Stochastic Processes

GPDS
GRUPO DE PROCESSAMENTO DIGITAL DE SINAIS

Prof. F. Assis

Sumary

- The Poisson counting process
- Definition of the Poisson process: average rate of events per unit of time, properties, PDF, CDF.
- Applications of Poisson process.
- First order Markov process.
- A Markov chain of order h.
- Markov process definitions and properties.
- Markov process: the set of states
- Discrete-Time Markov Chains

Sumary

- Markov Chains: transition probabilities
- Markov Chains: transition probability matrix
- The probability distribution of X_0 - $\pi^{(0)}$
- Discrete-Time Markov Chains and the law of total probability
- n-Step transition probabilities
- Discrete-time Markov chains: classes of states
- Markov Chain: Limiting Distributions
- Finite Markov Chain: Stationary Distributions
- Countably infinite Markov chains
- Continuous-time Markov process

Sumary

- Random walk
- Brownian Motion

Poisson Processes

- The Poisson process is an example of a continuous-time, discrete-value random process. A counting process $X(t)$ represents the total number of occurrences of random events in the interval $[0, t]$.

- In a counting process, events occur at random instants of time, such that the **average rate of events** (arrivals) per unit of time (second, for example) is equal to the constant λ.

Poisson Processes

- In a Poisson process we want to know the probability of counting a specific number of events in an observation time window.

Definition of the Poisson process

- The **average rate of events** per unit of time $\lambda > 0$ is a fixed constant.
- The counting process $\{N(t), t \in [0,\infty)\}$ is called a **Poisson process** with **rates** λ if all the following conditions hold:
  - 1. $N(0) = 0$;
  - 2. $N(t)$ has independent increments;
  - 3. The number of arrivals in any interval of length $t > 0$ has *Poisson* (λt) distribution.

Definition of the Poisson process

- Note that from the definition, we conclude that in a Poisson process, the distribution of the number of arrivals (events) in any interval depends only on the length of the interval, and not on the exact location of the interval on the real line.

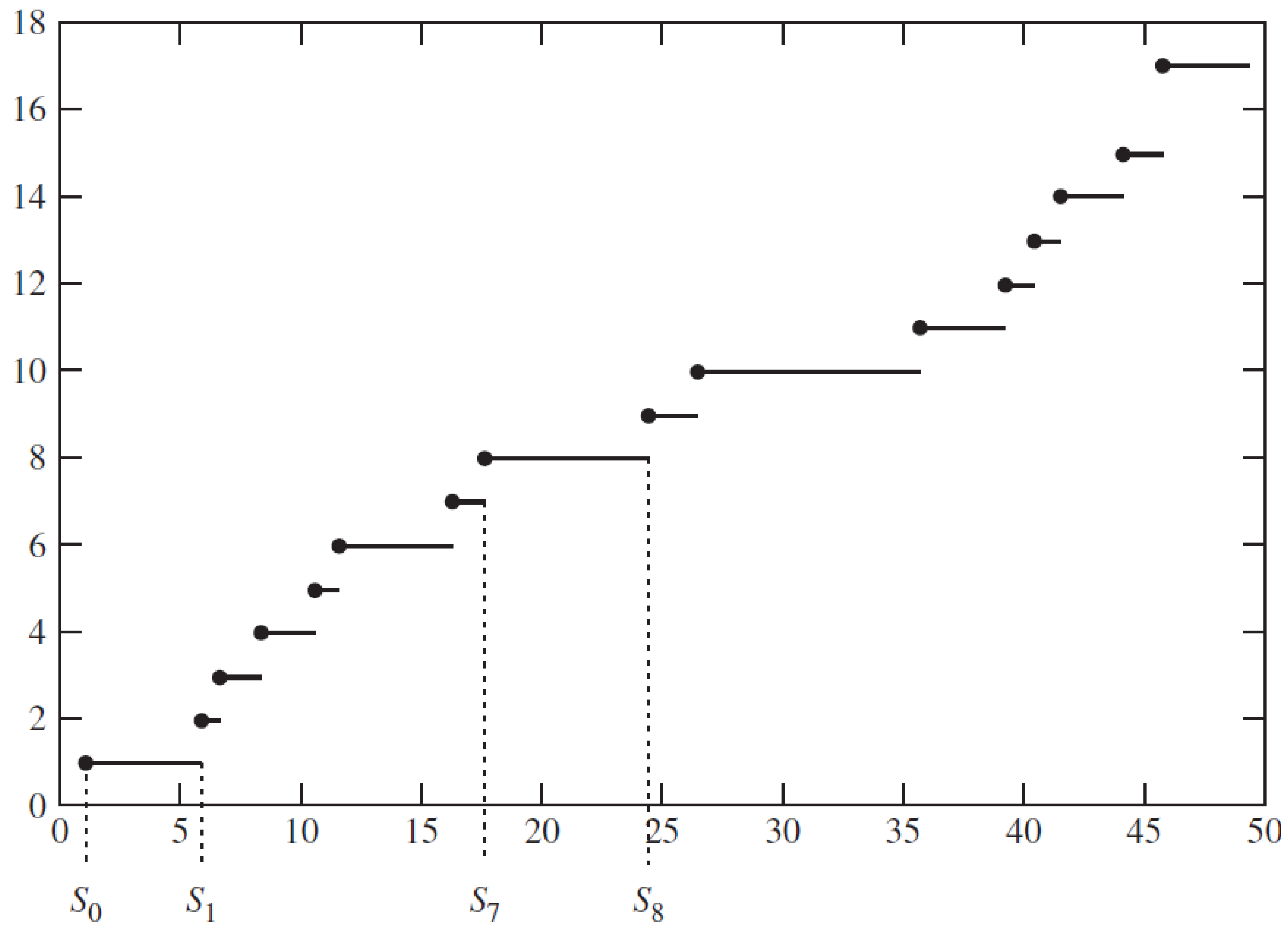
- Therefore, the *Poisson process* has ***stationary increments***.

Poisson Processes

- Note that $X(t)$ is a nondecreasing, integer-valued, continuous-time process, where we assume $X(0) = 0$.

- The event occurrence times are denoted by $S_1, S_2, \ldots$

- The j-th interevent time is denoted by $X_j = S_j - S_{j-1}$.



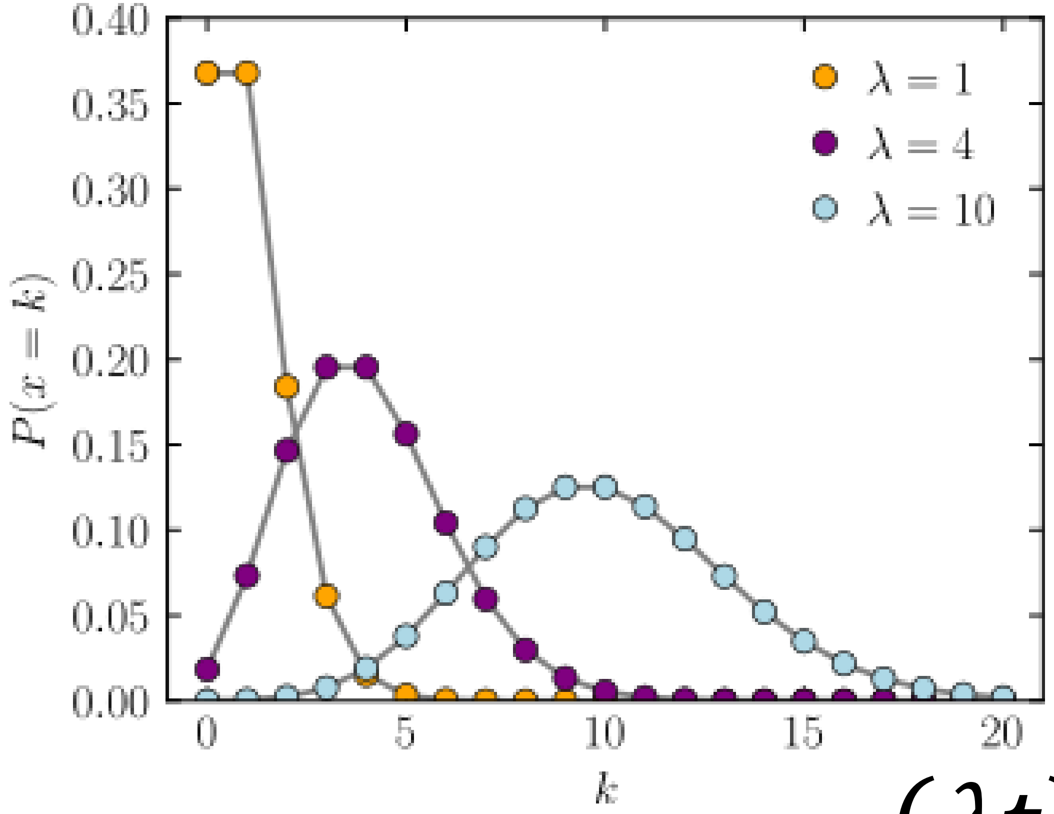
Poisson Processes

- Observe that a random counting process is said to be a ***Poisson process*** with average rate $\lambda > 0$, if it has the stationary independent increments property and the time between events is exponentially distributed, i.e. the number of event occurrences in the interval $[0, t]$ has a Poisson distribution with mean λt, as defined by

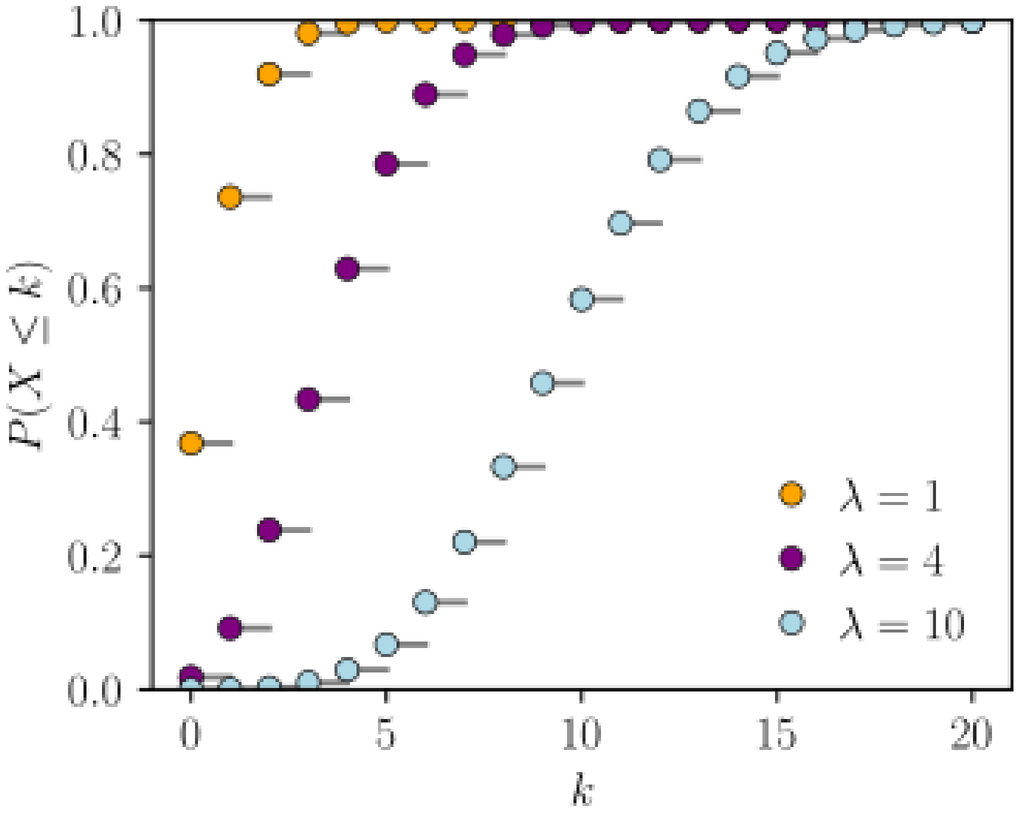
$$P[X(t) = k] = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \qquad k = 0, 1, 2, \ldots$$

Poisson Processes

- Probability Distribution Function



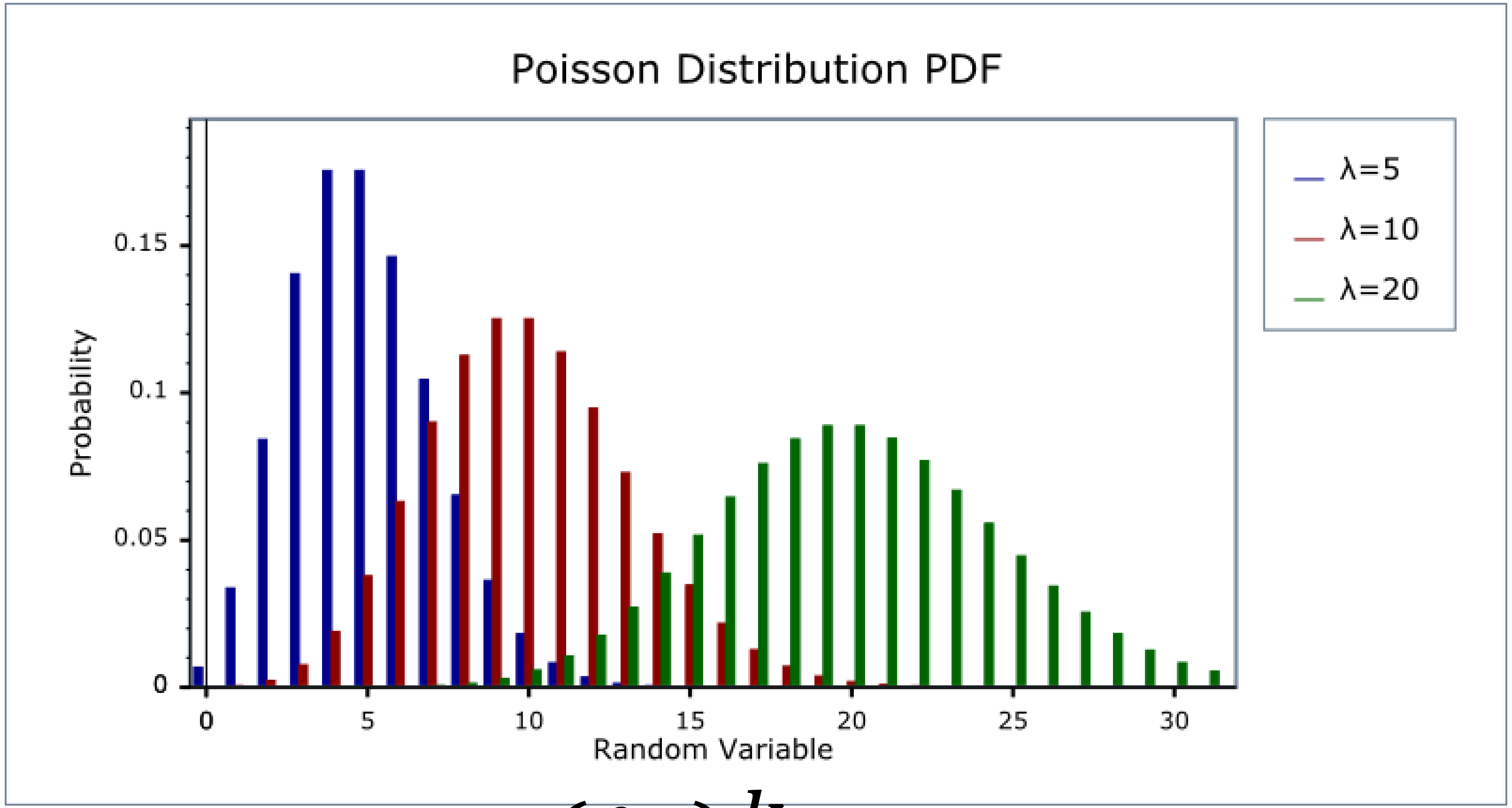
- Cumulative distribution function



$$P[X(t) = k] = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \qquad k = 0, 1, 2, \ldots$$

Poisson Processes

- Discrete evaluation



$$P[X(t) = k] = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \qquad k = 0, 1, 2, \dots$$

Poisson Processes

- The Poisson process is thus a **memoryless counting process** in which an arrival at a particular instant is **independent** of **an arrival at any other instant**.

- For any **pair of nonoverlapping intervals**, the number of arrivals in each interval are **independent random variables**.

- The number of arrivals in any interval $(t_0, t_1]$, that is $X(t_0)-X(t_1)$, is a Poisson random variable with expected value $\lambda\,(t_1 - t_0)$.

Poisson Processes

- The mean, variance, and autocorrelation functions of a Poisson process $X(t)$ with rate λ are as follows:

$$E[X(t)] = \lambda t$$

$$\text{Var}[X(t)] = \lambda t$$

$$R_X(t, s) = \lambda \min(t, s) + \lambda^2 ts$$

- The Poisson process is not Strict-Sense Stationary (SSS), nor it is even Wide-Sense Stationary (WSS).

Exercise 8.1

- Inquiries arrive at a recorded message device according to a Poisson process of rate 15 inquiries per minute.
- Find the probability that in a 1-minute period, 3 inquiries arrive during the first 10 seconds and 2 inquiries arrive during the last 15 seconds.

- As consultas observadas em um dispositivo de gravação de mensagens chegam à taxa de 15 consultas por minuto e, o sistema é modelado de acordo com um processo de Poisson.
- Encontre a probabilidade de que em um período total de 1(um) minuto, 3(três) consultas cheguem durante os primeiros 10 segundos de observação e 2(duas) consultas cheguem durante os últimos 15 segundos.

Exercise 8.1: solution

- The arrival rate in seconds is $\lambda = \frac{15(\text{inquiries})}{60(seconds)} = 1/4$ inquiries per second.
- Writing time in seconds, the probability of interest is

$$P[(X(10) = 3) and (X(60) - X(45) = 2)]$$

$$= P[(X(10) = 3), (X(60) - X(45) = 2)\,]$$

$$= P[X(10) = 3]P[X(60) - X(45) = 2]$$

Exercise 8.1: solution

- By applying first the **independent increments property**, and then the **stationary increments property**, we obtain

$$P[X(t_1) - X(t_0) = k] = P[X(t_1 - t_0) = k]$$

$$P[X(60) - X(45) = 2] = P[X(60 - 45) = 2]$$

$$= P[X(15) = 2]$$

Example 8.1: solution

- For $P[X(10) = 3]$, t = 10 seconds and k = 3 inquiries.
- For $P[X(15) = 2]$, t = 15 seconds and k = 2.

$$P[X(t) = k] = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \qquad k = 0, 1, 2, \dots$$

$$P[(X(10) = 3), (X(60) - X(45) = 2)]$$

$$= P[X(10) = 3] P[X(15) = 2]$$

$$= \left\{ \frac{\left(\frac{1}{4}10\right)^3}{3!} e^{-\frac{1}{4}10} \right\} \left\{ \frac{\left(\frac{1}{4}15\right)^2}{2!} e^{-\frac{1}{4}15} \right\}$$

Exercise 8.2

- The average number that a team kicks the ball in goal for each playing time of 45 minutes is equal to 5(five). The way of playing implemented by the team coach can be modeled by the Poisson process.
- a) Find the probability that this team kicks the ball in goal 4(four) times in the first half hour of a game.
- b) Find the probability that this team kicks the ball in goal 4(four) times between the 60th minute and the 90th minute.

Example 8.2

- c) Find the probability that this team kicks the ball in goal 2(two) times in the last 15 minutes of the first half time (the first 45 minutes) **and** 5(five) kicks between the 50th minute and the 70th minute.

Example 8.2: solution

- The ball kicks in goal rate in minutes is $\lambda = \frac{5(\text{kicks})}{45(minutes)} = \frac{1}{9}\ kicks/minute.$
- a) Find the probability that this team kicks the ball in goal 4(four) times in the first half hour of a game. For t = 30 min and k = 4 kicks we have

$$P[X(t) = k] = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \qquad k = 0, 1, 2, \ldots$$

$$P[X(30) = 4] = \frac{\left(\frac{1}{9} 30\right)^4}{4!} e^{-\frac{1}{9}30}$$

Example 8.2: solution

- b) Find the probability that this team kicks the ball in goal 4(four) times between the 60th minute and the 90th minute.

$$P[X(90) - X(60) = 4] = P[X(90 - 60) = 4]$$
$$= P[X(30) = 4]$$

- For t = 30 min and k = 4 kicks we have the same answer of letter (a).

Example 8.2

- c) Find the probability that this team kicks the ball in goal 2(two) times in the last 15 minutes of the first half time **and** 5(five) kicks between the 50th minute and the 70th minute.

$$P[(X(45) - X(30) = 2) and \ (X(70) - X(50) = 5)]$$

$$= P[\{X(15) = 2\} \ and \ \{X(20) = 5\}]$$

$$= P[X(15) = 2]P[X(20) = 5]$$

$$= \left\{\frac{\left(\frac{1}{9}15\right)^2}{2!} e^{-\frac{1}{9}15}\right\}\left\{\frac{\left(\frac{1}{9}20\right)^5}{5!} e^{-\frac{1}{9}20}\right\}$$

Introduction:Markov process

- A **discrete-valued** random process $X(t)$ is a (**first order**) **Markov process** if the future of the process given the present is independent of the past, that is, if for arbitrary times $t_1 < t_2 < t_3 \ldots t_k < t_{k+1}$

$$P[X(t_{k+1}) = x_{k+1} | X(t_k) = x_k, X(t_{k-1}) = x_{k-1}, \ldots, X(t_1) = x_1]$$

$$= P[X(t_{k+1}) = x_{k+1} | X(t_k) = x_k]$$

Introduction: **high-order** Markov chain

- A **Markov chain of order *h*** is defined as a sequence in which $X(t_{k+1})$ depends on its present and past only through its (*h*-1) previous values, $t_1 < t_2 < t_3 \ldots < t_{k-h-1} < \cdots < t_k < t_{k+1}$

$$P[X(t_{k+1}) = x_{k+1} | X(t_k) = x_k, \ldots, X(t_{k-h-1}) = x_{k-h-1}, \ldots, X(t_1) = x_1]$$

$$= P[X(t_{k+1}) = x_{k+1} | X(t_k) = x_k, \ldots, X(t_{k-h}) = x_{k-h}]$$

Introduction: Markov process

- The case $h = 1$ reduces to a simple Markov chain and $h = 0$ to an independent sequence.

- Any higher-order Markov chain with finite h defined over state space S can be transformed into a simple Markov chain by defining the state space $S^h = S \times S \cdots \times S$, the h-times cartesian product of S with itself.

Introduction: Markov process

- In the Markov processes expression is the "present," is the "future," and is the "past."

- Thus, in Markov processes, PMF's and PDF's that are conditioned on several time instants always reduce to a PMF/PDF that is conditioned only on the most recent time instant.

- For this reason, we refer to the value of $X(t)$ at time t as the **state** of the process at time t.

Markov process: the set of states

Denoting the simple Markov chain $\{X(t_k)\}=\{X_k\}$ defined above is referred to as a **discrete-time Markov chain** (DTMC).

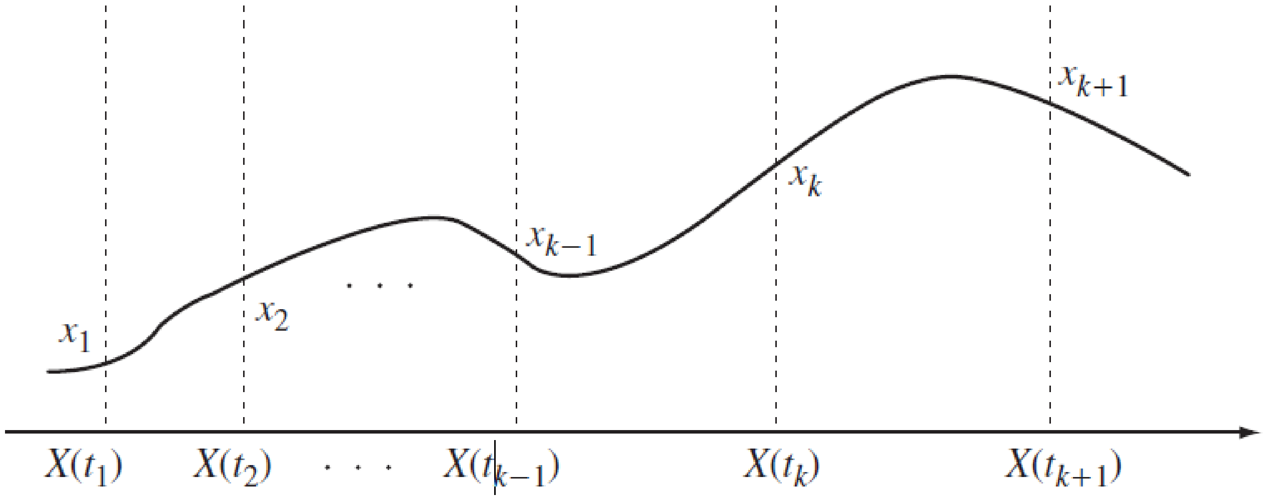
$$X_1, X_2, \ldots, X_{k-1}, X_k, X_{k+1} \ldots$$

If there are *M* **different states** that the Markov chain can take on, we can label them, without loss of generality, by integers,

$$0, 1, 2, \ldots, M-1.$$

We denote this **set of states** by *S*:

$$S = \{0, 1, 2, \ldots, M-1\}$$



Discrete-Time Markov Chains

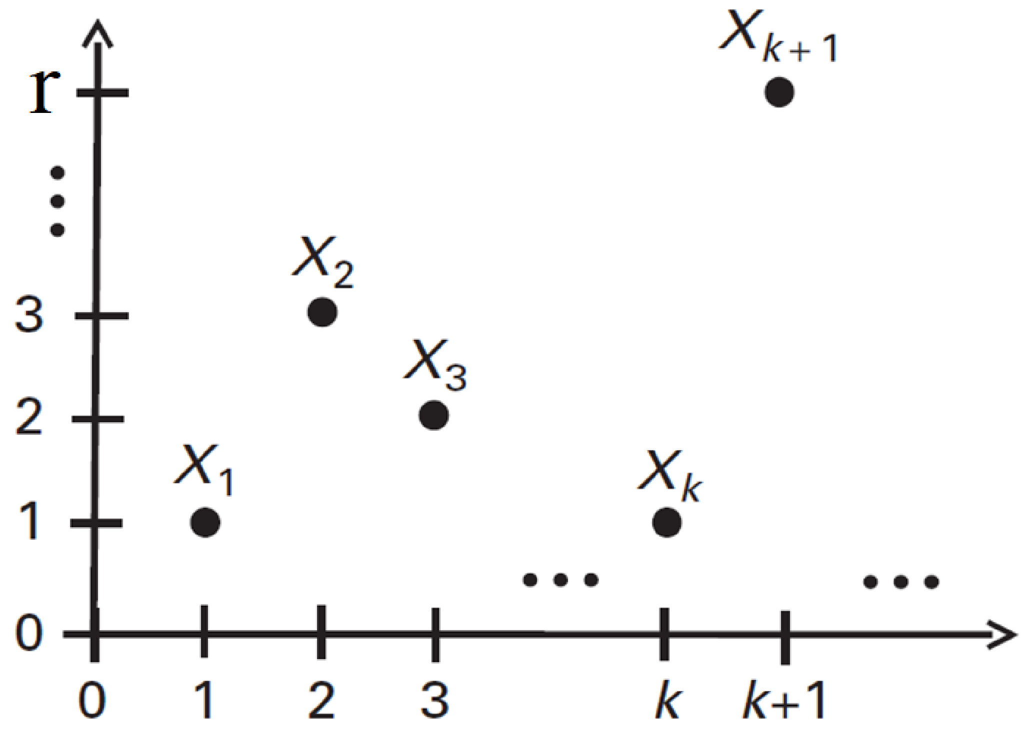
Definition: consider the random process $\{X_n, n = 0, 1, 2, \dots\}$,where $S \subset \{0, 1, 2, \dots\}$. We say that this process is a **Markov chain** if

$$P[X_{k+1} = j | X_k = i, X_{k-1} = i_{m-1}, \dots, X_0 = i_0] = P[X_{k+1} = j \, | X_k = i]$$

For all $m, j, i, i_0, i_1, \cdots i_{m-1}$. If the number of states is finite, e.g.,

$$S = \{0, 1, 2, \cdots, r\},$$

we call it **a finite Markov chain**.



Discrete-Time Markov Chains

- If X_k =j, we say that the process is in state *j*.

- The values $P[X_{k+1} = j \,|X_k = i]$ are called the **transition probabilities**.

- We assume that the transition probabilities do not depend on time (it is time-invariant).

- That means, $P[X_{k+1} = j \,|X_k = i]$ does not depend on *k*.

Discrete-Time Markov Chains

- Thus, we can define the **transition probabilities** as

$$p_{ij} = P[X_{k+1} = j \,| X_k = i]$$

In particular, we have

$$\begin{aligned} p_{ij} &= P[X_1 = j \,| X_0 = i] \\ &= P[X_2 = j \,| X_1 = i] \\ &= P[X_3 = j \,| X_2 = i] \ldots \end{aligned}$$

$$\ldots = P[X_{k+1} = j \,| X_k = i] = \cdots$$

In other words, **if the process is in state *i***, it will **next make a transition to state *j*** with **probability** p_{ij}.

Discrete-Time Markov Chains

- We often list the transition probabilities in a matrix. The matrix is called the **state transition matrix** or **transition probability matrix** and is usually shown by ***P***.
- Assuming the states are 1, 2, ⋯, *r*, then the state transition matrix is given by

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1r} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2r} \\ p_{31} & p_{32} & p_{33} & & p_{3r} \\ & \vdots & & \ddots & \vdots \\ p_{r1} & p_{r2} & p_{r3} & \cdots & p_{rr} \end{bmatrix}$$

Discrete-Time Markov Chains

- Note that $p_{ij} \geq 0$, and for all i, we have

$$\sum_{k=1}^{r} p_{ik} = \sum_{k=1}^{r} P[X_{m+1} = k \mid X_m = i] = 1$$

- This is because, given that we are in state i, the next state must be one of the possible states.
- Thus, when we sum over all the possible values of k, we should get one.
- That is, the rows of any state transition matrix must sum to one.

Exercise 8.3

- Consider a voice communication channel and you have two possible states speech activity or silence.
- A Markov model for packet speech assumes that if the *n*-th packet contains silence, then the probability of silence in the next packet is $(1 - \alpha)$ and the probability of speech activity is α.
- Similarly, if the *n*-th packet contains speech activity, then the probability of speech activity in the next packet is $(1 - \beta)$ and the probability of silence is β.
- Build the transition probability matrix and state diagram.

Exercise 8.3: solution

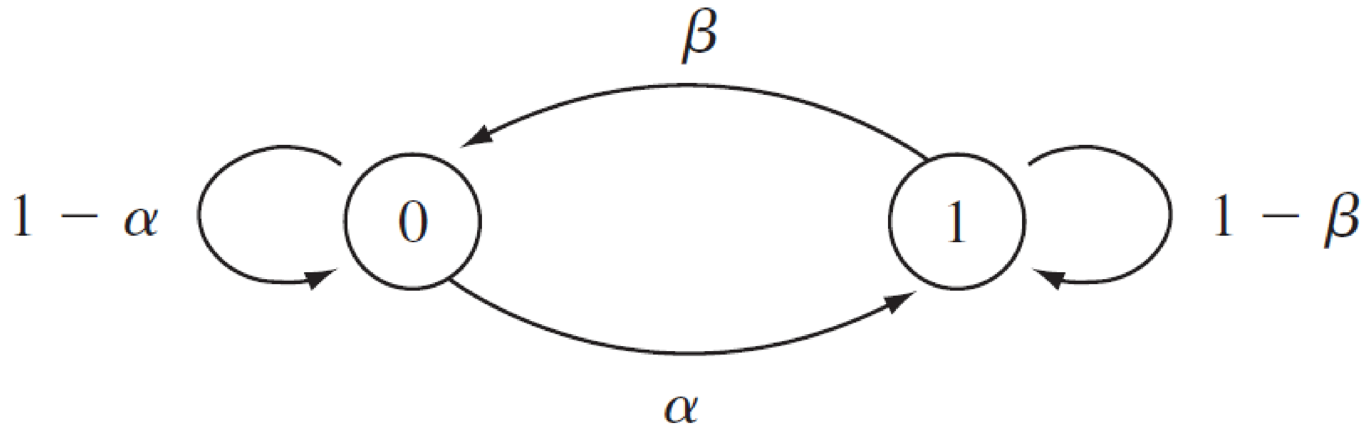
- The transition probability matrix

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}$$

Statistical concentrators

Silence = state 0

Voice = state 1



- State transition diagram for two-state Markov chain

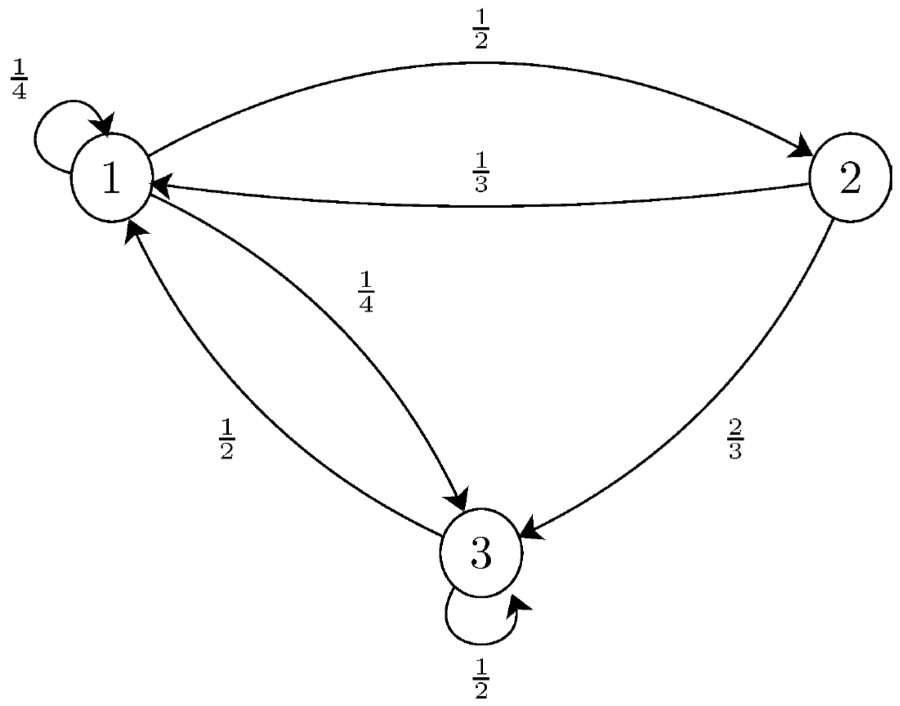
Exercise 8.4

- Consider a Markov chain with three possible states 1, 2, and 3 and the following transition probabilities give below

$$P = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

- The transition probability matrix

- State transition diagram for three-state Markov chain

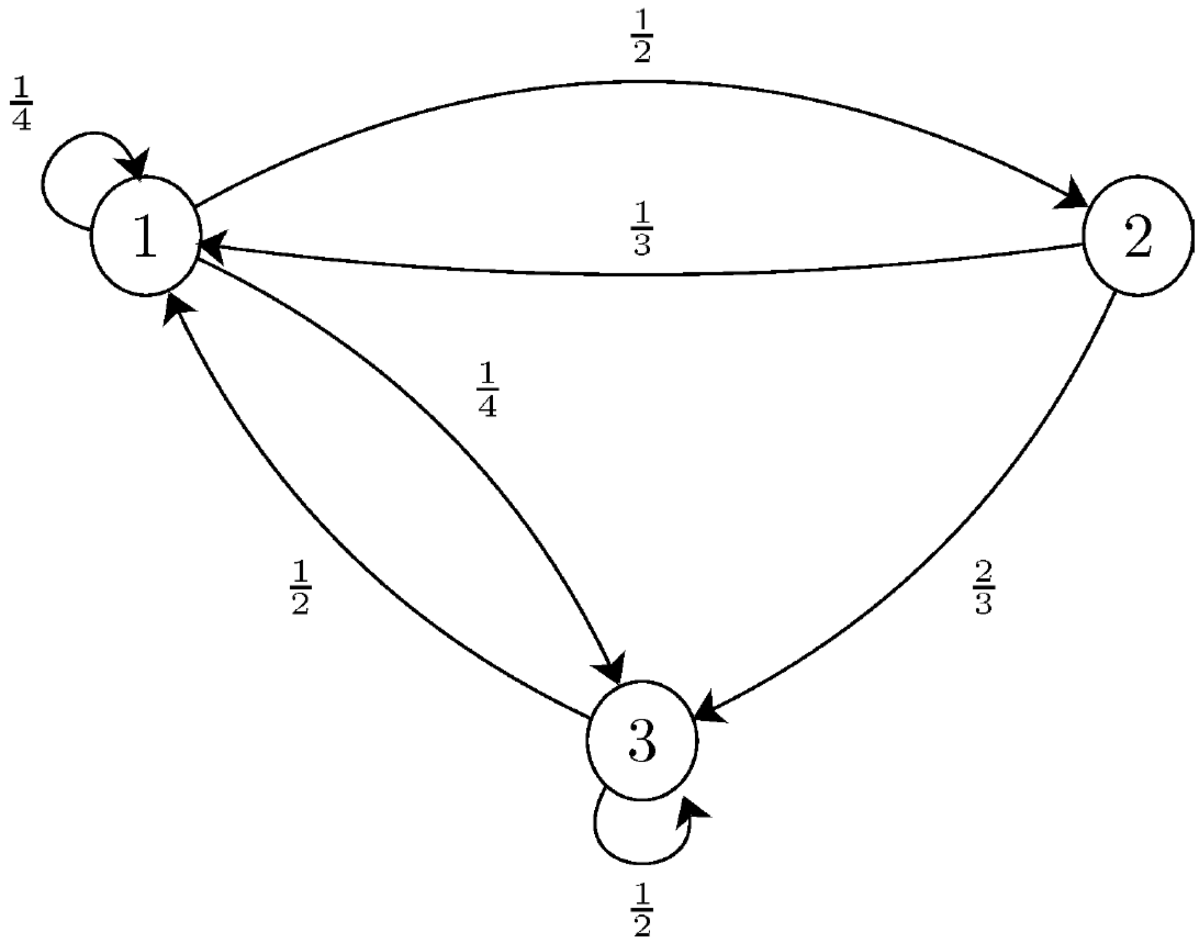


Exercise 8.4

- a) Find $P(X_4 = 3 | X_3 = 2)$.
- b) Find $P(X_3 = 1 | X_2 = 1)$.
- c) If we know $P(X_0 = 1) = 1/3$ (probability of initial state being 1), find $P(X_0 = 1, X_1 = 2)$.
- d) If we know $P(X_0 = 1) = 1/3$, find $P(X_0 = 1, X_1 = 2, X_2 = 3)$.

Exercise 8.4: Solution

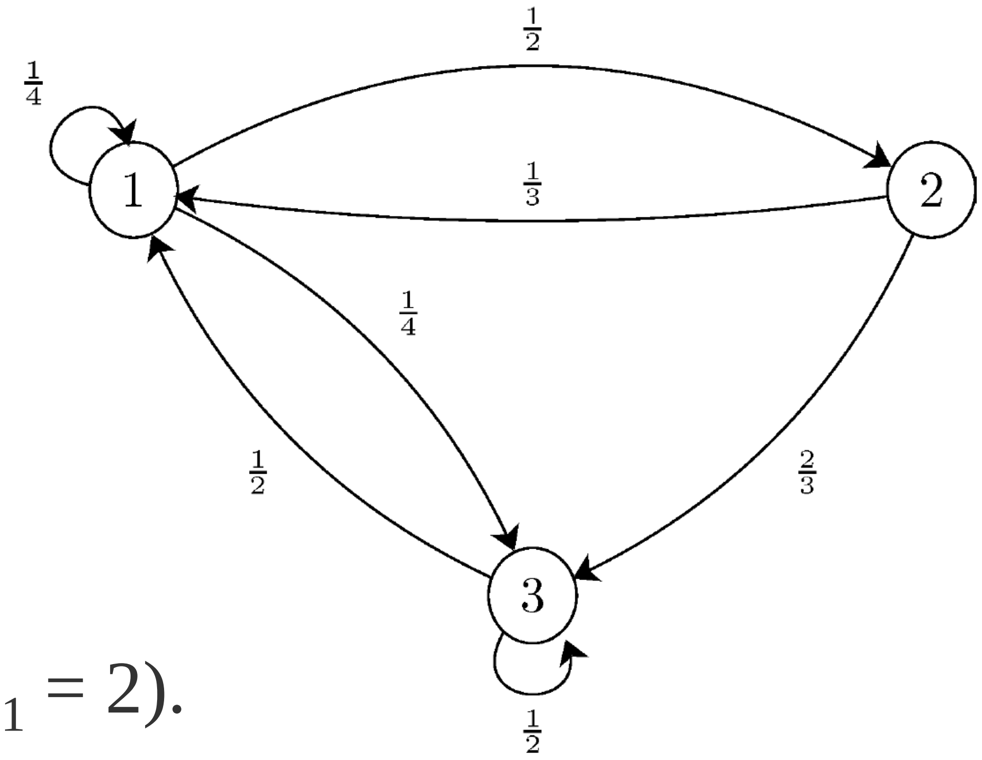
$$P = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$



- a) $P(X_4 = 3 | X_3 = 2) = p_{23} = \frac{2}{3}$
- b) $P(X_3 = 1 | X_2 = 1) = p_{11} = \frac{1}{4}$

Exercise 8.4: Solution

$$P = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$



- c) If we know $P(X_0 = 1) = 1/3$, find $P(X_0 = 1, X_1 = 2)$.
- X_0 corresponds to the Markov process start time. The boundary condition of the problem is: What is the probability of the Markov system to start in state 1 **and** then transit to state 2?
- The conditional probability can be put in the form

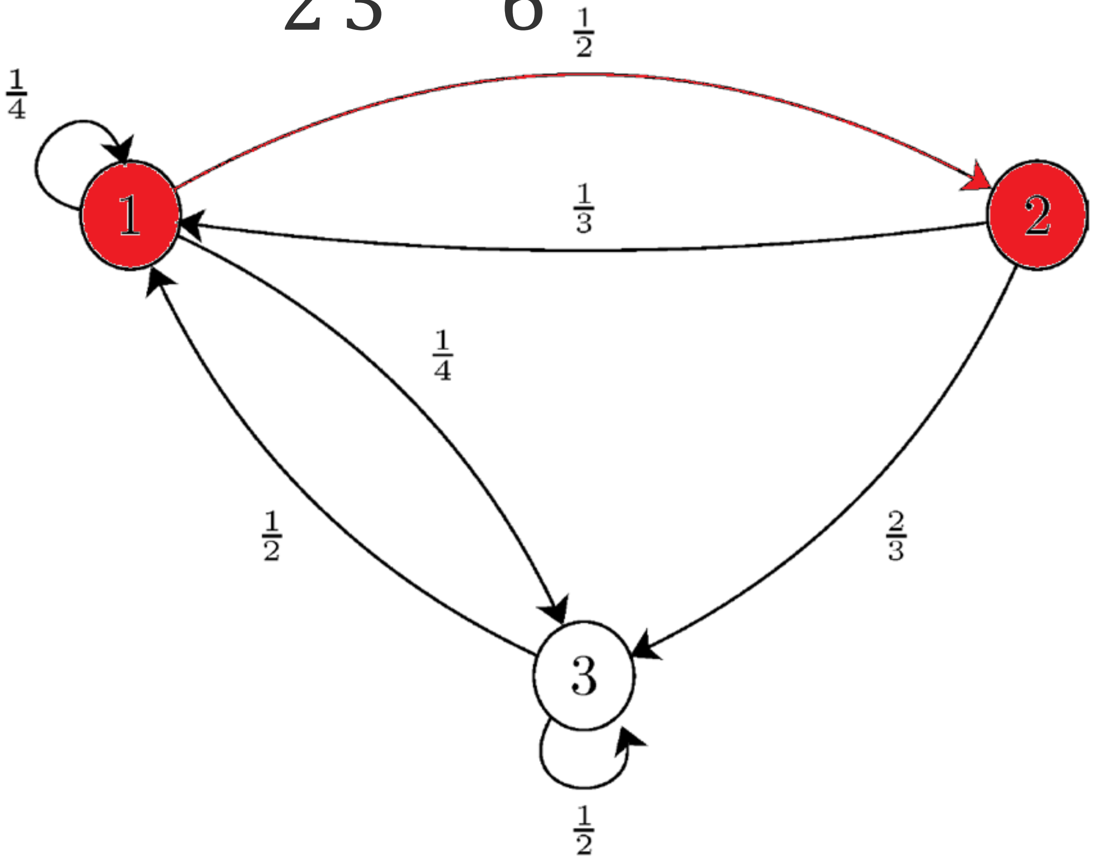
$$P(X_1 = 2 | X_0 = 1) = \frac{P(X_0 = 1, X_1 = 2)}{P(X_0 = 1)}$$

Exercise 8.4: Solution

- Thus $P(X_0 = 1, X_1 = 2) = P(X_1 = 2 | X_0 = 1)P(X_0 = 1)$

$$= p_{12}P(X_0 = 1) = \frac{1}{2}\frac{1}{3} = \frac{1}{6}$$

$$P = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$



Exercise 8.4: Solution

- d) If we know $P(X_0 = 1) = 1/3$, find $P(X_0 = 1, X_1 = 2, X_2 = 3)$.
- The boundary condition of the problem is: What is the probability of the Markov system to start in state 1 **and** then transit to state 2 **and** then transit to state 3?
- Using conditional probability, we have

$$P(X_2 = 3 | X_1 = 2, X_0 = 1) = \frac{P(X_0 = 1, X_1 = 2, X_2 = 3)}{P(X_0 = 1, X_1 = 2)}$$

$$P(X_1 = 2 | X_0 = 1) = \frac{P(X_0 = 1, X_1 = 2)}{P(X_0 = 1)}$$

Exercise 8.4: Solution

- Moving the probability of the intersection to the first member of the equations we get:

$$P(X_0 = 1, X_1 = 2, X_2 = 3)$$
$$= P(X_2 = 3|X_1 = 2, X_0 = 1\ )P(X_0 = 1, X_1 = 2)$$

$$P(X_0 = 1, X_1 = 2\ ) = P(X_1 = 2|\ X_0 = 1\ )P(X_0 = 1\ )$$

Exercise 8.4: Solution

- Substituting the second equation in the first, results:

$$P(X_0 = 1, X_1 = 2, X_2 = 3)$$

$$= P(X_2 = 3|X_1 = 2, X_0 = 1)P(X_1 = 2|X_0 = 1)P(X_0 = 1)$$

- From Markov property we know that

$$P(X_2 = 3|X_1 = 2, X_0 = 1) = P(X_2 = 3|X_1 = 2)$$

- Substituting in the first equation above we obtain

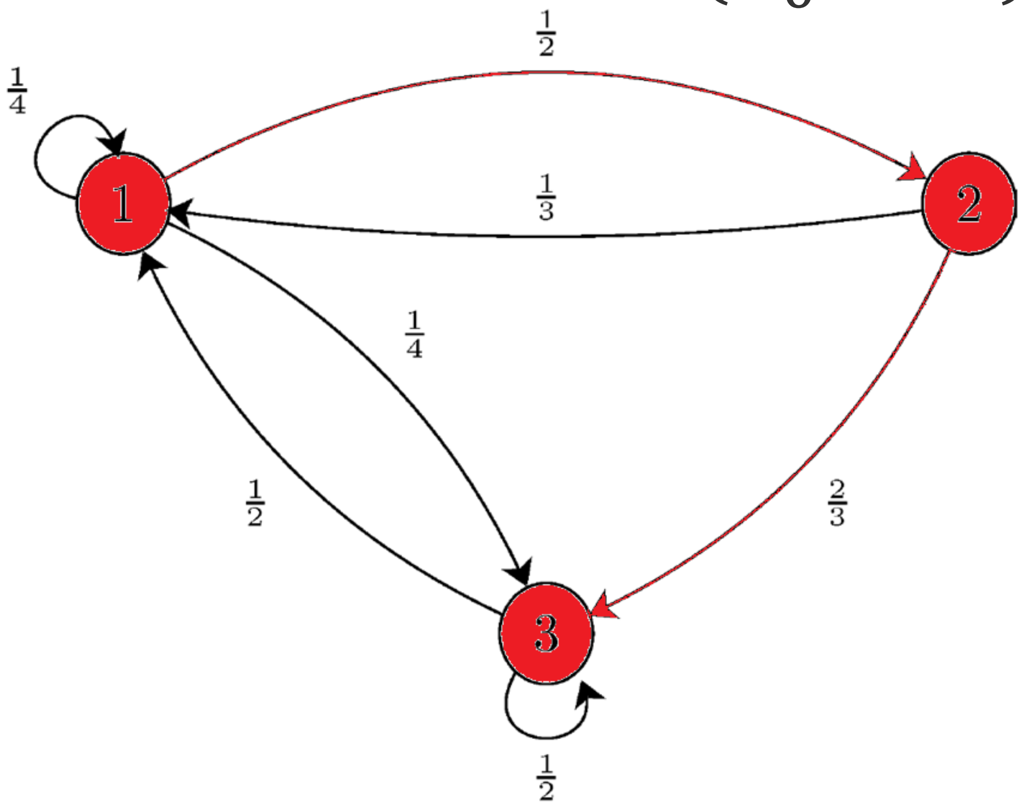
$$P(X_0 = 1, X_1 = 2, X_2 = 3)$$

$$= P(X_2 = 3|X_1 = 2)P(X_1 = 2|X_0 = 1)P(X_0 = 1)$$

Exercise 8.4: Solution

- $P(X_0 = 1, X_1 = 2, X_2 = 3)$ can be computed as

$$
\begin{aligned}
P(X_0 = 1, X_1 = 2, X_2 = 3)& \\
&= P(X_0 = 1\ )P(X_1 = 2|\ X_0 = 1\ )P(X_2 = 3|X_1 = 2\ ) \\
&= P(X_0 = 1\ )p_{12}p_{23} = \frac{1}{3}\frac{1}{2}\frac{2}{3} = \frac{1}{9}
\end{aligned}
$$



$$
P = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}
$$

Discrete-Time Markov Chains

- Consider a Markov chain $\{X_n, n = 0, 1, 2, . . . \}$, where $X_n \in S = \{1, 2, \cdots, r\}$.

- Suppose that we know the probability distribution of X_0. More specifically, define the row vector $\pi^{(0)}$ as

$$\pi^{(0)} = [P(X_0 = 1)\, P(X_0 = 2) \cdots P(X_0 = r)]$$

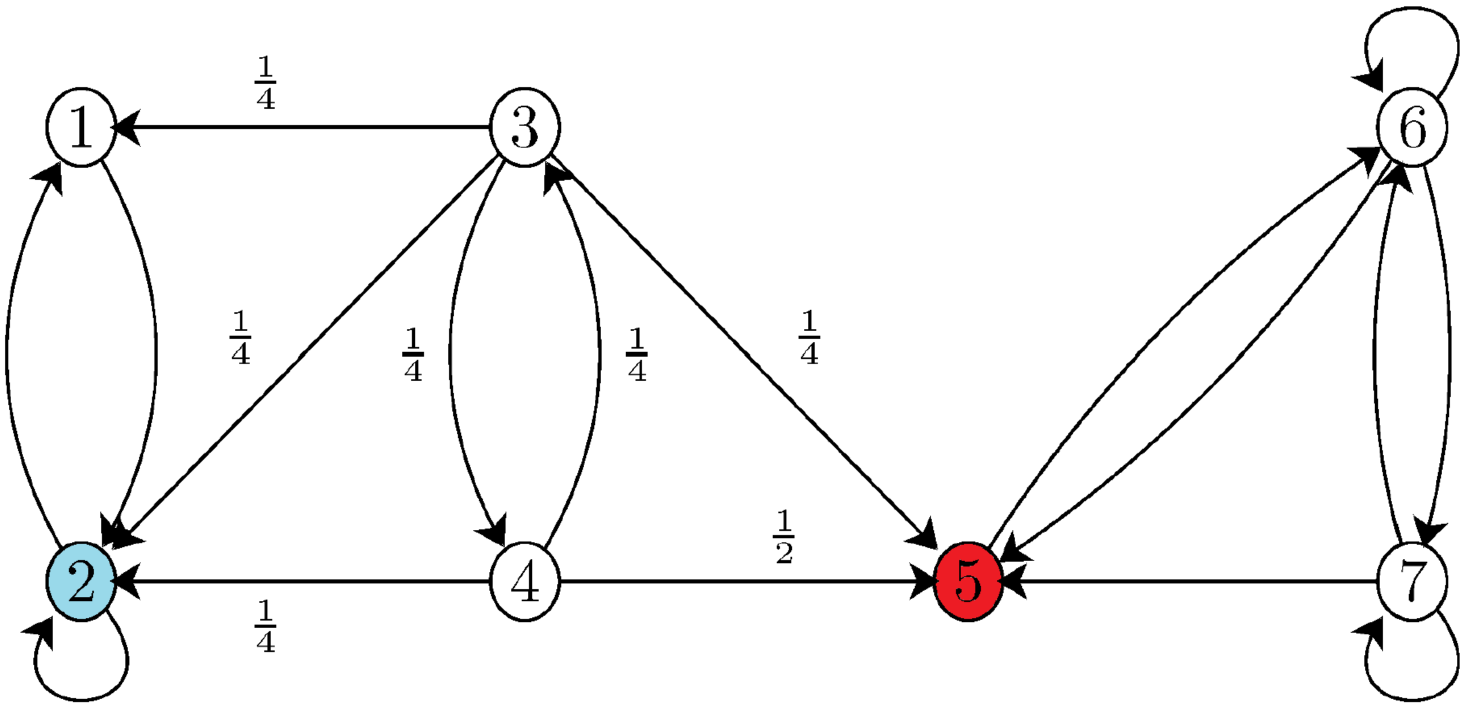
Discrete-Time Markov Chains

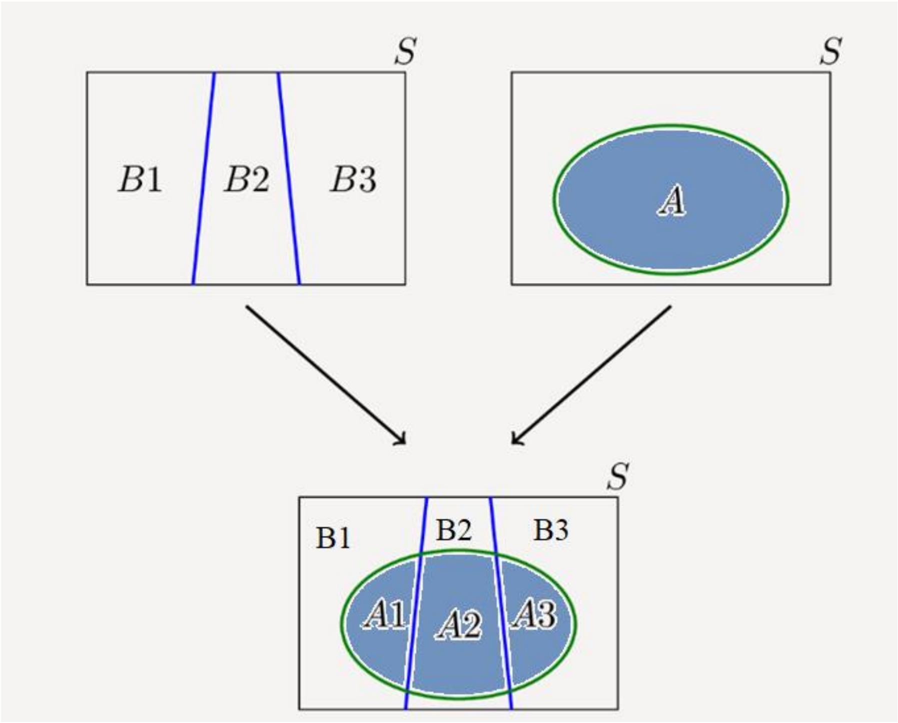
- We can use the law of total probability to obtain the probability distribution of $X_1, X_2, \cdots X_n \ldots$
- More specifically, for any $j \in S$, we can write

$$P(X_1 = j) = \sum_{k=1}^{r} P(X_1 = j | X_0 = k) P(X_0 = k) = \sum_{k=1}^{r} p_{kj} P(X_0 = k)$$

DISCRETE-TIME MARKOV CHAINS

$$P[X_1 = 2] = p_{12}P(X_0 = 1) + p_{22}P(X_0 = 2) + p_{32}P(X_0 = 3) + p_{42}P(X_0 = 4)$$





Law of total probability

$$P(A) = \sum_k P(A, B_k) = \sum_k P(A|B_k)P(B_k)$$

Discrete-Time Markov Chains

- We can generalize for each instant of time in the form

$$\pi^{(0)} = [P(X_0 = 1)\, P(X_0 = 2) \cdots P(X_0 = r)]$$

$$\pi^{(1)} = [P(X_1 = 1)\, P(X_1 = 2) \cdots P(X_1 = r)]$$

$$\vdots$$

$$\pi^{(n)} = [P(X_n = 1)\, P(X_n = 2) \cdots P(X_n = r)]$$

Discrete-Time Markov Chains

- It is possible to put the set of equations shown in the matrix form

$$\pi^{(0)} = [P(X_0 = 1)\, P(X_0 = 2) \cdots P(X_0 = r)]$$

$$\pi^{(1)} = \pi^{(0)}\mathbf{P}$$

$$\pi^{(2)} = \pi^{(1)}\mathbf{P} = \pi^{(0)}\mathbf{P}^2$$

$$\pi^{(3)} = \pi^{(2)}\mathbf{P} = \pi^{(1)}\mathbf{P}^2 = \pi^{(0)}\mathbf{P}^3$$

...

$$\pi^{(n)} = \pi^{(n-1)}\mathbf{P} = \pi^{(n-2)}\mathbf{P}^2 = \cdots = \pi^{(0)}\mathbf{P}^n$$

Exercise 8.5

- Consider a system that can be in one of two possible states, $S = \{0, 1\}$. In particular, suppose that the transition matrix is given by

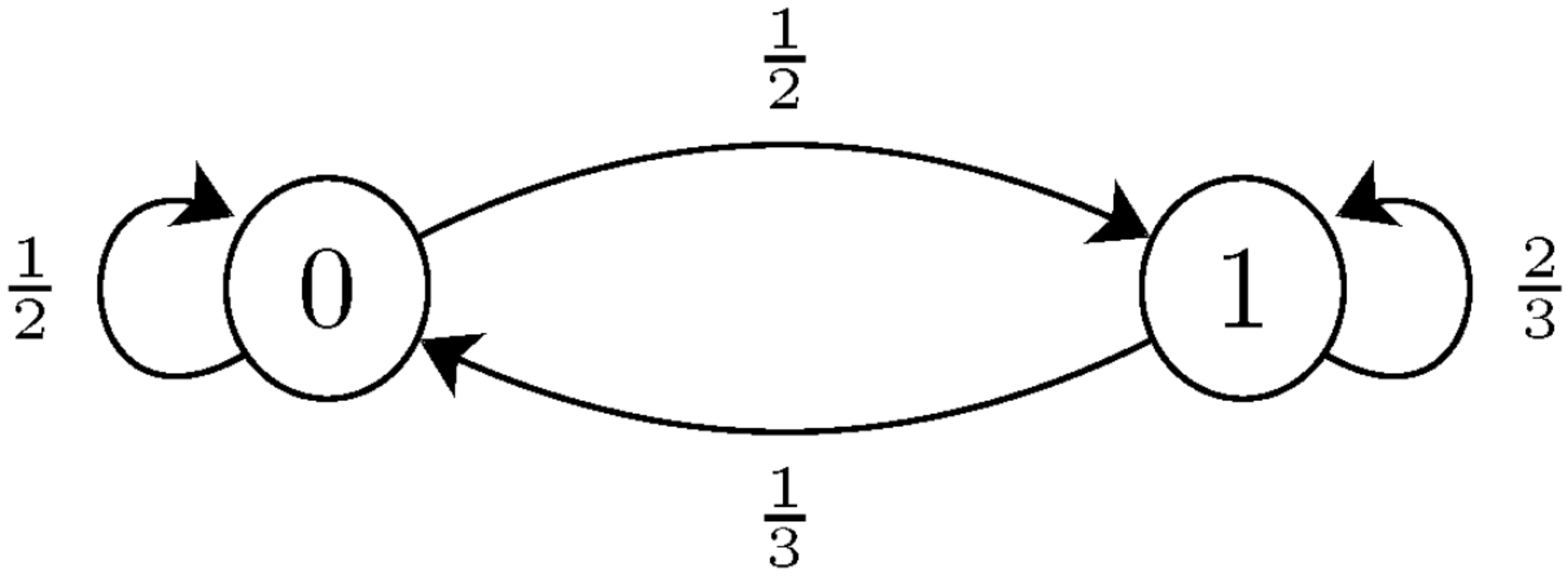
$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

- Suppose that the system is in state 0 at time $n = 0$, i.e., $X_0 = 0$.
- a) Draw the state transition diagram.
- b) Find the probability that the system is in state 1 at time $n = 3$.

Exercise 8.5: solution

- a) The state transition diagram

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$



Exercise 8.5: solution

- b) Find the probability that the system is in state 1 at time $n = 3$.

$$\pi^{(0)} = [P(X_0 = 0)P(X_0 = 1)] = [1\ 0]$$

- The probability that the system is in state **1** at time $n = 3$ is $\mathbf{\frac{43}{72}}$.

$$\pi^{(3)} = \pi^{(0)}\mathbf{P}^3 = [1\ 0]\left(\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}\right)^3 = \begin{bmatrix} \frac{29}{72} & \frac{43}{72} \end{bmatrix}$$

Discrete-Time Markov Chains

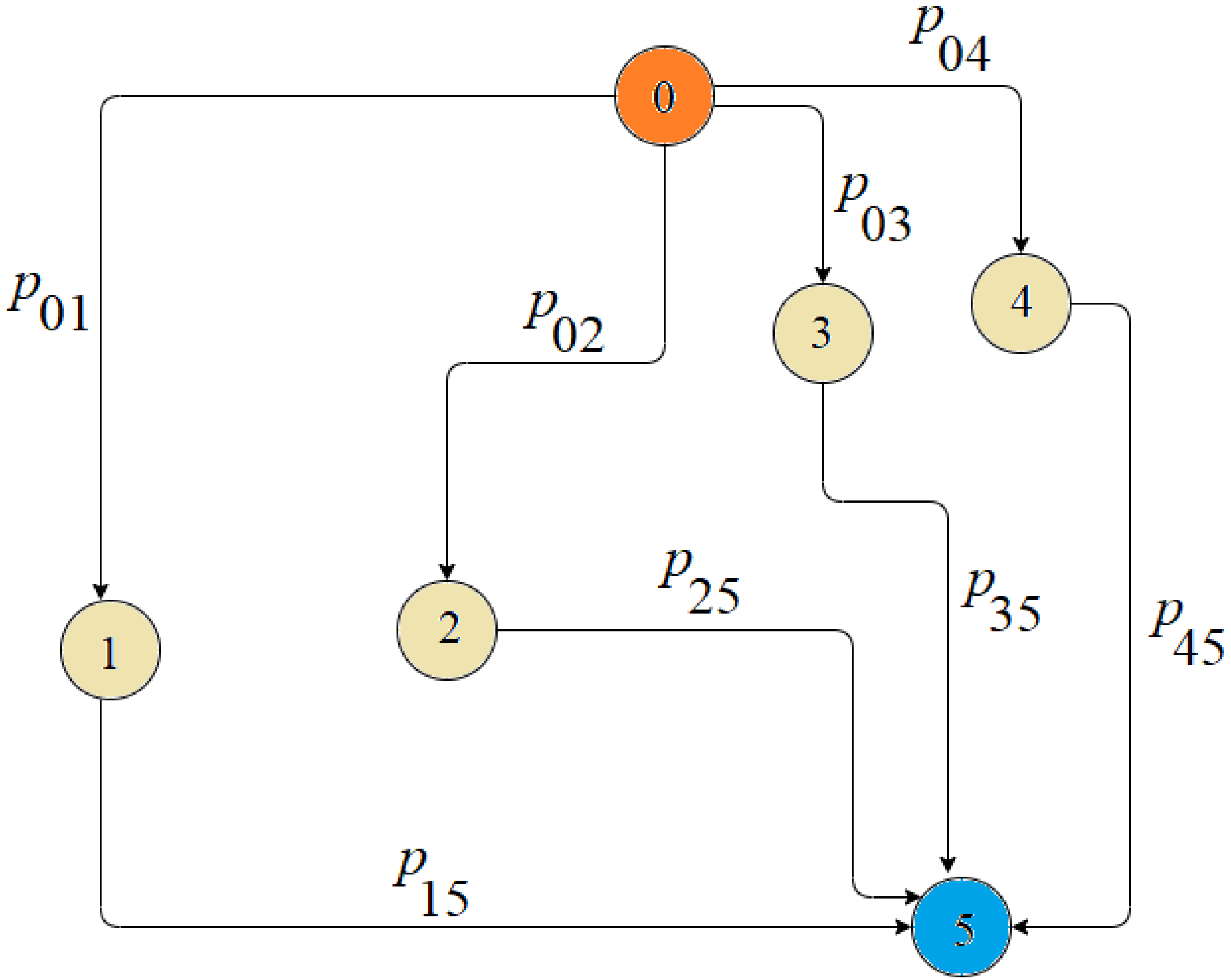
n-Step transition probabilities

- Consider a Markov chain $\{X_n, n = 0, 1, 2, \ldots\}$, where $X_n \in S$.
- If $X_0 = i$, then $X_1 = j$ with probability $\boldsymbol{p_{ij}}$. That is, $\boldsymbol{p_{ij}}$ gives us the probability of going from state ***i*** to state ***j*** in **one step**.
- Now suppose that we are interested in finding the probability of going from **state *i* to state *j* in two steps**

$$p_{ij}^{(2)} = P(X_2 = j | X_0 = i)$$

Discrete-Time Markov Chains

- Example:



$$p_{05}^{(2)} = p_{01}p_{15} + p_{02}p_{25} + p_{03}p_{35} + p_{04}p_{45}$$

Discrete-Time Markov Chains

- We can find this probability by applying the law of total probability. In particular, we argue that X_1 can take one of the possible values in S. Thus, we can write

$$p_{ij}^{(2)} = P(X_2 = j | X_0 = i)$$

$$= \sum_{k \in S} P(X_2 = j | X_1 = k, X_0 = \mathrm{i})P(X_1 = k | X_0 = \mathrm{i})$$

$$= \sum_{k \in S} P(X_2 = j | X_1 = k)P(X_1 = k | X_0 = \mathrm{i}) = \sum_{k \in S} p_{ik}\, p_{kj}$$

Discrete-Time Markov Chains

- We need to pass through some intermediate state k and the probability of this event is $p_{ik}p_{kj}$.

- Note that $p_{ij}^{(2)}$ corresponds to a linear combination of the probabilities of all paths.

- To recall the total probability theorem, $\boldsymbol{p_{ij}^{(2)}}$ can be seen as the **intersection of the event with each partition that defines the sample space**. In this case, **each of the partitions** would be **each of the paths** from state ***i*** to state ***j***.

Discrete-Time Markov Chains

- We can define the two-step transition matrix as follows:

$$\mathbf{P}^{(2)} = \begin{bmatrix} p_{11}^{(2)} & p_{12}^{(2)} & p_{13}^{(2)} & \cdots & p_{1r}^{(2)} \\ p_{21}^{(2)} & p_{22}^{(2)} & p_{23}^{(2)} & \cdots & p_{2r}^{(2)} \\ p_{31}^{(2)} & p_{32}^{(2)} & p_{33}^{(2)} & & p_{3r}^{(2)} \\ & \vdots & & \ddots & \vdots \\ p_{r1}^{(2)} & p_{r2}^{(2)} & p_{r3}^{(2)} & \cdots & p_{rr}^{(2)} \end{bmatrix}$$

Discrete-Time Markov Chains

- Notice that $\boldsymbol{p_{ij}^{(2)}}$ is in fact the element in the *i*-th row and *j*-th column of the matrix

$$\mathbf{P}^{(2)} = \mathbf{P}^2 = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1r} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2r} \\ p_{31} & p_{32} & p_{33} & & p_{3r} \\ & \vdots & & \ddots & \vdots \\ p_{r1} & p_{r2} & p_{r3} & \cdots & p_{rr} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1r} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2r} \\ p_{31} & p_{32} & p_{33} & & p_{3r} \\ & \vdots & & \ddots & \vdots \\ p_{r1} & p_{r2} & p_{r3} & \cdots & p_{rr} \end{bmatrix}$$

- Observe that

$$p_{11}^{(2)} = \sum_{k=1}^{r} p_{1k}\, p_{k1} \qquad p_{12}^{(2)} = \sum_{k=1}^{r} p_{1k}\, p_{k2}$$

Discrete-Time Markov Chains

- We can generalize this result to a **n-steps transition probabilities** $p_{ij}^{(n)}$ according to the trajectory in the state diagram.

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i), for\ n = 0,1,2, \ldots$$

Discrete-Time Markov Chains

- In the matrix form we can express $p_{ij}^{(n)}$ as

$$p_{ij}^{(n)} = \mathbf{P}^{(n)} = \begin{bmatrix} p_{11}^{(n)} & p_{12}^{(n)} & p_{13}^{(n)} & \cdots & p_{1r}^{(n)} \\ p_{21}^{(n)} & p_{22}^{(n)} & p_{23}^{(n)} & \cdots & p_{2r}^{(n)} \\ p_{31}^{(n)} & p_{32}^{(n)} & p_{33}^{(n)} & & p_{3r}^{(n)} \\ & \vdots & & \ddots & \vdots \\ p_{r1}^{(n)} & p_{r2}^{(n)} & p_{r3}^{(n)} & \cdots & p_{rr}^{(n)} \end{bmatrix}$$

Discrete-Time Markov Chains

- We can now generalize the last equation. Considert ***m*** and ***n*** be two positive integers and assume $\boldsymbol{X_0} = \boldsymbol{i}$.
- In order to get to state ***j*** in (***m***+***n***) steps, the chain will be at some intermediate state *k* after *m* steps. To obtain $p_{ij}^{(n+m)}$, we sum over all possible intermediate states:

$$p_{ij}^{(n+m)} = P(X_{n+m} = j | X_0 = i\ ) = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}$$

- This equation is called the **Chapman-Kolmogorov equation**

Discrete-Time Markov Chains

- **Classes of states**
- We say that state $\boldsymbol{j}$ is **accessible** from state $\boldsymbol{i}$, written as $\boldsymbol{i} \rightarrow \boldsymbol{j}$, if $p_{ij}^{(n)} > 0$ for some n.
- Two states $\boldsymbol{i}$ and $\boldsymbol{j}$ are said to **communicate**, written as $\boldsymbol{i} \leftrightarrow \boldsymbol{j}$, ($\boldsymbol{i} \leftrightarrow \boldsymbol{j}$ means $\boldsymbol{i} \rightarrow \boldsymbol{j}$ and $\boldsymbol{j} \rightarrow \boldsymbol{i}$) if they are **accessible** from each other.
- If $\boldsymbol{i} \leftrightarrow \boldsymbol{j}$ and $\boldsymbol{j} \leftrightarrow \boldsymbol{k}$, then $\boldsymbol{i} \leftrightarrow \boldsymbol{k}$.
- The long-term behavior of a Markov chain is related to the types of its **state classes**.

Example:

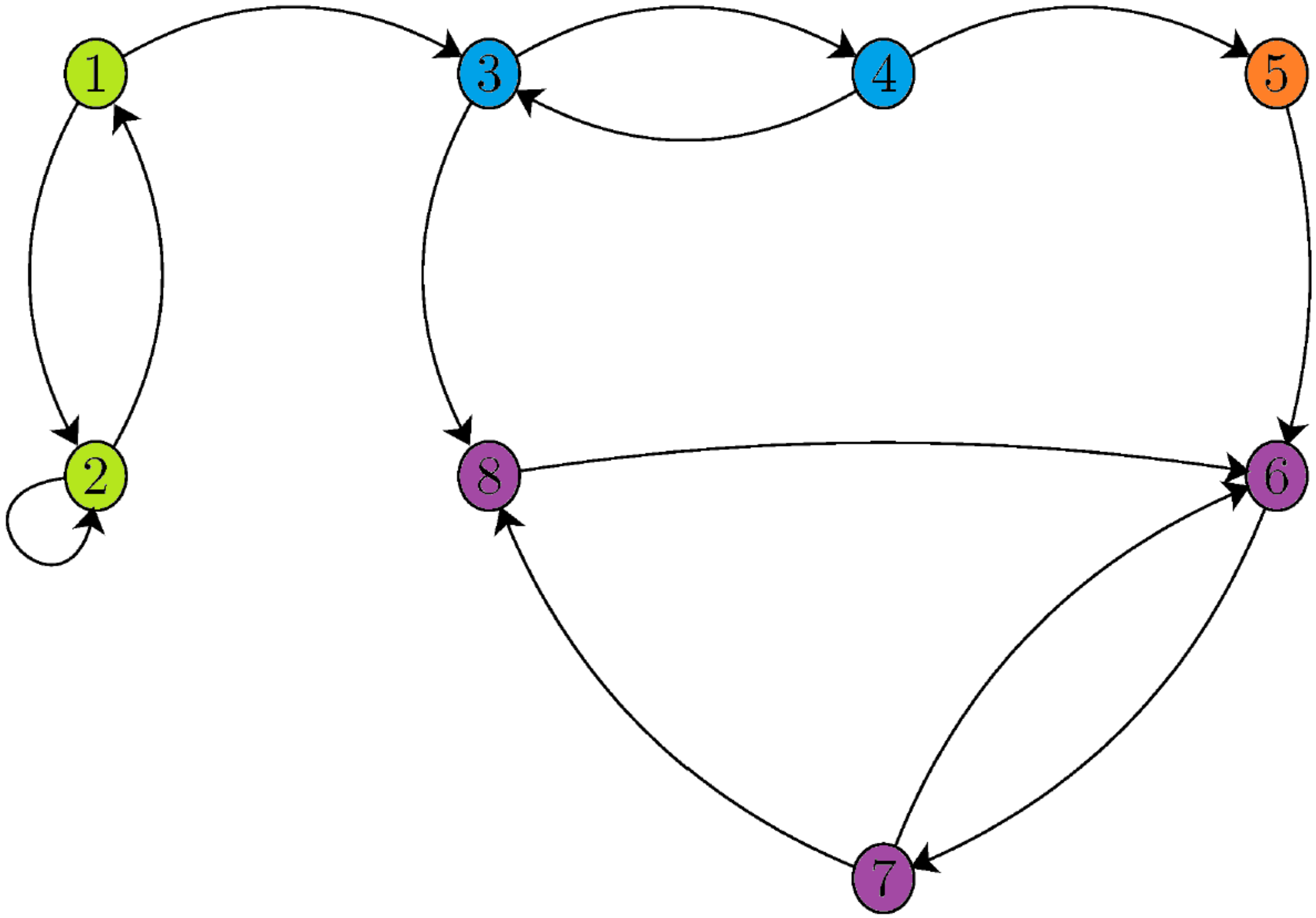
- Consider the markov chain shown in the state transition diagram. It is assumed that when there is an arrow from state ***i*** to state ***j***, then p_{ik}> 0.

- States 1 and 2 communicate with each other, but they do not communicate with any other nodes in the graph: **class 1** = {state 1, state 2}.

- States 3 and 4 communicate with each other, but they do not communicate with any other nodes in the graph: **class 2** = {state 3, state 4}.

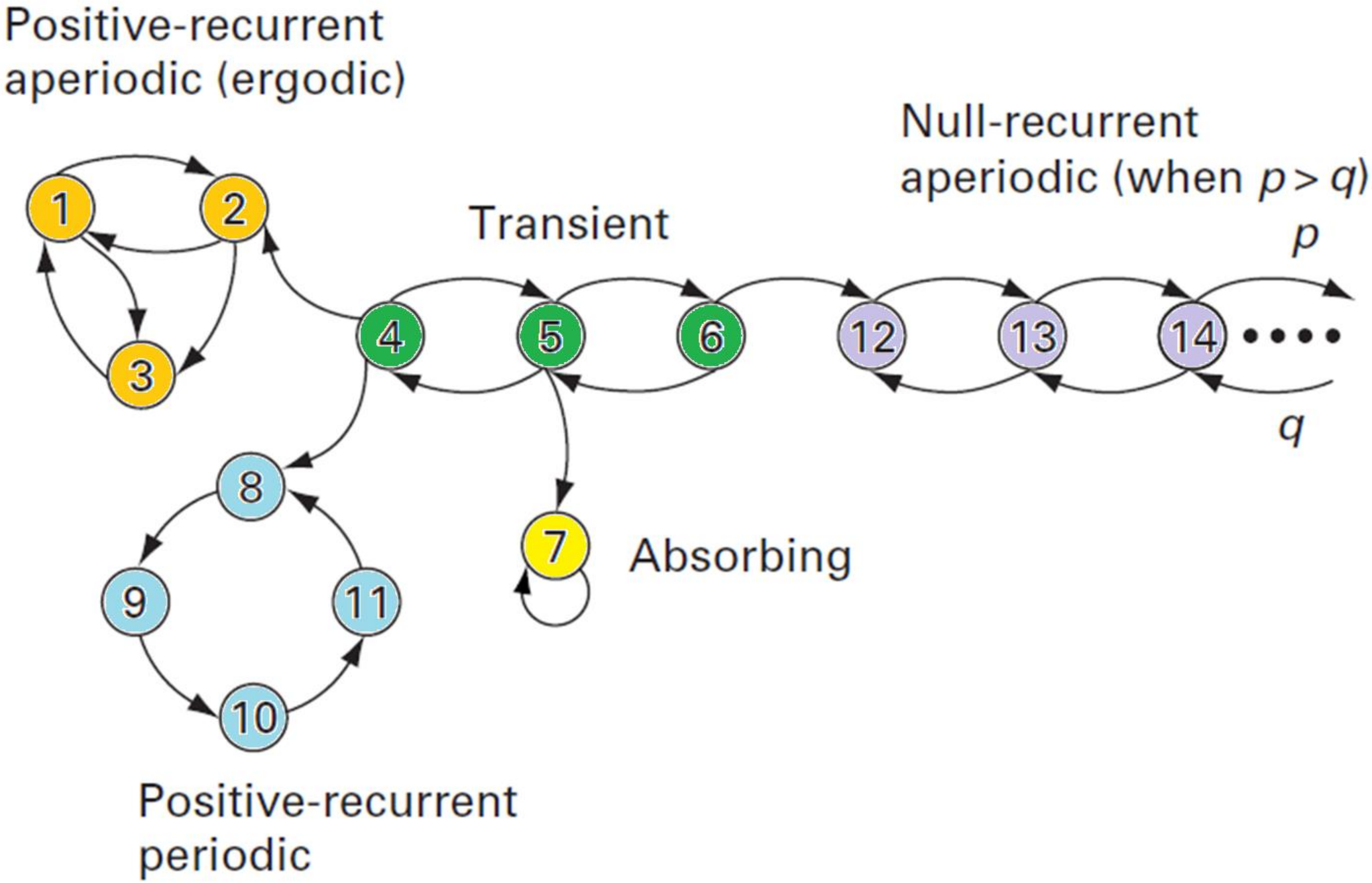
- State 5 does not communicate with any other states, so it by itself is a class: **class 3** = {state 5}.

- States 6 , 7, and 8 construct another class: **class 4** = {state 6, state 7, state 8}.



Discrete-Time Markov Chains

- An example of a Markov chain with various states and classes.



Discrete-Time Markov Chains

- A Markov chain is said to be **irreducible** if all states communicate with each other.
- For any state ***i***, we define

$$f_{ii} = P(X_n = i, \text{ for some } n \geq 1 | X_0 = i).$$

- State ***i*** is **recurrent** if $f_{ii} = 1$, and it is **transient** if $f_{ii} < 1$.
- The **absorbing** states:once you enter those states, you never leave them.

Discrete-Time Markov Chains

- Consider a discrete-time Markov chain. Let **V** be the total number of visits to state ***i***.
  - If ***i*** is a **recurrent** state, then

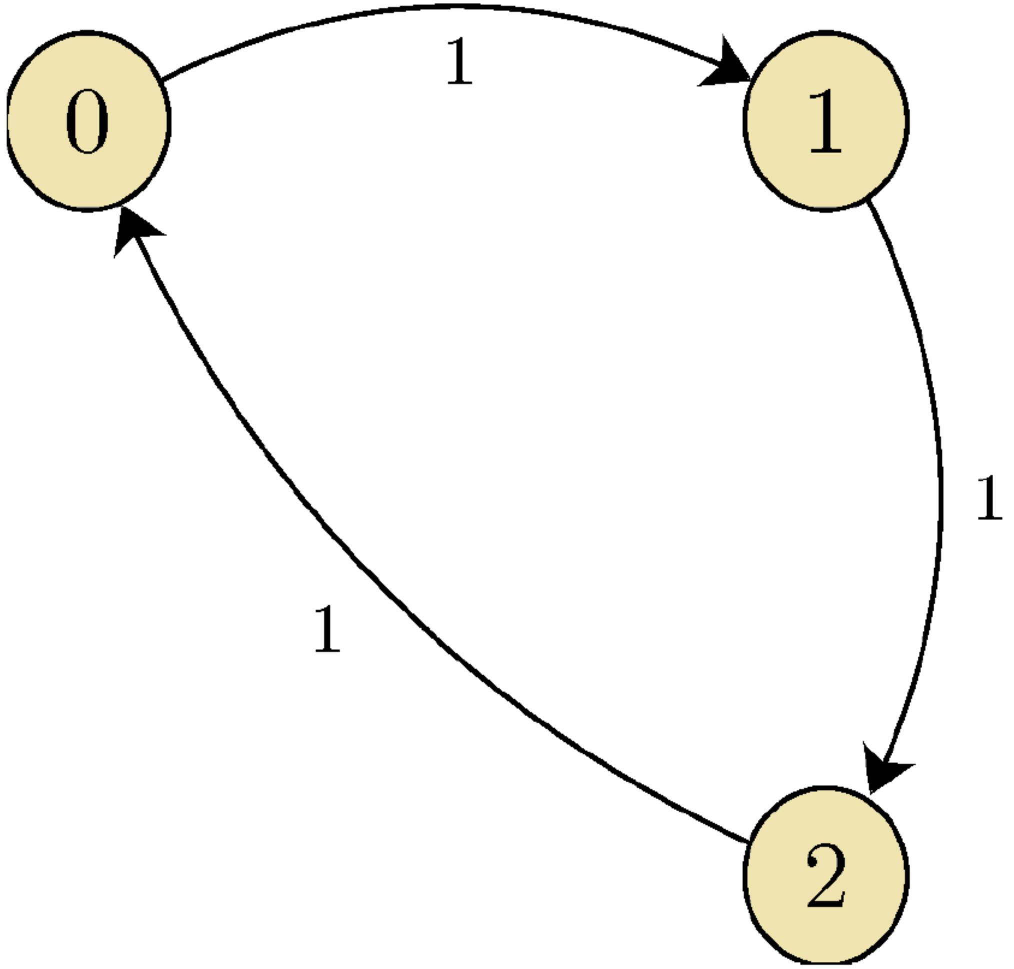
$$\boldsymbol{P\,(V = \infty | X_0 = i) = 1}.$$

  - If ***i*** is a **transient** state, then:

$$V\,|X_0 = \boldsymbol{i} \sim Geometric(1 - f_{ii}).$$

Periodicity of discrete-time Markov chain

- A periodic pattern is verified in this chain on the right.
- Starting from state 0, we only return to 0 at times $n = 3, 6, \cdots$.
- In other words, $p_{00}^{(n)} = 0$, if $\boldsymbol{n}$ is not divisible by 3. Such a state is called a ***periodic*** state with period $\mathbf{d(0) = 3}$.

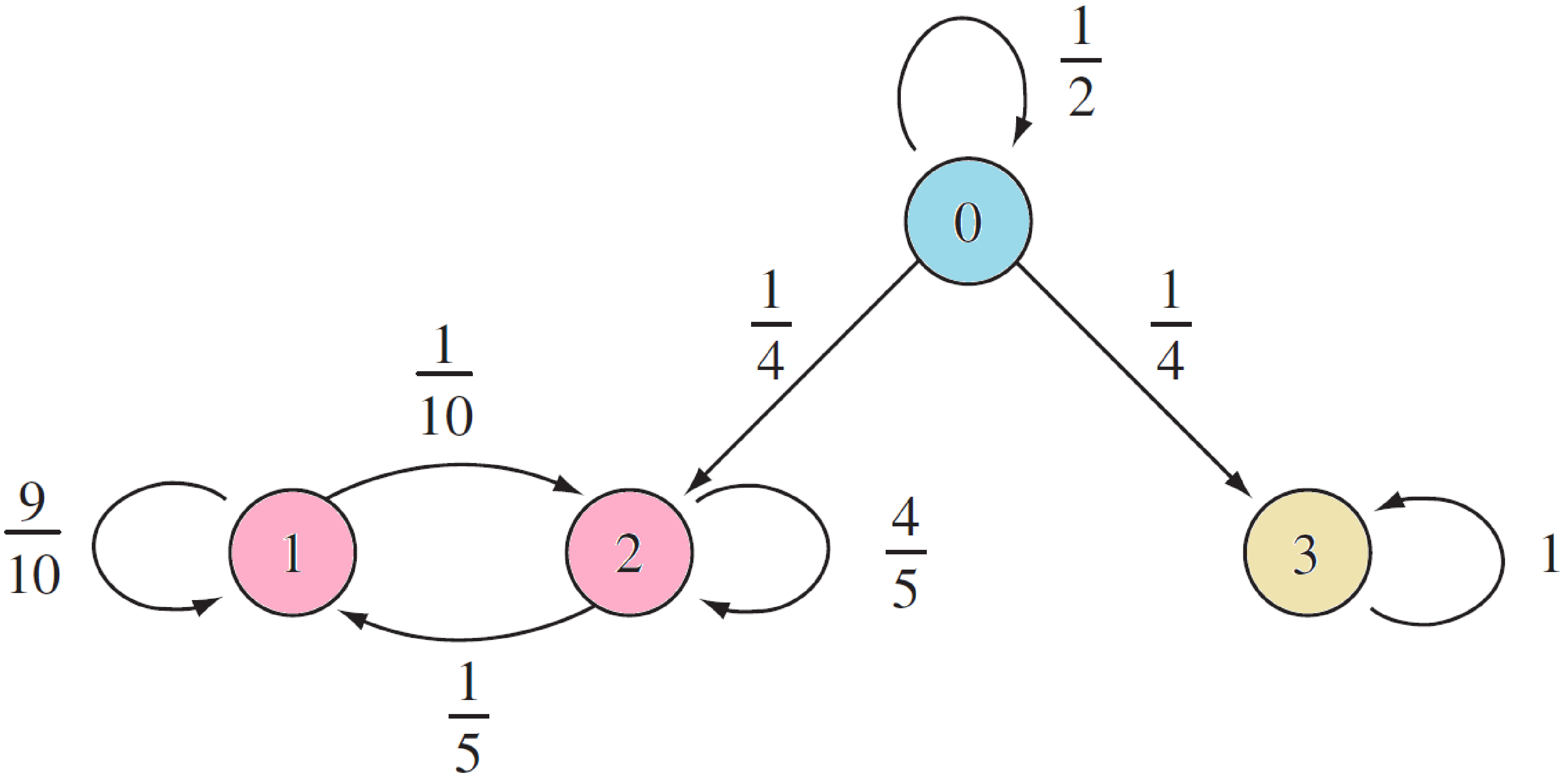


Discrete-Time Markov Chains

- The **period** of a state $\boldsymbol{i}$ is the largest integer $\boldsymbol{d}$ satisfying the following property: $p_{ii}^{(n)} = 0$, whenever $\boldsymbol{n}$ is not divisible by $\boldsymbol{d}$. The period of $\boldsymbol{i}$ is shown by $\boldsymbol{d}(\boldsymbol{i})$. If $p_{ii}^{(n)} = 0$, for all $\boldsymbol{n} > 0$, then we let $\boldsymbol{d}(\boldsymbol{i}) = \infty$.
  - If $\boldsymbol{d}(\boldsymbol{i}) > 1$, we say that state $\boldsymbol{i}$ is **periodic**.
  - If $\boldsymbol{d}(\boldsymbol{i}) = 1$, we say that state $\boldsymbol{i}$ is **aperiodic**.
  - If $\boldsymbol{i} \leftrightarrow \boldsymbol{j}$, then $\boldsymbol{d}(\boldsymbol{i}) = \boldsymbol{d}(\boldsymbol{j})$.

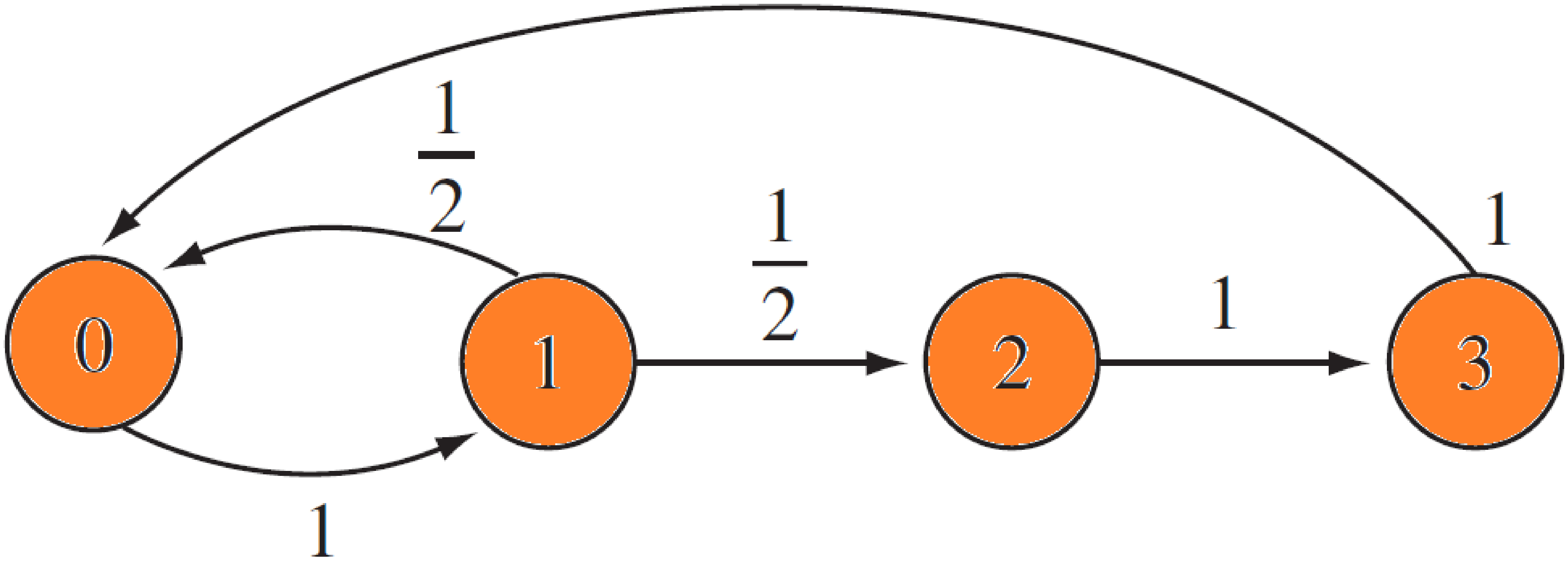
Discrete-Time Markov Chains

- A three-class Markov chain.



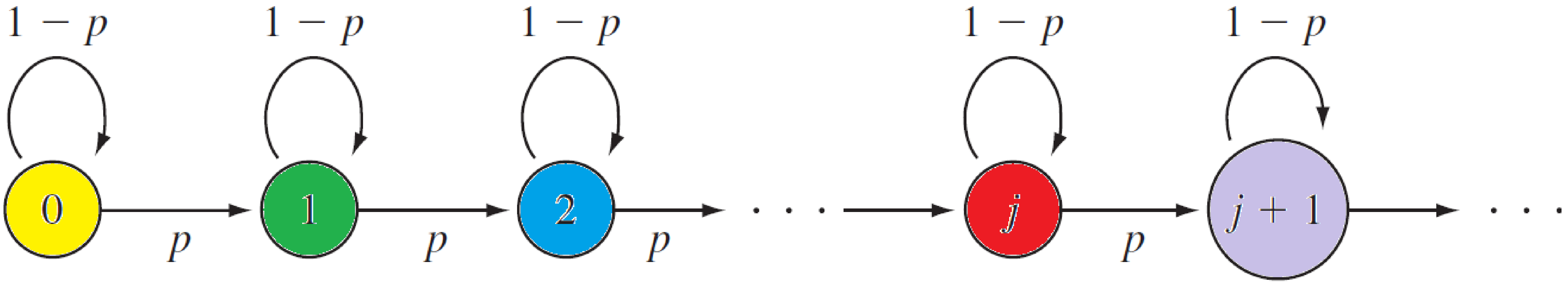
Discrete-Time Markov Chains

- Aperiodic Markov chain



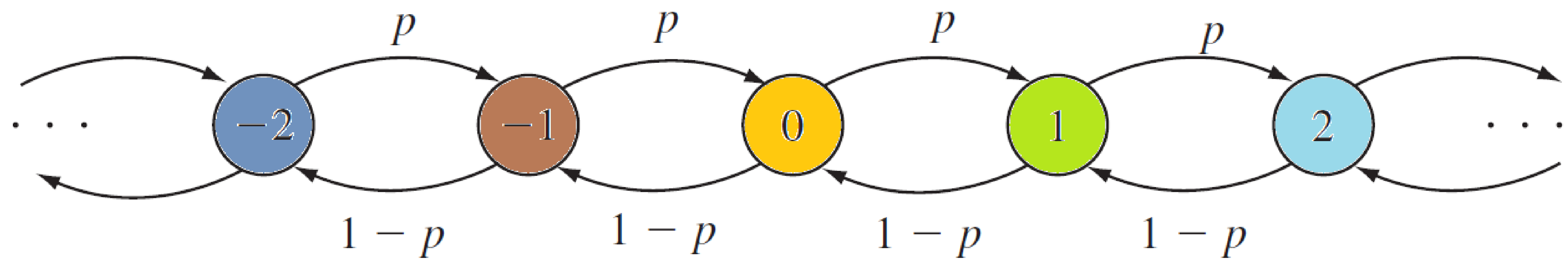
Discrete-Time Markov Chains

- A binomial counting process.



Discrete-Time Markov Chains

- The random walk process.



Markov Chain: Stationary and Limiting Distributions

- The **Long-term behavior of Markov** chains is an important issue to know.
- It means the fraction of times that the Markov chain spends in each state as n becomes large.
- More specifically, we would like to study the distributions

$$\pi^{(n)} = \pi^{(n-1)}\mathbf{P} = \pi^{(n-2)}\mathbf{P}^2 = \cdots = \pi^{(0)}\mathbf{P}^n$$

$$\pi^{(n)} = [p(X_n = 0)\ p(X_n = 1)\ \cdots\ ] \text{ as } n \to \infty.$$

Markov Chain: Limiting Distributions

- The probability distribution $\pi = [\pi_0, \pi_1, \pi_2, \cdots]$ is called the **limiting distribution** of the Markov chain X_n if

$$\pi_j \equiv \lim_{n \to \infty} P(X_n = j | X_0 = i)$$

- for all $i, j \in S$, and we have

$$\sum_{j \epsilon S} \pi_j = 1$$

Markov Chain: Limiting Distributions

- When a limiting distribution **exists**, it does not depend on the initial state ($X_0 = i$), so we can write

$$\pi_j = \lim_{n \to \infty} P(X_n = j)$$

- For all $j \in S$.

Finite Irredutible Markov Chain

- If there is a self-transition in the chain ($p_{ii} > 0$ for some i), then the chain is aperiodic.

- The chain is aperiodic if and only if there exists a positive integer ***n*** such that all elements of the matrix P^n are strictly positive, i.e.,

$$P_{ij}^{(n)} > 0, \text{ for all } i, j \in s.$$

Finite Markov Chain: Stationary Distributions

- We consider Markov chains with a finite number of states, a finite Markov chain can consist of several **transient** as well as **recurrent** states.

- As ***n*** becomes large the chain will enter a **recurrent class** and it will stay there forever. Therefore, when studying long-run behaviors we focus only on the recurrent classes.

Finite Markov Chain: Stationary Distributions

- If a finite Markov chain has more than one recurrent class, then the chain will get absorbed in one of the recurrent classes.
- We are going to limit our attention to the case where our Markov chain consists of one recurrent class.
- In finite **irreducible** Markov chains, all states are **recurrent**.
- It turns out that in this case the Markov chain has a well-defined limiting behavior if it is aperiodic.

Finite Markov Chain: Stationary Distributions

- If $\pi = [\pi_1, \pi_2, \cdots]$ is a limiting distribution for a Markov chain, then we have

$$\pi = \lim_{n\to\infty} \pi^{(n)} = \lim_{n\to\infty} \pi^{(0)} P^n$$

- Similarly, we can write

$$\pi = \lim_{n\to\infty} \pi^{(n+1)} = \lim_{n\to\infty} \left\{\pi^{(0)} P^{n+1}\right\}$$

$$= \lim_{n\to\infty} \left\{\pi^{(0)} P^n P\right\} = \left(\lim_{n\to\infty} \left\{\pi^{(0)} P^n\right\}\right) P = \pi P$$

Finite Markov Chain: Stationary Distributions

- The chain has reached its ***steady-state*** (limiting) distribution.
- We can equivalently write $\pi = \pi P$ as

$$\pi_j = \sum_{k \epsilon S} \pi_k p_{kj}$$

- If the chain is also **aperiodic**, we conclude that **the stationary distribution** is a **limiting distribution**.

Finite Markov Chain: Stationary Distributions

- Consider a finite Markov chain $\{X_n, n = 0, 1, 2, . . . \}$ where $X_n \in S = \{0, 1, 2, \cdots, r\}$.
- Assume that the chain is **irreducible and aperiodic**. Then, the set of equations

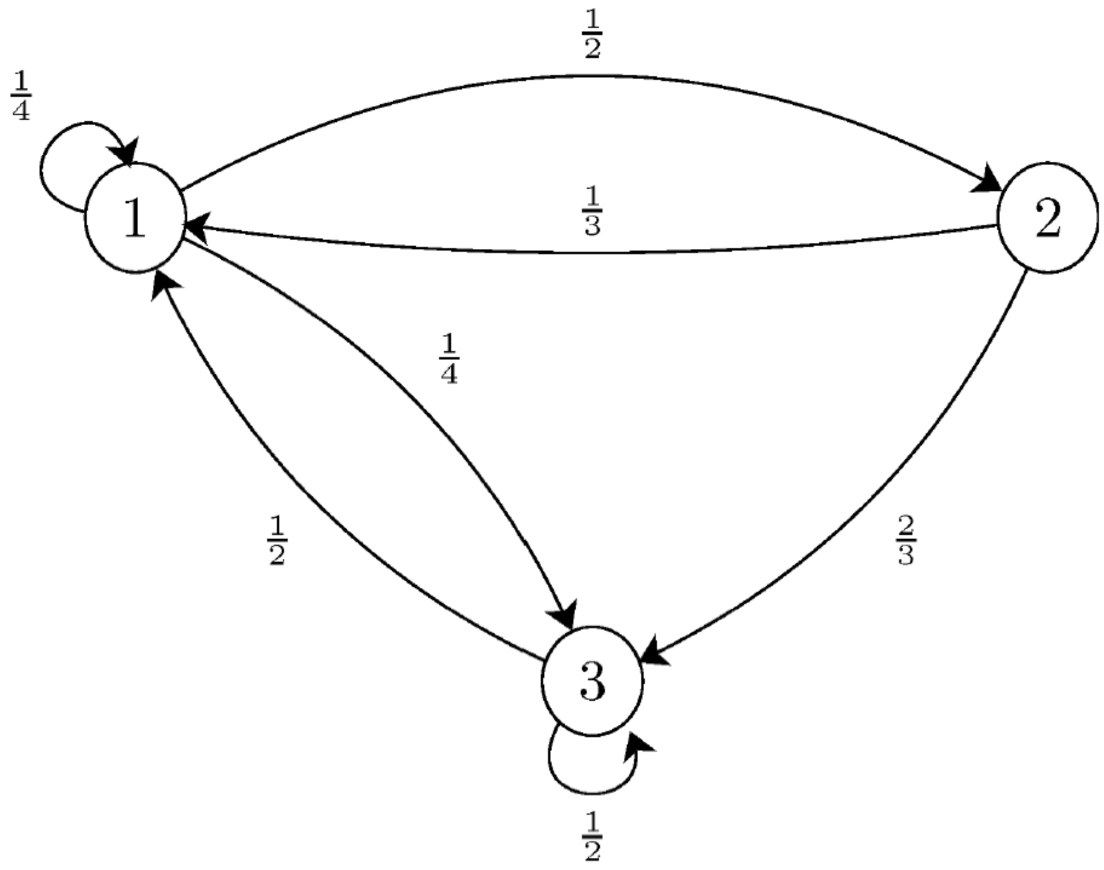
$$\pi = \pi P \qquad \sum_{j \epsilon S} \pi_j = 1$$

irreducible = all states communicate with each other.

has a unique solution.

Exercise 8.6

- Consider the Markov chain shown in figure below



- a) Find the stationary distribution for this chain
- b) Is the stationary distribution a limiting distribution for the chain?

Exercise 8.6: solution

- a) Find the stationary distribution for this chain
- The **transition probability** matrix can be written as

$$P = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

Exercise 8.6: solution

- We can build the system of equations from

$$\pi = \pi P \;\; or \; \pi_j = \sum_{k \epsilon S} \pi_k P_{kj} \qquad\qquad \sum_{j \epsilon S} \pi_j = 1$$

- As a result, we have

$$[\pi_1 \quad \pi_2 \quad \pi_3] = [\pi_1 \quad \pi_2 \quad \pi_3] \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}, \quad \pi_1 + \pi_2 + \pi_3 = 1$$

Exercise 8.6: solution

- Solving the system of equations, we obtain

$$\pi_1 = \frac{3}{8}, \pi_2 = \frac{3}{16}, \pi_3 = \frac{7}{16}$$

- b) Is the stationary distribution a limiting distribution for the chain?
- Since the chain is irreducible and aperiodic, we conclude that the above stationary distribution is a limiting distribution.

Countably Infinite Markov Chains

- Let ***i*** be a recurrent state. Assuming $X_0 = i$, let R_i be the number of transitions needed to return to state ***i***, i.e.,

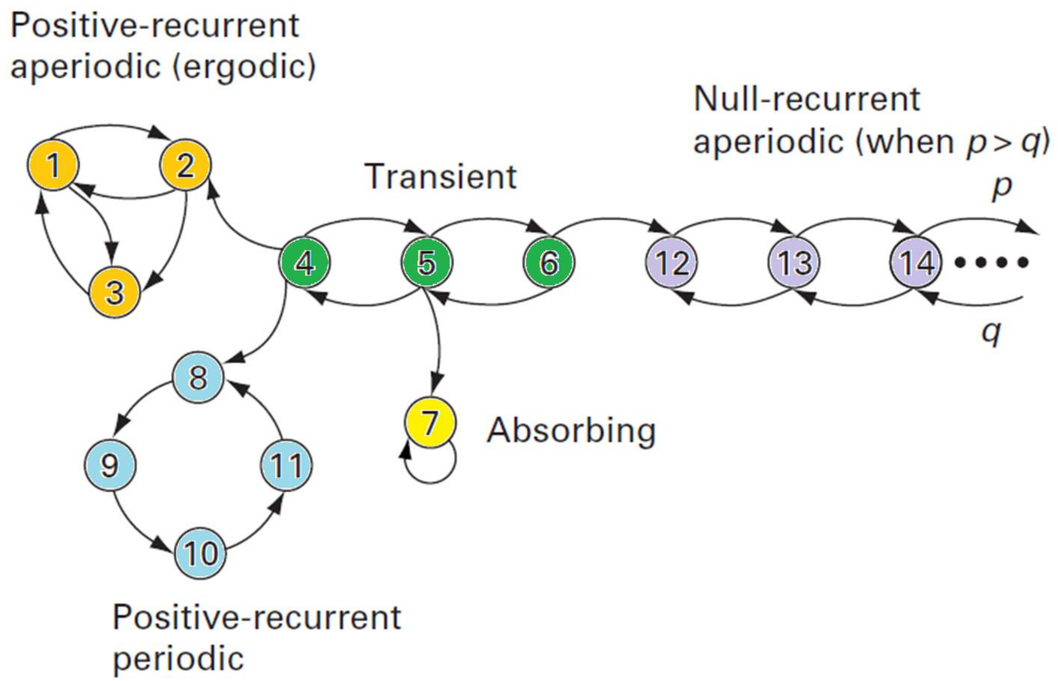
$$R_i = \min\{n \geq 1 : X_n = i\}.$$

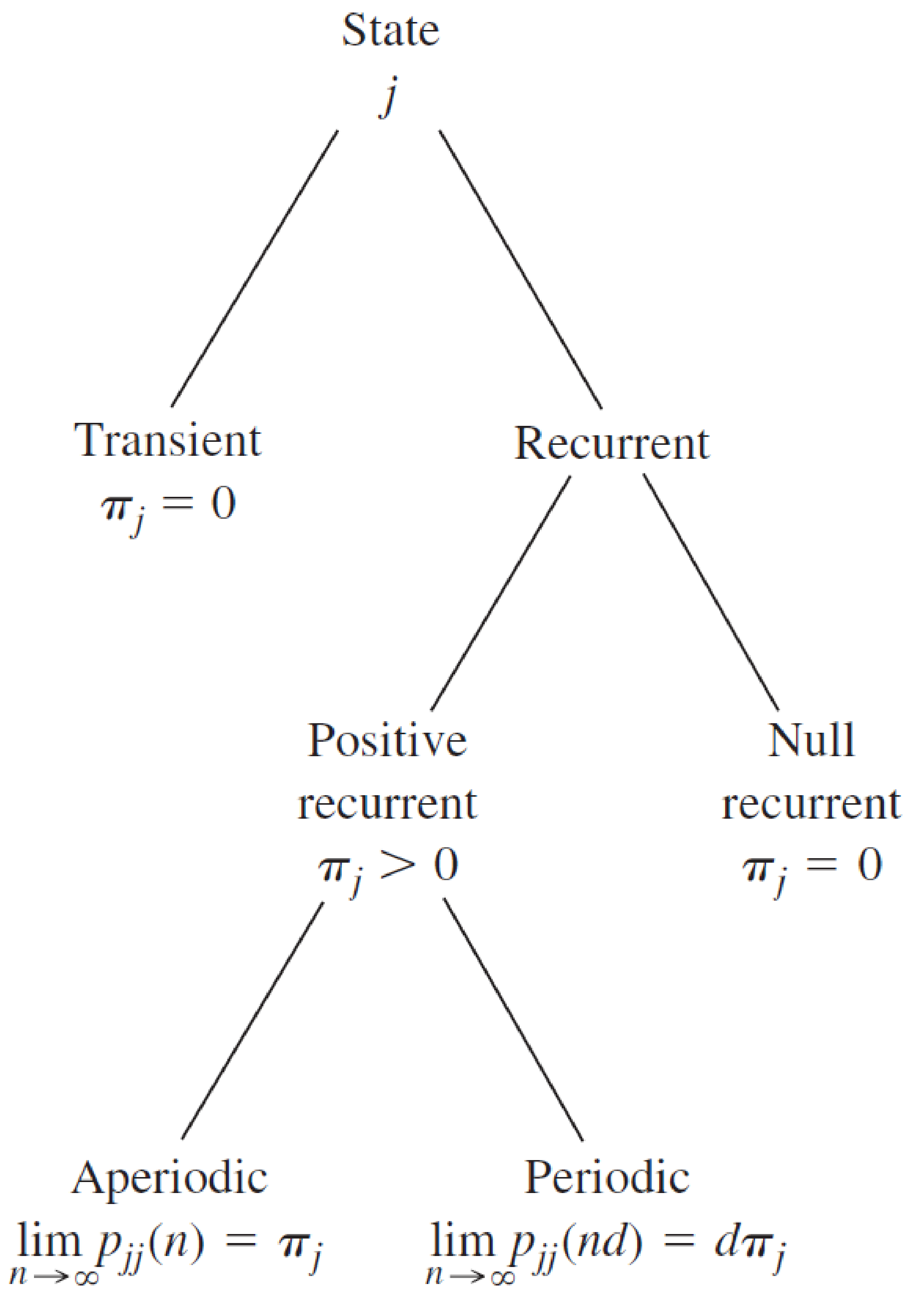
- If $\mathrm{R}_i = E[R_i|X_0 = i] < \infty$, then ***i*** is said to be **positive recurrent**.

- If $E[R_i|X_0 = i] = \infty$, then ***i*** is said to be **null recurrent**.

Discrete-Time Markov Chains

- Classification of states and associated longterm behavior. The proportion of “time” spent in state j is denoted by π_j.





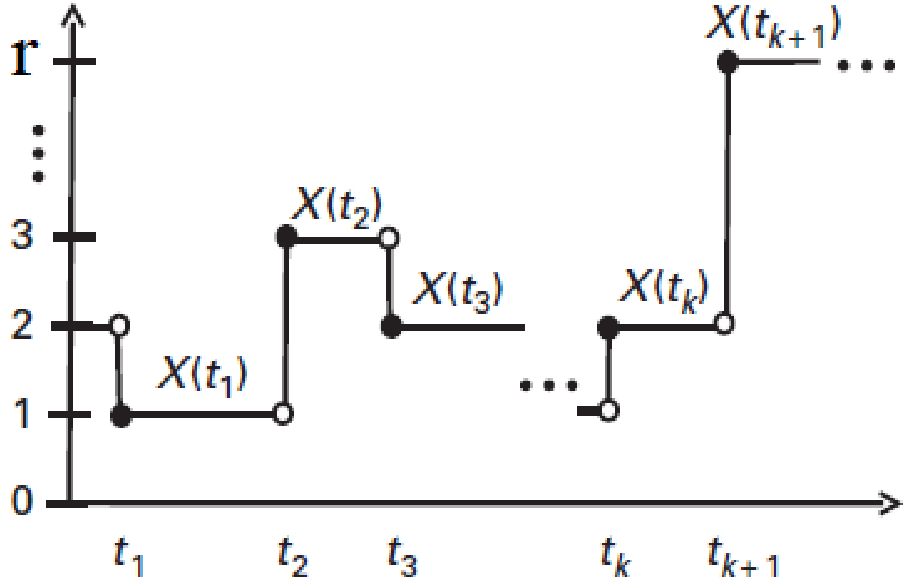
Continuous-Time Markov process

- If $X(t)$ is a **continuous-valued Markov process**, thus for $t_k < t < t_{k+1}$

$$P[X(t_{k+1})|X(t_k) = x_k, X(t_{k-1}) = x_{k-1}, \dots, X(t_1) = x_1] = P[X(t_{k+1})|X(t_k) = x_k]$$

- If the samples of $X(t)$ are jointly continuous, then it is equivalent to

$$f_{X(t_{k-1})}(x_{k+1}|X(t_k) = i_k, X(t_{k-1}) = i_{k-1}, \dots, X(t_1) = i_1)$$
$$= f_X(x_{k+1}|X(t_k) = i_k)$$



$$X(t) = i_k, \text{ for } t_k \le t < t_{k+1},$$

Continuous-Time Markov process

- Continuous-time Markov chain

A continuous-time Markov chain $X(t)$ is defined by two components: a *jump chain*, and a set of *holding time parameters* λ_i. The jump chain consists of a countable set of states $S \subset \{0,1,2,\cdots\}$ along with transition probabilities p_{ij}. We assume $p_{ii}=0$, for all non-absorbing states $i \in S$. We assume

1. if $X(t)=i$, the time until the state changes has $Exponential(\lambda_i)$ distribution;
2. if $X(t)=i$, the next state will be j with probability p_{ij}.

The process satisfies the Markov property. That is, for all $0 \leq t_1 < t_2 < \cdots < t_n < t_{n+1}$, we have

$$P\Big(X(t_{n+1})=j \Big| X(t_n)=i, X(t_{n-1})=i_{n-1}, \cdots, X(t_1)=i_1\Big) = P\big(X(t_{n+1})=j \big| X(t_n)=i\big).$$

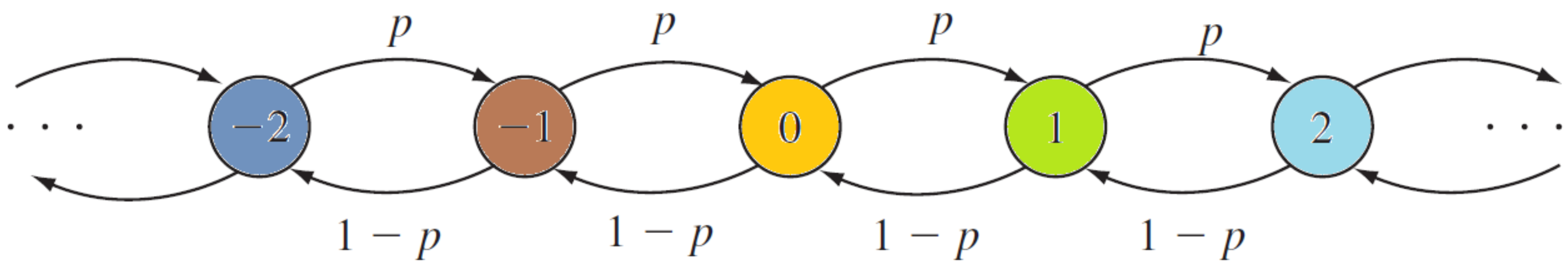
Continuous-Time Markov process

- We can then define the *transition matrix*, $p(t)$. Assuming the states are 1, 2, $\cdots$, r, then the state transition matrix for any $t \geq 0$ is given by

$$\mathbf{P}(t) = \begin{bmatrix} p_{11}(t) & p_{12}(t) & p_{13}(t) & \dots & p_{1r}(t) \\ p_{21}(t) & p_{22}(t) & p_{23}(t) & \dots & p_{2r}(t) \\ p_{31}(t) & p_{32}(t) & p_{33}(t) & & p_{3r}(t) \\ & \vdots & & \ddots & \vdots \\ p_{r1}(t) & p_{r2}(t) & p_{r3}(t) & \dots & p_{rr}(t) \end{bmatrix}$$

Random walk

- The random walk process as the discrete-time Markov chains.



- A **random walk** model appears in the context of many real-world problems, such as the motion of a particle, finance, and the dynamic change in network traffic.
- Let us imagine that we create a one-dimensional random process on the real line: we start at some initial position X_0 on the x-axis at time $t = 0$.
- At $t = 1$, we jump to position X_1, where the step size $S_1 = X_1 - X_0$ is a random variable with some distribution $f_X(s)$.

Random walk

- At time $t = 2$, we jump by another amount s_2, where s_2 is independent of s_1, but has the same distribution $f_X(s)$.

- The process continues and our position after n jumps, or at time $t = n$, is thus given by

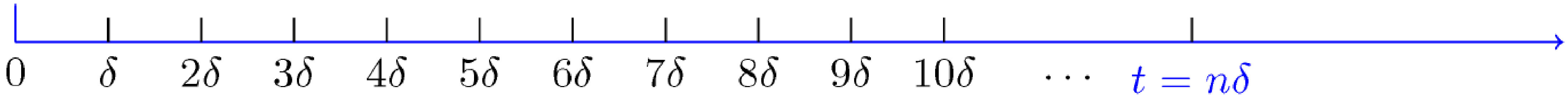
$$X_n = X_0 + s_1 + s_2 + \cdots + s_n = X_0 + \sum_{k=1}^{n} s_i$$

- where $\{s_i\}$ is a set of **i.i.d**. with the common distribution $f_X(s)$.
- This discrete-time random process $\{X_n\}$ is called a (one-dimensional) **random walk**.

Brownian Motion

- We are trying to construct a Brownian motion from a symmetric random walk.

- Divide the half-line $[0,\infty)$ to tiny subintervals of length δ as shown



Brownian Motion

- Each subinterval corresponds to a time slot of length δ. Thus, the intervals are $(0, \delta]$, $(\delta, 2\delta]$, $(2\delta, 3\delta]$, $\cdots$.

- More generally, the k-th interval is $((k-1)\delta, k\delta]$.

- We assume that in each time slot, as if we toss a fair coin.

Brownian Motion

- We define the random variables X_i as follows. $X_i = \sqrt{\delta}$ if the *k*-th coin toss results in heads, and $X_i = -\sqrt{\delta}$ if the *k*-th coin toss results in tails.

- Thus, we can write

$$X_i = \begin{cases} \sqrt{\delta}, with\ probability\ \dfrac{1}{2} \\ -\sqrt{\delta}, with\ probability\ \dfrac{1}{2} \end{cases}$$

Brownian Motion

- Moreover, the Xi's are independent. Note that

$$E[X_i] = 0; Var(X_i) = \delta$$

- Now, we would like to define the stochastic process $W(t)$ as follows. We let $W(0) = 0$.

- At time $t = n\delta$, the value of $W(t)$ is given by

$$W(t) = W(n\delta) = \sum_{i=1}^{n} X_i$$

Brownian Motion

- Since $W(t)$ is the sum of n i.i.d. random variables, we know how to find $E[W(t)]$ and $\mathrm{Var}(W(t))$.

- In particular

$$E[W(t)] = \sum_{i=1}^{n} E[X_i] = 0$$

$$Var\big(W(t)\big) = \sum_{i=1}^{n} Var(X_i) = nVar(X_1) = n\delta = t$$

Brownian Motion

- For any $t \in (0,\infty)$, as n goes to ∞, δ goes to 0. By the central limit theorem, $w(t)$ will become a normal random variable,

$$W(t) \sim N(0, t).$$

- $W(t)$ has *independente increments*. That is, for all

$$0 \leq t_1 < t_2 < t_3 \cdots < t_n,$$

the random variables

$$W(t_2) - W(t_1),\ W(t_3) - W(t_2),\ \cdots,\ W(t_n) - W(t_{n-1})$$

are independent.

Brownian Motion

- We conclude that the random process $W(t)$, defined above, has stationary increments.

- To see this, we argue as follows. For $0 \leq t_1 < t_2$, if we have $t_1 = n_1\delta$ and $t_2 = n_2\delta$, we obtain

$$W(t_1) = W(n_1\delta) = \sum_{i=1}^{n_1} X_i \qquad W(t_2) = W(n_2\delta) = \sum_{i=1}^{n_2} X_i$$

- Then we can write: $W(t_2) - W(t_1) = \sum_{i=n_1+1}^{n_2} X_i$

Brownian Motion

- We also can write

$$E[W(t_2) - W(t_1)] = \sum_{i=n_1+1}^{n_2} E[X_i] = 0$$

$$Var\big(W(t_2) - W(t_1)\big) = \sum_{i=n_1+1}^{n_2} Var(X_i)$$

$$= (n_2 - n_1)Var(X_1) = (n_2 - n_1)\delta = t_2 - t_1$$

Brownian Motion

- The random process $W(t)$ is called the standard *Brownian motion* or the standard *Wiener process*.

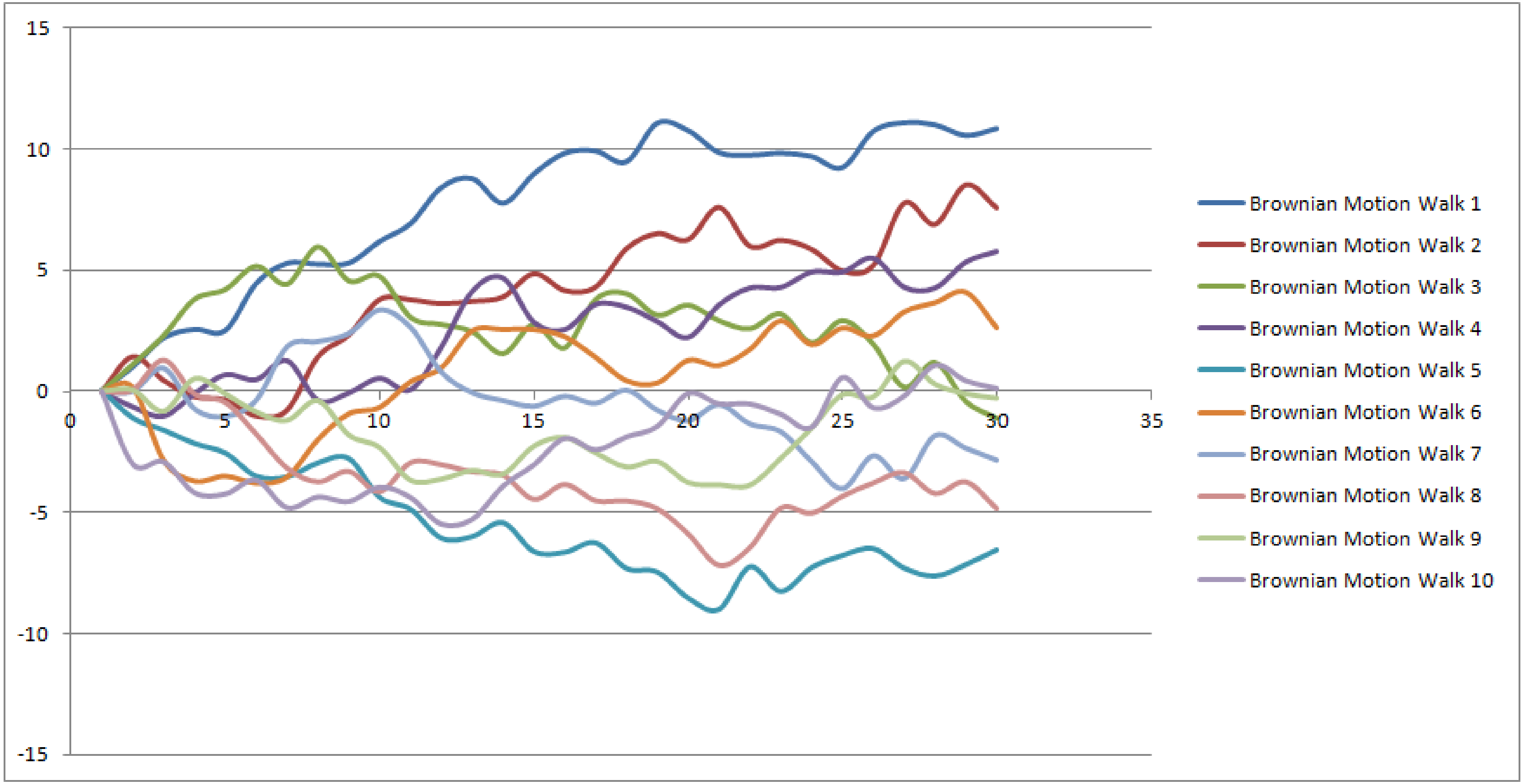
- A more general process is obtained if we define

$$X(t) = \mu + \sigma W(t)$$

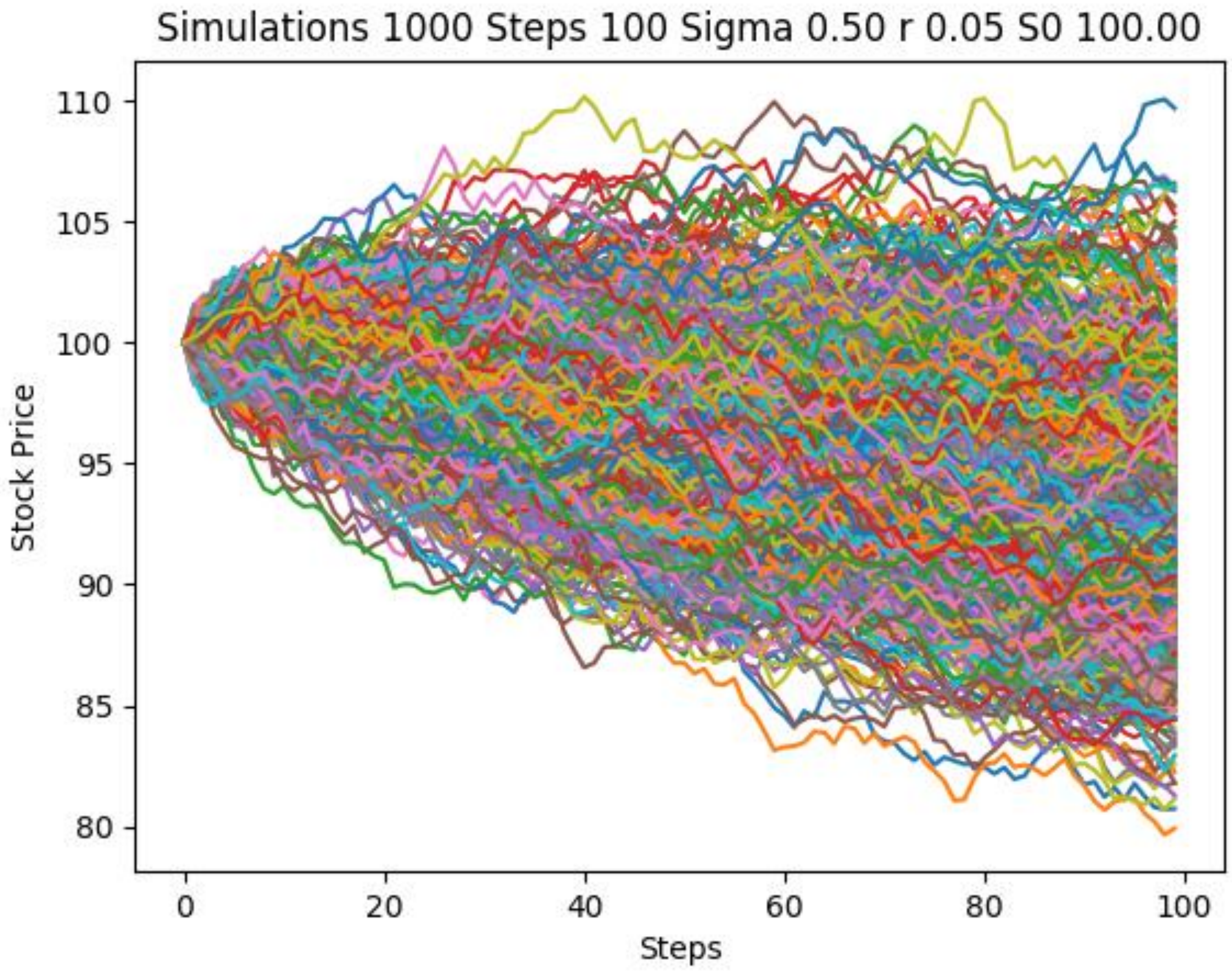
- In this case, $X(t)$ is a Brownian motion with

$$E[X(t)] = \mu \qquad Var\big(X(t)\big) = \sigma^2 t$$

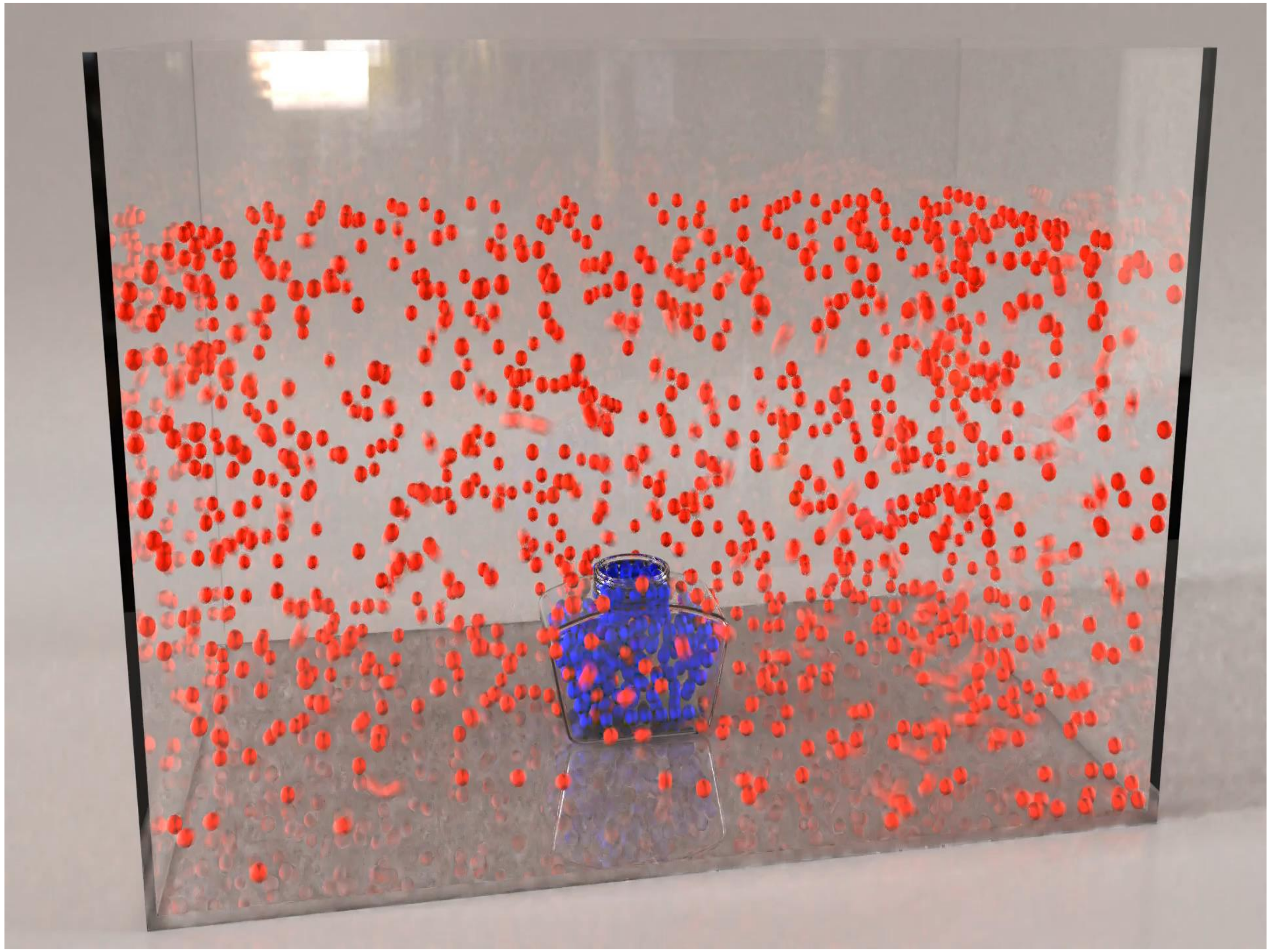
EXAMPLE:Brownian Motion



EXAMPLE:Brownian Motion



EXAMPLE:Brownian Motion




Fim do módulo 8