02/17/2026: Markov Chains (Part 3)

CSCI 546: Diffusion Models

Textbook reference: Sec 6.5, 6.6, 6.9

Announcement (Sign-in Sheet)

Please sign the sign-in sheet.

Announcement (Sign-in Sheet)

Please sign the sign-in sheet.

Intro Tutorial to Tempest

By Jess Kunke, Asst. Professor of Statistics

Wednesday, February 18th 4:10-5:00 PM

Wilson 1-154

Announcement (Sign-in Sheet)

Please sign the sign-in sheet.

Intro Tutorial to Tempest

By Jess Kunke, Asst. Professor of Statistics
Wednesday, February 18th 4:10-5:00 PM
Wilson 1-154

A look at the syllabus

- I've added some additional readings to the diffusion section which give some nice motivation and intuition.
- I'm considering splitting the final topic (stochastic calculus) up across 2 days. Thoughts?

Review Problem Set #9

Concepts for Problem Set #10

Outline for today's material

- Reversibility
- Chains with finitely many states
- Continuous-time Markov chains

Stationarity as “global balance”

Stationarity

Recall that $\boldsymbol{\pi}$ is a stationary distribution for a Markov chain if

$$\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P},$$

where $\boldsymbol{\pi}$ is a row vector and $\boldsymbol{P}$ is the transition probability matrix of the Markov chain.

Stationarity as "global balance"

Stationarity

Recall that $\boldsymbol{\pi}$ is a stationary distribution for a Markov chain if

$$\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P},$$

where $\boldsymbol{\pi}$ is a row vector and $\boldsymbol{P}$ is the transition probability matrix of the Markov chain.

The **most common interpretation** of stationarity: $\boldsymbol{\pi}$ gives the long-run frequencies of inhabiting each state.

Stationarity as "global balance"

Stationarity

Recall that $\boldsymbol{\pi}$ is a stationary distribution for a Markov chain if

$$\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P},$$

where $\boldsymbol{\pi}$ is a row vector and $\boldsymbol{P}$ is the transition probability matrix of the Markov chain.

The **most common interpretation** of stationarity: $\boldsymbol{\pi}$ gives the long-run frequencies of inhabiting each state. However, we can also give another interpretation.

Stationarity as "global balance"

Stationarity

Recall that $\boldsymbol{\pi}$ is a stationary distribution for a Markov chain if

$$\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P},$$

where $\boldsymbol{\pi}$ is a row vector and $\boldsymbol{P}$ is the transition probability matrix of the Markov chain.

Stationarity as "global balance"

$$\begin{aligned} \boldsymbol{\pi} &= \boldsymbol{\pi}\boldsymbol{P} \\ \implies \pi_j &= \sum_i \pi_i P_{ij} && \text{(matrix multiplication)} \\ \implies \sum_i \pi_j P_{ji} &= \sum_i \pi_i P_{ij} && \text{(probabilities sum to 1)} \end{aligned}$$

Stationarity as "global balance"

Stationarity

Recall that $\boldsymbol{\pi}$ is a stationary distribution for a Markov chain if

$$\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P},$$

where $\boldsymbol{\pi}$ is a row vector and $\boldsymbol{P}$ is the transition probability matrix of the Markov chain.

Stationarity as "global balance"

$$\underbrace{\sum_i \pi_j P_{ji}}_{\text{total flow } \underline{\textbf{out of}} \text{ state } j} = \underbrace{\sum_i \pi_i P_{ij}}_{\text{total flow } \underline{\textbf{into}} \text{ state } j}$$

Reversibility as "local balance"

"Local balance": a stronger condition than global balance

$$\underbrace{\pi_j P_{ji}}_{\text{total flow from } j \text{ to } i} = \underbrace{\pi_i P_{ij}}_{\text{total flow from } i \text{ to } j}$$

Reversibility as "local balance"

"Local balance": a stronger condition than global balance

$$\underbrace{\pi_j P_{ji}}_{\text{total flow from } j \text{ to } i} = \underbrace{\pi_i P_{ij}}_{\text{total flow from } i \text{ to } j}$$

Why do we care?

Reversibility as "local balance"

"Local balance": a stronger condition than global balance

$$\underbrace{\pi_j P_{ji}}_{\text{total flow from } j \text{ to } i} = \underbrace{\pi_i P_{ij}}_{\text{total flow from } i \text{ to } j}$$

Why do we care? This tells us when Markov chains are reversible.

Reversibility as "local balance"

"Local balance": a stronger condition than global balance

$$\underbrace{\pi_j P_{ji}}_{\text{total flow from } j \text{ to } i} = \underbrace{\pi_i P_{ij}}_{\text{total flow from } i \text{ to } j}$$

Why do we care? This tells us when Markov chains are reversible.

Definition

A Markov chain (X_n) is called (time) **reversible** if for each n

$$(X_0, X_1, \ldots, X_n) \stackrel{\mathcal{D}}{=} (X_n, X_{n-1}, \ldots X_0).$$

That is, the joint distribution of $(X_0, X_1, \ldots, X_n)$ is the same as the joint distribution of $(X_n, X_{n-1}, \ldots X_0)$.

Reversibility as "local balance"

Proposition. An irreducible Markov chain is reversible if and only the distribution π of X_0 is stationary and satisfies local balance.

Reversibility as "local balance"

Proposition. An irreducible Markov chain is reversible if and only the distribution $\boldsymbol{\pi}$ of X_0 is stationary and satisfies local balance.

Justification.

For example,

$$P(X_0 = i, X_1 = j, X_2 = k)$$

Reversibility as "local balance"

Proposition. An irreducible Markov chain is reversible if and only the distribution $\boldsymbol{\pi}$ of X_0 is stationary and satisfies local balance.

Justification.

For example,

$$P(X_0 = i, X_1 = j, X_2 = k) = [\pi_i p_{ij}]p_{jk}$$

Reversibility as "local balance"

Proposition. An irreducible Markov chain is reversible if and only the distribution $\boldsymbol{\pi}$ of X_0 is stationary and satisfies local balance.

Justification.

For example,

$$\begin{aligned} P(X_0 = i, X_1 = j, X_2 = k) &= [\pi_i p_{ij}] p_{jk} \\ &= [\pi_j p_{ji}] p_{jk} && \text{local balance} \end{aligned}$$

Reversibility as "local balance"

Proposition. An irreducible Markov chain is reversible if and only the distribution $\boldsymbol{\pi}$ of X_0 is stationary and satisfies local balance.

Justification.

For example,

$$\begin{aligned} P(X_0 = i, X_1 = j, X_2 = k) &= [\pi_i p_{ij}] p_{jk} \\ &= [\pi_j p_{ji}] p_{jk} && \text{local balance} \\ &= p_{ji} [\pi_j p_{jk}] \end{aligned}$$

Reversibility as "local balance"

Proposition. An irreducible Markov chain is reversible if and only the distribution $\boldsymbol{\pi}$ of X_0 is stationary and satisfies local balance.

Justification.

For example,

$$\begin{aligned}
P(X_0 = i, X_1 = j, X_2 = k) &= [\pi_i p_{ij}] p_{jk} \\
&= [\pi_j p_{ji}] p_{jk} && \text{local balance} \\
&= p_{ji} [\pi_j p_{jk}] \\
&= p_{ji} [\pi_k p_{kj}] && \text{local balance}
\end{aligned}$$

Reversibility as “local balance”

Proposition. An irreducible Markov chain is reversible if and only the distribution $\boldsymbol{\pi}$ of X_0 is stationary and satisfies local balance.

Justification.

For example,

$$\begin{aligned}
P(X_0 = i, X_1 = j, X_2 = k) &= [\pi_i p_{ij}] p_{jk} \\
&= [\pi_j p_{ji}] p_{jk} && \text{local balance} \\
&= p_{ji} [\pi_j p_{jk}] \\
&= p_{ji} [\pi_k p_{kj}] && \text{local balance} \\
&= \pi_k p_{kj} p_{ji}
\end{aligned}$$

Reversibility as "local balance"

Proposition. An irreducible Markov chain is reversible if and only the distribution $\boldsymbol{\pi}$ of X_0 is stationary and satisfies local balance.

Justification.

For example,

$$
\begin{aligned}
P(X_0 = i, X_1 = j, X_2 = k) &= [\pi_i p_{ij}]p_{jk} \\
&= [\pi_j p_{ji}]p_{jk} && \text{local balance} \\
&= p_{ji}[\pi_j p_{jk}] \\
&= p_{ji}[\pi_k p_{kj}] && \text{local balance} \\
&= \pi_k p_{kj} p_{ji} \\
&= P(X_0 = k, X_1 = j, X_2 = i)
\end{aligned}
$$

Reversibility as "local balance"

Proposition. An irreducible Markov chain is reversible if and only the distribution $\boldsymbol{\pi}$ of X_0 is stationary and satisfies local balance.

Justification.

For example,

$$\begin{aligned}
P(X_0 = i, X_1 = j, X_2 = k) &= [\pi_i p_{ij}] p_{jk} \\
&= [\pi_j p_{ji}] p_{jk} && \text{local balance} \\
&= p_{ji} [\pi_j p_{jk}] \\
&= p_{ji} [\pi_k p_{kj}] && \text{local balance} \\
&= \pi_k p_{kj} p_{ji} \\
&= P(X_0 = k, X_1 = j, X_2 = i)
\end{aligned}$$

Remark. Notice how local balance allowed the π factor to propagate through the product from the left end to the right, reversing the direction of all of the transitions along the way.

Ehrenfest model of diffusion

This chain can shed light on perplexing questions like "Why aren't people dying all the time due to the air molecules bunching up in some odd corner of their bedrooms while they sleep?" - Joseph Chang

Ehrenfest model of diffusion

This chain can shed light on perplexing questions like "Why aren't people dying all the time due to the air molecules bunching up in some odd corner of their bedrooms while they sleep?" - Joseph Chang

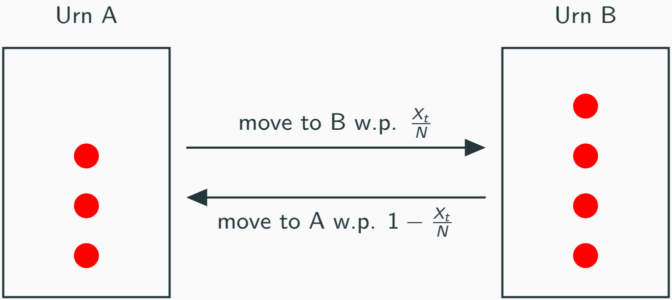
$$N = \text{number of balls in both urns}$$
$$X_t = \text{number of balls in Urn A at time } t$$

Ehrenfest model of diffusion

This chain can shed light on perplexing questions like "Why aren't people dying all the time due to the air molecules bunching up in some odd corner of their bedrooms while they sleep?" - Joseph Chang

$$N = \text{number of balls in both urns}$$

$$X_t = \text{number of balls in Urn A at time } t$$



Outline for today's material

- Reversibility
- Chains with finitely many states
- Continuous-time Markov chains

Chains with finitely many states

The theory of Markov chais is much simplified when the state space is finite.

Chains with finitely many states

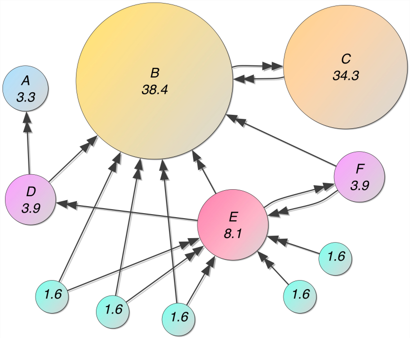
The theory of Markov chais is much simplified when the state space is finite.

In particular, we have:

Theorem (Perron-Frobenius)(Partial)

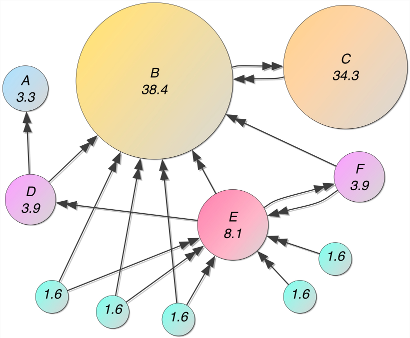
Any finite irreducible Markov chain has a unique stationary distribution $\boldsymbol{\pi}$. Furthermore, all components of $\boldsymbol{\pi}$ are strictly positive.

The 'Page Rank' Markov Chain



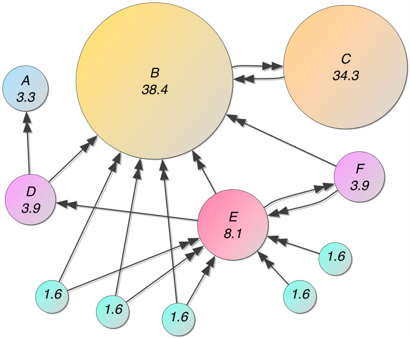
How does Google's patented 'Page Rank' algorithm describe the relative popularities of webpages?

The ‘Page Rank’ Markov Chain



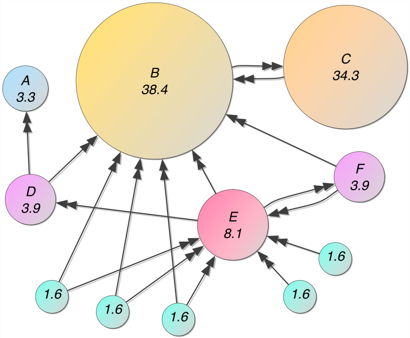
Webpages form a directed graph with n vertices (representing webpages). A link from page i to page j is designated by an directed edge.

The ‘Page Rank’ Markov Chain



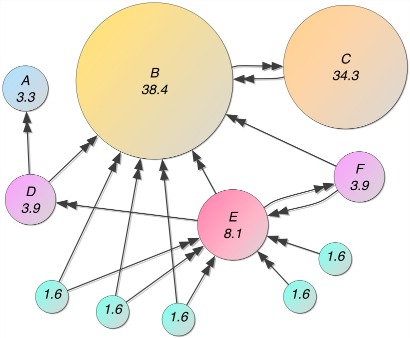
The behavior of a swiftly bored web surfer can be modeled by a random walk on the directed graph of webpages.

The 'Page Rank' Markov Chain



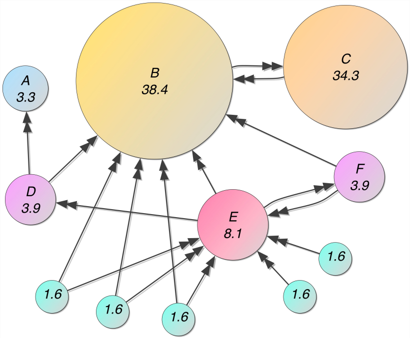
With probability b, the random walk moves to a randomly linked page from the current page (or to a random page in the graph, if a webpage has no outgoing links). With probability $1 - b$, the random walk moves to a random page in the graph.

The ‘Page Rank’ Markov Chain



Let d_i be the out-degree of vertex i.

The 'Page Rank' Markov Chain



Then the transition matrix $\boldsymbol{P} = b\boldsymbol{Q} + (1-b)\boldsymbol{u}^\top \boldsymbol{e}$, where $\boldsymbol{Q} = (q_{ij})$ with

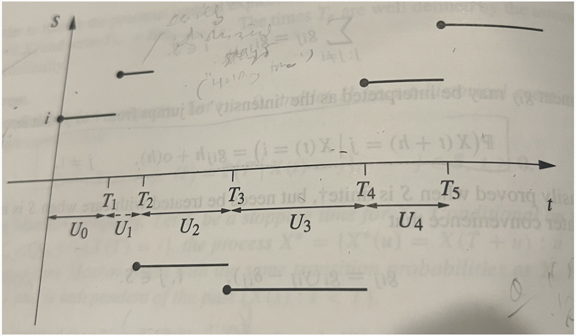
$$q_{ij} = \begin{cases} 1/d_i, & \text{if } i \to j \\ 1/n, & \text{if } i \text{ dangles } (d_i = 0) \\ 0, & \text{otherwise} \end{cases}$$

where $\boldsymbol{u}$ is a row vector with $u_i = 1/n$ and $\boldsymbol{e}$ is a row vector with entries 1.

Outline for today's material

- Reversibility
- Chains with finitely many states
- Continuous-time Markov chains

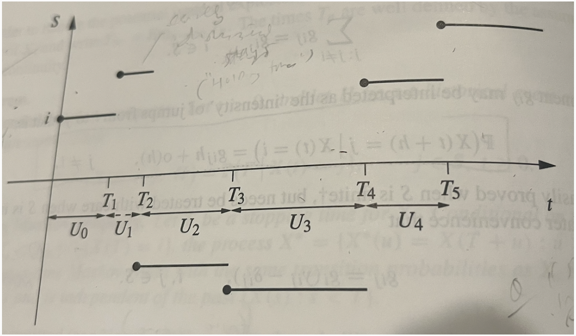
Continuous Time Markov Chains



A continuous time Markov chain is characterized by

1. A discrete-time **jump chain** (with some transition probability matrix).
2. A probability distribution over **holding times** at each state.

Continuous Time Markov Chains



A continuous time Markov chain is characterized by

1. A discrete-time **jump chain** (with some transition probability matrix).
2. A probability distribution over **holding times** at each state.

Examples. Traffic light cycles, inventory systems with bulk orders, popcorn popping, customer queues, website traffic,

Continuous Time Markov Chains

Let $X = \{X(t) : t \geq 0\}$ be a family of random variables taking values in a countable state space S and indexed by the half line $[0, \infty)$.

Continuous Time Markov Chains

Let $X = \{X(t) : t \geq 0\}$ be a family of random variables taking values in a countable state space S and indexed by the half line $[0, \infty)$.

Poll. How to characterize the Markov property for continuous chains?

Continuous Time Markov Chains

Let $X = \{X(t) : t \geq 0\}$ be a family of random variables taking values in a countable state space S and indexed by the half line $[0, \infty)$.

Poll. How to characterize the Markov property for continuous chains?

Definition

A continuous time chain X satisfies the **Markov property** if

$$P\big(X(t_n) = i_n \mid X(t_1) = i_1, X(t_1) = i_2, \ldots, X(t_{n-1}) = i_{n-1}\big)$$
$$= P\big(X(t_n) = i_n \mid X(t_{n-1}) = i_{n-1}\big)$$

for any increasing sequence of times $t_1 < t_2 < \ldots < t_n$ and states $i_1, i_2, \ldots i_n$.

Continuous Time Markov Chains

Let $X = \{X(t) : t \geq 0\}$ be a family of random variables taking values in a countable state space S and indexed by the half line $[0, \infty)$.

Poll. How to characterize the Markov property for continuous chains?

Definition

A continuous time chain X satisfies the **Markov property** if

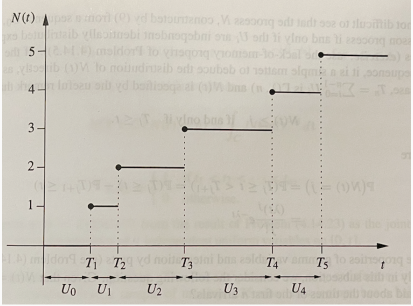
$$\begin{aligned} P\big(X(t_n) = i_n \mid X(t_1) = i_1, X(t_1) = i_2, \ldots, X(t_{n-1}) = i_{n-1}\big) \\ = P\big(X(t_n) = i_n \mid X(t_{n-1}) = i_{n-1}\big) \end{aligned}$$

for any increasing sequence of times $t_1 < t_2 < \ldots < t_n$ and states $i_1, i_2, \ldots i_n$.

Remark. For technical reasons, the text restricts consideration to *right continuous* chains – that is those which have *stepwise* sample paths.

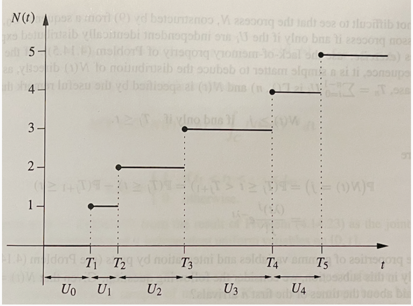
Example: Poisson Process

A Poisson process $N(t)$ can be used to model the number of 'arrivals' or 'occurrences' or 'events' by time t.



Example: Poisson Process

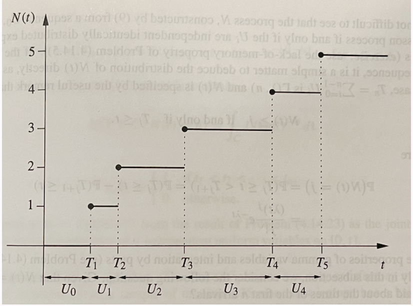
A Poisson process $N(t)$ can be used to model the number of ‘arrivals‘ or ‘occurrences‘ or ‘events‘ by time t.



Example. Emission of particles by a radioactive source.

Example: Poisson Process

A Poisson process $N(t)$ can be used to model the number of 'arrivals' or 'occurrences' or 'events' by time t.

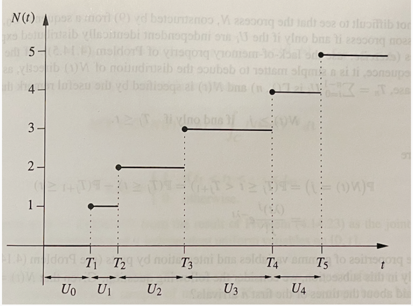


A Poisson process with intensity λ has pmf

$$P\big(N(t) = j\big) = \frac{(\lambda t)^j}{j!} e^{-\lambda t}, \qquad j = 0, 1, 2, \ldots$$

Example: Poisson Process

A Poisson process $N(t)$ can be used to model the number of ‘arrivals‘ or ‘occurrences‘ or ‘events‘ by time t.



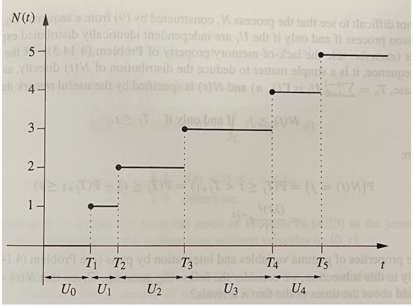
A Poisson process with intensity λ has pmf

$$P\big(N(t) = j\big) = \frac{(\lambda t)^j}{j!} e^{-\lambda t}, \qquad j = 0, 1, 2, \ldots$$

Hence, the expected value is $E[N(t)] = \lambda t$.

Example: Poisson Process

A Poisson process $N(t)$ can be used to model the number of 'arrivals' or 'occurrences' or 'events' by time t.



Theorem. The **interarrival times** $U_n = T_{n+1} - T_n$ are independent and exponentially distributed with parameter λ.

Popcorn as Poisson Process?

Poll. Can we consider popcorn popping a Poisson process?

Popcorn as Poisson Process?

Poll. Can we consider popcorn popping a Poisson process?

Potential issues.

1. **Non-constant rate.** The expected waiting time λ changes over time. It increases as heat rises, then decreases as kernels run out.

Popcorn as Poisson Process?

Poll. Can we consider popcorn popping a Poisson process?

Potential issues.

1. **Non-constant rate.** The expected waiting time λ changes over time. It increases as heat rises, then decreases as kernels run out.
2. **Finite population.** The number of kernels is limited, so there is not an unbounded number of possible pops.

Popcorn as Poisson Process?

Poll. Can we consider popcorn popping a Poisson process?

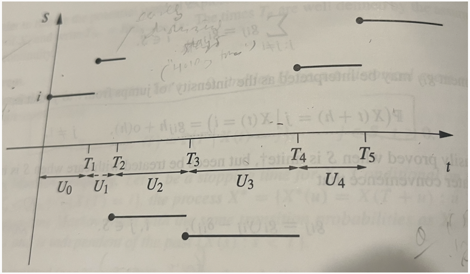
Potential issues.

1. **Non-constant rate.** The expected waiting time λ changes over time. It increases as heat rises, then decreases as kernels run out.
2. **Finite population.** The number of kernels is limited, so there is not an unbounded number of possible pops.
3. **Dependent events.** Although rare, a pop from one kernel could trigger another through physical collision.

Generalizing the Poisson Process

Poll. A continuous time Markov Chain generalizes the Poisson Process. How?

Poll. A continuous time Markov Chain generalizes the Poisson Process. How?

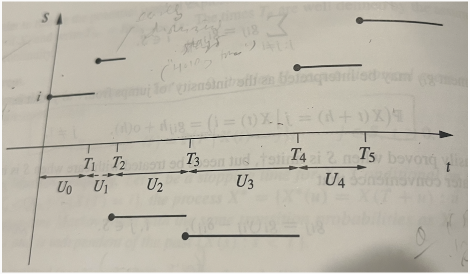


A **continuous time Markov chain** is characterized by

1. A discrete-time **jump chain** with any transition probability matrix (p_{ij}).

Generalizing the Poisson Process

Poll. A continuous time Markov Chain generalizes the Poisson Process. How?

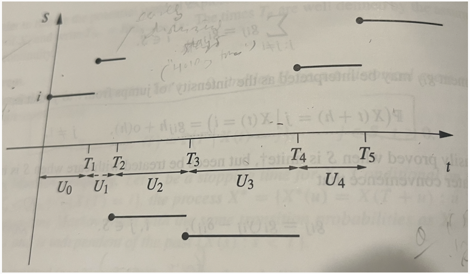


A **continuous time Markov chain** is characterized by

1. A discrete-time **jump chain** with any transition probability matrix (p_{ij}). (The state doesn't need to go up by one each time.)

Generalizing the Poisson Process

Poll. A continuous time Markov Chain generalizes the Poisson Process. How?

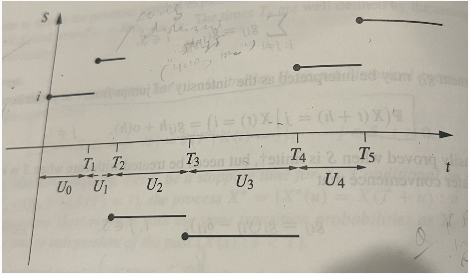


A **continuous time Markov chain** is characterized by

1. A discrete-time **jump chain** with any transition probability matrix (p_{ij}). (The state doesn't need to go up by one each time.)
2. A probability distribution over **holding times** at each state.

Generalizing the Poisson Process

Poll. A continuous time Markov Chain generalizes the Poisson Process. How?



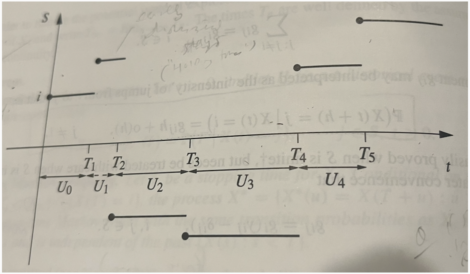
A **continuous time Markov chain** is characterized by

1. A discrete-time **jump chain** with any transition probability matrix (p_{ij}). (The state doesn't need to go up by one each time.)
2. A probability distribution over **holding times** at each state.

   (The waiting time doesn't need to have an identical probability distribution at each state.)

Generalizing the Poisson Process

Poll. A continuous time Markov Chain generalizes the Poisson Process. How?



A **continuous time Markov chain** is characterized by

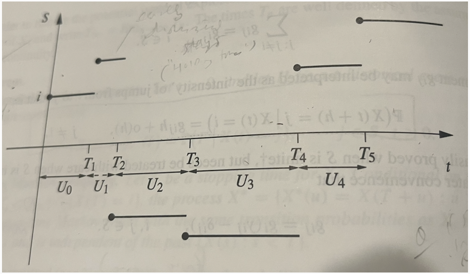
1. A discrete-time **jump chain** with any transition probability matrix (p_{ij}). (The state doesn't need to go up by one each time.)
2. A probability distribution over **holding times** at each state.

   (The waiting time doesn't need to have an identical probability distribution at each state.)

   (However, each such distribution **must** be exponential!

Generalizing the Poisson Process

Poll. A continuous time Markov Chain generalizes the Poisson Process. How?



A **continuous time Markov chain** is characterized by

1. A discrete-time **jump chain** with any transition probability matrix (p_{ij}). (The state doesn't need to go up by one each time.)
2. A probability distribution over **holding times** at each state.

   (The waiting time doesn't need to have an identical probability distribution at each state.)

   (However, each such distribution **must** be exponential! It is characterized by parameter g_i.)

Holding times for a continuous time Markov chain

In the group exercises, we show that the holding times in any continuous time Markov chain must be exponentially distributed!

Holding times for a continuous time Markov chain

In the group exercises, we show that the holding times in any continuous time Markov chain must be exponentially distributed!

This follows from the following characterization of exponential functions:

Fact from Problem 4.14.5a

Let $g : [0, \infty) \to [0, \infty)$ be such that $g(s+t) = g(s)g(t)$ for $s, t \geq 0$. If g is monotone, then $g(s) = e^{\mu s}$ for some μ.

Holding times for a continuous time Markov chain

In the group exercises, we show that the holding times in any continuous time Markov chain must be exponentially distributed!

This follows from the following characterization of exponential functions:

Fact from Problem 4.14.5a

Let $g : [0, \infty) \to [0, \infty)$ be such that $g(s+t) = g(s)g(t)$ for $s, t \geq 0$. If g is monotone, then $g(s) = e^{\mu s}$ for some μ.

which in turn implies:

[title

The **exponential distribution** is the only continuous distribution with the 'lack of memory' property

$$P(X > s + t \mid X > s) = P(X > t).$$

Random Groups

Aubrey Williams: 5
Austin Barton : 3
Blake Sigmundstad: 7
Diego Moylan: 6
Dillon Shaffer: 1
Ismoiljon Muzaffarov: 2
Jacob Tanner: 8

Josh Stoneback: 4
Joshua Bowen: 1
Joshua Culwell: 3
Laura Banaszewski: 1
Lina Hammel: 6
Logan Racz: 2

Matt Hall: 5
Micah Miller: 8
Mike Kadoshnikov: 4
Owen Cool: 2
Racquel Bowen: 4
Samuel Mocabee: 7
Tatiana Kirillova: 3

Group exercises - Problem Set 10

1. (6.5) **The Ehrenfest model.** Find the stationary distribution for the Ehrenfest model of diffusion. (Hint: Rather than solve the equation $\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P}$, consider that the diffusion model might be reversible in equilibrium. That is, look for solutions to the detailed balance equation $\pi_i p_{ij} = \pi_j p_{ji}$.)
2. (6.6.9) **The 'Page Rank' Markov chain.** Recall the construction of a random walk on the world wide web, whose transition matrix is $\boldsymbol{P} = b\boldsymbol{Q} + (1-b)\boldsymbol{u}^\top \boldsymbol{e}$, where $\boldsymbol{Q} = (q_{ij})$ with

$$q_{ij} = \begin{cases} 1/d_i, & \text{if } i \to j \\ 1/n, & \text{if } i \text{ dangles } (d_i = 0) \\ 0, & \text{otherwise} \end{cases}$$

   where n is the number of webpages, d_i is the outgoing degree of the i-th webpage, $\boldsymbol{u}$ is a row vector with $u_i = 1/n$, and $\boldsymbol{e}$ is a row vector of 1's.
   a) Deduce that the stationary distribution $\boldsymbol{\pi}$ is given by $\boldsymbol{\pi} = \{(1-b)/n\}\boldsymbol{e}(\boldsymbol{I} - b\boldsymbol{Q})^{-1}$, where $\boldsymbol{I}$ is the identity matrix.
   b) Explain why the elements of $\boldsymbol{\pi}$, when re-arranged in decreasing order, supply a description of the relative popularities of webpages.
3. (6.9) **Holding times.** Let X be a continuous time Markov chain. Let $X(0) = i$, and $U_0 = \inf\{t : X(t) \neq i\}$ be the time of the first change in value. Show that U_0 has the exponential distribution.