A Markov chain is a collection of random variables $X_0, X_1, \ldots$ over the same probability space with a common state space $S$ such that

\[
\mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n)
\]

for any $i_0, i_1, \ldots, i_{n+1} \in S$.

In this class, a Markov chain is specified by:

1. A finite state space $S$: $|S| < \infty$
2. A time-homogeneous transition matrix $P$:
   \[
P(i, j) = \mathbb{P}(X_{n+1} = j \mid X_n = i) = \mathbb{P}(X_{n+1} = j \mid X_n = i)\text{ for all } n \geq 0.
   \]
3. An initial distribution $\pi_0$:
   \[
   \sum_{k \in S} \pi_0(k) = 1
   \]

Thus, let $P_n$ be the $n$-step transition matrix $P_n(i, j) = \mathbb{P}(X_n = j \mid X_0 = i)$.

Let $\pi_n$ be the distribution over the states after $n$ steps.

Then

\[
P_n = P^n
\]

and

\[
\pi_n = \pi_0 P^n
\]
Then let \( \mathcal{A}, \mathcal{B} \subset \mathcal{S} \) be mutually exclusive. Let \( T_{\mathcal{A}}, T_{\mathcal{B}} \) be the times until we visit a state in \( \mathcal{A} \) and \( \mathcal{B} \), respectively.

Define
\[
\alpha(i) = \mathbb{P}(T_{\mathcal{A}} < T_{\mathcal{B}} \mid X_0 = i), \quad \beta(i) = \mathbb{E}[T_{\mathcal{A}} \mid X_0 = i].
\]

Then
\[
\alpha(i) = 1 \text{ if } i \in \mathcal{A} \\
\alpha(i) = 0 \text{ if } i \in \mathcal{B} \\
\alpha(i) = \sum_{j \in \mathcal{S}} \mathbb{P}(T_{\mathcal{A}} < T_{\mathcal{B}}, X_0 = i, X_1 = j) \cdot \mathbb{P}(X_1 = j \mid X_0 = i) \\
= \sum_{j \in \mathcal{S}} \mathbb{P}(T_{\mathcal{A}} < T_{\mathcal{B}} \mid X_0 = i, X_1 = j) \cdot \mathbb{P}(X_1 = j \mid X_0 = i) \\
= \sum_{j \in \mathcal{S}} \mathbb{P}(T_{\mathcal{A}} < T_{\mathcal{B}} \mid X_0 = j) \cdot \mathbb{P}(i, j) \text{ by homogeneity} \\
= \sum_{j \in \mathcal{S}} \alpha(j) \cdot \mathbb{P}(i, j) \text{ if } i \in \mathcal{A}, i \not\in \mathcal{B}.
\]

\[
\beta(i) = 0 \text{ if } i \in \mathcal{A} \\
\beta(i) = \sum_{j \in \mathcal{S}} \mathbb{E}[T_{\mathcal{A}} \mid X_0 = i, X_1 = j] \cdot \mathbb{P}(X_1 = j \mid X_0 = i) \\
= \sum_{j \in \mathcal{S}} (1 + \mathbb{E}[T_{\mathcal{A}} \mid X_0 = j]) \cdot \mathbb{P}(i, j) \text{ by homogeneity} \\
= \sum_{j \in \mathcal{S}} \mathbb{P}(i, j) + \sum_{j \in \mathcal{S}} \beta(j) \cdot \mathbb{P}(i, j) \\
= 1 + \sum_{j \in \mathcal{S}} \beta(j) \cdot \mathbb{P}(i, j) \text{ if } i \not\in \mathcal{A}.
\]

Then let \( V_A \) be the number of times a state in \( \mathcal{A} \) is visited.

Define
\[
\delta(i) = \mathbb{E}[V_A \mid X_0 = i].
\]

Then
\[
\delta(i) = 1 + \sum_{j \in \mathcal{S}} \delta(j) \cdot \mathbb{P}(i, j) \text{ if } i \in \mathcal{A} \\
\delta(i) = \sum_{j \in \mathcal{S}} \delta(j) \cdot \mathbb{P}(i, j) \text{ if } i \not\in \mathcal{A}.
\]

Note: \( \mathbb{E}[V_A] \) may be infinite.
Compartmental Model

Ex. Suppose that for some disease, the population can be split into three groups: susceptible, infected, and removed.
Each day,
- of the susceptible, 20% become infected;
- of the infected, 20% become susceptible;
- of the removed, \( p \) become susceptible
- 80% stay susceptible;
- 10% become removed;
- \( 1-p \) stay infected.
In the population prior to exposure, 95% are susceptible and 5% are removed.
We can model this as a Markov process:
\[
S = \{ S, I, R \}
\]
\[
\begin{bmatrix}
S & I & R \\
\end{bmatrix}
\]
\[
\begin{bmatrix}
0.8 & 0.2 & 0 \\
0.2 & 0.7 & 0.1 \\
p & 0 & 1-p \\
\end{bmatrix}
\]
\[
\pi_0 = \begin{bmatrix}
0.95 & 0 & 0.05 \\
\end{bmatrix}
\]
Invariant Distributions

**Def:** A distribution \( \pi \) is invariant for transition matrix \( \mathbf{P} \) if
\[
\pi = \pi \mathbf{P}
\]
These equations are called the balance equations.

**Thm:** \( \pi_n = \pi_0 \) for all \( n \geq 0 \) if and only if \( \pi_0 \) is invariant.

**Pf:** Suppose \( \pi_n = \pi_0 \) for all \( n \geq 0 \). Then
\[
\pi_i = \pi_0 \mathbf{P}^n = \pi_0,
\]
so 0 is invariant.

Suppose \( \pi_0 = \pi_0 \mathbf{P} \). Then
\[
\pi_i = \pi_0 \mathbf{P} = \pi_0
\]
\[
\pi_{in} = \pi_{in} \mathbf{P} = \pi_0
\]
By induction, \( \pi_n = \pi_0 \) for all \( n \geq 0 \).

**Note:** A Markov chain may have many invariant distributions. For example, \( \mathbf{P} = \mathbf{I} \) has infinitely many.

**Note:** Invariance means that the net flow in and out of states is equal.

\[
\pi(i) \text{ of the particles are leaving state } 0
\]
\[
\pi(0) \mathbf{P}(0,0) + \pi(1) \mathbf{P}(1,0) + \pi(2) \mathbf{P}(2,0) + \pi(3) \mathbf{P}(3,0) \text{ are entering state } 0
\]
The balance equations say these flows are equal:
\[
\pi(i) = \sum_{j \in S} \pi(j) \cdot \mathbf{P}(j,i)
\]

\[
\text{up}
\]
\[
\text{leaving i}
\]
\[
\text{entering i}
\]
**Irreducibility**

**Def.** We say i can reach j (i → j) if there is a path with positive probability from i to j.

A Markov chain is irreducible if

∀i,j i→j and j→i (you can get from any state to any state)

**Ex.** Which of the following chains is irreducible?

- [Diagram of chain 1 showing reducible with paths 0 → 1 (Reducible), 1 → 0 (Reducible), 0 → 2 (Reducible), Irreducible]

To show a chain is irreducible, construct a path that:

- starts at any state
- goes through all of the other states
- ends at the starting state
- has positive probability

- [Diagram of a path 0 → 1 → 3 → 2 → 3 → 0]
Long Run Proportion of Time

Then If a finite Markov chain is irreducible, then, for any initial distribution \( \pi_0 \),

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} 1\{X_i = i\} = \pi(i) \text{ for all } i \in S
\]

That is, the fraction of time spent in each state is given by \( \pi \).
Also, \( \pi \) is invariant; therefore the invariant distribution exists and is unique.

Proof: Because the chain is irreducible, each state will be visited infinitely many times.
For each state \( i \), let \( T_i \) be the number of steps to return to \( i \) starting with \( X_0 = i \). Then let

\[
\pi(i) = \frac{1}{E[T_i]}, \text{ the fraction of time spent in state } i
\]

So the fraction of time the chain spends in state \( i \) is \( \pi(i) \), as desired.
Now, \( \pi \) is invariant. Note that in \( n \) steps, the chain is in state \( i \) for \( n \pi(i) \) steps.
Over large \( n \), the chain transitions from any state \( j \) to \( i \)

\[
n \pi(j) \cdot P(j,i)
\]

in state transition from \( j \) to \( i \) times. The total visits to state \( i \) is the sum of all visits from the states:

\[
n \pi(i) = \sum_{j \in S} n \pi(j) P(j,i)
\]

This can be written as \( \pi = \pi P \).
Proof of uniqueness is more complicated.

Note: An irreducible Markov chain's distribution does not necessarily converge to \( \pi \):

The chain spends half the time in each state, but every initial distribution does not converge to \([1/2, 1/2]\), e.g.,

\[
[0,1] \to [1,0] \to [0,1] \to ...
\]

This is because of the periodic behavior of the chain.
**Periodicity**

**Def:** The period of a state $i$, denoted $d(i)$, is

$$d(i) = \gcd \{ n : P^n(i, i) > 0 \};$$

that is, the period of a state is the greatest common divisor of the lengths of paths from $i$ to $i$.

In this class we only consider the period of states in irreducible chains.

**Ex:** Find $d(i)$ for each of the following chains.

- $1 \to 2 \to 1$; $1 \to 2 \to 1 \to 2 \to 1$
  
  $d(1) = \gcd \{2, 4, 6, \ldots \} = 2$

- $1 \to 0 \to 1$; $1 \to 3 \to 0 \to 1$
  
  $d(1) = \gcd \{2, 3, \ldots \} = 1$

- $1 \to 2 \to 0 \to 1$; $1 \to 2 \to 0 \to 2 \to 1$
  
  $d(1) = \gcd \{3, 4, \ldots \} = 1$

**Claim:** If a Markov chain is irreducible, the periods of all states are the same.

$d(i) = d(j)$ for all $i, j \in S$.

**Proof:** As an exercise.

**Def:** If $d(i) = 1$ for any state in an irreducible Markov chain, we say the chain is aperiodic.

**Ex:** Which of the following chains is aperiodic?

- Periodic, since $d(1) = 2$

- Aperiodic, since $d(1) = 1$

- Aperiodic, since $d(1) = 1$

To show a chain is aperiodic, find two loops with coprime lengths.
Markov Chain Convergence Theorem

**Theorem (Finite Markov Chain Convergence)** If \( X_0, X_1, \ldots \) is a Markov chain on \( S \) with time-homogeneous transition matrix \( P \) and

1. \( |S| < \infty \)
2. \( P \) is irreducible: \( \forall i, j \in S, \exists n \in \mathbb{N} : P^n(i, j) > 0 \)
3. \( P \) is aperiodic: \( \forall i \in S, a(i) = 1 \),

then \( X_0, X_1, \ldots \) has an invariant distribution \( \pi = \pi P \) such that

\[
\lim_{n \to \infty} P^n(i, j) = \pi(j)
\]

i.e. the \( n \)-step transition probabilities converge to \( \pi \).

Moreover,

\[
\lim_{n \to \infty} \frac{1}{n} E[T_j] = \pi(j),
\]

i.e. \( \pi(j) \) is the long-run proportion of time spent in state \( j \).

**Corollary (Ergodicity)**

\[
\pi_n(i) = P(x_n = i) \to \pi(i) \text{ as } n \to \infty
\]

\[
\pi_n(i) = \sum_{j \in S} \pi_0(j) P_n(j, i)
\]

\[
\to \sum_{j \in S} \pi_0(j) \pi(i) \text{ as } n \to \infty
\]

\[
= \pi(i) \sum_{j \in S} \pi_0(j)
\]

\[
= \pi(i)
\]

**Note**: We can find the long-run probability of an event by conditioning on \( \pi \).

\[
\begin{array}{ccc}
\nu_2 & \circ & \nu_4 \\
0 \quad 0.75 \\
\end{array}
\]

\( \pi = \left[ \frac{1}{3}, \frac{2}{3} \right] \)

What is the long-run probability that we stay in the same state?

\[
P(\text{stay}) = P(\text{stay} | \text{state 0}) \cdot P(\text{state 0}) + P(\text{stay} | \text{state 1}) \cdot P(\text{state 1})
\]

\[
= \frac{1}{2} \cdot \frac{1}{3} + \frac{2}{3} \cdot \frac{2}{3}
\]

\[
= \frac{3}{3}
\]
In the Ehrenfest model, there are two containers, containing a total of \( N \) particles.

At each step:
- a container is selected uniformly at random,
- a particle is selected uniformly at random, independently of the container,
- and the selected particle is placed in the selected container; if the particle was already in the container, it remains in place.

Let \( X_n \) be the number of particles in the first container at time \( n \).

What are the transition probabilities of the chain?

The number of particles can either increase by 1, decrease by 1, or remain the same.

\[
P(X_{n+1} = i+1 \mid X_n = i) = \frac{1}{2} \cdot \frac{N-i}{N} = \frac{N-i}{2N}
\]

\[
P(X_{n+1} = i-1 \mid X_n = i) = \frac{1}{2} \cdot \frac{i}{N} = \frac{i}{2N}
\]

\[
P(X_{n+1} = i \mid X_n = i) = 1 - \frac{i}{2N} - \frac{N-i}{2N} = \frac{1}{2}
\]

So

\[
P(i,i) = \begin{cases} 
\frac{N-i}{2N} & \text{if } i = i+1 \\
\frac{i}{2N} & \text{if } i = i \\
\frac{i}{2N} & \text{if } i = i-1 \\
0 & \text{otherwise}
\end{cases}
\]

Prove that any distribution over the states converges to some distribution \( \pi \).

We must show that the chain is irreducible and aperiodic.

Irreducible: Consider the path \( 0 \to 1 \to 2 \to \ldots \to N \to N-1 \to \ldots \to 0 \).

Aperiodic: Note that \( a(i) = \gcd\{1, \ldots, i\} = 1 \). So all states have period 1.
In the Ehrenfest model, there are two containers, containing a total of \( N \) particles. At each step:
- A container is selected uniformly at random.
- A particle is selected uniformly at random, independently of the container.
- The selected particle is placed in the selected container; if the particle was already in the container, it remains in place.

Let \( X_n \) be the number of particles in the first container at time \( n \).

Find the stationary distribution of the chain.

The balance equations are:

\[
\pi(0) = \frac{1}{2} \pi(0) + \frac{1}{2N} \pi(1)
\]

\[
\pi(j) = \frac{N-(j-1)}{2N} \pi(j-1) + \frac{1}{2} \pi(j) + \frac{j}{2N} \pi(j+1) \quad \text{for} \quad 1 \leq j \leq N-1
\]

\[
\pi(N) = \frac{1}{2N} \pi(N-1) + \frac{1}{2} \pi(N)
\]

Rewrite the first few equations in terms of \( \pi(0) \)

\( \pi(0) = \frac{1}{2} \pi(0) + \frac{1}{2N} \pi(1) \Rightarrow \pi(1) = N\pi(0) = \binom{N}{1} \pi(0) \)

\( \pi(1) = \frac{N}{2N} \pi(0) + \frac{1}{2} \pi(1) + \frac{2}{2N} \pi(2) \Rightarrow \pi(2) = \frac{N}{2} (\pi(1) - \pi(0)) = \frac{N(N-1)}{2} \pi(0) = \binom{N}{2} \pi(0) \)

By induction, \( \pi(j) = \binom{N}{j} \pi(0) \).

Then:

\[
\sum_{j=0}^{N} \pi(j) = \sum_{j=0}^{N} \binom{N}{j} \pi(0) = \pi(0) \sum_{j=0}^{N} \binom{N}{j} = \pi(0) 2^N = 1 \Rightarrow \pi(0) = \frac{1}{2^N}
\]

So:

\( \pi(k) = P(X_n = k) = \binom{N}{k} \frac{1}{2^N} \),

i.e. \( \pi(k) \sim \text{Binomial}(N, \frac{1}{2}) \).