

CS70: Lecture 37.

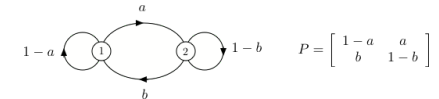
Markov Chains (contd.): First Passage Time: First Step Equations

1. Brief Recap of Markov Chains
2. First Passage Time – First Step Equations
3. Parting Thoughts

Brief Recap of Markov Chains

- ▶ Finite set \mathcal{X} ; $\pi_0; P = \{P(i, j), i, j \in \mathcal{X}\}$;
- ▶ $Pr[X_0 = i] = \pi_0(i), i \in \mathcal{X}$
- ▶ $Pr[X_{n+1} = j | X_0, \dots, X_n = i] = P(i, j), i, j \in \mathcal{X}, n \geq 0$.
- ▶ Note: $Pr[X_0 = i_0, X_1 = i_1, \dots, X_n = i_n] = \pi_0(i_0)P(i_0, i_1) \dots P(i_{n-1}, i_n)$.
- ▶ **Irreducible MC**: every state reachable from every other state
- ▶ State-transition equations: $\pi_{m+1}(j) = \sum_i \pi_m(i)P(i, j), \forall j \in \mathcal{X}$. With π_m, π_{m+1} as row vectors, these identities are written as $\pi_{m+1} = \pi_m P$.
- ▶ $\pi_1 = \pi_0 P, \pi_2 = \pi_1 P = \pi_0 P^2, \dots, \pi_n = \pi_0 P^n, n \geq 0$.
- ▶ **Invariant Distribution**: A distribution π_0 is invariant iff $\pi_0 P = \pi_0$. These equations are called the **balance equations**.
- ▶ **Finite irreducible** Markov Chains have **unique invariant distribution**.
- ▶ Balance Equations: **Prob. Flow In** = **Prob. Flow out** of every state.

Balance Equations: 2-state MC example



$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$$

$$\pi P = \pi \Leftrightarrow [\pi(1), \pi(2)] \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix} = [\pi(1), \pi(2)]$$

$$\Leftrightarrow \pi(1)(1-a) + \pi(2)b = \pi(1) \text{ and } \pi(1)a + \pi(2)(1-b) = \pi(2)$$

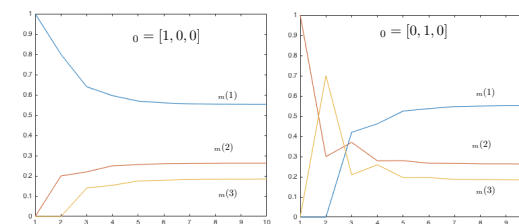
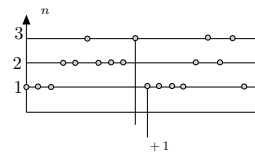
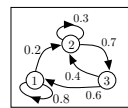
$$\Leftrightarrow \pi(1)a = \pi(2)b.$$

Prob. flow leaving state 1 = **Prob. flow entering state 1**

These equations are redundant! We have to add an equation: $\pi(1) + \pi(2) = 1$. Then we find

$$\pi = \left[\frac{b}{a+b}, \frac{a}{a+b} \right].$$

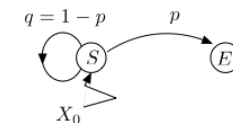
Distribution of X_n



As m increases, π_m converges to a vector that does not depend on π_0 .

First Passage Time - Example 1

Let's flip a coin with $Pr[H] = p$ until we get H . How many flips, on average?



Let $\beta(S)$ be the average time until E , starting from S . Then,

$$\beta(S) = 1 + (q \times \beta(S)) + (p \times 0).$$

(See next slide.) Hence,

$$p\beta(S) = 1, \text{ so that } \beta(S) = 1/p.$$

Note: Time until E is $G(p)$.

We have rediscovered that the mean of $G(p)$ is $1/p$.

OH Example: Balance Eqs.

$P = \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0 & 0.3 & 0.7 \\ 0.6 & 0.4 & 0 \end{bmatrix}$

$\pi_0 P = \pi_0$ (Balance Eqs.)

Let $\pi_0 = [x \ y \ z]$

"FLOW IN = FLOW OUT":

$S_0: 0.6z = 0.2x$

$S_1: 0.2x + 0.4z = 0.7y$

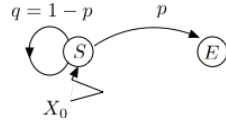
$S_2: 0.7y = z$

Need one more equation as one eqn. above is redundant.

$x + y + z = 1 \Rightarrow$ Solve for $x, y, z \Rightarrow \pi_0 = [0.55 \ 0.26 \ 0.19]$

First Passage Time - Example 1

Let's flip a coin with $Pr[H] = p$ until we get H . Average no. of flips?



Let $\beta(S)$ be the average time until E . Then,

$$\beta(S) = 1 + (q \times \beta(S)) + (p \times 0).$$

Justification:

Let N be the random number of steps until E , starting from S .

Let also N' be the number of steps until E , after the **second** visit to S .

Finally, let $Z = 1\{\text{first flip} = H\} = 1$ if first flip is H and 0 else. Then,

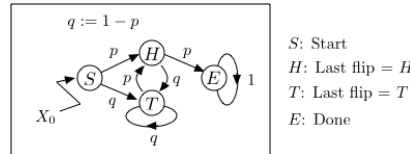
$$N = 1 + (1 - Z) \times N' + Z \times 0.$$

Now, Z and N' are independent. Also, $E[N'] = E[N] = \beta(S)$. Hence, taking expectation of both sides of the equation, we get:

$$\beta(S) = E[N] = 1 + ((1 - p) \times E[N']) + (p \times 0) = 1 + (q \times \beta(S)) + (p \times 0).$$

First Passage Time - Example 2

Let's flip a coin with $Pr[H] = p$ until we get two consecutive H s. How many flips, on average? Here is a picture:



Let $\beta(i)$ be the average time from state i until the MC hits state E .

We claim that (these are called the **first step equations**)

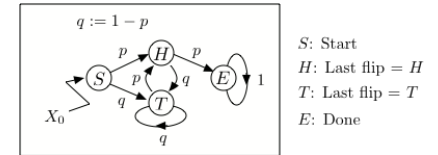
$$\beta(S) = 1 + p\beta(H) + q\beta(T)$$

$$\beta(H) = 1 + p \times 0 + q\beta(T)$$

$$\beta(T) = 1 + p\beta(H) + q\beta(T).$$

Solving, we find $\beta(S) = 2 + 3qp^{-1} + q^2p^{-2}$. (E.g., $\beta(S) = 6$ if $p = 1/2$.)

First Passage Time - Example 2



Let us justify the first step equation for $\beta(T)$. The others are similar.

Let $N(T)$ be the random number of steps, starting from T until the MC hits E . Let also $N(H)$ be defined similarly. Finally, let $N'(T)$ be the number of steps after the second visit to T until the MC hits E . Then,

$$N(T) = 1 + Z \times N(H) + (1 - Z) \times N'(T)$$

where $Z = 1\{\text{first flip in } T \text{ is } H\}$. Since Z and $N(H)$ are independent, and Z and $N'(T)$ are independent, taking expectations, we get

$$E[N(T)] = 1 + pE[N(H)] + qE[N'(T)],$$

i.e.,

$$\beta(T) = 1 + p\beta(H) + q\beta(T).$$

First Passage Time - Example 3: Practice Exercise

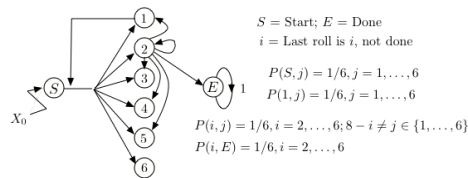
You keep rolling a fair six-sided die until **the sum of the last two rolls is 8**.

Question: How many times do you have to roll the die before you stop, on average?

Spoiler Alert: Solution on next slide

(but don't look: try to do it yourself first!)

Example 3: Practice Exercise Solution



The arrows out of 3, ..., 6 (not shown) are similar to those out of 2.

$$\beta(S) = 1 + \frac{1}{6} \sum_{j=1}^6 \beta(j); \beta(1) = 1 + \frac{1}{6} \sum_{j=1}^6 \beta(j); \beta(i) = 1 + \frac{1}{6} \sum_{j=1, \dots, 6, j \neq 8-i}^6 \beta(j), i=2, \dots, 6.$$

Symmetry: $\beta(2) = \dots = \beta(6) =: \gamma$. Also, $\beta(1) = \beta(S)$. Thus,

$$\beta(S) = 1 + (5/6)\gamma + \beta(S)/6; \quad \gamma = 1 + (4/6)\gamma + (1/6)\beta(S).$$

$$\Rightarrow \dots \beta(S) = 8.4.$$

First Step Equations

Let X_n be a MC on \mathcal{X} and $A \subset \mathcal{X}$. Define

$$T_A = \min\{n \geq 0 \mid X_n \in A\}.$$

Let

$$\beta(i) = E[T_A \mid X_0 = i], i \in \mathcal{X}.$$

The FSE are

$$\beta(i) = 0, i \in A$$

$$\beta(i) = 1 + \sum_j P(i, j) \beta(j), i \notin A$$

Summary

Markov Chains

- ▶ Markov Chain: $Pr[X_{n+1} = j | X_0, \dots, X_n = i] = P(i, j)$
- ▶ First Passage Time:
 - ▶ $A \subset \mathcal{X}; \beta(i) = E[T_A | X_0 = i]$;
 - ▶ $\beta(i) = 1 + \sum_j P(i, j)\beta(j)$;
- ▶ FSE: $\beta(i) = 1 + \sum_j P(i, j)\beta(j)$;
- ▶ $\pi_n = \pi_0 P^n$
- ▶ π is invariant iff $\pi P = \pi$
- ▶ Irreducible \Rightarrow one and only one invariant distribution π

Confusing Statistics: Simpson's Paradox

College	F. Appl.	F. Adm.	% F. Adm.	M. Appl.	M. Adm.	% M. Adm.
A	980	490	50%	200	80	40%
B	20	20	100%	800	720	90%
Total	1000	510	51%	1000	800	80%

The numbers are applications and admissions of males and females to the two colleges of a university.

Overall, the admission rate of male students is **80%** whereas it is only **51%** for female students.

A closer look shows that the admission rate is **larger** for female students **in both colleges**....

Female students happen to apply more to the college that admits fewer students.

Probability part of the course: key takeaways?

What should I take away about probability from this course?

I mean, **after the final**?

- ▶ Given the uncertainty around us, we should understand some probability. **"Being precise about being imprecise."**
- ▶ 4 key concepts:
 1. Learn from **observations** to revise our biases, given by the **role of the prior**; **Bayes' Theorem**;
 2. **Confidence Intervals**: CLT, Cheybyshv Bounds, WLLN.
 3. **Regression/Estimation**: $L[Y|X], E[Y|X]$
 4. **Markov Chains**: Sequence of RVs, $P[X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots] = P[X_{n+1} = x_{n+1} | X_n = x_n]$, Balance Equations.
- ▶ **Quantifying our degree of certainty**. This clear thinking invites us to question vague statements, and to convert them into precise ideas.

More on Confusing Statistics

Statistics are often confusing:

- ▶ The average household annual income in the US is \$72k. Yes, but the median is \$52k.
- ▶ The false alarm rate for prostate cancer is only 1%. Great, but only 1 person in 8,000 has that cancer. So, there are 80 false alarms for each actual case.
- ▶ The Texas sharpshooter fallacy. Look at people living close to power lines. You find clusters of cancers. You will also find such clusters when looking at people eating kale.
- ▶ False causation. Vaccines cause autism. Both vaccination and autism rates increased....
- ▶ Beware of statistics reported in the media!

Random Thoughts

Famous Quotes: French mathematician Pierre-Simon Laplace (Translated from French):

"The theory of probabilities is basically just common sense reduced to calculus"

Famous Quotes: Attributed by Mark Twain to British Prime Minister Benjamin Disraeli:

"There are three kinds of lies: lies, damned lies, and statistics."

Confirmation Bias

Confirmation bias is the tendency to search for, interpret, and recall information in a way that confirms one's beliefs or hypotheses, while giving disproportionately less consideration to alternative possibilities.

Confirmation biases contribute to **overconfidence in personal beliefs** and can maintain or strengthen beliefs **in the face of contrary evidence**.

Three aspects:

- ▶ **Biased search** for information. E.g., ignoring articles that dispute your beliefs.
- ▶ **Biased interpretation**. E.g., putting more weight on confirmation than on contrary evidence.
- ▶ **Biased memory**. E.g., remembering facts that confirm your beliefs and forgetting others.

Confirmation Bias: An experiment

There are two bags. One with 60% red balls and 40% blue balls; the other with the opposite fractions.

One selects one of the two bags.

As one draws balls one at time, one asks people to declare whether they think one draws from the first or second bag.

Surprisingly, people tend to be reinforced in their original belief, even when the evidence accumulates against it.

Parting Thoughts

You have worked hard and learned a lot in this course!

Proofs, Graphs, Stable Marriage, Mod(p), RSA, Reed-Solomon, Decidability, Probability, ... , HW option or Test-only option? how to handle stress, how to sleep less, how to keep smiling, ...
Difficult course? Yes! Useful? You bet!
Finally,

THANK YOU on behalf of Prof. Rao and me for persevering through this course! It has been an absolute pleasure!

Let us also not forget to thank the dedicated EECS70 Staff:

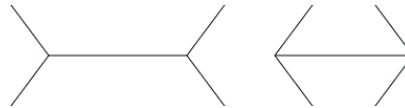
- ▶ The Thrilling TAs
- ▶ The Terrific Tutors
- ▶ The Rigorous Readers
- ▶ The Amazing Assistants

GOOD LUCK IN YOUR FINAL EXAM!!!

Being Rational: 'Thinking, Fast and Slow'

In this book, Daniel Kahneman discusses examples of our irrationality. Here are a few examples:

- ▶ A judge rolls a die in the morning. In the afternoon, he has to sentence a criminal. Statistically, the sentence tends to be heavier if the outcome of the morning roll was high.
- ▶ People tend to be more convinced by articles printed in Times Roman instead of Computer Modern Sans Serif.
- ▶ Perception illusions: Which horizontal line is longer?



It is difficult to think clearly!

What's Next?

Professors, I loved this course so much! I want to learn more about discrete math and probability!

Funny you should ask! How about

- ▶ CS170: Efficient Algorithms and Intractable Problems a.k.a. Introduction to CS Theory: Graphs, Dynamic Programming, Complexity.
- ▶ EECS126: Probability in EECS: An Applications-Driven Course: PageRank, Digital Links, Tracking, Speech Recognition, Planning, etc. Hands on labs with python experiments (GPS, Auctions, Kalman Filtering, RNA sequencing, ...).
- ▶ CS188: Artificial Intelligence: Hidden Markov Chains, Bayes Networks, Neural Networks.
- ▶ CS189: Introduction to Machine Learning: Regression, Neural Networks, Learning, etc. Programming experiments with real-world applications.