Coordinator
Daniel Contreras Salinas

Applied Stochastic Processes

Exercise sheet 9

Exercise 9.1 Let $(X_n)_{n\geq 0}$ be a homogeneous Markov chain with countable state space E and transition probabilities $(p_{x,y})_{x,y\in E}$. Let $C \subseteq E$ such that $E\backslash C$ is finite. Define $p_{x,C}(n) = \sum_{y\in C} p_{x,y}(n)$. Suppose that for each $x \in E\backslash C$ there exists an $n(x)$ such that $p_{x,C}(n(x)) > 0$. Let $\tau_C = \inf\{n \geq 0 \;:\; X_n \in C\}$, $\varepsilon = \min\{p_{x,C}(n(x)) \;:\; x \in E\backslash C\}$, and $N = \max\{n(x) \;:\; x \in E\backslash C\}$. Show that for all $k \in \mathbb{N}$,

$$\mathbf{P}_x[\tau_C > kN] \leq (1-\varepsilon)^k \quad \forall x \in E.$$

Exercise 9.2 Let $(X_n)_{n\geq 0}$ be a Markov chain with state space $E = \{0, 1, \ldots, N\}$. Let us fix $0 \leq x \leq N$ and suppose that under $\mathbf{P}_x$, X_n is a martingale for the canonical filtration $\mathcal{F}_n$. We define $\tau_y = \inf\{n \geq 0; X_n = y\}$. Suppose that $\mathbf{P}_x[\tau_0 \wedge \tau_N < \infty] > 0$ for every $x \in E$.

(a) Show that 0 and N are *absorbing states*, i.e., $p_{0,0} = p_{N,N} = 1$.

(b) Show that $\mathbf{P}_x[\tau_N < \tau_0] = \frac{x}{N}$.

(c) Consider the Gambler ruin chain. Assume that the gambler starts with $k > 0$ dollars. What is the probability that he/she finishes with 0 dollar?

Exercise 9.3 Wright–Fisher model.
Let us consider the following inheritance model for a particular gene with two alleles A and a. In each generation there are m individuals, each one having 2 alleles of the same gene. Each individual of generation $n+1$ chooses its alleles independently from the other individuals and uniformly among the $2m$ possible alleles of generation n. Let us suppose that there are $k \in \{0, \ldots, 2m\}$ alleles of type A in the generation 0. Let X_n be the number of alleles of type A in generation n.

(a) Prove that $(X_n)_{n\geq 0}$ is a Markov chain and find its transition probability $(p_{i,j})_{0\leq i,j\leq 2m}$.

(b) Show that the probability that the allele a disappears before allele A in some genration is $\frac{k}{2m}$.

Solution 9.1 For $k = 0$ the result is clear. If $x \in C$, then $\mathbf{P}_x[\tau_C > kN] = \mathbf{P}_x[0 > kN] = 0$ for all $k \geq 0$. We will prove the inequality for all $x \in E \backslash C$ and $k \geq 1$ by induction over k. For $x \in E \backslash C$, we have

$$\mathbf{P}_x[\tau_C > N] \leq \mathbf{P}_x[\tau_C > n(x)] \leq 1 - p_{x,C}(n(x)) \leq 1 - \varepsilon. \tag{1}$$

For $k \geq 2$, we have that

$$\mathbf{P}_x[\tau_C > kN] = \sum_{y_1,\dots,y_{kN} \in E \backslash C} \mathbf{P}_x[X_1 = y_1, \dots, X_{kN} = y_{kN}].$$

We know that the probability inside the sum equals

$$\mathbf{P}_x[X_{(k-1)N+1} = y_{(k-1)N+1}, \dots, X_{kN} = y_{kN} \mid X_1 = y_1, \dots, X_{(k-1)N} = y_{(k-1)N}] \\ \cdot \mathbf{P}_x[X_1 = y_1, \dots, X_{(k-1)N} = y_{(k-1)N}]$$

where we decalre this product to be 0 if $\mathbf{P}_x[X_1 = y_1, \dots, X_{(k-1)N} = y_{(k-1)N}] = 0$. By the simple Markov property this is equal to

$$\mathbf{P}_{y_{(k-1)N}}[X_1 = y_{(k-1)N+1}, \dots, X_N = y_{kN}] \cdot \mathbf{P}_x[X_1 = y_1, \dots, X_{(k-1)N} = y_{(k-1)N}]$$

Summing over $y_1, \dots, y_{kN} \in E \backslash C$ and setting $y = y_{(k-1)N}$ gives us

$$\begin{aligned} \mathbf{P}_x[\tau_C > kN] &= \sum_{y \in E \backslash C} \underbrace{\mathbf{P}_y[\tau_C > N]}_{\leq 1-\varepsilon \text{ by (1)}} \cdot \mathbf{P}_x[\tau_C > (k-1)N - 1, X_{(k-1)N} = y] \\ &\leq (1-\varepsilon) \underbrace{\mathbf{P}_x[\tau_C > (k-1)N]}_{\leq (1-\varepsilon)^{k-1} \text{ by ind. hyp.}} \leq (1-\varepsilon)^k. \end{aligned}$$

Solution 9.2

(a) Since X_n is a martingale under $\mathbf{P}_0$ we have that $0 = \mathbf{E}_0[X_0] = \mathbf{E}_0[X_1]$ and $X_1 \geq 0$ imply $p_{0,0} = \mathbf{P}_0[X_1 = 0] = 1$. Similarly, $N = \mathbf{E}_N[X_1]$ and $X_n \leq N$ imply $p_{N,N} = P_N[X_n = N] = 1$.

(b) Set $C = \{0, N\}$, then $\tau_C = \tau_0 \wedge \tau_N$. Since the set C is accessible from any point x (which means there is a path $x = x_0, x_1, \dots, x_m$, $x_m \in C$ such that $p_{x_i,x_{i+1}} > 0$), we know that the hypotheses in Exercise 9.1 hold. This implies that $\mathbf{P}_x[\tau_C = \infty] \leq \lim_{k\to\infty}(1-\varepsilon)^k = 0$. This means that $\mathbf{P}_x[\tau_C < \infty] = 1$ and then $(X_{n \wedge \tau_C})_{n \geq 0}$ is also a martingale. The martingale property and the bounded convergence theorem imply

$$x = \mathbf{E}_x[X_{\tau_C \wedge n}] \xrightarrow[n\to\infty]{} \mathbf{E}_x[X_{\tau_C}] = N \cdot \mathbf{P}_x[\tau_N < \tau_0] + 0 \cdot \mathbf{P}_x[\tau_N > \tau_0]$$

which gives us what we wanted (this last property is also known as Doob's Optional Stopping Theorem).

(c) Let us suppose that the Gambler stops when he/she has either N or 0 dollars. Let us denote by X_n the amount of dollars the Gambler has at time $n \geq 0$, and by Y_n the coin flipping at time n (which is independent of $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$). We see that for $k \in \{1, \dots, N-1\}$

$$\begin{aligned} \mathbf{E}_k[X_{n+1}|\mathcal{F}_n] &= \mathbf{E}_k[X_{n+1}1_{\{Y_n=0\}}|\mathcal{F}_n] + \mathbf{E}_k[X_{n+1}1_{\{Y_n=1\}}|\mathcal{F}_n] \\ &= \frac{1}{2}\mathbf{E}_k[X_n - 1|\mathcal{F}_n] + \frac{1}{2}\mathbf{E}_k[X_n + 1|\mathcal{F}_n] = X_n \end{aligned}$$

and for $k \in \{0, N\}$ we have $X_{n+1} = X_n$. This means that $(X_n)_{n\geq 0}$ is a martingale under $\mathbf{P}_k$, for every k. We know that $\mathbf{P}_k[\tau_0 \wedge \tau_N < \infty] \geq 2^{-N} > 0$, which implies that the hypothesis of the previous parts holds. Therefore, the probability that the Gambler finishes with 0 dollars is given by $\mathbf{P}_k[\tau_0 < \tau_N] = 1 - k/N$.

Solution 9.3

(a) We know from the definition of the model that the value of X_{n+1} depends only on the value of X_n, which means that $(X_n)_{n\geq 0}$ is a Markov chain. We can see it explicitly in the following way. For every $n \geq 1$ we let $(Y_i^{(n)})_{1\leq i\leq 2m}$ be i.i.d. uniform random variables on the set $\{1, \dots, 2m\}$. These random variables will represent he choices of the individuals of generation n. Therefore

$$X_{n+1} = \sum_{i=1}^{2m} 1_{\{Y_i^{(n+1)} \leq X_n\}} = \Phi\left(X_n, (Y_1^{(n+1)}, \dots, Y_{2m}^{(n+1)})\right),$$

where Φ is a measurable function. Notice that 0 and $2m$ are absorbing states. We can find explicitly the transition probability for i, j not both equal to 0 or $2m$ by observing that to have $X_{n+1} = j$ given that $X_n = i$ we need to choose j times the allele A among the i available, and $2m - j$ alleles of type a among the $2m - i$ in generation n. This choice has probability $(i/2m)^j((2m-i)/2m)^{2m-j}$. Now we need to assign this i alleles to individuals in generation $n+1$. This can be done in exactly $\binom{2m}{j}$ different ways. Therefore

$$p_{i,j} = \binom{2m}{j}\left(\frac{i}{2m}\right)^j\left(\frac{2m-i}{2m}\right)^{2m-j}, \; p_{0,0} = p_{2m,2m} = 1.$$

(b) Using the simple Markov property we have for $k \in \{1, \dots, 2m-1\}$ that

$$\begin{aligned}
\mathbf{E}_k[X_{n+1}|\mathcal{F}_n] &= \sum_{i=0}^{2m} \mathbf{E}_k[X_{n+1}|\mathcal{F}_n]1_{\{X_n=i\}} \\
&= \sum_{i=0}^{2m} \mathbf{E}_i[X_1]1_{\{X_n=i\}} \\
&= \sum_{i=0}^{2m} 1_{\{X_n=i\}} \sum_{j=1}^{2m} j\mathbf{P}_i[X_1 = j] \\
&= \sum_{i=0}^{2m} 1_{\{X_n=i\}} \sum_{j=1}^{2m} j \cdot \frac{2m}{j}\binom{2m-1}{j-1}\frac{i}{2m}\left(\frac{i}{2m}\right)^{j-1}\left(\frac{2m-i}{2m}\right)^{(2m-1)-(j-1)} \\
&= \sum_{i=0}^{2m} 1_{\{X_n=i\}} i \left(\frac{i}{2m} + \frac{2m-i}{2m}\right)^{2m-1} \\
&= \sum_{i=0}^{2m} X_n 1_{\{X_n=i\}} = X_n.
\end{aligned}$$

For $k \in \{0, N\}$ we have $X_{n+1} = X_n$. This implies that $(X_n)_{n\geq 0}$ is a martingale under $\mathbf{P}_k$, for every k. We also know that $\mathbf{P}_k[\tau_0 \wedge \tau_N < \infty] \geq p_{k,0} > 0$. Therefore, we can use Exercise 9.2 to conclude that the probability that the allele a disappears before allele A in some generation is $\mathbf{P}_k[\tau_{2m} < \tau_0] = k/2m$.