Solutions to Durrett's Probability: Theory and Examples

Lumiao

1 Martingales

1.1 Martingales, Almost Sure Convergence

Problem 1.1.1 *X_n is a martingale w.r.t. $\mathcal{G}_n$ and let $\mathcal{F}_n = \sigma(X_1, \cdots, X_n)$. Then $\mathcal{G}_n \supseteq \mathcal{F}_n$ and X_n is a martingale w.r.t. $\mathcal{F}_n$.*

Proof Note that $\sigma(X_m) \subseteq \mathcal{G}_m \subseteq \mathcal{G}_n$ if $m \leq n$. Thus:

$$\mathcal{F}_n = \sigma(X_1, \cdots, X_n) = \sigma(\bigcup_{k=1}^{n} \sigma(X_k)) \subseteq \mathcal{G}_n$$

Obviously $X_n \in \mathcal{F}_n$ and is integrable. Also, using the property of conditional expectation:

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \mathbb{E}[\mathbb{E}[X_{n+1}|\mathcal{G}_n]|\mathcal{F}_n] = \mathbb{E}[X_n|\mathcal{F}_n] = X_n, \ \forall n \geq 1$$

Thus $(X_n)_{n\geq 1}$ is a martingale w.r.t $\mathcal{F}_n$. □

Problem 1.1.2 *Give an example of a submartingale X_n so that X_n^2 is a supermartingale. Hint: X_n does not have to be random.*

Proof Let $X_n = -\frac{1}{n}$. □

Problem 1.1.3 *Generalize (i) of **Theorem** 4.2.7 by showing that if X_n and Y_n are submartingales w.r.t. $\mathcal{F}_n$ then $X_n \vee Y_n$ is also.*

Proof Obviously $X_n \vee Y_n \in \mathcal{F}_n$ and is integrable. By the definition of submartingale:

$$\mathbb{E}[X_{n+1} \vee Y_{n+1}|\mathcal{F}_n] \geq \mathbb{E}[X_{n+1}|\mathcal{F}_n] \geq X_n$$

$$\mathbb{E}[X_{n+1} \vee Y_{n+1}|\mathcal{F}_n] \geq \mathbb{E}[Y_{n+1}|\mathcal{F}_n] \geq Y_n$$

Thus $\mathbb{E}[X_{n+1} \vee Y_{n+1}|\mathcal{F}_n] \geq X_n \vee Y_n$, which proves the result. □

Problem 1.1.4 *Let X_n, $n \geq 0$, be a submartingale with $\sup X_n < \infty$. Let $\xi_n = X_n - X_{n-1}$ and suppose $\mathbb{E}[\sup \xi_n^+] < +\infty$. Show that X_n converges a.s.*

Proof Consider the stopping time $N_M = \inf\{n \geq 0 : X_n \geq M\}$ for $M \in \mathbb{N}$. $X_{N_M \wedge n}$ is a submartingale. Notice that $X^+_{N_M \wedge n} \leq M + \sup \xi_n^+$, we have:

$$\mathbb{E}[X^+_{N_M \wedge n}] \leq M + \mathbb{E}[\sup \xi_n^+] < +\infty$$

By **Theorem** 4.2.11, $X_{N_M \wedge n}$ converges to a limit a.s. Thus for each M, X_n converges a.s. to a limit on $\{N_M = \infty\}$. Since $\sup X_n < +\infty$, $\Omega = \bigcup_{M=1}^{\infty} \{N_M = \infty\}$, which means that X_n converges almost surely on the whole Ω. □

Problem 1.1.5 *Give an example of a martingale X_n with $X_n \to -\infty$ a.s. Hint: Let $X_n = \xi_1 + \cdots + \xi_n$, where the ξ_i are independent (but not identically distributed) with $\mathbb{E}[\xi_n] = 0$.*

Proof Following the hint, let $X_n = \xi_1 + \cdots + \xi_n$, where the ξ_i are independent. Each ξ_k satisfies $\mathbb{P}(\xi_k = -1) = \frac{1}{2k}$ and $\mathbb{P}(\xi_k = \frac{1}{2k-1}) = 1 - \frac{1}{2k}$. Then $\mathbb{E}[\xi_n] = 0$. Since ξ_k are independent and $\sum_{k=1}^{+\infty} \mathbb{P}(\xi_k = -1) = \sum_{k=1}^{+\infty} \frac{1}{2k} = +\infty$, we have $\mathbb{P}(\xi_n = -1,\ i.o.) = 1$, which means that $X_n \to -\infty$ a.s. $\square$

Problem 1.1.6 *Let $Y_1, Y_2, \cdots$ be nonnegative i.i.d. random variables with $\mathbb{E}[Y_m] = 1$ and $\mathbb{P}(Y_m = 1) < 1$. By **Example** 4.2.3 $X_n = \prod_{m \le n} Y_m$ defines a martingale.*
*(i) Use **Theorem** 4.2.12 and an argument by contradiction to show $X_n \to 0$ a.s.*
(ii) Use the strong law of large numbers to conclude $\frac{1}{n} log X_n \to c < 0$.

Proof (i) The convergence is obviously since $X_n \ge 0$. To prove the limit is 0, we only need to prove that $X_n \xrightarrow{\mathbb{P}} 0$. We have the following observations:

1. Since we have proved X_n converges almost everywhere, X_n also converges in probability, i.e. $\mathbb{P}(|X_n - X_m| > \epsilon) \to 0$ as $m, n \to 0$;

2. Since $\mathbb{P}(Y_m = 1) \le 1$, there exists $\epsilon_1 > 0$ and $\delta > 0$ such that $\mathbb{P}(|Y_m - 1| > \epsilon_1) > \delta$;

3. For any $\epsilon_2 > 0$, take $\epsilon = \epsilon_1 \cdot \epsilon_2$ and $m = n+1$ in the first observation. Since $|X_{n+1} - X_n| = |X_n| \cdot |Y_{n+1} - 1|$, we have:

$$\mathbb{P}(|X_n - X_m| > \epsilon_1 \cdot \epsilon_2) \ge \mathbb{P}(|X_n| > \epsilon_2,\ |Y_{n+1} - 1| > \epsilon_1) = \mathbb{P}(|X_n| > \epsilon_2) \cdot \mathbb{P}(|Y_{n+1} - 1| > \epsilon_1) \ge \delta \mathbb{P}(|X_n| > \epsilon_2)$$

This observations gives that X_n converges to 0 in probability
(ii) Notice that $\frac{1}{n} log X_n = \frac{1}{n} \sum_{k=1}^{n} log Y_m$. We need to discuss the existence of the expectation of $log(Y_m)$.
Case 1 If $\mathbb{E}[|log Y_m|] < +\infty$, then we can use the Jensen's Inequality:

$$0 = -log \mathbb{E}[Y_m] < \mathbb{E}[-log Y_m]$$

The equality could not hold since $\mathbb{P}(Y_m = 1) < 1$. Now by the strong law of large number, we have:

$$\frac{1}{n} \sum_{k=1}^{n} -log Y_m \xrightarrow{a.s.} \mathbb{E}[-log Y_m] > 0 \quad \Rightarrow \quad \frac{1}{n} log X_n \xrightarrow{a.s.} c < 0$$

Case 2 If $\mathbb{E}[|log Y_m|] = +\infty$. We need to check out only one of the expectation of $(-log Y_m)^+$ and $(-log Y_m)^-$ is infinity. In fact:

$$\mathbb{E}[(-log Y_m)^-] = \mathbb{E}[log Y_m \mathbf{1}_{Y_m \ge 1}] \le \mathbb{E}[(Y_m - 1)\mathbf{1}_{\{Y_m \ge 1\}}] \le 1 - \mathbb{P}(Y_m \ge 1) < +\infty$$

Thus only $\mathbb{E}[(-log Y_m)^+] = +\infty$. By **Theorem** 2.4.5, we have:

$$\frac{1}{n} \sum_{k=1}^{n} -log Y_m \xrightarrow{a.s.} +\infty \quad \Rightarrow \quad \frac{1}{n} log X_n \xrightarrow{a.s.} -\infty < 0$$

$\square$

Problem 1.1.7 *Let X_n and Y_n be positive integrable and adapted to $\mathcal{F}_n$. Suppose $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \le (1 + Y_n) X_n$ with $\sum_{n=1}^{+\infty} Y_n < +\infty$ a.s. Prove that X_n converges a.s. to a nite limit by nding a closely related supermartingale to which **Theorem** 4.2.12 can be applied.*

Proof Following the hint, define:

$$Z_n = \frac{X_n}{\prod_{m=1}^{n-1}(1 + Y_m)}, \quad n \ge 1$$

Obviously $Z_n \in \mathcal{F}_n$ and is integrable. Also:

$$\mathbb{E}[Z_{n+1}|\mathcal{F}_n] = \left(\prod_{m=1}^{n}(1 + Y_m)\right)^{-1} \cdot \mathbb{E}[X_{n+1}|\mathcal{F}_n] \le \left(\prod_{m=1}^{n-1}(1 + Y_m)\right)^{-1} \cdot X_n = Z_n,\ \forall n \ge 1$$

Thus Z_n is a supermartingale which is positive. By **Theorem** 4.2.12, Z_n converges to a limit which is finite almost everywhere. Denote it by Z_∞. Also, by $\sum_{n=1}^{+\infty} Y_n < +\infty$ a.s., we know $\prod_{m=1}^{\infty}(1+Y_m)$ converges almost everywhere. As a result:

$$X_n = Z_n \cdot \prod_{m=1}^{n-1}(1+Y_m) \xrightarrow{a.s.} Z_\infty \prod_{m=1}^{\infty}(1+Y_m)$$

where the limit is finite almost everywhere. $\square$

Problem 1.1.8 ***The switching principle*** *Suppose X_n^1 and X_n^2 are supermartingales with respect to $\mathcal{F}_n$, and N is a stopping time so that $X_N^1 \geq X_N^2$. Then:*

$$Y_n = X_n^1 \mathbf{1}_{\{N>n\}} + X_n^2 \mathbf{1}_{\{N\leq n\}} \text{ is a supermartingale}$$

$$Z_n = X_n^1 \mathbf{1}_{\{N\geq n\}} + X_n^2 \mathbf{1}_{\{N<n\}} \text{ is a supermartingale.}$$

Proof Obviously Y_n and Z_n are in $\mathcal{F}_n$ and is integrable. We have two opinions to make inequalities, one is the property of supermartingale and the other is the condition $X_N^1 \geq X_N^2$. Combine them properly. For Y_n, we have:

$$\begin{aligned}
\mathbb{E}[Y_{n+1}|\mathcal{F}_n] &= \mathbb{E}[X_{n+1}^1 \mathbf{1}_{\{N>n+1\}} + X_{n+1}^2 \mathbf{1}_{\{N\leq n+1\}}|\mathcal{F}_n] \\
&\leq \mathbb{E}[X_{n+1}^1 \mathbf{1}_{\{N>n\}} + X_{n+1}^2 \mathbf{1}_{\{N\leq n\}}|\mathcal{F}_n] \\
(\{N>n\},\ \{N\leq n\} \in \mathcal{F}_n) \quad &\leq \mathbf{1}_{\{N>n\}}\mathbb{E}[X_{n+1}^1|\mathcal{F}_n] + \mathbf{1}_{\{N\leq n\}}\mathbb{E}[X_{n+1}^2|\mathcal{F}_n] \\
&\leq \mathbf{1}_{\{N>n\}}X_n^1 + \mathbf{1}_{\{N\leq n\}}X_n^2 \\
&= Y_n
\end{aligned}$$

Similarly, for Z_n, we have:

$$\begin{aligned}
\mathbb{E}[Z_{n+1}|\mathcal{F}_n] &= \mathbb{E}[X_{n+1}^1 \mathbf{1}_{\{N\geq n+1\}} + X_{n+1}^2 \mathbf{1}_{\{N<n+1\}}|\mathcal{F}_n] \\
(\{N\geq n+1\},\ \{N<n+1\} \in \mathcal{F}_n) \quad &\leq \mathbf{1}_{\{N\geq n+1\}}\mathbb{E}[X_{n+1}^1|\mathcal{F}_n] + \mathbf{1}_{\{N<n+1\}}\mathbb{E}[X_{n+1}^2|\mathcal{F}_n] \\
&\leq \mathbf{1}_{\{N\geq n+1\}}X_n^1 + \mathbf{1}_{\{N<n+1\}}X_n^2 \\
&\leq \mathbf{1}_{\{N\geq n\}}X_n^1 + \mathbf{1}_{\{N<n\}}X_n^2 \\
&= Z_n
\end{aligned}$$

These inequalities show that both Y_n and Z_n are supermartingales! $\square$

1.2 Examples

Problem 1.2.1 *Let X_n and Y_n be positive integrable and adapted to $\mathcal{F}_n$. Suppose $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \leq X_n + Y_n$, with $\sum_{k=1}^{\infty} Y_n < \infty$ a.s. Prove that X_n converges a.s. to a finite limit. Hint: Let $N = \inf_k\{\sum_{m=1}^{k} Y_m > M\}$, and stop your supermartingale at time N.*

Proof The way to prove is similar to that of **Problem** 1.1.4, but here it is a little subtle in the index of summation. Let $Z_n = X_n - \sum_{k=1}^{n-1}$, $n \geq 1$. Obviously $Z_n \in \mathcal{F}_n$ and Z_n is integrable. Also, by the assumption, we have:

$$\mathbb{E}[Z_{n+1}|\mathcal{F}_n] = \mathbb{E}[X_{n+1}|\mathcal{F}_n] - \sum_{k=1}^{n} Y_k \leq X_n - \sum_{k=1}^{n-1} Y_k = Z_n, \quad \forall n \geq 1$$

Thus Z_n is a supermartingale. To bound the supermartingale, let $N_M = \inf_k\{\sum_{m=1}^{k} Y_m > M\}$ and consider the supermartingale $Z_{n\wedge N_M}$. We will find that:

$$Z_{n\wedge N_M} = X_{n\wedge N_M} - \sum_{k=1}^{n\wedge N_M - 1} Y_k \geq -M$$

Here the index of summation is very important! Thus we can apply **Theorem** 4.2.12 to $Z_{n\wedge N_M}+M$, which yields that Z_n converges to a finite limit a.s. on $\{N_M=\infty\}$. Notice that $\bigcup_{M=1}^{+\infty}\{N_M=\infty\}=\Omega$ a.s. since $\sum_{k=1}^{\infty}Y_n<\infty$ a.s. So Z_n converges to a finite limit a.s. on Ω. Finally, $X_n=Z_n+\sum_{k=1}^{n-1}Y_m$ also converges to a finite limit a.s. $\square$

Problem 1.2.2 *Show* $\sum_{n=2}^{\infty}\mathbb{P}(A_n|\cap_{m=1}^{n-1}A_m^c)=\infty$ *implies* $\mathbb{P}(\cap_{m=1}^{\infty}A_m^c)=0$

Proof 1 You may use **Exercise** 4.3.4 if you notice that:

$$\prod_{m=1}^{n}\left(1-\mathbb{P}(A_n|\cap_{k=1}^{m-1}A^c)\right)=\mathbb{P}(\cap_{m=1}^{n}A_m^c)$$

Proof 2 You can also use **Borel-Cantelli II**. Let $\mathcal{F}_n=\sigma(\{A_1,\cdots,A_n\})$. Notice that:

$$\mathbb{P}(A_{n+1}|\cap_{m=1}^{n}A_m^c)=\mathbb{P}(A_{n+1}|\mathcal{F}_n)\ \ on\ \bigcap_{m=1}^{n}A_m^c$$

Since for each $n\geq 1$, $\bigcap_{m=1}^{\infty}A_m^c\subseteq\bigcap_{m=1}^{n}A_m^c$, we have:

$$\mathbb{P}(A_{n+1}|\cap_{m=1}^{n}A_m^c)=\mathbb{P}(A_{n+1}|\mathcal{F}_n)\ \ on\ \bigcap_{m=1}^{\infty}A_m^c\ \ \forall n\geq 1$$

This and the assumption of the problem give:

$$\sum_{n=2}^{+\infty}\mathbb{P}(A_n|\mathcal{F}_{n-1})=+\infty\ \ on\ \bigcap_{m=1}^{\infty}A_m^c$$

By using **Borel-Cantelli II**, we have:

$$\bigcap_{m=1}^{\infty}A_m^c\subseteq\{A_n\ i.o.\}\ \ \mathbb{P}-a.s.$$

This happens only when $\mathbb{P}\cap_{m=1}^{\infty}A_m^c)=0$ $\square$

Problem 1.2.3 *Show that if* $\mathbb{P}(\lim\frac{Z_n}{\mu^n}=0)<1$*, then it is* $=\rho$ *and hence:*

$$\{\lim\frac{Z_n}{\mu^n}>0\}=\{Z_n>0\ for\ all\ n\}\ a.s.$$

Proof Generally, $Z_n>0$ does not indicate that the limit of Z_n/μ^n is not zero. This is why this problem meaningful! Notice that if $\mu\leq 1$, $\mathbb{P}(Z_n=0,\ \exists n\geq 1)=1$ and then $\mathbb{P}(\lim\frac{Z_n}{\mu^n}=0)=1$ which is contradictory to the assumption. So $\mu>1$ and $\mathbb{P}(Z_n>0,\ \forall n\geq 1)=1-\rho$. Since we have the relationship:

$$\{\lim\frac{Z_n}{\mu^n}>0\}\subseteq\{Z_n>0\ for\ all\ n\}$$

it is sufficient to prove $\mathbb{P}(\lim\frac{Z_n}{\mu^n}>0)=1-\rho$. Given $Z_1=k$, denote the number of child corresponding to each subtree by Z_n^i, $1\leq i\leq k$. Then $Z_n^1,\cdots,Z_n^k$ are independent and have the same distribution as Z_n. So:

$$\begin{aligned}\mathbb{P}(\lim\frac{Z_n}{\mu^n}=0)&=\sum_{k=1}^{+\infty}\mathbb{P}(\lim\frac{Z_n}{\mu^n}=0|Z_1=k)\mathbb{P}(Z_1=k)\\&=\sum_{k=0}^{+\infty}\mathbb{P}(\lim\sum_{i=1}^{k}\frac{Z_{n-1}^i}{\mu^n}=0)\cdot p_k\\&=\sum_{k=0}^{+\infty}\mathbb{P}(\bigcap_{i=1}^{k}\{\lim\frac{Z_{n-1}^i}{\mu^n}=0\})\cdot p_k\\ i.i.d\quad&=\sum_{k=0}^{+\infty}\mathbb{P}(\lim\frac{Z_n}{\mu^n}=0)p_k\end{aligned}$$

This means that $\mathbb{P}(\lim \frac{Z_n}{\mu^n} = 0)$ is the solution of $x = \varphi(x)$ in $[0, 1)$, which is just ρ. Thus $\mathbb{P}(\lim \frac{Z_n}{\mu^n} > 0) = 1 - \rho$ and the problem is solved. $\square$

Problem 1.2.4 *Let Z_n be a branching process with offspring distribution p_k, defined in part **(d)** of Section 4.3, and let $\varphi(\theta) = \sum_{k=1}^{\infty} p_k \theta^k$. Suppose $\rho < 1$ has $\varphi(\rho) = \rho$. Show that ρ^{Z_n} is a martingale and use this to conclude $\mathbb{P}(Z_n = 0 \ for \ some \ n \geq 1 | Z_0 = x) = \rho^x$.*

Proof Obviously $\rho^{Z_n} \in \mathcal{F}_n$. $|\rho^{Z_n}| < 1$ shows its integrability. Consider on the set $\{Z_n = k\}$, we have:

$$\begin{aligned}\mathbb{E}[\rho^{Z_{n+1}}|\mathcal{F}_n] =& \mathbb{E}[\rho^{\xi_1^{n+1}+\cdots+\xi_k^{n+1}}|\mathcal{F}_n] \\ =& \mathbb{E}[\rho^{\xi_1^{n+1}+\cdots+\xi_k^{n+1}}] \\ =& \left(\mathbb{E}[\rho^{\xi_1^{n+1}}]\right)^k \\ =& \varphi(\rho)^k = \rho^{Z_n}\end{aligned}$$

Here we have used the definition of $\varphi(\theta)$ and ρ. Thus $\mathbb{E}[\rho^{Z_{n+1}}|\mathcal{F}_n] = \rho^{Z_n}$ and ρ^{Z_n} is a martingale. For the second part of the problem, we will defer it until the proof of **Problem** 1.4.2. $\square$

1.3 Doob's Inequality, Convergence in $\mathbb{L}^p$, $p > 1$

Problem 1.3.1 *Show that if $j \leq k$ then $\mathbb{E}[X_j; N = j] \leq \mathbb{E}[X_k; N = j]$ and sum over j to get a second proof of $\mathbb{E}X_N \leq \mathbb{E}X_k$.*

Proof Since $\{N = j\} \in \mathcal{F}_j$, by the definition of submartingale we have:

$$\mathbb{E}[X_k; N = j] = \mathbb{E}[\mathbb{E}[X_k|\mathcal{F}_j]; N = j] \geq \mathbb{E}[X_j; N = j]$$

On $\{N = j\}$, $X_j = X_N$, so sum over j we get $\mathbb{E}X_N \leq \mathbb{E}X_k$. $\square$

Problem 1.3.2 *Generalize the proof of **Theorem** 4.4.1 to show that if X_n is a submartingale and $M \leq N$ are stopping times with $\mathbb{P}(N \leq k) = 1$, then $\mathbb{E}X_M \leq \mathbb{E}X_N$.*

Proof Imitate the proof in the textbook, let $K_n = \mathbf{1}_{\{M<n\leq N\}}$. Since $\{M < n \leq N\} = \{M \leq n-1\} \cap \{N \leq n-1\}^c \in \mathcal{F}_{n-1}$, K_n is predictable w.r.t $\mathcal{F}_n$. Consider the stochastic integration:

$$(K \cdot X)_n = X_{N\wedge n} - X_{M\wedge n}, \quad n \geq 0$$

It is still a submartingale and therefore $\mathbb{E}[(K \cdot X)_0] \leq \mathbb{E}[(K \cdot X)_k]$. This is just $\mathbb{E}[X_M] \leq \mathbb{E}[X_N]$ $\square$

Problem 1.3.3 *Suppose $M \leq N$ are stopping times. If $A \in \mathcal{F}_M$ then $L = M \cdot \mathbf{1}_A + N \cdot \mathbf{1}_{A^c}$ is a stopping time.*

Proof It is sufficient to show that $\{L \leq n\} \in \mathcal{F}_n$ for all $n \geq 0$. Notice that we have the decomposition:

$$\{L \leq n\} = (\{L \leq n\} \cap A) \cup (\{L \leq n\} \cap A^c) = (\{M \leq n\} \cap A) \cup (\{N \leq n\} \cap \{M \leq n\} \cap A^c) \in \mathcal{F}_n$$

Here we have used the fact that both $\{M \leq n\} \cap A$ and $\{M \leq n\} \cap A^c$ are in $\mathcal{F}_n$. $\square$

Problem 1.3.4 *Use the stopping times from the previous exercise to strengthen the conclusion of **Problem** 1.3.2 to $X_M \leq \mathbb{E}[X_N|\mathcal{F}_M]$.*

Proof By the definition of conditional distribution, we have to show that $\mathbb{E}[X_M\mathbf{1}_A] \leq \mathbb{E}[X_N\mathbf{1}_A]$ for ay $A \in \mathcal{F}_M$. Now for each $A \in \mathcal{F}_M$, define the stopping time L as in the previous problem. Obviously $L \leq N \leq k$ and we can use the result of **Problem** 1.3.2:

$$\mathbb{E}[X_M\mathbf{1}_A] + \mathbb{E}[X_N\mathbf{1}_{A^c}] = \mathbb{E}[X_L] \leq \mathbb{E}[X_N] = \mathbb{E}[X_N\mathbf{1}_A] + \mathbb{E}[X_N\mathbf{1}_{A^c}]$$

This indicates $\mathbb{E}[X_M\mathbf{1}_A] \leq \mathbb{E}[X_N\mathbf{1}_A]$ $\square$

Problem 1.3.5 *Prove the following variant of the conditional variance formula. If $\mathcal{F} \subseteq \mathcal{G}$, then:*

$$\mathbb{E}\left(\mathbb{E}[Y|\mathcal{G}] - \mathbb{E}[Y|\mathcal{F}]\right)^2 = \mathbb{E}\left(\mathbb{E}[Y|\mathcal{G}]\right)^2 - \mathbb{E}\left(\mathbb{E}[Y|\mathcal{F}]\right)^2$$

Proof The formula proves useful if you take $\mathcal{G}$ and $\mathcal{F}$ from a filtration $\mathcal{F}_n$ which gives an estimation of $\|Y_m - Y_n\|_{\mathbb{L}^2}$. Here Y_n denote the martingale $\mathbb{E}[Y|\mathcal{F}_n]$. The formula itself is easy to prove once you notice that:

$$\mathbb{E}[\mathbb{E}[Y|\mathcal{G}]\mathbb{E}[Y|\mathcal{F}]] = \mathbb{E}[\mathbb{E}[\mathbb{E}[Y|\mathcal{G}]\mathbb{E}[Y|\mathcal{F}]|\mathcal{F}]] = \mathbb{E}[\mathbb{E}[Y|\mathcal{F}]\mathbb{E}[\mathbb{E}[Y|\mathcal{G}]|\mathcal{F}]] = \mathbb{E}\left(\mathbb{E}[Y|\mathcal{F}]\right)^2$$

Take this back to the left hand side and the result follows. □

Problem 1.3.6 *If we let $S_n = \xi_1 + \cdots + \xi_n$ where the ξ_m are independent and have $\mathbb{E}\xi_m = 0$, $|\xi_m| \le K$, $\sigma_m^2 = \mathbb{E}\xi_m^2 < \infty$, then $S_n^2 - s_n^2 = S_n^2 - \sum_{m=1}^n \sigma_m^2$ is a martingale. Use this and **Theorem** 4.4.1 to conclude:*

$$\mathbb{P}\left(\max_{1\le m\le n} |S_m| \le x\right) \le \frac{(x+K)^2}{Var S_n}$$

Proof The way of proving is similar to that used in the section of **optional stopping time**. To get an accurate estimation of probability, you may set a stopping time to get another martingale (sub\supermartingale). Use the equality (inequality) of expectation and write it out at the stoping time. You might need the property of uniform integrability or the theorems of convergence to remove the changing index n. Here we do not. Let $N = \inf\{n \ge 0 : |S_n| > x\}$, A be the event that $\{\max_{1\le m\le} |S_m| \le x\}$. $S_n^2 - s_n^2$ is a martingale and so is $S_{n\wedge N}^2 - s_{n\wedge N}^2$. Use this we have:

$$0 = \mathbb{E}[S_{0\wedge N}^2 - s_{0\wedge N}^2] = \mathbb{E}[S_{n\wedge N}^2 - s_{n\wedge N}^2] = \mathbb{E}[S_{n\wedge N}^2] - \mathbb{E}[s_{n\wedge N}^2]$$

To split thing out, we have several observations. A^c implies $\{N \le n\}$. On A, we have $|S_{n\wedge N}| = |S_n| \le x$ and on A^c we have $|S_{n\wedge N}| = |S_N| \le K + x$. Combine these together, we have:

$$\mathbb{E}[S_{n\wedge N}^2] = \mathbb{E}[S_{n\wedge N}^2 \mathbf{1}_A] + \mathbb{E}[S_{n\wedge N}^2 \mathbf{1}_{A^c}] \le x^2\mathbb{P}(A) + (K+x)^2\left(1 - \mathbb{P}(A)\right)$$

For the other term, we have:

$$\mathbb{E}[s_{n\wedge N}^2] \ge \mathbb{E}[s_{n\wedge N}^2 \mathbf{1}_A] = Var[S_n^2]\mathbb{P}(A)$$

These two gives:

$$(K+x)^2 \ge \left(K^2 + 2Kx + Var[S_n^2]\right)\mathbb{P}(A) \ge Var[S_n^2]\mathbb{P}(A)$$

This proves the result! □

Problem 1.3.7 *The next result gives an extension of **Theorem** 4.4.2 to $p = 1$. Let X_n be a martingale with $X_0 = 0$ and $\mathbb{E}X_n^2 < \infty$. Show that:*

$$\mathbb{P}\left(\max_{1\le m\le n} X_m > \lambda\right) \le \frac{\mathbb{E}X_n^2}{\mathbb{E}X_n^2 + \lambda^2}$$

Proof Following the hint, for any $c > -\lambda$, consider the submartingale $(X_m + c)^2$. Using **Doob's Inequality**, we get:

$$\mathbb{P}\left(\max_{1\le m\le n} X_m > \lambda\right) = \mathbb{P}\left(\max_{1\le m\le n} X_m + c > \lambda + c\right) \le \mathbb{P}\left(\max_{1\le m\le n} (X_m + c)^2 > (\lambda + c)^2\right) \le \frac{E[(X_n + c)^2]}{(\lambda + c)^2}$$

Since $c > -\lambda$ is arbitrary, calculate the minimum of the right hand side and the result will follow! □

Problem 1.3.8 *Let X_n and Y_n be martingales with $\mathbb{E}X_n^2 < \infty$ and $\mathbb{E}Y_n^2 < \infty$. Then:*

$$\mathbb{E}X_nY_n - \mathbb{E}X_0Y_0 = \sum_{m=1}^n \mathbb{E}(X_m - X_{m-1})(Y_m - Y_{m-1})$$

Proof Note that for each $m \geq 1$:

$$\begin{aligned}\mathbb{E}(X_m - X_{m-1})(Y_m - Y_{m-1}) =&\mathbb{E}[X_m Y_m] + \mathbb{E}[X_{m-1}Y_{m-1}] - \mathbb{E}[X_m Y_{m-1}] - \mathbb{E}[X_{m-1}Y_m] \\ =&\mathbb{E}[X_m Y_m] + \mathbb{E}[X_{m-1}Y_{m-1}] - \mathbb{E}[\mathbb{E}[X_m Y_{m-1}|\mathcal{F}_{m-1}]] - \mathbb{E}[\mathbb{E}[X_{m-1}Y_m|\mathcal{F}_{m-1}]] \\ =&\mathbb{E}[X_m Y_m] + \mathbb{E}[X_{m-1}Y_{m-1}] - \mathbb{E}[Y_{m-1}\mathbb{E}[X_m|\mathcal{F}_{m-1}]] - \mathbb{E}[X_{m-1}\mathbb{E}[Y_m|\mathcal{F}_{m-1}]] \\ =&\mathbb{E}[X_m Y_m] - \mathbb{E}[X_{m-1}Y_{m-1}]\end{aligned}$$

Sum them up and the result follows. $\square$

Problem 1.3.9 *Let X_n, $n \geq 0$, be a martingale and $\xi_n = X_n - X_{n-1}$ for $n \geq 1$. If $\mathbb{E}X_0^2 < \infty$, $\sum_{m=1}^{\infty} \mathbb{E}\xi_m^2 < \infty$ then $X_n \to X_\infty$ a.s. and in $\mathbb{L}^2$.*

Proof Use the result from the previous problem, setting $X_n = Y_n$ and we will get:

$$\mathbb{E}[X_n^2] = \mathbb{E}[X_0^2] + \sum_{m=1}^{n} \mathbb{E}[\xi_m^2]$$

This gives:

$$\sup_{n\geq 0} \mathbb{E}[X_n^2] < +\infty$$

By the $\mathbb{L}^p$ **Convergence Theorem**, we have $X_n \to X_\infty$ a.s. and in $\mathbb{L}^2$. $\square$

1.4 Uniform Integrability, Convergence in $\mathbb{L}^1$

Problem 1.4.1 *Let X_n be r.v.'s taking values in $[0, \infty)$. Let $D = \{X_n = 0\ for\ some\ n \geq 1\}$ and assume:*

$$\mathbb{P}(D|X_1, \cdots, X_n) \geq \delta(x) > 0 \quad a.s.\ on\ \{X_n \leq x\}$$

*Use **Theorem** 4.6.9 to conclude that $\mathbb{P}(D \cup \{\lim_n X_n = \infty\}) = 1$.*

Proof We only need to prove that $\{\lim_n X_n = \infty\})^c \subseteq D$. Note that:

$$\{\lim_n X_n = \infty\}^c = \{\liminf_n X_n < \infty\} = \bigcup_{M=1}^{+\infty} \{\liminf_n X_n < M\}$$

It is true that:

$$\bigcup_{M=1}^{+\infty} \{\liminf_n X_n < M\} = \bigcup_{N=1}^{+\infty} \{X_n \leq N\ i.o.\} \subseteq \bigcup_{N=1}^{+\infty} \{\mathbb{P}(D|X_1, \cdots, X_n) \geq \delta(N)\ i.o.\}$$

By **Theorem** 4.6.9, we know that $\mathbb{P}(D|X_1, \cdots, X_n)$ converges to $\mathbf{1}_D$ almost surely. However, $\mathbf{1}_D$ is either 0 or 1, which indicates that on each $\{\mathbb{P}(D|X_1, \cdots, X_n) \geq \delta(N)\ i.o.\}$, $\mathbf{1}_D$ must be one. We can then conclude that: $\{\lim_n X_n = \infty\})^c \subseteq D$ a.s. This gives the desired result!. $\square$

Problem 1.4.2 *Let Z_n be a branching process with offspring distribution p_k (see the end of Section 5.3 for denitions). Use the last result to show that if $p_0 > 0$ then $\mathbb{P}(\lim_n Z_n = 0\ or\ \infty) = 1$.*

Proof It is a direct corollary of the previous result. Note that:

$$\mathbb{P}(D|Z_1, \cdots, Z_n) \geq \mathbb{P}(Z_{n+1} = 0|Z_1, \cdots, Z_n) \geq p_0^k \quad on\ \{Z_n \leq k\}$$

We can then use the previous result. Here $D = \{\lim_n Z_n = 0\}$ $\square$

Problem 1.2.4 Revisited Now we can finish the proof of **Problem** 1.2.4! We have proved that ρ^{Z_n} is actually a martingale. Since Z_n converges to either 0 or ∞, the limit of ρ^{Z_n} is 0 or 1. So using the **Bounded convergence Theorem**:

$$\rho^x = \mathbb{E}_x[\rho^{Z_n}] \to 1 \cdot \mathbb{P}_x(\lim_n Z_n = 0) + 0 \cdot \mathbb{P}_x(\lim_n Z_n = \infty)$$

(Here the subscript x denotes the initial state). Thus $\mathbb{P}(Z_n = 0\ for\ some\ n \geq 1|Z_0 = x) = \rho^x$. $\square$

Problem 1.4.3 *Show that if $\mathcal{F}_n \nearrow \mathcal{F}_\infty$ and $Y_n \to Y$ in $\mathbb{L}^1$ then $\mathbb{E}[Y_n|\mathcal{F}_n] \to \mathbb{E}[Y|F_\infty]$ in $\mathbb{L}^1$.*

Proof By triangular inequality:

$$\mathbb{E}[|\mathbb{E}[Y_n|\mathcal{F}_n] - \mathbb{E}[Y|\mathcal{F}_\infty]|] \leq \mathbb{E}[|\mathbb{E}[Y_n|\mathcal{F}_n] - \mathbb{E}[Y|\mathcal{F}_n]|] + \mathbb{E}[|\mathbb{E}[Y|\mathcal{F}_n] - \mathbb{E}[Y|\mathcal{F}_\infty]|]$$

For the first term:

$$\mathbb{E}[|\mathbb{E}[Y_n|\mathcal{F}_n] - \mathbb{E}[Y|\mathcal{F}_n]|] \leq \mathbb{E}[\mathbb{E}[|Y_n - Y||\mathcal{F}_n]] = \mathbb{E}[|Y_n - Y|] \to 0$$

For the second term, by the result of **Theorem 4.6.7**, $\mathbb{E}[Y|\mathcal{F}_n] \to \mathbb{E}[Y|\mathcal{F}_\infty]$ in $\mathbb{L}^1$, which indicates that the second term tends to 0. Thus we get $\mathbb{E}[Y_n|\mathcal{F}_n] \to \mathbb{E}[Y|F_\infty]$ in $\mathbb{L}^1$. □

1.5 Backwards Martingales

1.6 Optional Stopping Theorems

Problem 1.6.1 *If $L \leq M$ are stopping times and $Y_{M\wedge n}$ is a uniformly integrable submartingale, then $\mathbb{E}Y_L \leq \mathbb{E}Y_M$ and:*

$$Y_L \leq E[Y_M|\mathcal{F}_L]$$

Proof Before proving, we first need to justify the notation. On $\{M = \infty\}$, $Y_{M\wedge n} = Y_n$ and it converges to some limit, which we will denote by Y_∞. After noticing this, there will be no problem in the case when the stopping time index is infinity. Besides, there's always no problem in the definition of $\mathcal{F}_\infty$ since it equals $\bigvee_{n=1}^{+\infty} \mathcal{F}_n$.
Now use **Theorem 4.8.3** and let $X_n = Y_{M\wedge n}$, $N = L$. We immediately get $\mathbb{E}Y_L \leq \mathbb{E}Y_M$. To prove the last result, considering for any $A \in \mathcal{F}_L$, use the stopping time in **Problem 1.3.3** and let $N = L \cdot \mathbf{1}_A + M \cdot \mathbf{1}_{A^c}$. The first result now shows $\mathbb{E}Y_N \leq \mathbb{E}Y_M$. Since $N = M$ on A^c, we have:

$$\mathbb{E}[Y_L \mathbf{1}_A] \leq \mathbb{E}[Y_M \mathbf{1}_A]$$

Since A is arbitrary, use the definition of conditional expectation and we will get $Y_L \leq E[Y_M|\mathcal{F}_L]$. □

Problem 1.6.2 *If $X_n \geq 0$ is a supermartingale then $\mathbb{P}(\sup X_n > \lambda) \leq \mathbb{E}X_0/\lambda$.*

Proof Let $N = \inf\{n : X_n > \lambda\}$, $A = \{\sup_{n\geq 1} X_n > \lambda\}$. Using **Theorem 4.8.4** we get:

$$\mathbb{E}[X_0] \geq \mathbb{E}[X_N] \leq \mathbb{E}[X_N \mathbf{1}_A] \geq \lambda \mathbb{P}(A)$$

This gives $\mathbb{P}(A) \leq \mathbb{E}X_0/\lambda$. □

Problem 1.6.3 *Let $S_n = \xi_1 + \cdots + \xi_n$ where the ξ_i are independent with $\mathbb{E}\xi_i = 0$ and $Var(\xi_i) = \sigma^2$. $S_n^2 - n\sigma^2$ is a martingale. Let $T = \min\{n : |S_n| > a\}$. Use Theorem 4.8.2 to show that $\mathbb{E}T \geq a^2/\sigma^2$.*

Proof We might assume that $\mathbb{E}T$ is finite since otherwise the result is trivial. Using**Theorem 4.4.1**, we have:

$$0 = \mathbb{E}[S_0^2] - 0 \cdot \sigma^2 = \mathbb{E}[S_{n\wedge T}^2] - \mathbb{E}[n \wedge T]\sigma^2$$

This gives $\mathbb{E}[n \wedge T]\sigma^2 = \mathbb{E}[S_{n\wedge T}^2]$. Letting $n \to \infty$ and use **Fatou's Lemma** and **Monotone Convergence Theorem**, we get:

$$\sigma^2 \mathbb{E}[T] = \sigma^2 \lim_{n\to\infty} \mathbb{E}[n \wedge T] = \liminf_{n\to\infty} \mathbb{E}[S_{n\wedge T}^2] \geq \mathbb{E}[\lim_{n\to\infty} S_{n\wedge T}^2] = \mathbb{E}[S_T^2] > a^2$$

(Here S_T^2 is well defined since we have assumed that $\mathbb{E}T$ is finite.) Thus $\mathbb{E}T \geq a^2/\sigma^2$ □

Problem 1.6.4 ***(Wald's second equation)*** *Let $S_n = \xi_1 + \cdots \xi_n$, where the ξ_i are independent with $\mathbb{E}\xi_i = 0$ and $Var(\xi_i) = \sigma^2$. Use the martingale from the previous problem to show that if T is a stopping time with $\mathbb{E}T < \infty$ then $ES_T^2 = \sigma^2 \mathbb{E}T$.*

Proof By the same method in the previous problem, we have $\mathbb{E}[n \wedge T]\sigma^2 = \mathbb{E}[S^2_{n\wedge T}]$. This gives for any $n \geq 1$:

$$\mathbb{E}[S^2_{n\wedge T}] = \mathbb{E}[n \wedge T] \leq \mathbb{E}[T] < \infty$$

Thus $\sup_{n\geq 1} \mathbb{E}[S^2_{n\wedge T}] < \infty$. $\mathbb{L}^p$ **Convergence Theorem** gives that the martingale $S_{n\wedge T}$ converges to S_T in $\mathbb{L}^2$. As a result, $\lim_{n\to\infty} \mathbb{E}[S^2_{n\wedge T}] = \mathbb{E}[S^2_T]$. The result follows by letting $n \to \infty$ in the first equation.
Remark We can also prove directly that $\{S_{n\wedge T}\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{L}^2(\Omega)$. By the property of martingale, if $n > m$:

$$\mathbb{E}[(S_{n\wedge T} - S_{m\wedge T})^2] = \mathbb{E}S^2_{n\wedge T} - \mathbb{E}S^2_{m\wedge T} = \sigma^2 \left(\mathbb{E}[n \wedge T] - \mathbb{E}[m \wedge T]\right)$$

It is easy to check that the righthand side is cauchy sequence in $\mathbb{R}$ since we have known that $\mathbb{E}[n \wedge T]$ converges to $\mathbb{E}[T]$ which is finite. This proves our claim! □

Problem 1.6.5 ***Variance of the time of gambler's ruin*** *Let $\xi_1, \xi_2, \cdots$ be independent with $\mathbb{P}(\xi_i = 1) = p$ and $\mathbb{P}(\xi_i = -1) = q = 1 - p$ where $p < 1/2$. Let $S_n = S_0 + \xi_1 + \cdots + \xi_n$ and let $V_0 = \min\{n \geq 0 : S_n = 0\}$.* ***Theorem 4.8.9*** *tells us that $\mathbb{E}_x V_0 = x/(1-2p)$. The aim of this problem is to compute the variance of V_0. If we let $Y_i = \xi_i - (p - q)$ and note that $\mathbb{E}Y_i = 0$ and:*

$$\mathbf{Var}(Y_i) = \mathbf{Var}(\xi_i) = \mathbb{E}\xi_i^2 - (\mathbb{E}\xi_i)^2$$

then it follows that $(S_n - (p-q)n)^2 - n(1-(p-q)^2)$ is a martingale.
(a) *Use this to conclude that when $S_0 = x$ the variance of V_0 is:*

$$x \cdot \frac{1-(p-q)^2}{(q-p)^3}$$

(b) *Why must the answer in* ***(a)*** *be in the form cx?*

Proof Use the same method as in the previous problem. The martingale in the problem gives:

$$\mathbb{E}[(S_{n\wedge V_0} - (p-q)(n \wedge V_0))^2 - (n \wedge V_0)\sigma^2] = x^2$$

where $\sigma^2 = 1 - (p-q)^2$. This gives that:

$$\mathbb{E}[(S_{n\wedge V_0} - (p-q)(n \wedge V_0))^2] = x^2 + \sigma^2 \mathbb{E}[n \wedge V_0] \leq x^2 + \sigma^2 \mathbb{E}[V_0] < +\infty$$

Thus we get $\sup_{n\geq 1} \mathbb{E}[(S_{n\wedge V_0} - (p-q)(n \wedge V_0))^2] < +\infty$. $\mathbb{L}^p$ **Convergence Theorem** gives that the martingale $S_{n\wedge V_0} - (p-q)(n \wedge V_0)$ converges to $S_{V_0} - (p-q)V_0$ a.s. and in $\mathbb{L}^2$. As a result:

$$\lim_{n\to\infty} \mathbb{E}[(S_{n\wedge V_0} - (p-q)(n \wedge V_0))^2] = \mathbb{E}[(S_{V_0} - (p-q)V_0)^2] = (p-q)^2 E[V_0^2]$$

So letting $n \to \infty$ in the first equation we can get that $\mathbb{E}[V_0^2] = \left((p-q)x^2 - x + (p-q)^2 x\right)/(p-q)^3$. Notice that $\mathbb{E}[V_0] = x/(q-p)$, we come to:

$$\mathbf{Var}(V_0) = x \cdot \frac{1-(p-q)^2}{(q-p)^3}$$

Remark You can also exchange the limit and expectation by using **Dominated Convergence Theorem**. Consider:

$$(S_{n\wedge V_0} - (p-q)(n \wedge V_0))^2 = S^2_{n\wedge V_0} - 2(p-q)(n \wedge V_0)S_{n\wedge V_0} + (p-q)^2(n \wedge V_0)^2$$

For the first term, $0 \leq S^2_{n\wedge V_0} \leq (\sup_{n\geq 1} S_n)^2$. Using **Theorem 4.8.9** we can know that $\mathbb{P}(\sup_{n\geq 1} S_n \geq m) = (p/q)^m$, which implies that $\mathbb{E}[(\sup_{n\geq 1} S_n)^2] < +\infty$. For the last term, apply **Fatou's Lemma** to the very first equation in the first proof and we can get:

$$\mathbb{E}[(0 - (p-q)V_0)^2] \leq x^2 + \lim_{n\to\infty} \sigma^2 \mathbb{E}[n \wedge V_0] = x^2 + \sigma^2 \mathbb{E}[V_0] < +\infty$$

This shows that the third term $(n \wedge V_0)^2$ could be dominated by V_0^2 which is integrable. For the second term, just use **Cauchy Schwarz's Inequality** and we can find a dominator. Combining these observations, we can exchange the expectation and the limit. □

Problem 1.6.6 *Let $S_n = \xi_1 + \cdots + \xi_n$ be a random walk. Suppose $\varphi(\theta_o) = \mathbb{E}\exp(\theta_o \xi_1) = 1$ for some $\theta_o < 0$ and ξ_i is not constant. In this special case of the exponential martingale $X_n = \exp(\theta_o S_n)$ is a martingale. Let $\tau = \inf\{n : S_n \notin (a, b)\}$ and $Y_n = X_{n\wedge\tau}$. Use **Theorem 4.8.2** to conclude that $EX_\tau = 1$ and $\mathbb{P}(S_\tau \leq a) \leq \exp(-\theta_o a)$.*

Proof Following the hint, we need to check the condition in **Theorem4.8.2**. Obviously $X_n \mathbf{1}_{\{n<\tau\}}$ is bounded and thus uniformly integrable. Also, by using **Fatou's Lemma**, we get that:

$$\mathbb{E}[|X_\tau|] = \mathbb{E}[\lim_{n\to\infty} |X_{n\wedge\tau}|] \leq \liminf_{n\to\infty} \mathbb{E}[|X_{n\wedge\tau}|] < +\infty$$

Thus by the theorem, we conclude that $X_{n\wedge\tau}$ is uniformly integrable. By letting $n \to \infty$ in the equation:

$$1 = \mathbb{E}[X_{0\wedge\tau}] = \mathbb{E}[X_{n\wedge\tau}]$$

We can get $\mathbb{E}[X_\tau] = 1$. Finally, using **Chebyshev's Inequality** we can get $\mathbb{P}(S_\tau \leq a) \leq \exp(-\theta_o a)$. □

Problem 1.6.7 *Continuing with the set-up of the previous problem, suppose that the ξ_i are integer valued with $\mathbb{P}(\xi_i < -1) = 0$, $\mathbb{P}(\xi_1 = -1) > 0$ and $\mathbb{E}\xi_i > 0$. Let $T = \inf\{n : S_n = a\}$ with $a < 0$. Use the martingale $X_n = \exp(\theta_o S_n)$ to conclude that $\mathbb{P}(T < \infty) = \exp(-\theta_o a)$.*

2 Markov Chain

2.1 Examples

Problem 2.1.1 *Let $\xi_1, \xi_2, \cdots$ be i.i.d. $\in \{1, 2, \cdots, N\}$ and taking each value with probability $1/N$. Show that $X_n = |\{\xi_1, \cdots, \xi_n\}|$ is a Markov chain and compute its transition probability.*

Proof Obviously it is a Markov Chain, since the value of X_{n+1} depends only on X_n and ξ_{n+1}. It's transition probability could be represented as:

$$p(k, k+1) = 1 - \frac{k}{N}, \quad p(k, k) = \frac{k}{N}, \quad p(i, j) = 0 \text{ otherwise}$$

□

Problem 2.1.2 *Let $\xi_1, \xi_2, \cdots$ be i.i.d. $\in \{-1, 1\}$, taking each value with probability $1/2$. Let $S_0 = 0$, $S_n = \xi_1 + \cdots \xi_n$ and $X_n = \max\{S_m : 0 \leq m \leq n\}$. Show that X_n is not a Markov chain.*

Proof Since X_n might equal to S_n and might not, we find that:

$$\mathbb{P}(X_{n+1} = m + 1 | X_n = m) \neq \mathbb{P}(X_{n+1} = m + 1 | X_n = m, X_{n-1} = m)$$

Thus X_n is not a Markov chain. □

Problem 2.1.3 *Let $\xi_0, \xi_1, \cdots$ be i.i.d. $\in \{H, T\}$, taking each value with probability $1/2$. Show that $X_n = (\xi_n, \xi_{n+1})$ is a Markov chain and compute its transition probability p. What is p^2?*

Proof Since $X_n = (\xi_n, \xi_{n+1})$, we know that $X_{n-1} = (\xi_{n-2}, \xi_{n-1}), \cdots, X_0 = (\xi_0, \xi_1)$ are independent with X_n and thus X_n is a Markov chain. It's transition probability is:

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}_{4 \times 4}$$

For p^2, since it represents the two step transition probability and by our previous observation, X_{n+1} is independent with X_{n-1}, we immediately know that $p^2(i, j) = 1/4$ for any state i, j. □

2.2 Construction, Markov Properties

Problem 2.2.1 *Let $A \in \sigma(X_0, \cdots, X_n)$ and $B \in \sigma(X_n, X_{n+1}, \cdots)$. Use the Markov property to show that for any initial distribution :*

$$\mathbb{P}_\mu(A \cap B | X_n) = \mathbb{P}_\mu(A | X_n)\mathbb{P}_\mu(B | X_n)$$

Proof The left-hand side could be represented as:

$$\begin{aligned} \mathbb{E}_\mu[\mathbf{1}_A \mathbf{1}_B | X_n] &= \mathbb{E}_\mu[\mathbb{E}_\mu[\mathbf{1}_A \mathbf{1}_B | \mathcal{F}_n] | X_n] \\ &= \mathbb{E}_\mu[\mathbf{1}_A \mathbb{E}_\mu[\mathbf{1}_B | \mathcal{F}_n] | X_n] \\ (B \in \sigma(X_n, X_{n+1}, \cdots)) \quad &= \mathbb{E}_\mu[\mathbf{1}_A \mathbb{E}_\mu[\mathbf{1}_B | X_n] | X_n] \\ &= \mathbb{E}_\mu[\mathbf{1}_A | X_n] E_\mu[\mathbf{1}_B | X_n] \end{aligned}$$

This proves the equation. □

Problem 2.2.2 *Let X_n be a Markov chain and suppose*

$$\mathbb{P}(\cup_{m=n+1}^{\infty} \{X_m \in B_m\} | X_n) \geq \delta > 0 \quad on \quad \{X_n \in A_n\}$$

Then $\mathbb{P}(\{X_n \in A_n \ i.o.\} - \{X_n \in B_n \ i.o.\}) = 0$.

Proof Let $E_n = \cup_{m=n+1}^{\infty}\{X_m \in B_m\}$. Then $E \doteq \cap_{n=1}^{\infty} E_n = \{X_m \in B_m \ i.o.\}$. Also let $\mathcal{F}_n = \sigma(X_0, \cdots, X_n)$ and $\mathcal{F}_\infty$. Obviously $\mathbf{1}_{E_n} \in \sigma(X_n, X_{n+1}, \cdots)$ and $\mathbf{1}_{E_n}$ convergence almost surely to $\mathbf{1}_E \in \mathcal{F}_\infty$. So by the markov property and **Dominated Convergence Theorem for Conditional Expectation**, we get:

$$\mathbb{P}(E_n|X_n) = \mathbb{P}(E_n|\mathcal{F}_n) \to \mathbb{P}(E|\mathcal{F}_\infty) = \mathbf{1}_E, \quad n \to \infty$$

Since on $\{X_n \in A_n, \ i.o.\}$, $\mathbb{P}(E_n|X_n) \geq \delta > 0$ infinitely often, $\mathbf{1}_E$ must equal to one on $\{X_n \in A_n, \ i.o.\}$. This means that $\{X_n \in A_n \ i.o.\} \subseteq \{X_n \in B_n \ i.o.\}$. ☐

Problem 2.2.3 *A state a is called absorbing if $\mathbb{P}_a(X_1 = a) = 1$. Let $D = \{X_n = a \ for \ some \ n \geq 1\}$ and let $h(x) = \mathbb{P}_x(D)$. Use the result of the previous exercise to conclude that $h(X_n) \to 0$ a.s. on D^c. Here a.s. means $\mathbb{P}_\mu$ a.s. for any initial distribution μ.*

Proof For any $\epsilon > 0$, let $A_\epsilon = \{x : \mathbb{P}_x(D) \geq \epsilon\}$. By calculation:

$$\mathbb{P}_{X_n}(D) = \mathbb{E}_{X_n}[\mathbf{1}_D] = \mathbb{E}_\mu[\mathbf{1}_D \circ \theta_n|\mathcal{F}_n] = \mathbb{P}_\mu(\bigcup_{m=n+1}^{+\infty}\{X_m = a\}|X_n)$$

Thus on $\{X_n \in A_\epsilon\}$, we have $\mathbb{P}_\mu(\bigcup_{m=n+1}^{+\infty}\{X_m = a\}|X_n) \geq \epsilon$. So by using the previous problem, we get that:

$$\{X_n \in A_\epsilon \ i.o.\} \subseteq \{X_n = a \ i.o,\} \subseteq D$$

As a result, on D^C, for any $\epsilon > 0$, $X_n \in A_\epsilon$ for only finite many times. This means that on D^c, $h(X_n) \to 0$. ☐

Problem 2.2.4 ***First Entrance Decomposition*** *Let $T_y = \inf\{n \geq 1 : X_n = y\}$. Show that:*

$$p^n(x,y) = \sum_{m=1}^{n} \mathbb{P}_x(T_y = m)p^{n-m}(y,y)$$

Proof Notice that $\{X_n = y\} \subseteq \{T_y \leq n\}$. Thus:

$$\begin{aligned}
\mathbb{P}_x(X_n = y) =& \mathbb{E}_x[\mathbf{1}_{\{X_n=y\}}\mathbf{1}_{\{T_y \leq n\}}] \\
=& \mathbb{E}_x[\mathbf{1}_{\{T_y \leq n\}}\mathbb{E}_x[\mathbf{1}_{\{X_{n-T_y}=y\}} \circ \theta_{T_y}|\mathcal{F}_{T_y}]] \\
=& \mathbb{E}_x[\mathbf{1}_{\{T_y \leq n\}}\mathbb{E}_{X_{T_y}}[\mathbf{1}_{\{X_{n-T_y}=y\}}]] \\
=& \mathbb{E}_x[\mathbf{1}_{\{T_y \leq n\}}p^{n-T_y}(y,y)] \\
=& \sum_{m=1}^{n} \mathbb{P}_x(T_y = m)p^{n-m}(y,y)
\end{aligned}$$

This completes the result. ☐

Problem 2.2.5 *Show that:*

$$\sum_{m=0}^{n} \mathbb{P}_x(X_m = x) = \sum_{m=k}^{n+k} \mathbb{P}_x(X_m = x)$$

Proof Consider from the right hand side. Let $T_x = \inf\{n \geq k : X_n = x\}$ be the first time $X_n = x$ after time k. By the same method as in the previous problem, we can prove that:

$$\mathbb{P}_x(X_m = x) = \sum_{l=k}^{m} \mathbb{P}_x(T_x = l)p^{m-l}(x,x), \quad \forall m \geq k$$

Put the equation into the right hand side of our goal:

$$\begin{aligned}\sum_{m=k}^{n+k} p^m(x,x) &\leq \sum_{m=k}^{n+k}\sum_{l=k}^{m} \mathbb{P}_x(T_x = l)p^{m-l}(x,x) \\ &= \sum_{l=k}^{n+k} \mathbb{P}_x(T_x = l)\sum_{m=l}^{n+k} p^{m-l}(x,x) \\ &= \sum_{l=k}^{n+k} \mathbb{P}_x(T_x = l)\sum_{m=0}^{n+k-l} p^{m}(x,x) \\ &= \sum_{m=0}^{n} p^m(x,x)\sum_{l=k}^{n+k-m} \mathbb{P}_x(T_x = l) \\ &\leq \sum_{m=0}^{n} p^m(x,x)\end{aligned}$$

This proves the inequality. □

Problem 2.2.6 *Let $T_C = \inf\{n \geq 1 : X_n \in C\}$. Suppose that $|S \setminus C| < \infty$ and for $\forall x \in S \setminus C$, we have $\mathbb{P}_x(T_c < \infty) > 0$. Prove that there is an $N < \infty$ and $\epsilon > 0$ such that $\forall y \in S \setminus C$, $\mathbb{P}_y(T_C > kN) < (1-\epsilon)^k$.*

Proof By assumption, for any $x \in S \setminus C$, there exists $n(x) \in \mathbb{N}$ and $\epsilon_x > 0$ such that $\mathbb{P}_x(T_C \leq n(x)) > \epsilon_x$. Let $N = \max_{x\in S\setminus C}\{n(x)\}$ and $\epsilon = \min_{x\in S\setminus}\{\epsilon_x\}$. Then $\mathbb{P}_x(T_C \leq N) > \epsilon$ holds for all $x \in S \setminus C$. Now consider:

$$\begin{aligned}\mathbb{P}_y(T_C > kN) =&\mathbb{E}_y[\mathbf{1}_{\{T_C>kN\}}\mathbf{1}_{\{T_C>(k-1)N\}}] \\ =&\mathbb{E}_y[\mathbf{1}_{\{T_C>(k-1)N\}}\mathbb{E}_y[\mathbf{1}_{\{T_C>kN\}}|\mathbb{F}_{(k-1)N}]] \\ =&\mathbb{E}_y[\mathbf{1}_{\{T_C>(k-1)N\}}\mathbb{E}_y[\mathbf{1}_{\{T_C>N\}}\circ\theta_{(k-1)N}|\mathbb{F}_{(k-1)N}]] \\ =&\mathbb{E}_y[\mathbf{1}_{\{T_C>(k-1)N\}}\mathbb{E}_{X_{(k-1)N}}[\mathbf{1}_{\{T_C>N\}}]] \\ \leq&(1-\epsilon)\mathbb{E}_y[\mathbf{1}_{\{T_C>(k-1)N\}}] \\ \leq&\cdots \leq (1-\epsilon)^k\end{aligned}$$

This gives the result. □

Problem 2.2.7 ***Exit distributions*** *Let $V_C = \inf\{n \geq 0 : X_n \in C\}$ and let $h(x) = \mathbb{P}_x(V_A < V_B)$. Suppose $A \cap B \neq \emptyset$, $S-(A\cup B)$ is finite, and $\mathbb{P}_x(V_{A\cup B} < \infty) > 0$ for all $x \in S-(A\cup B)$.*
(i) *Show that:*

$$h(x) = \sum_y p(x,y)h(y) \quad \textit{for } x \notin A\cap B$$

(ii) *Show that if h satises the previous equation, then $h(X(n \wedge V_{A\cup B}))$ is a martingale.*
(iii) *Use this and **Problem** 2.2.6 to conclude that $h(x) = \mathbb{P}_x(V_A < V_B)$ is the only solution of the previous equation that is 1 on A and 0 on B.*

Proof (i) Define $Y(\omega) = 1$ if $V_A(\omega) < V_B(\omega)$ and $Y(\omega) = 0$ otherwise. In other words, $Y = \mathbf{1}_{\{V_A(\omega)<V_B(\omega)\}}$. By the Markov Property, we have:

$$\mathbb{E}_x[Y\circ\theta_1|\mathcal{F}_1] = \mathbb{E}_{X_1}[Y]$$

Notice that when $x \notin A\cup B$, $Y \circ \theta_1 = Y$. So by integration, we get that:

$$h(x) = \mathbb{P}_x(V_A < V_B) = \mathbb{E}_x[Y] = \mathbb{E}_x[\mathbb{E}_{X_1}[Y]] = \sum_{y\in S}\mathbb{P}_x(X_1 = y)\mathbb{P}_y(V_A < V_B) = \sum_{y\in S} p(x,y)h(y)$$

(ii) We have to check that $\mathbb{E}_\mu[h(X_{(n+1)\wedge V_{A\cup B}})|\mathcal{F}_n] = h((X_{n\wedge V_{A\cup B}})$ for all $n \geq 0$. For simplicity, we will omit the subscript μ and denote the stopping time $V_{A\cup B}$ by V. To prove this, it is sufficient to show it holds on both $\{V \leq n\}$ and $\{V > n\}$. On $\{V > n\}$, we have:

$$\mathbb{E}[h(X_{(n+1)\wedge V})|\mathcal{F}_n] = \mathbb{E}[h(X_{n+1})|\mathcal{F}_n] = \mathbb{E}_{X_n}[X_1] = h(X_n) = h(X_{n\wedge V})$$

Here we have used the fact that $X_n \notin A \cup B$ and the equation. On $\{V \leq n\}$, $h(X_{(n+1)\wedge V}) = h(X_{n\wedge V}) \in \mathcal{F}_n$. This gives:

$$\mathbb{E}[h(X_{(n+1)\wedge V})|\mathcal{F}_n] = \mathbb{E}[h(X_{n\wedge V})|\mathcal{F}_n] = h(X_{n\wedge V})$$

Thus $h(X_{n\wedge V})$ is a martingale.
(iii) Use the result of **Problem** 2.2.6, $\mathbb{P}(V_{A\cup B} < \infty)$=1. So $h(X_{n\wedge V_{A\cup B}}) \to h(X_{V_{A\cup B}})$ a.s. Also, notice that h must be a bounded function since there are only finite many variables in $S - A \cup B$ and outside it h is either 0 or 1. Thus by the **Bounded Convergence Theorem** and the property of martingale, we have:

$$h(x) = \mathbb{E}_x[h(X_{0\wedge V_{A\cup B}})] = \mathbb{E}_x[h(X_{n\wedge V_{A\cup B}})] \to \mathbb{E}_x[h(X_{V_{A\cup B}})] = 1 \cdot \mathbb{P}_x(V_A < V_B) + 0 \cdot \mathbb{P}_x(V_B < V_A)$$

This proves the result! $\square$

Problem 2.2.8 *Let X_n be a Markov chain with $S = \{0, 1, \cdots, N\}$ and suppose that X_n is a martingale. Let $V_x = \min\{n \geq 0 : X_n = x\}$ and suppose $\mathbb{P}_x(V_0 \wedge V_N < \infty) > 0$ for all x. Show $\mathbb{P}_x(V_N < V_0) = x/N$.*

Proof By **Problem 2.2.6** we know that $\mathbb{P}(V_0 \wedge V_N < \infty) = 1$. Let $T = V_0 \wedge V_N$, then $X_{n\wedge T} \xrightarrow{a.s.} X_T$. Since $X_{n\wedge T}$ is also a martingale and is bounded, by using **Dominated Convergence Theorem** we know that:

$$x = \mathbb{E}_x[X_{0\wedge T}] = \mathbb{E}_x[X_{n\wedge T}] \to \mathbb{E}[X_T] = N\mathbb{P}(V_0 > V_N) + 0 \cdot \mathbb{P}(V_0 < V_N), \quad n \to \infty$$

Thus we get $\mathbb{P}(V_N < V_0) = x/N$. $\square$

2.3 Recurrence and Transience

Problem 2.3.1 *Suppose y is recurrent and for $k \geq 0$, let $R_k = T_y^k$ be the time of the kth return to y, and for $k \geq 1$ let $r_k = R_k - R_{k1}$ be the kth interarrival time. Use the strong Markov property to conclude that under $\mathbb{P}_y$, the vectors $v_k = (r_k, X_{R_{k-1}}, \cdots, X_{R_k - 1})$, $k \geq 1$ are i.i.d.*

Proof To prove that $\{v_n\}_{n\geq 1}$ are independent, we need to show that for each $n \geq 2$, v_n is independent of $v_1, \cdots, v_{n-1}$. Using the shift operator, we know that:

$$v_n = v_1 \circ \theta_{R_{n-1}}, \quad n \geq 2$$

Since every v_i takes only countably possible values, we may just consider each $\mathbf{v}$ that v_i might take. By the strong Markov property, we know that:

$$\mathbb{P}_y(v_n = \mathbf{v}|\mathcal{F}_{R_{n-1}}) = \mathbb{P}_y(v_1 \circ \theta_{R_{n-1}} = \mathbf{v}|\mathcal{F}_{R_{n-1}}) = \mathbb{P}_{X_{R_{n-1}}}(v_1 = \mathbf{v}) = \mathbb{P}_y(v_1 = \mathbf{v})$$

This actually implies that v_n is independent of $\mathcal{F}_{R_{n-1}}$ because you can consider for any $A \in \mathcal{F}_{R_{n-1}}$, we have:

$$\mathbb{P}_y(v_n = \mathbf{v}, A) = \mathbb{E}_y[\mathbb{E}_y[\mathbf{1}_{\{v_n = \mathbf{v}\}}\mathbf{1}_A|\mathcal{F}_{R_{n-1}}]] = \mathbb{E}_y[\mathbf{1}_A\mathbb{P}(v_n = \mathbf{v}|\mathcal{F}_{R_{n-1}})] = \mathbb{P}(v_n = \mathbf{v})\mathbb{P}(A)$$

Thus v_n is independent of $\mathcal{F}_{R_{n-1}}$ and consequently independent of $v_1, \cdots, v_{n-1}$. So we have proved that $\{v_n\}_{n\geq 1}$ are independent. Finally by taking expectation over the first equation, we get that $\mathbb{P}_y(v_n = \mathbf{v}) = \mathbb{P}_y(v_1 = \mathbf{v})$. This shows that $\{v_n\}_{n\geq 1}$ are i.i.d. $\square$

Problem 2.3.2 *Use the strong Markov property to show that $\rho_{xz} \geq \rho_{xy}\rho_{yz}$.*

By the strong Markov property, we have:

$$\mathbb{P}_x(T_z \circ \theta_{T_y} < \infty | \mathcal{F}_{T_y}) = \mathbb{P}_y(T_z < \infty), \quad on\ \{T_y < \infty\}$$

Notice that:

$$\{T_y < \infty\} \cap \{T_z \circ \theta_{T_y} < \infty\} \subseteq \{T_z < \infty\}$$

Thus we have:

$$\begin{aligned}\mathbb{P}_x(T_z < \infty) \geq & \mathbb{P}(T_z \circ \theta_{T_y} < \infty, T_y < \infty) \\ = & \mathbb{E}_x[\mathbb{P}_x(T_z \circ \theta_{T_y} < \infty | \mathcal{F}_{T_y})\mathbf{1}_{\{T_y < \infty\}}] \\ = & \mathbb{E}_x[\mathbb{P}_y(T_z < \infty)\mathbf{1}_{\{T_y < \infty\}}] \\ = & \mathbb{P}_y(T_z < \infty)\mathbb{P}_x(T_y < \infty)\end{aligned}$$

This proves the inequality. □

Problem 2.3.3 *Show that in the Ehrenfest chain (**Example 5.1.3**), all states are recurrent.*

Proof Since for any $x, y \in S = \{0, 1, \cdots, r\}$, $\rho_{xy} > 0$, we know that S is irreducible. Thus all states are recurrent since $|S| < +\infty$. □

Problem 2.3.4 *f is said to be superharmonic if $f(x) \geq \sum_y p(x, y)f(y)$, or equivalently $f(X_n)$ is a supermartingale. Suppose p is irreducible. Show that p is recurrent if and only if every nonnegative superharmonic function is constant.*

Proof (⇒) Since $f(X_n)$ is a nonnegative supermartingale, it converges almost surely. However, X_n is a irreducible recurrent Markov chain, for each state i, X_n a.s. visits i for infinitely many times. This means that f could only be a constant.
(⇐) Suppose any nonnegative superharmonic function is constant. By using contradiction, if p is not recurrent, then all the states are transient. Fixing a state y, define:

$$f(x) = \mathbb{P}_x(T_y < \infty)$$

We first show that f is a nonnegative superharmonic function. By the Markov property:

$$\mathbb{P}_x(T_y < \infty) \geq \mathbb{P}_x(T_y \circ \theta_1 < \infty) = \mathbb{E}_x[\mathbb{P}_x(T_y \circ \theta_1 < \infty | \mathcal{F}_1)] = \mathbb{E}_x[\mathbb{P}_{X_1}(T_y < \infty)] = \sum_z p(x, z)\mathbb{P}_z(T_y < \infty)$$

This gives the result. Next we will show that it is impossible. Consider using the same method:

$$\begin{aligned} f(x) = & \mathbb{P}_x(X_1 = y) + \mathbb{P}_x(T_y < \infty, X_1 \neq y) \\ = & p(x, y) + \mathbb{P}_x(T_y \circ \theta_1 < \infty, X_1 \neq y) \\ = & p(x, y) + \sum_{z \neq y} p(x, z)\mathbb{P}_z(T_y < \infty) \\ (f \text{ is constant}) \quad = & p(x, y) + \sum_{z \neq y} p(x, z)f(x) \\ = & p(x, y) + f(x)(1 - p(x, y)) \end{aligned}$$

This means $p(x, y)(f(x) - 1) = 0$. Since we assume that the chain is irreducible, we know that $f(x) = f(y) = \mathbb{P}_y(T_y < \infty) < 1$. Thus $p(x, y) = 0$ for any state x. This is impossible since the chain is irreducible! As a result, p must be recurrent. □

2.4 Stationary Measure

Problem 2.4.1 *Let $w_{xy} = \mathbb{P}_x(T_y < T_x)$. Show that $\mu_x(y) = w_{xy}/w_{yx}$.*

Proof Recall that $\mu_x(y) = \mathbb{E}_x[N_y]$ denotes the expected times of visits to y before returning x. Here x is recurrent. Consider for any $k \geq 1$:

$$\{N_y = k\} = \{0 < T_y^1 < \cdots < T_y^k < T_x^1 < T_y^{k+1}\}$$

notice that since T_x^1 is finite, all the T_y^i are finite for $i = 1, \cdots, k$. So by using the strong Markov property, we get:

$$\begin{aligned}
\mathbb{P}_x(N_y = k) =& \mathbb{E}_x[\mathbf{1}_{\{0<T_y^1<\cdots<t_y^k<T_x^1<T_y^{k+1}\}}] \\
=& \mathbb{E}_x[\mathbf{1}_{\{0<T_y^1<\cdots<T_y^k<T_x^1\}}\mathbb{E}_x[\mathbf{1}_{\{T_x^1<T_y^{k+1}\}}|\mathcal{F}_{T_y^k}]] \\
=& \mathbb{E}_x[\mathbf{1}_{\{0<T_y^1<\cdots<T_y^k<T_x^1\}}\mathbb{E}_x[\mathbf{1}_{\{T_x^1<T_y^1\}\circ\theta_{T_y^k}}|\mathcal{F}_{T_y^k}]] \\
=& \mathbb{E}_x[\mathbf{1}_{\{0<T_y^1<\cdots<T_y^k<T_x^1\}}\mathbb{E}_y[\mathbf{1}_{\{T_x^1<T_y^1\}}]] \\
=& \mathbb{E}_x[\mathbf{1}_{\{0<T_y^1<\cdots<T_y^k<T_x^1\}}]\mathbb{P}_y(T_x < T_y) \\
=& \cdots \\
=& \mathbb{P}_x(T_y < T_x)\left(\mathbb{P}_y(T_y < T_x)\right)^{k-1}\mathbb{P}_y(T_x < T_y) \\
=& w_{xy}(1 - w_{yx})^{k-1}w_{yx}
\end{aligned}$$

The second equation above follows from the fact that $\{0 < T_y^1 < \cdots < T_y^k < T_x^1\} \in \mathcal{F}_{T_y^k}$. Thus we can get:

$$\mu_x(y) = \mathbb{E}_x[N_y] = \sum_{k=1}^{+\infty} k \cdot w_{xy}(1 - w_{yx})^{k-1}w_{yx} = \frac{w_{xy}}{w_{yx}}$$

This completes the proof! ☐

Remark Here is something subtle, here we could check directly that $\mathbb{P}_x(N_y = \infty) = 0$. If you use other methods, you should first check it out. An easy way (maybe) use the remarks and technique notes after **Theorem 5.5.7**

Problem 2.4.2 *Show that if p is irreducible and recurrent then:*

$$\mu_x(y)\mu_y(z) = \mu_x(z)$$

Proof From **Theorem 5.5.9**, we know that any two stationary measures differ only by a constant multiplier. Thus if $\mu_x(z)$ and $\mu_y(z)$ are simultaneously zero, the equation obviously holds. Otherwise, by the theorem, we have:

$$\frac{\mu_x(y)}{\mu_y(y)} = \frac{\mu_x(z)}{\mu_y(z)}$$

Notice that $\mu_y(y) = 1$ and the desired results follows. ☐

Problem 2.4.3 *Suppose p is irreducible and positive recurrent. Then $\mathbb{E}_xT_y < \infty$ for all x, y.*

Solution Since p is irreducible, any $x, y \in S$ satisfy $\rho_{xy} > 0$. Thus we can take $m = \inf\{n \in \mathbb{N} : p^n(x, y) > 0\}$ which is finite. By using C-K Equation, we can get $z_1, \cdots, z_{m-1}$ such that $p(x, z_1)p(z_1, z_2)\cdots p(z_{m-1}, y) = \mathbb{P}_x(X_1 = z_1, \cdots, X_{m-1} = z_{m-1}, X_m = y) > 0$. Here $z_i \neq x$ for all $1 \leq i \leq m-1$ by the definition of m. Now consider:

$$\begin{aligned}
\mathbb{E}_x[T_x] \geq& \mathbb{E}_x[T_x\mathbf{1}_{\{X_1=z_1,\cdots,X_{m-1}=z_{m-1},X_m=y\}}] \\
=& \mathbb{E}_x[\mathbb{E}_x[T_x|\mathcal{F}_m]\mathbf{1}_{\{X_1=z_1,\cdots,X_{m-1}=z_{m-1},X_m=y\}}] \\
(z_i \neq x) \quad \geq& \mathbb{E}_x[\mathbb{E}_x[T_x \circ \theta_m|\mathcal{F}_m]\mathbf{1}_{\{X_1=z_1,\cdots,X_{m-1}=z_{m-1},X_m=y\}}] \\
=& \mathbb{E}_x[\mathbb{E}_y[T_x]\mathbf{1}_{\{X_1=z_1,\cdots,X_{m-1}=z_{m-1},X_m=y\}}] \\
=& \mathbb{E}_y[T_x]\mathbb{P}_x(X_1 = z_1, \cdots, X_{m-1} = z_{m-1}, X_m = y) > 0
\end{aligned}$$

Since $\mathbb{E}_x[T_x] < \infty$ and $\mathbb{P}_x(X_1 = z_1, \cdots, X_{m-1} = z_{m-1}, X_m = y) > 0$, we get that $\mathbb{E}_y[T_x] < \infty$. This proves the result. ☐

Problem 2.4.4 *Suppose p is irreducible and has a stationary measure with $\sum_x \mu(x) = \infty$. Then p is not positive recurrent.*

Proof If the chain is irreducible and positive recurrent, then there exists a stationary distribution by **Theorem 5.5.12**. However **Theorem 5.5.9** tells us any two stationary measures differ only by a constant multiplier. But this could not happen since the stationary distribution has total mass 1 while the given stationary measure has total mass ∞. Thus p is not positive recurrent. □

Problem 2.4.5 *Use Theorems 5.5.7 and 5.5.9 to show that for simple random walk:*
(a) *The expected number of visits to k between successive visits to 0 is 1 for all k.*
(b) *If we start from k the expected number of visits to k before hitting 0 is $2k$.*

Proof (a) By the strong markov property, the desired expected number is equal to $\mu_0(k)$. Since we have known that $\mu(x) \equiv 1$ is a stationary measure, **Theorem 5.5.9** gives that $\mu_0(k) = \mu_0(0) = 1$.
Proof 1 of (b) Let N_k denote the times of visits before hitting 0. Without the loss of generality, we might assume that $k > 0$. Using the same method as in **Problem 2.4.1**, we can get for any $j \geq 1$:

$$\mathbb{P}_k(N_k = j) = \mathbb{P}_k(T_k < T_0)^{j-1}\mathbb{P}_k(T_0 < T_k) = (1 - \mathbb{P}_k(T_0 < T_k))^{j-1}\,\mathbb{P}_k(T_0 < T_k)$$

We only need to calculate $\mathbb{P}_k(T_0 < T_k)$. If $k \geq 2$, conditioning on the first step and use the markov property, we can get:

$$\begin{aligned}
\mathbb{P}_k(T_0 < T_k) =&\mathbb{E}_k[\mathbb{E}_k[\mathbf{1}_{\{T_0<T_k\}}|\mathcal{F}_1]] \\
=&\mathbb{E}_k[\mathbf{1}_{\{X_1=k-1\}}\mathbb{E}_k[\mathbf{1}_{\{T_0<T_k\}}|\mathcal{F}_1]] + \mathbb{E}_k[\mathbf{1}_{\{X_1=k+1\}}[\mathbb{E}_k[\mathbf{1}_{\{T_0<T_k\}}|\mathcal{F}_1]] \\
=&\mathbb{E}_k[\mathbf{1}_{\{X_1=k-1\}}\mathbb{P}_{k-1}(T_0 < T_k)] + 0 \\
=&\frac{1}{2}\mathbb{P}_{k-1}(T_0 < T_k)
\end{aligned}$$

The second is zero since if $X_1 = k + 1$, it will always first reach k before 0. Now use **Theorem 4.8.7** in the textbook, we know that $\mathbb{P}_{k-1}(T_0 < T_k) = 1/k$ and thus the desired probability is $1/2k$. If $k = 1$, this obviously holds. Finally we get:

$$\mathbb{E}_k[N_k] = \sum_{j=1}^{+\infty} j \cdot \frac{1}{2k} \cdot \left(1 - \frac{1}{2k}\right)^{j-1} = 2k$$

This completes the proof! □
Proof 2 of (b) For any $k \geq 1$, we have:

$$\mu_0(k) = \mathbb{E}_0[\sum_{n=0}^{T_0-1} \mathbf{1}_{\{X_n=k\}}] = \mathbb{E}_0[\sum_{n=1}^{T_0-1} \mathbf{1}_{\{X_n=k\}}] = \mathbb{E}_0[\sum_{n=1}^{T-1} \mathbf{1}_{\{X_n=k,X_1=1\}}] + \mathbb{E}_0[\sum_{n=1}^{T-1} \mathbf{1}_{\{X_n=k,X_1=-1\}}]$$

Since $\mathbb{E}_0[\sum_{n=1}^{T-1} \mathbf{1}_{\{X_n=k,X_1=-1\}}] = 0$, using markov property (the same as in the previous method):

$$\mu_0(k) = \mathbb{E}_0[\sum_{n=1}^{T-1} \mathbf{1}_{\{X_n=k,X_1=1\}}] = \mathbb{P}_0(X_1 = 1)\mathbb{E}_1[\sum_{n=0}^{T-1} \mathbf{1}_{\{X_n=k\}}] = \frac{1}{2}\mathbb{E}_1[\sum_{n=0}^{T-1} \mathbf{1}_{\{X_n=k\}}] = 1$$

Thus $\mathbb{E}_1[\sum_{n=0}^{T-1} \mathbf{1}_{X_n=k}] = 2$. Now consider starting from k, before hitting $k-1$, we visit k twice on average. Then starting from $k-1$, before hitting $k-2$, we visit k twice on average. Repeat the argument, we would get that starting from k, before hitting 0, we will visit k on average $2k$ times. □

2.5 Asymptotic Behaviour

Problem 2.5.1 *Suppose $S = \{0, 1\}$ and:*

$$p = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

Use induction to show that:

$$\mathbb{P}_\mu(X_n = 0) = \frac{\beta}{\alpha+\beta} + (1-\alpha-\beta)^n \left\{\mu(0) - \frac{\beta}{\alpha+\beta}\right\}$$

Proof It's easy to check that the equation holds when $n = 0$. Suppose it is right when $n = k-1$. For $n = k$, use the markov property, we get:

$$\begin{aligned}
&\mathbb{P}_\mu(X_k = 0) = \mathbb{E}_\mu[p(X_{k-1}, 0)] = \mathbb{P}_\mu(X_{k-1} = 0)\cdot p(0,0) + \mathbb{P}_\mu(X_{k-1} = 1)\cdot p(1,0)\\
=&\left(\frac{\beta}{\alpha+\beta} + (1-\alpha-\beta)^{k-1}\left\{\mu(0) - \frac{\beta}{\alpha+\beta}\right\}\right)(1-\alpha) + \left(1 - \frac{\beta}{\alpha+\beta} - (1-\alpha-\beta)^{k-1}\left\{\mu(0) - \frac{\beta}{\alpha+\beta}\right\}\right)\beta\\
=&\frac{\beta}{\alpha+\beta}(1-\alpha-\beta) + \beta + (1-\alpha-\beta)^k\left\{\mu(0) - \frac{\beta}{\alpha+\beta}\right\}\\
=&\frac{\beta}{\alpha+\beta} + (1-\alpha-\beta)^k\left\{\mu(0) - \frac{\beta}{\alpha+\beta}\right\}
\end{aligned}$$

This shows that the equation holds when $n = k$ and the proof is then complete. □

Problem 2.5.2 *Show that if S is finite and p is irreducible and aperiodic, then there is an m so that $p^m(x, y) > 0$ for all x, y.*

Proof By **Lemma 5.6.5** in the textbook, we know that for any $x \in S$, there is $N_x \in \mathbb{N}$ such that $p^n(x, x) > 0$ whenever $n \geq N_x$. Moreover, by the condition of irreducible, for any pair $x, y \in S$, there is $n(x, y) \in \mathbb{N}$ such that $p^{n(x,y)}(x, y) > 0$. Now let:

$$N = \max\left\{\max_{x\in S}\{N_x\}, \max_{x,y\in S}\{n(x,y)\}\right\}$$

Notice that N is finite since $|S| < \infty$. Consider if $n = 2N$, we have:

$$p^{2N}(x,y) \geq p^{n(x,y)}(x,y)p^{2N-n(x,y)}(y,y) > 0$$

The first inequality follows from C-K equation. This gives the desired $m = 2N$. □

Problem 2.5.3 *Show that if S is nite, p is irreducible and aperiodic, and T is the coupling time dened in the proof of **Theorem 5.6.6** then $\mathbb{P}(T > n) \leq Cr^n$ for some $r < 1$ and $C < \infty$. So the convergence to equilibrium occurs exponentially rapidly in this case. Hint: First consider the case in which $p(x, y) > 0$ for all x and y and reduce the general case to this one by looking at a power of p.*

Proof Case 1 If $p(x, y) > 0$ for all x and y, set $\epsilon = \min_{x,y}\{p(x, y)\} > 0$. Also, let $N = |S|$. Consider for any $x \neq y$, we have:

$$\mathbb{P}(T = n+1|X_n = x, Y_n = y, T > n) = \mathbb{P}(X_n + 1 = Y_{n+1}|X_n = x, Y_n = y) = \sum_{z\in S} p(x,z)p(y,z) \geq \epsilon^2 N$$

Notice that:

$$\{T > n\} = \bigsqcup_{x\neq y}\{T > n\} \cap \{X_n = x, Y_n = y\}$$

From this we can deduce that for all $n \in \mathbb{N}$ (just basic calculations of probability):

$$\mathbb{P}(T = n+1|T > n) \geq \epsilon^2 N \quad \Rightarrow \quad \mathbb{P}(T > n+1|T > n) < 1 - \epsilon^2 N$$

Thus by induction, we get:

$$\mathbb{P}(T > n) = \mathbb{P}(T > n|T > n-1)\mathbb{P}(T > n-1) < (1-\epsilon^2 N)\mathbb{P}(T > n-1) < \cdots < (1-\epsilon^2 N)^n \mathbb{P}(T > 0)$$

Take $r = 1 - \epsilon^2 N$ and $C = \mathbb{P}(T > 0)$, the desired result follows.
Case 2 In the general case, we can use the result of **Problem 2.5.2**. It says there exists $m \in \mathbb{N}$ such that

$p^m(x,y) > 0$ for all x and y. Now take $\epsilon = \min_{x,y}\{p^m(x,y)\} > 0$. Use the same method as in case1, we can prove that for all $n \in \mathbb{N}$:

$$\mathbb{P}(T > n+m | T > n) < 1 - \epsilon^2 N$$

Thus by induction, we come to:

$$\mathbb{P}(T > n) < (1-\epsilon^2 N)^{\lfloor \frac{n}{m} \rfloor}\mathbb{P}(T > n(\operatorname{mod} m)) \leq (1-\epsilon^2 N)^{\frac{n}{m}-1}\mathbb{P}(T > n(\operatorname{mod} m))$$

Take $r = (1-\epsilon^2 N)^{1/m}$ and $C = \frac{1}{r}\mathbb{P}(T > n(\operatorname{mod} m))$. The desired result follows. $\square$