Week #6: Markov Processes

Stochastic Methods

Notes by: Francisco Richter

May 15, 2025

Definition. Stochastic Process

A *stochastic process* is a collection of random variables $\{X_n\}_{n\geq 0}$ defined on a common probability space, where the index n represents time.

Definition. Markov Chain

A stochastic process $\{X_n\}_{n\geq 0}$ is called a *Markov chain* if for all $n \geq 0$ and for all states $s_0, s_1, \ldots, s_{n+1}$ we have

$$P(X_{n+1} = s_{n+1} \mid X_n = s_n, X_{n-1} = s_{n-1}, \ldots, X_0 = s_0) = P(X_{n+1} = s_{n+1} \mid X_n = s_n).$$

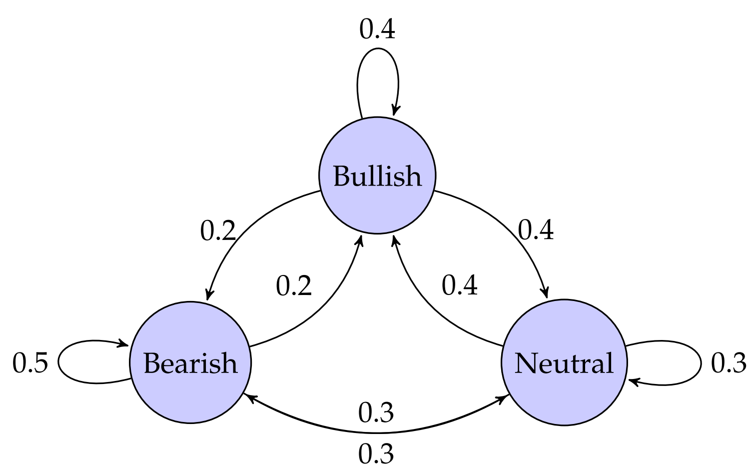

Figure 1: Market State Transition Diagram: Circles represent the states of the system, and the arrows symbolize the potential transitions between these states. The numbers adjacent to the arrows indicate the respective probabilities for these transitions. Given two states i and j, the probability of transition from state i to state j is equal to p_{ij}.

Let T denote the waiting time (in number of steps) until a transition occurs, that is, the first time the chain leaves its current state. The key consequence of the Markov property is that, if no transition occurs in the first n steps, the process effectively *restarts* at time n. Hence, the additional waiting time is independent of the past. Mathematically, we can show this as follows.

Assume that no transition has occurred by time n; then the probability of waiting at least an additional m steps is

$$P(T > n + m \mid T > n) = \frac{P(T > n + m)}{P(T > n)}.$$

Since the process "restarts" after n steps, the probability of waiting an additional m steps is the same as the probability of waiting m steps from the start. That is,

$$\frac{P(T > n+m)}{P(T > n)} = P(T > m).$$

Multiplying both sides by $P(T > n)$ yields

$$P(T > n+m) = P(T > n)\, P(T > m), \quad \forall\, n, m \in \mathbb{N}.$$

Define

$$G(k) = P(T > k), \quad \text{with } G(0) = 1.$$

Then the above equation becomes the functional equation

$$G(n+m) = G(n)G(m), \quad \forall\, n, m \in \mathbb{N}.$$

It is well known that the only solution to this equation is of the form

$$G(k) = q^k, \quad \text{with } 0 < q < 1.$$

Therefore, the probability mass function for T is given by

$$P(T = k) = G(k-1) - G(k) = q^{k-1} - q^k = q^{k-1}(1-q), \quad k \geq 1.$$

This derivation shows that the only discrete distribution satisfying the memoryless property is the *geometric distribution*.

For a Markov chain with a finite state space of size N, the transition probabilities are compactly represented by the **transition matrix** P, an $N \times N$ matrix:

$$P = [p_{ij}], \quad \text{with} \quad p_{ij} = P(X_{t+1} = j \mid X_t = i).$$

Each row of P sums to 1:

$$\sum_{j=1}^{N} p_{ij} = 1 \quad \text{for all } i.$$

Example 0.1 (Financial Market Dynamics)**.** *Financial markets can be modeled as stochastic systems that evolve over time. In this example, we assume that the market can be in one of three distinct states:* ***Bullish***, ***Bearish***, *or* ***Neutral***.

Figure 1 shows a state diagram where:

- *Each circle represents one of the three market states.*
- *The directed arrows indicate the possible transitions between states.*
- *The numbers on the arrows represent the probability of moving from one state to another in a single time step. For instance, an arrow labeled 0.2 from* ***Bullish*** *to* ***Bearish*** *means that if the market is currently bullish, there is a 20% chance it will be bearish in the next period.*

These one-step transition probabilities are summarized in the transition matrix P, *which for a system with three states is a* 3×3 *matrix where the entry* p_{ij} *represents the probability of transitioning from state i to state j. In our example, the transition matrix is:*

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

Here:

- *The first row corresponds to the **Bullish** state: there is a 40% chance that the market remains bullish, a 20% chance it transitions to bearish, and a 40% chance it transitions to neutral.*
- *The second row corresponds to the **Bearish** state.*
- *The third row corresponds to the **Neutral** state.*

This transition matrix succinctly encapsulates the dynamics of the market, and it serves as the basis for further analysis of the system's long-term behavior, which will be discussed in later sections.

Example 0.2 (Random Walk)**.** *A **Random walk** is a classic example of a stochastic process and is often used to describe systems or sequences of events where the next state depends only on the current state and some random element. It can be visualized as a path taken by a particle that moves in random directions.*

The simplest random walk is the one-dimensional walk. At each step, the walker takes a step either to the right (+1) or to the left (-1) with equal probability.

1. *Start at position 0.*
2. *At each time step, flip a coin:*
   - *If heads, move +1 step to the right.*
   - *If tails, move -1 step to the left.*
3. *Record the position after each step.*
4. *Repeat for a desired number of steps.*

Figure 2: 1D Random Walk over 1000 Steps

The particle starts at position 0. At each step, it moves either one step to the right (+1) or one step to the left (-1) with equal probability. The path appears quite random, and the particle can drift far from the starting position but can also return close to its starting position at various times.

A 2D random walk can be visualized on a grid or plane. At each step, the walker takes a step either up, down, left, or right with equal probability.

1. *Start at position* $(0, 0)$.

2. *At each time step, randomly choose one of the four directions:*
   - *Up:* $(+0, +1)$
   - *Down:* $(+0, -1)$
   - *Left:* $(-1, +0)$
   - *Right:* $(+1, +0)$
3. *Record the position after each step.*
4. *Repeat for a desired number of steps.*

Figure 3: 2D Random Walk over 1000 Steps in Greys

In the 2D random walk, the particle moves in a plane, starting from the origin. It takes steps in one of four possible directions: up, down, left, or right. The resulting path is a series of connected line segments in the plane, illustrating the random journey of the particle over time. The path is visualized using the 'Greys' color map, where the starting point is black and the ending point is white, representing the time progression.

Example 0.3 (Infection Dynamics)**.** *Consider a population that can be in one of three states: Susceptible (S), Infected (I), and Recovered (R). Individuals transition between these states according to the following Markov chain.*

The transition probabilities are defined as: β*, the probability of a Susceptible individual becoming Infected;* γ*, the probability of an Infected individual recovering and moving to the Recovered state;* σ*, the probability of an Infected individual transitioning back to the Susceptible state without recovering.*

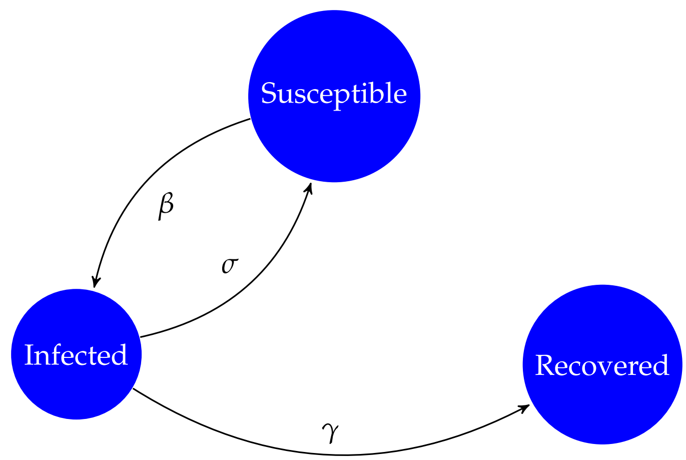

Figure 4: Example of a Markov Chain with three states and transition probabilities.

Given these probabilities, the transition matrix P for this Markov chain can be constructed as:

$$P = \begin{bmatrix} 1-\beta & \beta & 0 \\ \sigma & 1-\sigma-\gamma & \gamma \\ 0 & 0 & 1 \end{bmatrix}$$

Where the rows represent the current state and the columns represent the next state.

This Markov chain captures the essential dynamics of many infectious diseases. The future state of each individual depends only on their current state, satisfying the Markov property. For example, if a person is currently Infected, the probability that they will be Recovered in the next time step is γ, irrespective of their past states.

In a democratic setting, conversations can be modeled as a Markov chain, where each entry p_{ij} in the transition matrix P represents the probability that person i will speak immediately after person j. We now define three key structural properties—irreducibility, aperiodicity, and positive recurrence—and illustrate each with examples drawn from models of democratic discussion.

Definition. Irreducibility

A Markov chain is *irreducible* if, for every pair of states i and j, there exists an integer $n \geq 1$ such that

$$P_{ij}^{(n)} > 0.$$

In a conversational model, this means that starting from any speaker, there is a positive probability that any other person will eventually have a chance to speak.

Example 0.4 (Disconnected Factions in a Democracy)**.** *Imagine a system where people must reach consensus on various topics through dialogue. Suppose that due to ideological divides, the participants split into two factions that never interact. Label the speakers so that persons* $1, 2, \ldots, k$ *belong to faction A and persons* $k+1, k+2, \ldots, N$ *to faction B. The transition matrix may then take the block-diagonal form*

$$P = \begin{pmatrix} P_A & 0 \\ 0 & P_B \end{pmatrix},$$

where P_A *and* P_B *are the transition matrices within each faction. Since there is zero probability of transitioning from a speaker in one faction to another, the chain is not irreducible.*

Definition. Aperiodicity

A state i is *aperiodic* if

$$\gcd\{n \geq 1 : P_{ii}^{(n)} > 0\} = 1.$$

If every state is aperiodic, then the entire chain is said to be aperiodic. In our conversation model, aperiodicity implies that there is no fixed cyclic pattern dictating who speaks next.

Example 0.5 (Periodic Power Exchange Between Two Parties)**.** *Consider a simplified model of democratic conversation between two individuals, where turns alternate strictly. The transition matrix is:*

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

If person 1 speaks at time t, they can only speak again at even time steps $t+2, t+4, \ldots$. Since the possible return times are all multiples of 2, the greatest common divisor is 2. Thus, the system is periodic rather than aperiodic.

Definition. Positive Recurrence

A state i is *positive recurrent* if the expected return time to i, starting from i, is finite:

$$\mathbb{E}_i[T_i] = \sum_{n=1}^{\infty} n \cdot \mathbb{P}(T_i = n) < \infty,$$

where T_i is the first return time to state i. A Markov chain is positive recurrent if all states are positive recurrent. This is a key condition for the existence of a stationary distribution.

Example 0.6 (Null-Recurrent Opinion Dynamics)**.** *Consider a model of opinion dynamics where speaker changes follow a symmetric random walk on the integers $\mathbb{Z}$. The transition matrix is infinite and given by*

$$P = \begin{pmatrix} \ddots & \ddots & \ddots & & & \\ & 0.5 & 0 & 0.5 & & \\ & & 0.5 & 0 & 0.5 & \\ & & & \ddots & \ddots & \ddots \end{pmatrix},$$

with $p_{i,i\pm1} = 0.5$ for all $i \in \mathbb{Z}$. This chain is recurrent (returns to every state with probability 1) but not positive recurrent: the expected return time to any state is infinite. Therefore, it is null recurrent *and lacks a stationary distribution.*

Example 0.7 (Hierarchical Influence Game)**.** *Now consider a hierarchical conversational model where a central leader exerts strong influence. Regardless of who just spoke, the probability that the next speaker is:*

- *the leader is 0.5,*
- *the first subordinate is 0.25,*
- *the second subordinate is 0.125,*
- *and so on.*

This gives rise to the transition matrix:

$$P = \begin{pmatrix} 0.5 & 0.25 & 0.125 & 0.0625 & \cdots \\ 0.5 & 0.25 & 0.125 & 0.0625 & \cdots \\ 0.5 & 0.25 & 0.125 & 0.0625 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Here:

- *The chain is* irreducible *because every person can be reached from any other with positive probability.*
- *It is* aperiodic *since there is a nonzero probability (0.5) of remaining in the same state (the same speaker speaking again).*
- *It is* positive recurrent *since each state has finite expected return time, and a stationary distribution exists (identical to the common row of P).*

Thus, the hierarchical influence game is an example of an ergodic *Markov chain.*

Definition. Ergodicity

A Markov chain is *ergodic* if it is irreducible, aperiodic, and positive recurrent. In this case, the chain has a unique stationary distribution π satisfying:

$$\pi P = \pi \quad \text{and} \quad \sum_j \pi_j = 1,$$

and the long-run behavior is given by:

$$\lim_{n\to\infty} P_{ij}^{(n)} = \pi_j \quad \text{for all states } i.$$

Example 0.8 (Financial Market Dynamics: Stationary Distribution via Eigenvalue Analysis)**.** *Consider the financial market model with three states:* ***Bullish****,* ***Bearish****, and* ***Neutral****. The one-step transition matrix is*

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

We wish to find the stationary distribution $\pi = (\pi_1, \pi_2, \pi_3)$ such that

$$\pi P = \pi \quad \textit{and} \quad \pi_1 + \pi_2 + \pi_3 = 1.$$

Writing the stationary condition in components, we have

$$\begin{cases} \pi_1 = 0.4\pi_1 + 0.2\pi_2 + 0.4\pi_3, \\ \pi_2 = 0.2\pi_1 + 0.5\pi_2 + 0.3\pi_3, \\ \pi_3 = 0.4\pi_1 + 0.3\pi_2 + 0.3\pi_3. \end{cases}$$

Subtract π_1, π_2, and π_3 from the respective equations to obtain:

$$0 = -0.6\pi_1 + 0.2\pi_2 + 0.4\pi_3, \quad (1)$$

$$0 = 0.2\pi_1 - 0.5\pi_2 + 0.3\pi_3, \quad (2)$$

$$0 = 0.4\pi_1 + 0.3\pi_2 - 0.7\pi_3. \quad (3)$$

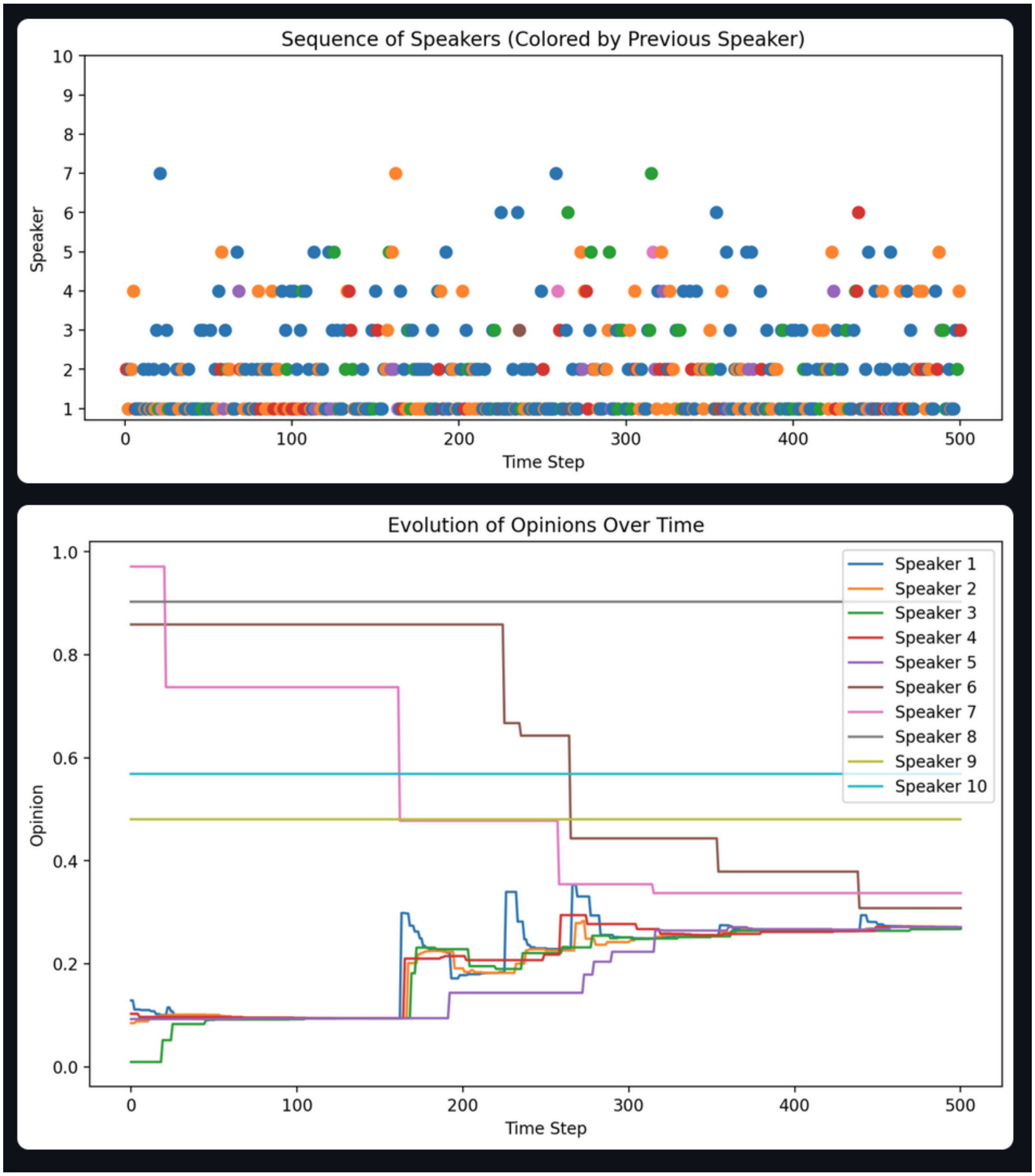

Figure 5: A Markov chain representation of speaker transitions (top) and evolving opinions (bottom). The top layer encodes who speaks next, while the bottom layer tracks how opinions shift based on influence and structure.

It is common to work with only two of these equations along with the normalization condition.

From equation (1):

$$0.6\pi_1 = 0.2\pi_2 + 0.4\pi_3 \quad \Longrightarrow \quad \pi_1 = \frac{1}{3}\pi_2 + \frac{2}{3}\pi_3. \quad (4)$$

From equation (2):

$$0.2\pi_1 = 0.5\pi_2 - 0.3\pi_3 \quad \Longrightarrow \quad \pi_1 = 2.5\pi_2 - 1.5\pi_3. \quad (5)$$

Equate (4) and (5):

$$\frac{1}{3}\pi_2 + \frac{2}{3}\pi_3 = 2.5\pi_2 - 1.5\pi_3.$$

Multiply both sides by 3:

$$\pi_2 + 2\pi_3 = 7.5\pi_2 - 4.5\pi_3.$$

Rearrange terms:

$$0 = 6.5\pi_2 - 6.5\pi_3 \quad \Longrightarrow \quad \pi_2 = \pi_3.$$

Substitute $\pi_2 = \pi_3$ in (4):

$$\pi_1 = \frac{1}{3}\pi_2 + \frac{2}{3}\pi_2 = \pi_2.$$

Thus, $\pi_1 = \pi_2 = \pi_3$. Using the normalization condition:

$$3\pi_1 = 1 \quad \Longrightarrow \quad \pi_1 = \pi_2 = \pi_3 = \frac{1}{3}.$$

Therefore, the stationary distribution is

$$\pi = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right).$$

In the context of Markov chains, the terms "invariant distribution" and "stationary distribution" are usually used interchangeably. Let $P = [p_{ij}]$ be the transition matrix of a Markov chain with state space S. A probability distribution $\pi = (\pi_i)_{i \in S}$ is called an *invariant distribution* if

$$\pi P = \pi,$$

that is, for every $j \in S$,

$$\pi_j = \sum_{i \in S} \pi_i \, p_{ij},$$

and

$$\sum_{i \in S} \pi_i = 1.$$

1 Exercises

1. **Telephone System Dynamics.**
   A telephone exchange operates in discrete time with two possible states: **Busy** and **Idle**. Assume that if the system is **Busy** at time n, then at time $n+1$ it remains **Busy** with probability $1-q$ and becomes **Idle** with probability q. If it is **Idle** at time n, it becomes **Busy** at time $n+1$ with probability p and stays **Idle** with probability $1-p$.

   (a) Write down the 2×2 transition matrix P for this Markov chain.

   (b) Find the invariant probability distribution $\pi = (\pi_{\text{Busy}}, \pi_{\text{Idle}})$.

2. **Doubly Stochastic Chain.**
   Let $S = \{1, 2, \ldots, N\}$ and suppose that the transition matrix $P = [p(j \mid i)]_{i,j \in S}$ of a Markov chain satisfies

   $$\sum_{i=1}^{N} p(j \mid i) = 1 \quad \text{for every } j \in S.$$

   Prove that the uniform distribution

   $$\pi_i = \frac{1}{N}, \quad i = 1, \ldots, N,$$

   is invariant; that is, show that $\pi P = \pi$.

3. **Branching Process Survival.**
   Consider a branching process where each individual produces a random number of offspring according to a distribution $\{p_k\}_{k\geq 0}$ with $p_k > 0$ for every k and $\sum_{k\geq 0} p_k = 1$. Let Z_n denote the number of individuals in generation n. Show that $\{Z_n\}_{n\geq 0}$ is a Markov chain.

4. **Offspring Distributions in Branching Processes.**
   Consider two cases for the offspring distribution:

   (a) Suppose the number of offspring is binomially distributed with parameters $N \in \mathbb{N}^*$ and $p \in (0,1)$:
   $$p_k = \binom{N}{k} p^k (1-p)^{N-k}, \quad k = 0, 1, \ldots, N,$$
   and $p_k = 0$ for $k > N$. Show that if there are i individuals in generation n, then it is impossible to have more than Ni individuals in generation $n+1$.

   (b) Suppose instead that the offspring distribution is Poisson with mean $\lambda > 0$, i.e.,
   $$p_k = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \geq 0.$$
   Explain how, if there are i individuals in generation n, the total number of offspring in generation $n+1$ is distributed.

5. **Dice Game: Maximum Outcome.**
   A fair six-sided die is rolled repeatedly. Define X_n to be the maximum outcome observed in the first n rolls.

   (a) Prove that $\{X_n\}_{n\geq 0}$ is a Markov chain.

   (b) Determine its transition probabilities.

6. **Pyramid Game.**
   A fair four-faced die is rolled repeatedly. Define Y_n as the maximum outcome observed in the first n rolls.

   (a) Identify the state space for $\{Y_n\}$.

   (b) Find the transition probabilities.

7. **Random Walk with Absorbing Barriers.**
   Consider a random walk on the state space
   $$S = \{0, 1, 2, \ldots, a\},$$
   where 0 and a are absorbing states. For $i \in \{1, \ldots, a-1\}$, the walk moves to $i+1$ with probability p and to $i-1$ with probability $q = 1-p$.

   (a) Determine all invariant measures for this Markov chain.

   (b) For a given initial state i, compute the probability that the process is absorbed at state a before state 0.