STA447 Notes

Ian Zhang

1 Markov Chains

1.1 Markov Chain Definitions and Examples

Definition. A discrete-time, discrete-state, time-homogeneous Markov chain has 3 components:

1. State space S (finite or countably infinite)

2. Initial distribution $(\nu_i)_{i \in S}$ where $\nu_i = P(X_0 = i)$

3. Transitition probabilities $(p_{ij})_{i,j \in S}$ where

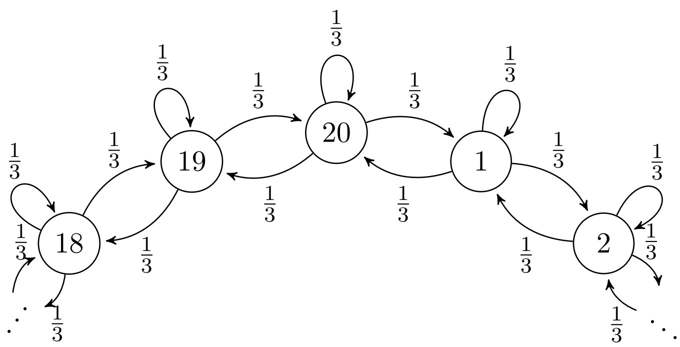
$$p_{ij} = P(X_{t+1} = j \mid X_t = i) = \frac{P(X_{t+1} = j, X_t = i)}{P(X_t = i)}$$

We now look at the most common Markov chains.

Frog Walk

Consider $S = \{1, \ldots, 20\}$. The Frog Walk is the Markov chain defined over S with initial probabilities defined as $\nu_{20} = 1$ and $\nu_i = 0$ for all $i \neq 20$, and transition probabilities p_{ij} defined as

$$p_{ij} = \begin{cases} \frac{1}{3} & \text{if } |i-j| \leq 1 \text{ or } |i-j| = 19 \\ 0 & \text{otherwise} \end{cases}$$



By how a Markov chain is structured, we have

$$P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = P(X_0 = i_0)P(X_1 = i_i \mid X_0 = i_0)P(X_2 = i_2 \mid X_1 = i_1, X_0 = i_0) \cdots P(X_n = i_n \mid X_0 = i_0, \cdots, X_{n-1} = i_{n-1})$$

What makes a MC "Markov" is the **Markov property**:

$$P(X_j = i_j \mid X_0 = i_0, \cdots, X_{j-1} = i_{j-1}) = P(X_j = i_j \mid X_{j-1} = i_{j-1})$$

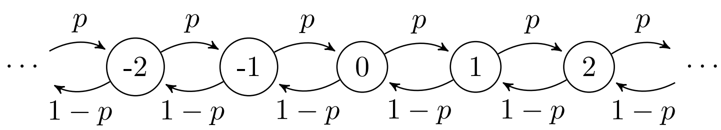
In other words, the state of the chain at time $t+1$ depends only on the state at time t. This property implies that

$$P(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \nu_{i_0} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}$$

Simple Random Walk

Let $0 < p < 1$ and suppose we repeatedly gamble \$1. Each time, we have a probability p of winning \$1 and a probability $1-p$ of losing the dollar. Let X_n represent the net gain after n bets. In this case, $S = \mathbb{Z}$ and the transition probabilities are

$$p_{ij} = \begin{cases} p & j = i+1 \\ 1-p & j = i-1 \\ 0 & \text{otherwise} \end{cases}$$



Ehrenfest's Urn

Suppose we have 2 urns and d balls in total. At each time t, we randomly select one ball and move it to the other urn. Let X_n be the number of balls in the left side at time n. In this case, $S = \{0, \ldots, d\}$ and the transition probabilities are

$$p_{i,i-1} = \frac{i}{d} \qquad p_{i,i+1} = \frac{d-i}{d}$$

1.2 Multi-Step Transitions

Let $\{X_n\}$ be a Markov chain with state space S, transition probabilities p_{ij}, and initial probabilities ν_i. By the Markov property, we know that

$$P(X_0 = i_0, X_1 = i_1, X_2 = i_2) = \nu_{i_0} p_{i_0 i_1} p_{i_1 i_2}$$

By the Law of Total Probability,

$$P(X_0 = i_0, X_2 = i_2) = \nu_{i_0} \sum_{i_1 \in S} p_{i_0 i_1} p_{i_1 i_2}$$

and

$$P(X_2 = i_2) = \sum_{\substack{i_0 \in S \\ i_1 \in S}} \nu_{i_0} p_{i_0 i_1} p_{i_1 i_2}$$

Let $m = |S|$, where $m \leq \infty$. Write $\nu = (\nu_1, \dots,)$ and

$$P := \begin{pmatrix} p_{11} & p_{12} & \cdots \\ p_{21} & p_{22} & \cdots \\ \vdots & \ddots & \vdots \end{pmatrix}$$

where P is a $m \times m$ matrix.
Define $\nu_i^{(2)} = P(X_2 = i)$. Then $\nu^{(2)} = \nu P^2$ and so on.

Definition. Let $p_{ij}^{(n)} = P(X_n = j \mid X_0 = i)$ for all $i, j \in S$. If $\nu_i = 1$ and $\nu_j = 0$ for all $j \neq i$, then νP^m is the m-step transition probability from state i. For the new chain, it has transition matrix $P^n = (p_{ij}^{(n)})_{i,j \in S}$.

Chapman-Kolmogorov Equations

$$(p_{ij}^{(m+n)})_{i,j \in S} = P^{m+n} = (p_{ij}^{(m)})_{i,j \in S} (p_{ij}^{(n)})_{i,j \in S} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}$$
$$(p_{ij}^{(m+s+n)})_{i,j \in S} = P^{m+s+n} = (p_{ij}^{(m)})_{i,j \in S} (p_{ij}^{(s)})_{i,j \in S} (p_{ij}^{(n)})_{i,j \in S} = \sum_{k \in S} \sum_{l \in S} p_{ik}^{(m)} p_{kl}^{(s)} p_{lj}^{(n)}$$

The Chapman-Kolmogorov inequality follows:

$$p_{ij}^{(m+n)} \geq p_{ik}^{(m)} p_{jk}^{(n)}$$

for any fixed $k \in S$ and

$$p_{ij}^{(m+s+n)} \geq p_{ik}^{(m)} p_{kl}^{(s)} p_{lj}^{(n)}$$

Basically, to compute $p_{ij}^{(n)}$, just compute P^n and observe the (i, j)th item in the resulting matrix.

1.3 Recurrence and Transcience

Definition. Let $N(i) :=$ total number of times for a Markov chain to visit i, so

$$N(i) = \sum_{t=1}^{\infty} I(x_t = i)$$

Let $f_{ij} := P(N(j) \geq 1 \mid X_0 = i) = P_i(N(j) \geq 1)$ be the probability that the Markov chain visits j eventually, starting from i.

In general, $P_i(N(i) \geq k) = (f_{ii})^k$ since

$$P_i(N(i) \geq k) = P_i(N(i) \geq k \mid N(i) \geq k-1) \cdot P_i(N(i) \geq k-1)$$

Let $\tau_i^{(k-1)}$ be the time step that hits i for the $k-1$th time. Then $X_0 = i$, $X_{\tau_i} = i$, $X_{\tau_i^{(k-1)}} = i$, and so on. Let τ be a hitting time for i. The above implies that

$$(X_\tau, X_{\tau+1}, \dots) \stackrel{d}{=} (X_0, X_1, \dots,)$$

It follows that

$$\begin{aligned} P_i(N(i) \geq k \mid N(i) \geq k-1) &= P_i(\tau_i^{(k)} < \infty \mid \tau_i^{(k-1)} < \infty) \\ &= P_i(\tau_i^{(1)} < \infty) \\ &= f_{ii} \end{aligned}$$

By induction, this implies $P_i(N(i) \geq k) = (f_{ii})^k$.

Corollary. $P_i(N(j) \geq k) = f_{ij}(f_{jj})^{k-1}$

Corollary. $E_i[N(j)] = \sum_{k=1}^{\infty} P_i(N(j) \geq k) = \begin{cases} \frac{f_{ij}}{1-f_{jj}} & \text{if } f_{jj} < 1 \\ 0 & \text{otherwise} \end{cases}$

Definition. A state i of a Markov chain is **recurrent** if $f_{ii} = 1$. State i is **transient** if $f_{ii} < 1$.

Corollary. A state i is recurrent if, and only if, $P_i(N(i) = \infty) = 1$.

Theorem (Recurrent State Theorem)**.** A state i is recurrent if, and only if, $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$.

Proof.

$$\begin{aligned} \sum_{n=1}^{\infty} p_{ii}^{(n)} = \sum_{n=1}^{\infty} P_i(X_n = i) &= \sum_{n=1}^{\infty} E_i[I(X_n = i)] \\ &= E_i\left[\sum_{n=1}^{\infty} I(X_n = i)\right] && \text{by Fubini-Tonelli} \\ &= E_i[N(i)] \\ &= \begin{cases} \infty & f_{ii} = 1 \\ \frac{f_{ii}}{1-f_{ii}} & f_{ii} < 1 \end{cases} \end{aligned}$$

■

Lemma (Borel-Cantelli)**.** Let $(E_i)_{i=1}^{\infty}$ be a sequence of events. If $\sum_{i=1}^{\infty} P(E_i) < \infty$, then

$$P\left((E_i)_{i=1}^{\infty} \text{ happens finite times}\right) = 1$$

Consider the simple random walk. Is state 0 recurrent?

- Need to check $\sum_{n=1}^{\infty} p_{00}^{(n)} = \infty$ as per the Recurrent State Theorem

For odd n, $p_{00}^{(n)}$ (obviously). For even n, we have

$$\begin{aligned} p_{00}^{(n)} &= P\left(\frac{n}{2} \text{ heads and } \frac{n}{2} \text{ losses in the first } n \text{ tosses}\right) \\ &= \binom{n}{\frac{n}{2}} p^{\frac{n}{2}}(1-p)^{\frac{n}{2}} \\ &= \frac{n!}{\left[\left(\frac{n}{2}\right)!\right]^2} p^{\frac{n}{2}}(1-p)^{\frac{n}{2}} \end{aligned}$$

Sterling's approximation: If n is large, then $n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$.
Thus,

$$p_{00}^{(n)} \approx \frac{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}}{\left[\left(\frac{n}{2e}\right)^{\frac{n}{2}} \sqrt{\pi n}\right]^2} p^{\frac{n}{2}}(1-p)^{\frac{n}{2}}$$
$$= [4p(1-p)]^{\frac{n}{2}} \sqrt{\frac{2}{\pi n}}$$

If $p = \frac{1}{2}$, then $4p(1-p) = 1$, so

$$\sum_{n=1}^{\infty} p_{00}^{(n)} \approx \sum_{n=2,4,\ldots} \sqrt{\frac{2}{\pi n}} = \left(\sqrt{\frac{2}{\pi}}\right) \sum_{n=2,4,\ldots} n^{-\frac{1}{2}} \to \infty$$

so state 0 is recurrent if $p = \frac{1}{2}$.
If $p \neq \frac{1}{2}$, then $4p(1-p) < 1$, then

$$\sum_{n=1}^{\infty} p_{00}^{(n)} \approx \sum_{n=2,4,\ldots} [4p(1-p)]^{\frac{n}{2}} \sqrt{\frac{2}{\pi n}} < \sum_{n=2,4,\ldots} [4p(1-p)]^{\frac{n}{2}} = \frac{4p(1-p)}{1-4p(1-p)} < \infty$$

This means that for the simple random walk, state 0 is recurrent iff $p = \frac{1}{2}$.
f-expansion:

$$f_{ij} = p_{ij} + \sum_{\substack{k \in S \\ k \neq j}} p_{ik} f_{kj}$$

1.3.1 Gambler's Ruin

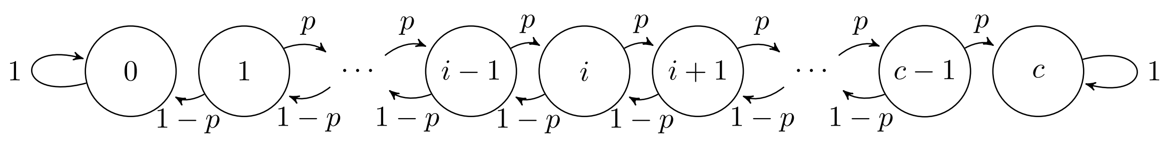
Suppose one starts with initial money amount $a \in \mathbb{N}$, and each round, the gain in money is captured by

$$\begin{cases} +1 & \text{with probability } p \\ -1 & \text{with probability } 1-p \end{cases}$$

The game stops if either one of the following happens:

1. All money is lost
2. The amount of money reaches c for some c

Let X_t be the amount of money the player has at time t. The state space is $S = \{0, \ldots, c\}$, the initial probabilities $\nu_a = 1$, $\nu_i = 0$ if $i \neq a$. The chain is captured as below:



Some of characteristics of the series:

- $f_{00} = f_{cc} = 1$, which means states 0 and c are recurrent (obviously since if we start at either then the game has already ended so the only option is to go back to the same state).

- $f_{ii} < 1$ for all $i \neq 0, c$, which means all the other states are transient

To compute the probability of losing all money given the player starts at state i, we compute f_{i0}.

$$\begin{aligned} f_{i0} &= p_{i0} + \sum_{\substack{k \in S \\ k \neq 0}} p_{ik} f_{k0} \\ &= \begin{cases} 1 - p + p f_{20} & i = 1 \\ (1-p) f_{(i-1)0} + p f_{(i+1)0} & i \geq 2 \end{cases} \\ &= (1-p) f_{(i-1)0} + p f_{(i+1)0} \end{aligned}$$

Obviously, $f_{c0} = 0$ since if the player starts at c, then he's already won.
Special case: $p = \frac{1}{2}$
If this is the case, then

$$f_{i0} = \frac{1}{2} f_{(i-1)0} + \frac{1}{2} f_{(i+1)0} = \frac{c-i}{c}$$

If $p \neq \frac{1}{2}$, then

$$f_{i0} = (1-p) f_{(i-1)0} + p f_{(i+1)0}$$

However,

$$\begin{aligned} f_{(i+1)0} - f_{(i-1)0} &= \frac{1}{p} f_{i0} + \frac{1-p}{p} f_{(i-1)0} - f_{i0} \\ &= \frac{1-p}{p} \left(f_{i0} - f_{(i-1)0} \right) \\ &= \cdots \\ &= \left(\frac{1-p}{p} \right)^i (f_{10} - f_{00}) \end{aligned}$$

So, given the player starts with a amount of money, then the probability the game ends with the player have nothing left is

$$f_{a0} = \begin{cases} \frac{\left(\frac{1-p}{p}\right)^c - \left(\frac{1-p}{p}\right)^a}{\left(\frac{1-p}{p}\right)^c - 1} & p \neq \frac{1}{2} \\ \frac{c-a}{c} & p = \frac{1}{2} \end{cases}$$

1.4 Communicating States and Irreducibility

Definition (Communication)**.** State i communicates with state j if $f_{ij} > 0$ (i.e.: if it's possible for the chain to visit j at least once starting from i). If so, then we say $i \to j$.

Definition (Irreducibility)**.** A Markov chain is irreducible if $i \to j$ for all $i, j \in S$.

- $i \leftrightarrow j$ if $i \to j$ and $j \to i$

Note that from above, Gambler's ruin is obviously reducible since $f_{0c} = f_{c0} = 0$.
Fact: If $i \leftrightarrow k$, then i is recurrent iff k is.

Corollary. For an irreducible MC, either

(a) All states are recurrent

(b) All states are transient

Lemma (Sum)**.** If $i \to k$, $l \to j$, and $\sum_{n=1}^{\infty} p_{kl}^{(n)} = \infty$, then $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$.

Proof. By definition, we know there exists m and r such that $p_{ik}^{(m)} > 0$ and $p_{lj}^{(r)} > 0$. By the Chapman-Kolmogorov inequality, we have

$$p_{ij}^{(m+s+r)} \geq p_{ik}^{(m)} p_{kl}^{(s)} p_{lj}^{(r)} > 0$$

Since each $p_{ij}^{(n)} \geq 0$, then

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} \geq \sum_{n=m+r+1}^{\infty} p_{ij}^{(n)} = \sum_{s=1}^{\infty} p_{ij}^{(m+s+r)} \geq p_{ik}^{(m)} p_{lj}^{(r)} \sum_{s=1}^{\infty} p_{kl}^{(s)} = \infty$$

■

Theorem (Finite Space)**.** An irreducible Markov chain on a finite state space always falls into case (a) of the above corollary.

Proof. Choose any state i. Then

$$\sum_{j \in S} \sum_{n=1}^{\infty} p_{ij}^{(n)} = \sum_{n=1}^{\infty} \sum_{j \in S} p_{ij}^{(n)} = \sum_{n=1}^{\infty} 1 = \infty$$

Since S is finite, there must exist some $j \in S$ such that $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$. ■

Lemma (Hit)**.** Define $H_{ij} = \{\text{MC hits state } i \text{ before returning to } j\}$. If j communicates with i with $j \neq i$, then $P_j(H_{ij}) > 0$.

Lemma (f)**.** If $j \to i$ and $f_{jj} = 1$, then $f_{ij} = 1$.

Proof. We know that $P_j(H_{ij}) > 0$ by the Hit Lemma. Then

$$P_j(H_{ij}) P_i(\text{never returns to } j) \leq P_j\left(\text{never returns to } j\right)$$

since one way to never return to j is to first visit i then never return to j. But since $f_{jj} = 1$ and by the Hit Lemma,

$$1 - f_{ij} = P_i(\text{never return to } j) = 0$$

which means $f_{ij} = 1$. ■

Lemma (Infinite Returns)**.** For irreducible MC, if recurrent, then for all $i, j \in S$,

$$P_i(N(j) = \infty) = 1$$

If transient, then for all $i, j \in S$,

$$P_i(N(j) = \infty) = 0$$

Proof. If recurrent, then $f_{ij} = f_{jj} = 1$ by the f-Lemma. So, for all k,

$$P_i(N(j) = k) = f_{ij}(f_{jj})^{k-1} = (1)(1)^{k-1} = 1$$

thus $P_i(N(j) = \infty) = \lim_{k\to\infty} P_i(N(j) = k) = \lim_{k\to\infty} 1 = 1$.
If transient, then

$$\lim_{k\to\infty} P_i(N(j) = k) = \lim_{k\to\infty} f_{ij}(f_{jj})^{k-1} = 0$$

since $f_{jj} < 1$. ■

Theorem (Recurrence Equivalences Theorem)**.** If a MC is irreducible, then the following are equivalent:

1. There exists $k, l \in S$ such that $\sum_{n=1}^{\infty} p_{kl}^{(n)} = \infty$
2. For all $i, j \in S$, $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$
3. There exists k such that $f_{kk} = 1$
4. For all i, $f_{ii} = 1$
5. For all i, j, $f_{ij} = 1$
6. There exists k, l such that $P_k(N(l) = \infty) = 1$
7. For all i, j, $P_i(N(j) = \infty) = 1$

All equivalences can be proven with the lemmas above. For transience, there's a similar theorem:

Theorem (Transience Equivalences Theorem)**.** If a MC is irreducible, then the following are equivalent:

1. For all $i, j \in S$, $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$
2. There exists i, j such that $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$
3. For all k, $f_{kk} < 1$
4. There exists i such that $f_{ii} < 1$
5. There exists i, j such that $f_{ij} < 1$
6. For all k, l, $P_k(N(l) = \infty) < 1$
7. There exists i, j such that $P_i(N(j) = \infty) < 1$

Proposition. There exists irreducible MC such that it's transient but there exists k, l such that $f_{kl} = 1$.

Take the simple random walk for example with $p > \frac{1}{2}$. We know that

$$p_{00}^{(n)} = \begin{cases} 0 & n \equiv 0 \mod 2 \\ \approx (4p(1-p))^{\frac{n}{2}} \sqrt{\frac{2}{\pi n}} & \text{otherwise} \end{cases}$$

so $\sum_{n=1}^{\infty} p_{00}^{(n)} < \infty$.
Fact: If i is transcient, j recurrent, then $j \not\to i$.

Proof. Suppose $j \to i$. Then

$$\begin{aligned} 0 &= P_j(\text{never return to } i) \\ &\geq P_j(\text{visit } i) P_i(\text{not return to } j) \end{aligned}$$

However, $P_j(\text{visit } i) > 0$, so $P_i(\text{not return to } j)$. Thus, $i \leftrightarrow j$, so since j is recurrent, then so is i, a contradiction. Thus, $j \not\to i$. ■

2 Markov Chain Convergence

2.1 Stationary Distributions

Suppose $\mu_j^{(n)} := P(X_n = j)$ with $\mu_j^{(n)} \to q_j$ for all states j. Then since

$$\mu_j^{(n+1)} \to q_j$$

$$\mu^{(n+1)} = \mu^{(n)} P$$

we have

$$q = qP$$

Definition. If π is a probability distribution on S, then π is stationary for a MC with transition probabilities (p_{ij}) if

$$\sum_{i \in S} \pi_i p_{ij} = \pi_j \quad \forall j \in S$$

If we write $\pi = \begin{bmatrix} \pi_1 & \pi_2 & \cdots & \end{bmatrix}^T$, then $\pi P = \pi$.

For example, take the frog walk and let π be a 1×20 vector with $\frac{1}{20}$ in all of its entries. Is π a stationary distribution? For all $j \in S$,

$$\sum_{i \in S} \pi_i p_{ij} = \frac{1}{20}\left(\frac{1}{3} + \frac{1}{3} + \frac{1}{3}\right) = \frac{1}{20} = \pi_j$$

thus π is stationary.

Definition (Doubly stochastic)**.** If $\sum_{i \in S} p_{ij} = 1$ in addition to $\sum_{j \in S} p_{ij} = 1$, then the MC is doubly stochastic.

Let π be a uniform distribution for a doubly stochastic chain on S, so $\pi_i = \frac{1}{|S|}$. Then,

$$\sum_{i\in S} \pi_i p_{ij} = \frac{1}{|S|}\sum_{i\in S} p_{ij} = \frac{1}{|S|} = \pi_j$$

Definition (Reversible)**.** A MC is reversible wrt distribution $\{\pi_i\}$ if $\pi_i p_{ij} = \pi_j p_{ji}$ for all $i, j \in S$.

Proposition. If a chain is reversible wrt π, then π is a stationary distribution.

Proof. Reversibility means $\pi_i p_{ij} = \pi_j p_{ji}$, thus

$$\sum_{i\in S} \pi_i p_{ij} = \sum_{i\in S} \pi_j p_{ji} = \pi_j(1) = \pi_j$$

■

Fact: There exists a MC P with stationary distribution π such that P is not reversible wrt π.

M-test

Consider a sequence $\{x_{nk}\}_{n,k\in\mathbb{N}}$. Suppose $\lim_{n\to\infty} x_{nk}$ exists for all $k \in \mathbb{N}$, and $\sum_{k=1}^{\infty} \sup_{n\geq 1}|x_{nk}| < \infty$. Then,

$$\lim_{n\to\infty}\sum_{k=1}^{\infty} x_{nk} = \sum_{k=1}^{\infty}\lim_{n\to\infty} x_{nk}$$

Proposition (Vanishing Probabilities)**.** If $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ for all $i, j \in S$, then a stationary distribution does not exist.

Proof. Suppose π is stationary, so $\pi_j = \sum_{i\in S}\pi_i p_{ij}^{(n)}$ for any n, thus

$$\pi_j = \lim_{n\to\infty}\pi_j = \lim_{n\to\infty}\sum_{i\in S}\pi_i p_{ij}^{(n)}$$

Notice that

$$\sum_{i\in S}\sup_{n\geq 1}|\pi_i p_{ij}^{(n)}| \leq \sum_{i\in S}\pi_i = 1 < \infty$$

so by the M-test,

$$\pi_j = \lim_{n\to\infty}\pi_j = \lim_{n\to\infty}\sum_{i\in S}\pi_i p_{ij}^{(n)} = \sum_{i\in S}\lim_{n\to\infty}\pi_i p_{ij}^{(n)} = \sum_{i\in S} 0 = 0$$

which is a contradiction since $\sum_{j\in S}\pi_j = 0$. ■

Lemma (Vanishing)**.** If a MC has some $k, l \in S$ such that $\lim_{n\to\infty} p_{kl}^{(n)} = 0$, then for all $i, j \in S$ such that $k \to i$ and $j \to l$, then $\lim_{n\to\infty} p_{ij}^{(n)} = 0$.

Proof. There exists $r, s \in \mathbb{N}$ such that $p_{ki}^{(r)} > 0$ and $p_{jl}^{(s)} > 0$. By Chapman-Kolmogorov,

$$p_{kl}^{(r+n+s)} \geq p_{ki}^{(r)} p_{ij}^{(n)} p_{jl}^{(s)}$$

Thus,

$$p_{ij}^{(n)} \leq \frac{p_{kl}^{(r+n+s)}}{p_{ki}^{(r)} p_{jl}^{(s)}}$$

By the assumption, we know that $\lim_{n\to\infty} \frac{p_{kl}^{(r+n+s)}}{p_{ki}^{(r)} p_{jl}^{(s)}} = \infty$, thus since $p_{ij}^{(n)} \geq 0$, by the Squeeze Theorem, $\lim_{n\to\infty} p_{ij}^{(n)} = 0$. ∎

Corollary. For an irreducible MC, either

(i) $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ for all $i, j \in S$ (the MC is *transient*)

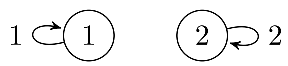
(ii) $\lim_{n\to\infty} p_{ij}^{(n)} \neq 0$ for all $i, j \in S$ (the MC is *recurrent*)

Corollary. If an irreducible MC has $\lim_{n\to\infty} p_{kl}^{(n)} = 0$ for some $k, l \in S$, then it does not a stationary distribution.

Proof. By the above corollary, if there exists k, l such that $\lim_{n\to\infty} p_{kl}^{(n)} = 0$, then since the chain is irreducible, we have $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ for all $i, j \in S$. By the Vanishing Probabilities proposition, the chain does not have a stationary distribution. ∎

2.2 Obstacles to Convergence

Let $S = \{1, 2\}$, $\nu_1 = 1$, and $(p_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.



Let $\pi_1 = \pi_2 = \frac{1}{2}$, so $\{\pi_i\}$ is stationary. However,

$$\lim_{n\to\infty} P(X_n = 1) = 1 \neq \frac{1}{2} = \pi_1$$

so the chain does not converge to stationarity.

Definition (Period)**.** The period of a state i is the gcd of $\{n \geq 1 : p_{ii}^{(n)} > 0\}$. If the period of every state i is 1, then the MC is aperiodic. Otherwise, it is periodic.

Its entirely possible to have a MC be aperiodic despite all $p_{ii} = 0$. Take $S = \{1, 2, 3\}$ and consider the transition probabilities

$$(p_{ij}) = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

Clearly, we can go $1 \to 2 \to 1$ or $1 \to 2 \to 3 \to 2 \to 1$, and so on, so the period of state 1 is $\gcd\{2, 3, \ldots\} = 1$. Similarly, the period of states 2 and 3 are each 1, so the chain is aperiodic. However, $p_{ii} = 0$ for all $i \in S$.

- Since $\gcd\{1, \ldots, \} = 1$, then if $p_{ii} > 0$, state i has period 1

- Since $\gcd\{n, n+1, \ldots, \} = 1$, then if both $p_{ii}^{(n)} > 0$ and $p_{ii}^{(n+1)} > 0$, state i has period 1

Take the frog walk for example. We know that $p_{ii} = \frac{1}{3}$ for all i, so the chain is aperiodic.
For the simple random walk, we can only return to a state after an even number of moves, thus the period of each state is 2.

Lemma (Equal Periods)**.** If $i \leftrightarrow j$, then the periods of i and j are equal.

Proof. Let t_i and t_j be the periods of states i and j respectively. We know there exists $r, s \in \mathbb{N}$ such that $p_{ij}^{(r)} > 0$ and $p_{ji}^{(s)} > 0$, thus by Chapman-Kolmogorov,

$$p_{ii}^{(r+s)} \geq p_{ij}^{(r)} p_{ji}^{(s)}$$

thus $t_i \mid r + s$. Suppose for some n that $p_{jj}^{(n)} > 0$. Then by Chapman-Kolmogorov again,

$$p_{ii}^{(r+n+s)} \geq p_{ij}^{(r)} p_{jj}^{(n)} p_{ji}^{(s)}$$

thus $t_i \mid r + n + s$. Since $t_i \mid r + s$, then we must have $t_i \mid n$, so t_i is a common divisor of the set $A = \{n \geq 1 : p_{jj}^{(n)} > 0\}$. But since $t_j = \gcd(A)$, then t_i and t_j divide each other, which implies $t_i = t_j$. ■

2.3 Markov Chain Convergence Theorem

Theorem (Markov Chain Convergence)**.** If a MC is irreducible, aperiodic, and has a stationary distribution π, then for all $i, j \in S$,

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$$

and for any initial distribution $\{\nu_i\}$,

$$\lim_{n\to\infty} P(X_n = j) = \pi_j$$

Theorem (Stationary Recurrence)**.** If a MC is irreducible and has a stationary distribution π, then it's recurrent.

Proposition. If state i is aperiodic and $f_{ii} > 0$, then there exists some $n_0(i) \in \mathbb{N}$ such that $p_{ii}^{(n)} > 0$ for all $n \geq n_0(i)$.

Proof. Let $A = \{n \geq 1 : p_{ii}^{(n)} > 0\} \neq \emptyset$ since $f_{ii} > 0$. If $m, n \in A$, then by Chapman-Kolmogorov, $p_{ii}^{(m+n)} \geq p_{ii}^{(m)} p_{ii}^{(n)} > 0$ thus $m + n \in A$ so A satisfies additivity. Citing Bézout's Identity completes the proof. ■

Corollary. If a MC is irreducible and aperiodic, then for all $i, j \in S$, there exists some $n_0(i, j) \in \mathbb{N}$ such that for all $n \geq n_0(i, j)$, $p_{ij}^{(n)} > 0$.

Lemma (Markov Forgetting)**.** If a MC is irreducible and aperiodic and has stationary distribution $\{\pi_i\}_i$, then for all $i, j, k \in S$,

$$\lim_{n\to\infty} \left| p_{ik}^{(n)} - p_{jk}^{(n)} \right| = 0$$

2.3.1 Proof of Markov Chain Convergence Theorem

For all $i, j \in S$, by definition of a stationary distribution, we have

$$\left|p_{ij}^{(n)} - \pi_j\right| = \left|\sum_{k\in S} \pi_k \left(p_{ij}^{(n)} - p_{kj}^{(n)}\right)\right| \leq \sum_{k\in S} \pi_k \left|p_{ij}^{(n)} - p_{kj}^{(n)}\right|$$

By the Markov Forgetting Lemma, we have

$$\lim_{n\to\infty} \pi_k \left|p_{ij}^{(n)} - p_{kj}^{(n)}\right| = 0$$

Furthermore, $\sup_{n\geq 1} \left|p_{ij}^{(n)} - p_{kj}^{(n)}\right| \leq 2$ which implies

$$\sum_{k\in S} \sup_{n\geq 1} \pi_k \left|p_{ij}^{(n)} - p_{kj}^{(n)}\right| \leq \sum_{k\in S} 2\pi_k = 2 < \infty$$

Thus, by the M-test,

$$\lim_{n\to\infty} \left|p_{ij}^{(n)} - \pi_j\right| \leq \lim_{n\to\infty} \sum_{k\in S} \pi_k \left|p_{ij}^{(n)} - \pi_j\right| = \sum_{k\in S} \lim_{n\to\infty} \pi_k \left|p_{ij}^{(n)} - \pi_j\right| = \sum_{k\in S} 0 = 0$$

which implies

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$$

as required. For any initial distribution, $\{\nu_i\}$, we have

$$\lim_{n\to\infty} P(X_n = j) = \lim_{n\to\infty} \sum_{i\in S} P(X_n = j, X_0 = i) = \lim_{n\to\infty} \sum_{i\in S} \nu_i p_{ij}^{(n)} = \sum_{i\in S} \nu_i \lim_{n\to\infty} p_{ij}^{(n)} = \sum_{i\in S} \nu_i \pi_j = \pi_j$$

2.4 Periodic Convergence

Theorem (Periodic Convergence)**.** Suppose a MC is irreducible with period $b \geq 2$ and stationary distribution $\{\pi_i\}$. Then for all $i, j \in S$,

$$\lim_{n\to\infty} \frac{1}{b}[p_{ij}^{(n)} + \cdots + p_{ij}^{(n+b-1)}] = \pi_j$$

and

$$\lim_{n\to\infty} \frac{1}{b} \sum_{i=0}^{b-1} P(X_{n+i} = j) = \pi_j$$

and

$$\lim_{n\to\infty} \frac{1}{b} P[X_n = j \text{ or } X_{n+1} = j \text{ or } \cdots \text{ or } X_{n+b-1} = j] = \pi_j$$

Corollary (Cesàro Sum)**.** For any irreducible MC with stationary distribution $\{\pi_j\}$, for all $i, j \in S$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} p_{ij}^{(t)} = \pi_j$$

Corollary. An irreducible MC has at most one stationary distribution.

Lemma (Cyclic Decomposition)**.** If a MC has period $b \geq 2$, then $S = S_0 \cup S_1 \cup \ldots \cup S_{b-1}$ for $S_i \cap S_j = \emptyset$ for all $i \neq j$, where if $i \in S_r$, then $\{j \in S : p_{ij} > 0\} \subseteq S_{(r+1) \bmod b}$. Furthermore, $P^{(b)}$ restricted to S_0 forms an irreducible and aperiodic transition matrix.

2.5 Mean Recurrences Times

The mean recurrence time of a state i is $m_i = E_i(\inf\{n \geq 1 : X_n = i\}) = E_i(T_i)$.

- If the chain never returns to i, then $T_i = \infty$
  - If i is transient, then $m_i = \infty$
  - If $m_i < \infty$, then i is recurrent

Definition. A state is positive recurrent if $m_i < \infty$, null recurrent if recurrent but $m_i = \infty$.

Theorem. For an irreducible MC, either

(a) $m_i < \infty$ for all $i \in S$ and there exists a unique stationary distribution given by $\pi_i = \frac{1}{m_i}$

(b) $m_i = \infty$ for all $i \in S$ and there does not exist a stationary distribution

2.6 Stationary Measures

A stationary measure is a measure μ such that $\mu = \mu P$.

Theorem. For any irreducible and recurrent MC, for $i_0 \in S$,

$$\mu_{i_0}(y) = \sum_{n=0}^{\infty} P_{i_0}(X_n = y, T_{i_0} > n)$$

defines a stationary measure μ_{i_0} such that $0 < \mu_{i_0}(y) < \infty$.

If $E_i(T_i) < \infty$, we can normalize and define stationary distribution

$$\pi_i = \frac{\mu_i}{E_i(T_i)}$$

Corollary. If MC is irreducible and there exists i such that i is positive recurrent, then a stationary distribution exists.

Theorem. Suppose X is irreducible and recurrent. Let $N_n(i) := \sum_{t=1}^{n} I(X_t = i)$. Then

$$\frac{N_n(i)}{n} \to \frac{1}{E_i(T_i)} \quad a.s.$$

Corollary. If a MC is irreducible with stationary distribution π, then

$$\pi_i = \frac{1}{E_i(T_i)}$$

for all $i \in S$.

3 Martingales

3.1 Martingale Definitions

Definition. A sequence $(X_n)_{n\geq 0}$ is a martingale if

$$E(X_{n+1} \mid X_1, \dots, X_n) = X_n$$

For a discrete sequence, the sequence is a martingale if $E(X_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) = i_n$. If $(X_n)_n$ is a Markov chain, then

$$\begin{aligned} E(X_{n+1} \mid X_0 = i_0, \dots, X_n = i_n) &= \sum_{j \in S} jP(X_{n+1} = j \mid X_0 = i_0, \dots, X_n = i_n) \\ &= \sum_{j \in S} jP(X_{n+1} = j \mid X_n = i_n) \\ &= \sum_{j \in S} jp_{i_n,j} \\ &= i_n \text{ if } (X_n)_n \text{ is a martingale} \end{aligned}$$

So, a Markov chain is a martingale if $\sum_{j \in S} jp_{ij} = i$ for all $i \in S$.
Alternate notation: Let $\{\mathscr{F}_n\}$ denote an increasing collection of information (so $\mathscr{F}_m \subseteq \mathscr{F}_n$ if $m < n$). Then if X_n is measurable wrt $\mathscr{F}_n$, $(X_n)_n$ is a martingale if $\forall m < n$,

$$E(X_n \mid \mathscr{F}_m) = X_m$$

If $(X_n)_n$ is a martingale, then

$$E(X_{n+1}) = E[E[X_{n+1} \mid \mathscr{F}_n]] = E(X_n)$$

This implies that $E(X_n) = E(X_0)$ for all n.

3.2 Stopping Times and Optional Stopping

Definition. $T \in \mathbb{Z}_{\geq 0}$ is a stopping time if the event $\{T = n\}$ is determined by $X_0, \dots, X_n$.

- $\mathbb{1}_{T=n} = \varphi(X_0, \dots, X_n)$

Optional Stopping Lemma: If $(X_n)_n$ is a martingale and T a bounded stopping time (i.e.: $\exists M < \infty$ s.t. $P(T < M) = 1$), then $E(X_T) = E(X_0)$.

Proof.

$$\begin{aligned} E(X_T) - E(X_0) &= E\left[\sum_{k=1}^{T} (X_k - X_{k-1})\right] \\ &= E\left[\sum_{k=1}^{M} (X_k - X_{k-1})\mathbb{1}_{k \leq T}\right] \\ &= \sum_{k=1}^{M} E[(X_k - X_{k-1})\mathbb{1}_{k \leq T}] \end{aligned}$$

For each k, since $(X_n)_n$ is a martingale, then

$$\begin{aligned} E[(X_k - X_{k-1})\mathbb{1}_{k \leq T}] &= E[(X_k - X_{k-1})(1 - \mathbb{1}_{k-1 \geq T})] \\ &= E[(1 - \mathbb{1}_{k-1 \leq T})E(X_k - X_{k-1} \mid \mathscr{F}_{k-1})] \\ &= 0 \end{aligned}$$

■

Optional Stopping Theorem: If $(X_n)_n$ is a martingale with stopping time T and $P(T < \infty) = 1$, $E[|X_T|] < \infty$, and if $\lim_{n\to\infty} E(X_n \mathbb{1}_{T>n}) = 0$, then $E(X_T) = E(X_0)$.

Corollary. If $(X_n)_n$ is a martingale with stopping time T, which is "bounded up to time T" (i.e.: $\exists M < \infty$ s.t. $P(|X_n|\mathbb{1}_{T\geq n} \leq M) = 1$ for all n), and $P(T < \infty) = 1$, then $E(X_T) = E(X_0)$.

3.3 Uniform Integrability

For a fixed X,

$$E[|X|\mathbb{1}_A] = E[|X|\mathbb{1}_{A\cap\{|X|>K\}}] + E[|X|\mathbb{1}_{A\cap\{|X|\leq K\}}] \leq E[|X|\mathbb{1}_{|X|>K}] + KP(A)$$

If F is the cdf of $|X|$, then

$$\lim_{K\to\infty} E[|X|\mathbb{1}_{|X|>K}] = \lim_{K\to\infty} \int_K^\infty |x| dF(x) = 0$$

Definition. A sequence of random variables $(X_n)_n$ is uniform integrable if

$$\forall \varepsilon > 0, \exists K \text{ s.t. } \forall n, E[|X_n|\mathbb{1}_{|X_n|>K}] < \varepsilon$$

We can use uniform integrability to restate the Optional Stopping Theorem.
Optional Stopping Theorem V2: If $(X_n)_n$ is uniform integrable, T a stopping time with $T < \infty$ almost surely and $E[|X_T|] < \infty$, then $E(X_T) = E(X_0)$.
Fact: If there exists some $C < \infty$ such that $E(X_n^2) < C$ for each n, then the sequence is uniform integrable.
For example, let

$$Z_j = \begin{cases} 1 & w.p.\ \frac{1}{2} \\ -1 & w.p.\ \frac{1}{2} \end{cases}$$

and define

$$X_n = \sum_{j=1}^n \frac{1}{j} Z_j$$

where the Z_j are i.i.d., thus $E(X_n) = 0$. Then

$$E(X_n^2) = \text{Var}(X_n) = \sum_{j=1}^n \text{Var}\left(\frac{1}{j} Z_j\right) = \sum_{j=1}^n \frac{1}{j^2} \leq \sum_{j=1}^\infty \frac{1}{j^2} < \infty$$

so X_n is uniform integrable.

3.4 Wald's Theorem

Wald's Theorem: If $X_n = \sum_{i=1}^n Z_i$ where Z_i are i.i.d. with finite mean m, T is a stopping time for X_n such that $E(T) < \infty$, then $E(X_T) = mE(T)$.

By the optional stopping lemma, $E[X_{n\wedge T} - m(n \wedge T)] = 0$ for all n (where $n \wedge T = \min(n, T)$). This implies that

$$\lim_{n\to\infty} E(n \wedge T) = E(T) < \infty$$

thus

$$\begin{aligned}|E(X_{n\wedge T})-E(X_T)| &\leq E\left[\sum_{m=n+1}^{T}|Z_m|\mathbb{1}_{T>n}\right]\\ &= E\left[\sum_{m=n+1}^{\infty}|Z_m|\mathbb{1}_{T\geq m}\right]\\ &= \sum_{m=n+1}^{\infty}E[|Z_m|\mathbb{1}_{T\geq m}]\end{aligned}$$

Since the event $\{T\geq m\}$ is determined by $Z_1,\ldots,Z_{m-1}$, and thus independent with Z_m, then

$$E[|Z_m|\mathbb{1}_{T\geq m}]=E[|Z_m|]P(T\geq m)$$

Thus,

$$\sum_{m=n+1}^{\infty}E[|Z_m|\mathbb{1}_{T\geq m}]\leq E[|Z_1|]\sum_{m=n+1}^{\infty}P(T\geq m)$$

However, since $\sum_{m=1}^{\infty}P(T\geq m)=E(T)<\infty$, then $\sum_{m=n+1}^{\infty}P(T\geq m)$ converges to 0. So,

$$|E(X_{n\wedge T})-E(X_T)|\to 0$$

3.5 Martingale Convergence

Martingale Convergence Theorem: For a martingale $(M_n)_n$, if $E[|M_n|]\leq C<\infty$ almost surely, then $M_n\to M_\infty$ almost surely where M_∞ is some random variable.

- Note that $E[|M_n|]\leq C$ can be replaced by one of two options:
  1. $M_n\geq C$ for some $C\in\mathbb{R}$
  2. $M_n\leq C$ for some $C\in\mathbb{R}$

For example, let $(X_n)_n$ be a simple random walk and define stopping time $T=\inf\{t\geq 0: X_t=-1\}$ with $T<\infty$ almost surely. Define $Y_n=X_{n\wedge T}$, so Y_n is a martingale. Furthermore, we know that $Y_n\geq -1$ almost surely by definition of T. So, by the Martingale Convergence Theorem, Y_n converges almost surely to some Y_∞. In fact, $Y_\infty=-1$ almost surely. Note, however, that $E[Y_\infty]\neq E[Y_0]$.

Fact: If $(X_n)_{n\geq 0}$ is a martingale and is uniform integrable with $E[|X_\infty|]<\infty$, then $E[X_\infty]=E[X_0]$.

- Uniform integrable suffices as an assumption for the martingale convergence theorem since there must exist some K_1 such that

$$E[|X_n|]\leq E[|X_n|\mathbb{1}_{|X_n|>K_1}]+E[|X_n|\mathbb{1}_{|X_n|\leq K_1}]\leq K_1+1<\infty$$

Suppose $(X_n)_n$ is an irreducible Markov chain. A function f is **harmonic** if

$$f(x)=\sum_{y\in S}p(x,y)f(y)$$

When f is integrable, $(f(X_n))_n$ forms a martingale. Let T be the hitting time of $Z \in S$. Then, $M_n := f(X_{n\wedge T})$ is also a martingale.

Fact: If f is harmonic and bounded, and P (the Markov transition kernel) is a recurrent, then f is constant.

Proof. Since recurrent, then $P(T < \infty) = 1$. As defined above, $(M_n)_{n\geq 0}$ is a uniformly bounded martingale, so $M_n \to M_\infty$ almost surely. So,

$$E[M_n] = E[M_\infty] = f(Z)$$

thus $f(x) = E[M_0]$, which is constant. ■

In the case of transience, fix $Z \in S$. Then

$$f(x) := \begin{cases} P_x(T_z < \infty) = f_{xz} & x = z \\ 1 & x = z \end{cases}$$

From the f-expansion, f is harmonic.

3.6 Branching Processes

Let X_n be the number of individuals who are present at time n. Start with $X_0 = a$ for some $0 < a < \infty$. At time n, each of the X_n individuals creates a random number of offspring to appear at time $n+1$. The number of offspring of each individual is $\overset{iid}{\sim} \mu$ where μ is the offspring distribution on $\{0, 1, \ldots, \}$. Thus, $X_{n+1} = Z_{n,1} + Z_{n,2} + \cdots + Z_{n,X_n}$ where $(Z_{n,i})_{i=1}^{X_n} \overset{iid}{\sim} \mu$. X_n is a Markov chain.

$$E[X_{n+1} \mid \mathscr{F}_n] = \sum_{i=1}^{X_n} E[Z_{n,i} \mid \mathscr{F}_n] = X_n E_\mu[Z] =: X_n m$$

Thus, $Y_n = m^{-1} X_n$ forms a martingale.
If $m < 1$, then $E[X_n] = m^n E[X_0] \to 0$ as $n \to \infty$, so $X_n \xrightarrow{p} 0$. This implies that extinction is certain if $m < 1$.
If $m > 1$, then $E[X_n] = m^n E[X_0] \to \infty$. This means that $P(X_n \to \infty) > 0$, so the probability of flourishing exists. Assuming $\mu_0 > 0$, $P(X_n \to 0) > 0$ still. So, $P(X_n \to \infty) \leq 1 - P(X_n \to 0) < 1$. Thus, at $m > 1$, extinction and flourishing are both possible.
If $m = 1$, assuming $\mu_1 < 1$, then $E(X_0) = E(X_n) = a$, so X_n is a martingale. By the Martingale Convergence Theorem, $X_n \to X_\infty$ almost surely since $X_n \geq 0$. Since $X_n \in \mathbb{Z}$, then there exists some $T < \infty$ such that $X_n = X_\infty$ for all $n \geq T$. Thus, the only logical solution is $X_n \to 0$ almost surely.

4 Brownian Motion

4.1 Brownian Motion Definitions

Definition. A continuous process $(B_t)_{t\geq 0}$ is Brownian motion if it satisfies the following:

1. $B_0 = 0$
2. $B_t \sim \mathcal{N}(0, t)$
3. Independent normal increments: For $t > s$, $B_t - B_s \sim \mathcal{N}(0, t-s)$ and is independent with B_s
4. $\text{Cov}(B_t, B_s) = \min(s, t)$
5. The mapping $t \mapsto B_t$ is continuous

Fact: Brownian motion is Markov.
This is due to the **strong Markov property:** $(B_t - B_s)_{t \geq s}$ is a Brownian motion and independent of its past.

4.1.1 Stopping Times in a Continuous-Time Case

Definition. T is a stopping time if the event $\{T \leq t\}$ is determined by $\mathscr{F}_t$ (i.e.: by $(B_s)_{0 \leq s \leq t}$)

Fact: If T is a stopping time with $P(T > \infty) < 1$, then $(B_{t+T} - B_t)_t$ is a Brownian motion independent of $(B_s)_{0 \leq s \leq t}$.

4.2 Reflection principle

Let $T = \inf\{t : B_t = 1\}$. What is $P(T \leq 1)$?
We know at time $t = 1$, either

- $B_t \leq 1$
- $B_t > 1$ (has already hit 1)

This means

$$P(B_1 \geq 1) = P(T \leq 1)P(B \geq 1 \mid T \leq 1)$$

$P(B_1 \geq 1)$ is computable, and by the strong Markov property, $(B_t - 1)_{t \geq T}$ is a Brownian motion. Conditionally on $(B_t)_{0 \leq t \leq T}$,

$$P(B_1 - 1 \geq 0 \mid (B_t)_{0 \leq t \leq T}, T \leq 1) = P(N(0, 1-t) \geq 0) = \frac{1}{2}$$

Thus,

$$P(T \leq 1) = 2P(B_1 \geq 1) = 2\int_1^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

since $B_1 \sim \mathcal{N}(0, 1)$. This implies the **Reflection Principle**.
Reflection Principle: For a stopping time $T_a = \inf\{t : B_t = a\}$,

$$P(T_a \leq t) = 2P(B_t \geq a) = 2\int_a^\infty \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}\, dx = P\left(\max_{0 \leq s \leq t} B_s \geq a\right)$$

4.3 Brownian Motion as a Martingale

Fact: Brownian motion is a martingale.
To see this, we know $E(|B_t| < \infty)$ and $E[B_t \mid \mathscr{F}_s] = B_s$ for all $0 \leq s < t$ because $B_t = B_s + (B_t - B_s)$, so $B_t \mid B_s \sim \mathcal{N}(B_s, t-s)$.
Using this fact, we can also apply the Optional Stopping Theorem and Martingale Convergence Theorem.
For example, let $a, b > 0$ and $T = \inf\{t \geq \theta : B_t = -a \text{ or } B_t = b\}$. What is $P(B_t = -a)$?
We know $(B_t)_t$ is a martingale and is bounded up to time T because $|B_t|\mathbb{1}_{T \geq t} \leq \max(a, b)$. Thus, by the Optional Stopping Theorem, $E[B_T] = E[B_0] = 0$. However, we also know that $E[B_T] = -aP(B_T = -a) + bP(B_T = b)$, and $P(B_T = -a) + P(B_T = b) = 1$, thus $P(B_T = -a) = \frac{b}{a+b}$.
What is $E(T)$?
Let $Y_t = B_t^2 - t$. For $0 \leq s \leq t$

$$E[B_t^2 \mid \mathscr{F}_t] = E[B_t^2 + (B_t - B_s)^2 + 2B_s(B_t - B_s) \mid \mathscr{F}_s] = B_s^2 + (t - s)$$

thus Y_t is a martingale bounded up to time T. By the Optional Stopping Theorem,

$$0 = E[B_T^2 - T] = -E[T] + P(B_T = -a)a^2 + P(B_T = b)b^2$$

so $E(T) = ab$.

4.4 Zero Set of Brownian Motion

Scaling Properties: If $(B_t)_t$ is a standard Browninan motion, then

1. If $a > 0$ and $Y_t = a^{-\frac{1}{2}} B_{at}$, then Y_t is also a Brownian motion
2. If $Y_t = tB_{\frac{1}{t}}$, then Y_t is also a Brownian motion

5 Stochastic Calculus

5.1 Stochastic Integration With Respects to Brownian Motion

The goal: To define

$$Z_t = \int_0^t Y_s dB_s$$

where B_s is a Brownian motion. Think of B_s as a gambling game and Y_s is the amount bet at time s. There are 3 important properties to know for Z_t:

1. Z_t is linear
2. Martingale Property: Z_t is a martingale. In particular, $E(Z_t) = 0$.
3. Itô's isometry: $E[Z_t^2] = \int_0^t E[Y_s^2]\, ds$

5.2 Itô's Calculus

Note that for stochastic integrals the FTC does **not** hold:

$$\int_0^t B_s dB_s \neq \frac{1}{2}(B_t^2 - B_0^2)$$

since the LHS has expectation 0 but RHS has expectation $\frac{t^2}{2}$.
Itô's Formula I: If $f \in C^2$ and B_t is a standard Brownian motion, then

$$f(B_t) - f(B_0) = \int_0^t f'(B_s)\, dB_s + \frac{1}{2}\int_0^t f''(B_s)\, ds$$

Itô's Formula implies

$$\left(f(B_t) - f(B_0) - \frac{1}{2}\int_0^t f''(B_s)\, ds \right)_{t \geq 0}$$

is a martingale.

5.2.1 Extensions of Itô's Formula

Suppose

$$dZ_t = X_t dt + Y_t dB_t \tag{1}$$

Then for some $f \in C^2$,

$$df(Z_t) = f'(Z_t)dZ_t + \frac{1}{2}f''(Z_t)Y_t^2 dt = f'(Z_t)X_t dt + f'(Z_t)Y_t dt + \frac{1}{2}f''(Z_t)Y_t^2 dt$$

Definition. The quadratic variation of Z_t is

$$\langle Z \rangle_t = \lim_{n\to\infty} \sum_{j=0}^{n-1} \left[Z_{\frac{j+1}{n}t} - Z_{\frac{j}{n}t} \right]^2$$

Thus,

$$\langle Z \rangle_t = \int_0^t Y_s^2\, ds$$

This implies that

$$\int_0^t Y_s^2 f''(Z_s)\, ds = \int_0^t f''(Z_s) d\langle Z \rangle_s$$

Thus, we get
Itô's Formula II: If $f \in C^2$ and Z_t satisifes (1), then

$$\begin{aligned} f(Z_t) - f(Z_0) &= \int_0^t f'(Z_s)dZ_s + \frac{1}{2}\int_0^t f''(Z_s)d\langle Z \rangle_s \\ &= \int_0^t f'(Z_s)Y_s dB_s + \int_0^t \frac{1}{2}f''(Z_s)Y_s^2 + f'(Z_s)X_s\, ds \end{aligned}$$

In differential form,

$$df(Z_t) = f'(Z_t)Y_t dB_t + f'(Z_t)X_t dt + \frac{1}{2}f''(Z_t)Y_t^2 dt$$

The **product rule** is

$$d(Z_t^{(1)} Z_t^{(2)}) = Z_t^{(1)} dZ_t^{(2)} + Z_t^{(2)} dZ_t^{(1)} + d\langle Z^{(1)}, Z^{(2)} \rangle_t$$

where

$$\langle Z^{(1)}, Z^{(2)} \rangle_t = \lim_{n\to\infty} \sum_{j=0}^{n-1} \left(Z^{(1)}_{\frac{j+1}{n}t} - Z^{(1)}_{\frac{j}{n}t} \right) \left(Z^{(2)}_{\frac{j+1}{n}t} - Z^{(2)}_{\frac{j}{n}t} \right)$$

If both $Z_t^{(1)}$ and $Z_t^{(2)}$ satisfy (1) respectively, then

$$\langle Z^{(1)}, Z^{(2)} \rangle_t = \int_0^t Y_s^{(1)} Y_s^{(2)}\, ds$$

Itô's Formula III: If $f(t,x)$ is C^1 in t, C^2 in x, and Z_t satisfies (1), then

$$\partial f(t, Z_t) = \partial_t f(t, Z_t)dt + \partial_x f(t, Z_t) dZ_t + \frac{1}{2}\partial_x^2 f(t, Z_t) d\langle Z \rangle_t$$

6 Other Processes

6.1 Poisson Processes

Consider the positive real line and suppose a sequence of marked points exist on the line. Let τ_n be the time between the $n-1$ and nth mark and let $T_n = \tau_1 + \ldots + \tau_n$ be the arrival time of the nth mark. Let $N(t)$ be the number of marked points on $[0,t]$, thus $N(t) = \max\{n \geq 0 : T_n \leq t\}$. If we suppose $\tau_n \sim \text{Exponential}(\lambda)$, then $N(t)$ is a *Poisson process*.
If $X \sim \text{Binomial}(n,p)$ where $p = \frac{\lambda}{n}$, then as $n \to \infty$,

$$P(X = k) = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \to \frac{\lambda^k}{k!} e^{-\lambda}$$

As a fact, $N(t) \sim \text{Poisson}(\lambda t)$.

Proof. Let $S_n = \sum_{t=1}^n T_t$ and let f_{S_n} be the density of S_n. Since the $T_1, \ldots, T_n \overset{iid}{\sim} \text{Exponential}(\lambda)$, then

$$f_{S_n}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \quad \forall t$$

So,

$$\begin{aligned} P(N(t) = n) &= P(S_n \leq t \leq S_{n+1}) \\ &= \int_0^t f_{S_n}(s) P(T_{n+1} > t - s)\, ds \\ &= \int_0^t \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda(t-s)}\, ds \\ &= e^{-\lambda t} \frac{(\lambda t)^n}{n!} \end{aligned}$$

thus $N(t) \sim \text{Poisson}(\lambda t)$. ■

Another fact: $(N(t))_{t\geq 0}$ has independent Poisson increments. This means that if $t_0 < t_1 < \cdots < t_n$, then $N(t_i) - N(t_{i-1}) \sim \text{Poisson}(\lambda(t_i - t_{i-1}))$ and are all independent.

Definition. A Poisson process of intensity $\lambda > 0$ is a collection $\{N(t)\}_{t\geq 0}$ of non-decreasing integer-valued random variables satisfying

1. $N(0) = 0$
2. $N(t) \sim \text{Poisson}(\lambda t)$ for all $t \geq 0$
3. $N(t)$ has independent Poisson increments

Generalizing, let A be some subset of $\mathbb{R}^n$ and let $N(A)$ be the number of marked points in A. Then

$$N(A) \sim \text{Poisson}\left(\int_A \lambda(x)dx\right)$$

defines the Poisson process, and for all disjoint A, B, $N(A)$ and $N(B)$ are independent.

Proposition. As $h \to 0$,

(i) $P(N(t+h) - N(t) = 1) = \lambda h + o(h)$

(ii) $P(N(t+h) - N(t) \geq 2) = o(h)$

Fact: Any stochastic process with independent Poisson increments and satisfying (i) and (ii) above is a Poisson process with rate λ.
Superposition property: If $(N_1(t))_{t\geq 0}$ and $(N_2(t))_{t\geq 0}$ are independent Poisson processes with intensity λ_1 and λ_2 respectively, then $(N_1(t) + N_2(t))_{t\geq 0}$ is a Poisson process with intensity $\lambda_1 + \lambda_2$.
Thinning property: Let $(N(t))_{t\geq 0}$ be a Poisson process with intensity λ. Suppose each arrival is independently of type i with probability p_i for all i with $\sum_i p_i = 1$. Let $N_i(t)$ be the number of arrivals of type i up to time t. Then, then $(N_i(t))_{t\geq 0}$ are all independent Poisson proceses with intensity λp_i.
Claim. Conditionally on $N(t) = N$, the number of marked points is $\overset{iid}{\sim}$ Uniform$[0, t]$.

6.2 Continuous-time Discrete-space Markov Processes

Definition. A continuous-time Markov process on a countable (discrete) state space S is a collection $\{X(t)\}_{t\geq 0}$ of random variables such that

$$P(X_0 = i_0, X_{t_1} = i_1, \ldots, X_{t_n} = i_n) = v_{i_0} p_{i_0 i_1}^{(t_1)} p_{i_1 i_2}^{(t_2 - t_1)} \cdots p_{i_{n-1} i_n}^{(t_n - t_{n-1})}$$

for all $i_0, \ldots, i_n \in S$ and times $0 < t_1 < t_2 < \cdots < t_n$.

- A Poisson process with intensity λ is a continuous-time discrete-space Markov process with transition probabilities

$$p_{ij}^{(t)} = \begin{cases} 0 & j < i \\ \frac{e^{-\lambda t}(\lambda t)^{j-i}}{(j-i)!} & j \geq i \end{cases}$$

A "standard Markov process" is characterized by

$$\lim_{t\to\infty} p_{ij}^{(t)} = p_{ij}^{(0)}$$

The characteristics of discrete-time MCs apply here:

Kolmogorov-Chapman (continuous): For $s, t \geq 0$, $P^{(s)}P^{(t)} = P^{(s+t)}$. Additionally, $p_{ij}^{(t+s)}$ is continuous if the process is standard Markov.

Definition. A generator of a standard Markov procss is

$$g_{ij} := \lim_{t \to 0} \frac{p_{ij}^t - p_{ij}^{(0)}}{t}$$

The idea is if t is small, then $P^{(t)} \approx I + tG$ where $G = (g_{ij})_{i,j \in S}$.
Properties:

- $g_{ii} = \lim_{t \to 0} \frac{p_{ii}^{(t)} - 1}{t} \leq 0$
- $g_{ij} \geq 0$
- $\sum_{j \in S} g_{ij} = 0$
- $-g_{ii} = \sum_{\substack{j \in S \\ j \neq i}} g_{ij}$

Theorem (Continuous-time Transitions Theorem)**.**

$$P^{(t)} = \exp(tG) := I + tG + \frac{t^2G^2}{2!} + \cdots$$

To compute $P^{(t)}$: suppose G is diagonalizable, so $G = P\Lambda P^{-1}$. Then

$$\begin{aligned} \exp(tG) &= \sum_{n=0}^{\infty} \frac{t^n G^n}{n!} \\ &= P \left(\sum_{n=0}^{\infty} t^n \Lambda^n \right) P^{-1} \\ &= P\text{diag}[e^{t\lambda_1}, \ldots]P^{-1} \end{aligned}$$

Definition. $(\pi_i)_{i \in S}$ is a stationary distribution if $\pi G = 0$ (or equivalently, $\pi P^{(t)} = \pi$).

- $\sum_{i \in S} \pi_i g_{ij} = 0$

Definition. A Markov process is reversible wrt $(\pi_i)_{i \in S}$ if

$$\pi_i g_{ij} = \pi_j g_{ji}$$

Note that reversible implies stationary since $\sum_j g_{ij} = 0$.

Theorem. If a Markov process is irreducible and has stationary distribution π, then

$$\lim_{t \to \infty} p_{ij}^{(t)} = \pi_j$$

for all $i, j \in S$.

6.2.1 Constructing continuous-time Markov processes

Given a generator $(g_{ij})_{ij}$, sample the times τ_i from

$$\text{Time} = \begin{cases} \tau_i \sim \text{Exponential}(-g_{ij}) & g_{ij} > 0 \\ \text{Absorbing state} & g_{ij} = 0 \end{cases}$$

Define the next-step transitions as

$$\tilde{p}_{ij} = \begin{cases} \frac{g_{ij}}{-g_i i} & j \neq i \\ 0 & \text{otherwise} \end{cases}$$

A continuous-time Markov process with generator G is equal to in distribution to the process above.