Review of probability

1. It's sufficient to show that if $A_1, \cdots, A_n$ are independent events, then $A_1, \cdots, A_{n-1}, A_n^c$ are independent events. Suppose $A_1, A_2, \cdots, A_n$ are independent. Then

$$
\begin{aligned}
P(A_1 \cdots A_n^c) &= P(A_1 \cdots A_{n-1}) - P(A_1 \cdots A_n) = \prod_{i=1}^{n-1} P(A_i) - \prod_{i=1}^{n} P(A_i) \\
&= \prod_{i=1}^{n-1} P(A_i)(1 - P(A_n)) = \prod_{i=1}^{n-1} P(A_i) P(A_n^c).
\end{aligned}
$$

So $A_1, \cdots, A_{n-1}, A_n^c$ are independent. □

2. $P(A) = 1/2$ since there are 2 faces out of 4 faces that has color A. The pairwise-independent follows by $P(B, A) = 1/4 = P(A)P(B)$. But A, B and C are not independent since $P(A, B, C) = 1/4 \neq 1/8$. □

3. (**Exercise 1.5.4**) Let N be the number of squares in $G(n, p)$. Then

$$
\begin{aligned}
\mathbb{E}N &= \mathbb{E} \sum_{\{v_1,v_2,v_3,v_4\} \subset V(G)} \mathbb{1}_{\{\{v_1,v_2,v_3,v_4\} \text{ forms a square}\}} \\
&= \sum_{\{v_1,v_2,v_3,v_4\} \subset V(G)} \mathbb{P}\left(\{v_1, v_2, v_3, v_4\} \text{ forms a square}\right) \\
&= \binom{n}{4} \frac{4!}{4 \cdot 2} p^4 (1-p)^2.
\end{aligned}
$$

The reason of $\frac{4!}{4 \cdot 2}$ is that the number of cycle permutations of $v_1, \cdots, v_4$ is $4!/4$. However since $G(n, p)$ is an undirected graph, we consider two cycle permutations that have opposite directions as the same kind. So there are $4!/(4 \cdot 2)$ ways.

□

4. (**Exercise 1.5.7**) Define N_k to be the number of k-cycle in a random permutation. Then

$$
N_k = \sum_{\substack{I_k \subset \{1, \cdots, n\} \\ |I_k| = k}} \mathbb{1}_{\{I_k \text{ forms a } k\text{-cycle}\}}.
$$

Therefore

$$
\mathbb{E}N_k = \sum_{\substack{I_k \subset \{1, \cdots, n\} \\ |I_k| = k}} \mathbb{P}(I_k \text{ forms a } k\text{-cycle}) = \binom{n}{k} \times \frac{(n-k)!(k-1)!}{n!} = \frac{1}{k}.
$$

Hence the expected number of cycles is $\sum_{k=1}^{n} \mathbb{E}N_k = \sum_{k=1}^{n} \frac{1}{k}$.

□

5. (**Exercise 1.5.8**) Let N be the number of records in a random permutation. Then

$$\mathbb{E}N = \mathbb{E}\sum_{k=2}^{n} \mathbb{1}_{\{\pi_k \text{ is a record}\}} = \sum_{k=2}^{n} \mathbb{P}(\pi_k \text{ is a record}).$$

Define $p_k = \mathbb{E}N$. Then

$$\begin{aligned} p_k = \sum_{m=k}^{n} \mathbb{P}(\pi_k = m,\ \pi_k \text{ is a record}) &= \sum_{m=k}^{n} \frac{\binom{m-1}{k-1}(k-1)!(n-k)!}{n!} \\ &= \frac{1}{k\binom{n}{k}} \sum_{m=k}^{n} \binom{m-1}{k-1}. \end{aligned} \tag{1}$$

The summation $\sum_{m=k}^{n} \binom{m-1}{k-1}$ can be calculated via mathematical induction. For positive integer k, define $S_{n,k}$ as

$$S_{n,k} = \sum_{\ell=0}^{n-1} \binom{k+\ell}{k}.$$

Then we claim that $S_{n,k} = \binom{k+n}{k+1}$. Indeed, when $n = 1$, $S_{n,k} = S_{1,k} = \binom{k}{k} = \binom{k+1}{k+1}$. Next, we make the induction hypothesis that $S_{n,k} = \binom{k+n}{k+1}$ for $n \geq 1$. Suppose the hypothesis is true. Then

$$S_{n+1} = S_n + \binom{k+n}{k} = \binom{k+n}{k+1} + \binom{k+n}{k} = \binom{k+n+1}{k+1}.$$

Therefore by mathematical induction we have proved the claim.

Hence by (1),

$$\begin{aligned} p_k = \frac{1}{k\binom{n}{k}} \sum_{m=k}^{n} \binom{m-1}{k-1} &= \frac{1}{k\binom{n}{k}} \sum_{\ell=0}^{n-k} \binom{k-1+\ell}{k-1} \\ &= \frac{1}{k\binom{n}{k}} S_{n-k+1,k-1} = \frac{1}{k\binom{n}{k}} \binom{n}{k} = \frac{1}{k}. \end{aligned}$$

Therefore $\mathbb{E}N = \sum_{k=2}^{n} \frac{1}{k}$.

□

The maximum clique size. For the random graph $G(n,p)$, define $N_k(n)$ as the number of the size-k cliques in $G(n,p)$. The maximum clique size of $G(n,p)$ is the random variable

$$\omega(n) = \max\{k : N_k(n) \geq 1\}.$$

Use the property of $k_0(n)$ to verify that (see **Exercise 2.4.1**)

$$\limsup_{n\to\infty} \frac{\mathbb{E}\omega(n)}{\log n} \leq \frac{2}{\log(1/p)}.$$

2.4

Exercise 2.4.1. Denote $\omega(n)$ to be the clique number of $G(n,p)$, $N_q(n)$ to be the number of size-q cliques of $G(n,p)$. Recall that $k_0(n) := \max\{q : \mathbb{E}N_q(n) > 0\}$. Given $\epsilon > 0$, we know that for sufficiently large n,

$$\frac{(2-\epsilon)\log n}{\log(1/p)} < k_0(n) < \frac{(2+\epsilon)\log n}{\log(1/p)}$$

by **Lemma 1.9**. Therefore, for sufficiently large n we can take some $k \in \mathbb{N}$ such that

$$\frac{(2-\epsilon)\log n}{\log(1/p)} \le k < k_0(n).$$

Then by **Theorem 2.4** and the fact that $\mathbb{E}N_{k_0-m} \ge n^{m(1-\epsilon)}$ (eq. 2.23) for $m \ge 1$,

$$\mathbb{P}\left(\left|\frac{N_k}{\mathbb{E}N_k} - 1\right| \ge 1\right) \le 2\left(c_p\frac{(\log n)^4}{n^2} + \frac{1}{n^{1-\epsilon}}\right), \tag{2}$$

where $c_p := \frac{2}{\log(1/p)}$. Since $\omega(n) > 0$, we thus have

$$\begin{aligned}
\mathbb{E}\omega(n) &\ge k\mathbb{P}(\omega(n) \ge k) \\
&= k\mathbb{P}(N_k > 0) \\
&= k\mathbb{P}\left(\frac{N_k}{\mathbb{E}N_k} > 0\right) \\
&\ge k\mathbb{P}\left(\left|\frac{N_k}{\mathbb{E}N_k} - 1\right| < 1\right) \\
&\ge \frac{(2-\epsilon)\log n}{\log(1/p)}\left(1 - 2c_p\frac{(\log n)^4}{n^2} - \frac{2}{n^{1-\epsilon}}\right) \text{ by (2).}
\end{aligned}$$

That is,

$$\liminf_{n\to\infty}\frac{\mathbb{E}\omega(n)}{\log n} \ge \lim_{n\to\infty}\frac{2-\epsilon}{\log(1/p)}\left(1 - 2c_p\frac{(\log n)^4}{n^2} - \frac{2}{n^{1-\epsilon}}\right) = \frac{2-\epsilon}{\log(1/p)}. \tag{3}$$

For the other direction, recall from (1.114) that states when $k_0(n) > \frac{(2-\epsilon)\log n}{\log(1/p)}$, we have

$$\mathbb{P}\left(N_{k_0(n)+m+1} > 0\right) \le \mathbb{E}N_{k_0(n)+m+1} \le \frac{1}{n^{m(1-\epsilon)}}$$

for $m \ge 1$. Therefore if n is sufficiently large, we can write

$$\begin{aligned}
\mathbb{E}\omega(n) &\le n\mathbb{P}\left(\omega(n) \ge k_0(n) + 3\right) + k_0(n) + 2 \\
&= n\mathbb{P}\left(N_{k_0(n)+3}(n) > 0\right) + k_0(n) + 2 \\
&\le \frac{n}{n^{2-2\epsilon}} + \frac{(2+\epsilon)\log n}{\log(1/p)} + 2.
\end{aligned}$$

That is,

$$\limsup_{n\to\infty}\frac{\mathbb{E}\omega(n)}{\log n} \le \lim_{n\to\infty}\left[\frac{2+\epsilon}{\log(1/p)} + \frac{1}{n^{1-2\epsilon}\log n} + \frac{2}{\log n}\right] = \frac{2+\epsilon}{\log(1/p)}. \tag{4}$$

Since $\epsilon > 0$ is arbitrary, by (3) and (4) we concludes

$$\lim_{n\to\infty} \frac{\mathbb{E}\omega(n)}{\log n} = \frac{2}{\log(1/p)}.$$

$\square$

Second moment calculations

1. (Exercise 2.2.4) Since $X_1, X_2, \cdots, X_n$ are uncorrelated, $\text{Cov}(X_i, X_j) = 0$ if $i \neq j$. Therefore,

$$\text{Var}(\overline{X}_n) = \frac{1}{n^2} \sum_{1 \le i,j \le n} \text{Cov}(X_i, X_j) = \frac{1}{n^2} \sum_{i=1}^{n} \text{Var}(X_i) \le \frac{\sigma^2}{n}.$$

Then by Markov inequality,

$$\mathbb{P}(|\overline{X}_n - \mu| \ge \epsilon) \le \frac{\mathbb{E}|\overline{X}_n - \mathbb{E}\overline{X}_n|^2}{\epsilon^2} \le \frac{\sigma^2}{n\epsilon^2}.$$

□

2. (Exercise 2.2.5) Note that $\mathbb{E}X_i^2 = \frac{i}{i \log(2i)}$. And $\mathbb{E}\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}X_k = 0$. By the i.i.d. assumption of $X_1, X_2, \cdots$, when n is sufficiently large,

$$\begin{aligned} \text{Var}(\overline{X}_n) = \mathbb{E}\overline{X}_n^2 = \frac{1}{n^2} \sum_{1 \le i,j \le n} \mathbb{E}X_iX_j = \frac{1}{n^2} \sum_{i=1}^{n} \mathbb{E}X_i^2 &\le \frac{1}{n^2} n \max_{1 \le i \le n} \frac{i}{\log(2i)} \\ &= \frac{1}{\log(2n)}. \end{aligned}$$

Therefore $\text{Var}(\overline{X}_n) \to 0$ as $n \to \infty$. And by Markov's inequality, $\mathbb{P}(|\overline{X}_n| \ge \epsilon) \to 0$ as $n \to \infty$.

□

3. (Exercise 2.2.6) By the i.i.d. assumption of $X_1, \cdots, X_n$,

$$\mathbb{E}U = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} \mathbb{E}X_iX_j = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} \mu^2 = \mu^2.$$

Therefore, $\mathbb{E}|U - \mu^2| = \text{Var}(U)$. And

$$\text{Var}(U) = \binom{n}{2}^{-2} \sum_{\substack{1 \le i < j \le n \\ 1 \le i' < j' \le n}} \text{Cov}(X_iX_j, X_{i'}X_{j'}). \tag{5}$$

There are 3 cases in the sum in the RHS of (5).

(1) $\{i, j\} = \{i', j'\}$: $\text{Cov}(X_iX_j, X_{i'}X_{j'}) = \mathbb{E}X_i^2\mathbb{E}X_j^2 - \mu^4 = (\sigma^2 + \mu^2)^2 - \mu^4$.

(2) $\{i, j\} \neq \{i', j'\}$ and $\{i, j\} \cap \{i', j'\} \neq \emptyset$: This means $\{i, j\}$ and $\{i', j'\}$ share one common index. And

$$\text{Cov}(X_iX_j, X_{i'}X_{j'}) = \mathbb{E}X_i^2\mathbb{E}X_j\mathbb{E}X_{j'} - \mu^4 = \mu^2(\mathbb{E}X_i^2 - \mu^2) = \sigma^2\mu^2.$$

(3) $\{i, j\} \cap \{i', j'\} = \emptyset$: $\text{Cov}(X_iX_j, X_{i'}X_{j'}) = \mu^4 - \mu^4 = 0$

Note that the case (1) has $\binom{n}{2}$ choices. Case (3) has $\binom{n}{2}\binom{n-2}{2}$ choices. For case (2), first pick 3 different indices $i < j < k$. One can choose one of these 3 indices as the common index of $\{i, j\}$ and $\{i', j'\}$. So there are $\binom{n}{3} \times 3 \times 2$ choices. Therefore,

$$\mathrm{Var}(U) = \binom{n}{2}^{-2} \left[\binom{n}{2} (\sigma^4 + 2\sigma^2\mu^2) + 6\binom{n}{3}\sigma^2\mu^2 \right].$$

And $\mathrm{Var}(U) = O(1/n)$ as $n \to \infty$. Hence by Markov inequality, $\mathbb{P}(|U - \mu^2| \geq \epsilon) \to 0$ as $n \to \infty$.

□

4. (Exercise 2.2.7) The following inequality is immediate:

$$\begin{aligned}
\mathbb{P}(\max_{i \leq n} |X_i| \geq \epsilon n^{1/p}) = \mathbb{P}(\max_{i \leq n} |X_i|^p \geq \epsilon^p n) &\leq \mathbb{P}\left(\bigcup_{i \leq n} |X_i|^p \geq \epsilon^p n \right) \\
&\leq \sum_{1 \leq i \leq n} \mathbb{P}(|X_i|^p \geq \epsilon^p n) \leq \sum_{1 \leq i \leq n} \frac{\mathbb{E}|X_i|^p \mathbb{1}\{|X_i|^p \geq \epsilon^p n\}}{\epsilon^p n} \\
&= \frac{\mathbb{E}|X_1|^p \mathbb{1}\{|X_1|^p \geq \epsilon^p n\}}{\epsilon^p}.
\end{aligned}$$

Since $\mathbb{E}|X_1|^p < \infty$, by monotone convergence we know that $\lim_{n\to\infty} \mathbb{E}|X_1|^p \mathbb{1}\{|X_1|^p \geq \epsilon^p n\} = 0$. Therefore $\mathbb{P}(\max_{i \leq n} |X_i| \geq \epsilon n^{1/p}) \to 0$ as $n \to \infty$.

□

5. (Exercise 2.3.1) Let $x \in [0, 1]$. Consider i.i.d. Bernoulli(x) random variables $X_1, X_2 \cdots$. Then, the Bernstein polynomial of $f(x) = x^2$ can be wrote as

$$\mathbb{E}f(\overline{X}_n) = \mathbb{E}\overline{X}_n^2 = \frac{1}{n^2}\left[n\mathbb{E}X_1^2 + n(n-1)\left(\mathbb{E}X_1\right)^2\right] = \frac{x}{n} + \left(1 - \frac{1}{n}\right)x^2.$$

Therefore

$$\left|\mathbb{E}f(\overline{X}_n) - x^2\right| = \frac{x(1-x)}{n} \leq \frac{1}{4n}.$$

□

6. (Exercise 2.3.2) Consider $X_1, X_2, \cdots$ be the i.i.d. sequence of Bernoulli(x) random variables. Then $B_n(x) = \mathbb{E}f(\overline{X}_n)$, and for any $\delta > 0$,

$$\begin{aligned}
|B_n(x) - f(x)| &\leq \mathbb{E}|f(\overline{X}_n) - f(x)| \\
&= \mathbb{E}\left[\left|f(\overline{X}_n) - f(x)\right| \mathbb{1}_{\{|\overline{X}_n - x| \geq \delta\}} + \left|f(\overline{X}_n) - f(x)\right| \mathbb{1}_{\{|\overline{X}_n - x| < \delta\}}\right] \\
&\leq \mathbb{E}\left[2C\mathbb{1}_{\{|\overline{X}_n - x| \geq \delta\}}\right] + \mathbb{E}\left[D\delta\mathbb{1}_{\{|\overline{X}_n - x| < \delta\}}\right].
\end{aligned}$$

The last line was due to the mean value theorem, $|f(u) - f(v)| = |f'(c)||u - v|$ for some $c \in (u, v)$, and D is the uniform bound of $|f'|$ on $[0, 1]$. Then we have

$$|B_n(x) - f(x)| \leq 2C\mathbb{P}(|\overline{X}_n - x| \geq \delta) + D\delta \leq \frac{C}{2n\delta^2} + D\delta.$$

Since the RHS doesn't depend on x,

$$\max_{x \in [0,1]} |B_n(x) - f(x)| \leq \frac{C}{2n\delta^2} + D\delta = F(\delta).$$

By taking derivative for $F(\delta)$ over $\delta > 0$, we found that F attains minimum when $\delta = \left(\frac{C}{nD}\right)^{1/3}$. Plug into $F(\delta)$, the minimum of F is

$$F\left(\left(\frac{C}{nD}\right)^{1/3}\right) = \frac{3D^{2/3}C^{1/3}}{2n^{1/3}}.$$

This is one of the uniform bounds of $\max_{x\in[0,1]} |B_n(x) - f(x)|$.

$\square$

7. (Exercise 2.3.3) Let $x_1, \cdots, x_m \in [0,1]$. And let $X_1, X_2, \cdots$ be the i.i.d. random vectors with coordinates $X_1^{(1)}, \cdots, X_1^{(m)}$ independently distributed, with the law

$$\mathbb{P}(X_1^{(i)} = 1) = x_i;\ \mathbb{P}(X_1^{(i)} = 0) = 1 - x_i$$

for each $i = 1, \cdots, m$. Note that the sum $S_n = \sum_{i=1}^n X_i$ also has independent coordinates, and its p.m.f. is

$$\mathbb{P}\left(S_n = (k_1, \cdots, k_m)^T\right) = \prod_{i=1}^m \binom{n}{k_i} x_i^{k_i}(1 - x_i)^{n-k_i}.$$

for $0 \le k_i \le n$ for each $i = 1, \cdots, m$. Therefore for $\overline{X}_n := S_n/n$,

$$\mathbb{E}f(\overline{X}_n) = \sum_{0\le k_1,\cdots,k_m\le n} f\left(\frac{k_1}{n}, \cdots, \frac{k_m}{n}\right) \mathbb{P}\left(S_n = (k_1, \cdots, k_m)^T\right).$$

Note that it is a function of $x = (x_1, \cdots, x_n) \in [0,1]^m$. Since f is continuous on $[0,1]^m$, f is bounded by some $C > 0$ and uniformly continuous on $[0,1]^m$. Given $\epsilon > 0$, take $\delta > 0$ s.t. $|f(x_1) - f(x_2)| < \epsilon$ for any $x_1, x_2 \in [0,1]^m$ with $\|x_1 - x_2\|_2 < \delta$. Next,

$$\begin{aligned}
\mathbb{E}\left|f(\overline{X}_n) - f(x)\right| &= \mathbb{E}\left[\left|f(\overline{X}_n) - f(x)\right| \mathbb{1}_{\{\|\overline{X}_n - x\|_2 \ge \delta\}} + \left|f(\overline{X}_n) - f(x)\right| \mathbb{1}_{\{\|\overline{X}_n - x\|_2 < \delta\}}\right] \\
&\le 2C\mathbb{P}(\|\overline{X}_n - x\|_2 \ge \delta) + \epsilon \cdot 1 \\
&= 2C\mathbb{P}\left(\frac{1}{n^2}\sum_{i=1}^m (S_n^{(i)} - nx_i)^2 \ge \delta^2\right) + \epsilon \\
&\le \frac{2C}{n^2\delta^2}\sum_{i=1}^m \mathbb{E}(S_n^{(i)} - nx_i)^2 + \epsilon,
\end{aligned}$$

where the last inequality holds because $S_n^{(1)}, \cdots, S_n^{(m)}$ are independent. We should mention that the notation $\|X\|_2$ here is a real-valued random variable, not the integral norm of X. Now since $X_1, \cdots, X_n$ are i.i.d.,

$$\mathbb{E}(S_n^{(i)} - nx_i)^2 = \mathbb{E}\left[\sum_{j=1}^n (X_j^{(i)} - x_i)\right]^2 = \sum_{j=1}^n \mathbb{E}(X_j^{(i)} - x_i)^2 = nx_i(1 - x_i).$$

Therefore

$$\begin{aligned}
\left|\mathbb{E}f(\overline{X}_n) - f(x)\right| &\le \frac{2C}{n\delta^2}\sum_{i=1}^m x_i(1 - x_i) + \epsilon \\
&\le \frac{Cm}{2n\delta^2} + \epsilon
\end{aligned}$$

The last inequality holds by the fact $x(1-x) \le \frac{1}{4}$ for $x \in [0,1]$. Hence $\mathbb{E}f(\overline{X}_n)$ is uniformly bounded, and

$$\max_{x\in[0,1]^m} \left|\mathbb{E}f(\overline{X}_n) - f(x)\right| \le \frac{Cm}{2n\delta^2} + \epsilon \to \epsilon \text{ as } n \to \infty.$$

Since $\epsilon > 0$ is arbitrary, $\mathbb{E}f(\overline{X}_n)$ converges to f uniformly on $[0,1]^m$.

□

8. (Exercise 2.3.4) Let $[a,b]$ be a finite interval in $[0,\infty)$. Let $x \in [a,b]$, and let $X_1, X_2, \cdots$ be a sequence of i.i.d. Poisson(x) random variables. Then

$$\mathbb{E}f(\overline{X}_n) = \sum_{k=0}^{\infty} f\left(\frac{k}{n}\right)\frac{(nx)^k}{k!}e^{-nx} = P_n(x).$$

Note that $S_n := \sum_{i=1}^n X_i$ has Poisson(nx) distribution, and its p.m.f. is

$$\mathbb{P}(S_n = k) = \frac{(nx)^k}{k!}e^{-nx}$$

for $k \in \mathbb{N}\cup\{0\}$.

Given $\epsilon > 0$. Since f is continuous on $[0,\infty)$, f is uniformly continuous and bounded on the compact set $[a,b]$ by some $C > 0$. Therefore there exists $\delta > 0$ such that $|f(x_1) - f(x_2)| < \epsilon$ for all $x_1, x_2 \in [a,b]$ that satisfy $|x_1 - x_2| < \delta$. Then,

$$\begin{aligned}
\left|\mathbb{E}f(\overline{X}_n) - f(x)\right| &\le \mathbb{E}\left|f(\overline{X}_n) - f(x)\right| \\
&= \mathbb{E}\left[\left|f(\overline{X}_n) - f(x)\right| \mathbb{1}_{\{|\overline{X}_n - x| \ge \delta\}} + \left|f(\overline{X}_n) - f(x)\right| \mathbb{1}_{\{|\overline{X}_n - x| < \delta\}}\right] \\
&\le 2C\mathbb{P}(|\overline{X}_n - x| \ge \delta) + \epsilon \cdot 1 \\
&\le 2C\frac{\text{Var}(\overline{X}_n)}{\delta^2} + \epsilon \\
&= \frac{2Cx}{n\delta^2} + \epsilon
\end{aligned}$$

The last equality holds because $\{X_i\}$ are i.i.d. and $\text{Var}(X_1) = x$. Also note that since $x \in [a,b]$, we have $\max_{a\le x\le b} |P_n(x) - f(x)| \le \frac{2Cb}{n\delta^2} + \epsilon$. Thus

$$\lim_{n\to\infty} \max_{a\le x\le b} |P_n(x) - f(x)| \le \epsilon.$$

Since $\epsilon > 0$ is arbitrary, this completes the proof.

□

3.1 Hoeffding Inequality

Exercise 3.1.1 Let $X_1, \cdots, X_{100000}$ be IID Bernoulli(1/2) random variables. Then the problem is to estimate $P(|\sum_{i=1}^{100000} X_i - 50000| \geq 500)$. Let $Y_i = 2X_i - 1$ for each i. Then Y_i is a sequence of IID Rademacher random variables. Hence by Hoeffding's inequality, for any $t \geq 0$,

$$P\left(\left|\sum_{i=1}^{100000} Y_i\right| \geq t\right) = P\left(\left|\sum_{i=1}^{100000} X_i - 50000\right| \geq \frac{t}{2}\right) \leq 2\exp\left(\frac{-t^2}{200000}\right).$$

Take $t = 1000$, we then get a probability bound e^{-5}.

□

Exercise 3.1.4 Suppose $\lambda \in [0, \frac{1}{4}]$. By the inequality

$$e^x \leq 1 + x + \frac{x^2}{2} + \sum_{k\geq 3} \frac{(x_+)^k}{k!},\ x \in \mathbb{R},$$

we have

$$\mathbb{E}e^{-\lambda Y} \leq 1 - \lambda\mathbb{E}Y + \frac{\lambda^2}{2}\mathbb{E}Y^2 + \sum_{k\geq 3} \frac{\lambda^k}{k!}\mathbb{E}(Y_-)^k,$$

where $Y_- := -\min\{Y, 0\} = -\min\{X^2 - 1, 0\} = \max\{1 - X^2, 0\} \leq 1$. So we have

$$\begin{aligned}
\mathbb{E}e^{-\lambda Y} &\leq 1 - \lambda\mathbb{E}Y + \frac{\lambda^2}{2}\mathbb{E}Y^2 + \sum_{k\geq 3}\frac{\lambda^k}{k!} \\
&= 1 + \frac{\lambda^2}{2}(\mathbb{E}X^4 - 2\mathbb{E}X^2 + 1) + \left(e^\lambda - 1 - \lambda - \frac{\lambda^2}{2}\right) \\
&= \frac{\lambda^2}{2}\mathbb{E}X^4 + e^\lambda - \lambda - \lambda^2 \\
&\leq 7\lambda^2 + e^\lambda - \lambda,
\end{aligned}$$

where the last inequality is due to (3.13) in the textbook, that is, $\mathbb{E}X^{2k} \leq 2^{k+1}k!$ for $k \geq 1$. Now, the remaining task is to bound the RHS. First, observe that for any $t \in [0, 1]$,

$$e^t = 1 + t + \sum_{k\geq 2}\frac{t^k}{k!} \leq 1 + t + \sum_{k\geq 2}\frac{t^k}{2^{k-1}} = 1 + t + \frac{t^2}{2-t} \leq 1 + t + t^2.$$

In the first inequality we have used the fact $k! \geq 2^{k-1}$ for $k \geq 1$. Therefore,

$$e^t - t \leq 1 + t^2 \leq e^{t^2}. \tag{6}$$

This means when $\lambda \in [0, \frac{1}{4}]$,

$$\begin{aligned}
e^{16\lambda^2} &\geq e^{4\lambda} - 4\lambda = 1 + 8\lambda^2 + \sum_{k\geq 3}\frac{(4\lambda)^k}{k!} \\
&\geq 1 + \frac{15}{2}\lambda^2 + \sum_{k\geq 3}\frac{\lambda^k}{k!} = 7\lambda^2 - \lambda + e^\lambda.
\end{aligned} \tag{7}$$

One shall notice that (6) also holds for $t > 1$. That is, (7) holds for all $\lambda \geq 0$. Hence we have shown that $\mathbb{E}e^{-\lambda Y} \leq e^{16\lambda^2}$ for all $\lambda \geq 0$.

$\square$

Exercise 3.1.5. Let $X_1, X_1', \cdots, X_n, X_n'$ be independent random variables with X_i and X_i' have the same law. Set $Y_i := X_i - X_i'$ for $1 \leq i \leq n$, then the distribution of Y_i is symmetric about zero. Further, we set ϵ_i be the Rademacher random variables independent to Y_i. Then for each i, Y_i and $\epsilon_i|Y_i|$ have the same distribution.

Proof of $Y_i \overset{d}{\sim} \epsilon_i|Y_i|$. Suppose $t > 0$. Then

$$
\begin{aligned}
P(\epsilon_i|Y_i| \leq t) &= P(\epsilon_i = -1) + P(\epsilon_i = 1,\ |Y_i| \leq t) \\
&= \frac{1}{2} + \frac{1}{2}(2P(Y_i \leq t) - 1) = P(Y_i \leq t).
\end{aligned}
$$

If $t \leq 0$, then

$$
P(\epsilon_i|Y_i| \leq t) = P(\epsilon_i = -1, |Y_i| \geq |t|) = \frac{1}{2} \cdot 2P(Y_i \leq t) = P(Y_i \leq t).
$$

Therefore Y_i and $\epsilon_i|Y_i|$ have the same distribution.

$\square$

Denote $Y := (Y_1, \cdots, Y_n)^t$, then

$$
\begin{aligned}
P\left(\sum_{i=1}^n (X_i - X_i') > \left[2t\sum_{i=1}^n (X_i - X_i')^2\right]^{1/2}\right) &= P\left(\sum_{i=1}^n \frac{Y_i}{\|Y\|_2} > \sqrt{2t}\right) \\
&= P\left(\sum_{i=1}^n \frac{\epsilon_i|Y_i|}{\|Y\|_2} > \sqrt{2t}\right).
\end{aligned}
$$

Note that the last equality holds because $\|(Y_1, \cdots, Y_n)\|_2 = \|(\epsilon_1|Y_1|, \cdots, \epsilon_n|Y_n|)\|_2$, and $Y_i \overset{d}{\sim} \epsilon_i|Y_i|$. Now define random vector $a = (a_1, \cdots, a_n)$ with $a_i = |Y_i|/\|Y\|_2$. Let $\mu_{(\epsilon,a)}$ be the law of the random element (ϵ, a), then the last probability can be written as

$$
\begin{aligned}
P\left(\sum_{i=1}^n \epsilon_i a_i > \sqrt{2t}\right) &= \int_{\{\pm 1\}^n \times S^{n-1}} \mathbb{1}\left\{\sum_{i=1}^n e_i\alpha_i > \sqrt{2t}\right\} d\mu_{(\epsilon,a)}(e, \alpha) \\
(\epsilon \text{ and } a \text{ are independent}) &= \int_{S^{n-1}} \int_{\{\pm 1\}^n} \mathbb{1}\left\{\sum_{i=1}^n e_i\alpha_i > \sqrt{2t}\right\} d\mu_\epsilon(e) d\mu_a(\alpha) \\
&= \int_{S^{n-1}} P_\epsilon\left(\sum_{i=1}^n \epsilon_i\alpha_i > \sqrt{2t}\right) d\mu_a(\alpha) \\
(\text{Hoeffding's inequality}) &\leq \int_{S_{n-1}} e^{-t} d\mu_a(\alpha) = e^{-t}.
\end{aligned}
$$

Remark. For now, you may think $d\mu_{\epsilon,a}(e, \alpha)$ in the sense of the *joint p.d.f.*, $p_{\epsilon,a}(e, \alpha)d(e, \alpha)$, and you may think P_ϵ as the marginal distribution of ϵ given α.

$\square$

3.2 Johnson-Lindenstrauss Lemma

Exercise 3.2.1. In the proof of **Theorem 3.2**, the author has shown that

$$\mathbb{P}\left(\bigcap_{1\le k<\ell\le m}\left\{\left|\frac{1}{n}\sum_{i=1}^{n}Y_i(a_{k\ell})\right|\le\epsilon\right\}\right)\ge 1-m^2e^{-n\epsilon^2/64}.$$

The RHS $\ge 1-\delta$ if and only if $m^2e^{-n\epsilon^2/64}\le\delta$. By taking logarithms, we can see that this is equivalent to $n>\frac{64}{\epsilon^2}\log(m^2/\delta)$. This implies that the LHS $\ge 1-\delta$. □

3.5 Applications of Azuma Inequality

Exercise 3.5.1. For each X_i, $1\le i\le n$, define a functional $\mu_i:\mathfrak{F}\to\mathbb{R}$ as $\mu_i(f)=\mathbb{E}f(X_i)$. Then consider the function

$$Z(x_1,\cdots,x_n)=\sup_{f\in\mathfrak{F}}\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\big(f(x_i)-\mu_i(f)\big)\Big|.$$

Then

$$\begin{aligned}
Z(x_1',\cdots,x_n)&=\sup_{f\in\mathfrak{F}}\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\big(f(x_i)-\mu_i(f)\big)+f(x_1')-f(x_1)\Big|\\
&\le\sup_{f\in\mathfrak{F}}\Big(\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\big(f(x_i)-\mu_i(f)\big)\Big|+\frac{1}{\sqrt{n}}\Big|f(x_1')-f(x_1)\Big|\Big)\\
&\le\sup_{f\in\mathfrak{F}}\frac{1}{\sqrt{n}}\Big|\sum_{i=1}^{n}\big(f(x_i)-\mu_i(f)\big)\Big|+\sup_{f\in\mathfrak{F}}\frac{1}{\sqrt{n}}\Big|f(x_1')-f(x_1)\Big|\\
&\le Z(x_1,\cdots,x_n)+\frac{1}{\sqrt{n}}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
Z(x_1,\cdots,x_n)\le&\sup_{f\in\mathfrak{F}}\frac{1}{\sqrt{n}}\Big|\sum_{i=2}^{n}\big(f(x_i)-\mu_i(f)\big)+f(x_1')-\mu_1(f)\Big|\\
&+\sup_{f\in\mathfrak{F}}\frac{1}{\sqrt{n}}\Big|f(x_1)-f(x_1')\Big|\\
\le&\, Z(x_1',x_2,\cdots,x_n)+\frac{1}{\sqrt{n}}.
\end{aligned}$$

So $|Z(x_1,x_2,\cdots,x_n)-Z(x_1',x_2,\cdots,x_n)|\le\frac{1}{\sqrt{n}}$. This is also true for any index $1\le i\le n$. Then by Azuma's inequality we have for all $t\ge 0$,

$$\begin{aligned}
\mathbb{P}\Big(\big|Z(X_1,\cdots,X_n)-\mathbb{E}Z(X_1,\cdots,X_n)\big|\ge t\Big)&\le 2\exp\Big(-\frac{t^2}{2\sum_{i=1}^{n}\big(\frac{1}{\sqrt{n}}\big)^2}\Big)\\
&=2\exp\Big(-\frac{t^2}{2}\Big).
\end{aligned}$$

□

Exercise 3.5.2. Since each edges connect exactly two vertices, each edges will be counted twice when traverse all the vertices and sum all the edges connected to it. Therefore $\sum_{i=1}^{n}d_i=2m=2nd$.

(Application to **Exercise 3.5.3**: There are m teams, each composed of 3 members, and the degree $d_i = 6$ for all $i \in [n]$. Then by the same argument, $3m = \sum_{i=1}^{n} d_i = 6n \Rightarrow m = 2n$.)

$\square$

Exercise 3.5.4. By definition, $N = \sum_{i=1}^{m} \mathbb{1}_{\{\text{at least two balls in } i\text{th box}\}}$. Then

$$\begin{aligned}
\mathbb{E}N &= \sum_{i=1}^{m} \mathbb{P}(\text{at least two balls in } i\text{th box}) \\
&= m\big(1 - \mathbb{P}(\text{only } 1 \text{ or } 0 \text{ ball in the box})\big) \\
&= m\Big(1 - \frac{(m-1)^n}{m^n} - \frac{n(m-1)^{n-1}}{m^n}\Big) \\
&= m\Big[1 - \Big(1 - \frac{1}{m}\Big)^n\Big] - n\Big(1 - \frac{1}{m}\Big)^{n-1}
\end{aligned}.$$

When $m = \alpha n$ for some $\alpha > 0$,

$$\lim_{n\to\infty} \frac{\mathbb{E}N}{n} = \alpha(1 - e^{-\frac{1}{\alpha}}) - e^{-\frac{1}{\alpha}} = \alpha - (\alpha + 1)e^{-\frac{1}{\alpha}}.$$

To apply Azuma inequality, Let $N = N(X_1, \cdots, X_n)$ where X_i is the label of the box containing the ith ball. Then $X_1, \cdots, X_n$ are independent random variables. Furthermore, let $(x_1 \cdots x_i' \cdots x_n)$ and $(x_1 \cdots x_i \cdots x_n)$ to be two realizations of balls distributions, then $\big|N(x_1 \cdots x_i \cdots x_n) - N(x_1 \cdots x_i' \cdots x_n)\big| \in \{0, 1\}$, since changing the position of one ball can only increase or decrease the set {box: there are more than two ball by at most one box. Finally, by Azuma inequality we have

$$\mathbb{P}\big(|N(X_1, \cdots, X_n) - \mathbb{E}N(X_1, \cdots, X_n)| \geq t\big) \leq 2\exp\Big(-\frac{t^2}{2n}\Big).$$

By taking $t = \sqrt{2n\log(2n)}$,

$$\mathbb{P}\left(|N(X_1, \cdots, X_n) - \mathbb{E}N(X_1, \cdots, X_n)| \geq \sqrt{2n\log(2n)}\right) \leq \frac{1}{n},$$

which shows that the deviation is of a smaller order $O(\sqrt{n\log n})$ than $\mathbb{E}N$, which is $O(n)$.

$\square$

5.2 Stationary distribution

Exercise 5.2.5.

$$P(Y_n = f(s_n), Y_{n-1} = f(s_{n-1}), \cdots, Y_1 = f(s_1))$$

Exercise 5.2.6.

$$ET_1^p = \sum_{k\geq 1} P(T_1^p \geq k) \leq \int_0^\infty P(T_1^p \geq t)dt = \int_0^\infty P(T_1 \geq u)pu^{p-1}du,$$

where the last equality holds by substituting $u = t^p$. Then by the property of T_1,

$$P(T_1 \geq u) \leq \delta^{u-m},$$

we have

$$ET_1 \leq p\delta^{-m}\int_0^\infty u^{p-1}e^{-u\log(\delta^{-1})}du = \frac{p}{\delta^m}\frac{\Gamma(p)}{(\log(\delta^{-1}))^p} < \infty.$$

Exercise 5.2.8. Suppose there are m states. Let $n \in \mathbb{N}$. Then

$$\begin{aligned}
P(T > nm) &= P\left((X_1, \cdots, X_{nm}) \in S_1^{nm}\right) \\
&= P\left((X_1, \cdots, X_{(n-1)m}) \in S_1^{(n-1)m}\right) \\
&\quad \times P\left((X_{(n-1)m+1}, \cdots, X_{nm}) \in S_1^m \middle| (X_1, \cdots, X_{(n-1)m}) \in S_1^{(n-1)m}\right),
\end{aligned}$$

Write $X^{(1)} = (X_1, \cdots, X_{(n-1)m})$ and $X^{(2)} = (X_{(n-1)m+1}, \cdots, X_{nm})$. Then the conditional probability reads

$$\begin{aligned}
&P\left(X^{(2)} \in S_1^m \middle| X^{(1)} \in S_1^{(n-1)m}\right) \\
&= \sum_{\mathbf{s}\in S_1^{(n-1)m}} \frac{P\left(X^{(2)} \in S_1^m,\, X^{(1)} = \mathbf{s}\right)}{P\left(X^{(1)} \in S_1^{(n-1)m}\right)} \\
&= \sum_{\mathbf{s}\in S_1^{(n-1)m}} \frac{P\left(X^{(2)} \in S_1^m \middle| X^{(1)} = \mathbf{s}\right) P\left(X^{(1)} = \mathbf{s}\right)}{P\left(X^{(1)} \in S_1^{(n-1)m}\right)} \\
&= \sum_{\mathbf{s}\in S_1^{(n-1)m}} \frac{P\left(X^{(2)} \in S_1^m \middle| X_{(n-1)m} = s_{(n-1)m}\right) P\left(X^{(1)} = \mathbf{s}\right)}{P\left(X^{(1)} \in S_1^{(n-1)m}\right)},
\end{aligned}$$

where $s_{(n-1)m} \in S_1$ is the $(n-1)m$-th component of s. Since s_1 is inessential and $S_1 \ni s_1$ is a communication set,

$$
\begin{aligned}
&P\left(X^{(2)} \in S_1^m \middle| X_{(n-1)m} = s_{(n-1)m}\right) \\
&= 1 - P\left(X_{(n-1)m+\ell} \notin S_1 \text{ for some } \ell \le m \middle| X_{(n-1)m} = s_{(n-1)m}\right) \\
&\le 1 - P\left(X_{(n-1)m+2} \notin S_1 \middle| X_{(n-1)m+1} = s_1\right) P\left(X_{(n-1)m+1} = s_1 \middle| X_{(n-1)m} = s_{(n-1)m}\right) \\
&\le 1 - \left(\sum_{j \notin S_1} p_{s_1,j}(1)\right) \min_{i \in S_1} p_{i,s_1}(1) < 1.
\end{aligned}
$$

And we denote δ to be the RHS, $\delta := 1 - \left(\sum_{j \notin S_1} p_{s_1,j}(1)\right) \min_{i \in S_1} p_{i,s_1}(1)$. Consequently,

$$
\begin{aligned}
P(T > nm) &= P\left((X_1, \cdots, X_{nm}) \in S_1^{nm}\right) \\
&\le \delta P\left((X_1, \cdots, X_{(n-1)m}) \in S_1^{(n-1)m}\right) = \delta P(T > (n-1)m),
\end{aligned}
$$

and hence

$$P(T > nm) \le \delta^n.$$

Therefore,

$$ET = \sum_{k \ge 1} P(T \ge k) \le \sum_{k \ge 1} P\left(T \ge \lfloor k/m \rfloor m\right) \le m \sum_{\ell \ge 0} \delta^\ell = \frac{m}{1-\delta} < \infty.$$

□

Exercise 5.2.9.

(c) Suppose $t' = (t'_1, \cdots, t'_m)$ and $t = (t_1, \cdots, t_m)$ are both the solution to equations (a) and (b). Then for i such that $s_i \notin A$,

$$t_i - t'_i = \sum_{j : s_j \notin A} p_{ij}(t_j - t'_j).$$

Let I be the index such that $|t_I - t'_I| \ge |t_i - t'_i|$ for all i such that $s_i \notin A$. WLOG, we can assume $t_I - t'_I \ge 0$. Then

$$t_I - t'_I = \sum_{j : s_j \notin A} p_{1j}(t_j - t'_j) \le (t_I - t'_I) \sum_{j : s_j \notin A} p_{1j}.$$

Note that the RHS is strictly less then $t_I - t'_I$ if $t_I - t'_I \ne 0$. Thus it is only possible to have $t_I - t'_I = 0$. Therefore $|t_i - t'_i| = 0$ for all i such that $s_i \notin A$, and hence $t = t'$.

Exercise 5.3.1.

$$\frac{1}{2n}\left(P^{n+1}+\cdots+P^{2n}\right)\to\frac{1}{2}\mu$$
$$\frac{1}{2n}\left(1+P+\cdots+P^{n}\right)\to\frac{1}{2}A$$
$$\frac{1}{2}A+\frac{1}{2}\mu=\mu\Rightarrow\mu=A.$$

Exercise 5.3.2. The row vector of P^n is

$$\left(1,\frac{p}{q},\frac{p^2}{q^2},\cdots,\frac{p^{m-1}}{q^{m-1}}\right).$$

Exercise 5.3.3.

$$\lim_{n\to\infty}Ef(X_n)=\lim_{n\to\infty}\sum_{i\in S}f(i)P(X_n=i)=\sum_{i\in S}f(i)\mu(i)$$

Exercise 5.3.4. $\mu(i)=\frac{1}{m}\sum_{j=1}^{m}s_j$ for all $i\le m$.
Exercise 5.3.5. For each $i\le m$,

$$(Px)_i-(Py)_i=(P(x-y))_i=\sum_{j\le m}P_{ij}(x-y)_j\le P_{ij}\|x-y\|_\infty=\|x-y\|_\infty$$

So $\max_{i\le m}|(Px)_i-(Py)_i|\le\|x-y\|_\infty$.

Exercise 13.11.
(1) $(X_n)_{n\ge0}$ actually follows

$$X_{n+1}=\begin{cases}Y_1-1 & \text{if } X_n=0\\ x-1 & \text{if } X_n=x\ge1.\end{cases}$$

The state space is $E=\mathsf{Range}(Y-1)\subseteq\mathbb{Z}_+$. This process is irreducible since for any $x,y\in E$,

$$0<\mathbb{P}(Y-1=y)\le Q_{x+1}(x,y)\le U(x,y).$$

To show the positive recurrence, we evaluate

$$\mathbb{E}_0H_0=\sum_{k\ge0}(k+1)\mathbb{P}(Y-1=k)=\sum_{k\ge1}k\mathbb{P}(Y=k)=\mathbb{E}Y=a<\infty,$$

and for $x\in\mathsf{Range}(Y-1)\setminus\{0\}$,

$$\mathbb{E}_xH_x=x+\mathbb{E}_0H_x,$$

where

$$\begin{aligned}\mathbb{E}_0H_x&\le\mathbb{P}(Y-1<x)\left(x+\mathbb{E}_0H_x\right)+\sum_{\ell\ge x}\mathbb{P}(Y-1=\ell)(\ell-x+1)\\&\le\mathbb{P}(Y\le x)\left(x+\mathbb{E}_0H_x\right)+\sum_{\ell\ge x}\mathbb{P}(Y=\ell+1)(\ell+1),\end{aligned}$$

and therefore

$$\mathbb{E}_0H_x\le\frac{x\mathbb{P}(Y\le x)+\mathbb{E}Y}{\mathbb{P}(Y-1\ge x)}<\infty.$$

This shows $\mathbb{E}_x H_x = x + \mathbb{E}_0 H_x < \infty$. Thus X_n is positive recurrent.

Now we've known that X_n is recurrent irreducible. Then according to **Proposition 13.32**, all the states in E have the same period. In particular, by the given condition,

$$\mathsf{period}(0) = \gcd\{n : Q_n(0,0) > 0\} = \gcd\left(\mathsf{Range}(Y)\right) = 1.$$

Therefore the chain is aperiodic.

(2) Notice that

$$\mathbb{P}(n \in \mathcal{Z}) = \mathbb{P}(Z_{k(n)} = n) = \mathbb{P}(X_n = 0).$$

Also, **Corollary 13.25** guarantees that

$$\mathbb{E}_0 H_0 = \frac{1}{\mu(0)}.$$

Therefore by the previous calculation,

$$\lim_{n\to\infty} \mathbb{P}(n \in \mathcal{Z}) = \lim_{n\to\infty} \mathbb{P}(X_n = 0) = \mu(0) = \frac{1}{\mathbb{E}_0 H_0} = \frac{1}{a}.$$

□

Exercise 13.12.

(1) Recall **Theorem 13.17**: *Consider a random walk $(Y_n)_{n\in\mathbb{Z}_+}$ on $\mathbb{Z}$ whose jump distribution μ is such that $\sum_{k\in\mathbb{Z}} |k|\mu(k) < \infty$ and $\sum_{k\in\mathbb{Z}} k\mu(k) = 0$. Then all states are recurrent. Moreover the chain is irreducible if and only if the subgroup generated by $\{x \in \mathbb{Z} : \mu(x) > 0\}$ is $\mathbb{Z}$.*

Therefore by the theorem we know that $(S_n)_{n\in\mathbb{Z}_+}$ is recurrent. Also note that $\mu(0) < 1$ and $\mu(k) = 0$ for $k \le -2$ and $\sum_{k\in\mathbb{Z}} k\mu(k) = 0$ implies $\mu(-1) > 0$. Hence $\langle -1 \rangle$ is contained in the subgroup generated by the set $\{x : \mu(x) > 0\}$, and therefore this subgroup is $\mathbb{Z}$. Hence $(S_n)_{n\in\mathbb{Z}_+}$ is irreducible.

(2) The measure $\mu(\cdot) = 1$ is an invariant measure, since for any $k \in \mathbb{Z}$,

$$1 = \sum_{\ell\in\mathbb{Z}} 1 \cdot Q(\ell, k) = \sum_{\ell\in\mathbb{Z}} \mu(k - \ell) = \sum_{\ell'\in\mathbb{Z}} \mu(\ell') = 1.$$

Also, since (S_n) is recurrent irreducible by (1), the invariant measure is unique up to a multiplicative constant. Therefore

$$\nu(k) = \mathbb{E} \sum_{n=0}^{H-1} \mathbf{1}_{\{S_n = k\}} = C.$$

Since $\nu(0) = 1$, $C = 1$ and thus $\nu(k) = 1$ for all $k \in \mathbb{Z}$.

Suppose $k \le 0$, then

$$\sum_{n=R}^{H-1} \mathbf{1}_{\{S_n = k\}} = \mathbf{1}_{\{S_R > 0\}} \sum_{n=R}^{H-1} \mathbf{1}_{\{S_n = k\}} = 0,$$

since $S_n > 0$ when $R \leq n \leq H-1$ in the set $\{S_R > 0\}$. Therefore

$$\mathbb{E} \sum_{n=0}^{R-1} \mathbf{1}_{\{S_n = k\}} = 1.$$

(3)

Exercise 13.13. According to $Q(\cdot, \cdot)$, we have

$$\mu_1 + \mu_4 = \alpha_1$$
$$\mu_2 + \mu_5 = \alpha_2$$
$$\mu_3 + \mu_6 = \alpha_3$$

Thus $\mu_1 = \alpha_1(\mu_1 + \mu_2 + \mu_5) = \alpha_1(\mu_1 + \alpha_2)$, and therefore

$$\mu_1 = \frac{\alpha_1 \alpha_2}{1 - \alpha_1} = \lim_{n \to \infty} P(X_n = \text{state } 1).$$

$\square$

Exercise 13.14. Let j be a state, and $i \in N(j)$. Then $Q(i,j) = \frac{1}{\deg(i)}$ and $Q(j,i) = \frac{1}{\deg(j)}$. Thus a measure on the state space μ is reversible if

$$\mu(i)Q(i,j) = \frac{\mu(i)}{\deg(i)} = \frac{\mu(j)}{\deg(j)} = \mu(j)Q(j,i), \ \forall i \in N(j).$$

This suggests that the reversible measure μ satisfies $\frac{\mu(i)}{\deg(i)} = A$ for some constant A for all states i. Taking $A = \frac{1}{\sum_i \deg(i)}$, then

$$\sum_i \mu(i) = A \sum_i \deg(i) = 1.$$

Then $\mu(i) = \frac{\deg(i)}{\sum_j \deg(j)}$. Since μ is also an invariant measure,

$$\mathbb{E}_0 H_0 = \frac{1}{\mu(0)} = \frac{\sum_i \deg(i)}{\deg(0)} = \frac{336}{2} = 168.$$

$\square$

Countable Markov Chain

1. (Exercise 2.2) The chain is positive recurrent if and only if

$$\mathbb{E}_0 H_0 = \sum_{k\geq 1}(k+1)p_k = 1 + \sum_{k\geq 1} kp_k < \infty.$$

In that case, $\pi(0) = \frac{1}{\mathbb{E}_0 H_0}$ and

$$\begin{aligned}\pi(x) = \sum_{k\geq 0}\pi(k)p(k,x) &= \pi(x+1)p(x+1,x) + \pi(0)p(0,x) \\ &= \begin{cases} \pi(x+1) + p_x\pi(0), & \text{if } x \geq 1 \\ \pi(1), & \text{if } x = 0 \end{cases}.\end{aligned}$$

Thus

$$\begin{aligned}\pi(1) &= \pi(0) = \pi(0)\sum_{k=1}^{\infty} p_k, \\ \pi(x+1) &= \pi(0)\left(1 - \sum_{k=1}^{x} p_x\right) = \pi(0)\sum_{k=x+1}^{\infty} p_k \text{ for } x \geq 1.\end{aligned}$$

□

2. (Exercise 2.3) The expected time for returning 0 is

$$\mathbb{E}_0 H_0 = \frac{1}{3}\sum_{k\geq 1} k\left(\frac{2}{3}\right)^{k-1} = \frac{1}{3}\frac{d}{dp}\sum_{k\geq 0} p^k\Big|_{p=\frac{2}{3}} = \frac{1}{3}\cdot\frac{1}{(1-\frac{2}{3})^2} = 3.$$

Therefore $\pi(0) = \frac{1}{3}$. Moreover, for $x \in \mathbb{N}$,

$$\pi(x) = \sum_{k\geq 0}\pi(k)p(k,x) = \pi(x-1)p(x-1,x),$$

thus

$$\pi(x) = \frac{2}{3}\pi(x-1)\ \forall x \in \mathbb{N}.$$

That is, $\pi(x) = \pi(0)\left(\frac{2}{3}\right)^x = \frac{1}{3}\left(\frac{2}{3}\right)^x$ for all $x \in \mathbb{Z}_+$. □

3. (Exercise 2.4) Since the chain is irreducible, it is sufficient to check one of the states. For $x \in \mathbb{Z}_+$, define $h(x) = \mathbb{P}_x(T_0 < \infty)$. Then clearly $h(0) = 1$, and for any $x \geq 1$,

$$\begin{aligned}h(x) &= p\mathbb{P}_{x+2}(T_0 < \infty) + (1-p)\mathbb{P}_{x-1}(T_0 < \infty) \\ &= ph(x+2) + (1-p)h(x-1).\end{aligned}$$

Next we solve this recursion equation. By considering $h(x) = t^x$, then t satisfies

$$t = pt^3 + (1-p).$$

The equation $pt^3 - t + 1 - p = (t-1)(pt^2 + pt - 1 + p)$ has roots

$$t_1 = 1,\ t_2 = \frac{-p + \sqrt{-3p^2 + 4p}}{2p},\ t_3 = \frac{-p - \sqrt{-3p^2 + 4p}}{2p}.$$

Therefore

$$h(x) = c_1 + c_2 t_2^x + c_3 t_3^x$$

for some constants c_1, c_2 and c_3. When $\frac{1}{3} < p \leq 1$,

$$\begin{aligned} -2 < t_3 &= -\frac{1}{2}\left(1 + \sqrt{\frac{4}{p} - 3}\right) \leq -1, \\ 0 \leq t_2 &= -\frac{1}{2}\left(1 - \sqrt{\frac{4}{p} - 3}\right) < 1. \end{aligned}$$

Thus $h(x) = t_2^x$ is a solution, since $h(0) = 1$, $0 \leq h(x) \leq 1$ on $\mathbb{Z}_+$ and $\inf_x h(x) = 0$. Therefore it is transient when $\frac{1}{3} < p \leq 1$.

When $p \leq \frac{1}{3}$, $t_2 \geq 1$ and $t_3 \leq -2$. To keep $h(x) \geq 0$, we must have $c_3 = 0$. To keep $h(x) \leq 1$, we must have $c_2 = 0$. Thus the only solution is $h(x) = 1$ on $x \in \mathbb{Z}_+$. That is, $\mathbb{P}_x(T_0 < \infty) = 1$ for all $x \in \mathbb{Z}_+$ when $p \leq \frac{1}{3}$. Therefore it is recurrent when $p \leq \frac{1}{3}$. $\square$

4. (Exercise 2.5)

(a) For any $k \in \mathbb{N}$, define $\alpha(-k) = P(Y \leq -k)$. Then

$$\alpha(-k) = p\alpha(-k-1) + (1-p)\alpha(-k+1).$$

Further define $\alpha(0) = P(Y \leq 0) = 1$, then

$$\begin{aligned} \alpha(-k-1) - \alpha(-k) &= \frac{1-p}{p}[\alpha(-k) - \alpha(-k+1)] \\ &= \left(\frac{1-p}{p}\right)^k [\alpha(-1) - \alpha(0)]. \end{aligned}$$

Therefore

$$\begin{aligned} &\sum_{k \geq 0} \alpha(-k-1) - \alpha(-k) = \alpha(-k-1) - \alpha(0) = [\alpha(-1) - \alpha(0)]\frac{1 - \beta^{k+1}}{1 - \beta} \\ &\Rightarrow \alpha(-k-1) = 1 - \frac{1 - \beta^{k+1}}{1 - \beta}[1 - \alpha(-1)] \\ &\Rightarrow 0 = 1 - \frac{1 - \alpha(-1)}{1 - \beta} \qquad (P(Y \leq -k) \to 0 \text{ as } k \to \infty) \end{aligned}$$

Then we have $P(Y \leq -k) = \alpha(-k) = \beta^k = \left(\frac{1-p}{p}\right)^k$ for all $k \in \mathbb{N} \cup \{0\}$.

(b) We verify this by induction. Clearly, $e(1) = 1 \cdot e(1)$. Suppose $e(k) = ke(1)$. Then

$$\begin{aligned} e(k+1) = ET_{k+1} &= \sum_{\ell \geq 1} \ell \sum_{s=1}^{\ell} P(T_1 = s) P(T_k = \ell - s) \\ &= \sum_{s \geq 1} \sum_{\ell \geq s} P(T_1 = s) \cdot \ell P(T_k = \ell - s) \\ &= \sum_{s \geq 1} P(T_1 = s)[ET_k + s] \\ &= ET_k + ET_1 = (k+1) ET_1. \end{aligned}$$

(c) Since

$$e(2) = ET_2 = p[1 + ET_1] + (1-p)[1 + ET_3],$$

by (b) we have

$$2e(1) = p[1 + e(1)] + (1-p)[1 + 3e(1)].$$

Therefore $e(1) = \frac{1}{2p-1}$.
(d) From the result in (c) we immediately have $e(1) = \infty$ when $p = \frac{1}{2}$.

$\square$

5. (Exercise 2.7)

(a)

$$\mathbb{P}_0(H_0 = \infty) = \prod_{x=0}^{\infty} \frac{x+1}{x+2} = \frac{1}{2} \cdot \frac{2}{3} \cdots = \lim_{n \to \infty} \frac{1}{n} = 0.$$

Therefore $\mathbb{P}_0(H_0 < \infty) = 1$ and it is recurrent. However,

$$\mathbb{E}_0 H_0 = \sum_{k \geq 0} (k+1) \cdot \prod_{\ell=0}^{k-1} \frac{\ell+1}{\ell+2} \cdot \frac{1}{k+2} = \sum_{k \geq 0} \frac{1}{k+2} = \infty.$$

Therefore this chain is null recurrence.
(b) One can directly compute

$$\begin{aligned} \mathbb{E}_0 H_0 &= \sum_{k \geq 0} (k+1) \cdot \prod_{\ell=0}^{k-1} \frac{1}{\ell+2} \cdot \frac{k+1}{k+2} \\ &= \sum_{k \geq 0} \frac{1}{k!} \cdot \frac{k+1}{k+2} \leq \sum_{k \geq 0} \frac{1}{k!} = e < \infty. \end{aligned}$$

Therefore this chain is positive recurrent.
(c)

$$\mathbb{P}_0(H_0 = \infty) = \prod_{k=0}^{\infty} \left(1 - \frac{1}{k^2+2}\right) \geq \frac{1}{3} \prod_{k=2}^{\infty} \left(1 - \frac{1}{k^2}\right) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6} > 0.$$

Therefore this chain is transient. $\square$

8. (Exercise 2.10)

(a) Let H_0 be the first time that the chain be in the state 0. Then $H_0 < \infty$ if and only if the population dies out. That is, $\mathbb{P}_1(H_0 < \infty) = a = P\{\text{population dies out} \mid X_0 = 1\}$. Also, since

$$\begin{aligned}\mathbb{P}_0(H_0 < \infty) &= P(H_0 < \infty \mid X_0 = 0)\\ &= \sum_{k\geq 0} P(H_0 < \infty \mid X_1 = k)P(X_1 = k \mid X_0 = 0)\\ &= P(H_0 < \infty \mid X_1 = 1)\\ &= P(H_0 - 1 < \infty \mid X_0 = 1) = \mathbb{P}_1(H_0 < \infty),\end{aligned}$$

if $\mu \leq 1$ then $a = 1$, and the chain is recurrent; if $\mu > 1$ then $a < 1$, and the chain is transient.
($\mu < 1$) One can directly compute

$$\begin{aligned}EX_n &= \sum_{k\geq 0} kP(X_n = k) = \sum_{k\geq 0} k \sum_{\ell\geq 1} P(X_{n-1} = \ell)P(S_\ell = k)\\ &= \sum_{\ell\geq 1}\sum_{k\geq 0} kP(S_\ell = k)P(X_{n-1} = \ell) = \sum_{\ell\geq 1} \ell\mu P(X_{n-1} = \ell) = \mu EX_{n-1},\end{aligned}$$

and get that $EX_n = \mu^n EX_0 = \mu^n$. Since $H_0 > n$ if and only if $X_n \geq 1$. Therefore by the Markov inequality,

$$P(H_0 > n) = P(X_n \geq 1) \leq EX_n = \mu^n.$$

Hence

$$EH_0 = \sum_{n\geq 0} P(H_0 > n) \leq \sum_{n\geq 0} \mu^n,$$

which is finite when $\mu < 1$. Therefore the chain is positive recurrent when $\mu < 1$.
($\mu = 1$) Since $P(H_0 > n) = 1 - P(H_0 \leq n) = 1 - \phi^{(n)}(0)$, If $\{p_n\}$ such that $\mu = 1$ and

$$EH_0 = \sum_{n\geq 0} 1 - \phi^{(n)}(0) < \infty,$$

then it forms a positive recurrent chain. Otherwise, it is null recurrent. (TBD)

9. (Exercise 2.14)
(a) Since $\mu > 1$, $p_2 > 0$. Therefore $\phi''(s) = \sum_{k\geq 2} k(k-1)s^{k-2}p_k > 0$ on $[0, 1]$. Consider $g(s) = s - \phi(s)$, then $g''(s) < 0$, and $g(a) = 0 = g(1)$.

Suppose $g'(a) \leq 0$. Then since $g''(s) < 0$ on $[0, 1]$, $g'(s) < 0$ for $s \in (a, 1]$. Thus g is monotonically decreasing on $(a, 1]$. But we have $g(1) = g(a)$, which leads to a contradiction. Therefore $g'(a) = 1 - \phi'(a) > 0$. □

(b) Since $\phi'(a) < 1$ and $a_n \uparrow a$, there exists an $\varepsilon > 0$ sufficiently small, and thus a sufficiently large $N \geq 1$, such that $a - a_n < \varepsilon$ for all $n > N$, and $\phi'(x) < \rho < 1$ whenever $|x - a| < \varepsilon$. Then

$$a - a_{n+1} = \phi(a) - \phi(a_n) \leq \rho(a - a_n).$$

□

(c) Since $\{X_n = 0\} \subset \{H_0 < \infty\}$, $P(X_n > 0, H_0 < \infty) = P(H_0 < \infty) - P(X_n = 0)$. Therefore

$$\begin{aligned}\frac{P(\text{ extinction } \mid X_{n+1} > 0)}{P(\text{ extinction } \mid X_n > 0)} &= \frac{P(H_0 < \infty) - P(X_{n+1} = 0)}{P(H_0 < \infty) - P(X_n = 0)} \cdot \frac{P(X_n > 0)}{P(X_{n+1} > 0)} \\ &= \frac{a - a_{n+1}}{a - a_n} \cdot \frac{1 - a_n}{1 - a_{n+1}}.\end{aligned}$$

Since $a_n \uparrow a$, and by (b), there exists $N \geq 1$ such that for all $n > N$,

$$\rho\left(1 + \frac{a_{n+1} - a_n}{1 - a_{n+1}}\right) \leq \rho\left(1 + \frac{\varepsilon}{1 - a}\right),$$

where $\varepsilon > 0$ is chosen so that the RHS is strictly less than 1. Denote the bound in RHS as $\delta < 1$. Then we have for all $n > N$,

$$\frac{P(\text{ extinction } \mid X_{n+1} > 0)}{P(\text{ extinction } \mid X_n > 0)} < \delta,$$

thus we can find $C > 0$ such that

$$P(\text{ extinction } \mid X_n > 0) \leq C\delta^n = Ce^{-n\log(1/\delta)}$$

for all $n \in \mathbb{N}$. $\square$