Lecture 5
Stochastic Process

- Markov Chain
- Martingale and Random Walk

Contents

- Markov Chain
- Martingale and Random Walk

Definition

- A Markov chain is a sequence of random variables $X_0, X_1, \ldots, X_n, \ldots$ with the Markov property that **given the present state, the future states and the past states are independent**:

$$\mathbf{Pr}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0)$$
$$= p_{i,j} = \mathbf{Pr}(X_{n+1} = j \mid X_n = i).$$

- State space: a set containing all possible values that X_n can take.
- Once the current state is known, past history has no bearing on the future.
- Homogeneous Markov chain: transition probability from state i to state j **does not** depend on n. (We will focus on homogeneous Markov chains in this lecture).

Example

$$
\begin{array}{c|c}
 & k_{n+1} \\
\begin{array}{c} k_n = 1 \\ k_n = 2 \\ k_n = 3 \\ k_n = 4 \\ k_n = 5 \\ k_n = 6 \end{array} &
\begin{pmatrix}
1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\
0 & 2/6 & 1/6 & 1/6 & 1/6 & 1/6 \\
0 & 0 & 3/6 & 1/6 & 1/6 & 1/6 \\
0 & 0 & & \ddots & & \\
0 & 0 & & & & \\
 & & & & & 6/6
\end{pmatrix}
\end{array}
$$

- A six-sided die is rolled repeatedly.
- After each roll $n \geq 1$, let X_n be the largest number rolled in the first n rolls.
- Is $\{X_n : n \geq 1\}$ a discrete-time Markov chain?

$S = \{1, 2, 3, 4, 5, 6\}$

$$
\Pr\{X_{n+1} = k_{n+1} \mid X_n = k_n) = \begin{cases} 0, & \text{when } k_n > k_{n+1} \\ \frac{k_n}{6}, & \text{when } k_n = k_{n+1} \\ \frac{1}{6}, & \text{when } k_n < k_{n+1} \end{cases}
$$

Transition Matrix

- Transition matrix:

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1M} \\ p_{21} & p_{22} & \cdots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MM} \end{pmatrix}$$

  where p_{ij} is the transition probability from state i to state j.

- Initial probabilities: $P(X_0) = (\mathbf{Pr}(X_0 = 1), \ldots, \mathbf{Pr}(X_0 = M))$, where $\sum_{i=1}^{M} \mathbf{Pr}(X_0 = i) = 1$.

- Probability of a path:

  $= \Pr(X_1 = i_1 \mid X_0 = i_0) \Pr(X_2 = i_2, \cdots, X_n = i_n \mid \cancel{X_0 = i_0}, X_1 = i_1)$

$$\mathbf{Pr}(X_1 = i_1, X_2 = i_2, \ldots, X_n = i_n \mid X_0 = i_0) = p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n}.$$

Example

- Suppose that your daily mood only depends on your mood on the previous day.
- Your mood has three distinctive states:

Happy (0), So-so (1), Gloomy (2).

- Transition probability matrix

$$P = \begin{pmatrix} 0.7 & 0.3 & 0 \\ 0.3 & 0.5 & 0.2 \\ 0 & 0.9 & 0.1 \end{pmatrix}$$

- Compute $\mathbf{Pr}\{X_{n+1} = 2 \mid X_n = 0\}$ and $\mathbf{Pr}\{X_{n+2} = 2 \mid X_n = 0\}$.

Example

Suppose you have a discrete time Markov chain with state space $\mathcal{S} = \{1, 2, 3\}$ and with the following transition probability matrix

$$P = \begin{pmatrix} 0.3 & 0.2 & 0.5 \\ 0.6 & 0 & 0.4 \\ 0 & 0.8 & 0.2 \end{pmatrix}.$$

$a = (0.5, 0.2, 0.3)$

Suppose we know the initial probability distribution is $(0.5, 0.2, 0.3)$.

How can we compute

$(P^{10})_{3,1}$

$= \sum_{i=1}^{3} Pr\{X_{10} = 1, X_0 = i\}$

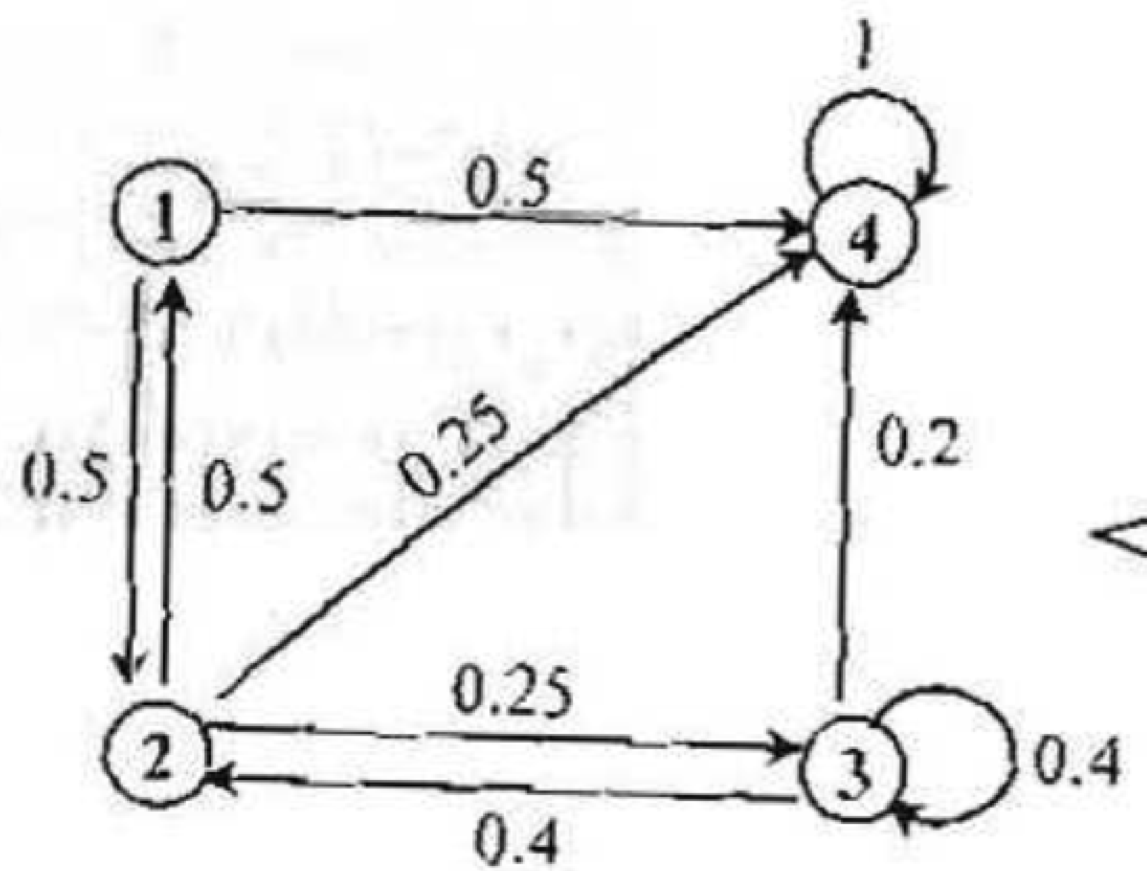
$$\mathbf{Pr}\{X_{10} = 1 \mid X_0 = 3\}, \quad \mathbf{Pr}\{X_{10} = 1\}?$$

$= \sum_{i=1}^{3} Pr\{X_{10} = 1 \mid X_0 = i\} \, Pr\{X_0 = i\}$

$= \sum_{i=1}^{3} (P^{10})_{i,1} \, a_i$

Transition graph

- Transition graph is used to express the transition matrix graphically:



$$
\begin{array}{cc}
 & \begin{array}{cccc} i = 1 & 2 & 3 & 4 \end{array} \quad j = \\
P = & \begin{bmatrix} 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0.25 & 0.25 \\ 0 & 0.4 & 0.4 & 0.2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array}
\end{array}
$$

Figure 5.1 Transition graph and transition matrix of the Play

Classification of states

- We say state j is **accessible** from state i if either one of the following conditions hold:
  1. There is a directed path in the transition graph from i to j;
  2. $(P^n)_{ij} > 0$;
  3. Denote by $T_{ij} = \min(n : \ X_n = j \mid X_0 = i)$ and $\mathbf{Pr}(T_{ij} < \infty) > 0$.
- We say state i and j communicate if i is accessible from j and j is accessible from i.

Classification of states

- We say state j is **accessible** from state i if either one of the following conditions hold:

  1. There is a directed path in the transition graph from i to j;
  2. $(P^n)_{ij} > 0$;
  3. Denote by $T_{ij} = \min(n: \ X_n = j \mid X_0 = i)$ and $\mathbf{Pr}(T_{ij} < \infty) > 0$.

- We say state i and j communicate if i is accessible from j and j is accessible from i.

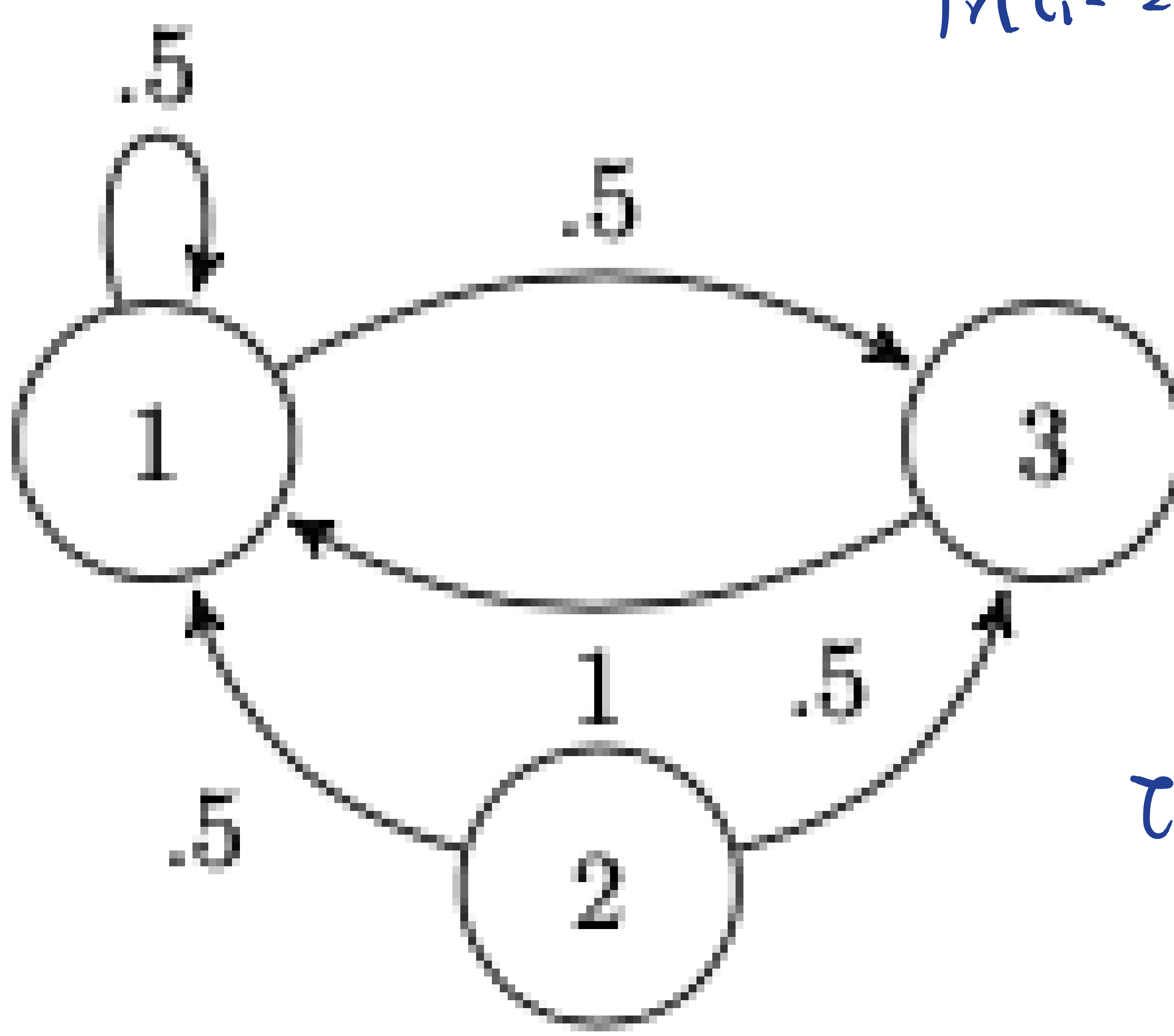
Classification of states

- For each state $i \in \mathcal{S}$, let τ_i denote the first $n \geq 1$ such that $X_n = i$.
- State i is said to be *recurrent* if $\mathbf{Pr}(\tau_i < \infty \mid X_0 = i) = 1$.
- State i is said to be *transient* if it is not recurrent, i.e.,

  $\mathbf{Pr}(\tau_i < \infty \mid X_0 = i) < 1$.

Example

Is state $1, 2, 3$ recurrent or transient?



τ_1: first $n \geq 1$ s.t. $X_n = 1$

$Pr(\tau_1 = 1 \mid X_0 = 1) = \frac{1}{2}$

$Pr(\tau_1 = 2 \mid X_0 = 1) = \frac{1}{2}$

$Pr(\tau_1 < \infty \mid X_0 = 1)$

$= Pr(\tau_1 = 1 \mid X_0 = 1) + Pr(\tau_1 = 2 \mid X_0 = 1)$

$= 1$

τ_3: first $n \geq 1$ s.t. $X_n = 3$

$Pr(\tau_3 = 1 \mid X_0 = 3) = 0$

$Pr(\tau_3 = 2 \mid X_0 = 3) = \frac{1}{2}$

$Pr(\tau_3 = 3 \mid X_0 = 3) = (\frac{1}{2})^2$

$Pr(\tau_3 = m \mid X_0 = 3) = (\frac{1}{2})^{m-1}$

$3 \to 1 \to 1 \to 1 \to \cdots \to 3$

$Pr(\tau_3 < \infty \mid X_0 = 3)$

$= \sum_{m \geq 1} Pr(\tau_3 = m \mid X_0 = 3)$

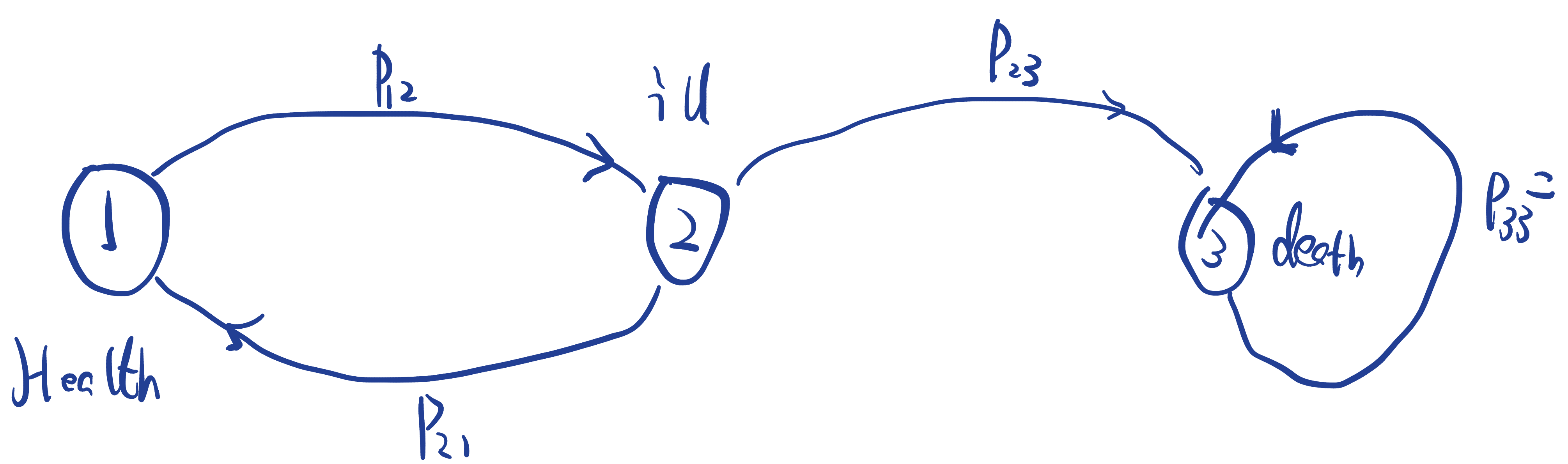
$= \sum_{m \geq 1} (\frac{1}{2})^{m-1} = 1$

τ_2: first $n \geq 1$ s.t. $X_n = 2$

$Pr(\tau_2 < \infty \mid X_0 = 2) = 0$

Absorbing Markov Chains

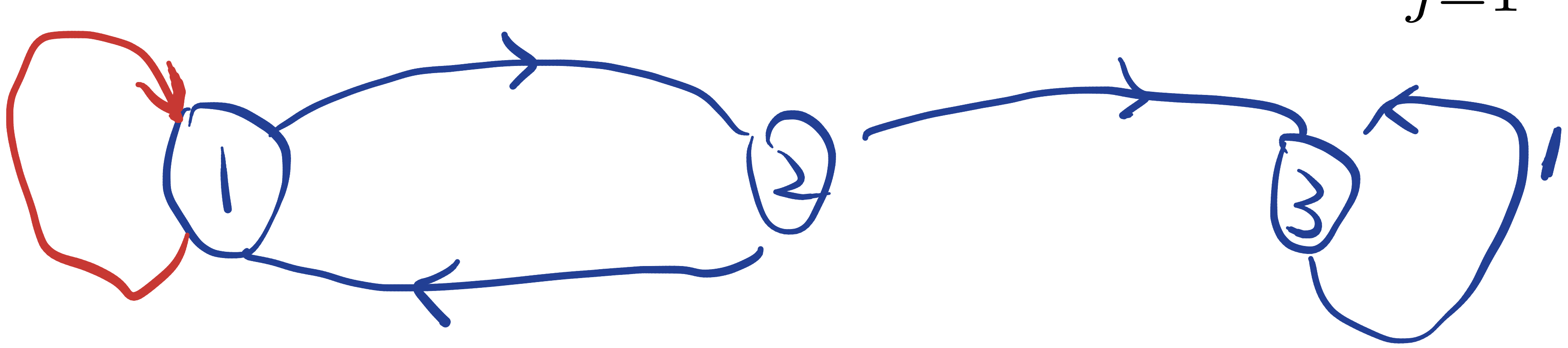
- A state i is called absorbing if it is impossible to leave this state $(p_{ii} = 1, p_{ij} = 0, \forall j \neq i)$.
- A Markov chain is absorbing if
  1. it has at least one absorbing state
  2. and if from every state it is possible to go to an absorbing state.



Probability and expected time to absorption

- Suppose the state space $\mathcal{S} = \{1, 2, \ldots, M\}$
- From state $i \in \mathcal{S}$, denote the probability to reach a specific absorbing state s as a_i.
- It holds that $a_s = 1$ and for all absorbing states $i \neq s$, $a_i = 0$.
- For all transient states i,

$$a_i = \sum_{j=1}^{M} a_j p_{ij}.$$



$a_3 = 1$

$a_1 = a_2 P_{12} + a_1 P_{11}$

$a_2 = a_1 P_{21} + a_3 P_{23}$

$= P_{23} + a_1 P_{21}$

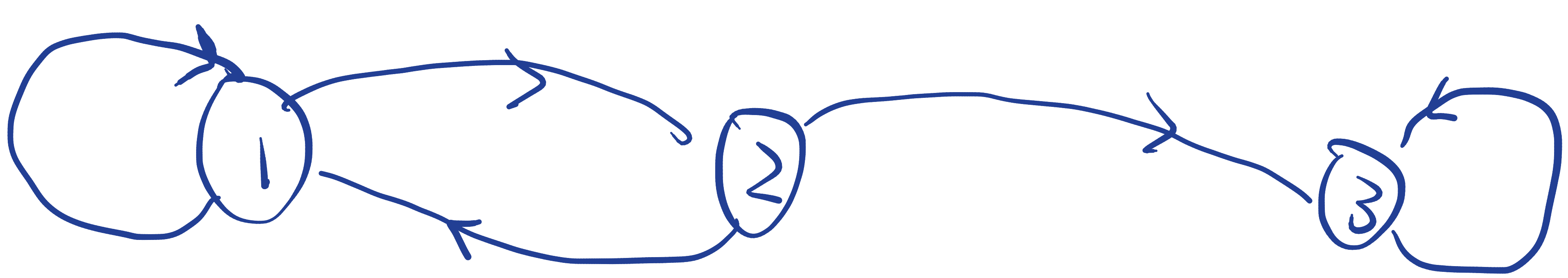
Probability and expected time to absorption

- Suppose the state space $\mathcal{S} = \{1, 2, \ldots, M\}$
- From state $i \in \mathcal{S}$, denote the expected times to absorption as $\mu_1, \ldots, \mu_M$.
- $\{\mu_i\}_{i \in \mathcal{S}}$ is the unique solution to the system of equations
  - $\mu_i = 0$ for all absorbing state(s) i
  - For all transient states i, $\cancel{\mu_i = 1 + \sum_{j=1}^{M} P_{ij}\mu_i}$.

$\mu_3 = 0$

$\mu_i = 1 + \sum_{j=1}^{M} P_{ij}\mu_j$

$\mu_1 = 1 + \mu_1 P_{11} + \mu_2 P_{12}$



$\mu_2 = 1 + \mu_1 P_{21} + \cancel{\mu_3 P_{23}}$

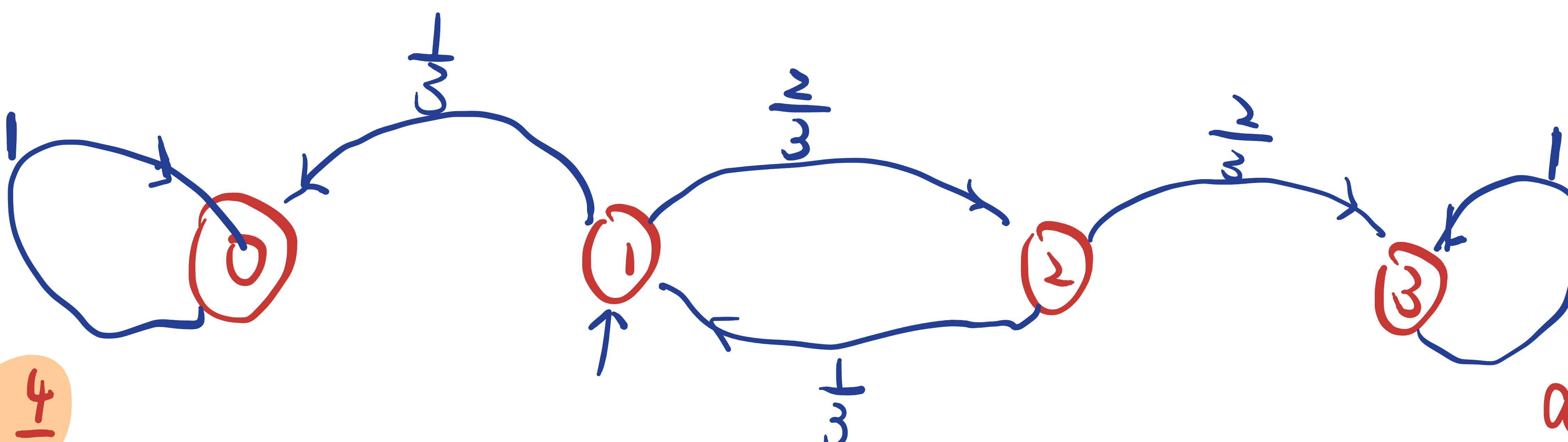
Example

$S = \{(3,0), (2,1), (1,2), (0,3)\}$

$S = \{\#\text{ of money M have}\} = \{0, 1, 2, 3\}$.

a_i: Prob that M reaches state 3 starting from state i.

Player M has \$1 and player N has \$2. Each game gives the winner \$1 from the other. As a better player, M wins $2/3$ of the games. They play until one of them is bankrupt. What is the probability that M wins?



$a_0 = 0$

$a_3 = 1$

$a_1 = \frac{1}{3} a_0 + \frac{2}{3} a_2 = \frac{2}{3} a_2$

$a_2 = \frac{1}{3} a_1 + \frac{2}{3} a_3$

$= \frac{2}{3} + \frac{1}{3} a_1$

$a_2 = \frac{2}{3} + \frac{1}{3}\cdot\frac{2}{3} a_2 \iff a_2 - \frac{2}{9} a_2 = \frac{2}{3} \iff \frac{7}{9} a_2 = \frac{2}{3}$

$\iff a_2 = \frac{2}{3} \times \frac{9}{7} = \frac{6}{7}$

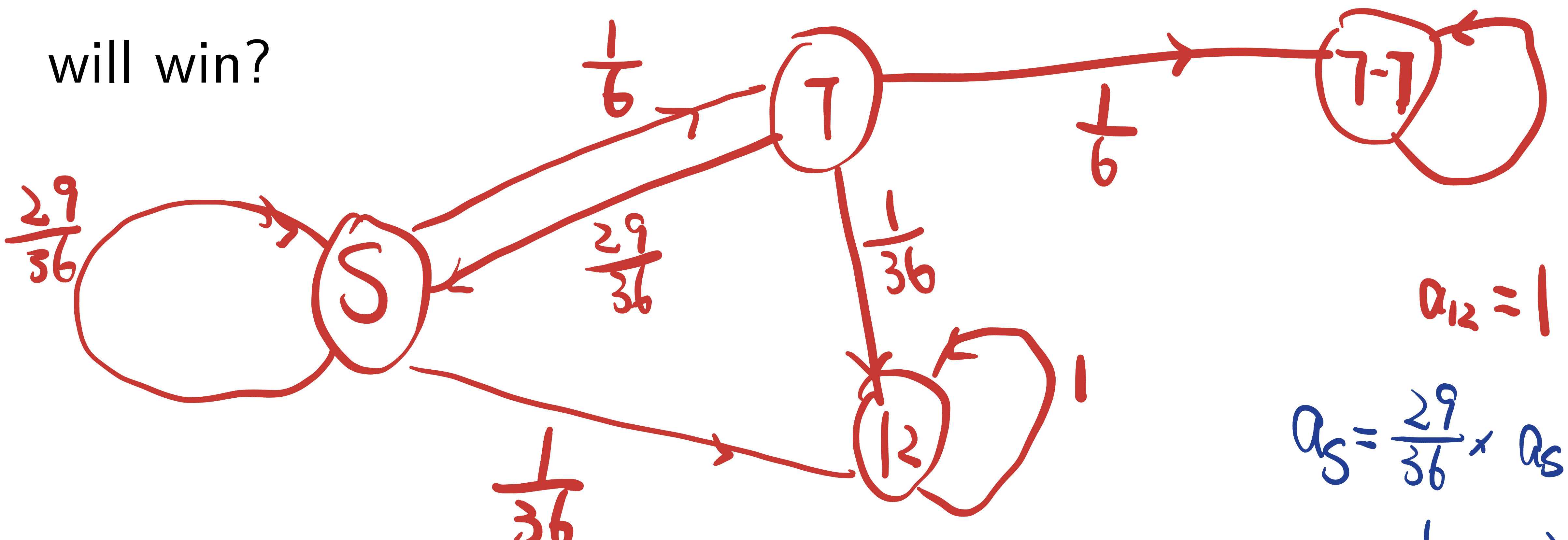
$a_1 = \frac{2}{3} a_2 = \frac{4}{7}$

Example

Catan Island.

$a_S = \frac{7}{13}$

Two players bet on roll(s) of the total of two standard six-face dice. Player A bets that a sum of 12 will occur first. Player B bets that two consecutive 7s will occur first. The players keep rolling the dice and record the sums until one player wins. What is the probability that A will win?

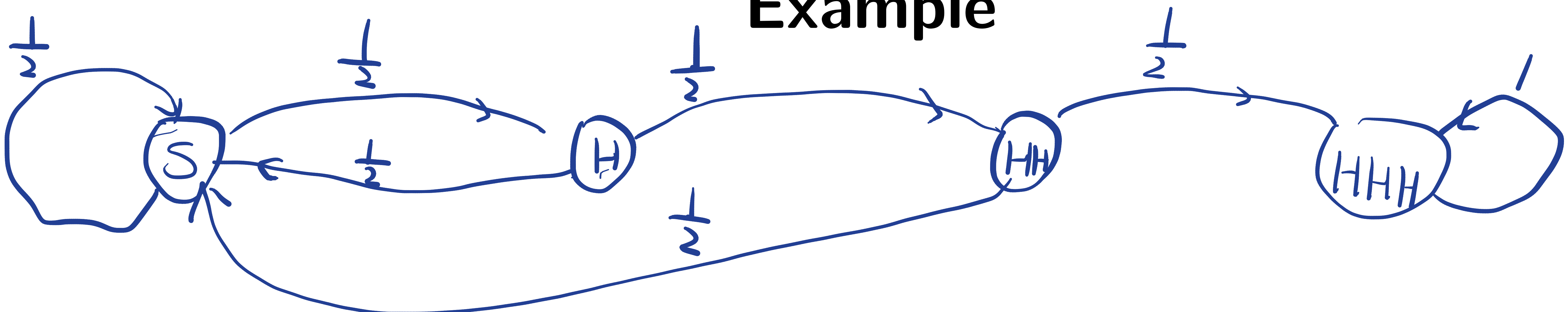


$a_{12} = 1 \qquad a_{7\text{-}7} = 0$

$a_S = \frac{29}{36} \times a_S + \frac{1}{36} + \frac{1}{6} a_7$

$a_7 = \frac{1}{36} + \frac{29}{36} \times a_S$

Example



- Suppose you keep on tossing a fair coin. What is the expected number of tosses such that you can have HHH (heads heads heads) in a row? What is the expected number of tosses to have THH (tails heads heads) in a row?

μ_i: expected # of time to HHH starting from i.

- Keep flipping a fair coin until either HHH or THH occurs in the sequence. What is the probability that you get an HHH subsequence before THH?

$\mu_{HHH} = 0$

$\mu_S = 1 + \frac{1}{2}\mu_S + \frac{1}{2}\mu_H \Rightarrow$

$\mu_H = 1 + \frac{1}{2}\mu_S + \frac{1}{2}\mu_{HH}$

$\mu_{HH} = 1 + \frac{1}{2}\mu_S + \frac{1}{2}\mu_{HHH}$

$\mu_{HH} = 1 + \frac{1}{2}\mu_S$

$\mu_H = 1 + \frac{1}{2}\mu_S + \frac{1}{2}(1 + \frac{1}{2}\mu_S)$

$= \frac{3}{2} + \frac{3}{4}\mu_S$

$\mu_S = 1 + \frac{1}{2}\mu_S + \frac{1}{2}(\frac{3}{2} + \frac{3}{4}\mu_S)$

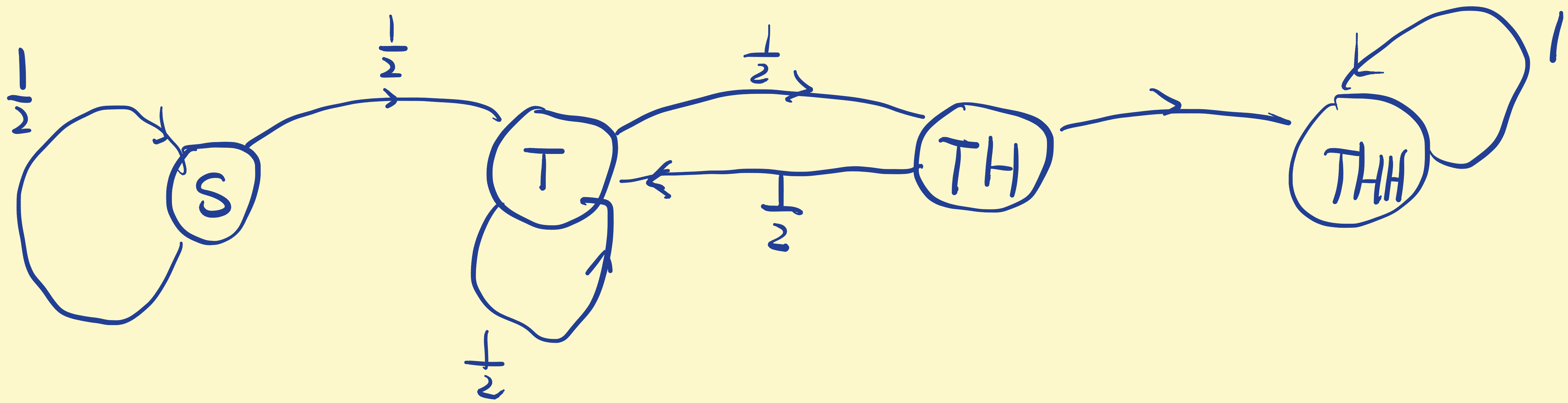
$= 1 + \frac{1}{2}\mu_S + \frac{3}{4} + \frac{3}{8}\mu_S$

$\Rightarrow$

$\mu_S = \frac{7}{4} + \frac{7}{8}\mu_S$

$\frac{1}{8}\mu_S = \frac{7}{4} \Rightarrow \mu_S = 14$

of tosses to have THH ?



$$\mu_{THH} = 0$$

$$\left.\begin{aligned}
\mu_S &= 1 + \frac{1}{2}\mu_S + \frac{1}{2}\mu_T \\
\mu_T &= 1 + \frac{1}{2}\mu_T + \frac{1}{2}\mu_{TH} \\
\mu_{TH} &= 1 + \frac{1}{2}\mu_T + \frac{1}{2}\mu_{THH}.
\end{aligned}\right\} \Rightarrow \quad \mu_S = 8$$

$a_{HHH}=1 \qquad a_{THH}=0$

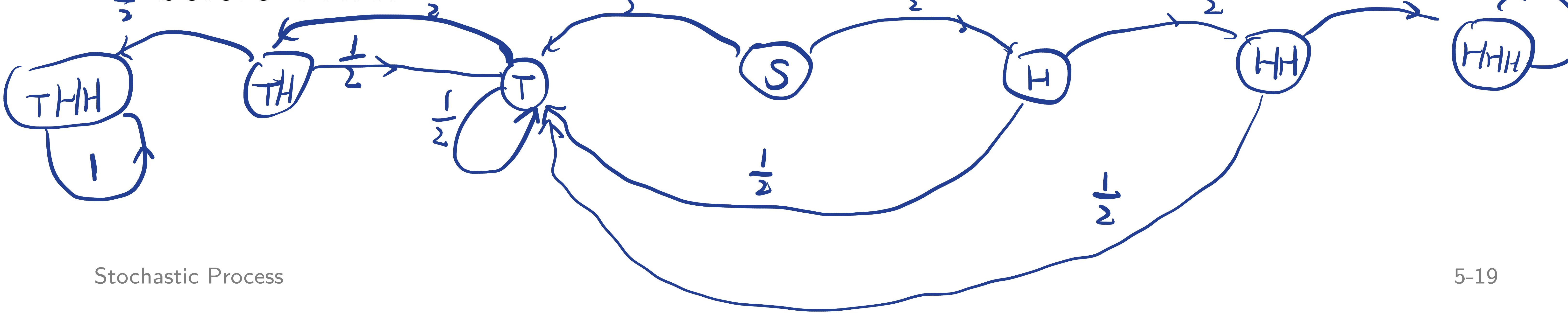
$a_S = \frac{1}{2} a_T + \frac{1}{2} a_H$

$a_T = \frac{1}{2} a_T + \frac{1}{2} a_{TH} \qquad a_{TH} = \frac{1}{2} a_T + \frac{1}{2} a_{THH} \qquad \Rightarrow \qquad a_S = \frac{1}{8}$

$a_H = \frac{1}{2} a_T + \frac{1}{2} a_{HH} \qquad a_{HH} = \frac{1}{2} a_T + \frac{1}{2} a_{HHH}$

Example

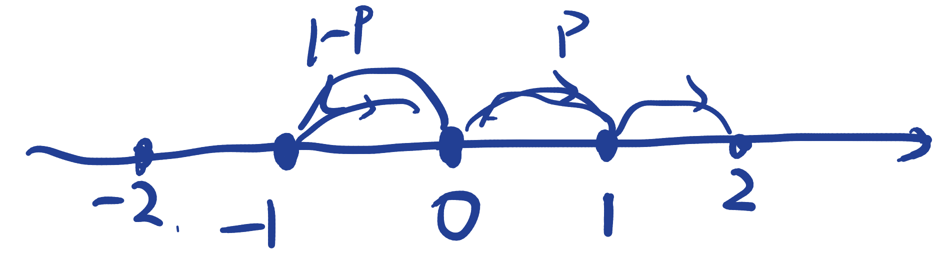
- Suppose you keep on tossing a fair coin. What is the expected number of tosses such that you can have HHH (heads heads heads) in a row? What is the expected number of tosses to have THH (tails heads heads) in a row?
- Keep flipping a fair coin until either HHH or THH occurs in the sequence. What is the probability that you get an HHH subsequence before THH?



Contents

- Markov Chain
- Martingale and Random Walk

Definition



$S_n = X_1 + \cdots + X_n.$

- Random Walk: The process $\{S_n\}_{n\geq 1}$ is called a random walk if $\{X_i\}_{i\geq 1}$ are i.i.d. random variables and $S_n = X_1 + \cdots + X_n$.
- We can imagine that S_n is the position at time n for a walker who makes successive random steps $X_1, X_2, \ldots$.
- Simple random walk: X_i takes values $1, -1$ with probabilities $p, 1-p$, respectively.
- For $p = 1/2$, S_n is called a symmetric random walk. Then,

$$\mathbb{E}[S_n] = 0, \quad \mathrm{Var}(S_n) = \mathbb{E}[S_n^2] = n.$$

$\mathrm{Var}(S_n) = E[S_n^2] - \cancel{(E[S_n])^2}$

$= E[(X_1 + \cdots + X_n)^2]$

$= \sum_{k=1}^{n} E[X_k^2] + \sum_{i=1}^{n} \sum_{j \neq i} \cancel{E[X_i X_j]}$

$= n$

- Typical interview questions: Find first n that S_n reaches a defined threshold α, or probability that S_n reaches α for a given value of n.

Martingale

- A martingale $\{Z_n\}_{n\geq 1}$ is a stochastic process with properties that $\mathbb{E}[|Z_n|] < \infty$ for all n and $\mathbb{E}[Z_{n+1} \mid Z_n = z_n, Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1] = z_n$.
- Martingale also satisfies

$$\mathbb{E}[Z_m \mid Z_n = z_n, Z_{n-1} = z_{n-1}, \ldots, Z_1 = z_1] = z_n, \quad m > n,$$

which means the conditional expected value of future Z_m is the current value Z_n.

Example

$S_n = X_1 + \cdots + X_n$

$X_i = \begin{cases} 1, & \text{w.p. } 1/2 \\ -1, & \text{w.p. } 1/2, \end{cases}$

- A symmetric random walk is a martingale.

$\{S_n\}_n$ is a martingale.

$E[|S_n|] = \sum_{i=1}^{n} E[|X_i|] = n < \infty$

- We have $S_{n+1} = \begin{cases} S_n + 1, & \text{w.p. } 1/2 \\ S_n - 1, & \text{w.p. } 1/2. \end{cases}$
- Thus $\mathbb{E}[S_{n+1} \mid S_n = s_n, \ldots, S_1 = s_1] = s_n$.
- We can verify that $S_n^2 - n$ is also a martingale.

$S_{n+1} = \begin{cases} S_n + 1, & \text{w.p. } 1/2 \\ S_n - 1 & \text{w.p. } 1/2, \end{cases}$

$\Rightarrow E[\, S_{n+1} \mid S_n = s_n, \ldots, S_1 = s_1] = 1/2(s_n+1) + 1/2(s_n - 1) = s_n$

$S_n = X_1 + X_2 + \cdots + X_n$. $\quad \{S_n^2 - n\}_n$ is a martingale.

- $E[|S_n^2 - n|] \leq E[S_n^2] + n = \sum_{i=1}^{n} E[X_i^2] + \sum_{i \neq j} E[|X_i X_j|] + n$

$$\leq n + (n^2 - n) + n = n^2 + n < +\infty$$

- $E[S_{n+1}^2 - (n+1) \mid S_n^2 - n = y_n,\ S_{n-1}^2 - (n-1) = y_{n-1},\ \ldots,\ S_1^2 - 1 = y_1]$

$$S_{n+1} = \begin{cases} S_n + 1, & \text{w.p. } 1/2 \\ S_n - 1, & \text{w.p. } 1/2. \end{cases}$$

$$= \frac{1}{2}\left[(S_n+1)^2 + (S_n-1)^2\right] - (n+1) = S_n^2 - n = y_n.$$

Stopping rule

- For an experiment with a set of i.i.d. random variables $X_1, X_2, \ldots$, a stopping rule for $\{X_i\}_{i\geq 1}$ is a positive integer-valued random variable N (stopping time) such that for each $n > 1$, the event $\{N \leq n\}$ is independent of $X_{n+1}, X_{n+2}, \ldots$.
- Stopping time says that whether to stop at n depends only on $X_1, \ldots, X_n$, but not look ahead.
- Wald's equality: Let N be a stopping rule for i.i.d. random variables $X_1, X_2, \ldots$ and let $S_N = X_1 + \cdots + X_N$. Then $\mathbb{E}[S_N] = \mathbb{E}[X]\mathbb{E}[N]$.
- A martingale stopped at a stopping time is also a martingale.

- Let N be a stopping time.
- $S_N = X_1 + X_2 + \cdots + X_N$
- $E[S_N] = E[X]\,E[N]$.

Proof:

$$S_N = \sum_{n=1}^{\infty} X_n I_n, \qquad I_n = \begin{cases} 1, & \text{if } N \geq n \\ 0, & \text{if } N \leq n-1. \end{cases}$$

$$E[S_N] = E\left[\sum_{n=1}^{\infty} X_n I_n\right] = \sum_{n=1}^{\infty} E[X_n I_n]$$

I_n is independent of X_n ac. to stopping time.

$$= \sum_{n=1}^{\infty} E[X_n]\,E[I_n]$$

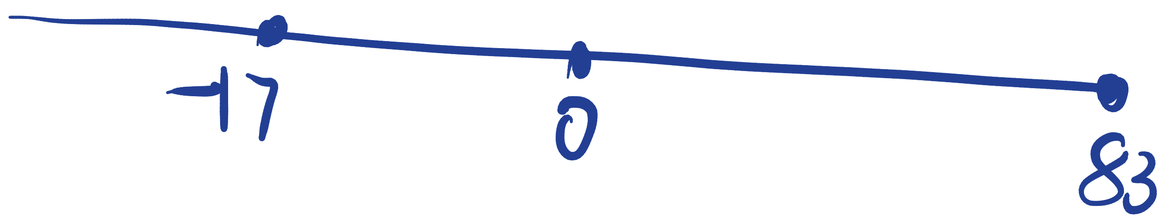
$$= E[X]\left(\sum_{n=1}^{\infty} E[I_n]\right) = E[X] \underbrace{\sum_{n=1}^{\infty} \Pr\{N \geq n\}}_{E[N]}$$

$$= E[X] \cdot E[N]$$

Example

$S_n = X_1 + X_2 + \cdots + X_n$.

$\{S_n\}$ is a martingale.



N: stopping time where he go to endpoints

A drunk man is at the 17th meter of a 100-meter-long bridge. He has a 50% probability of staggering forward or backward one meter each step.

- What is the probability that he will make it to the end of the bridge (the 100th meter) before the beginning (the 0th meter)?
- What is the expected number of steps he takes to reach either the beginning or the end of the bridge?

$\{S_{N}\}_N$ is also martingale.

P_1: prob. that it stops at 83 before -17.

$$E[S_N] = 0$$
$$= P_1 \times 83 + (1-P_1) \times (-17)$$

$$83P_1 = 17 - 17P_1 \Rightarrow P_1 = \frac{17}{100} = 0.17$$

Example

• $\{S_n\}$ is martingale.

• $\{S_n^2 - n\}$ is martingale $\Rightarrow$ $\{S_N^2 - N\}$ is martingale.

A drunk man is at the 17th meter of a 100-meter-long bridge. He has a 50% probability of staggering forward or backward one meter each step.

- What is the probability that he will make it to the end of the bridge (the 100th meter) before the beginning (the 0th meter)?
- What is the expected number of steps he takes to reach either the beginning or the end of the bridge?

$$E[S_N^2 - N] = 0 \Rightarrow E[N] = E[S_N^2]$$

$$= P_1 \times 83^2 + (1-P_1) \times 17^2$$

$$= 1411$$

Example

Suppose you roll a dice. For each roll, you are paid the face value. If a roll gives $4, 5$, or 6, you can roll the dice again. If you get $1, 2$, or 3, the game stops. What is the expected payoff of this game?

X_i: value of face at i-th roll

$S_n = X_1 + \cdots + X_n$

N: a stopping time when we obtain 1, 2, 3

$E[S_N] = E[X]\,E[N] = \frac{7}{2} \times E[N] = \frac{7}{2} \times 2 = 7$