \\ &= X_1 + X_2 + \cdots + X_k - kp + (p - \alpha)n \\ &= X_1 + X_2 + \cdots + X_k - k\alpha + (n - k)(p - \alpha) \\ &\geq S_k - k\alpha, \quad k = 0, 1, \dots, n, \end{aligned}$$

if and only if  $p \geq \alpha$ .

#### Exercise 10.9

a) We have

$$\phi(M_k) = \phi(\mathbb{E}[M_n \mid \mathcal{F}_k]) \leq \mathbb{E}[\phi(M_n) \mid \mathcal{F}_k], \quad k = 0, 1, \dots, n$$

b) We have

$$\phi(M_k) \leq \phi(\mathbb{E}[M_n \mid \mathcal{F}_k]) \leq \mathbb{E}[\phi(M_n) \mid \mathcal{F}_k], \quad k = 0, 1, \dots, n.$$

#### Problem 10.10

a) We have

$$\{\tau_x > n\} = \bigcap_{k=0}^n \{M_k < x\}.$$

On the other hand, for all  $k = 0, 1, \dots, n$  we have  $\{M_k < x\} \in \mathcal{F}_k \subset \mathcal{F}_n$ , hence  $\{\tau_x > n\} \in \mathcal{F}_n$  by the stability property of  $\sigma$ -algebras under intersection, cf. (1.1.1).

b) We have

$$\begin{aligned} x\mathbb{P}\left(\max_{k=0,1,\dots,n} M_k \geq x\right) &= x\mathbb{P}(\tau_x \leq n) \\ &= x\mathbb{E}[\mathbb{1}_{\{\tau_x \leq n\}}] \\ &\leq \mathbb{E}[M_{\tau_x} \mathbb{1}_{\{\tau_x \leq n\}}] \\ &= \mathbb{E}[M_{\tau_x \wedge n} \mathbb{1}_{\{\tau_x \leq n\}}] \quad (\text{S.59}) \\ &\leq \mathbb{E}[M_{\tau_x \wedge n}] \quad (\text{S.60}) \\ &= \mathbb{E}[M_{\tau_x \wedge 0}] \\ &= \mathbb{E}[M_0] \\ &= \mathbb{E}[M_n], \quad (\text{S.61}) \end{aligned}$$

where we used the condition  $M_{\tau_x \wedge n} \geq 0$  from (S.59) to (S.60), hence

$$\mathbb{P}\left(\max_{k=0,1,\dots,n} M_k \geq x\right) \leq \frac{\mathbb{E}[M_n]}{x}, \quad x > 0. \quad (\text{S.62})$$

*Remark.* The nonnegativity of  $(M_n)_{n \geq 0}$  is used to reach (S.60), and the Doob Stopping Time Theorem 10.9 is used to derive (S.61).

- c) When  $(M_n)_{n \geq 0}$  is a *submartingale*, by the Doob Stopping Time Theorem 10.9 for *submartingales* we have

$$\mathbb{E}[M_{\tau_x \wedge n}] \leq \mathbb{E}[M_n],$$

see Exercise 10.6-(d), showing that (S.61) and (S.62) above still hold in this case.

- d) Since  $x \mapsto x^2$  is a convex function,  $((M_n)^2)_{n \geq 0}$  is a *submartingale* by Question (a), hence by Question (c) we have

$$\mathbb{P}\left(\max_{k=0,1,\dots,n} M_k \geq x\right) = \mathbb{P}\left(\max_{k=0,1,\dots,n} (M_k)^2 \geq x^2\right) \leq \frac{\mathbb{E}[(M_n)^2]}{x^2}, \quad x > 0.$$

- e) Similarly to Question (d),  $x \mapsto x^p$  is a convex function for all  $p \geq 1$  hence  $((M_n)^p)_{n \geq 0}$  is a *submartingale* by Question (a), and by Question (c) we find

$$\mathbb{P}\left(\max_{k=0,1,\dots,n} M_k \geq x\right) = \mathbb{P}\left(\max_{k=0,1,\dots,n} (M_k)^p \geq x^p\right) \leq \frac{\mathbb{E}[(M_n)^p]}{x^p}, \quad x > 0.$$

- f) We note that  $(S_n)_{n \geq 0}$  is a martingale because it has centered and independent increments, with

$$\mathbb{E}[(S_n)^2] = \text{Var}[S_n] = n \text{Var}[Y_1] = n\sigma^2,$$

hence by Question (d) we have

$$\mathbb{P}\left(\max_{k=0,1,\dots,n} S_k \geq x\right) \leq \frac{\mathbb{E}[(S_n)^2]}{x^2} = \frac{n\sigma^2}{x^2}, \quad x > 0.$$

- g) When  $(M_n)_{n \geq 0}$  is a (not necessarily nonnegative) *submartingale* we can modify the answer to Question (b) using the Doob Stopping Time Theorem 10.9 for *submartingales*, see Exercise 10.6-(d), as follows:

$$\begin{aligned} x\mathbb{P}\left(\max_{k=0,1,\dots,n} M_k \geq x\right) &= x\mathbb{E}[\mathbb{1}_{\{\tau_x \leq n\}}] \\ &\leq \mathbb{E}[M_{\tau_x \wedge n} \mathbb{1}_{\{\tau_x \leq n\}}] \\ &\leq \mathbb{E}[(M_{\tau_x \wedge n})^+] \\ &\leq \mathbb{E}[(M_n)^+], \end{aligned}$$

since  $((M_k)^+)_{k \in \mathbb{N}}$  is a *submartingale* because  $x \mapsto x^+$  is a non-decreasing convex function, cf. Exercise 10.9-(b), hence

$$\mathbb{P}\left(\max_{k=0,1,\dots,n} M_k \geq x\right) \leq \frac{\mathbb{E}[(M_n)^+]}{x}, \quad x > 0.$$

h) We have

$$\begin{aligned} x\mathbb{P}\left(\max_{k=0,1,\dots,n} M_k \geq x\right) &= x\mathbb{P}(\tau_x \leq n) \\ &= x\mathbb{E}[\mathbb{1}_{\{\tau_x \leq n\}}] \\ &\leq \mathbb{E}[M_{\tau_x \wedge n} \mathbb{1}_{\{\tau_x \leq n\}}] \\ &\leq \mathbb{E}[M_{\tau_x \wedge n}] \\ &\leq \mathbb{E}[M_0]. \end{aligned}$$

i) We have

$$\begin{aligned} x\mathbb{P}\left(\max_{k=0,1,\dots,n} \phi(M_k) \geq x\right) &= x\mathbb{E}[\mathbb{1}_{\{\max_{k=0,1,\dots,n} \phi(M_k) \geq x\}}] \\ &= x\mathbb{E}[\mathbb{1}_{\{\tau_x^\phi \leq n\}}] \\ &= \mathbb{E}\left[\phi\left(M_{\tau_x^\phi \wedge n}\right) \mathbb{1}_{\{\tau_x^\phi \leq n\}}\right] \\ &= \mathbb{E}[\phi(M_{\tau_x^\phi \wedge n})] \\ &\leq \mathbb{E}[\phi(M_n)], \end{aligned}$$

where the last inequality follows from Exercise 10.6-(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  $S_n = X_1 + \dots + X_n$  and  $(X_k)_{k \geq 1}$  is a sequence of independent identically distributed Bernoulli random variables with  $p = \mathbb{P}(X_k = 1)$  and  $q = 1 - p = \mathbb{P}(X_k = -1)$ ,  $k \geq 1$ . Then  $Z_n := e^{-n} M_n$  will be a *supermartingale*, since

$$\begin{aligned} \mathbb{E}[Z_n \mid \mathcal{F}_k] &= e^{-n} \mathbb{E}[M_n \mid \mathcal{F}_k] \\ &= e^{-n} M_k \\ &\leq e^{-k} M_k \\ &= Z_k, \quad k = 0, 1, \dots, n. \end{aligned}$$

Exercise 10.11 We have

$$\mathbb{E}[(M_n)^r] = \mathbb{E}\left[\left(\frac{q}{p}\right)^{rS_n}\right]$$

$$\begin{aligned}
&= \mathbb{E} \left[ \prod_{k=1}^n \left( \frac{q}{p} \right)^{r(S_k - S_{k-1})} \right] \\
&= \prod_{k=1}^n \mathbb{E} \left[ \left( \frac{q}{p} \right)^{r(S_k - S_{k-1})} \right] \\
&= \prod_{k=1}^n \left( p \left( \frac{q}{p} \right)^r + q \left( \frac{q}{p} \right)^{-r} \right) \\
&= \left( p \left( \frac{q}{p} \right)^r + q \left( \frac{p}{q} \right)^r \right)^n \\
&= \left( \frac{pq^{2r} + qp^{2r}}{(pq)^r} \right)^n \\
&= (\mathbb{E}[(M_1)^r])^n.
\end{aligned}$$

We note that by Jensen's inequality we have

$$\mathbb{E}[(M_1)^r] \geq (\mathbb{E}[M_1])^r = 1,$$

and  $\mathbb{E}[(M_n)^r]$  is non-decreasing in  $n \geq 0$  as  $((M_n)^r)_{n \geq 0}$  is a *submartingale*. By Problem 10.10-(e), for all  $n \geq 0$  and  $r \geq 1$  we have

$$\begin{aligned}
\mathbb{P} \left( \max_{k=0,1,\dots,n} M_k \geq x \right) &\leq \frac{\mathbb{E}[(M_n)^r]}{x^r} \\
&= \frac{(p(q/p)^r + q(p/q)^r)^n}{x^r}, \quad x > 0. \quad (\text{S.63})
\end{aligned}$$

By (3.3.3), we also check that

$$\begin{aligned}
\mathbb{E}[(M_{2n})^r] &= \mathbb{E} \left[ \left( \frac{q}{p} \right)^{rS_{2n}} \right] \\
&= \sum_{k=-n}^n \left( \frac{q}{p} \right)^{2kr} \mathbb{P}(S_{2n} = 2k) \\
&= \sum_{k=-n}^n \binom{2n}{n+k} \left( \frac{q}{p} \right)^{2kr} p^{n+k} q^{n-k} \\
&= \sum_{k=0}^{2n} \binom{2n}{k} \left( \frac{q}{p} \right)^{2(k-n)r} p^k q^{2n-k} \\
&= \left( \frac{q}{p} \right)^{-2nr} q^{2n} \sum_{k=0}^{2n} \binom{2n}{k} \left( \left( \frac{q}{p} \right)^{2r-1} \right)^k \\
&= \left( p \left( \frac{q}{p} \right)^r + q \left( \frac{p}{q} \right)^r \right)^{2n}.
\end{aligned}$$

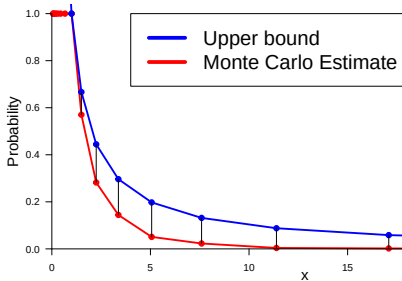


Fig. S.23: Supremum deviation probability with  $n = 7$  and  $p = 0.4$ .

Note also that when  $r = 1$  we find

$$\mathbb{P}\left(\max_{k=0,1,\dots,n} M_k \geq x\right) \leq \frac{\mathbb{E}[M_1]}{x} = \frac{1}{x}, \quad x > 0.$$

The following **R** code together with Figures S.23-S.24 provide a numerical confirmation of the upper bound (S.63) when  $r = 1$ .

```

1  nSim=99999;p=0.4;q=1-p;n=7;a=q/p;r=1;prob=rep(0,2*n+2)
2  par(mar=c(2,3,2,2));par(mgp=c(1,1,0)); for (i in (:-n):(n+1)){for (j in
   1:nSim){
3    M<-a^cumsum(2*rbinom(n,1,p)-1);color="blue"
4    if (max(c(1,M))>=a^i) {prob[n+1+i]=prob[n+1+i]+1;color="darkred"}
5    if ((j%%10000)==0){
6      plot(seq(0,n),c(1,M), xlab = "Time", xlim=c(0,n), type='o',ylim =
       c(0,a^(n-1)), lwd = 3, ylab = "", col = color,
7        main="",xaxs='i',yaxs='i',yaxp = c(0, 11, 11), las = 1)
8      text(3,8,paste("x=",format(round(a^i,2)),",",
9        ",prob[n+1+i],"/",j,"=",format(prob[n+1+i]/j,digits=4)),col=color,cex=2)
10     axis(2, at=c(a^i), labels=c("x"), las=1)
11     lines(seq(0,n),rep(a^i,n+1),col = "black",lty = 2, lwd = 2); text(-0.1,a^i,
12       paste("x"));Sys.sleep(0.9)}
13   prob[n+1+i]=prob[n+1+i]/nSim;x=a^seq(-n,n+1);
14   par(mar=c(2,3,2,2));par(mgp=c(1,1,0))
15   plot(x,prob,type="o",pch=19,lwd=3,col="red",xlab="x",ylim=c(0,1),
16     ylab="Probability",main="",axes=FALSE,cex.axis=1.5, cex.lab=1.5)
17   lines(x,(p*a^r+q*(p/q)^r)^n/x^r,type="o",pch=19,lwd=3,col="blue",main="")
18   axis(1, pos=0,las =1);axis(2, pos=0,las =1);
19   legend(4, 1, legend=c("Upper bound", "Monte Carlo estimate"),
20     col=c("blue","red"), lty=1, lwd=6,cex=2)
21   for (i in 0:(n+1)){segments(x0=a^i, y0=prob[n+1+i],x1=a^i,
22     (p*a^r+q*a^(-r))^n/a^(i*r),col="black")}
```

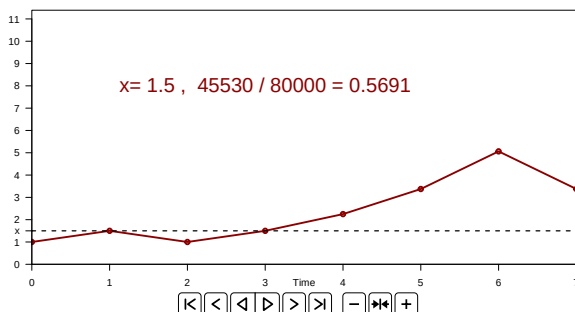


Fig. S.24: Martingale supremum as a function of time.\*

## 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 (11.1.1) we find that given  $\{N_T = k\}$ , the distribution of  $(T_1, T_2, \dots, T_k)$  is uniform on

$$\{(t_1, t_2, \dots, t_k) : 0 \leq t_1 \leq \dots \leq t_k \leq T\}.$$

This yields

$$\mathbb{E}[T_1 \mid N_T = 2] = \frac{2}{T^2} \int_0^T \int_0^y x dx dy = \frac{1}{T^2} \int_0^T y^2 dy = \frac{T}{3},$$

and

$$\mathbb{E}[T_2 \mid N_T = 2] = \frac{2}{T^2} \int_0^T y \int_0^y x dx dy = \frac{2}{T^2} \int_0^T y^2 dy = \frac{2T}{3},$$

and we check that

$$\mathbb{E}[T_1 + T_2 \mid N_T = 2] = T = \frac{2}{T} \int_0^T x dx,$$

which is consistent with the fact that the (unordered) locations of the two jump times are uniformly distributed on  $[0, T]$  given that  $N_T = 2$ , see also (11.1.1) page 334.

\* The animation works in Acrobat Reader.

On the other hand, by the memoryless property of the standard Poisson process we have

$$\mathbb{E}[T_3 \mid N_T = 2] = T + \mathbb{E}[T_1 \mid N_T = 0] = T + \frac{1}{\lambda}.$$

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

$$\begin{aligned}\mathbb{P}(N \geq 3) &= 1 - \mathbb{P}(N \leq 2) = 1 - e^{-10\theta} \left( 1 + 10\theta + \frac{(10\theta)^2}{2} \right) \\ &= 1 - e^{-6} \left( 1 + 6 + \frac{6^2}{2} \right) = 1 - 25e^{-6} \simeq 0.938.\end{aligned}$$

Exercise 11.4 Let  $X_A$ , resp.  $X_B$ , the number of defects found by the first, resp. second, inspection. We know that  $X_A$  and  $X_B$  are independent Poisson random variables with intensities 0.5, hence the probability that both inspections yield defects is

$$\begin{aligned}\mathbb{P}(X_A \geq 1 \text{ and } X_B \geq 1) &= \mathbb{P}(X_A \geq 1)\mathbb{P}(X_B \geq 1) \\ &= (1 - \mathbb{P}(X_A = 0))(1 - \mathbb{P}(X_B = 0)) \\ &= (1 - e^{-0.5})^2 \simeq 0.2212.\end{aligned}$$

Exercise 11.5 The number  $X_N$  of points in the interval  $[0, \lambda]$  has a binomial distribution with parameter  $(N, \lambda/N)$ , *i.e.*

$$\mathbb{P}(X_N = k) = \binom{N}{k} \left( \frac{\lambda}{N} \right)^k \left( 1 - \frac{\lambda}{N} \right)^{N-k}, \quad k = 0, 1, \dots, N,$$

and we find

$$\lim_{N \rightarrow \infty} \mathbb{P}(X_N = k) = \frac{\lambda^k}{k!} \lim_{N \rightarrow \infty} \left( \left( 1 - \frac{\lambda}{N} \right)^{N-k} \prod_{i=0}^{k-1} \frac{N-i}{N} \right) = e^{-\lambda} \frac{\lambda^k}{k!},$$

which is the Poisson distribution with parameter  $\lambda > 0$ .

Exercise 11.6

a) Based on the area  $\pi r^2 = 9\pi$ , this probability is given by

$$e^{-9\pi/2} \frac{(9\pi/2)^{10}}{10!}.$$

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

$$e^{-9\pi} \frac{(9\pi)^8}{8!}.$$

d) Since the location of points are uniformly distributed by (11.1.1), 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. (S.10) in the solution of Exercise 1.7 and Exercise 9.2-(d), and we find the probability

$$\binom{5}{3} \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^2 = \frac{45}{512} \simeq 0.08789.$$

Exercise 11.7 (Wang et al. (2012)) By the moment identity Relation (11.4.1) we have

$$\begin{aligned} \mathbb{E} \left[ \left| \frac{S_n - \lambda n}{\sqrt{n}} \right|^p \right] &= n^{-p/2} \sum_{k=0}^p (n\lambda)^k S_2(p, k) = n^{-p/2} \sum_{k=0}^{p/2} (n\lambda)^k S_2(p, k) \\ &= n^{-p/2} \sum_{k=0}^{p/2} (n\lambda)^k p^k = n^{-p/2} \frac{(np\lambda)^{1+p/2} - 1}{np\lambda - 1} \\ &\leq \frac{(p\lambda)^{1+p/2}}{p\lambda - 1/n} < C_p, \end{aligned}$$

where  $C_p > 0$  is a finite constant.

Exercise 11.8

a) We have

$$\begin{aligned} M'(s) &= \int_X f(x) (e^{sf(x)} - 1) \sigma(dx) \mathbb{E}_{\mathbb{P}_\sigma^X} \left[ \exp \left( s \int_0^\infty f(y) (dN_y - dy) \right) \right] \\ &= s \int_X |f(x)|^2 \frac{e^{sf(x)} - 1}{sf(x)} \sigma(dx) \mathbb{E}_{\mathbb{P}_\sigma^X} \left[ \exp \left( s \int_0^\infty f(y) (dN_y - dy) \right) \right] \\ &\leq \frac{e^{sK} - 1}{K} \int_X |f(x)|^2 \sigma(dx) \mathbb{E}_{\mathbb{P}_\sigma^X} \left[ \exp \left( s \int_0^\infty f(y) (dN_y - dy) \right) \right] \end{aligned}$$

$$= \alpha^2 \frac{e^{sK} - 1}{K} \mathbb{E}_{\mathbb{P}_\sigma^X} \left[ \exp \left( s \int_0^\infty f(y)(dN_y - dy) \right) \right] = \alpha^2 \frac{e^{sK} - 1}{K} M(s),$$

which shows that

$$\frac{M'(s)}{M(s)} \leq h(s) := \alpha^2 \frac{e^{sK} - 1}{K}, \quad s \geq 0.$$

b) We have

$$\begin{aligned} \log M(t) &= \log M(0) + \int_0^t d \log M(s) \\ &\leq \int_0^t \frac{M'(s)}{M(s)} ds \\ &\leq \int_0^t h(s) ds, \end{aligned}$$

hence

$$M(t) \leq \exp \left( \int_0^t h(s) ds \right) = \exp \left( \alpha^2 \int_0^t \frac{e^{sK} - 1}{K} ds \right), \quad t \geq 0.$$

c) By Markov's inequality, we have

$$\begin{aligned} \mathbb{P}_\sigma^X \left( \int_0^\infty f(y)(dN_y - dy) \geq x \right) &= \mathbb{E}_{\mathbb{P}_\sigma^X} \left[ \mathbb{1}_{\{\int_0^\infty f(y)(dN_y - dy) \geq x\}} \right] \\ &\leq e^{-tx} \mathbb{E}_{\mathbb{P}_\sigma^X} \left[ \mathbb{1}_{\{\int_0^\infty f(y)(dN_y - dy) \geq x\}} \exp \left( t \int_0^\infty f(y) dN_y \right) \right] \\ &\leq e^{-tx} \mathbb{E}_{\mathbb{P}_\sigma^X} \left[ \exp \left( t \int_0^\infty f(y) dN_y \right) \right] \\ &\leq \exp \left( -tx + \int_0^t h(s) ds \right) \\ &\leq \exp \left( -tx + \alpha^2 \int_0^t \frac{e^{sK} - 1}{K} ds \right) \\ &= \exp \left( -tx + \frac{\alpha^2}{K^2} (e^{tK} - tK - 1) \right), \end{aligned}$$

which also yields

$$\mathbb{P} \left( \int_0^\infty f(y)(dN_y - dy) \geq x \right) \leq \exp \left( -tx + \alpha^2 \int_0^t \frac{e^{sK} - 1}{K} ds \right).$$

d) Minimizing the above term in  $t$  with  $t = K^{-1} \log(1 + Kx/\alpha^2)$  shows that

$$\mathbb{P}_\sigma^X \left( \int_0^\infty f(y)(dN_y - dy) \geq x \right) \leq \exp \left( \frac{x}{K} - \left( \frac{x}{K} + \frac{\alpha^2}{K^2} \right) \log \left( 1 + \frac{xK}{\alpha^2} \right) \right)$$

$$\begin{aligned}
&\leq \exp\left(-\frac{x}{2K} \log\left(1 + \frac{xK}{\alpha^2}\right)\right) \\
&= \left(1 + \frac{xK}{\alpha^2}\right)^{-x/2K}.
\end{aligned}$$

### Exercise 11.9

a) We have

$$\begin{aligned}
\mathcal{G}_\Phi(f) &= e^{-\sigma(\mathbb{X})} \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{X}^n} \prod_{i=1}^n f(x_i) \sigma(dx_1) \cdots \sigma(dx_n) \\
&= e^{-\sigma(\mathbb{X})} \sum_{n=0}^{\infty} \frac{1}{n!} \left( \int_{\mathbb{X}} f(x_1) \sigma(dx_1) \right)^n \\
&= \exp\left(\int_{\mathbb{X}} f(x) \sigma(dx) - \sigma(\mathbb{X})\right) \\
&= \exp\left(\int_{\mathbb{X}} (f(x) - 1) \sigma(dx)\right), \quad f \in L^1(\mathbb{X}, \mu).
\end{aligned}$$

b) We have

$$\begin{aligned}
\mathbb{P}(\Phi \cap A = \emptyset) &= \mathbb{E}\left[\mathbf{1}_{\{\Phi \cap A = \emptyset\}}\right] \\
&= \mathbb{E}\left[\prod_{x \in \Phi} \mathbf{1}_{A^c}(x)\right] \\
&= \mathcal{G}_\Phi(\mathbf{1}_{A^c}) \\
&= \exp\left(\int_{\mathbb{X}} (\mathbf{1}_{A^c}(x) - 1) \sigma(dx)\right) \\
&= \exp\left(-\int_{\mathbb{X}} \mathbf{1}_A(x) \sigma(dx)\right) \\
&= e^{-\sigma(A)}.
\end{aligned}$$

c) Letting

$$\mathcal{C} := \{(x, r) \in \mathbb{R}^d \times \mathbb{R}_+ : \|x\| \leq r\},$$

we have

$$\begin{aligned}
\mathbb{P}(0 \in \mathcal{B}) &= 1 - \mathbb{P}(0 \notin \mathcal{B}) \\
&= 1 - \mathbb{P}(\Phi \cap \mathcal{C} = \emptyset) \\
&= 1 - \mathcal{G}_\Phi(\mathbf{1}_{\mathcal{C}^c}) \\
&= 1 - \exp\left(-\int_{\mathbb{R}^d \times \mathbb{R}_+} \mathbf{1}_{\mathcal{C}}(x) \sigma(dx)\right) \\
&= 1 - \exp\left(-\lambda \int_{\mathbb{R}^d \times \mathbb{R}_+} \mathbf{1}_{\{y \in B(0, r)\}} dy \rho(r) dr\right) \\
&= 1 - \exp\left(-\lambda \int_{\mathbb{R}^d} \rho(r) \int_0^\infty \mathbf{1}_{\{y \in B(0, r)\}} dy dr\right)
\end{aligned}$$

$$= 1 - \exp\left(-\lambda v_d \int_{\mathbb{R}^d} \rho(r) r^d dr\right).$$

d) By translation invariance, this probability is given by

$$\mathbb{P}(0 \in \mathcal{B}) = 1 - \exp\left(-\lambda v_d \int_{\mathbb{R}^d} \rho(r) r^d dr\right).$$

Exercise 11.10

a) This probability is given by

$$\exp\left(-\int_{[0,1]^d} \sigma(dy) \int_0^{1/2} e^{-r} dr\right) = e^{-\sigma([0,1]^d)(1-e^{-1/2})}.$$

b) The mean is given by

$$\int_{[0,1]^d} \sigma(dy) \int_0^{1/2} e^{-r} dr = \sigma([0,1]^d)(1-e^{-1/2}).$$

## Chapter 12 - Reliability Theory

Exercise 12.1

a) We have

$$F_\beta(t) = \mathbb{P}(\tau < t) = \int_0^t f_\beta(x) dx = \beta \int_0^t x^{\beta-1} e^{-x^\beta} dx = -\left[e^{-x^\beta}\right]_0^t = 1 - e^{-t^\beta},$$

$$t \geq 0.$$

b) We have

$$R(t) = \mathbb{P}(\tau > t) = 1 - F_\beta(t) = e^{-t^\beta}, \quad t \geq 0.$$

c) We have

$$\lambda(t) = -\frac{d}{dt} \log R(t) = \beta t^{\beta-1}, \quad t \geq 0.$$

d) By (12.3.1), we have

$$\begin{aligned} \mathbb{E}[\tau] &= \int_0^\infty R(t) dt \\ &= \int_0^\infty e^{-t^\beta} dt \\ &= \frac{1}{\beta} \int_0^\infty x^{-1+1/\beta} e^{-x} dx \\ &= \frac{1}{\beta} \Gamma\left(\frac{1}{\beta}\right) \end{aligned}$$

$$= \Gamma\left(1 + \frac{1}{\beta}\right).$$

In particular, this yields  $\mathbb{E}[\tau] = \sqrt{\pi}/2$  when  $\beta = 2$ .

### Exercise 12.2

a) We have

$$F_{\beta}(t) = \mathbb{P}(\tau < t) = \int_0^t g_{\beta}(x) dx = \int_0^t \frac{\beta}{x^{\beta+1}} e^{-x^{-\beta}} dx = \left[ e^{-x^{-\beta}} \right]_0^t = e^{-t^{-\beta}},$$

$t \geq 0$ .

b) We have

$$R(t) = \mathbb{P}(\tau > t) = 1 - F_{\beta}(t) = 1 - e^{-t^{-\beta}}, \quad t \geq 0.$$

c) We have

$$\lambda(t) = -\frac{d}{dt} \log R(t) = -\frac{d}{dt} \log(1 - e^{-t^{-\beta}}) = \frac{\beta t^{-\beta-1}}{e^{t^{-\beta}} - 1}, \quad t \geq 0.$$

d) By (12.3.1), we have

$$\begin{aligned} \mathbb{E}[\tau] &= \int_0^{\infty} R(t) dt \\ &= \int_0^{\infty} (1 - e^{-t^{-\beta}}) dt \\ &= \beta \int_0^{\infty} t^{-\beta} e^{-t^{-\beta}} dt \\ &= \int_0^{\infty} x^{1-(1-1/\beta)} e^{-x} dx \\ &= \Gamma\left(1 - \frac{1}{\beta}\right). \end{aligned}$$

In particular, this yields  $\mathbb{E}[\tau] = \sqrt{\pi}$  when  $\beta = 2$ .

### Exercise 12.3 We have

$$\begin{aligned} \mathbb{P}(\tau > t \mid \mathcal{F}_t) &= \mathbb{P}\left(\int_0^t X_u du < L \mid \mathcal{F}_t\right) \\ &= \exp\left(-\mu \int_0^t X_u du\right) \\ &= \exp\left(-\int_0^t \lambda_u du\right), \quad t \geq 0. \end{aligned}$$

# References

- [1] B.B. Bhattacharya and S. Mukherjee. Inference in Ising models. *Bernoulli*, 24(1):493–525, 2018. (Cited on page 159.)
- [2] D. Bosq and H.T. Nguyen. *A Course in Stochastic Processes: Stochastic Models and Statistical Inference*. Mathematical and Statistical Methods. Kluwer, 1996. (Cited on page 200.)
- [3] W.L. Champion, T.M. Mills, and S.J. Smith. Lost in space. *Math. Scientist*, 32:88–96, 2007. (Cited on page 43.)
- [4] H. Finner, P. Kern, and M. Scheer. On some compound distributions with Borel summands. *Insurance Math. Econom.*, 62:234–244, 2015. (Cited on page 171.)
- [5] 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 page 100.)
- [6] F.A. Haight and M.A. Breuer. The Borel-Tanner distribution. *Biometrika*, 47:143–150, 1960. (Cited on page 171.)
- [7] J. Jacod and P. Protter. *Probability essentials*. Springer-Verlag, Berlin, 2000. (Cited on page 11.)
- [8] N.L. Johnson, A.W. Kemp, and S. Kotz. *Univariate discrete distributions*. Wiley Series in Probability and Statistics. Wiley-Interscience, Hoboken, NJ, third edition, 2005. (Cited on page 172.)
- [9] A. Laforgia. Bounds for modified Bessel functions. *J. Comput. Appl. Math.*, 34(3):263–267, 1991. (Cited on page 199.)
- [10] 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 page 180.)
- [11] G. Pólya. Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Strassenetz. *Mathematische Annalen*, 84:149–160, 1921. (Cited on page 43.)
- [12] G. Pólya and G. Szegő. *Problems and Theorems in Analysis I*. Springer, 1998. Reprint of the 1978 Edition. (Cited on page 171.)
- [13] 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 page 202.)
- [14] 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 218.)