Solution to Homework 2

Exercise 1.13
Suppose \( x \in S \) are such that it attains the minimum for \( \frac{\pi_1(x)}{\pi_2(x)} \). For any other state \( y \), \( \exists t > 0 \) such that \( P^t(y, x) > 0 \) since the chain is irreducible. We have

\[
\frac{\pi_1(x)}{\pi_2(x)} = \frac{\sum_{z \in S} P^t(z, x) \pi_1(z)}{\sum_{z \in S} P^t(z, x) \pi_2(z)} = \frac{\sum_{z \in S} P^t(z, x) \pi_2(z) \pi_1(z) / \pi_2(z)}{\sum_{z \in S} P^t(z, x) \pi_2(z)} = \sum_{z \in S} \left( \frac{P^t(z, x) \pi_2(z)}{\sum_{z \in S} P^t(z, x) \pi_2(z)} \right) \pi_1(z) = \sum_{z \in S} \left( \frac{P^t(z, x) \pi_2(z)}{\sum_{z \in S} P^t(z, x) \pi_2(z)} \pi_2(z) \right) = \pi_1(x) / \pi_2(x),
\]

where equality holds if and only if \( \frac{\pi_1(z)}{\pi_2(z)} = \frac{\pi_1(x)}{\pi_2(x)} \) for all \( z \) with \( P^t(z, x) > 0 \), therefore \( \frac{\pi_1(z)}{\pi_2(z)} = \frac{\pi_1(y)}{\pi_2(y)} \).

Therefore the ratio on every element between any two stationary distributions is constant, however, both sum to 1, so they are identical. This gives an alternative proof for the uniqueness of stationary distribution.

Exercise 2.7
Let \( U \) be the uniform distribution on finite group \( G \). If it is reversible, then for any \( g, h \in G \),

\[ U(g)P(g, h) = \frac{\mu(hg^{-1})}{|G|} = \frac{\mu(g^{-1}h)}{|G|} = U(h)P(h, g); \]

therefore \( \mu(hg^{-1}) = \mu(g^{-1}h) \), take \( h = id \) we have \( \mu(g^{-1}) = \mu(g) \) for any \( g \in G \) or the increment distribution is symmetric.

A Markov chain solution to the secretary problem

(a) The transition probability is \( P(i, j) = \frac{i}{j(j-1)} \) for \( 1 \leq i < j \leq m \), while \( P(i, j) = \frac{i}{m} \) for \( 1 \leq i \leq m \) and \( j = m + 1 \), and of course \( P(m + 1, m + 1) = 1 \).

Note that \( P(1, j) = P(X_1 = j) = P(j \text{ the best}, 1 \text{ the second best among } \{1, \ldots, j\}) = \frac{1}{j(j-1)} \), for \( 1 \leq j \leq m \), and \( P(1, m + 1) = P_1(X_1 = m + 1) = P(1 \text{ the overall best}) = \frac{1}{m} \).

In general, for \( 1 \leq i < j \leq m \),

\[
P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1) \]

\[
= \frac{P(j \text{ the best}, i \text{ the second best among } \{1, \ldots, j\}, \ldots, 1 \text{ the second best among } \{1, \ldots, i_1\})}{P(i \text{ the best}, i_{n-1} \text{ the second best among } \{1, \ldots, i\}, \ldots, 1 \text{ the second best among } \{1, \ldots, i_1\})} = \frac{1/j(j-1) \cdot 1/(i-1) \cdot 1/(i_{n-1} - 1) \cdots 1/(i_1 - 1)}{1/i(i-1) \cdot 1/(i_{n-1} - 1) \cdots 1/(i_1 - 1)} = \frac{i}{j(j-1)}.
\]
For $j = m + 1$, $1 \leq i \leq m$,

$P(i, j) = P(X_{n+1} = m + 1, X_n = i, X_{n-1} = i_{n-1}, ..., X_1 = i_1)$

$= P(i$ the best among all, ..., $1$ the second best among $\{1, ..., i_1\})$

$= 1/m \cdot 1/(i - 1) \cdot 1/(i_{n-1} - 1) ... 1/(i_1 - 1)$

Finally, $P(m + 1, m + 1) = 1$. This shows that $\{X_n\}$ is a Markov chain with the required transition matrix.

(b) If we stop at any positions other than some $X_r$, the probability of getting the best candidate is 0, therefore the optimal rule must stop at some $X_r$. Recall state $X_r = j$ implies $j$ is the best among $\{1, ..., r\}$. The probability that it is best overall is $P(j, m + 1) = j/m$ and we must compare this probability with the probability of all other paths to $m + 1$. By induction, the comparison favors stopping monotonically in the state space; i.e. $d(X_r = m) = ... = d(X_r = k + 1) = 1$ while $d(X_r = k) = ... = d(X_r = 1) = 0$, and $k$ is the largest integer with

$$\frac{1}{k} + \frac{1}{m - 1} > 1.$$

Finally using Riemann integration approximation,

$$\frac{1}{k} + \cdots \frac{1}{m - 1} \approx \int_k^m \frac{dy}{y} = \ln(m/k) = 1 \Rightarrow k/m \sim e^{-1}$$

when $m \to \infty$. The interpretation is to reject all candidates before $k + 1$ (overall position, not $X_{k+1}$), where $k/m \sim e^{-1}$, then pick the first candidate after $k$ who is the best among all previous candidates.

(c) Similar to one pick situation, one must stop at some $X_r$, but now we can stop 2 times instead of 1. Therefore the optimal strategy should first reject $r - 1$ candidates, after that, pick the first candidate who is better than anyone earlier, if the first such candidate is before round $(s - 1)$, we reject all candidates until round $s - 1$. After that, the second candidate is chosen to be the first to be better than anyone before him/her. If the first candidate happens after $(s - 1)$, then we directly start to pick the second one who is better than anyone before. When $m$ is large, it can be shown that the optimal choice for $r, s$ should be $r/m \sim e^{-1.5}$, $s/m \sim e^{-1}$.

**Harmonic function**

(a) A finite-state Markov Chain must contain at least one recurrent state $x$. Any other state $y$ communicates with $x, y \leftrightarrow x$. This implies $\exists$ finite $t_1, t_2 > 0$, s.t. $P_x(X_{t_1} = y) > 0, P_y(X_{t_2} = x) > 0$, therefore it follows that for all $n > 0$,

$$P_y(X_{t_1 + t_2 + n} = y) \geq P_x(X_{t_1} = y)P_y(X_{t_2} = x)P_x(X_n = x) > c_0P_x(X_n = x),$$

Now since $x$ is recurrent,

$$\sum_n P_x(X_n = x) = \infty,$$

therefore

$$\sum_n P_y(X_{t_1 + t_2 + n} = y) = \infty,$$

so state $y$ is recurrent as well.
(b) This should inherit the condition from part (a), otherwise we can have $h$ take different values on groups of states that do not communicate with each other.

Using the Markov property of $\{X_n\}$,

$$E[h(X_{n+1})|X_1, ..., X_n] = E[h(X_{n+1})|X_n] = \sum_{j \in S} p(X_n, j)h(j) = h(X_n)$$

also $h(\cdot)$ must be bounded since it takes values on a finite set, therefore $h(X_n)$ forms a bounded martingale, and must a.s. converge to some $Z_\infty$ by Doob’s martingale convergence theorem. Now using irreducibility and recurrence of $X_n$, $X_n = i$, infinitely often (i.o.) for all $i \in S$, this implies $h(X_n) = h(i)$, i.o. Note that $h(i)$ is a non random constant, this is possible if and only if $h(i) = h(j)$ for all $i, j \in S$, i.e., $h(\cdot)$ is constant.

This gives a different proof from that in the textbook.

(c) For any state $i$,

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

therefore $h(\cdot)$ is a harmonic function, it is not a constant as $h(i) \neq h(j)$ for $i \neq j$. 