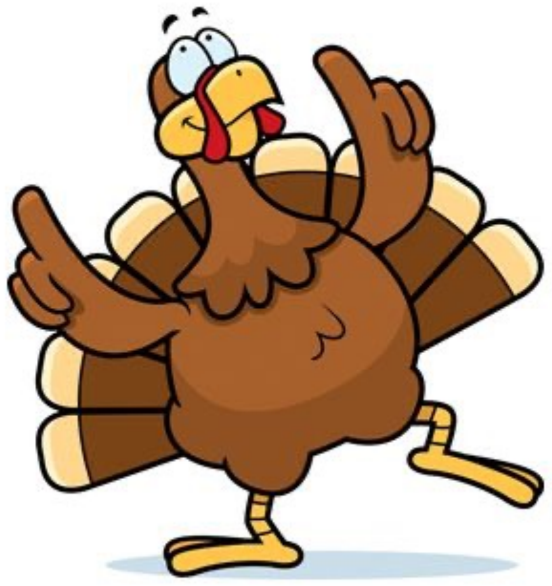
Class 15

Mixing Times and Coupling

Announcements

- Tomorrow (Friday) is deadline for changing to/from CR/NC, and for withdrawing.
- HW7 is out now, due the Friday after break.
- Have a fantastic fall break!



Recap: Mixing Times

$$\Delta(t) = \max_s ||\pi - P_s^t||$$

$$\tau_{mix} = \min\left\{t : \Delta(t) < \frac{1}{2e}\right\}$$

Recap: Coupling

- Motivation:

$$\Delta(t) \leq \max_{s,s'} \left\| P_s^t - P_{s'}^t \right\| \leq 2\Delta(t)$$

- Def of Coupling:
  - X_t, Y_t both individually follow the Markov chain (but may be correlated!)
  - If $X_t = Y_t$, then $X_{t+1} = Y_{t+1}$

- Why we care:

$$\Delta(t) \leq \max_{s,s'} \Pr[X_t \neq Y_t]$$

Aka, $\tau_{mix} \leq$ time for $\{X_t\}$ and $\{Y_t\}$ to meet with probability at least $\frac{1}{2e}$.

Questions?

Mixing times, coupling, quiz?

Q1

- What is τ_{mix} for the MC with transition matrix $\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$?
- Answer: 1. We mix immediately.

Q2

- What about $\begin{pmatrix} 1/4 & 3/4 \\ 1/4 & 3/4 \end{pmatrix}$?

- Still 1. The stationary distribution is (1/4, ¾), and we achieve it immediately.

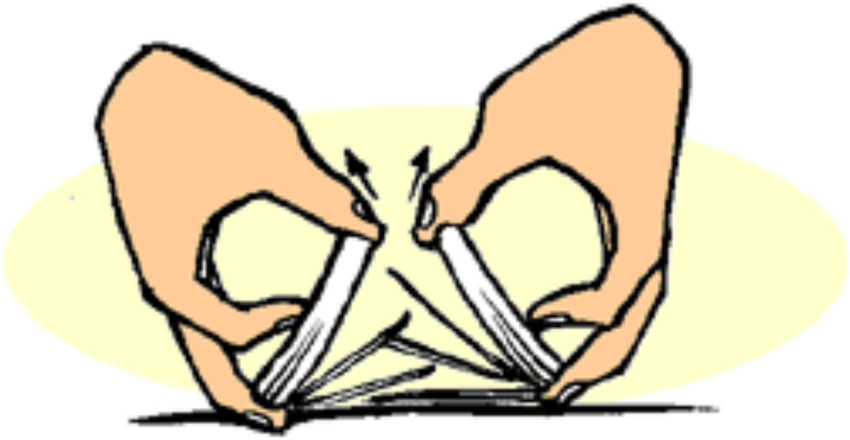
Q3

- In a deck of 52 cards, pick two independently random cards and swap them. Which is a valid coupling for this chain?

- A: Two completely independent random walks.
- B: Two completely independent random walks, UNTIL they hit the same thing, and then they step together.
- C: If we swap index a and index b in deck 1, do the same in deck 2.
  - (Note that this is a terrible coupling, since we NEVER couple if we don’t start at the same point, but it’s still legit).

Today: Shuffling!

- Shuffling procedure:
  - Pick a random card.
  - Move it to the top of the deck.



First Group Work: practice with coupling

- Shuffling procedure:
  - Pick a random card.
  - Move it to the top of the deck.

1. Is this irreducible and aperiodic? What's the stationary dist?
2. Come up with a good coupling.
3. How long until your coupling couples whp?
4. Bound the mixing time.

Q1

- This is aperiodic, irreducible.

- The stationary distribution is ___________________.

Q1

- This is aperiodic, irreducible.
  - Aperiodic since there are self-loops.
  - Irreducible since you can get to any ordering by picking cards to move in the order that you want.
- The stationary distribution is uniform.
  - If you start with a uniformly random deck, and move a random card to the top, it's still uniformly random.

Q2: A good coupling?

A coupling

- We need to come up with $(X_0, Y_0), (X_1, Y_1), \ldots$
- Choose the **same** card in both decks (both the “X” and the “Y” deck).
  - A.k.a., if we move the Ace of Spades in the X deck, also move the Ace of Spades in the Y deck.
- This is a valid coupling:
  - Looking at just one deck, it looks like we are always choosing a uniformly random card and moving it to the top.
  - Once the two decks are the same, they will always stay the same.

Q3: How long until the two chains couple?

How long until the chains couple?

- Observation: The chains have coupled once every card has been chosen.

eg.	DECK 1	DECK 2
	1, 2, 3, 4	4, 3, 1, 2
Pick "1"	2, 3, 4, 1	4, 3, 2, 1
Pick "3"	2, 4, 1, 3	4, 2, 1, 3
Pick "1"	2, 4, 3, 1	4, 2, 3, 1
Pick "4"	2, 3, 1, 4	2, 3, 1, 4
Pick "2"	3, 1, 4, 2	3, 1, 4, 2

COUPLED!

How long until the chains couple?

- Observation: The chains have coupled once every card has been chosen.
- How long until every card has been chosen?
  - This is coupon collecting! About $n \log n$.

$$\max_{s,s'} \mathbb{P}\left[T_{s,s'} \geq 2n\log n\right] \leq \mathbb{P}\left[\text{time to choose all } n \text{ cards} \geq 2n\log n\right] = o(1)$$

time until the chains couple if they start at s, s'.

Translate this to a bound on the mixing time

Translate this to a bound on the mixing time

$$\Delta(2n\log n) \le \max_{s,s'} \mathbb{P}\left[T_{s,s'} \ge 2n\log n\right] = o(1)$$

Where $\Delta(t) = \max_s || \pi - P_s^t ||$ is the total variation distance between the distribution after t steps and the stationary distribution.

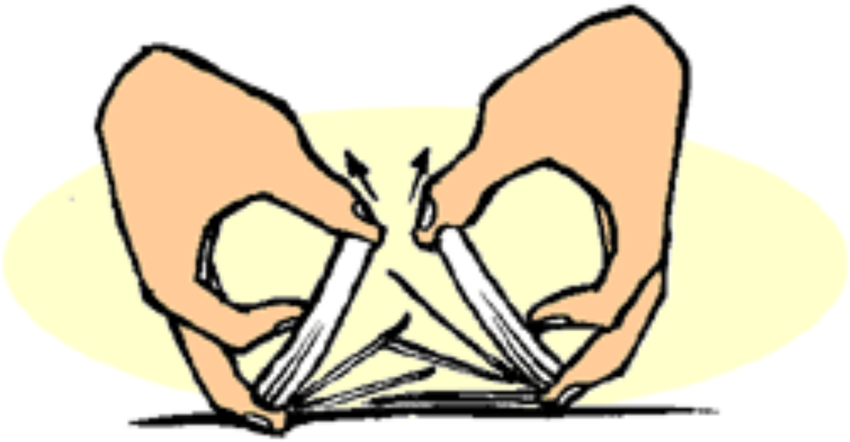
This implies that $\tau_{mix} \le 2n\log n$.

as long as n is big enough that the $o(1)$ term is at most 1/(2e).

Group Work II: More shuffling!

- Another shuffling procedure:
  - Choose the top card.
  - Put it somewhere random.

- We'll analyze this one with a different technique.



Group work!

- Another shuffling procedure:
  - Choose the top card.
  - Put it somewhere random.

1. Convince yourself that this chain is aperiodic and irreducible, with uniform stationary distribution.
2. Let T be the first time the original bottom card is placed randomly somewhere. Show that the deck is completely uniform at any time $t > T$.
3. What is E[T]?
4. Bound the mixing time.

Q1

- This is irreducible and aperiodic.
  - Actually, this was a quiz question last time! It's aperiodic since there are self-loops, and irreducible since you can get any deck you want by building it on the bottom of the deck.
- The stationary distribution is uniform.
  - If you take a uniformly random deck and put the top card somewhere random, it's still a uniformly random deck.

Q1

- This is irreducible and aperiodic.
- The stationary distribution is uniform.

Q2: Show that X_t is uniform if $t > T$

Q2: Show that X_t is uniform if $t > T$

- Let T be the first time at which the original bottom card is randomly placed somewhere in the deck.
  - For convenience, from now on "Ace of Spades" refers to "the original bottom card."
  - So $T - 1$ is the first time that the Ace of Spades gets to the top.

- When $t \geq T$, X_t is uniform.
  - Anything below the Ace of Spades is in a uniformly random order, since the only way it could have gotten there was to be placed completely at random.
  - Once the Ace of Spades is at the top, at time T-1, the rest of the deck is uniform.
  - Once you place the Ace of Spades, at time T, the whole deck is now uniform.

Q3: What is $E[T]$?

Q3: What is $\mathbf{E}[T]$?

$$\mathbb{E}[T] = \mathbb{E}\left[\begin{array}{l}\text{time to move}\\ \text{from bottom}\\ \text{to 2nd-from-bottom}\end{array}\right] + \mathbb{E}\left[\begin{array}{l}\text{time to move from}\\ \text{2nd-from-bottom to}\\ \text{3rd-from-bottom}\end{array}\right] + \cdots$$

↑ move if the top card gets placed below the bottom card.

$\mathbb{P}[\swarrow] = 1/n$, so

$$\mathbb{E}\left[\begin{array}{l}\text{time until}\\ \text{it happens}\end{array}\right] = n$$

↑ Move if the top card gets placed below the 2nd-to-bottom card.

$\mathbb{P}[\swarrow] = 2/n$, so

$$\mathbb{E}\left[\begin{array}{l}\text{time until}\\ \text{it happens}\end{array}\right] = n/2$$

$$= n + \frac{n}{2} + \frac{n}{3} + \cdots + \frac{n}{(n-1)} + 1 \approx n\log(n)$$

Q4: Bound the mixing time by $O(n \log n)$

- Markov's inequality: $\Pr[\, T \geq 2e \cdot E[T]] \leq \frac{1}{2e}$
- Hint: $P_s^t = p \cdot \sigma + (1 - p) \cdot \pi$, where $p = \Pr[\, T > t \,]$

- $\Delta(t) = \max_s ||\pi - P_s^t||_{TV} =$

$$\leq \frac{1}{2e}$$

The mixing time is at most $2e\mathbf{E}[X] = O(n \log n)$

$$P_s^t = \pi \cdot \mathbb{P}[t \geq T] + \sigma \cdot \mathbb{P}[t < T]$$

dist. of X_t given $X_0 = s$

some distribution

$$\Delta(t) = \| \pi - P_s^t \| = \| \pi - \pi \cdot \mathbb{P}[t \geq T] - \sigma \mathbb{P}[t < T] \| = \mathbb{P}[t < T] \, \| \pi - \sigma \|$$

$\pi - \pi \cdot \mathbb{P}[t \geq T] = \pi \cdot \mathbb{P}[t < T]$

$$\leq \mathbb{P}[t < T]$$

The mixing time is at most $2e\mathbf{E}[X] = O(n \log n)$

$$\Delta(t) = \| \pi - P_s^t \| = \| \pi - \pi \cdot \mathbb{P}[t \geq T] - \sigma \mathbb{P}[t < T] \| = \mathbb{P}[t < T] \| \pi - \sigma \| \leq \mathbb{P}[t < T]$$

$$\underbrace{\pi - \pi \cdot \mathbb{P}[t \geq T]}_{\pi \cdot \mathbb{P}[t < T]}$$

$$\Delta(2en \log n) \leq \mathbb{P}[T \geq 2en\log n] \leq 1/2e$$

$$\tau_{mix} \leq 2en \log n.$$

Strong Stationary Stopping Times

- A random variable T is a **strong stationary stopping time** if:
  1. The event that $T = t$ depends only on $X_0, \ldots, X_t$
  2. For all states s, $\Pr[X_t = s \mid T \geq t] = \pi(s)$

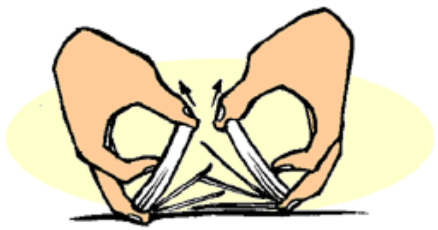
Strong Stationary Stopping Times

- A random variable T is a **strong stationary stopping time** if:
  1. The event that $T = t$ depends only on $X_0, \ldots, X_t$
  2. For all states s, $\Pr[X_t = s \mid T \geq t] = \pi(s)$

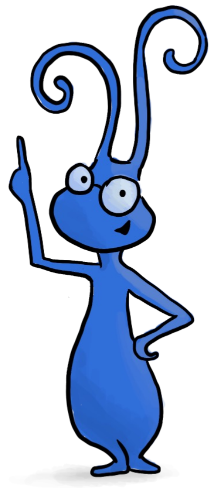
- **Theorem:** For any strong stationary stopping time, $\Delta(t) \leq P[T > t]$.

Bonus group work if time!

- Shuffling procedure:
  - Assign each card “L” or “R” independently, with probability 1/2
  - Put all ”L” cards to the left, preserving their relative order
  - Put all “R” cards to the right, preserving their relative order
  - Put the “L” stack on top of the “R” stack.

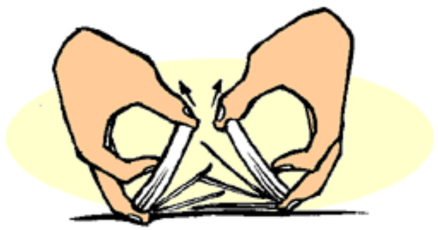


This is the **inverse** of a standard riffle shuffle!



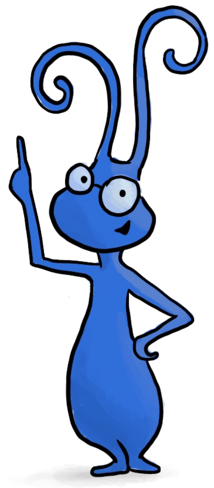
Bonus group work if time!

- Shuffling procedure:
  - Assign each card “L” or “R” independently, with probability 1/2
  - Put all ”L” cards to the left, preserving their relative order
  - Put all “R” cards to the right, preserving their relative order
  - Put the “L” stack on top of the “R” stack.



This is the **inverse** of a standard riffle shuffle!

Group work:
Use the method of strong stationary stopping times to show that the mixing time of this shuffle is $O(\log n)$.



Solution

Solution

- Each card gets a string in $\{L, R\}^t$
- If two cards have different strings, their order relative to each other will be random.
- Let T be the first time that all the cards have different strings.

Solution

- Each card gets a string in $\{L, R\}^t$
- If two cards have different strings, their order relative to each other will be random.
- Let T be the first time that all the cards have different strings.

- $\Pr[\text{ card x and card y have the same string }] = 1/2^t$
- $\Pr[\text{ any two cards have the same string }] \leq n^2 \cdot 2^{-t}$
- Choose $t \geq 3 \log n$ (say) and this is really small.

Recap

- Two ways to bound mixing times:
  1. Coupling
  2. Strong Stationary Mixing Times

- (Plus spectral techniques, which we saw last week)

Have a great fall break!