UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Discussion 08
Spring 2023

1. **Moving Books Around**

   You have N books labelled $1, \ldots, N$ on your shelf. At each time step, you pick a book i with probability $\frac{1}{N}$, place it on the left of all others on the shelf, then repeat this process, each step independent of any other step. Construct a suitable Markov chain which takes values in the set of all $N!$ permutations of the books.

   a. Find the transition probabilities of the Markov chain.

   b. Find its stationary distribution.

      Hint: You can guess the stationary distribution before computing it.

   Solution:

   a. The state space consists of all $N!$ permutations on N books. The transition probabilities are then

   $$P((\sigma_1, \ldots, \sigma_{i-1}, \sigma_i, \sigma_{i+1}, \ldots, \sigma_N), (\sigma_i, \sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \sigma_N)) = \frac{1}{N}$$

   for $i = 1, \ldots, N$, and 0 otherwise.

   b. By symmetry, every state $\sigma \in S_N$ should have the same stationary probability,

   $$\pi(\sigma) = \frac{1}{N!}.$$

   We can verify that this probability distribution satisfies the balance equations. Let $\sigma^{(1)} = (\sigma_1, \ldots, \sigma_{i-1}, \sigma_i, \ldots, \sigma_N)$ be a permutation, and for $i = 2, \ldots, n$, let $\sigma^{(i)}$ be the permutation with σ_1 in the ith position, $(\sigma_2, \ldots, \sigma_{i-1}, \sigma_1, \sigma_i, \ldots, \sigma_N)$. With this notation,

   $$\pi(\sigma^{(1)}) = \sum_{i=1}^{N} \pi(\sigma^{(i)}) P(\sigma^{(i)}, \sigma^{(1)}) = \sum_{i=1}^{N} \frac{1}{N!} \cdot \frac{1}{N} = \frac{1}{N!}.$$

2. **Markov Chain Practice**

Consider a Markov chain with three states 0, 1, 2, and suppose its transition probabilities are $P(0,1) = P(0,2) = \frac{1}{2}$, $P(1,0) = P(1,1) = \frac{1}{2}$, $P(2,0) = \frac{2}{3}$, and $P(2,2) = \frac{1}{3}$.

a. Classify the states in the chain. Is this chain periodic or aperiodic?

b. In the long run, what fraction of time does the chain spend in state 1?

c. Suppose that X_0 is chosen according to the steady-state or stationary distribution. What is $\mathbb{P}(X_0 = 0 \mid X_2 = 2)$?

Solution:

a. The Markov chain is one recurrent, aperiodic class.

b. By solving $\pi P = \pi$, we have

$$\pi = \frac{1}{11}\begin{bmatrix} 4 & 4 & 3 \end{bmatrix}.$$

Thus $\pi(1) = 4/11$.

c. By the definition of conditional probability,

$$\mathbb{P}(X_0 = 0 \mid X_2 = 2) = \frac{\mathbb{P}(X_0 = 0, X_2 = 2)}{\mathbb{P}(X_2 = 2)} = \frac{\mathbb{P}(X_0 = 0, X_1 = 2, X_2 = 2)}{\mathbb{P}(X_2 = 2)}.$$

Note that we used the fact that the only possible two-step path from $X_0 = 0$ to $X_2 = 2$ in this chain is $0 \to 2 \to 2$. Now, $\mathbb{P}(X_2 = 2) = \mathbb{P}(X_0 = 2)$ because X_0 is chosen according to the stationary distribution π, so

$$\frac{\mathbb{P}(X_0 = 0, X_1 = 2, X_2 = 2)}{\mathbb{P}(X_2 = 2)} = \frac{\pi(0) \cdot (1/2) \cdot (1/3)}{\pi(2)} = \frac{2}{9}.$$

3. **Hitting Time with Coins**

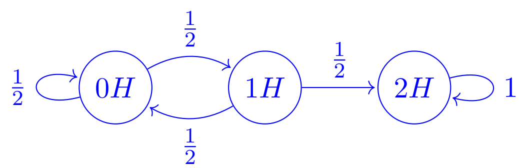
Consider a sequence of fair coin flips.

a. What is the expected number of flips until we first see two heads in a row?

b. What is the expected number of flips until we see a head followed immediately by a tail?

Solution:

a. We can create a Markov chain to compute the expected hitting time. $2H$ represents all sequences with HH as a subsequence, $1H$ all sequences that end in H but do not contain HH, and $0H$ all other sequences, including the initial empty sequence.

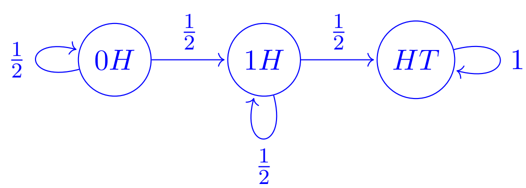


From here, we can set up our hitting-time equations, letting $\beta(i)$ denote the expected number of flips until two consecutive heads, given that we are in state i right now:

$$\begin{aligned}
\beta(0H) &= 1 + \mathbb{P}(H) \cdot \beta(1H) + \mathbb{P}(T) \cdot \beta(0H) \\
&= 1 + \tfrac{1}{2}\beta(1H) + \tfrac{1}{2}\beta(0H) \\
\beta(1H) &= 1 + \mathbb{P}(H) \cdot \beta(2H) + \mathbb{P}(T) \cdot \beta(0H) \\
&= 1 + \tfrac{1}{2}\beta(2H) + \tfrac{1}{2}\beta(0H) \\
\beta(2H) &= 0.
\end{aligned}$$

Solving this system of equations gives us $\beta(1H) = 4$ and $\beta(0H) = 6$. Thus, it takes 6 flips on average until we first see two heads in a row.

b. This part has a slightly different setup: if we flips heads after we just flipped a head, we do not need to reset to the initial state.



Letting $\beta(i)$ be the expected number of flips until we see HT, we have the equations

$$\begin{aligned}
\beta(0H) &= 1 + \mathbb{P}(H) \cdot \beta(1H) + \mathbb{P}(T) \cdot \beta(0H) \\
&= 1 + \tfrac{1}{2}\beta(1H) + \tfrac{1}{2}\beta(0H) \\
\beta(1H) &= 1 + \mathbb{P}(H) \cdot \beta(HT) + \mathbb{P}(T) \cdot \beta(1H) \\
&= 1 + \tfrac{1}{2}\beta(HT) + \tfrac{1}{2}\beta(1H) \\
\beta(HT) &= 0.
\end{aligned}$$

Solving this system gives $\beta(1H) = 2$ and $\beta(0H) = 4$.