UChicago Markov Chains, Martingales, and Brownian Motion Notes: 23500

Notes by Agustín Esteva, Lectures by Stephen Yearwood, Book by Greg Lawler

Academic Year 2024-2025

Contents

1 Lectures

1.1 Monday, Mar 24: Markov Processes: Basic Definitions and Examples

He is from Trinidad and Tobago.

Definition 1. A **stochastic process** is a collection of random variables $\{X_t\}_{t\in T}$ indexed by time, where each X_t takes values in S

We call S our *state space*. In discrete time, it should be obvious that $T = \mathbb{N}_0^*$ (where $\mathbb{N}_0^*$ is when on goscale the naturals by a constant) and $T = [0, \mathbb{R})$ in continuous time. We say that S is a discrete space if it is countable and continuous if it is $\mathbb{R}^d$.

Remark 1. In order to characterize the distribution of $\{X_n\}$, we must specify $\mathbb{P}\{X_0 = S_0, \dots, X_n = S_n\}$ for all $n \in \mathbb{N}$ and for all $S_0, \dots, S_n \in S$. It is much easier work with conditional probability. We will see that with Markov Chains, the Markov Property guarantees that we only need to worry about the distribution of X_{n-1} to figure out X_n.

Definition 2. If E, F are events with $\mathbb{P}\{F\} > 0$, then the **conditional probability** of E given F is

$$\mathbb{P}(E|F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}$$

Proposition 1. (Law of Total Probability) Recall that if $(B_n) \in S$ is a sequence of mutually exclusive and exhaustive events, then

$$P(A) = \sum_{n=1}^{\infty} P(A \cap B_n) = \sum_{n=1}^{\infty} P(A \mid B_n)P(B_n)$$

Definition 3. A stochastic process is called a **Markov Chain** if for all $n \in \mathbb{N}$, and for all $S_0, \dots, S_n \in S$, we have that

$$\mathbb{P}(X_n = S_n | X_0 = S_0, \dots, X_{n-1} = S_{n-1}) = \mathbb{P}(X_n = S_n | X_{n-1} = S_{n-1})$$

Definition 4. A Markov Chain $\{X_n\}$ is **time-homogeneous** if for all $n \in N$, for all $x, y \in S$,

$$\mathbb{P}(X_n = y | X_{n-1} = x) = \mathbb{P}(X_1 = y | X_0 = x)$$

Thus, it does not matter when you get to the states, and so we can just specify the distribution of X_0 and the transition probability;

$$p(x, y) := \mathbb{P}(X_1 = y | X_0 = x)$$

and then scale to find the rest.

Example 1.1. Let $S = \{0, 1\}$. A Markov chain taking values in S specified by $p = p(0, 1)$ and $q = p(1, 0)$. It is obvious that $p(0, 0) = 1 - p$ and $q(1, 1) = 1 - q$.

Example 1.2. Let $S = \mathbb{Z}$. Let $\{X_n\}_{n\geq 0}$ be defined by $X_0 = X_1 = X_2 = 0$ and for $n \geq 3$, then

$$X_n = \begin{cases} X_{n-1} + 1, \\ X_{n-1} - 1, \\ X_{n-3} \end{cases}$$

each with with probability 1/3. This is NOT a Markov Process since X_n can depend on X_{n-3}.

Definition 5. The $n-$step transition probabilities

$$p^n(x, y) = \mathbb{P}(X_n = y | X_0 = x).$$

That is, if we start at x, what is the probability that the nth step is y?

Proposition 2. For all $m, n \in \mathbb{N}$ and for all $x, y \in S$, then

$$p^{n+m}(x,y) = \sum_{z \in S} p^n(x,z)p^m(z,y)$$

Proof. We have that by time homogeneity,

$$p^n(x,z)p^m(z,y) = \mathbb{P}\{X_n = z \mid X_0 = x\}\mathbb{P}\{X_{n+m} = y \mid X_n = z\}$$

By the Markov Property,

$$\mathbb{P}\{X_{n+m} = y \mid X_n = z\} = \mathbb{P}\{X_{n+m} = y \mid X_n = z, X_0 = x\}$$

By Bayes' rule:

$$\mathbb{P}\{X_n = z \mid X_0 = x\}\mathbb{P}\{X_{n+m} = y \mid X_n = z, X_0 = x\} = \mathbb{P}\{X_{n+m} = y, X_n = z \mid X_0 = x\}$$

Thus,

$$\sum_{z \in S} p^n(x,z)p^m(z,y) = \sum_{z \in S} \mathbb{P}\{X_{n+m} = y, X_n = z \mid X_0 = x\} = \mathbb{P}\{X_{n+m} = y \mid X_0 = x\} = p^{n+m}(x,y)$$

□

Definition 6. The **transition matrix** of a Markov Chain the the $N \times N$ matrix P whose ij entry is $p(i,j)$.

Proposition 3. For each $n \in \mathbb{N}$, P^n is the matrix whose ij entry is $p^n(i,j)$.

Proof. We prove by induction. It is trivial for $n = 1$. Assume it is true for any general $n - 1$. Thus, the ij entry of $P^n = P^{n-1}P$ is

$$\sum_{k=1}^{n} P_{ik}^{n-1} P_{kj} = \sum_{k=1}^{n} p^{n-1}(i,k)p(k,j) = p^n(i,j),$$

where the last equality comes from Proposition 1. □

1.2 Wednesday, Mar 26: Recurrence and Transience for Finite State Space

We illustrate the stuff from last class with a simple example:

Example 1.3. Consider the two state Markov chain with $S = \{0, 1\}$ and

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

then

$$P = \begin{pmatrix} p(0,0) & p(0,1) \\ p(1,0) & p(1,1) \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

then

$$\mathbb{P}\{X_3 = 0 | X_0 = 0\} = p^3(0,0) = \frac{65}{108}$$

For the following, we consider a Markov chain $\{X_n\}$ with state space S.

Definition 7. Two states $x, y \in S$ **communicate** if there exists $m, n > 0$ such that $p^n(x,y) > 0$ and $p^m(y,x) > 0$. We write $x \leftrightarrow y$.

Remark 2. Communication is an equivalence relation, and so we can partition S into a disjoint union of communication classes by modding out the communication classes.

Definition 8. A communication class c is **recurrent** if $p(x,y) = 0$ for all $x \in C$ and $y \in S \setminus \{C\}$. Otherwise, we say that the communication class is **transient**.

In other words, if C is recurrent, then the chain never leaves. If it is transcient, then there is no problem leaving.

Definition 9. A Markov chain is **irreducible** if there is only one communication class.

Example 1.4. Consider a Markov chain with $S = \{1, 2, 3, 4, 5\}$ and

$$P = \begin{pmatrix} \frac{1}{5} & \frac{1}{5} & 0 & 0 & \frac{3}{5} \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} & \frac{3}{4} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \end{pmatrix}$$

then

$$C_1 = \{1, 2, 5\}, \qquad C_2 = \{3, 4\}$$

and C_1 is transient and C_2 is recurrent

Example 1.5. (Gambler's ruin) Consider the random walk on $S = \{0, 1, 2, \dots, N\}$ with absorbing boundary, and transition probability

$$p(x, x+1) = p(x, x-1) = \frac{1}{2}, \quad x \in \{1, 2, \dots, N-1\}$$

and

$$p(0,0) = p(N,N) = 1$$

$\{1, \dots, N-1\}$ is a transient communication class, while $\{0\}, \{N\}$ are both recurrent communication classes.

Proposition 4. Suppose S is finite. If C is a recurrent communication class, then if $\{X_n\}$ starts in C, with probability 1, $\{X_n\}$ visits every state in C infinitely often. That is, for each $x, y \in C$, $\mathbb{P}\{X_n = y \text{ i.o.} \mid X_0 = x\} = 1$.

Proof. Since C is a communication class, then for every $z \in C$, there exists some $n_z \in \mathbb{N}$ such that $p^{n_z}(z, y) > 0$. Let $n = \max\{n_z \ \forall z \in C\}$, and $q = \min\{p^{n_z}(z, y) \mid z \in C\}$. Note that this quantities necessarily exist because S is finite. Let $E_k = \{X_i = y, \mid i \in \{(k-1)n + 1, (k-1)n + 2 \ldots (k-1)nk\}\}$. Then for states $s_0, s_1, \ldots, s_{nk} \in S$, we have that

$$\mathbb{P}\{E_{k+1} \mid X_0 = s_0, X_1 = s_1, \ldots, X_{nk} = S_{nk}\} = \mathbb{P}\{E_1 \mid X_0 = s_{nk}\} \geq q.$$

For $M, N \in \mathbb{N}$ with $M > N$, we have that

$$\begin{aligned}
\mathbb{P}\{E_k \text{ does not occur for any any } k \in \{N, N+1, \ldots, M\}\} &= \mathbb{P}\{\bigcap_{k=N}^{M} E_k^c\} \\
&= \mathbb{P}\{E_M^c \mid \bigcap_{k=N}^{M-1} E_k\}\mathbb{P}\{\bigcap_{K=N}^{M-1} E_k\} \\
&\leq (1-q)\mathbb{P}\{\bigcap_{K=N}^{M-1} E_k\} \\
&\leq \cdots \leq (1-q)^{M-N} \to 0
\end{aligned}$$

□

Proposition 5. Suppose S is finite. If C is a transient communication class, then w.p. 1, X eventually leaves C and never returns.

Remark 3. In fact, there exists a $c > 0$ such that if $X_0 = x$ and C is transient, then

$$\mathbb{P}\{X_n \text{ exits } C \text{ before time } n \geq 1 - e^{-cn}\}$$

Thus, if we let

$$\tau = \inf\{n \geq 0 : X_n \notin C\},$$

then

$$\mathbb{P}\{\tau \geq n\} = 1 - \mathbb{P}\{X_n \text{ exits } C \text{ before time } n \leq 1 - (1 - e^{-cn}) = e^{-cn} \to 0$$

as $n \to \infty$.

1.3 Friday, Mar 28: The Strong Markov Property

Definition 10. A random time $\tau \in \mathbb{N}_0 \cup \{\infty\}$ is called a **stopping time** if, for all $n \in \mathbb{N}$, the event $\{\tau = n\}$ is determined by $X_0, X_1, \ldots, X_n$.

Example 1.6. (a) (The Hitting Time)

$$\tau = \min\{n \mid X_n = x\}, \quad x \in S.$$

In the Gambler's ruin model, where the gambler starts with $k-$dollars and gambles all the way to N or 0 dollars, the hitting time is the first the the gambler reaches \$1.

(b) $\tau = k^{\text{th}}$ time for which $X_n \in A$, $k \in \mathbb{N}$, $A \subset S$.

(c) $\tau = \min\{\tau_1, \tau_2\}$, where τ_1 and τ_2 are stopping times.

(d) Let $N \in \mathbb{N}$, $x \in S$, τ be the last time $n \leq N$ for which $X_n = x$, and $\tau = 0$ if no such time exists. τ is NOT a stopping time because it depends on stuff in the future.

Theorem 1. (Strong Markov Property) Let τ be a stopping time for $\{X_n\}$. Let $n \geq 0$, $m \geq 1$, $x_n \ldots, x_n \in S$ such that

$$\mathbb{P}\{X_0 = x_0, \ldots, X_\tau = X_n\} > 0,$$

and let $y_1, \ldots, y_m \in S$. Then

$$\mathbb{P}\{X_{\tau+1} = y, \ldots, X_{\tau+m} = y_m \mid X_0 = x_0, \ldots, X_\tau = x_n\} = \mathbb{P}\{X_1 = y, \ldots, X_m = y_m \mid X_0 = x_n\}$$

Proof. The event $\{X_0 = x_0, \ldots, X_\tau = x_n\}$ is the same as the event $\{X_0 = x_0, \ldots, X_n = x_n\}$ and $\{\tau = n\}$. Since τ is a stopping time determined where the event $\{\tau = n\}$ by $\{X_0, \ldots, X_n\}$. Thus, we get that the events $\{X_0 = x_0, \ldots, X_\tau = x_n\} = \{X_0 = x_0, \ldots, X_n = n\} \cap \{\tau = n\}$, then

$$\begin{aligned}
\mathbb{P}\{X_{\tau+1} = y_1, \ldots, X_{\tau+m} = y_m \mid X_0 = x_0, \ldots, X_\tau = x_n\} &= \mathbb{P}\{X_{n+1} = y_1, \ldots, X_{n+m} = y_m \mid X_0 = x_0, \ldots, X_n = x_n\} \\
&= \mathbb{P}\{X_{n+1} = y_1, \ldots, X_{n+m} = y_m \mid X_n = x_n\} \\
&= \mathbb{P}\{X_1 = y_1, \ldots, X_m = y_m \mid X_0 = x_n\}
\end{aligned}$$

and conclude with the regular Markov property and $n-$step invariance. □

Example 1.7. Let $x \in S$ with S finite and let $\tau = \min\{n \geq 0 \mid X_n = x\}$. Assume that $\mathbb{P}\{\tau < \infty\} = 1$. For any $x_0, \ldots, x_n \in S$ such that $\mathbb{P}\{X_0 = x_0, \ldots, X_\tau = x_n\} > 0$, we have $x_n = x$, and so the conditional distribution of $\{X_{\tau+j}\}_{j \geq 0}$ given everything up to τ is the same as the original chain starting at x.

Thus, the Strong Markov property essentially guarantees that the Markov chain resets after hitting τ!

Definition 11. Let X_n be a Markov chain on a countable state space S. For any $x \in S$, the **period** of x is the greatest common divisor of $J_x = \{n \geq 1 \mid p^n(x, x) > 0\}$.

Example 1.8. If $p(x, x) > 0$, then $1 \in J_x$, and so the period of x is 1.

Theorem 2. If $x \leftrightarrow y$, the periods of x and y, denoted by d_x and d_y respectively, are the same.

Proof. Choose n, m such that $p^n(x, y) > 0$ and $p^m(y, x) > 0$. Then $p^{n+m}(x, x) > 0$ and $p^{m+n}(y, y) > 0$. Thus, $n + m \in J_x \cap J_y$. Assume that $d_x < d_y$. Then there exists some $k \in K_x$ not divisible by d_y. Then $n + m + k \in J_y$, and so d_y divides $n + m + k$ and $n + m$, and so d_y divides k, which is a contradiction. □

Definition 12. A Markov chain is aperiodic if every state has period 1.

Example 1.9. A knight on an 8×8 chessboard selects one of the next legal moves with equal probability, independently of the past.

$$S = \{(x, y) \in \mathbb{R}^2 \mid x \in \{1, \ldots, 8\}, y \in \{1, \ldots, 8\}\}.$$

The period is 2.

1.4 Monday, Mar 31: Stationary Distributions

Definition 13. Let $\tilde{v} = (\tilde{v}_1, \tilde{v}_2, \dots, \tilde{v}_n)$ be a vector such that $\tilde{v}_j = \mathbb{P}\{X_0 = j\}$. We say that $\tilde{v}$ is the **initial distribution** of the Markov chain.

Proposition 6. For each $i \in [n]$, then ith entry of the row vector $\tilde{v}P$ is $\mathbb{P}\{X_1 = i\}$.

Example 1.10. Consider the Markov chain with $S = \{1, 2, 3, \}$ and

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

For n large,

$$P^n = \begin{pmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \end{pmatrix}.$$

We encounter the phenomena that the limit of P^n, if it exists, has identical rows.

Call this row π. For any probability vector $\tilde{v}$, $\lim_{n\to\infty} \tilde{v}P^n = \pi$

Suppose π is a limiting probability vector. That is, for any initial distribution $\tilde{v}$, $\lim_{n\to\infty} \tilde{v}P^n = \pi$. Then

$$\pi = \lim_{n\to\infty} \tilde{v}P^{n+1} = (\lim_{n\to\infty} \tilde{v}P^n)P = \pi P$$

We say that π is a stationary/invariant/equilibrium/steady-state distribution for the Markov chain.

Definition 14. Let $\pi : S \to [0, 1]$ be a probability distribution on S such that $\sum_{x\in S} \pi_x = 1$. We say that π is a **stationary distribution** of $\{X_n\}$ if

$$\pi_y = \sum_{x\in S} \pi_x p(x, y), \quad \forall y \in S.$$

That is, $\pi P = \pi$.

Theorem 3. If S is finite and $\{X_n\}$ is an irreducible and aperiodic chain, then there exists a unique stationary distribution π for $\{X_n\}$. Moreover, for any $x, y \in S$,

$$\lim_{n\to\infty} p^n(y, x) = \pi_y$$

Proof. (Existence) Fix $z \in S$. Suppose $X_0 = z$. Let $\tau = \min\{n \geq 1 \mid X_n = z\}$ be the first return time to z. Note that $\tau < \infty$ by proposition 4. For any $x \in S$, define

$$\tilde{\pi}_x := \mathbb{E}[\#\{n \in \{0, 1, \dots \tau - 1\} \mid X_n = x\}]$$

We claim that

$$x \mapsto \frac{\pi_x}{\mathbb{E}[\tau]}$$

is a stationary distribution for $\{X_n\}$. We want to show that for all $y \in S$, $\tilde{\pi}_y = \sum_{x\in S} \tilde{\pi}_x p(x, y)$. Evidently,

$$\tilde{\pi}_x = \mathbb{E}\left[\sum_{x\in S} \mathbb{1}_{\{X_n = x\}}\right] = \mathbb{E}\left[\sum_{n=0}^{\infty} \mathbb{1}_{\{X_n = x, \tau > n\}}\right] = \sum_{n=0}^{\infty} \mathbb{P}\{X_n = x, \tau > n\}.$$

Thus,

$$\sum_{x\in S}\tilde{\pi}_x p(x,y) = \sum_{x\in S}\sum_{n=0}^{\infty}\mathbb{P}\{X_n = x, \tau > n\}p(x,y) = \sum_{n=0}^{\infty}\sum_{x\in S}\mathbb{P}\{X_n = x, T > nX_{n+1} = y\} = \sum_{n=0}^{\infty}\mathbb{P}\{\tau > n+1X_{n+1} = y\} = \tilde{\pi}_y$$

In order to find a stationary probability distribution, we need that

$$\sum_{x\in S}\frac{\tilde{\pi}_x}{\mathbb{E}[\tau]} = 1.$$

But then

$$\tilde{pi}_x = \sum_{n=0}^{\infty}\mathbb{P}\{X_n = x, \tau > n\} = \sum_{n=0}^{\infty}\mathbb{P}\{\tau > n, X_n = x \mid X_n = x\}.$$

So then

$$\tilde{\pi}_x = \sum_{n=0}^{\infty}\sum_{x\in S}\mathbb{P}\{\tau > n \mid X_n = x\}\mathbb{P}\{\tau > n \mid X_n = x\}\mathbb{P}\{X_n = x\} = \sum_{n=0}^{\infty}\mathbb{P}\{\tau > n\} = \mathbb{E}[\tau]$$

□

Example 1.11. Consider a Markov chain $\{X_n\}$ with

$$P = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{5} & \frac{1}{5} & \frac{3}{5} \end{pmatrix}.$$

$\{X_n\}$ is clearly aperiodic and irreducible. To compute π, we solve the system $\pi P = \pi$, with $\pi_1+\pi_2+\pi_3 = 1$.

$$(\pi_1,\pi_2,\pi_3)P = \pi \iff \frac{1}{3}\pi_1 + \frac{1}{2}\pi_2 + \frac{1}{5}\pi_3 = \pi_1, \quad \frac{1}{3}\pi_1 + \frac{1}{5}\pi_3 = \pi_2, \quad \frac{1}{3}\pi_1 + \frac{1}{2}\pi_2 + \frac{3}{5}\pi_3.$$

Alternatively, we know that

$$\pi P = \pi \iff P^T\pi^T = \pi,$$

and so we solve for the eigenvectors of P^T corresponding to $\lambda = 1$ and then normalizing so that $\pi_1+\pi_2+\pi_3 = 1$. Solving,

$$\pi = (\frac{3}{10}, \frac{1}{5}, \frac{1}{2})$$

1.5 Wednesday, Apr 2: Uniqueness of Stationary Distributions

Theorem 4. If π is a stationary distribution for $\{X_n\}$, then for any $x, y \in S$, then $\lim_{n\to\infty}\{\mathbb{P}\{X_n = y \mid X_0 = x\}\} = \pi_y$. Then π is unique.

Proof. Let $\{X_n\}$ and $\{Y_n\}$ be two Markov chains starting at x and y, respectively. Consider the Markov chain (X_n, Y_n) in $S \times S$ with transition probability

$$\overline{p}((x,y),(x',y')) = \begin{cases} p(x,x')p(y,y'), & x \neq y \\ p(x,x'), & x' = y' \\ 0, & x' \neq y' \end{cases}$$

Note that $\mathbb{P}\{X_1 = x' \mid X_0 = x, Y_0 = y\} = \sum_{y\in S} p(x,x')p(y,y') = p(x,x')$. Then $\{X_n\}$ and $\{Y_n\}$ have the same as our original Markov chain. Let

$$\tau = \min\{n \mid X_n = Y_n\}.$$

We claim that $\mathbb{P}\{\tau < \infty \mid X_0 = x, Y_0 = y\} = 1$. We prove this on the homework.

Consider (X_n, Y_n) where where $X_0 = x$ and $Y_0 = \pi$, then Y_n has distance π for all n, and for large n, $X_n = Y_n$. FOr any y,

$$\lim_{n\to\infty} \mathbb{P}\{X_n = y \mid X_0 = x\} - \pi_y = \lim_{n\to\infty} (\mathbb{P}\{X_n = y\} - \mathbb{P}\{Y_n - y\}) = 0.$$

Thus, if $\pi, \tilde{\pi}$ are two stationary distributions, then

$$\pi_y = \lim_{n\to\infty} \mathbb{P}\{X_n = y \mid X_0 = x\} = \tilde{\pi}_y$$

□

Proposition 7. Let π be a stationary distribution for $\{X_n\}$ and let $T_x := \min\{n \geq 1 \mid X_n = x\}$. Then

$$\pi_x = \frac{1}{\mathbb{E}[T_x \mid X_0 = x]}$$

Proof. Assume $X_0 = x$. Then the stationary distribution is given by

$$\pi_y = \frac{1}{\mathbb{E}[T_x]}\mathbb{E}[|\{n \in [T_x - 1] \mid X_n = y|].$$

By definition,

$$|n \in \{0, 1, \ldots, T_x - 1\} \mid X_n = x| = 1$$

so then we are done. □

Example 1.12. Let $G = (V, E)$ be a finite connected, non-bipartite graph. Then $\{X_n\}$ is irreducible and aperiodic. It is not hard to show that the sum of all the degrees is $2|E|$ since you need to count each edge twice. The stationary distribution for $\{X_n\}$ is

$$\pi_X = \frac{\deg x}{2|E|}.$$

Note that π is indeed a valid probability distribution since

$$\sum_{x\in V} \pi_x = \sum_{x\in V} \frac{\deg x}{2|E|} = \frac{2|E|}{2|E|} = 1.$$

If $p(x,y) = \frac{1}{\deg x}$ and $x \sim y$ when they are joined by an edge and 0 else, then for any $y \in V$,

$$\sum_{x \in V} \pi_x p(x,y) = \sum_{x | x \sim y} \frac{1}{\deg x} \frac{\deg x}{2|E|} = \frac{\deg y}{2|E|} = \pi_y$$

Example 1.13. Consider the Knight's Tour, where a knight chooses a legal move randomly on an 8×8 chessboard. What is expected time to return to the bottom left corner. The degrees are given by:

$$\begin{bmatrix} 2 & 3 & 4 & 4 & 4 & 4 & 3 & 2 \\ 3 & 4 & 6 & 6 & 6 & 6 & 4 & 3 \\ 4 & 6 & 8 & 8 & 8 & 8 & 6 & 4 \\ \vdots \\ 2 & 3 & 4 & 4 & 4 & 4 & 3 & 2 \end{bmatrix}$$

Where the sum of degrees is 336, and so the expected time to return to $(1,1)$ is $\frac{336}{2} = 168$.

Example 1.14. (King's) Tour

Example 1.15.

$$\begin{bmatrix} 3 & 5 & 5 & 5 & 5 & 5 & 5 & 3 \\ 5 & 8 & 8 & 8 & 8 & 8 & 8 & 5 \\ \vdots \\ 3 & 5 & 5 & 5 & 5 & 5 & 5 & 3 \end{bmatrix}$$

So then the expected number of moves till return is 140.

1.6 Friday, Apr 4: First Step Analysis

Our goal is to calculate the expected duration (that is, the hitting probability) conditioned on a first step. Let's illustrate using an example! Before, we do a little aside to talk about a useful technique that comes up naturally:

Remark 4. Suppose f satisfies

$$f(n) = af(n-1) + bf(n+1), \qquad 0 \le n \le N,$$

where $f(0)$ and $f(N)$ are known. Suppose $a \neq b$, then guess the solution to be

$$f(n) = \alpha^n,$$

then

$$\alpha^n = a\alpha^{n-1} + b\alpha^{n+1} \implies \alpha = a + b\alpha^2 \implies \alpha = \frac{1 \pm \sqrt{1-4ab}}{2b}.$$

Thus the general solution is

$$f(n) = \lambda_1 \alpha_+^n + \lambda_2 \alpha_-^n$$

If $a = b$, then

$$f(n) = \lambda_1 n + \lambda_2,$$

since we are left with a simple linear system.

Example 1.16. Consider the Gambler's Ruin where $S = \{0, 1, 2, \ldots, N-1, N\}$ with transition probabilities

$$p(x, x+1) = p, \qquad p(x, x-1) = q, \quad x \notin \{0, N\}$$

$$p(0,0) = p(N,N) = 1.$$

That is, we have two absorbing boundaries and so we have three communication classes:

$$C_1 = \{0\}, \quad C_2 = \{1, 2, \ldots, N-1\}, \quad C_3 = \{N\}.$$

Note that C_1 and C_3 are recurring and C_2 is transient. Define

$$\tau_k := \min\{n > 0 : X_n \in C_1 \cup C_3 \mid X_0 = k\}, \quad P_k = \mathbb{P}\{X_{\tau_k} = N\}.$$

We defined P_k to be the probability that the gambler wins starting with k dollars. Clearly, $P_0 = 0$, $P_N = 1$. Using the law of total probability

$$P_k = \mathbb{P}\{X_{\tau_k} = N\} = \mathbb{P}\{X_{\tau_k} = N \mid X_1 = k+1\}\}\mathbb{P}\{X_1 = k+1\} + \mathbb{P}\{X_{\tau_k} = N \mid X_1 = k-1\}\}\mathbb{P}\{X_1 = k-1\} = pP_{k+1} + qP_{k-1}.$$

Using the above remark we see that

$$\alpha_+ = 1, \quad \alpha_- = \frac{1-p}{p}$$

so

$$P_n = \lambda_1 + \lambda_2 (\frac{1-p}{p})^n \implies P_0 = \lambda_1 + \lambda_2 = 0, \quad P_N = \lambda_1 + \lambda_2 (\frac{1-p}{p})^N = 1$$

Solving:

$$P_n = \frac{1 - (\frac{1-p}{p})^n}{1 - (\frac{1-p}{p})^N}, \quad p \neq \frac{1}{2},$$

$$P_n = \frac{n}{N}, \quad p = \frac{1}{2}$$

1.7 Monday, Apr 7: Countable State Space

Suppose the state space S is countably infinite.

Definition 15. We say a Markov Chain $\{X_n\}$ is **irreducible** if for all $x, y \in S$, there exists some $n \in \mathbb{N}$ such that $p^n(x,y) > 0$

Definition 16. We say $\{X_n\}$ is **recurrent** if

$$\mathbb{P}\{X_n = x \text{ i.o.} \mid X_0 = x\} = 1.$$

We say $\{X_n\}$ is **transient** if it is not recurrent.

Proposition 8. Suppose $\{X_n\}$ is irreducible. Then either every state is recurrent or every state is transient.

Proof. Suppose x is a recurrent state. Without loss of generality, let $X_0 = x$. Let $\tau_1, \tau_2, \ldots$ be times to successive visits to x. Note that $\tau_k < \infty$ since x is recurrent. Note also that $\{X_{\tau_k}, \ldots, X_{\tau_{k+1}}\}_{k \geq 0}$ are all i.i.d by the strong Markov property. Let $y \in S$. There is some $n \in \mathbb{N}$ such that $p^n(x,y) > 0$ by the irreducibility of $\{X_n\}$. Thus, there is some k such that

$$q := \mathbb{P}\{y \in \{X_{\tau_k}, \ldots, X_{\tau_{k+1}}\}\} > 0.$$

Since the increments are identical, each increment has probability q of containing y. Thus, there is a probability 1 that infinitely many of them contain y. That is, if $\sigma = \min\{n \geq 0 \mid X_n = y\}$, then

$$\mathbb{P}\{\sigma < \infty \mid X_0 = x\} = 1.$$

By the strong Markov property, $\{X_{\sigma+j}\}$ has the same distribution as X_n started at y. But we know that $\{X_{\sigma+j}\}$ visits y infinitely many times starting at x, and so $\{X_n\}$ will visit y infinitely many times starting at y. $\square$

Proposition 9. A state $x \in S$ is recurrent if and only if $\displaystyle\sum_{n=0}^{\infty} p^n(x,x) = \infty$.

Proof. We will prove the contrapositive of the forward direction. Define the total number of visits to x by

$$R_x := \sum_{n=0}^{\infty} \mathbb{1}_{X_n = x}.$$

Then using linearity of expectation we fine that

$$\begin{aligned}
\mathbb{E}[R_x] &= \sum_{n=0}^{\infty} \mathbb{E}[\mathbb{1}_{X_n = x}] \\
&= \sum_{n=0}^{\infty} p^n(x,x) \\
&< \infty.
\end{aligned}$$

Since $\mathbb{E}[R_x] < \infty$, then $R_x < \infty$ almost surely, and thus x is transient.

Suppose x is transient. Define $\tau_1, \tau_2, \ldots$ as successive visits to x. By transience, with probability 1, there is some k such that $\tau_k = \infty$. Let $q = \mathbb{P}\{\tau_k = \infty\}$ By the strong Markov property, $\tau_k - \tau_{k-1}$ are i.i.d. Thus, each increment has probability q of being infinite. Then R_x is the smallest k such that $\tau_{k+1} = \infty$. Then $R_x \sim \text{Geometric}(q)$ with the 'success' of $\tau_{k+1} = \infty$. Thus, $\mathbb{E}[R_x] = \frac{1}{q} < \infty$ and we are done since

$$\sum_{n=0}^{\infty} p^n(x,x) = \mathbb{E}[R_x \mid X_0 = x] = \frac{1}{q} < \infty$$

$\square$

Example 1.17. Consider the Markov Chain $\{X_n\}$ with $S = \{0, 1, 2, \dots\}$. Then if the transition probabilities are given by

$$p(x,0) = \frac{1}{x+2}, \quad p(x,x+1) = 1 - \frac{1}{x+2},$$

we claim that p^n is recurrent. To see this, it suffices to see that the series of $p^n(0,0)$ diverges.

1.8 Wednesday, Apr 9: Queuing and Stationary Distributions for Countable State Spaces

Example 1.18. Let's continue the queue example from last class.

Let $\{X_n\}$ be the number of people at time n. Let $p \in (0,1)$ be the probability each person arrives and $q \in (0,1)$ be the probability each person leaves the queue.

If $p(0,1) = 0$ and $p(0,0) = 1-p$ and

$$p(x,x-1) = q(1-p), \quad p(x,x+1) = p(1-q), \quad p(x,x) = pq + (1-q)(1-p).$$

Thus, p is probability a person enters the line and q is the probability a person leaves (or is served).

Proposition 10. The queue is recurrent if and only if $q \geq p$.

Proof. Let τ_k be the kth time for which $X_n \neq X_{n-1}$. For $x \geq 1$,

$$\begin{aligned}
\mathbb{P}\{X_{\tau_k} = x+1 \mid X_{\tau_{k-1}} = x\} &= \mathbb{P}\{X_{\tau_k} = x+1 \mid X_{\tau_k - 1} = x\} \\
&= \mathbb{P}\{X_1 = x+1 \mid X_0 = x, X_1 \neq X_0\} \\
&= \frac{p(x,x+1)}{1-p(x,x)} \\
&= \frac{p(1-q)}{q(1-p)+p(1-q)}
\end{aligned}$$

and by the same logic,

$$\mathbb{P}\{X_{\tau_k} = x-1 \mid X_{\tau_{k-1}} = x\} = 1 - \frac{p(1-q)}{q(1-p)+p(1-q)}$$

Thus, $\{X_{\tau_k}\}_{k\geq 0}$ is a biased random walk with parameter $\frac{p(1-q)}{q(1-p)+p(1-q)}$. Thus, $\{X_{\tau_k}\}$ hits 0 with probability 1 if and only if $\{X_n\}$ started from $x \geq 1$ hit 0 with probability 1. From last class, happens if and only if

$$\frac{p(1-q)}{q(1-p)+p(1-q)} \leq \frac{1}{2} \iff p(1-q) \leq q(1-p) \iff p \leq q$$

$\square$

Remark 5. By the Brower fixed point theorem, we have the existence of a stationary distribution in finite space spaces.

Recall that if S is finite, then for an irreducible, aperiodic Markov chain, there exists a unique stationary distribution π that satisfies

$$\sum_{x\ inS} \pi_x = 1, \qquad \pi P = \pi \iff \sum_{x\in S} \pi_x p(x,y) = \pi_y.$$

In the countably infinite space, we note that a transient state space cannot yield a stationary distribution since the expected return times are all ∞.

Definition 17. Suppose $\{X_n\}$ is irreducible and recurrent. We say that $\{X_n\}$ is **null recurrent** if

$$\lim_{n\to\infty} p^n(x,y) = 0, \quad \forall\, x,y \in S.$$

Otherwise, $\{X_n\}$ is **positive recurrent**.

Remark 6. Clearly, null recurrent Markov chains also have infinite return times, and thus have no stationary distribution.

Proposition 11. Suppose $\{X_n\}$ is irreducible. The following are equivalent:

(a) $\{X_n\}$ is positive recurrent;

(b) $\{X_n\}$ has a stationary distribution;

(c) For any $x, y \in S$,
$$\limsup_{n\to\infty} p^n(x,y) > 0.$$

(d) If $T_x = \min\{n \geq 1 \mid X_n = x\}$, then
$$\mathbb{E}[T_x \mid X_0 = x] < \infty$$
for any $x \in S$.

Furthermore, if $\{X_n\}$ is aperiodic and positive recurrent, then π is unique and for any $x \in S$,

$$\pi_x = \frac{1}{\mathbb{E}[T_x \mid X_0 = x]}$$

Remark 7. By condition 3, checking that a recurrent Markov chain is null recurrent amount to checking that $\lim_{n\to\infty} p^n(x,y) = 0$ for some states $x, y \in S$.

Example 1.19. Consider the biased random walk on $\{0, 1, \dots\}$ with partially reflecting boundary:

$$p(x, x-1) = q = 1 - p, \quad p(x, x+1) = p, \quad p(0,0) = q, \quad p(0,1) = p.$$

Is this positive recurrent?

We know that $\{X_n\}$ is transient for $p > \frac{1}{2}$. For some π to possibly exist, we need $p < \frac{1}{2}$. The stationary distribution must satisfy:
$$\pi(x) = p\pi(x-1) + q\pi(x+1), \quad \forall x \geq 1.$$

Moreover,
$$\pi(0) = q\pi(0) + q\pi(1).$$

This is because
$$\pi P = \pi \begin{pmatrix} q & p & 0 & 0 & \cdots \\ q & 0 & p & 0 & \cdots \\ 0 & q & 0 & p & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} \pi_0 \\ \pi_1 \\ \vdots \end{pmatrix}$$

For $p < \frac{1}{2}$, the solution takes the form

$$\pi(x) = \lambda_1 + \lambda_2 (\frac{p}{1-p})^x$$

We need
$$\sum_{x=0}^{\infty} \pi(x) = 1 \implies \lambda_1 = 0.$$

After some algebra, we find that $\lambda_1 = 1 - \frac{p}{1-p}$, and thus

$$\pi(x) = (1 - \frac{p}{1-p})(\frac{p}{1-p})^x.$$

Thus, for $p < \frac{1}{2}$, $\{X_n\}$ is positive recurrent.

For $p = \frac{1}{2}$, the general solution is

$$\pi(x) = \lambda_1 + \lambda_2 x.$$

But

$$\sum_{n=0}^{\infty} \pi(x) < \infty \implies \lambda_1 = \lambda_2 = 0,$$

and thus $\{X_n\}$ is null recurrent.

1.9 Friday, Apr 11: Random Walks on $\mathbb{Z}^d$

Consider $\mathbb{Z}^d$ as a graph with edges joining each $x, y \in \mathbb{Z}^d$ such that $\|z - y\|_1 = 1$. Then the random walk $\mathbb{Z}^d$ is a Markov chain moving distance 1 in any of the $2d$ possible directions at each step.

Remark 8. (Stirling's Approximation) For n large,

$$n! \sim \sqrt{2\pi n}(\frac{n}{e})^n$$

Proposition 12. The random walk is recurrent for $d = 1, 2$ and transient for $d \geq 3$

Proof. For $d = 1$, we can graph the random walk by every time we go to the right, we take a step up, and every time we go left, we take a step down.

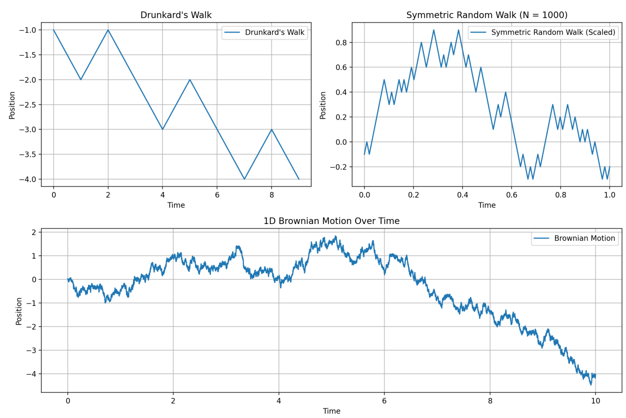


Figure 1: From My REU Paper

We wish to show that $\sum_{n=1}^{\infty} p^n(0,0) = \infty$. Since the parity of n must be even, it suffices to show that $\sum_{n=1}^{\infty} p^{2n}(0,0)$. After a bit of thought, and by using Stirling's Formula above,

$$p^{2n}(0,0) = \mathbb{P}\{X_{2n} = 0 \mid X_0 = 0\} = \frac{\binom{2n}{n}}{2^{2n}} \sim \frac{\sqrt{4\pi n}(\frac{2n}{e})^{2n}}{2\pi n(\frac{n}{e})^{2n}} 2^{-2n} = \frac{1}{\sqrt{\pi n}}.$$

Hence, the infinite series diverges and so the random walk on $\mathbb{Z}$ is recurrent. Indeed, it is null recurrent.

For $d = 2$, we want to do something similar. Fix j. There must be j steps to the right and j steps to the left in order to get back to the origin. Thus, there must be $n - j$ steps up and $n - j$ steps down. Thus, there are

$$\frac{2n!}{j!^2(n-j)!^2}$$

of such combinations.

$$p^{2n}(0,0) = \frac{\sum_{j=0}^{n} \frac{(2n)!}{(j!)^2(n-j)!^2}}{4^{2n}} = \frac{1}{4^{2n}} \frac{(2n)!}{n!^2} \sum_{j=0}^{n} \left(\frac{n!}{j!(n-j)!}\right)^2 = \frac{1}{4^{2n}} \binom{2n}{n} \sum_{j=0}^{n} \binom{n}{j}\binom{n}{n-j} = \frac{1}{4^{2n}} \binom{2n}{n}^2$$

By Stirling's approximation,

$$p^{2n}(0,0) = \frac{1}{4^{2n}} \binom{2n}{n}^2 \approx \frac{1}{\pi n} \implies \sum_{n=0}^{\infty} p^{2n}(0,0) = \infty,$$

and thus the random walk is null recurrent.

For $d \geq 3$, it can be shown that $p^{2n} = \left(\frac{1}{\sqrt{\frac{2\pi n}{d}}} \right)^d$

$\square$

1.10 Monday, Apr 14: Branching Processes

Let $\{X_n\}$ denote the size of the population at time n. Independent of the rest, each individual produces some number of offspring according to an offspring distribution $\{p_k\}$. We remark that $\{p_k\}_{k\geq 0}$ satisfies $p_k \geq 0$ for all k and $\sum_{k=0}^{\infty} p_k = 1$. Informally, p_k represents the probability of an individual producing k children. Then if each n represents a new generation where each previous generation dies of, then

$$X_{n+1} = \sum_{j=1}^{X_n} \xi_j,$$

where the ξ_j are conditionally independent given X_n with

$$\mathbb{P}\{\xi_j = k \mid X_n = x\} = p_k.$$

Thus, ξ_j is how many offspring individual j produces. We are interested in the extinction probability, which is formally denoted by

$$a := \mathbb{P}\{\exists n \geq 1 \text{ s.t. } X_n = 0 \mid X_0 = 1\}.$$

Denote the mean of the offspring distribution of each individual by

$$\mu := \sum_{k=0}^{\infty} kp_k.$$

Note that

$$\mathbb{E}[X_{n+1} \mid X_n = m] = \sum_{j=1}^{m} \mathbb{E}[\xi_j] = \mu m.$$

Then using the law of total expectation,

$$\mathbb{E}[X_{n+1}] = \sum_{m=0}^{\infty} \mathbb{E}[X_{n+1} \mid X_n]\mathbb{P}\{X_n = m\} = \mu \sum_{m=0}^{\infty} m\mathbb{P}\{X_n = m\} = \mu\mathbb{E}[X_n],$$

and so $\mathbb{E}[X_n] = \mu^n \mathbb{E}[X_0]$.

Proposition 13. If $\mu < 1$, then $a = 1$.

Proof. Since $\mathbb{E}[X_n] \leq \mu^n \to 0$ as $n \to \infty$, then $\mathbb{E}[X_n] \to 0$. Consider that by Markov's inequality

$$\mathbb{P}\{X_n \neq 0\} = \mathbb{P}\{X_n \geq 1\} \leq \mathbb{E}[X_n] \to 0,$$

and so $\mathbb{P}\{X_n = 0\} \to 1$ as $n \to \infty$. $\square$

Remark 9. If $X_1 = k$, then by independence, $\mathbb{P}\{\text{extinction} \mid X_1 = k\} = a^k$. Thus,

$$a = \mathbb{P}\{\text{extinction} \mid X_0 = 1\} = \sum_{k=0}^{\infty} \mathbb{P}\{X_1 = k\}\mathbb{P}\{\text{extinction} \mid X_1 = k\} = \sum_{k=0}^{\infty} p_k a^k.$$

Definition 18. Let Y be a random variable taking values in $\{0, 1, 2, \dots\}$. Then the **generating function** for Y is given by

$$\phi : [0, \infty) \to [0, \infty), \quad \phi(s) = \phi_Y(s) = \mathbb{E}[s^y] := \sum_{k=0}^{\infty} \mathbb{P}\{Y = k\}s^k$$

Remark 10. (Basic Properties of ϕ):

(a) We allow $\phi(s) = \infty$, but we have that $\phi(s) < \infty$ for $s \in [0,1]$ since it is bounded above by a converging geometric series.

(b) $\phi(1) = 1$ and $\phi(0) = p_0$.

(c) The derivative has meaning:

$$\phi'(s) = \sum_{k=1}^{\infty} \mathbb{P}\{Y = k\} k s^{k-1} \implies \phi'(1) = \mathbb{E}[Y].$$

(d) If $Y_1, \ldots Y_m$ are independent random variables, then

$$\phi_{Y_1+\cdots+Y_m}(s) = \mathbb{E}[S^{Y_1+\cdots+Y_m}] = \prod_{j=1}^{m} \phi_{Y_j}(s)$$

Proposition 14. Let ϕ be the generating function for p_k. Let $\phi^{(n)} = \phi \circ \phi \cdots \circ \phi$ (n times). Then

$$\phi_{X_n}(s) = \phi^{(n)}(s)$$

Proof. It suffices to show that $\phi_{X_{n+1}}(s) = \phi_{X_n}(\phi(s))$. By definition,

$$\begin{aligned}
\phi_{X_{n+1}}(s) &= \sum_{k=0}^{\infty} \mathbb{P}\{X_{n+1} = k\} s^k \\
&= \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \mathbb{P}\{X_{n+1} = k \mid X_n = j\} \mathbb{P}\{X_n = j\} s^k \\
&= \sum_{j=0}^{\infty} \mathbb{P}\{X_n = j\} \sum_{k=0}^{\infty} \mathbb{P}\{X_{n+1} = k \mid X_n = j\} s^k \\
&= \sum_{j=0}^{\infty} \mathbb{P}\{X_n = j\} \phi(s)^j \\
&= \sum_{j=0}^{\infty} \mathbb{P}\{X_n = j\} \phi(s)^j \\
&= \phi_{X_n}(\phi(s))
\end{aligned}$$

□

1.11 Wednesday, Apr 16: Extinction Probabilities

Recall the generating function

$$\phi_X(s) = \sum_{k=0}^{\infty} p_k s^k$$

Proposition 15. The extinction probability is the smallest positive solution for which $\phi(s) = s$ given that $0 < p_0 < 1$.

Proof. Since the set $\{s : \phi(s) = s\}$ is closed by continuity, then since $p_0 = \phi(0) = 0$, then $0 \notin S$. Thus, the set has a smallest positive element s_0. Since $\phi(1) = 1$, then $s_0 \leq 1$. We claim that

$$\phi_{X_n}(0) < s_0, \quad n \geq 0$$

This is clear for $n = 0$ since $\phi_{X_0}(0) = 0$. If this is true for n, that $\phi_{X_n}(0) < s_0$, then by our proposition 14, and the fact that ϕ is strictly increasing (the derivative is strictly increasing) we have that

$$\phi_{X_{n+1}}(0) = \phi(\phi_{X_n}(0)) < \phi(s_0) = s_0.$$

Note that since $\phi_{X_n}(0) = \mathbb{P}\{X_n = 0\}$, then

$$a = \lim_{n\to\infty} \phi_{X_n}(0) \leq s_0.$$

Since $\phi(a) = a$, then $a = s_0$. □

Proposition 16. If $\mu > 1$, then $a < 1$. If $\mu \leq 1$ and $p_0 \neq 0$, then $a = 1$. In particular, the population will die with probability 1 if, and only if, $\mu \leq 1$ and $p_0 \neq 0$.

Proof. Suppose $\mu = 1$ and $p_0 \neq 0$. Note that

$$\phi''(s) = \sum_{k=2}^{\infty} k(k-1)p_k s^{k-2}.$$

If $p_0 \neq 0$ and $\mu = 1$, then we claim that $p_k > 0$ for some $k \geq 2$. If not, then $\mu = p_1 < 1$. Then $\phi''(s) > 0$. Recall that $\phi(1) = 1$ and $\phi'(1) = \mu$. Thus, $\phi'(s) < \mu$ for all $s \in [0, 1)$. If $\mu = 1$, then

$$1 - \phi(s) = \int_s^1 \phi'(t)dt < \int_s^1 \mu dt = 1 - s.$$

Thus, $\phi(s) > s$ for all $s \in [0, 1)$. Thus, $a = 1$.

If $\mu > 1$, then $\phi'(1) > 1$, $\phi(1) = 1$. Then there exists $s < 1$ such that $\phi(s) < s$. Since $\phi(0) > 0$, the continuity of $t \mapsto \phi(t) - t$ implies by the IVT there is a $t \in (0, 1)$ such that $\phi(t) = t$. Hence, $a < 1$. □

Definition 19. We say that a branching process is **subcritical** if $\mu < 1$, **critical** if $\mu = 1$, and **supercritical** if $\mu > 1$.

Example 1.20. Suppose $p_0 = \frac{1}{10}$, $p_1 = \frac{3}{5}$, and $p_2 = \frac{3}{10}$, then

$$\mu = \frac{1}{10} + \frac{3}{5} + \frac{6}{10} = \frac{12}{10} > 1.$$

Note that

$$\phi(s) = \sum_{k=0}^{3} p_k s^k = \frac{1}{10} + \frac{3}{5}s + \frac{3}{10}s^2.$$

We want the smallest positive solution to $\phi(s) = s$. Solving gives $a = \frac{1}{3}$.

Example 1.21. Bacteria reproduce by cell division. In one unit of time, bacterium will either die (w.p. $\frac{1}{4}$,) stay the same (w.p. $\frac{1}{4}$), or split in 2 (w.p. $\frac{1}{2}$). At time $n = 0$, the population starts with 100 bacteria.

We can interpret this as the following: If a bacteria dies, then it produces three offspring? Then

$$\phi(s) = \sum_{k=0}^{\infty} p_k s^k,$$

where

$$p_1 = \frac{1}{4}, p_2 = \frac{1}{4}, p_3 = \frac{1}{2}$$

$$\phi(s) = \frac{1}{4} + \frac{1}{4}s + \frac{1}{2}s^2, \quad Y_1, \ldots, Y_{100}, \quad \phi_{Y_1 + \cdots Y_{100}} = (\tilde{\phi})^{100}.$$

Then

$$\mu = \frac{1}{4} + 1 > 1.$$

Thus, we need to solve

$$\phi(s) - s = 0 \implies s = \frac{1}{2}$$

for each one. Thus, by independence,

$$a = \frac{1}{2^{100}}$$

1.12 Friday, Apr 18: Poisson Processes

Definition 20. A random variable Y is **Poisson** with parameter λ if for all $k \geq 0$, we have that

$$\mathbb{P}\{Y = k\} = e^{-\lambda}\frac{\lambda^k}{k!}$$

Remark 11. Recall that $\mathbb{E}[Y] = \mathrm{Var}[Y] = \lambda$.

Proposition 17. Suppose Y_1, Y_2 are independent Poisson random variables with parameters λ_1, λ_2, respectively. Then $Y_1 + Y_2 \sim \mathrm{Poisson}(\lambda_1 + \lambda_2)$.

We motivate the Poisson process with an example:

Example 1.22. Suppose that in a phone line, calls occur at the same rate λ at all hours in a day. The number of calls during disjoint time intervals are independent. Let X_t be the number of calls at (or before) time t.

Definition 21. The **Poisson process** with rate λ is the continuous time stochastic process $\{X_t\}_{t\geq 0}$ such that:

- $X_0 = 0$;
- For any times $0 \leq s_1 \leq t_1 \leq s_2 \leq \cdots \leq s_k \leq t_k$, the random increments

  $$X_{t_j} - X_{s_j} \sim \mathrm{Poisson}(\lambda(t_j - s_j))$$

  are independent

Definition 22. A random variable $T \in [0, \infty)$ has the exponential distribution with parameter λ if

$$\mathbb{P}\{T \geq t\} = e^{-\lambda}$$

Recall that $\mathbb{E}[T] = \frac{1}{\lambda}$ and $\mathrm{Var}[T] = \frac{1}{\lambda^2}$.

Remark 12. We could alternatively describe a Poisson process in terms of arrival times. Let $\tau_0 = 0$ and let $\tau_j = \inf\{t \geq 0 \mid X_t = j\}$ be the time of the jth call. Then $X_t = \max\{j \mid \tau_j \leq t\}$ is the number of calls before time t.

Proposition 18. The arrival times $\tau_j - \tau_{j-1}$ for $j = 1, 2, 3, \ldots$ are i.i.d. and each has exponential distribution with parameter λ.

Proof. Note that $\tau_0 = 0$, and

$$\mathbb{P}\{\tau_1 > t\} = \mathbb{P}\{X_t = 0\} = \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda}.$$

Then $\tau_1 \sim \mathrm{Exp}(\lambda)$. Note that τ_1 is a stopping time for $\{X_t\}$ and $\{T_{s+t} - X_t\}_{s\geq 0}$ is a Poisson process independent of $\{X_s\}_{s\leq t}$. By the strong Markov property,

$$\{X_{s+\tau_1} - X_{\tau_1}\}_{s\geq 0}$$

is Poisson and independent of $\{X_s\}_{s\leq \tau_1}$. In particular, $\tau_2 - \tau_1$ is independent of τ_1 and has the same distribution. $\square$

Example 1.23. Suppose $\{X_t\}_{t\geq 0}$ is a Poisson process with rate λ. Then

$$\mathbb{E}[X_2 \mid X_1] = X_1 + \lambda$$

1.13 Monday, Apr 21: Continuous Time Markov Process

Proposition 19. (Minimum Property) Suppose $T_1, \dots, T_n$ are independent exponential r.v. with parameters $\lambda_1, \dots, \lambda_n$ (respectively). Then $\min\{T_1, \dots, T_n\}$ is an exponential random variable with parameter $\lambda_1 + \cdots + \lambda_n$. Moreover, for any $j = 1, 2, 3 \dots, n$, we have that

$$\mathbb{P}\{T_j = \min\{T_1, \dots, T_n\}\} = \frac{\lambda_j}{\lambda_1 + \cdots + \lambda_n}$$

Proof. We have that

$$\begin{aligned}
\mathbb{P}\{T_{(1)} \geq t\} &= \mathbb{P}\{T_1, \geq t, T_2 \geq t, \dots, T_n \geq t\} \\
&= \mathbb{P}\{T_1 \geq t\} \cdots \mathbb{P}\{T_n \geq t\} \\
&= e^{-\lambda_1 t} \cdots e^{-\lambda_n t} \\
&= e^{-(\lambda_1 + \cdots + \lambda_n)t}.
\end{aligned}$$

For the second claim, we have that if $j = 1$, then

$$\begin{aligned}
\mathbb{P}\{T_1 = T_{(1)}\} &= \mathbb{P}\{T_j \geq T_1,\ \forall j = 1, 2, \dots, n\} \\
&= \mathbb{E}[\mathbb{P}\{T_j \geq T_1,\ \forall j = 1, 2, \dots, n \mid T_1\}] \\
&= \mathbb{E}[e^{-(\lambda_1 + \dots, \lambda_n)T_1}] \\
&= \lambda_1 \int_0^\infty e^{-(\lambda_1 + \dots, \lambda_n)t} dt \\
&= \frac{\lambda_1}{\sum_{j=1}^n \lambda_j}
\end{aligned}$$

□

Suppose S is a finite state space. We wish to define a stochastic process $\{X_t\}_{t \geq 0}$ taking values in x. For distinct $x, y \in S$, there exists some rates $\alpha(x, y) \geq 0$ that specify how frequently we jump from x to y.

Remark 13. For $y \in S$ with $\alpha(x, y) \neq 0$, let T_1 be an exponential r.v. with parameter $\alpha(x, y)$. Let $T = \min_y T_y$. By the minimum property, $T \sim \exp(\alpha(x))$, where $\alpha(x) = \sum_{y \in S\{x\}} \alpha(x, y)$. Moreover, we have that

$$\mathbb{P}\{X_T = y\} = \frac{\alpha(x, y)}{\alpha(x)}$$

Proposition 20. (The Markov Property) Let $t \geq 0$ and $x \in S$. The conditional distribution of $\{X_{t+s}\}_{s \geq 0}$ given $\{X_t = x\}$ (and everything before time t), is the same as the distribution of $\{X_s\}$ started from x.

Now, we aim to write down a transition matrix for this process.

Definition 23. Define the $t-$**time transition probabilities** to be

$$p_t(x, y) = \mathbb{P}\{X_t = y \mid X_0 = x\}$$

and label $S = \{1, 2, \dots, N\}$ and define the $t-$**time transition matrix** as

$$P_t = \big(p_t(x, y)\big)_{x, y = 1, 2, \dots, N}$$

Remark 14. For $x \neq y$, we know that by the Markov Property,

$$\mathbb{P}\{X_{t+\epsilon} = y \mid X_t = x\} = \mathbb{P}\{X_\epsilon = y \mid X_0 = x\} \approx \mathbb{P}\{T_y < \epsilon\} = 1 - e^{\alpha(x,y)\epsilon} = \alpha(x,y)\epsilon + O(\epsilon^2).$$

Thus,

$$\mathbb{P}\{X_{t+\epsilon} \neq x \mid X_t = x\} = \sum_{y\{x\}} \alpha(x,y)\epsilon + O(\epsilon^2) = \alpha(x)\epsilon + O(\epsilon^2).$$

Informally,

$$\frac{d}{dt}p_t(x,y) = \lim_{\epsilon\to 0} \frac{1}{\epsilon}\left(\mathbb{P}\{X_{t+\epsilon} = y \mid X_0 = x\} - \mathbb{P}\{X_t = y \mid X_0 = x\}\right) = \cdots = \sum_{z\in S\{y\}} \alpha(x,y)p_t(x,z) - \alpha(y)p_t(x,y)$$

In matrix form,

$$\frac{d}{dt}P_t = P_t A,$$

where A is the $N \times N$ whose x, y entry is $\alpha(x,y)$ if $x \neq y$ and $-\alpha(x)$ if $x = y$. Note that $p_0(x,x) = 1$ and $p_0(x,y) = 0$. Thus, $P_0 = I$. The solution to this ODE gives

$$P_t = e^{tA}$$

1.14 Wednesday, Apr 23: Midterm

1.15 Friday, Apr 25: $t-$Time Transition Matrix

Remark 15. Since $P_t = e^{tA}$, we have that

$$e^{tA} = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!}.$$

Thus, if $A = QDQ^{-1}$ and so

$$e^{tA} = Qe^{tD}Q^{-1}.$$

Note that if D is diagonal with entries $\lambda_1, \ldots, \lambda_m$, then e^{tD} is the diagonal matrix with diagonal entries $e^{t\lambda_1}, \ldots, e^{t\lambda_m}$.

The matrix A is called the **infinitesimal generator** for the Markov Chain.

Example 1.24. Consider the Markov Chain with state space $S = \{0, 1\}$ and rates

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

Then

$$A = \begin{pmatrix} -2 & 2 \\ 3 & -3 \end{pmatrix}$$

Then

$$A - \lambda I = \begin{pmatrix} -2-\lambda & 2 \\ 3 & -3-\lambda \end{pmatrix}$$

Then

$$|A - \lambda I| = (-2-\lambda)(-3-\lambda) - 6 \implies \lambda_1 = -5, \lambda_2 = 0.$$

Then finding the null space E_{-5} is

$$E_{-5} = \{v \mid (A - \lambda_{-5})v = 0\}$$

which is formed by $v_5 = (-2, 3)$. Similarly, $v_0 = (1, 1)$. Thus,

$$D = \begin{pmatrix} -5 & 0 \\ 0 & 0 \end{pmatrix} \quad Q = \begin{pmatrix} -2 & 1 \\ -3 & 1 \end{pmatrix} \quad Q^{-1} =$$

and so

$$P_t = QDQ^{-1} = \begin{pmatrix} -2 & 1 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} e^{-5t} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ -3 & 1 \end{pmatrix}^{-1} = \frac{1}{5} \begin{pmatrix} 3+2e^{-5t} & 2-2e^{-5t} \\ 3-3e^{-5t} & 2+3e^{-5t} \end{pmatrix}$$

Then

$$\mathbb{P}\{X_1 = 1 \mid X_0 = 0\} = \frac{2-2e^{-5t}}{5}$$

Definition 24. A continuous time Markov Chain is **irreducible** if $p_t(x, y) > 0$ for all $x, y \in S$. and for all $t > 0$.

Definition 25. A function $\pi : S \to [0, 1]$ with $\sum_{x \in S} \pi_x = 1$ is an **invariant/stationary distribution** for $\{X_t\}$ if the following is true:

- If X_0 has distribution π, then X_t has distribution π for all $t \geq 0$.
- It satisfies

$$\frac{d}{dt}\pi P_t = 0 \iff \frac{d}{dt}\pi e^{tA} = 0 \iff \pi A = 0.$$

Proposition 21. If $\{X_t\}$ is irreducible, then there exists a unique stationary distribution π for $\{X_t\}$ which satisfies, for any $x, y \in S$,

$$\lim_{t\to\infty} p_t(x, y) = \pi_y$$

Example 1.25. It is easy to see that in the example above,

$$\pi = \begin{pmatrix} \frac{3}{5} & \frac{2}{5} \end{pmatrix}$$

1.16 Monday, Apr 28: Conditional Expectation

Remark 16. Suppose X and Y are r.v. taking values in countable sets $S, T \subset \mathbb{R}$. For $x \in S$ and $y \in T$, let

$$f(x, y) = \mathbb{P}\{X = x, Y = y\}.$$

Then the *conditional expectation* is

$$\mathbb{E}[X \mid Y = y] = \sum_{x \in S} x\mathbb{P}\{X = x \mid Y = y\} = \frac{\sum_{x \in S} sf(x, y)}{\sum_{x \in S} f(x, y)}.$$

Moreover, we define the random variable

$$\mathbb{E}[X \mid Y] = \frac{\sum_{x \in S} xf(x, y)}{\sum_{x \in S} f(x, y)}$$

Definition 26. Let X, Y be random variables with X taking values in $\mathbb{R}$ and with $\mathbb{E}[|X|] < \infty$. The **conditional expectation** $\mathbb{E}[X \mid Y]$ is the unique random variable that satisfies the following:

(a) $\mathbb{E}[X \mid Y]$ is a function of Y.

(b) If $F(y)$ is a function of y taking values in $\mathbb{R}$ and $\mathbb{E}[|F(y)|] < \infty$, then

$$\mathbb{E}[X\, F(Y)] = \mathbb{E}[\mathbb{E}[X \mid Y]F(Y)]$$

Suppose $\mathbb{E}[|X|] < \infty$ for any random variable talked about in this class.

Definition 27. Let $\{X_t\}_{tT}$ be a stochastic process. We say that the **natural filtration** of X_t is the sigma algebra generated by Y_t. That is, $\mathscr{F}_t = \sigma(Y_1, Y_2, ..., Y_t)$.

We note that the natural filtration of $X_0, \ldots, X_n$ is the information contained in these random variables.

Remark 17. Moreover,

$$\mathbb{E}[X \mid \mathscr{F}_n] = \mathbb{E}[Y_0, \ldots, Y_n]$$

and we say that X is $\mathscr{F}_n-$measurable if it is a function of $Y_0, \ldots, Y_n$

Proposition 22. $\mathbb{E}[X \mid \mathscr{F}_n]$ is the unique real valued random variable satisfying the following:

(a) $\mathbb{E}[X \mid \mathscr{F}_n]$ is $\mathscr{F}_n$ measurable.

(b) $\mathbb{E}[XZ] = \mathbb{E}[\mathbb{E}[X \mid \mathscr{F}_n]Z]$

Proposition 23. (Linearity) Suppose X_1, X_2 are real valued r.v.s and $a, b \in \mathbb{R}$. Then

$$\mathbb{E}[aX_1 + bX_2 \mid \mathscr{F}_n] = a\mathbb{E}[X_1 \mid \mathscr{F}_n] + b\mathbb{E}[X_2|\mathscr{F}_n]$$

Proof. If Z is $\mathscr{F}_n-$measurable, then $\mathbb{E}[(aX_1 + bX_2)Z] = \mathbb{E}[\mathbb{E}[aX_1 + bX_2 \mid \mathscr{F}_n]Z]$. Let

$$W := a\mathbb{E}[X_1 \mid \mathscr{F}_n] + b\mathbb{E}[X_2 \mid \mathscr{F}_n]$$

Then

$$\begin{aligned}
\mathbb{E}[WZ] &= \mathbb{E}[(a\mathbb{E}[X_1 \mid \mathscr{F}_n] + b\mathbb{E}[X_2 \mid \mathscr{F}_n])Z] \\
&= \mathbb{E}[a\mathbb{E}[X_1 \mid \mathscr{F}_n]Z] + \mathbb{E}[b\mathbb{E}[X_2 \mid \mathscr{F}_n]Z] \\
&= a\mathbb{E}[X_1 Z] + b\mathbb{E}[X_2 Z] \\
&= \mathbb{E}[(aX_1 + bX_2)Z].
\end{aligned}$$

Because conditional expectation is unique, we are done. □

Proposition 24. Suppose X is $\mathscr{F}_n-$measurable, then $\mathbb{E}[X \mid \mathscr{F}_n] = X$.

Proposition 25. If X is independent from $\mathscr{F}_n$, then

$$\mathbb{E}[X \mid \mathscr{F}_n] = \mathbb{E}[X]$$

Proposition 26. If Z is $\mathscr{F}_n-$measurable, then $\mathbb{E}[ZX \mid \mathscr{F}_n] = Z\mathbb{E}[X \mid \mathscr{F}_n]$.

Proposition 27. If $m \leq n$, then

$$\mathbb{E}[\mathbb{E}[X \mid \mathscr{F}_n] \mid \mathscr{F}_m] = \mathbb{E}[X \mid \mathscr{F}_m]$$

Example 1.26. Let $X_1, X_2, \ldots$ be i.i.d r.vs with mean μ. Then if $m \geq n$,

$$\mathbb{E}[X_1 + \cdots + X_n \mid \mathscr{F}_m] = X_1 + \cdots + X_n.$$

If $m < n$, then

$$\begin{aligned}
\mathbb{E}[X_1 + \cdots + X_n \mid \mathscr{F}_m] &= \mathbb{E}[X_1 + \cdots + X_m \mid \mathscr{F}_m] + \mathbb{E}[X_{m+1} + \cdots + X_n \mid \mathscr{F}_m] \\
&= X_1 + \cdots + X_m + \mathbb{E}[X_{m+1} \mid \mathscr{F}_m] + \cdots + \mathbb{E}[X_n \mid \mathscr{F}_m] \\
&= X_1 + \cdots + X_m + \mathbb{E}[X_{m+1}] + \cdots + \mathbb{E}[X_n] \\
&= X_1 + \cdots + X_m + (n-m)\mu
\end{aligned}$$

Example 1.27. Let $X_1, \ldots, X_n$ be i.i.d. with mean 0 and variance σ^2. Then

$$\mathbb{E}[(X_1 + \cdots + X_n)^2 \mid \mathscr{F}_m] = S_m^2 + (n-m)\sigma^2.$$

1.17 Wednesday, Apr 30: Martingales

Definition 28. Let $\{X_t\}$ be a stochastic process. The **natural filtration** of X_t is defined to be

$$\mathscr{F}_n = (X_1, \dots, X_n)$$

Definition 29. A stochastic process $\{M_n\}$ is called a martingale w.r.t. $\{\mathscr{F}_n\}$ if:

(a) Each M_n is $\mathscr{F}_n$ measurable.

(b) $\mathbb{E}[|M_n|] < \infty$ for all n.

(c) For each $m < n$,
$$\mathbb{E}[M_n \mid \mathscr{F}_m] = M_m.$$

Remark 18. In particular, we can rewrite the Martingale Property to be $\mathbb{E}[M_n - M_m \mid \mathscr{F}_m] = 0$. By the tower property,

$$\mathbb{E}[M_n] = \mathbb{E}[\mathbb{E}[M_n \mid \mathscr{F}_0]] = \mathbb{E}[M_0].$$

Proposition 28. To prove $\mathbb{E}[M_n \mid F_m] = M_m$, it suffices to show that $\mathbb{E}[M_n \mid \mathscr{F}_{n-1}] = M_{n-1}$

Proof.
$$\mathbb{E}[M_n \mid \mathscr{F}_{n-2}] = \mathbb{E}[\mathbb{E}[M_n \mid \mathscr{F}_{n-1}] \mid \mathscr{F}_{n-2}] = \mathbb{E}[M_{n-1} \mid \mathscr{F}_{n-2}] = M_{n-2}$$
□

Example 1.28. Let $X_1, X_2, \dots$ be i.id. with mean μ, and let $\mathscr{F}_n$ be the natural filtration of X_n. Let $S_0 = 0$ and $S_n = X_1 + \cdots + X_n$. Let $M_n = S_n - n\mu$.

(a) The first property is inmediate.

(b) For any n,
$$\mathbb{E}[|M_n|] \le \mathbb{E}[|S_n|] + \mu n < \infty.$$

(c) For any n,
$$\begin{aligned}\mathbb{E}[M_n \mid \mathscr{F}_{n-1}] &= \mathbb{E}[(X_1 + \cdots + X_n) - n\mu \mid \mathscr{F}_{n-1}] \\ &= \mathbb{E}[M_{n-1} + (X_n - \mu) \mid \mathscr{F}_{n-1}] \\ &= \mathbb{E}[M_{n-1} \mid \mathscr{F}_{n-1}] + \mathbb{E}[X_n - \mu \mid \mathscr{F}_{n-1}] \\ &= M_{n-1}\end{aligned}$$

Example 1.29. Let $X_1, X_2, \dots$ be i.i.d. with mean 0 and variance σ^2. Let $\mathscr{F}_n$ be the natural filtration of X_n. Let $M_n := S_n^2 - n\sigma^2$.

(a) Again, the first property is clear.

(b) For any n,
$$\mathbb{E}[|M_n|] \le \mathbb{E}[S_n^2] + n\sigma^2 = \mathrm{Var}[S_n] - \mathbb{E}[S_n]^2 + n\sigma^2 = n\sigma^2 + n\sigma^2 < \infty$$

(c) For any n, we have (from Example 1.27)
$$\begin{aligned}\mathbb{E}[M_n \mid \mathscr{F}_{n-1}] &= \mathbb{E}[S_n^2 - n\sigma^2 \mid \mathscr{F}_{n-1}] \\ &= \mathbb{E}[S_n^2 \mid \mathscr{F}_{n-1}] - n\sigma^2 \\ &= S_{n-1}^2 + \sigma^2 - \sigma^2 n \\ &= M_{n-1}\end{aligned}$$

Example 1.30. Let $Y_0, Y_1, \dots$ be a sequence of random variables and let $\mathscr{F}_n$ be the natural filtration. Let X be a random variable such that $\mathbb{E}[|X|] < \infty$, and let

$$M_n = \mathbb{E}[X \mid \mathscr{F}_n]$$

(a) Clealry, M_n is $\mathscr{F}_n$ measurable.

(b) For any n,
$$\mathbb{E}[|M_n|] \leq \mathbb{E}[\mathbb{E}[|X| \mid \mathscr{F}_n]] = \mathbb{E}[|X|] < \infty.$$

(c) By the tower property,

$$\begin{aligned} \mathbb{E}[M_n \mid \mathscr{F}_{n-1}] &= \mathbb{E}[\mathbb{E}[X \mid \mathscr{F}_n] \mid \mathscr{F}_{n-1}] \\ &= \mathbb{E}[X \mid \mathscr{F}_{n-1}] \\ &= M_{n-1} \end{aligned}$$

We know $\mathbb{E}[M_n] = \mathbb{E}[M_0]$ for any n. Is this still true for a random time τ.

Remark 19. Let τ be a stopping time for $\{\mathscr{F}_n\}$. Doob's Optional Stopping Theorem states that, under certain conditions,
$$\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$$

Example 1.31. (Gambler's ruin) Let $X_1, X_2, \dots$ be i.i.d with $\mathbb{P}\{X_i = 1\} = \mathbb{P}\{X_i = -1\} = \frac{1}{2}$. Then $S_n = X_1 + \dots X_n$ is a martingale by Example 1.28. Let $\tau := \min\{n \geq 0 \mid S_n = a \text{ or } S_n = -b\}$. DOST tells us that
$$\mathbb{E}[S_\tau] = \mathbb{E}[S_0] = 0,$$

and thus

$$\begin{aligned} 0 &= \mathbb{E}[S_\tau] \\ &= a\mathbb{P}\{S_n = a\} + (-b)\mathbb{P}\{S_n = -b\} \\ &= ap_W - bp_L \\ &= ap_W - b(1 - p_W) \end{aligned}$$

and thus
$$p_W = \frac{b}{a+b}$$

1.18 Friday, May 2: Optional Stopping Theorem

Theorem 5. (OST I) Suppose $\{M_n\}$ is a Martingale with respect to some filtration $\{\mathscr{F}_n\}$, and let τ be a stopping time such that $\mathbb{P}\{\tau < K\} = 1$ for some constant K. Then $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$.

Proof. Note that $M_\tau = \sum_{n=0}^{K} M_n \mathbb{1}_{\tau=n}$. Then conditioning,

$$\begin{aligned}
\mathbb{E}[M_\tau \mid \mathscr{F}_{K-1}] &= \mathbb{E}\left[\sum_{n=0}^{K} M_n \mathbb{1}_{\tau=n} \mid \mathscr{F}_{K-1}\right] \\
&= \sum_{n=0}^{K} \mathbb{E}\left[M_n \mathbb{1}_{\tau=n} \mid \mathscr{F}_{K-1}\right] \\
&= \sum_{n=0}^{K-1} \mathbb{E}\left[M_n \mathbb{1}_{\tau=n} \mid \mathscr{F}_{K-1}\right] + \mathbb{E}\left[M_K \mathbb{1}_{\tau=K} \mid \mathscr{F}_{K-1}\right] \\
&= \sum_{n=0}^{K-1} M_n \mathbb{1}_{\tau=n} + \mathbb{E}[\mathbb{1}_{\tau=K}]\mathbb{E}[M_K \mid \mathscr{F}_{K-1}] \\
&= \sum_{n=0}^{K-1} M_n \mathbb{1}_{\tau=n} + \mathbb{E}[\mathbb{1}_{\tau>K}] M_{K-1} \\
&= \sum_{0}^{K-2} M_n \mathbb{1}_{\tau=n} + M_{K-1} \mathbb{1}_{K>K-2}
\end{aligned}$$

Similarly,

$$\mathbb{E}[M_\tau \mid \mathscr{F}_{K-2}] = \sum_{0}^{K-3} M_n \mathbb{1}_{\tau=n} + M_{K-3} \mathbb{1}_{K>K-3}$$

and so on until

$$\mathbb{E}[M_\tau \mid \mathscr{F}_0] = M_0 \mathbb{1}_{\tau\geq 1} = M_0.$$

Using the tower property,

$$\mathbb{E}[M_\tau] = \mathbb{E}[\mathbb{E}[M_\tau \mid \mathscr{F}_0]] = \mathbb{E}[M_0]$$

□

Remark 20. Let $\{X_n\}$ be a simple random walk started at 0. We know X_n is a martingale by Example 1.28. Let $\tau := \min\{n \geq 1 \mid X_n = 1\}$. Then $\mathbb{P}\{\tau < \infty\} = 1$, but τ is not bounded. Note that $\mathbb{E}[X_\tau] = 1 \neq 0 = \mathbb{E}[X_0]$.

Recall that $\tau \wedge n = \min\{\tau, n\}$. Since $\tau \wedge n$ is a bounded stopping time, then by the optional stopping theorem,

$$\mathbb{E}[M_0] = \mathbb{E}[M_{\tau\wedge n}] = \mathbb{E}[M_\tau \mathbb{1}_{\tau\leq n}] + \mathbb{E}[M_n \mathbb{1}_{\tau>n}].$$

We would like to show that in the limit, the second term dies.

Since τ is finite, then $1_{\tau>n} \to 0$ and $\mathbb{1}_{\tau\leq n} \to 1$ and so $M_{\tau\wedge n} \to M_\tau$. If $\mathbb{E}[|M_\tau|] < \infty$ and $\mathbb{P}\{\tau < \infty\} = 1$, then we apply the Dominated convergence theorem with $X_n = M_\tau \mathbb{1}_{\tau\leq n}$ and $Y = M_\tau$. We know Y dominates the X_n and that $\mathbb{1}_{\tau\leq n} \to 1$ and by the DCT,

$$\mathbb{E}[M_\tau \mathbb{1}_{\tau\leq n}] \to \mathbb{E}[M_\tau].$$

Similarly, $\mathbb{E}[M_n \mathbb{1}_{\tau>n}] \to 0$

This remark leads to the following result:

Theorem 6. (OST II) Let τ be a stopping time and assume that

(a) $\mathbb{P}\{\tau < \infty\} = 1$.

(b) $\mathbb{E}[|M_\tau|] < \infty$

(c) $\mathbb{E}[M_n \mathbb{1}_{\tau > n}] \to 0$.

Then $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$.

Remark 21. If M_n is bounded, i.e, there exists a c such that $|M_n| < c$ for all n, then by Jensen's inequality

$$\mathbb{E}[M_n \mathbb{1}_{\tau > n}] \leq c\mathbb{P}\{\tau > n\} \to 0$$

Example 1.32. We are going to formally provide the solution to the Gambler's ruin model in example 1.31. Let X_n be the martingale modeling the random walk started at 0, and $\tau := \min\{n \geq 0 \mid X_n = a \text{ or } S_n = b\}$. Since the random walk is recurrent, we have that $\mathbb{P}\{\tau < \infty\} = 1$. Also, we have that $\mathbb{E}[|X_\tau|] < \infty$ since X_τ is either a or b. Then

$$\mathbb{E}[|X_n| \mathbb{1}_{\tau > n}] \leq \max\{a, b\}\mathbb{P}\{\tau > n\} \to 0.$$

Now and only now can we proceed as in Example 1.31.

1.19 Monday, May 5: Applications of the OST

Example 1.33. Let $\{X_n\}$ be a symmetric random walk on $\mathbb{Z}$, and let $\tau := \min\{n > 0 : X_n = a \text{ or } X_n = b \mid X_0 = 0\}$. Then let $M_n := X_n^2 - n$. Recall from Example 1.29 that M_n is a martingale (noting that $\sigma^2 = 1$) with respect to $\mathscr{F}_n$.

We claim that $\mathbb{P}\{\tau > n\} \leq e^{-cn}$ for some $c > 0$. To see this, note that the random walk stopped when we hit a or $-b$ is a Markov chain on $S = \{-b, -b+1, \dots, a-1, a\}$ with absorbing ends. Then $\{-b+1, \dots, a-1\}$ is a transient communication class. It is a simple consequence of Proposition 5 that the probability that we stay in $\{-b+1, \dots, a-1\}$ is bounded above by e^{-cn} for some $c > 0$. Using this bound, we have that

$$\mathbb{E}[\tau] = \sum_{n=0}^{\infty} \mathbb{P}\{\tau \geq n\} \leq \sum_{n=0}^{\infty} e^{-cn} < \infty$$

and thus

$$\mathbb{E}[|X_\tau|] \leq \mathbb{E}[X_\tau^2] + \mathbb{E}[\tau] \leq \max\{a^2, b^2\} + \mathbb{E}[\tau] < \infty$$

Moreover,

$$\mathbb{E}[|M_n| \mathbb{1}_{\tau > n}] \leq (\max\{a, b\}^2 + n) e^{-cn} \to 0.$$

We can now apply the optional stopping theorem (ii), we have that

$$0 = \mathbb{E}[M_0] = \mathbb{E}[M_\tau] = \mathbb{E}[X_\tau^2] - \mathbb{E}[\tau] = \mathbb{P}\{X_\tau = a\}a^2 + \mathbb{P}\{X_\tau = -b\}b^2 - \mathbb{E}[\tau].$$

From Example 1.31,

$$0 = \frac{a^2 b + b^2 a}{a+b} - \mathbb{E}[T] = ab - \mathbb{E}[T] \implies \mathbb{E}[T] = ab.$$

Example 1.34. At each toss of a coin, you win \$1 if T and lose a buck if H. Let X_n be the winnings on round n. Then

$$\mathbb{P}\{X_n = 1\} = \mathbb{P}\{X_n = -1\} = \frac{1}{2}.$$

Then if $Y_n = \sum_{i=1}^n X_i$, we have shown in Example 1.28 that Y_n is a martingale. By OST, we know that if $\tau := \min\{n > 0 \mid X_n = T\}$, then

$$0 = \mathbb{E}[Y_0] = \mathbb{E}[Y_\tau] = \mathbb{E}[\sum_{n=0}^{\tau} X_n] = \mathbb{E}[(-1)(\tau - 1) + 1] \implies \mathbb{E}[\tau] = 2.$$

Note that if the coin were biased with $\mathbb{P}\{H\} = p$, then for every \$1 bet you make, you are compensate $\$(1-2p)$.

Example 1.35. I am a monkey at a typewriter with only capital letters. How long will it take for me, the monkey, to type out

ABRACADABRA

Some rich assholes bet on my typing such that just before time n, a new gambler arrives and bets \$1 that the nth letter will be A. If he loses, he dies. If he wins, he wins \$26 and bets all of it that the $n+1$th letter will be B. Repeat with the rest of the letters in ABRACADABRA. Denote the winnings of the jth better out of n after n rounds by M_n^j. This is a martingale since

$$\mathbb{E}[M_n^j \mid \mathscr{F}_{n-1}] = \begin{cases} 0, & \text{if he loses sometime before } n \\ 26^n \frac{1}{26} + 0 \frac{25}{26} = 26^{n-1} & \end{cases} = M_{n-1}^j.$$

Let $M_n = \sum_{j=1}^{n} M_n^j$ be total winnings on all the betters before round $n+1$. We will see on a PSET that M_n is a martingale and that we can apply the OST. Thus, if $\tau = \min\{n > 0 \mid \text{ we see ABRACADABRA}\}$, then

$$\mathbb{E}[M_0] = 0 = \mathbb{E}[M_\tau] = \mathbb{E}[\sum_{j=1}^{\tau} M_n^j] = \mathbb{E}[(26^{11} - 1) + (26^4) + (26 - 1) + (-1)(\tau - 3).]$$

Thus,

$$\mathbb{E}[\tau] = 26^{11} + 26^4 + 26$$

1.20 Wednesday, May 7: Martingale Convergence Theorem

Example 1.36. Let $\mathscr{F}_n$ be a filtration, and X be a r.v. with $\mathbb{E}[|X|] < \infty$. Define

$$M_n := \mathbb{E}[X \mid \mathscr{F}_n].$$

We've shown that M_n is a martingale. The MCT will tell us that M_n converges as $n \to \infty$

Proposition 29. Let M_n be a martingale with respect to $\mathscr{F}_n$, let τ be a stopping time. Then $M_{n\wedge\tau}$ is a Martingale.

Proof.

- To show $M_{n\wedge\tau}$ is $\mathscr{F}_n$ measurable, noe that

$$M_{n\wedge\tau} = M_n \mathbb{1}_{n<\tau} + \sum_{k=0}^{n} M_k \mathbb{1}_{\tau=k}$$

is clearly $\mathscr{F}_n$ measurable.

- We have

$$|M_{n\wedge t}| \leq |M_n| + \sum_{k=0}^{n} |M_k|,$$

and the result follows from taking expectation.

-

$$\begin{aligned}
\mathbb{E}[M_{n\wedge\tau} \mid \mathscr{F}_{n-1}] &= \mathbb{E}[M_{n\wedge\tau}\mathbb{1}_{\tau>n} \mid \mathscr{F}_{n-1}] + \mathbb{E}[M_n \wedge \tau \mathbb{1}_{\tau\leq n} \mid \mathscr{F}_{n-1}] \\
&= \mathbb{E}[M_n \mathbb{1}_{\tau>n} \mid \mathscr{F}_{n-1}] + \mathbb{E}[M_\tau \mathbb{1}_{\tau\leq n} \mid \mathscr{F}_{n-1}] \\
&= \mathbb{E}[M_n \mid \mathscr{F}_{n-1}]\mathbb{1}_{\tau>n} + M_\tau \mathbb{1}_{\tau\leq n} \\
&= M_{n-1}\mathbb{1}_{\tau>n} + M_\tau \mathbb{1}_{\tau\leq n} \\
&= M_{_{n-1\wedge\tau}}
\end{aligned}$$

$\square$

Theorem 7. (Martingale Convergence Theorem) Let M_n be a martingale and suppose there exists some $C > 0$ such that $\mathbb{E}[|M_n|] \leq C$ for all $n > 0$. Then with probability 1, there exists a random variable M_∞ such that

$$\lim_{n\to\infty} M_n = M_\infty$$

Proof. Suppose M_n doesn't converge. By necessity, it must be true that

$$\liminf_{n\to\infty} M_n < \liminf_{n\to\infty} M_n.$$

Take $a, b \in \mathbb{R}$ such that there exist times $m_1 < n_1 < m_2 < n_2 < \cdots <$ so that

$$M_{m_j} \leq a, \quad M_{n_j} \geq b \qquad \forall j \in \mathbb{N}.$$

We say that an **upcrossing** of $[a, b]$ is an interval $\{m, \ldots, n\}$ such that $M_m \leq a$ and $M_n \geq b$ and $M_k \in (a, b)$ for all k in the uncrossing. Note that since M_n doesn't converge, we must have that M_n has an infinite number of upcrossings for some interval $[a, b]$.

Define $B_0 := 1$, and

$$B_n := \begin{cases} 1 \qquad B_{n-1} = 0, \quad M_{n-1} \leq a \\ 0, \qquad B_{n-1} = 1, \quad M_{n-1} \geq b \\ B_{n-1} \end{cases}$$

to be the betting strategy where you buy, to the best of your ability, a 'stock' when M_n is less than a and sell a 'stock' when M_n is greater than b. Then we can define W_n to be our total winnings at time n by

$$W_n = \sum_{k=0}^{n} B_k(M_k - M_{k-1}).$$

W_n is a **discrete stochastic integral** and is clearly $\mathscr{F}_n$−measurable with $\mathbb{E}[|W_n|] < \infty$. Moreover, we can say that because B_n is $\mathscr{F}_{n-1}$−measurable,

$$\begin{aligned}
\mathbb{E}[W_n \mid \mathscr{F}_{n-1}] &= \mathbb{E}\left[\sum_{k=0}^{n} B_k(M_k - M_{k-1}) \mid \mathscr{F}_{n-1}\right] \\
&= \sum_{k=0}^{n-1} B_k(M_k - M_{k-1}) + \mathbb{E}[B_n(M_n - M_{n-1}) \mid \mathscr{F}_{n-1}] \\
&= W_{n-1} + B_n\mathbb{E}[(M_n - M_{n-1}) \mid \mathscr{F}_{n-1}] \\
&= W_{n-1} + B_n\mathbb{E}[M_n \mid \mathscr{F}_{n-1}] - B_n M_{n-1} \\
&= W_{n-1}
\end{aligned}$$

Thus, W_n is a martingale. Let U_n be the number of upcrossings by time n. Each U_n results in a profit of at least $(b-a)$, and so

$$W_n \geq (b-a)U_n + (M_n - a),$$

where the last term represents the loss of holding on to the asset at the present. Thus,

$$0 = \mathbb{E}[W_0] = \mathbb{E}[W_n] \geq (b-a)\mathbb{E}[U_n] + \mathbb{E}[M_n] - a.$$

Rearranging,

$$\mathbb{E}[U_n] \leq \frac{a - \mathbb{E}[M_n]}{b-a} \leq \frac{|a| + \mathbb{E}[|M_n|]}{b-a} < \frac{|a| + C}{b-a}.$$

Because this holds for al n, then

$$\lim_{n\to\infty} \mathbb{E}[U_n] < K$$

for some K. Since for any n, $U_n \geq 0$ and $U_{n-1} \leq U_n$, then

$$\lim_{n\to\infty} \mathbb{E}[U_n] = \mathbb{E}[\lim_{n\to\infty} U_n] < K.$$

But this in turn implies that $\lim_{n\to\infty} U_n < \infty$, which is a contradiction. □

Remark 22. Consider the following assumptions: If $M_n \geq -k$ for all n, then

$$\begin{aligned}
\mathbb{E}[|M_n|] &= \mathbb{E}[M_n \mathbb{1}_{M_n \geq 0}] + \mathbb{E}[M_n \mathbb{1}_{M_n < 0}] \\
&\leq \mathbb{E}[M_n \mathbb{1}_{M_n \geq 0}] + k \\
&\leq \mathbb{E}[M_n + k] + k \\
&\leq \mathbb{E}[M_0] + 2k
\end{aligned}$$

If $M_n \in L_2$, then

$$\mathbb{E}[|M_n|] = \mathbb{E}[M_n \cdot 1] \leq \left(\mathbb{E}[|M_n|^2]\right)^{\frac{1}{2}} (\mathbb{E}[1])^{\frac{1}{2}} \leq C^{\frac{1}{2}}$$

The limit M_∞ is a random variable: Let X_n be a random walk on $\mathbb{Z}$, and let $\tau := \inf\{x \mid X_n \in \{a,b\}\}$. Then by Proposition 29, $X_{n\wedge\tau}$ is a martingale that is uniformly bounded. Clearly,

$$\lim_{n\to\infty} X_{n\wedge\tau} = X_\tau = \begin{cases} -a, & \text{w.p.} \frac{1}{2} \\ b, & \text{w.p.} \frac{1}{2} \end{cases}.$$

While we might expect that $\mathbb{E}[M_\infty] = M_0$ by the optional stopping theorem, this is usually not the case. To see this, let X_n be a random walk on $\mathbb{Z}$ and $\tau = \inf\{n : X_n = 1 \mid X_0 = 0\}$. Then

$$\mathbb{E}[\lim_{n\to\infty} X_{n\wedge\tau}] = \mathbb{E}[X_\tau] = 1 \neq \mathbb{E}[X_0] = 0.$$

Why do we need the assumption of boundedness? Let X_n be a random walk on $\mathbb{Z}$. We know that X_n is a martingale, but clearly $\mathbb{E}[|X_n|]$ is unbounded. In particular, $\lim\limits_{n\to\infty} X_n$ does not exist.

Example 1.37. Recall that $\sum \frac{1}{n} = \infty$ and $\sum \frac{(-1)^n}{n}$ converges. Suppose Y_n is a random variable such that $Y_n \in \{-1, 1\}$ w.p. $\frac{1}{2}$. What is $\sum \frac{Y_n}{n}$. It is not hard to see that

$$M_n = \sum_{k=0}^{n} \frac{Y_k}{k}$$

is a martingale. To apply MCT, we saw in Remark 21 that it suffices to show that $M_n \in L_2$. Note that

$$\mathbb{E}[|M_n|^2] = \sum_{k=0}^{n}\sum_{j=0}^{n} \mathbb{E}[Y_k Y_j]\frac{1}{kj} = \sum_{k=0}^{n} \mathbb{E}[Y_k^2]\frac{1}{k^2} < \infty.$$

Thus, we apply MCT to note that $M_n \to M_\infty = \displaystyle\sum_{k=0}^{\infty} \frac{Y_k}{k}$ and so the series converges.

Example 1.38. (Polya's Urn) We have an urn with red and green balls, starting with a single red and a single green ball. A time n, we pick a ball, put it back, and add another of the same color. Let $X_n = \#\{\text{red balls}\}$. Note that are $N + 2$ balls at time N. Thus,

$$\mathbb{P}\{X_{n+1} = x + 1 \mid X_n = x\} = \frac{x}{n+2}, \quad \mathbb{P}\{X_{n+1} = x \mid X_n = x\} = \frac{n+2-x}{n+2}$$

Define

$$M_n = \frac{X_n}{n+2}$$

to be the proportion of red bals in the urn. We claim that this is a martingale since

$$\mathbb{E}[M_{n+1]} \mid \mathscr{F}_n] = \frac{X_n}{n+2}\frac{X_n+1}{n+3} + \frac{n+2-x}{n+2}\frac{X_n}{n+3} = \frac{X_n}{n+2} = M_n.$$

Since this martingale is bounded by 1, then

$$\lim_{n\to\infty} M_n = M_\infty \sim U([0,1]).$$

1.21 Monday, May 12: Introduction to Brownian Motion

Informally, a Brownian motion $\{B_t\}_{t\geq 0}$ is a random function indexed by continuous time taking values in continuous space.

Definition 30. A **Brownian motion** is a a random function $B : [0, \infty) \to \mathbb{R}$ that satisfies:

(a) $B_0 = 0$;

(b) (Stationary increments) For every $0 < s < t$, $B_t - B_s$ has the same distribution as B_{t-s};

(c) (Independent increments) For every $s_1 \leq t_1 \dots s_k \leq t_k$, the increments $B_{t_j} - B_{s_j}$ are independent for $j = 1, 2, \dots, k$;

(d) (Continuity) $t \mapsto B_t$ is continuous.

Theorem 8. Let $B : [0, \infty) \to \mathbb{R}$ be a random continuous function with independent, stationary increments and $B_0 = 0$. Then there exists some $\mu \in \mathbb{R}$ and $\sigma^2 > 0$ such that $B_t \sim N(\mu t, \sigma^2 t)$, where μ and σ^2 uniquely characterize the distribution of $\{B_t\}_{t\geq 0}$

Remark 23. There are two way of building B_t :

(i) Suppose X_n is our coin flip random variable and $S_n = \sum X_i$. Then by the CLT, if Φ is the CDF of the standard normal,

$$\lim_{n\to\infty} \mathbb{P}\{a < \frac{S_n}{\sqrt{n}} < b\} = \Phi(b) - \Phi(a).$$

Now let

$$B_n^{(t)} := \frac{S_{\lfloor nt \rfloor}}{\sqrt{nt}}.$$

A theorem by Donsker states that $B_n^{(t)} \xrightarrow[n\to\infty]{} B_t$.

(ii) Levy's Construction is done by linearly interpolating a family of Normal r.v.'s indexed by a countably dense subset of $[0, 1]$.

Definition 31. **(Standard) Brownian motion** is the process $\{B_t\}_{t\geq 0}$ with $B_0 = 0$ satisfying:

(a) B is continuous

(b) For each $s < t$, $B_t - B_s$ has the normal distribution with mean 0 and variance $t - s$.

(c) For every $s_1 \leq t_1 \leq \dots s_k \leq t_k$, the increments $B_{t_j} - B_{s_j}$ are independent.

Example 1.39. Suppose B is a standard Brownian Motion. Then

$$\mathbb{P}\{B_1 \geq 1, B_3 \geq B_1 + 1\} = \mathbb{P}\{B_1 \geq 1, B_3 - B_1 \geq 1\}$$

By (c) and then (b),

$$\mathbb{P}\{B_1 \geq 1, B_3 - B_1 \geq 1\} = \mathbb{P}\{B_1 \geq 1\}\mathbb{P}\{B_3 - B_1 \geq 1\} = \int_1^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \int_1^\infty \frac{1}{2\sqrt{\pi}} e^{-\frac{x^2}{4}}\, dx$$

Proposition 30. Let B_t be a standard Brownian motion and let $c > 0$. Then $\frac{1}{\sqrt{c}} B_{ct}$ is a standard Brownian motion.

Proof. Clearly, the new motion is continuous and has independent increments. For any $s \leq t$, we have that

$$B_{ct} - B_{cs} \sim N(0, c(t-s)) \implies \frac{1}{\sqrt{c}}(B_{ct} - B_{cs}) \sim N(0, t-s)$$

$\square$

Proposition 31. For each fixed $t \geq 0$, it holds w.p. 1 that B is not differentiable at t.

Proof. Let $\epsilon > 0$, we have by the previous proposition that $\frac{1}{\sqrt{\epsilon}}(B_{t+\epsilon} - B_t) \sim N(0,1)$, and so

$$\frac{B_{t+\epsilon} - B_t}{\epsilon} \sim \frac{1}{\sqrt{\epsilon}} Z,$$

where $Z \sim N(0,1)$. Thus,

$$\mathbb{P}\{\frac{|B_{t+\epsilon} - B_t|}{\epsilon} \geq c\} = \mathbb{P}\{|Z| \geq \epsilon^{\frac{1}{2}} c\} \to 1,$$

and so

$$\limsup_{n \to \infty} \frac{B_{t+\epsilon} - B_t}{\epsilon} = \infty.$$

$\square$

1.22 Wednesday, May 14: Applications of Brownian Motion

Remark 24. Suppose $\{X_n\}$ is a discrete stochastic process, and $A \subset S$ has $\mathbb{P}\{X_n \in A\} = 0$ for any n. Then

$$\mathbb{P}\{\exists\, n : X_n \in A\} \leq \sum_{n=0}^{\infty} \mathbb{P}\{X_n \in A\} = 0.$$

This is not true for B.M. We know that $\mathbb{P}\{B_t = 0\}$, but we will see in this class that

$$\mathbb{P}\{\exists\, t : B_t = 1\}$$

Example 1.40. Suppose B is S.B.M. Then

(a)
$$\mathbb{E}[B_4 \mid B_2 = 6] = \mathbb{E}[B_4 - B_2 + B_2 \mid B_2 = 6] = \mathbb{E}[B_4 - B_2] + \mathbb{E}[B_2 \mid B_2] = 6$$

(b) Suppose $s \leq t$, then

$$\begin{aligned}
\mathbb{E}[B_s^2 B_t^2] &= \mathbb{E}[B_s^2(B_t - B_s + B_s)^2] \\
&= \mathbb{E}[\big(B_s(B_t - B_s) + B_s^2\big)^2] \\
&= \mathbb{E}[B_s^2(B_t - B_s)^2 + 2B_s^3(B_t - B_s) + B_s^4] \\
&= \mathbb{E}[B_s]^2\mathbb{E}[B_t - B_s]^2 + 2\mathbb{E}[B_s^3]\mathbb{E}[B_t - B_s] + \mathbb{E}[B_s^4] \\
&= s(t-s) + 0 + 3s^2 \\
&= s(t-s) + 3s^2
\end{aligned}$$

Where the fourth moment can be calculated using the moment-generating function of $X \sim N(0,1)$.

$$\mathbb{E}[X^{2n}] = (2n-1)!!\sigma^{2n}$$

(c) Find

$$\mathbb{P}\{B_2 > B_1 > B_3\}.$$

Let $X = B_1$, $Y = B_2 - B_1$ and $Z = B_3 - B_2$. It suffices to find

$$\begin{aligned}
\mathbb{P}\{X + Y > X > X + Y + Z\} &= \mathbb{P}\{Y > 0 > Y + Z\} \\
&= \mathbb{P}\{B_1 > 0, B_2 < 0\} \\
&= \int_0^\infty \mathbb{P}\{B_2 < 0 \mid B_1 = x\} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \\
&= \int_0^\infty \mathbb{P}\{B_2 < -x \mid B_1 = 0\} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \\
&= \int_0^\infty \mathbb{P}\{B_2 > x \mid B_1 = 0\} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \\
&= \int_0^\infty \left(\int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \right) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \\
&= \frac{1}{2\pi} \int_0^\infty \int_0^\infty e^{-(\frac{y^2+x^2}{2})}\, dxdy \\
&= \frac{1}{2\pi} \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \int_0^\infty e^{\frac{r^2}{2}} r\, dr d\theta \\
&= \frac{1}{8}
\end{aligned}$$

1.23 Friday, May 16: Properties of Brownian Motion

Definition 32. A continuous-time stochastic process M_t is a **martingale** with respect to $\mathscr{F}_t$ if:

(a) M_t is $\mathscr{F}-$measurable.

(b) $\mathbb{E}[|M_t|] < \infty$ for all t.

(c) For all $u > t$,
$$\mathbb{E}[M_u \mid \mathscr{F}_t] = M_t.$$

Proposition 32. (Markov Property) Let $\{B_t\}$ be a SBM. Then for each $t \geq 0$, the process $\{B_{s+t} - B_t\}_{s\geq 0}$ is an SBM independent of $\mathscr{F}_t$.

(Martingale Property) B is a martingale w.r.t. $\mathscr{F}_t$

Proof. We show the martingale property. Take $u > t$. Then

$$\begin{aligned}\mathbb{E}[B_u \mid \mathscr{F}_t] &= \mathbb{E}[B_{u-t} \mid \mathscr{F}_t] + \mathbb{E}[B_t \mid \mathscr{F}_t] \\ &= \mathbb{E}[B_{u-t}] + B_t \\ &= B_t\end{aligned}$$

□

Definition 33. A random variable $\tau \in [0, \infty]$ is a **stopping time** for $\mathscr{F}_t$ if, for any $t \geq 0$, we have that $\{\tau \leq t\}$ is $\mathscr{F}_t-$measurable.

Example 1.41. For $a \in \mathbb{R}$, let $\tau = \min\{t \geq 0 \mid B_t = a\}$ is a stopping time. $\tau' = \max\{t \leq 1 \mid B_t = a\}$ is *not* a stopping time.

Theorem 9. (OST) Let $\{M_t\}_{t\geq 0}$ be a martingale w.r.t $\{\mathscr{F}_t\}$ such that $t \mapsto M_t$ is continuous. Let τ be a stopping time and suppose there is a $c > 0$ such that

$$\mathbb{P}\{\tau < \infty, |M_t| \leq C, \forall\, t \leq \tau\} = 1.$$

Then
$$\mathbb{E}[M_\tau] = \mathbb{E}[M_0].$$

Proposition 33. Suppose B is an SBM and define

$$\tau := \min\{t \geq 0 \mid B_t = a \text{ or } B_t = -b\}.$$

Then
$$\mathbb{P}\{B_\tau = a\} = \frac{b}{a+b}$$

Proof. We know that since B_t has not struck a or b when $\tau > t$, we have that

$$\mathbb{P}\{\tau > t\} \leq \mathbb{P}\{B_t \in [-b, a]\} \to 0 \implies \mathbb{P}\{\tau < \infty\} = 1.$$

We also know that for $t \leq \tau$,
$$|B_t| \leq \max\{|a|, |b|\}.$$

Hence, by the OST
$$\mathbb{E}[B_\tau] = \mathbb{E}[B_0] = 0$$

but
$$\mathbb{E}[B_\tau] = \mathbb{P}\{\tau = a\}a + \mathbb{P}\{\tau = -b\}(-b) = \mathbb{P}\{\tau = a\}a + (1 - \mathbb{P}\{\tau = a\})(-b) = 0$$

□

Proposition 34. Brownian motion is recurrent in one dimension. That is, for each $T \geq 0$,

$$\mathbb{P}\{\exists t > T : B_t = 0\} = 1$$

Proof. The process $\{B_{T+t} - B_T\}_{t\geq 0}$ is an SBM independent of $\mathscr{F}_T$ by the Markov property. It suffices to show that there exists some t such that $B_{T+t} - B_T = -B_T$ almost surely. We use the previous result with $b = -B_T$ and $a > 0$ to find that

$$\mathbb{P}\{B_{T+t} - B_T = a \text{ before } -B_T\} = \frac{B_T}{a + B_T} \xrightarrow[a\to\infty]{} 0.$$

$\square$

1.24 Monday, May 19: The Reflection Principle

Proposition 35. (Strong Markov Property) Let τ be a stopping time for B_t. The process $\{B_{s+\tau} - B_\tau\}_{s\geq 0}$ is a Brownian motion and is independent of $\mathscr{F}_\tau$.

Proposition 36. Let B_t be a SBM and let $a > 0$. Then

$$\mathbb{P}\{\max_{0\leq s\leq t} B_s \geq a\} = 2\mathbb{P}\{B_t \geq a\}$$

Proof. Define

$$\tau := \min\{t \geq 0 : B_t = a\}.$$

Then τ is a stopping time w.r.t. $\mathscr{F}_t$. Thus,

$$\{\max_{0\leq s\leq t} B_s \geq a\} = \{\tau \leq t\}$$

and so using the law of total probability and then the strong markov property on the event $\tau \leq t$, we know that

$$B_t - B_\tau = B_t - a \sim N(0, t-\tau)$$

$$\begin{aligned}
\mathbb{P}\{B_t \geq a\} &= \mathbb{P}\{B_t \geq a \mid \tau \leq t\}\mathbb{P}\{\tau \leq t\} + \mathbb{P}\{B_t \geq a \mid \tau \geq t\}\mathbb{P}\{\tau > t\} \\
&= \mathbb{P}\{B_t \geq a \mid \tau \leq t\}\mathbb{P}\{\tau \leq t\} + 0 \\
&= \mathbb{P}\{B_t - a \geq 0 \mid \tau \leq t\}\mathbb{P}\{\tau \leq t\} \\
&= \frac{1}{2}\mathbb{P}\{\tau \leq t\}
\end{aligned}$$

$\square$

Example 1.42. Let $\{B_t\}$ be a SBM. For $t > 1$, we wish to compute $\mathbb{P}\{B_s = 0 \text{ for some } 1 \leq s \leq t\}$. Suppose $B_1 = a > 0$. Then the probability $B_s = 0$ for some $s \in [1, t]$ is the same as probability that $B_s \leq a$ for some $0 \leq s \leq t-1$. By symmetry, this is the same as the probability that $B_s \geq a$ for some $0 \leq s \leq t-1$. Thus,

$$\begin{aligned}
\mathbb{P}\{B_s = 0 \text{ for some } 1 \leq s \leq t \mid B_1 = a\} &= \mathbb{P}\{B_s \geq a \text{ for some } 0 \leq s \leq t-1\} \\
&= 2\mathbb{P}\{B_{t-1} \geq a\} \\
&= 2\int_a^\infty \frac{1}{\sqrt{2\pi(t-1)}} e^{-\frac{x^2}{2(t-1)}}
\end{aligned}$$

Averaging over all possible values of a,

$$\mathbb{P}\{B_s = 0 \text{ for some } 1 \leq s \leq t\} = 2\int_{-\infty}^\infty \left(\int_a^\infty \int_a^\infty \frac{1}{\sqrt{2\pi(t-1)}} e^{-\frac{x^2}{2(t-1)}}\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{a^2}{2}}\, da = 1 - \frac{2}{\pi} \arctan \frac{1}{\sqrt{t-1}}$$

1.25 Wednesday, May 21: Multi-Dimensional Brownian Motion

Definition 34. Let $d \in \mathbb{N}$. The **Gaussian (normal) distribution in** $\mathbb{R}^d$ with mean 0 and covariance matrix $\sigma^2 I$ is the distribution of

$$\overline{X} = (X_1, \ldots, X_d),$$

where each $X_i \sim N(0, \sigma^2)$ i.i.d.

Remark 25. The density of $\overline{X}$ is given by

$$\prod_{j=1}^{d} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-x_j^2}{2\sigma^2}} = \frac{1}{(2\pi\sigma^2)^{\frac{d}{2}}} \exp\{\frac{-|\overline{x}|^2}{2\sigma^2}\}$$

Proposition 37. Let $\overline{X}$ be a normal r.v. in $\mathbb{R}^d$ with mean 0 and covariance matrix $\sigma^2 I$ and let Q be an $n \times n$ orthonormal matrix. Then $Q\overline{X} \stackrel{d}{=} \overline{X}$

Definition 35. The **standard Brownian motion in** $\mathbb{R}^d$ is the stochastic process $\{\overline{B}_t\}_{t\geq 0} = \{(B_t^1, B_t^2, \ldots, B_t^d)\}_{t\geq 0}$ where $B_t^1, \ldots, B_t^d$ are independent one dimensional standard Brownian motions.

Proposition 38. The S.B.M. in $\mathbb{R}^d$ is the unique stochastic process taking values in $\mathbb{R}^d$ satisfying:

(a) $\overline{B}_0 = 0$;

(b) $\overline{B}$ is continuous;

(c) For each $s < t$, $(\overline{B}_t - \overline{B}_s)$ has a multivariate Gaussian distribution with mean 0 and covariance matrix $(t-s)I$.

(d) For $s_1 \leq t_1 \leq \cdots s_k \leq t_k \leq \cdots$ the increments $\overline{B}_{t_j} - \overline{B}_{s_j}$ are independent.

Remark 26. Multi-dimensional B.M. satisfies some things that the one-dimensional B.M did:

- $\overline{B}$ satisfies Brownian scaling. I.e, $\{\sqrt{c}\overline{B}_{ct}\} =^d \{\overline{B}\}$
- $\overline{B}$ satisfies the Markov property and the strong Markov property.

Example 1.43. Suppose $\{W_t\}$ is a 2D B.M starting at $(1,1)$. What is the probability that W_t hits the positive real axis before it hits the negative real axis. We can write

$$W_t = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} B_t^1 \\ B_t^2 \end{pmatrix}$$

where B_t^1, B_t^2 are S.B.M. Let

$$\tau = \inf\{t \geq 0 : (1,1) + (B_t^1, B_t^2) \in \mathbb{R} \times (-\infty, 0]\} = \inf\{\tau \geq 0 : 1 + B_t^2 \leq 0\}$$

Then τ is a stopping time w.r.t. $\mathscr{F}_t^2$. We seek

$$\mathbb{P}\{1 + B_\tau^1 \geq 0\} = 1 - \mathbb{P}\{B_\tau^1 \leq -1\}.$$

Conditioning,

$$\begin{aligned}
\mathbb{P}\{B_\tau^1 \leq -1\} &= \int_0^\infty \mathbb{P}\{B_t^1 \leq -1 \mid \tau = t\} \mathbb{P}\{\tau = t\}\, dt \\
&= \int_0^\infty \mathbb{P}\{B_t^1 \leq -1 \mid \tau = t\} f_\tau(t)\, dt \\
&= \int_0^\infty \mathbb{P}\{B_t^1 \leq -1\} f_\tau(t)\, dt
\end{aligned}$$

We know by the Reflection Principle that

$$\mathbb{P}\{\tau \leq t\} = 2\mathbb{P}\{B_t^1 \leq -1\}$$

and thus

$$\begin{aligned}
\int_0^\infty \mathbb{P}\{B_t^1 \leq -1\} f_\tau(t)\, dt &= 2\int_0^\infty \mathbb{P}\{B_t^1 \leq -1\} d\mathbb{P}\{B_t^1 \leq -1\}\, dt \\
&= (\mathbb{P}\{B_t^1 \leq -1\}^2)_0^\infty \\
&= \lim_{t\to\infty} \left(\mathbb{P}\{B_t^1 \leq -1\}\right)^2 \\
&= \lim_{t\to\infty} \left(B_1^1 \leq \frac{-1}{\sqrt{t}}\right)^2 \\
&= \lim_{t\to\infty} (\Phi(\frac{-1}{\sqrt{t}}))^2 \\
&= \frac{1}{4}
\end{aligned}$$

and so our answer is $\frac{3}{4}$.

1.26 Friday, May 23: The Heat Equation

Definition 36. Let $f \in C^2(\mathbb{R}^d, \mathbb{R})$. The **Laplacian** of f is defined by

$$\Delta f(\overline{x}) = \sum_{i=1}^{d} \frac{\partial^2 f(x_i)}{dx_i^2}.$$

Theorem 10. Let $f \in C^2(\mathbb{R}^d \times [0,\infty), \mathbb{R})$ and assume that for each $t > 0$, there is some C_t such that

$$|f(\overline{x}, t)| \leq C_t e^{C_t |\overline{x}|}$$

and so are its partials. Assume f also satisfies the **heat equation**, i.e,

$$\frac{\partial}{\partial t} f(\overline{x}, t) + \frac{1}{2} \Delta_{\overline{x}} f(\overline{x}, t) = 0, \tag{1}$$

where $\Delta_{\overline{x}}$ is only taken in $\overline{x}$. Let $\overline{B}$ be a Brownian motion in $\mathbb{R}^d$. Then $f(\overline{B}_t, t)$ is a Martingale w.r.t. $\mathscr{F}_t$.

Proof. We sketch the proof. We claim that for any $\overline{x}_0 \in \mathbb{R}^d$ and for any $t > 0$,

$$\mathbb{E}[f(\overline{B}_t + \overline{x}_0, t)] = f(\overline{x}, 0).$$

Note that

$$p_t(\overline{x}) := \frac{1}{(\sqrt{2\pi t})^d} e^{\frac{-\overline{x}^2}{2t}}$$

satisfies the heat equation with the sign flipped. That is,

$$\partial_t p(t) - \frac{1}{2} \Delta_{\overline{x}} p_t(\overline{x}) = 0.$$

Note that

$$\begin{aligned} \partial_t \mathbb{E}[f(\overline{B}_t + \overline{x}_0, t)] &= \partial_t \int_{\mathbb{R}^d} p_t(\overline{x} - \overline{x}_0) f(\overline{x}, t)\, d\overline{x} \\ &= \cdots \text{[integrating by parts 3x]} \\ &= 0 \end{aligned}$$

Hence, we get the martingale property with $\overline{B}_s$ instead of $\overline{x}_0$ and $t - s$ instead of t:

$$\begin{aligned} \mathbb{E}[f(\overline{B}_t + \overline{x}_0, t) \mid \mathscr{F}_s] &= \mathbb{E}[f(\overline{B}_t - \overline{B}_s + \overline{B}_s, t - s + s) \mid \mathscr{F}_s] \\ &= f(\overline{B}_s, s) \end{aligned}$$

□

Definition 37. Let $f \in C^2(\mathbb{R}^d, \mathbb{R})$. We say that f is **harmonic** if $\Delta f = 0$.

Remark 27. Harmonic functions immediately imply (1) since they are not dependent on time.

Example 1.44. Let $f(x, y) = e^x \cos y$. We have that

$$\frac{\partial^2}{\partial x^2} = e^x \cos y$$

$$\frac{\partial^2}{\partial y^2} = \frac{\partial}{\partial y}(-e^x \sin y) = -e^x \cos y.$$

Hence, f is harmonic. We also have that

$$|f(x,y)| \leq |e^x \cos x| \leq e^x$$

Then if $\overline{B}_t$ is a standard 2D B.M, we know that $f(\overline{B}_t)$ is a martingale w.r.t $\mathscr{F}_t$.

Let $\tau = \inf\{t \geq 0 \mid |\overline{B}_t - (1,0)| > 5\}$ be a stopping time for $\overline{B}_t$. Suppose the conditions of OST are met. Then

$$\mathbb{E}[f(\overline{B}_\tau)] = \mathbb{E}[f(\overline{B}_0)] = 1$$

But

$$\mathbb{E}[f(\overline{B}_\tau)] > \mathbb{P}\{\tau < \infty\} f(\overline{B}_\tau)$$

2 Problem Sessions

2.1 Monday, Mar 31: Problem Session 1

When do we run into the issue that $\lim_{n\to\infty} \pi P^n$ does not exist?

(1) If Period > 1

(2) If there are multiple recurrence classes that are transient.

Problems:

(a) Let $X_0, X_1, \ldots$ be a Markov chain with state space $S = \{0, 1, 2, 3\}$, and with transition matrix

$$\begin{bmatrix} \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{5} & 0 & \frac{1}{5} \\ 0 & 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \end{bmatrix}.$$

A new process is defined by $Z_n = 0$ if $X_n = 0$ or 1 and $Z_n = X_n$ if $X_n = 2$ or 3. Find $P(Z_{n+1} = 2|Z_n = 0, Z_{n-1} = 2)$ and $P(Z_{n+1} = 2|Z_n = 0, Z_{n-1} = 3)$. Is Z_n a Markov chain?

SOLUTION:

$$\begin{aligned} \mathbb{P}\{Z_{n+1} = 2 \mid Z_n = 0, Z_{n-1} = 2\} &= \frac{\mathbb{P}\{Z_{n+1} = 2, Z_n = 0, Z_{n-1} = 2\}}{\mathbb{P}\{Z_n = 0, Z_{n-1} = 2\}} \\ &= \frac{\mathbb{P}\{Z_{n+1} = 2, Z_n = 0 \mid Z_{n-1} = 2\}}{\mathbb{P}\{Z_n = 0 \mid Z_{n-1} = 2\}} \\ &= 0 \end{aligned}$$

$$\begin{aligned} \mathbb{P}\{Z_{n+1} = 2 \mid Z_n = 0, Z_{n-1} = 3\} &= \frac{\mathbb{P}\{Z_{n+1} = 2, Z_n = 0 \mid Z_{n-1} = 3\}}{\mathbb{P}\{Z_n = 0 \mid Z_{n-1} = 3\}} \\ &= \frac{\mathbb{P}\{X_n = 2, X_n = 0 \mid X_{n-1} = 3\} + \mathbb{P}\{X_n = 2, X_n = 1 \mid X_{n-1} = 3\}}{\mathbb{P}\{X_n = 0 \mid X_{n-1} = 3\} + \mathbb{P}\{X_n = 1 \mid X_{n-1} = 3\}} \\ &= \frac{\frac{1}{3}\frac{1}{2} + 0}{\frac{1}{6} + \frac{1}{6}} \\ &= \frac{1}{4} \end{aligned}$$

Thus, $\{Z_n\}$ is not a Markov chain ■

(b) We repeatedly roll two four-sided dice with numbers 1, 2, 3, and 4 on them. Let Y_k be the sum on the k-th roll, $S_n = Y_1 + Y_2 + \cdots + Y_n$ be the total of the first n rolls, and $X_n = S_n (\bmod 6)$. Find the transition matrix for $\{X_n\}$.

SOLUTION: $\{X_n\}$ is a Markov process because it only depends on the current state S_n, as you can figure out the next turn only from this, since

$$X_{n+1} = S_{n+1} \bmod 6 = (S_n + Y_{n+1}) \bmod 6 = S_n \bmod 6 + Y_{n+1} \bmod 6 = X_n + Y_{n+1} \bmod 6$$

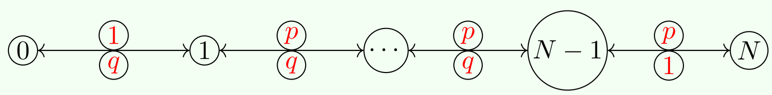
$$P = \begin{bmatrix} \frac{3}{16} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{3}{16} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{16} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{3}{16} \\ \frac{3}{16} & \frac{1}{4} & \frac{3}{16} & \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{3}{16} & \frac{1}{4} & \frac{3}{16} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{3}{16} & \frac{1}{4} & \frac{3}{16} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \frac{3}{16} & \frac{1}{4} & \frac{3}{16} \end{bmatrix}$$

■

(c) Consider a Markov chain with states $S = \{0, \dots, N\}$ and transition probabilities $p(i, i+1) = p$, $p(i, i-1) = q$, for $1 \leq i \leq N-1$, where $p + q = 1$, $0 < p < 1$. Assume $p(0,1) = p(N, N-1) = 1$.

(i) Draw a transition diagram for this chain.

SOLUTION: In the following diagram, if you the probability is above, then it is going to the right:



■

(ii) Is the Markov chain irreducible?

SOLUTION: Let $i, j \in [N]$, then we claim that $i \leftrightarrow j$. Without loss of generality, suppose that $i < j$. Suppose $j - i = N$. Then $p(i,j) \geq p^N > 0$ and $p(j,i) \geq q^N > 0$. ■

(iii) What is the period of this chain?

SOLUTION: The period has to be 2. ■

(d) A taxicab driver moves between the airport A and two hotels B and C according to the following rules. If he is at the airport, he will be at one of the two hotels next with equal probability. If at a hotel, then he returns to the airport with probability $\frac{3}{4}$ and goes to the other hotel with probability $\frac{1}{4}$.

(i) Find the transition matrix for the chain.

SOLUTION:

$$P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{3}{4} & 0 & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} & 0 \end{bmatrix}$$

■

(ii) Suppose the driver begins at the airport at time 0. Find the probability for each of his three possible locations at time 2 and the probability he is at hotel B at time 3.

SOLUTION: Squaring the matrix gives

$$P^2 = \begin{bmatrix} \frac{3}{4} & \frac{1}{8} & \frac{1}{8} \\ \frac{3}{16} & \frac{7}{16} & \frac{6}{16} \\ \frac{3}{16} & \frac{6}{16} & \frac{7}{16} \end{bmatrix}$$

and so

$$p^2(A,A) = \frac{3}{4}, \quad p^2(B,A) = \frac{3}{16}, \quad p^2(C,A) = \frac{3}{16}$$

Similarly,

$$p^3(B,A) = \frac{13}{32}$$

■

(e) At time $n = 0$, two ladybirds are placed at vertices i and j of a regular hexagon, whose vertices are labeled $1, \ldots, 6$. At time $n = 1$, each of them moves, independently of the other, to one of the two adjacent vertices with probability $\frac{1}{2}$, and so on at each time $n = 2, 3, \ldots$.

(i) Denote X_n the distance between the two ladybirds at time $n \geq 0$, i.e., the minimum number of edges between them. Find the transition matrix for $\{X_n\}$.

SOLUTION: $S = \{0, 1, 2, 3\}$

$$P = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{3}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

■

(ii) Identify the communication classes. Are they recurrent or transient?

$$C_1 = \{0, 2\}, \quad C_2 = \{1, 3\}.$$

Both are recurrent.

(f) Find the invariant (stationary) distributions for the following Markov chains with given transition matrices: (a)

$$\begin{bmatrix} \frac{1}{2} & \frac{2}{5} & \frac{1}{10} \\ \frac{1}{5} & \frac{1}{2} & \frac{3}{10} \\ \frac{1}{10} & \frac{3}{10} & \frac{3}{5} \end{bmatrix}$$

SOLUTION: This will be the only time I type out a full solution to this. We need to solve for the normalized eigenvector of P^T corresponding to $\lambda = 1$. It suffices to find π^T such that

$$10P^T - 10\pi^T = 0$$

That is, we need to find the eigen basis of

$$E_{10}(10P^T) = \begin{pmatrix} 5 & 2 & 1 \\ 4 & 5 & 3 \\ 1 & 3 & 6 \end{pmatrix} - 10I$$

$$
\begin{aligned}
&= \begin{pmatrix} -5 & 2 & 1 \\ 4 & -5 & 3 \\ 1 & 3 & -4 \end{pmatrix} \\
&\simeq \begin{pmatrix} -20 & 8 & 4 \\ 20 & -25 & 15 \\ 20 & 60 & -80 \end{pmatrix} \\
&\simeq \begin{pmatrix} -20 & 8 & 4 \\ 0 & -17 & 19 \\ 0 & 68 & -76 \end{pmatrix} \\
&\simeq \begin{pmatrix} 1 & -\frac{2}{5} & -\frac{1}{5} \\ 0 & 1 & -\frac{19}{17} \\ 0 & -68 & 76 \end{pmatrix} \\
&\simeq \begin{pmatrix} 1 & -\frac{2}{5} & -\frac{1}{5} \\ 0 & 1 & -\frac{19}{17} \\ 0 & -68 & 76 \end{pmatrix} \\
&\simeq \begin{pmatrix} 1 & 0 & -\frac{11}{17} \\ 0 & 1 & -\frac{19}{17} \\ 0 & 0 & 0 \end{pmatrix}
\end{aligned}
$$

Thus,

$$\pi_1^T = \frac{11}{17}\pi_3^T, \qquad \pi_2^T = \frac{19}{17}\pi_3^T, \qquad \pi_1^T + \pi_2^T + \pi_3^T = 1$$

Solving:

$$
\begin{aligned}
\left[\begin{array}{ccc|c} 1 & 0 & -\frac{11}{17} & 0 \\ 0 & 1 & -\frac{19}{17} & 0 \\ 1 & 1 & 1 & 1 \end{array}\right] &\simeq \left[\begin{array}{ccc|c} 1 & 0 & -\frac{11}{17} & 0 \\ 0 & 1 & -\frac{19}{17} & 0 \\ 0 & 0 & \frac{28}{17} & 1 \end{array}\right] \\
&\simeq \left[\begin{array}{ccc|c} 17 & 0 & -11 & 0 \\ 0 & 17 & -19 & 0 \\ 0 & 0 & 28 & 47 \end{array}\right] \\
&\simeq \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{11}{47} \\ 0 & 1 & 0 & \frac{19}{47} \\ 0 & 0 & 1 & \frac{17}{47} \end{array}\right]
\end{aligned}
$$

So then

$$\pi = (\frac{11}{47}, \frac{19}{47}, \frac{17}{47})$$

■

(b)

$$\begin{bmatrix} \frac{1}{2} & \frac{2}{5} & \frac{1}{10} \\ \frac{3}{10} & \frac{2}{5} & \frac{3}{10} \\ \frac{1}{5} & \frac{1}{5} & \frac{3}{5} \end{bmatrix}.$$

SOLUTION: $\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ ■

2.2 Monday, Apr 7: Problem Session 2

(a) A queen can move any number of squares horizontally, vertically, or diagonally on an 8×8 chessboard. Let $\{X_n\}$ be the sequence of squares that results if we pick one of the queen's legal moves uniformly at random. Find

(i) the stationary distribution and

SOLUTION: We use Example 1.12. First, consider that the following 4x4 grid of the degrees for each positions on the upper left of the board: $\begin{bmatrix} 21 & 21 & 21 & 21 \\ 21 & 23 & 23 & 23 \\ 21 & 23 & 25 & 25 \\ 21 & 23 & 25 & 27 \end{bmatrix}$ Thus,

$$\pi_{(0,0)} = \frac{\deg(1,1)}{2|E|} = \frac{21}{1456}$$

■

(ii) the expected number of moves needed to return to the bottom left corner when we start there.

SOLUTION: If we define

$$T_{i,j} = \inf[n : X_n = (i,j) \mid X_0 = (i,j)],$$

then by Example 1.12,

$$\mathbb{E}[T_{0,0}] = \frac{1}{\pi_{0,0}} = \frac{1456}{21}$$

■

(b) Let $\{X_n\}$ be the random walk on $\{-10, -9, \ldots, 9, 10\}$ with reflected boundary, i.e., the Markov chain with transition probabilities

$$p(x, x+1) = p(x, x-1) = \frac{1}{2}, \quad \forall x \in \{-9, \ldots, 9\}, \quad p(-10, -9) = p(10, 9) = 1.$$

Assume that $X_0 = 0$ and let

$$T = \min\{n \geq 1 : X_n = 1\}.$$

Find the expected number of times that the walk hits 0 before time T, i.e., compute

$$\mathbb{E}\left[\#\{n \in \{0, \ldots, T\} : X_n = 0\}\right].$$

(Hint: use the strong Markov property applied at the times for which $X_n = 0$).

SOLUTION: Call

$$T_1 = \inf\{n : X_n = 1 \mid X_0 = 0\}$$

Call

$$K(i,j) = \sum_{n=1}^{j} \mathbb{1}_{X_n=1}$$

Then using the law of total expectation:

$$\begin{aligned}\mathbb{E}[K(0,T_1)] &= \mathbb{E}[\mathbb{E}[K(0,T_1) \mid X_1]] \\ &= \frac{1}{2}(\mathbb{E}[K(0,T_1) \mid X_1 = -1] + 1) + \frac{1}{2}(\mathbb{E}[K(0,T_1) \mid X_1 = 1] + 1) \\ &\text{By definition:} \\ &= \frac{1}{2}(\mathbb{E}[K(0,T_1) \mid X_1 = -1] + 1) + \frac{1}{2}(0+1) \\ &\text{Strong Markov Property:} \\ &= \frac{1}{2}(\mathbb{E}[K(T_0,T_1)] + 1) + \frac{1}{2}(0+1) \\ &= \frac{1}{2} + \frac{1}{2}(\mathbb{E}[K(0,T_1)] + 1)\end{aligned}$$

Solving:

$$\mathbb{E}[K(0,T_1)] = 2.$$

■

(c) Consider the numbers $\{1, 2, 3, \ldots, 12\}$ written around a ring as they usually are on a clock. Consider a Markov chain $\{X_n\}$ that at any point jumps with equal probability to the two adjacent numbers.

(i) What is the expected number of steps that $\{X_n\}$ will take to return to its starting position?

Solution: Suppose WLOG $X_n = 12$. Then $2|E| = 24$ and $\deg 12 = 2$. Thus, if

$$T_{12} = \inf\{n > 0 : X_n = 12 \mid X_0 = 12\},$$

then

$$\mathbb{E}[T_{12}] = \frac{1}{\pi_{12}} = \frac{2|E|}{\deg(12)} = \frac{24}{2} = 2.$$

■

(ii) What is the probability that $\{X_n\}$ will visit all the other states before returning to its starting position?

Solution: We will use the law of total probability. Again, WLOG, assume $X_0 = 12$. Define

$$E_{12} = \text{running through everything before returning to 12}$$

$$\mathbb{P}\{E_{12}\} = 2(\mathbb{P}\{E_{12} \mid X_1 = 1\}\mathbb{P}\{X_1 = 1\})$$

Thus, we can make $\mathbb{P}\{E_{12} \mid X_1 = 1\}$ a gambler's ruin problem!

Figure 2: Gambling With a Clock

the probability of reaching 12 on the left is given by $1 - P_{11} = 1 - \frac{11}{12} = \frac{1}{12}$ ■

(d) Before you are six light bulbs (all off), numbered $\{1, 2, \ldots, 6\}$. At each time $n \geq 1$, you roll a fair six-sided die and flip the switch of the bulb corresponding to the number you have rolled, turning it on if it is off, and off if it is on. What is the expected number of rolls needed until every light bulb is turned on?

(e) Find the probability that, in the process of repeatedly flipping a fair coin, one will encounter a run of 5 heads in a row before one encounters a run of 2 tails in a row.

SOLUTION: ■

(f) In a game similar to three card monte, the dealer places three cards on the table: the queen of spades and two red cards. The cards are placed in a row, and the queen starts in the center; the card configuration is thus RQR. The dealer proceeds to move. With each move, the dealer randomly switches the center card with one of the two edge cards (so the configuration after the first move is either RRQ or QRR). What is the probability that, after 2025 moves, the center card is the queen?

2.3 Monday, Apr 14: Problem Session 3

(a) Let $\{X_n\}$ be a Markov chain with state space $S = \{0, 1, 2, \ldots\}$. For each of the following transition probabilities, state whether the chain is positive recurrent, null recurrent, or transient. If it is positive recurrent, give the stationary distribution.

(i) $p(x, x+1) = \frac{5}{7}$, $p(x, 0) = \frac{2}{7}$.

Solution: We claim that $\{X_n\}$ is positive recurrent. Indeed, we can see that the stationary distribution must satisfy

$$\pi_0 = \frac{2}{7}(\sum_{n=0}^{\infty} \pi_n), \pi_1 = \frac{5}{7}\pi_0, \quad \ldots, \pi_n = (\frac{5}{7})^n \pi_0, \quad \sum_{n=0}^{\infty} \pi_n = 1.$$

Then

$$\pi_n = \left(\frac{5}{7}\right)^n \frac{2}{7}$$

is a valid stationary distribution.

■

(ii) $p(x, 0) = \frac{x+1}{x+2}$, $p(x, x+1) = \frac{1}{x+2}$.

Solution: Let $\tau_0 := \{n > 0 | X_n = 0\}$. Then

$$\mathbb{P}\{\tau_0 = \infty \mid X_0 = 0\} \geq \prod_{k=1}^{\infty} \frac{1}{(k+2)} = \lim_{n \to \infty} \frac{1}{(n+2)!} = 0.$$

Thus, $\{X_n\}$ is recurrent. Let's check for a stationary distribution!

■

(iii) $p(x, x+1) = \frac{x+1}{x+2}$, $p(x, 0) = \frac{1}{x+2}$.

(iv) $p(x, 0) = \frac{1}{x^2+2}$, $p(x, x+1) = \frac{x^2+1}{x^2+2}$.

(b) Let $\{X_n\}$ be a biased random walk on $\mathbb{Z}$ with probability $p > 1/2$ to move to the right, i.e., the transition probabilities are

$$p(x, x-1) = 1 - p, \quad p(x, x+1) = p, \quad \forall x \in \mathbb{Z}.$$

Assume that $X_0 = 0$.

(i) Find $\mathbb{P}[X_n = 0]$. **(Hint:** $X_n = 0$ if and only if we have the same number of $+1$ steps as -1 steps before time n.)

Solution:

$$p^{2n}(0, 0) = \binom{2n}{n} p^n (1-p)^n$$

■

(ii) For $K \geq 1$, let $T_K = \min\{n \geq 1 : X_n = K\}$. Find $\mathbb{P}[X_n \geq 0, \forall n = 0, \ldots, T_K]$.

SOLUTION: This is a shifted gambler's ruin. We can think of it as shifting it by 1 to the right and the probability of winning $K+1$ dollars before losing it all (starting with one dollar). We have found that

$$P_1 = \frac{1-(\frac{1-p}{p})}{1-(\frac{1-p}{p})^{K+1}}$$

■

(iii) Find the probability that X_n returns to 0 after time T_K.

SOLUTION: Let $A = \{\exists n > T_K \mid X_n = 0\}$. Then by the strong Markov property and the law of total probability, if we call

$$\alpha(k) := \mathbb{P}\{A \mid X_0 = k\},$$

then

$$\begin{aligned}\alpha(k) &= \alpha(k-1)\mathbb{P}\{X_1 = K-1\} + \alpha(k+1)\mathbb{P}\{X_1 = K+1\} \\ &= \alpha(k-1)(1-p) + \alpha(k+1)p\end{aligned}$$

Solving gives

$$\alpha = \frac{1 \pm \sqrt{1-4p(1-p)}}{2p} = \frac{1 \pm (1-2p)}{2p} \in \{1, \frac{1-p}{p}\}.$$

Thus,

$$\alpha(n) = \lambda_1 + \lambda_2(\frac{1-p}{p})^n.$$

We know that $\alpha(0) = 1$. Thus,

$$\lambda_1 + \lambda_2 = 1$$

Moreover, we know that

$$\lim_{n\to\infty} \alpha(n) = 0 = \lambda_1.$$

Thus, $\lambda_2 = 1$ and so

$$\alpha(K) = (\frac{1-p}{p})^K.$$

■

(iv) Find the probability that $X_n \geq 0$ for all $n = 0, \ldots, T_K$ and X_n returns to 0 after time T_K.

SOLUTION: By the strong Markov Property, these are independent, so simply multiply the results from (ii) and (iii) ■

(c) Consider the Markov chain with state space $\mathbb{Z}$ whose transition probabilities are given by $p(x, x-1) = 2/3$ and $p(x, x+2) = 1/3$ ($p(x, y) = 0$ for all other y). Determine whether this Markov chain is positive recurrent, null recurrent, or transient. **(Hint:** use a calculation with the binomial distribution similar to what we did for random walk on $\mathbb{Z}$.)

SOLUTION:

$$
\begin{aligned}
p^{3k}(0,0) &= \binom{3k}{k}(\frac{1}{3})^k(\frac{2}{3})^{2k} \\
&= \frac{(3k)!}{(k)!(2k)!}(\frac{1}{3})^k(\frac{2}{3})^{2k} \\
&\sim \frac{\sqrt{6\pi k}(\frac{3k}{e})^{3k}}{\sqrt{2\pi k}(\frac{k}{e})^k\sqrt{4\pi k}(\frac{2k}{e})^{2k}}(\frac{1}{3})^k(\frac{2}{3})^{2k} \\
&= \frac{\sqrt{6\pi k}(\frac{3k}{e})^{3k}}{\sqrt{2\pi k}(\frac{k}{e})^k\sqrt{4\pi k}(\frac{2k}{e})^{2k}}(\frac{1}{3})^k(\frac{2}{3})^{2k} \\
&= \sqrt{\frac{3}{4\pi k}}
\end{aligned}
$$

We have that

$$\sum_{k=0}^{\infty} p^{3k}(0,0) = \infty,$$

and so the process is recurrent. It is null recurrent since $p^n(0,0) \to 0$.

■

(d) The *infinite binary tree* is the graph T whose vertex set $V(T)$ consists of the empty sequence $\emptyset$ together with all k-tuples of the form $(a_1, a_2, \ldots, a_k)$, where $k \in \mathbb{N}$ and $a_j \in \{0,1\}$ for $j \in \{1, \ldots, k\}$. There are edges joining $\emptyset$ to (0) and (1) and edges joining $(a_1, \ldots, a_k)$ to $(a_1, \ldots, a_k, 0)$ and $(a_1, \ldots, a_k, 1)$ for each $(a_1, \ldots, a_k)$, see the figure below. Let $\{X_n\}$ be the random walk on T. Compute

$$\mathbb{P}[X_n = \emptyset \text{ for some } n \geq 1 \,|\, X_0 = \emptyset],$$

and conclude that $\{X_n\}$ is transient. **(Hint:** let H_n be the "height" for X_n, i.e., if $X_n = \emptyset$ then $H_n = 0$ and if $X_n = (a_1, \ldots, a_k)$ then $H_n = k$. Relate H_n to a biased random walk.)