Random Processes

- A random process (also called stochastic process) \( \{ X(t) : t \in \mathcal{T} \} \) is an infinite collection of random variables, one for each value of time \( t \in \mathcal{T} \) (or, in some cases distance)

- Random processes are used to model random experiments that evolve in time:
  - Received sequence/waveform at the output of a communication channel
  - Packet arrival times at a node in a communication network
  - Thermal noise in a resistor
  - Scores of an NBA team in consecutive games
  - Daily price of a stock
  - Winnings or losses of a gambler
  - Earth movement around a fault line
Questions Involving Random Processes

- Dependencies of the random variables of the process:
  - How do future received values depend on past received values?
  - How do future prices of a stock depend on its past values?
  - How well do past earth movements predict an earthquake?

- Long term averages:
  - What is the proportion of time a queue is empty?
  - What is the average noise power generated by a resistor?

- Extreme or boundary events:
  - What is the probability that a link in a communication network is congested?
  - What is the probability that the maximum power in a power distribution line is exceeded?
  - What is the probability that a gambler will lose all his capital?

Discrete vs. Continuous-Time Processes

- The random process \( \{X(t) : t \in T\} \) is said to be \textit{discrete-time} if the index set \( T \) is countably infinite, e.g., \( \{1, 2, \ldots\} \) or \( \{\ldots, -2, -1, 0, +1, +2, \ldots\} \):
  - The process is simply an infinite sequence of r.v.s \( X_1, X_2, \ldots \)
  - An outcome of the process is simply a sequence of numbers

- The random process \( \{X(t) : t \in T\} \) is said to be \textit{continuous-time} if the index set \( T \) is a continuous set, e.g., \((0, \infty)\) or \((\infty, \infty)\)
  - The outcomes are random \textit{waveforms} or random occurrences in continuous time

- We only discuss discrete-time random processes:
  - IID processes
  - Bernoulli process and associated processes
  - Markov processes
  - Markov chains
IID Processes

• A process $X_1, X_2, \ldots$ is said to be *independent and identically distributed* (IID, or i.i.d.) if it consists of an infinite sequence of independent and identically distributed random variables.

• Two important examples:
  - Bernoulli process: $X_1, X_2, \ldots$ are i.i.d. $\text{Bern}(p)$, $0 < p < 1$, r.v.s. Model for random phenomena with binary outcomes, such as:
    - Sequence of coin flips
    - Noise sequence in a binary symmetric channel
    - The occurrence of random events such as packets (1 corresponding to an event and 0 to a non-event) in discrete-time
    - Binary expansion of a random number between 0 and 1
  - Discrete-time white Gaussian noise (WGN) process: $X_1, X_2, \ldots$ are i.i.d. $\mathcal{N}(0, N)$ r.v.s. Model for:
    - Receiver noise in a communication system
    - Fluctuations in a stock price

• Useful properties of an IID process:
  - *Independence*: Since the r.v.s in an IID process are independent, any two events defined on sets of random variables with *non-overlapping* indices are independent.
  - *Memorylessness*: The independence property implies that the IID process is memoryless in the sense that for any time $n$, the future $X_{n+1}, X_{n+2}, \ldots$ is independent of the past $X_1, X_2, \ldots, X_n$.
  - *Fresh start*: Starting from any time $n$, the random process $X_n, X_{n+1}, \ldots$ behaves identically to the process $X_1, X_2, \ldots$, i.e., it is also an IID process with the same distribution. This property follows from the fact that the r.v.s are identically distributed (in addition to being independent).
The Bernoulli Process

- The Bernoulli process is an infinite sequence $X_1, X_2, \ldots$ of i.i.d. $\text{Bern}(p)$ r.v.s
- The outcome from a Bernoulli process is an infinite sequence of 0s and 1s
- A Bernoulli process is often used to model occurrences of random events; $X_n = 1$ if an event occurs at time $n$, and 0, otherwise
- Three associated random processes of interest:
  - Binomial counting process: The number of events in the interval $[1, n]$  
  - Arrival time process: The time of event arrivals  
  - Interarrival time process: The time between consecutive event arrivals
- We discuss these processes and their relationships

Binomial Counting Process

- Consider a Bernoulli process $X_1, X_2, \ldots$ with parameter $p$
- We are often interested in the number of events occurring in some time interval
- For the time interval $[1, n]$, i.e., $i = 1, 2, \ldots, n$, we know that the number of occurrences  
  $$ W_n = \left( \sum_{i=1}^{n} X_i \right) \sim \text{B}(n, p) $$
- The sequence of r.v.s $W_1, W_2, \ldots$ is referred to as a Binomial counting process
- The Bernoulli process can be obtained from the Binomial counting process as:  
  $$ X_n = W_n - W_{n-1}, \text{ for } n = 1, 2, \ldots, $$
  where $W_0 = 0$
- Outcomes of a Binomial process are integer valued stair-case functions
• Note that the Binomial counting process is not IID

• By the fresh-start property of the Bernoulli process, for any \( n \geq 1 \) and \( k \geq 1 \), the distribution of the number of events in the interval \([k + 1, n + k]\) is identical to that of \([1, n]\), i.e., \( W_n \) and \((W_{k+n} - W_k)\) are identically distributed.

Example: Packet arrivals at a node in a communication network can be modeled by a Bernoulli process with \( p = 0.09 \).

1. What is the probability that 3 packets arrive in the interval \([1, 20]\), 6 packets arrive in \([1, 40]\) and 12 packets arrive in \([1, 80]\)?

2. The input queue at the node has a capacity of \(10^3\) packets. A packet is dropped if the queue is full. What is the probability that one or more packets are dropped in a time interval of length \( n = 10^4\)?

Solution: Let \( W_n \) be the number of packets arriving in interval \([1, n]\).

1. We want to find the following probability

\[
P\{W_{20} = 3, W_{40} = 6, W_{80} = 12\},
\]

which is equal to

\[
P\{W_{20} = 3, W_{40} - W_{20} = 3, W_{80} - W_{40} = 6\}
\]

By the independence property of the Bernoulli process this is equal to

\[
P\{W_{20} = 3\}P\{W_{40} - W_{20} = 3\}P\{W_{80} - W_{40} = 6\}
\]
Now, by the fresh start property of the Bernoulli process
\[
P{W_{40} - W_{20} = 3} = P{W_{20} = 3}, \text{ and } P{W_{80} - W_{40} = 6} = P{W_{40} = 6}
\]
Thus
\[
P{W_{20} = 3, W_{40} = 6, W_{80} = 12} = (P{W_{20} = 3})^2 \times P{W_{40} = 6}
\]
Now, using the Poisson approximation of Binomial, we have
\[
P{W_{20} = 3} = \binom{20}{3}(0.09)^3(0.91)^{17} \approx \frac{(1.8)^3}{3!} e^{-1.8} = 0.1607
\]
\[
P{W_{40} = 6} = \binom{40}{6}(0.09)^6(0.91)^{34} \approx \frac{(3.6)^6}{6!} e^{-3.6} = 0.0826
\]
Thus
\[
P{W_{20} = 3, W_{40} = 6, W_{80} = 12} \approx (0.1607)^2 \times 0.0826 = 0.0021
\]

2. The probability that one or more packets are dropped in a time interval of length \( n = 10^4 \) is
\[
P{W_{10^4} > 10^3} = \sum_{n=1001}^{10^4} \binom{10^4}{n}(0.09)^n(0.91)^{10^4-n}
\]
Difficult to compute, but we can use the CLT!
Since \( W_{10^4} = \sum_{i=1}^{10^4} X_i \) and \( E(X) = 0.09 \) and \( \sigma^2_X = 0.09 \times 0.91 = 0.0819 \),
we have
\[
P\left\{ \sum_{i=1}^{10^4} X_i > 10^3 \right\} = P\left\{ \frac{1}{100} \sum_{i=1}^{10^4} (X_i - 0.09) > \frac{10^3 - 900}{100 \sqrt{0.0819}} \right\}
\]
\[
= P\left\{ \frac{1}{100} \sum_{i=1}^{10^4} \frac{(X_i - 0.09)}{0.286} > 3.5 \right\}
\]
\[
\approx Q(3.5) = 2 \times 10^{-4}
\]
Arrival and Interarrival Time Processes

- Again consider a Bernoulli process $X_1, X_2, \ldots$ as a model for random arrivals of events
- Let $Y_k$ be the time index of the $k$th arrival, or the $k$th arrival time, i.e., smallest $n$ such that $W_n = k$
- Define the \textit{interarrival time} process associated with the Bernoulli process as $T_1 = Y_1$ and $T_k = Y_k - Y_{k-1}$, for $k = 2, 3, \ldots$

Thus the $k$th arrival time is given by: $Y_k = T_1 + T_2 + \ldots + T_k$

\[ \begin{array}{cccccccccccc}
0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & \cdots & \cdots & \cdots \vspace{5mm} \\
T_1 (Y_1) & T_2 & T_3 & T_4 & T_5 & \vspace{5mm} \\
Y_2 & & & & & \vspace{5mm} \\
Y_3 & & & & & \end{array} \]

- Let’s find the pmf of $T_k$:

First, the pmf of $T_1$ is the same as the number of coin flips until a head (i.e, a 1) appears. We know that this is $\text{Geom}(p)$. Thus $T_1 \sim \text{Geom}(p)$

Now, having an event at time $T_1$, the future is a fresh starting Bernoulli process. Thus, the number of trials $T_2$ until the next event has the same pmf as $T_1$

Moreover, $T_1$ and $T_2$ are independent, since the trials from 1 to $T_1$ are independent of the trials from $T_1 + 1$ onward. Since $T_2$ is determined exclusively by what happens in these future trials, it’s independent of $T_1$

Continuing similarly, we conclude that $T_1, T_2, \ldots$ are i.i.d., i.e., the interarrival process is an IID $\text{Geom}(p)$ process

- The interarrival process gives us an alternate definition of a Bernoulli process:

Start with an IID $\text{Geom}(p)$ process $T_1, T_2, \ldots$. Record the arrival of an event at time $T_1$, $T_1 + T_2$, $T_1 + T_2 + T_3$, \ldots
• Arrival time process: The sequence of r.v.s $Y_1, Y_2, \ldots$ is denoted by the arrival time process. From its relationship to the interarrival time process $Y_1 = T_1$, $Y_k = \sum_{i=1}^{k} T_i$, we can easily find the mean and variance of $Y_k$ for any $k$

$$E(Y_k) = E\left( \sum_{i=1}^{k} T_i \right) = \sum_{i=1}^{k} E(T_i) = k \times \frac{1}{p}$$

$$Var(Y_k) = Var\left( \sum_{i=1}^{k} T_i \right) = \sum_{i=1}^{k} Var(T_i) = k \times \frac{1-p}{p^2}$$

Note that, $Y_1, Y_2, \ldots$ is not an IID process

It is also not difficult to show that the pmf of $Y_k$ is

$$p_{Y_k}(n) = \binom{n-1}{k-1} p^k (1-p)^{n-k} \text{ for } n = k, k+1, k+2, \ldots,$$

which is called the Pascal pmf of order $k$

---

• Example: In each minute of a basketball game, Alicia commits a foul independently with probability $p$ and no foul with probability $1-p$. She stops playing if she commits her sixth foul or plays a total of 30 minutes. What is the pmf of of Alicia’s playing time?

Solution: We model the foul events as a Bernoulli process with parameter $p$

Let $Z$ be the time Alicia plays. Then

$$Z = \min\{Y_6, 30\}$$

The pmf of $Y_6$ is

$$p_{Y_6}(n) = \binom{n-1}{5} p^6 (1-p)^{n-6}, \text{ for } n = 6, 7, \ldots$$

Thus the pmf of $Z$ is

$$p_Z(z) = \begin{cases} 
\binom{z-1}{5} p^6 (1-p)^{z-6}, & \text{for } z = 6, 7, \ldots, 29 \\
1 - \sum_{z=6}^{29} p_Z(z), & \text{for } z = 30 \\
0, & \text{otherwise}
\end{cases}$$
Markov Processes

A discrete-time random process \( X_0, X_1, X_2, \ldots \), where the \( X_n \)'s are discrete-valued r.v.s, is said to be a Markov process if for all \( n \geq 0 \) and all \((x_0, x_1, x_2, \ldots, x_n, x_{n+1})\)

\[
P\{X_{n+1} = x_{n+1} | X_n = x_n, \ldots, X_0 = x_0\} = P\{X_{n+1} = x_{n+1} | X_n = x_n\},
\]

i.e., the past, \( X_{n-1}, \ldots, X_0 \), and the future, \( X_{n+1} \), are conditionally independent given the present \( X_n \).

A similar definition for continuous-valued Markov processes can be provided in terms of pdfs.

Examples:
- Any IID process is Markov
- The Binomial counting process is Markov

Markov Chains

A discrete-time Markov process \( X_0, X_1, X_2, \ldots \) is called a Markov chain if

- For all \( n \geq 0 \), \( X_n \in S \), where \( S \) is a finite set called the state space. We often assume that \( S \in \{1, 2, \ldots, m\} \)
- For \( n \geq 0 \) and \( i, j \in S \)

\[
P\{X_{n+1} = j | X_n = i\} = p_{ij}, \text{ independent of } n
\]

So, a Markov chain is specified by a transition probability matrix

\[
P = \begin{bmatrix}
p_{11} & p_{12} & \cdots & p_{1m} 
p_{21} & p_{22} & \cdots & p_{2m} 
\vdots & \vdots & \ddots & \vdots 
p_{m1} & p_{m2} & \cdots & p_{mm}
\end{bmatrix}
\]

Clearly \( \sum_{j=1}^{m} p_{ij} = 1 \), for all \( i \), i.e., the sum of any row is 1
• By the Markov property, for all $n \geq 0$ and all states

$$P\{X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0\} = P\{X_{n+1} = j | X_n = i\} = p_{ij}$$

• Markov chains arise in many real world applications:
  ○ Computer networks
  ○ Computer system reliability
  ○ Machine learning
  ○ Pattern recognition
  ○ Physics
  ○ Biology
  ○ Economics
  ○ Linguistics

**Examples**

• Any IID process with discrete and finite-valued r.v.s is a Markov chain

• *Binary Symmetric Markov Chain*: Consider a sequence of coin flips, where each flip has probability of $1 - p$ of having the same outcome as the previous coin flip, regardless of all previous flips

  The probability transition matrix (head is 1 and tail is 0) is

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

  A Markov chain can are be specified by a *transition probability graph*

  ![Transition Probability Graph]

  Nodes are states, arcs are state transitions; $(i, j)$ from state $i$ to state $j$ (only draw transitions with $p_{ij} > 0$)
Can construct this process from an IID process: Let $Z_1, Z_2, \ldots$ be a Bernoulli process with parameter $p$

The Binary symmetric Markov chain $X_n$ can be defined as

$$X_{n+1} = X_n + Z_n \mod 2,$$

for $n = 1, 2, \ldots$.

So, each transition corresponds to passing the r.v. $X_n$ through a binary symmetric channel with additive noise $Z_n \sim \text{Bern}(p)$

- Example: *Asymmetric binary Markov chain*
  
  State 0 = Machine is working, State 1 = Machine is broken down
  
  The probability transition matrix is

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

  and the transition probability graph is:

  ![Transition Probability Graph](image)

- Example: *Two spiders and a fly*
  
  A fly’s possible positions are represented by four states
  
  States 2, 3: safely flying in the left or right half of a room
  
  State 1: A spider’s web on the left wall
  
  State 4: A spider’s web on the right wall
  
  The probability transition matrix is:

  $$P = \begin{bmatrix} 1.0 & 0 & 0 & 0 \\ 0.3 & 0.4 & 0.3 & 0 \\ 0 & 0.3 & 0.4 & 0.3 \\ 0 & 0 & 0 & 1.0 \end{bmatrix}$$

  The transition probability graph is:

  ![Transition Probability Graph](image)
• Given a Markov chain model, we can compute the probability of a sequence of states given an initial state \(X_0 = i_0\) using the chain rule as:

\[
P\{X_1 = i_1, X_2 = i_2, \ldots, X_n = i_n | X_0 = i_0\} = p_{i_0i_1}p_{i_1i_2}\ldots p_{i_{n-1}i_n}
\]

Example: For the spider and fly example,

\[
P\{X_1 = 2, X_2 = 2, X_3 = 3 | X_0 = 2\} = p_{22}p_{22}p_{23}
\]

\[
= 0.4 \times 0.4 \times 0.3
\]

\[
= 0.048
\]

• There are many other questions of interest, including:
  
  ◦ \(n\)-state transition probabilities: Beginning from some state \(i\) what is the probability that in \(n\) steps we end up in state \(j\)
  
  ◦ Steady state probabilities: What is the expected fraction of time spent in state \(i\) as \(n \to \infty\) ?

\[n\text{-State Transition Probabilities}\]

• Consider an \(m\)-state Markov Chain. Define the \(n\)-step transition probabilities as

\[
r_{ij}(n) = P\{X_n = j | X_0 = i\} \text{ for } 1 \leq i, j \leq m
\]

• The \(n\)-step transition probabilities can be computed using the Chapman-Kolmogorov recursive equation:

\[
r_{ij}(n) = \sum_{k=1}^{m} r_{ik}(n-1)p_{kj} \text{ for } n > 1, \text{ and all } 1 \leq i, j \leq m,
\]

starting with \(r_{ij}(1) = p_{ij}\)

This can be readily verified using the law of total probability
We can view the $r_{ij}(n), 1 \leq i, j \leq m,$ as the elements of a matrix $R(n)$, called the $n$-step transition probability matrix, then we can view the Kolmogorov-Chapman equations as a sequence of matrix multiplications:

$$R(1) = \mathcal{P}$$
$$R(2) = R(1)\mathcal{P} = \mathcal{P}\mathcal{P} = \mathcal{P}^2$$
$$R(3) = R(2)\mathcal{P} = \mathcal{P}^3$$
$$\vdots$$
$$R(n) = R(n-1)\mathcal{P} = \mathcal{P}^n$$

Example: For the binary asymmetric Markov chain

$$R(1) = \mathcal{P} = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix}$$
$$R(2) = \mathcal{P}^2 = \begin{bmatrix} 0.76 & 0.24 \\ 0.72 & 0.28 \end{bmatrix}$$
$$R(3) = \mathcal{P}^3 = \begin{bmatrix} 0.752 & 0.248 \\ 0.744 & 0.256 \end{bmatrix}$$
$$R(4) = \mathcal{P}^4 = \begin{bmatrix} 0.7504 & 0.2496 \\ 0.7488 & 0.2512 \end{bmatrix}$$
$$R(5) = \mathcal{P}^5 = \begin{bmatrix} 0.7501 & 0.2499 \\ 0.7498 & 0.2502 \end{bmatrix}$$

In this example, each $r_{ij}$ seems to converge to a non-zero limit independent of the initial state, i.e., each state has a steady state probability of being occupied as $n \to \infty$. 
Example: Consider the spiders-and-fly example

\[
P = \begin{bmatrix} 1.0 & 0 & 0 & 0 \\ 0.3 & 0.4 & 0.3 & 0 \\ 0 & 0.3 & 0.4 & 0.3 \\ 0 & 0 & 0 & 1.0 \end{bmatrix}
\]

\[
P^2 = \begin{bmatrix} 1.0 & 0 & 0 & 0 \\ 0.42 & 0.25 & 0.24 & 0.09 \\ 0.09 & 0.24 & 0.25 & 0.42 \\ 0 & 0 & 0 & 1.0 \end{bmatrix}
\]

\[
P^3 = \begin{bmatrix} 1.0 & 0 & 0 & 0 \\ 0.50 & 0.17 & 0.17 & 0.16 \\ 0.16 & 0.17 & 0.17 & 0.50 \\ 0 & 0 & 0 & 1.0 \end{bmatrix}
\]

\[
P^4 = \begin{bmatrix} 1.0 & 0 & 0 & 0 \\ 0.55 & 0.12 & 0.12 & 0.21 \\ 0.21 & 0.12 & 0.12 & 0.55 \\ 0 & 0 & 0 & 1.0 \end{bmatrix}
\]

As \( n \to \infty \), we obtain

\[
P^\infty = \begin{bmatrix} 1.0 & 0 & 0 & 0 \\ 2/3 & 0 & 0 & 1/3 \\ 1/3 & 0 & 0 & 2/3 \\ 0 & 0 & 0 & 1.0 \end{bmatrix}
\]

Here \( r_{ij} \) converges, but the limit depends on the initial state and can be 0 for some states.

Note that states 1 and 4 corresponding to capturing the fly by one of the spiders are absorbing states, i.e., they are infinitely repeated once visited.

The probability of being in non-absorbing states 2 and 3 diminishes as time increases.
Classification of States

- As we have seen, various states of a Markov chain can have different characteristics.

- We wish to classify the states by the long-term frequency with which they are visited.

- Let $A(i)$ be the set of states that are accessible from state $i$ (may include $i$ itself), i.e., can be reached from $i$ in $n$ steps, for some $n$.

- State $i$ is said to be recurrent if starting from $i$, any accessible state $j$ must be such that $i$ is accessible from $j$, i.e., $j \in A(i)$ iff $i \in A(j)$. Clearly, this implies that if $i$ is recurrent then it must be in $A(i)$.

- A state is said to be transient if it is not recurrent.

- Note that recurrence/transience is determined by the arcs (transitions with nonzero probability), not by actual values of probabilities.

Example: Classify the states of the following Markov chain

- The set of accessible states $A(i)$ from some recurrent state $i$ is called a recurrent class.

- Every state $k$ in a recurrent class $A(i)$ is recurrent and $A(k) = A(i)$.

Proof: Suppose $i$ is recurrent and $k \in A(i)$. Then $k$ is accessible from $i$ and hence, since $i$ is recurrent, $k$ can access $i$ and hence, through $i$, any state in $A(i)$. Thus $A(i) \subseteq A(k)$.

Since $i$ can access $k$ and hence any state in $A(k)$, $A(k) \subseteq A(i)$. Thus $A(i) = A(k)$. This argument also proves that any $k \in A(i)$ is recurrent.
• Two recurrent classes are either identical or disjoint

• Summary:
  ○ A Markov chain can be decomposed into one or more recurrent classes plus possibly some transient states
  ○ A recurrent state is accessible from all states in its class, but it is not accessible from states in other recurrent classes
  ○ A transient state is not accessible from any recurrent state
  ○ At least one recurrent state must be accessible from a given transient state

• Example: Find the recurrent classes in the following Markov chains

![Diagram of Markov chains with states 1, 2, 3, and 4 connected in different patterns]
Periodic Classes

- A recurrent class $A$ is called **periodic** if its states can be grouped into $d > 1$ disjoint subsets $S_1, S_2, \ldots, S_d$, $\bigcup_{i=1}^{d} S_i = A$, such that all transitions from one subset lead to the next subset.

- Example: Consider a Markov chain with probability transition graph

Note that the recurrent class $\{1, 2\}$ is periodic and for $i = 1, 2$

$$r_{ii}(n) = \begin{cases} 
1 & \text{if } n \text{ is even} \\
0 & \text{otherwise}
\end{cases}$$

- Note that if the class is periodic, $r_{ii}(n)$ never converges to a steady-state.
Steady State Probabilities

- **Steady-state convergence theorem**: If a Markov chain has only one recurrent class and it is not periodic, then \( r_{ij}(n) \) tends to a steady-state \( \pi_j \) independent of \( i \), i.e.,
  \[
  \lim_{n \to \infty} r_{ij}(n) = \pi_j \text{ for all } i
  \]

- **Steady-state equations**: Taking the limit as \( n \to \infty \) of the Chapman-Kolmogorov equations
  \[
  r_{ij}(n+1) = \sum_{k=1}^{m} r_{ik}(n)p_{kj} \text{ for } 1 \leq i, j \leq m,
  \]
  we obtain the set of linear equations, called the *balance equations*:
  \[
  \pi_j = \sum_{k=1}^{m} \pi_k p_{kj} \text{ for } j = 1, 2, \ldots, m
  \]

  The balance equations together with the *normalization* equation
  \[
  \sum_{j=1}^{m} \pi_j = 1,
  \]
  uniquely determine the steady state probabilities \( \pi_1, \pi_2, \ldots, \pi_m \)

- The balance equations can be expressed in a matrix form as:
  \[
  \Pi P = \Pi, \text{ where } \Pi = \begin{bmatrix} \pi_1 & \pi_2 & \ldots & \pi_m \end{bmatrix}
  \]

- In general there are \( m - 1 \) linearly independent balance equations

- The steady state probabilities form a probability distribution over the state space, called the *stationary distribution* of the chain. If we set \( \text{P}\{X_0 = j\} = \pi_j \) for \( j = 1, 2, \ldots, m \), we have \( \text{P}\{X_n = j\} = \pi_j \) for all \( n \geq 1 \) and \( j = 1, 2, \ldots, m \)

- **Example: Binary Symmetric Markov Chain**
  \[
  P = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}
  \]

  Find the steady-state probabilities

  **Solution**: We need to solve the balance and normalization equations
  \[
  \pi_1 = p_{11} \pi_1 + p_{21} \pi_2 = (1-p) \pi_1 + p \pi_2
  \]
  \[
  \pi_2 = p_{12} \pi_1 + p_{22} \pi_2 = p \pi_1 + (1-p) \pi_2
  \]
  \[
  1 = \pi_1 + \pi_2
  \]
Note that the first two equations are linearly dependent. Both yield $\pi_1 = \pi_2$. Substituting in the last equation, we obtain $\pi_1 = \pi_2 = 1/2$

- Example: *Asymmetric binary Markov chain*

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

The steady state equations are:

$$\pi_1 = p_{11}\pi_1 + p_{21}\pi_2 = 0.8\pi_1 + 0.6\pi_2, \text{ and } 1 = \pi_1 + \pi_2$$

Solving the two equations yields $\pi_1 = 3/4$ and $\pi_2 = 1/4$

- Example: Consider the Markov chain defined by the following probability transition graph

```
1  2  3  4
0.9 0.1 0.5 0.8 0.2 0.4
```

- Note: The solution of the steady state equations yields $\pi_j = 0$ if a state is transient, and $\pi_j > 0$ if a state is recurrent
Long Term Frequency Interpretations

- Let $v_{ij}(n)$ be the # of times state $j$ is visited beginning from state $i$ in $n$ steps.
- For a Markov chain with a single aperiodic recurrent class, can show that
  \[
  \lim_{n \to \infty} \frac{v_{ij}(n)}{n} = \pi_j
  \]
  Since this result doesn’t depend on the starting state $i$, $\pi_j$ can be interpreted as the long-term frequency of visiting state $j$.
- Since each time state $j$ is visited, there is a probability $p_{jk}$ that the next transition is to state $k$, $\pi_j p_{jk}$ can be interpreted as the long-term frequency of transitions from $j$ to $k$.
- These frequency interpretations allow for a simple interpretation of the balance equations, that is, the long-term frequency $\pi_j$ is the sum of the long-term frequencies $\pi_k p_{kj}$ of transitions that lead to $j$
  \[
  \pi_j = \sum_{k=1}^{m} \pi_k p_{kj}
  \]

- Another interpretation of the balance equations: Rewrite the LHS of the balance equation as
  \[
  \pi_j = \pi_j \sum_{k=1}^{m} p_{jk} = \sum_{k=1}^{m} \pi_j p_{jk} = \pi_j p_{jj} + \sum_{k=1, k \neq j}^{m} \pi_j p_{jk}
  \]
  The RHS can be written as
  \[
  \sum_{k=1}^{m} \pi_k p_{kj} = \pi_j p_{jj} + \sum_{k=1, k \neq j}^{m} \pi_k p_{kj}
  \]
  Subtracting the $p_{jj} \pi_j$ from both sides yields
  \[
  \sum_{k=1, k \neq j}^{m} \pi_k p_{kj} = \sum_{k=1, k \neq j}^{m} \pi_j p_{jk}, \ j = 1, 2, \ldots, m
  \]
The long-term frequency of transitions into $j$ is equal to the long-term frequency of transitions out of $j$

This interpretation is similar to Kirkoff’s current law. In general, if we partition a chain (with single aperiodic recurrent class) into two sets of states, the long-term frequency of transitions from the first set to the second is equal to the long-term frequency of transitions from the second to the first.

**Birth-Death Processes**

- A *birth-death* process is a Markov chain in which the states are linearly arranged and transitions can only occur to a neighboring state, or else leave the state unchanged.

- For a birth-death Markov chain use the notation:
  
  $b_i = P\{X_{n+1} = i + 1 | X_n = i\}$, *birth* probability at state $i$
  
  $d_i = P\{X_{n+1} = i - 1 | X_n = i\}$, *death* probability at state $i$
For a birth-death process the balance equations can be greatly simplified: Cut the chain between states \( i - 1 \) and \( i \). The long-term frequency of transitions from right to left must be equal to the long-term frequency of transitions from left to right, thus:

\[
\pi_i d_i = \pi_{i-1} b_{i-1}, \text{ or } \pi_i = \pi_{i-1} \frac{b_{i-1}}{d_i}
\]

By recursive substitution, we obtain

\[
\pi_i = \pi_0 b_1 \ldots b_{i-1} \frac{d_1}{d_2 \ldots d_i}, \quad i = 1, 2, \ldots, m
\]

To obtain the steady state probabilities we use these equations together with the normalization equation

\[
1 = \sum_{j=0}^{m} \pi_j
\]

Examples

Queuing: Packets arrive at a communication node with buffer size \( m \) packets. Time is discretized in small periods. At each period:

- If the buffer has less than \( m \) packets, the probability of 1 packet added to it is \( b \), and if it has \( m \) packets, the probability of adding another packet is 0.
- If there is at least 1 packet in the buffer, the probability of 1 packet leaving it is \( d > b \), and if it has 0 packets, this probability is 0.
- If the number of packets in the buffer is from 1 to \( m - 1 \), the probability of no change in the state of the buffer is \( 1 - b - d \). If the buffer has no packets, the probability of no change in the state is \( 1 - b \), and if there are \( m \) packets in the buffer, this probability is \( 1 - d \).

We wish to find the long-term frequency of having \( i \) packets in the queue.
We introduce a birth-death Markov chain with states $0, 1, \ldots, m$, corresponding to the number of packets in the buffer.

The local balance equations are

$$\pi_i d = \pi_{i-1} b, \; i = 1, \ldots, m$$

Define $\rho = b/d < 1$, then $\pi_i = \rho \pi_{i-1}$, which leads to

$$\pi_i = \rho^i \pi_0$$

Using the normalizing equation: $\sum_{i=0}^{m} \pi_i = 1$, we obtain

$$\pi_0 (1 + \rho + \rho^2 + \cdots + \rho^m) = 1$$

Hence for $i = 1, \ldots, m$

$$\pi_i = \frac{\rho^i}{1 + \rho + \rho^2 + \cdots + \rho^m}$$

Using the geometric progression formula, we obtain

$$\pi_i = \rho^i \frac{1 - \rho}{1 - \rho^{m+1}}$$

Since $\rho < 1$, $\pi_i \to \rho^i (1 - \rho)$ as $m \to \infty$, i.e., $\{\pi_i\}$ converges to Geometric pmf

- **The Ehrenfest model**: This is a Markov chain arising in statistical physics. It models the diffusion through a membrane between two containers. Assume that the two containers have a total of $2a$ molecules. At each step a molecule is selected at random and moved to the other container (so a molecule diffuses at random through the membrane). Let $Y_n$ be the number of molecules in container 1 at time $n$ and $X_n = Y_n - a$. Then $X_n$ is a birth-death Markov chain with $2a + 1$ states; $i = -a, -a + 1, \ldots, -1, 0, 1, 2, \ldots, a$ and probability transitions

$$p_{ij} = \begin{cases} b_i = (a - i)/2a, & \text{if } j = i + 1 \\ d_i = (a + i)/2a, & \text{if } j = i - 1 \\ 0, & \text{otherwise} \end{cases}$$
The steady state probabilities are given by:

\[ \pi_i = \pi_{-a} \frac{b_{-a} b_{-a+1} \ldots b_{i-1}}{d_{-a+1} d_{-a+2} \ldots d_i}, \quad i = -a, -a + 1, \ldots, -1, 0, 1, 2, \ldots, a \]

\[ = \pi_{-a} \frac{2a (2a - 1) \ldots (a - i + 1)}{1 \times 2 \times \ldots \times (a + i)} \]

\[ = \pi_{-a} \frac{2a!}{(a + i)! (2a - (a + i))!} = \left( \frac{2a}{a + i} \right) \pi_{-a} \]

Now the normalization equation gives

\[ \sum_{i=-a}^{a} \left( \frac{2a}{a + i} \right) \pi_{-a} = 1 \]

Thus, \( \pi_{-a} = 2^{-2a} \)

Substituting, we obtain

\[ \pi_i = \left( \frac{2a}{a + i} \right) 2^{-2a} \]