MICT-5101: Probability and Stochastic Process[1]

Dr. Md. Rezaul Karim
PhD(KULeuven & UHasselt), MS(Biostatistics), MS(Statistics)
Professor, Department of Statistics and Data Science
Jahangirnagar University (JU), Savar, Dhaka - 1342, Bangladesh

MS-2024

[1]These course slides should not be reproduced nor used by others (without permission).

Lecture Outline I

Lecture Outline II

Introduction

1 Introduction

1.1 Text & Reference Book List

Text Book

1. Ross, S. (2010): *Introduction to Probability Models*, 10th edition, Pearson, Prentice Hall.
2. Anthony J. Hayter (2012): *Probability and Statistics for Engineers and Scientists* 4th Edition, Duxbury Press.

Reference Book List

1. Mehdhi, J. (2009): *Stochastic Processes* , 3rd Revised Edition, New Age Science.
2. Beichelt F. (2016): *Applied probability and stochastic processes*, 7th edition, CRC Press.
3. Ross, S. (2020): *Introduction to Probability and Statistics for Engineers and Scientists*, 6th Edition, Pearson Education Inc.

Fundamentals of Probability Models

1. **Part I:** Probability Theory
   - Basic Concepts of Probability
   - Random Variable
   - Expectation
   - Some Probability Distributions
     - Bernoulli
     - Binomial
     - Poisson
     - Uniform and
     - Normal
     - exponential
     - ...
2. **Part II:** Stochastic Processes
   - Basics of Stochastic Processes
   - Random Point Processes
   - Discrete-Time Markov Chains
   - ...

Chapter 1: Introduction to Stochastic Process

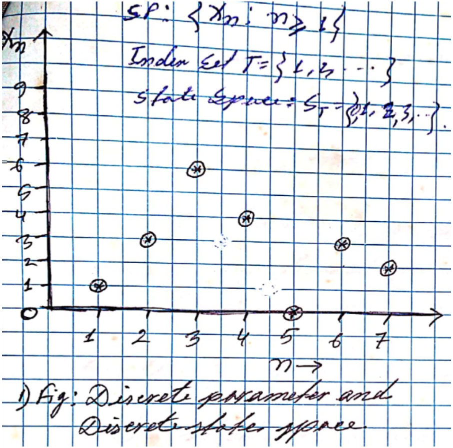
Stochastic Processes

- Stochastic Process: A stochastic process can be defined as a collection (or family) of random variables indexed by time or space.
- Examples:
  - Suppose $X(t)$ be represent the number of customers in the Supermarket at time t. Then $\{X(t); t \in \mathcal{T}\}$ is a family of random variables indexed by the time parameter t, hence the process $X(t)$ is a stochastic process.
  - The daily temperature can be treated as a stochastic process. While there are trends (like seasonal changes), the specific temperature on any given day is influenced by random atmospheric conditions.
  - The daily closing price of a stock can be modeled as a stochastic process. The price changes from day to day are influenced by a variety of random factors, such as market sentiment, news, and economic indicators.
- The indexing parameter can be either continuous or discrete. When the indexing parameter is discrete, generally we denote it by n and represent the stochastic process as $\{X_n; n = 0, 1, 2, \ldots\}$.

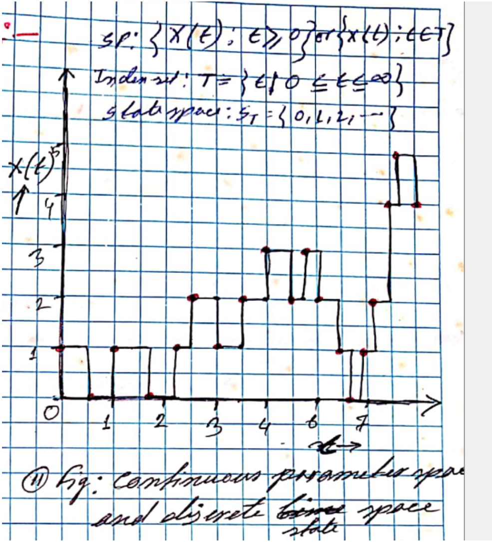
- index set:
  - when $\mathcal{T}$ is a countable set the stochastic process is said to be a discrete-time stochastic process
  - if $\mathcal{T}$ is an interval of the real line, the stochastic process is said to be a continuous-time stochastic process
- State Space: The set of possible values that the random variables $X(t)$ can take, which can be finite, countably infinite, or continuous.
- classification of stochastic process
  - discrete index set and state space
  - continuous index set and discrete state space
  - discrete index set and continuous state space
  - continuous index set and state space

Examples of Stochastic Process

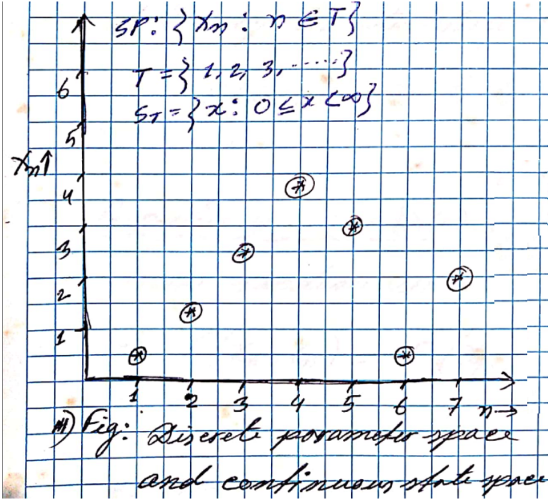
1. **discrete index set and state space**
   - number of the student getting scholarship every year in JU
   - consumer preferences observed on a monthly basis
   - number of the total customer in a supermarket everyday



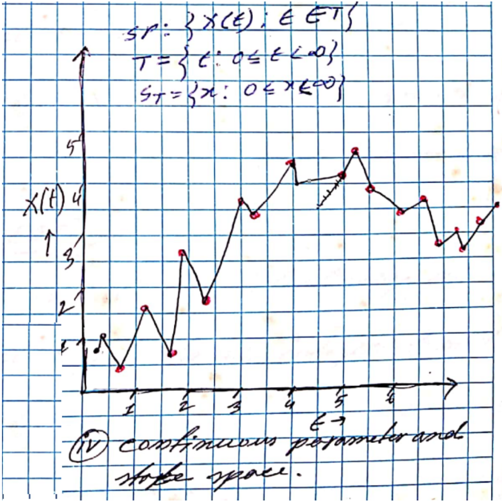
1. **continuous index set and discrete state space**
   - number of incoming calls at a switch board at any time of day
   - number of students waiting for the bus at any time of day



1. **discrete index set and continuous state space**
    - waiting time of the nth student arriving at the bus stop
    - maximum temperature at a fixed place every day



1. **continuous index set and state space**
   - maximum temperature in JU at any time of the day
   - content of a dam observed over an interval of time



Stationary Process

- A stationary process is a type of stochastic process whose statistical properties do not change over time. This means that its behavior and characteristics are consistent when observed at different time intervals. Stationarity is a crucial concept in time series analysis and probability theory.
  - Strict Stationarity
  - Weak Stationarity:

Strictly Stationary Process

- A process is strictly stationary if its joint distribution remains the same regardless of shifts in time. This means that all statistical properties, including all moments, are invariant to time shifts. That is, the stochastic process $\{X(t), t \in \mathcal{T}\}$ is said to be stationary in the strict sense if and only if

  $$F(x_1, x_2, \ldots, x_n; t_1, t_2, \ldots, t_n) = F(x_1, x_2, \ldots, x_n; t_1 + h, t_2 + h, \ldots, t_n + h)$$

  for all integer h and $n \geq 1$
- Example: For example, if you look at the returns from a stock over different time periods (say, from 2000 to 2005 and from 2010 to 2015) and find that they exhibit the same distribution characteristics, this would indicate a strictly stationary process

Weak Stationarity

- A process is weakly stationary (or second-order stationary) if its mean and variance are constant and the autocovariance only depends on the time difference (lag) h. That is, the stochastic process $\{X(t); t \in \mathcal{T}\}$ is said to be weakly stationary (or covariance stationary) if and only if
  - $E[X(t)] = \mu =$ Constant over time
  - $Var[X(t)] = \sigma^2 =$ Constant over time
  - $\gamma(h) = \gamma(t, t + h) =$ Constant over time
- **Monthly Average Temperature:** The monthly average temperature in a city can be modeled as a weakly stationary process.

Example of Weak Stationary: Monthly Average Temperature

- **Mean:** The average temperature over a month is relatively stable from year to year, assuming no significant climate changes.
- **Variance:** The variability of temperatures within each month is also consistent over time. For instance, January may consistently have a certain range of temperatures, and this range doesn't change drastically from year to year.
- **Covariance:** The covariance between the average temperatures of two different months depends only on the time difference between them (e.g., the correlation between January and February would be similar regardless of the specific years being analyzed).

Evolutionary

If the process is not stationary (in any sense), then it is called evolutionary.

Problem 1

Suppose $[X(t), t \in T]$ be a stochastic process where

$$\Pr[X(t) = n] = \frac{e^{-at}(at)^n}{n!} \quad ; \quad n = 0, 1, 2, \ldots, a > 0.$$

Is this process stationary?

Solution

We know that the mean function

$$
\begin{aligned}
m(t) = E[X(t)] &= \sum_{n=0}^{\infty} n \frac{e^{-at}(at)^n}{n!} \\
&= at \sum_{n=1}^{\infty} \frac{e^{-at}(at)^{n-1}}{(n-1)!} \\
&= at\, e^{-at} \sum_{n=1}^{\infty} \frac{(at)^{n-1}}{(n-1)!} \qquad = at\, e^{-at}\, e^{at} \qquad = at
\end{aligned}
$$

and

$$
\begin{aligned}
E\left[\{X(t)\}^2\right] &= \sum_{n=0}^{\infty} n^2 \frac{e^{-at}(at)^n}{n!} = \sum_{n=0}^{\infty} \{n(n-1)+n\} \frac{e^{-at}(at)^n}{n!} \\
&= \sum_{n=0}^{\infty} n(n-1) \frac{e^{-at}(at)^n}{n!} + \sum_{n=0}^{\infty} n \frac{e^{-at}(at)^n}{n!} \\
&= (at)^2\; e^{-at} \sum_{n=2}^{\infty} \frac{(at)^{n-2}}{(n-2)!} + at\, e^{-at} \sum_{n=1}^{\infty} \frac{(at)^{n-1}}{(n-1)!} \\
&= (at)^2\; e^{-at}\, e^{at} + at\, e^{-at}\, e^{at} \\
&= (at)^2 + at
\end{aligned}
$$

$$
\begin{aligned}
\therefore V[X(t)] &= E\left[\{X(t)\}^2\right] - [E\{X(t)\}]^2 \\
&= (at)^2 + at - (at)^2 = at
\end{aligned}
$$

Comment: Since the mean and variance functions of this stochastic process are dependent on time t, so the stochastic process $[X(t), t \in T]$ is evolutionary process.

Problem 2

Consider the process $X(t) = A_1 + A_2 t$ where A_1 *and* A_2 are independent random variables with $E(A_i) = a_i$, $V(A_i) = \sigma_i^2$; $i = 1, 2$. Test the stationarity.

Solution:

We know that the mean function is

$$m(t) = E[X(t)] = E[A_1 + A_2 t] = a_1 + a_2 t$$

$$\begin{aligned} E[X(t)X(s)] &= E[(A_1 + A_2 t)(A_1 + A_2 s)] \\ &= E\left[A_1^2 + A_2^2 ts + (t+s) A_1 A_2\right] \\ &= \sigma_1^2 + a_1^2 + ts\left(\sigma_2^2 + a_2^2\right) + (t+s) a_1 a_2 \end{aligned}$$

$$\text{Sicne } E\left(A_1^2\right) = [E(A_1)]^2 + V(A_1) = a_1^2 + \sigma_1^2$$

$$\begin{aligned} E[X(t)]^2 &= E\left[(A_1 + A_2 t)^2\right] \\ &= E\left[A_1^2 + A_2^2 t^2 + 2A_1 A_2 t\right] \\ &= E\left[A_1^2\right] + t^2 E\left[A_2^2\right] + 2t\, E[A_1 A_2] \\ &= \sigma_1^2 + a_1^2 + t^2\left(\sigma_2^2 + a_2^2\right) + 2t a_1 a_2 \end{aligned}$$

$$
\begin{aligned}
V\left[X(t)\right] &= E\left[X(t)\right]^2 - \left[E\left[X(t)\right]\right]^2 \\
&= \sigma_1^2 + a_1^2 + t^2\left(\sigma_2^2 + a_2^2\right) + 2ta_1a_2 - \left(a_1 + a_2t\right)^2 \\
&= \sigma_1^2 + a_1^2 + t^2\sigma_2^2 + t^2a_2^2 + 2ta_1a_2 - a_1^2 - 2ta_1a_2 - t^2a_2^2 \\
&= \sigma_1^2 + t^2a_2^2
\end{aligned}
$$

$$
\begin{aligned}
C(s,t) &= cov\left[X(s), X(t)\right] \\
&= E\left[X(s), X(t)\right] - E\left[X(s)\right]E\left[X(t)\right] \\
&= \sigma_1^2 + a_1^2 + (s+t)a_1a_2 + ts\left(\sigma_2^2 + a_2^2\right) - \left(a_1 + a_2s\right)\left(a_1 + a_2t\right) \\
&= \sigma_1^2 + a_1^2 + sa_1a_2 + ta_1a_2 + ts\sigma_2^2 + tsa_2^2 - a_1^2 - sa_1a_2 - ta_1a_2 - tsa_2^2 \\
&= \sigma_1^2 + ts\sigma_2^2
\end{aligned}
$$

Comment: Since $m(t) = a_1 + a_2t$ and $V\left[X(t)\right] = \sigma_1^2 + t^2\sigma_2^2$ are functions of t, so that the process is evolutionary.

Problem 3

Consider the process $X(t) = A\cos\omega t + B\sin\omega t$, where A, B uncorrelated random variables each with are mean 0 and variance 1 and ω is a positive constant. Show that the process is covariance stationary.

Solution:

We have the mean function

$$\begin{aligned} M(t) = E[X(t)] &= \cos\omega t E(A) + \sin\omega t E(B) \\ &= 0 \end{aligned}$$

$$\begin{aligned} &E[X(t)X(s)] = E[(A\cos\omega t + B\sin\omega t)(A\cos\omega s + B\sin\omega s)] \\ &= E\left[A^2\cos\omega t\cos\omega s + AB\cos\omega t\sin\omega s + AB\sin\omega t\cos\omega s + B^2\sin\omega t\sin\omega s\right] \\ &= \cos\omega t\cos\omega s + \sin\omega t\sin\omega s \\ &= \cos(s-t)\omega \end{aligned}$$

$$\begin{aligned} \text{since, } E\left[\{X(t)\}^2\right] &= Var[X(t)] + [E\{X(t)\}]^2 \\ &= 1 \end{aligned}$$

and,

$$\begin{aligned} C(s,t) &= Cov\left\{X(t), X(s)\right\} \\ &= \cos(s-t)\,\omega \end{aligned}$$

Comment: Here the first two moments are finite and the covariance function is a function of $(s-t)$. Thus the process is covariance stationary.

Problem 4

Consider the process $[X(t), t \in T]$ whose probability distribution, under a certain condition, is given by

$$\Pr[X(t) = n] = \frac{(at)^{n-1}}{(1+at)^{n+1}}, \quad n = 1, 2, \dots \dots$$
$$= \frac{at}{1+at}, \quad n = 0$$

Test the stationarity.

Solution:

We have the mean function

$$m(t) = E[X(t)] = \sum_{n=0}^{\infty} n \Pr[X(t) = n] = \sum_{n=1}^{\infty} n \frac{(at)^{n-1}}{(1+at)^{n+1}}$$
$$= \frac{1}{(1+at)^2} \sum_{n=1}^{\infty} n \left(\frac{at}{1+at}\right)^{n-1}$$
$$= \frac{1}{(1+at)^2} \left[1 + 2\frac{at}{1+at} + 3\left(\frac{at}{1+at}\right)^2 + \dots \dots \dots\right]$$
$$= \frac{1}{(1+at)^2} \left[1 - \frac{at}{1+at}\right]^{-2} = \frac{1}{(1+at)^2} \left[\frac{1}{1+at}\right]^{-2} = 1$$

and,

$$E\left\{[X(t)]^2\right\} = \sum_{n=0}^{\infty} n^2 \ \Pr[X(t) = n] = \sum_{n=1}^{\infty} n^2 \ \frac{(at)^{n-1}}{(1+at)^{n+1}}$$
$$= \sum_{n=1}^{\infty} n(n-1) \ \frac{(at)^{n-1}}{(1+at)^{n+1}} + \sum_{n=1}^{\infty} n \frac{(at)^{n-1}}{(1+at)^{n+1}}$$
$$= \frac{at}{(1+at)^3} \sum_{n=2}^{\infty} n(n-1) \left(\frac{at}{1+at}\right)^{n-2} + 1$$
$$= \frac{at}{(1+at)^3} \left[2 + 3 \cdot 2 \frac{at}{1+at} + 4 \cdot 3 \left(\frac{at}{1+at}\right)^2 + 5 \cdot 4 \left(\frac{at}{1+at}\right)^3 + \ ..\right.$$
$$= \frac{2at}{(1+at)^3} \left[1 + 3 \frac{at}{1+at} + 6 \left(\frac{at}{1+at}\right)^2 + 10 \left(\frac{at}{1+at}\right)^3 + \ \dots \ \dots \ \dots\right.$$
$$= \frac{2at}{(1+at)^3} \left[1 - \frac{at}{1+at}\right]^{-3} + 1$$
$$= \frac{2at}{(1+at)^3} \left[\frac{1}{1+at}\right]^{-3} + 1 = 2at + 1$$

$$\therefore \ Var[X(t)] = 2at$$

Comment: Here the first moment is constant but the second moment (and the variance) increases with t. Thus the process $[X(t), \ t \in T]$ is evolutionary.

What is a Gaussian Process?

- A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian (normal) distribution. That is, if the distribution of $[X(t_1), X(t_2), \ldots, X(t_n)]$ for all $t_1, t_2, \ldots, t_n$ is multivariate normal then $\{X(t); t \in T\}$ is said to be a Gaussian Process.
- Gaussian processes are often used in machine learning and statistics for regression, classification, and spatial modeling.

Martingales Process:

A martingale is a specific type of stochastic process that satisfies certain properties related to conditional expectations. That is, a stochastic process $\{X_n\}_{n=1}^{\infty}$ is called a martingale (or a Martingale process) if,

(i) the expected value of the absolute value of X_n is finite. That is

$$E\left[|X_n|\right] < \infty$$

(ii) the expected value of the next observation, given all previous observations, is equal to the current observation. That is,

$$E\left[X_{n+1}|X_n,\ X_{n-1},\ \cdots,\ X_0\right] = X_n$$

Example:

Let $\{Z_i\}$; $i = 1,\ 2,\ \cdots$ be a sequence of $i.i.d.$ random variables with mean 0 and let $X_n = \sum_{i=1}^{n} Z_i$, then show that $\{X_n\}_{n=1}^{\infty}$ is a martingale.

Solution: We have,

$$E\left(X_n\right) = E\left\{\sum_{i=1}^{n} Z_i\right\} = \sum_{i=1}^{n} E\left\{Z_i\right\} = 0$$

Since, Z_i's are i.i.d. and with mean 0. We have,

$$\begin{aligned} & X_n = Z_1 + Z_2 + \cdots + Z_n \\ \Rightarrow \quad & X_{n+1} = Z_1 + Z_2 + \cdots + Z_n + Z_{n+1} = X_n + Z_{n+1} \end{aligned}$$

So that,

$$\begin{aligned} & E\left\{X_{n+1} | X_n,\ X_{n-1},\ \cdots,\ X_1\right\} = E\left\{\left(X_n + Z_{n+1}\right) | X_n,\ X_{n-1},\ \cdots,\ X_1\right\} \\ & = E\left\{X_n | X_n,\ X_{n-1},\ \cdots,\ X_1\right\} + E\left\{Z_{n+1} | X_n,\ X_{n-1},\ \cdots,\ X_1\right\} \\ & = X_n + E\left\{Z_{n+1}\right\} \text{ ; Since } Z_i \text{ are independent and } E\left\{Z_{n+1}\right\} = 0 \\ & = X_n \end{aligned}$$

So, the process is martingale process because it satisfy the two condition of martingale process. (*Shown*)

Example:

Let $\{Z_i\}$; $i = 1, 2, \cdots$ be a sequence of $i.i.d.$ random variables with $E\{Z_i\} = 1$ and let $X_n = \prod_{i=1}^{n} Z_i$, then show that $\{X_n\}_{n=1}^{\infty}$ is a martingale.

Solution:

We have,

$$\begin{aligned} E(X_n) &= E\{Z_1 \cdot Z_2 \cdot \cdots \cdot Z_n\} \\ &= E\{Z_1\} \cdot E\{Z_2\} \cdot \cdots \cdot E\{Z_n\} \quad ; \quad \textit{Since they are independent} \\ &= 1 \cdot 1 \cdot \cdots \cdot 1 = 1 < \infty \end{aligned}$$

Again We have,

$$\begin{aligned} & X_n = Z_1 \cdot Z_2 \cdot \cdots \cdot Z_n \\ \Rightarrow \quad & X_{n+1} = Z_1 \cdot Z_2 \cdot \cdots \cdot Z_n \cdot Z_{n+1} \\ & \quad = X_n \cdot Z_{n+1} \end{aligned}$$

So that,

$$\begin{aligned} E\{X_{n+1}|X_n, X_{n-1}, \cdots, X_1\} &= E\{(X_n \cdot Z_{n+1})|X_n, X_{n-1}, \cdots, X_1\} \\ &= E\{X_n|X_n, X_{n-1}, \cdots, X_1\} \cdot E\{Z_{n+1}|X_n, X_{n-1}, \cdots, X_1\} \\ &= X_n \cdot E\{Z_{n+1}\} \quad ; \quad \textit{Since } Z_{i+1} \textit{ and } X_n, X_{n-1}, \cdots, X_1 \\ &= X_n \quad ; \quad \textit{Since } E\{Z_{n+1}\} = 1 \end{aligned}$$

Chapter 2: Discrete Markov Chain

3 Chapter 2: Discrete Markov Chain

Markov Chain

A stochastic process is said to be **Markov chain** if the the conditional probability distribution of future states depends only on the present state, not on past states. A discrete-time stochastic process $\{X_n\}_{n\geq 0}$ is said to be discrete Markov chain if the conditional distribution of any future state X_{n+1}, give the past states $X_0, X_1, \ldots, X_{n-1}$ and the present state X_n, is independent of the past states and depends on only on the current state. That is

$$P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1, X_0 = i_0) = P(X_{n+1} = j \mid X_n = i)$$

for all $n \geq 0$ and all $i_0, i_1, \ldots, i_{n-1}, i, j \in S$,

Time-Homogeneous Markov Chain

A time-homogeneous Markov chain is a type of Markov chain in which the transition probabilities are constant over time. This means that the probability of transitioning from one state to another depends only on the current state and the next state, not on the specific time at which the transition occurs. For a time-homogeneous Markov chain, the transition probabilities can be expressed as:

$$P(X_{n+1} = j \mid X_n = i) = p_{ij} \tag{1}$$

where p_{ij} is the probability of transitioning from state i to state j. This probability is the same for all n.

Transition Probability

A transition probability refers to the probability of moving from one state to another in a Markov chain. It quantifies how likely it is for a system to transition from a current state i to a next state j in a given time step. The transition probability from state i to state j is denoted as:

$$P_{ij} = P(X_{n+1} = j \mid X_n = i)$$

Properties:

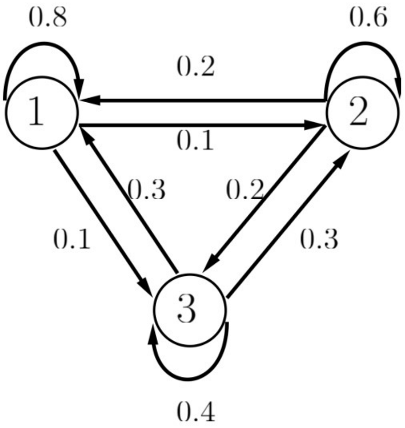
- Non-negativity: $P_{ij} \geq 0$ for all states i and j.
- Normalization:

$$\sum_{j \in S} P_{ij} = 1$$

  This ensures that from any state i, the system will transition to some state j with certainty.

Example 0.2 (Social mobility). Let X_n be a family's social class: 1 (lower), 2 (middle), 3 (upper) in the nth generation. We can model this process as a Markov chain with certain kind of transition probabilities such as

$$\mathbf{P} = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.2 & 0.6 & 0.2 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}$$



Example: Weather Model

Consider a simple weather model with three states:

- Sunny
- Cloudy
- Rainy

The transition probabilities can be represented in the following matrix P:

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

where:

- Row 1 represents transitions from Sunny
- Row 2 represents transitions from Cloudy
- Row 3 represents transitions from Rainy

This system is a time-homogeneous Markov chain because these transition probabilities do not change over time.

Understanding the Transition Matrix

Each entry P_{ij} in the matrix represents the probability of transitioning from state i to state j:

- $P_{11} = 0.6$: Probability of staying Sunny.
- $P_{12} = 0.3$: Probability of transitioning from Sunny to Cloudy.
- $P_{13} = 0.1$: Probability of transitioning from Sunny to Rainy.

(Similar interpretations apply for the other rows.)

Forecasting Example

Suppose today is Sunny. We can use the transition matrix to forecast the weather for the next day:

- Probability of Sunny tomorrow:

$$P_{11} = 0.6$$

- Probability of Cloudy tomorrow:

$$P_{12} = 0.3$$

- Probability of Rainy tomorrow:

$$P_{13} = 0.1$$

Thus, the forecast for tomorrow would be:

- Sunny: 60%
- Cloudy: 30%
- Rainy: 10%

Multiple Days Forecast

To forecast multiple days ahead, you can raise the transition matrix to a power n corresponding to the number of days:

$$P^n$$

For example, to forecast the weather for two days ahead, calculate P^2 to see the probabilities for the states after two transitions.

Time-Homogeneous vs. Time-Inhomogeneous

- Time-Homogeneous: Transition probabilities do not change over time (constant P_{ij}).
- Time-Inhomogeneous: Transition probabilities can vary with time and may depend on the time step n.

Time-Homogeneous Markov Chain

Example: Weather Model

Consider a weather model with three states: Sunny (S), Cloudy (C), and Rainy (R).

- Transition Probability Matrix:

$$P = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.4 & 0.2 \\ 0.2 & 0.5 & 0.3 \end{pmatrix}$$

- Interpretation:
  - If today is Sunny, there is a 60% chance it will be Sunny tomorrow.
  - If today is Cloudy, there is a 40% chance it will be Sunny tomorrow.
  - These probabilities remain constant over time.

Time-Nonhomogeneous Markov Chain

Example: Weather Model Consider the same weather model but with changing probabilities over time.

- **Day 1 Transition Matrix**:

$$P^{(1)} = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.4 & 0.2 \\ 0.2 & 0.5 & 0.3 \end{pmatrix}$$

- **Day 2 Transition Matrix**:

$$P^{(2)} = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.1 & 0.4 & 0.5 \end{pmatrix}$$

- Interpretation:
  - On Day 1, if it is Sunny, there is a 60% chance it will be Sunny again.
  - On Day 2, that chance increases to 70%.
  - Transition probabilities change based on the day.

Example 4.1: Forecasting the Weather (See textbook)

Suppose that the chance of rain tomorrow depends on previous weather conditions only through whether or not it is raining today, and not on past weather conditions.

- If it rains today, then it will rain tomorrow with probability α; and if it does not rain today, then it will rain tomorrow with probability β.
- If we say that the process is in state 0 when it rains and state 1 when it does not rain, then the preceding is a two-state Markov chain whose transition probabilities are given by:

$$P = \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & \alpha & 1-\alpha \\ 1 & \beta & 1-\beta \\ \hline \end{array}$$

Example 4.3 On any given day Gary is either cheerful (C), so-so (S), or glum (G). If he is cheerful today, then he will be C, S, or G tomorrow with respective probabilities 0.5, 0.4, 0.1. If he is feeling so-so today, then he will be C, S, or G tomorrow with probabilities 0.3, 0.4, 0.3. If he is glum today, then he will be C, S, or G tomorrow with probabilities 0.2, 0.3, 0.5.

Letting X_n denote Gary's mood on the nth day, then $\{X_n, n \geqslant 0\}$ is a three-state Markov chain (state $0 = C$, state $1 = S$, state $2 = G$) with transition probability matrix

$$\mathbf{P} = \begin{Vmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{Vmatrix} \qquad \blacksquare$$

Example 4.4 (Transforming a Process into a Markov Chain) Suppose that whether or not it rains today depends on previous weather conditions through the last two days. Specifically, suppose that if it has rained for the past two days, then it will rain tomorrow with probability 0.7; if it rained today but not yesterday, then it will rain tomorrow with probability 0.5; if it rained yesterday but not today, then it will rain tomorrow with probability 0.4; if it has not rained in the past two days, then it will rain tomorrow with probability 0.2.

If we let the state at time n depend only on whether or not it is raining at time n, then the preceding model is not a Markov chain (why not?). However, we can transform this model into a Markov chain by saying that the state at any time is determined by the weather conditions during both that day and the previous day. In other words, we can say that the process is in

state 0	if it rained both today and yesterday,
state 1	if it rained today but not yesterday,
state 2	if it rained yesterday but not today,
state 3	if it did not rain either yesterday or today.

The preceding would then represent a four-state Markov chain having a transition probability matrix

The Weather Prediction Problem

How to model this problem as a Markov Process ?

The state space: 0 = (RR) 1 = (NR) 2 = (RN) 3 = (NN)

The transition matrix:

$$
\mathbf{P} = \begin{array}{c} \text{YT} \\ 0\ (\text{RR}) \\ 1\ (\text{NR}) \\ 2\ (\text{RN}) \\ 3\ (\text{NN}) \end{array}
\begin{array}{c} \begin{array}{cccc} \text{Tt} & & & \\ 0(\text{RR}) & 1(\text{NR}) & 2(\text{RN}) & 3(\text{NN}) \end{array} \\
\boxed{\begin{array}{cccc} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{array}} \end{array}
$$

This is a discrete-time Markov process.

Random Walk

- A **random walk** is a stochastic process involving a sequence of steps in a discrete state space, where each step is determined by random choices. It can be defined as a Markov chain where the next state depends only on the current state. For a state i, the probabilities are given by:

$$P(X_{n+1} = i + 1 \mid X_n = i) = p \quad \text{(move right)}$$

$$P(X_{n+1} = i - 1 \mid X_n = i) = 1 - p \quad \text{(move left)}.$$

  Each step is independent, and the process usually starts from an initial state, such as $X_0 = 0$. Random walks are fundamental in various fields, modeling phenomena like diffusion and market fluctuations.
- It is called a symmetric random walk if $p = 0.5$.
- Imagine a person taking steps either forward or backward with equal probability. The position of the person after a number of steps is a classic example of a random walk.

Gambler's Ruin Problem

In the **Gambler's Ruin** problem, a gambler starts with A dollars and plays a series of bets. The outcomes of each bet are determined by random choices. The gambler can either win or lose money in each round.

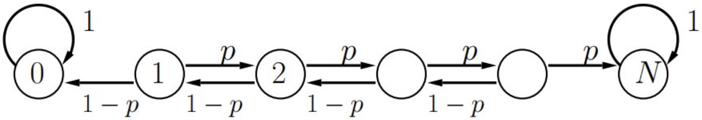
- **Initial State**: The gambler starts with A dollars.
- **Target State**: The gambler aims to reach $N = 5$ dollars.
- **States**: The states of the process are $0, 1, 2, 3, 4, 5$:
  - 0: Absorbing state representing ruin.
  - 5: Absorbing state representing winning.

Transition Probabilities: At each step, the transition probabilities are:

$$P(X_{n+1} = i + 1 \mid X_n = i) = p \quad \text{(win)}$$
$$P(X_{n+1} = i - 1 \mid X_n = i) = 1 - p \quad \text{(lose)}$$

Example 0.1 (Gambler's Ruin). This can be modeled as a Markov chain with state space $S = \{0, 1, 2, \ldots, N\}$ and transition probabilities

$$P_{i,i+1} = p = 1 - P_{i,i-1}, \quad 1 \leq i \leq N-1$$
$$P_{00} = 1 = P_{NN} \quad \text{(absorbing states)}$$



It is a random walk on a finite state space and with two absorbing barriers.

Transition Matrix

For $N = 5$, the transition matrix P can be represented as follows:

$$P = \begin{array}{cc} & \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ p & 0 & 1-p & 0 & 0 & 0 \\ 0 & p & 0 & 1-p & 0 & 0 \\ 0 & 0 & p & 0 & 1-p & 0 \\ 0 & 0 & 0 & p & 0 & 1-p \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Where the rows and columns represent the states $0, 1, 2, 3, 4, 5$ respectively.

Ehrenfest Diffusion Model

- Imagine two urns (or boxes), A and B, that can hold a certain number of balls (particles).
- Initially, some balls are placed in one of the urns.
- Let K be the total number of balls, and let X_n be the number of balls in urn A at step n.

Transition:

- At each step, one ball is chosen uniformly at random and moved to the other urn.
- This process continues indefinitely.

States:

- The state space consists of the possible configurations of the balls in the urns.
- For example, if there are 4 balls, the possible states for X_n (number of balls in urn A) range from 0 to 4.

Transition Probabilities of Ehrenfest Diffusion Model

If there are i balls in urn A (i.e., $X_n = i$), the transition probabilities are:

$$P_{i,i+1} = P(X_{n+1} = i + 1 | X_n = i) = \frac{K - i}{K}$$

(probability of moving a ball from urn B to urn A)

$$P_{i,i-1} = P(X_{n+1} = i - 1 | X_n = i) = \frac{i}{K}$$

(probability of moving a ball from urn A to urn B)

$$P_{ii} = P(X_{n+1} = i | X_n = i) = 0$$

(no self-transition)

Transition Probability Matrix

For example, if there are 4 balls, the possible states for X_n (number of balls in urn A) range from 0 to 4. For $K = 4$ balls, the transition probability matrix P is given by:

$$P = \begin{array}{c|ccccc} & 0 & 1 & 2 & 3 & 4 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & \frac{1}{4} & 0 & \frac{3}{4} & 0 & 0 \\ 2 & 0 & \frac{2}{4} & 0 & \frac{2}{4} & 0 \\ 3 & 0 & 0 & \frac{3}{4} & 0 & \frac{1}{4} \\ 4 & 0 & 0 & 0 & 1 & 0 \end{array}$$

Joint Probability in Terms of Conditional Probability

The joint probability $P(A, B, C)$ can be expressed as:

$$P(A, B, C) = P(A) \cdot P(B|A) \cdot P(C|A, B)$$

Where:

- $P(A)$ is the probability of event A.
- $P(B|A)$ is the conditional probability of event B given that event A has occurred.
- $P(C|A, B)$ is the conditional probability of event C given that both events A and B have occurred.

Alternative Expressions:

$$P(A, B, C) = P(B) \cdot P(A|B) \cdot P(C|A, B)$$
$$P(A, B, C) = P(C) \cdot P(A|C) \cdot P(B|A, C)$$

Joint Distribution of Random Variables in a Markov Chain

Suppose $\{X_n : \ n = 0, 1, 2, \ldots\}$ is a stationary Markov chain with

- state space $\mathfrak{X}$ and
- transition probabilities $\{P_{ij} : i, j \in \mathfrak{X}\}$.

Define $\pi_0(i) = \mathrm{P}(X_0 = i)$, $i \in \mathfrak{X}$ to be the distribution of X_0.

What is the joint distribution of X_0, X_1, X_2?

$$
\begin{aligned}
&\mathrm{P}(X_0 = i_0, X_1 = i_1, X_2 = i_2) \\
&= \mathrm{P}(X_0 = i_0)\mathrm{P}(X_1 = i_1 | X_0 = i_0)\mathrm{P}(X_2 = i_2 | X_1 = i_1, X_0 = i_0) \\
&= \mathrm{P}(X_0 = i_0)\mathrm{P}(X_1 = i_1 | X_0 = i_0)\mathrm{P}(X_2 = i_2 | X_1 = i_1) \quad (\because \text{ Markov}) \\
&= \pi_0(i_0) P_{i_0 i_1} P_{i_1 i_2}
\end{aligned}
$$

In general,

$$
\begin{aligned}
&\mathrm{P}(X_0 = i_0, X_1 = i_1, X_2 = i_2, \ldots, X_{n-1} = i_{n-1}, X_n = i_n) \\
&= \pi_0(i_0) P_{i_0 i_1} P_{i_1 i_2} \ldots P_{i_{n-1} i_n}
\end{aligned}
$$

n-Step Transition Probabilities

Suppose $\{X_n\}$ is a stationary Markov chain with state space $\mathfrak{X}$.
Define the n-step transition probabilities

$$P_{ij}^{(n)} = \mathrm{P}(X_{n+k} = j | X_k = i) \quad \text{for } i, j \in \mathfrak{X} \text{ and } n, k = 0, 1, 2, \ldots$$

How to calculate $P_{ij}^{(n)}$?

n-step Transition Probability

In a Markov chain, the n-step transition probability is defined as:

$$P_{ij}^{(n)} = P(X_n = j | X_0 = i) \text{ or, } P_{ij}^{(n)} = P(X_{n+k} = j | X_k = i)$$

This represents the probability of transitioning from state i to state j in n steps.

How to calculate $P_{ij}^{(n)}$?

Example: Ehrenfest Model, 4 Balls

$$\mathbb{P} = \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 & 0 \\ 0 & 2/4 & 0 & 2/4 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

Q1 Find $P^{(2)}_{4,2} = \mathrm{P}(X_2 = 2|X_0 = 4)$.
Only one possible path: $4 \to 3 \to 2$,
so $P^{(2)}_{4,2} = P_{4,3}P_{3,2} = 1 \cdot (3/4) = 3/4$.

Q2 Find $P^{(3)}_{4,2} = \mathrm{P}(X_3 = 2|X_0 = 4)$.
Impossible to go from 4 to 2 in odd number of steps,
so $P^{(3)}_{4,2} = 0$.

$$\mathbb{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 & 0 \\ 0 & 2/4 & 0 & 2/4 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

Q3 Find $P_{4,2}^{(4)} = \mathrm{P}(X_4 = 2 | X_0 = 4)$.

Possible paths:

$$\begin{array}{ccccccccc} 4 & \rightarrow & 3 & \rightarrow & 4 & \rightarrow & 3 & \rightarrow & 2 \\ & & & \searrow & & \nearrow & & \nearrow & \\ & & & & 2 & \rightarrow & 1 & & \end{array}$$

$$P_{4,2}^{(4)} = P_{4,3}P_{3,4}P_{4,3}P_{3,2} + P_{4,3}P_{3,2}P_{2,3}P_{3,2} + P_{4,3}P_{3,2}P_{2,1}P_{1,2}$$
$$= 1 \cdot \frac{1}{4} \cdot 1 \cdot \frac{3}{4} + 1 \cdot \frac{3}{4} \cdot \frac{2}{4} \cdot \frac{3}{4} + 1 \cdot \frac{3}{4} \cdot \frac{2}{4} \cdot \frac{3}{4} = \frac{3}{4}$$

Q4 Find $P_{4,2}^{(10)} = \mathrm{P}(X_{10} = 2 | X_0 = 4)$.

Too many paths to list, likely to miss a few.

Chapman-Kolmogorov Equation for Higher Order Transition Probabilities

The Chapman-Kolmogorov equation relates n-step and m-step transition probabilities as follows:

$$P_{ij}^{(n+m)} = \sum_{k \in S} P_{ik}^{(n)} P_{kj}^{(m)}$$

Where:

- $P_{ij}^{(n+m)}$: Probability of transitioning from state i to state j in $n + m$ steps.
- $P_{ik}^{(n)}$: Probability of transitioning from state i to intermediate state k in n steps.
- $P_{kj}^{(m)}$: Probability of transitioning from intermediate state k to state j in m steps.
- The sum is taken over all possible intermediate states k in the state space S.

Proof of the Chapman-Kolmogorov Equation

We know,

$$P_{ik}^{(n)} = P(X_n = k | X_0 = i), \quad P_{kj}^{(m)} = P(X_m = j | X_0 = k)$$

Using the Law of Total Probability:

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

Applying the Markov Property:

$$P(X_{n+m} = j | X_n = k) = P_{kj}^{(m)}$$

Thus:

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

Therefore:

$$P_{ij}^{(n+m)} = \sum_{k \in S} P_{ik}^{(n)} P_{kj}^{(m)}$$

Chapman-Kolmogorov Equation in Matrix Notation

For $n = 1, 2, 3, \ldots$, let

$$\mathbb{P}^{(n)} = \begin{pmatrix} P_{00}^{(n)} & P_{01}^{(n)} & P_{02}^{(n)} & \cdots & P_{0j}^{(n)} & \cdots \\ P_{10}^{(n)} & P_{11}^{(n)} & P_{12}^{(n)} & \cdots & P_{1j}^{(n)} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \\ P_{i0}^{(n)} & P_{i1}^{(n)} & P_{i2}^{(n)} & \cdots & P_{ij}^{(n)} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \end{pmatrix}$$

be the n-step transition probability matrix.

The Chapman-Kolmogorov equation just asserts that

$$\mathbb{P}^{(m+n)} = \mathbb{P}^{(m)} \times \mathbb{P}^{(n)}$$

Note $\mathbb{P}^{(1)} = \mathbb{P}$, $\Rightarrow \mathbb{P}^{(2)} = \mathbb{P}^{(1)} \times \mathbb{P}^{(1)} = \mathbb{P} \times \mathbb{P} = \mathbb{P}^2$.

By induction,

$$\mathbb{P}^{(n)} = \mathbb{P}^{(n-1)} \times \mathbb{P}^{(1)} = \mathbb{P}^{n-1} \times \mathbb{P} = \mathbb{P}^n$$

Example:

Given the transition matrix:

$$P = \begin{bmatrix} 0.5 & 0.5 \\ 0.2 & 0.8 \end{bmatrix}$$

To find the 2-step transition probabilities:

$$P^{(2)} = P \cdot P = \begin{bmatrix} 0.5 & 0.5 \\ 0.2 & 0.8 \end{bmatrix} \cdot \begin{bmatrix} 0.5 & 0.5 \\ 0.2 & 0.8 \end{bmatrix} = \begin{bmatrix} 0.35 & 0.65 \\ 0.34 & 0.66 \end{bmatrix}$$

Example 4.8 Consider Example 4.1 in which the weather is considered as a two-state Markov chain. If $\alpha = 0.7$ and $\beta = 0.4$, then calculate the probability that it will rain four days from today given that it is raining today.

Solution: The one-step transition probability matrix is given by

$$\mathbf{P} = \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix}$$

Hence,

$$\mathbf{P}^{(2)} = \mathbf{P}^2 = \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix} \cdot \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix}$$

$$= \begin{Vmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{Vmatrix},$$

$$\mathbf{P}^{(4)} = (\mathbf{P}^2)^2 = \begin{Vmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{Vmatrix} \cdot \begin{Vmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{Vmatrix}$$

$$= \begin{Vmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{Vmatrix}$$

and the desired probability P_{00}^4 equals 0.5749. ■

Example 4.9 Consider Example 4.4. Given that it rained on Monday and Tuesday, what is the probability that it will rain on Thursday?

Solution: The two-step transition matrix is given by

$$\mathbf{P}^{(2)} = \mathbf{P}^2 = \begin{Vmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{Vmatrix} \cdot \begin{Vmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{Vmatrix}$$

$$= \begin{Vmatrix} 0.49 & 0.12 & 0.21 & 0.18 \\ 0.35 & 0.20 & 0.15 & 0.30 \\ 0.20 & 0.12 & 0.20 & 0.48 \\ 0.10 & 0.16 & 0.10 & 0.64 \end{Vmatrix}$$

RR NR

Since rain on Thursday is equivalent to the process being in either state 0 or state 1 on Thursday, the desired probability is given by $P_{00}^2 + P_{01}^2 = 0.49 + 0.12 = 0.61$. ■

Law of Total Probability i n-th Step

Define $\pi_n(i) = \mathrm{P}(X_n = i)$, $i \in \mathfrak{X}$ to be the marginal distribution of X_n, $n = 1, 2, \ldots$. Then again by the law of total probabilities,

$$\begin{aligned} \pi_n(j) &= \mathrm{P}(X_n = j) \\ &= \sum_{k \in \mathfrak{X}} \mathrm{P}(X_0 = k)\mathrm{P}(X_n = j | X_0 = k) \qquad (1) \\ &= \sum_{k \in \mathfrak{X}} \pi_0(k) P_{kj}^{(n)} \end{aligned}$$

Suppose the state space $\mathfrak{X}$ is $\{0, 1, 2, \ldots\}$.
If we write the marginal distribution of X_n as a row vector

$$\pi_n = (\pi_n(0), \pi_n(1), \pi_n(2), \ldots),$$

then the equation (1) is

$$\pi_n = \pi_0 \mathbb{P}^{(n)} = \pi_0 \mathbb{P}^n$$

Example: Ehrenfest Model, 4 Balls

$$\mathbb{P} = \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0 & 4/4 & 0 & 0 & 0 \\ 1/4 & 0 & 3/4 & 0 & 0 \\ 0 & 2/4 & 0 & 2/4 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \\ 0 & 0 & 0 & 4/4 & 0 \end{pmatrix} \end{array}$$

Q3 Find $P_{4,2}^{(4)} = \mathrm{P}(X_4 = 2|X_0 = 4)$.

Q4 Find $P_{4,2}^{(10)} = \mathrm{P}(X_{10} = 2|X_0 = 4)$.

Q5 Given $\mathrm{P}(X_0 = i) = 1/5$ for $i = 0, 1, 2, 3, 4$, find $\mathrm{P}(X_4 = 2)$

$$\mathbb{P}^2 = \mathbb{P} \times \mathbb{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1/4 & 0 & 3/4 & 0 & 0 \\ 0 & 5/8 & 0 & 3/8 & 0 \\ 1/8 & 0 & 3/4 & 0 & 1/8 \\ 0 & 3/8 & 0 & 5/8 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \end{pmatrix} \end{array}$$

$$\mathbb{P}^3 = \mathbb{P} \times \mathbb{P}^2 = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 5/8 & 0 & 3/8 & 0 \\ 5/32 & 0 & 3/4 & 0 & 3/32 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 3/32 & 0 & 3/4 & 0 & 5/32 \\ 0 & 3/8 & 0 & 5/8 & 0 \end{pmatrix} \end{array}$$

$$\mathbb{P}^4 = \mathbb{P}^2 \times \mathbb{P}^2 = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 5/32 & 0 & 3/4 & 0 & 3/32 \\ 0 & 17/32 & 0 & 15/32 & 0 \\ 1/8 & 0 & 3/4 & 0 & 1/8 \\ 0 & 15/32 & 0 & 5/32 & 0 \\ 3/32 & 0 & 3/4 & 0 & 5/32 \end{pmatrix} \end{array}$$

Example: Ehrenfest Model, 4 Balls (Cont'd)

$$\mathbb{P}^4 = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 5/32 & 0 & 3/4 & 0 & 3/32 \\ 0 & 17/32 & 0 & 15/32 & 0 \\ 1/8 & 0 & 3/4 & 0 & 1/8 \\ 0 & 15/32 & 0 & 5/32 & 0 \\ 3/32 & 0 & \boxed{3/4} & 0 & 5/32 \end{pmatrix} \end{array}$$

For Q3, $\mathrm{P}(X_4 = 2 | X_0 = 4) = P_{42}^{(4)} = 3/4.$
which agrees with our previous calculation.

Example: Ehrenfest Model, 4 Balls (Cont'd)

To find $P_{4,2}^{(10)}$ for Q4, it's awful lots of work to compute $\mathbb{P}^{10}$...

There are ways to save some work. By the C-K equation,

$$\mathbb{P}_{4,2}^{(10)} = \underbrace{\mathbb{P}_{4,0}^{(5)}\mathbb{P}_{0,2}^{(5)}}_{=0} + \mathbb{P}_{4,1}^{(5)}\mathbb{P}_{1,2}^{(5)} + \underbrace{\mathbb{P}_{4,2}^{(5)}\mathbb{P}_{2,2}^{(5)}}_{=0} + \mathbb{P}_{4,3}^{(5)}\mathbb{P}_{3,2}^{(5)} + \underbrace{\mathbb{P}_{4,4}^{(5)}\mathbb{P}_{4,2}^{(5)}}_{=0}$$

because it's impossible to move between even states in odd number of moves.

We just need to find $\mathbb{P}_{4,1}^{(5)}$, $\mathbb{P}_{4,3}^{(5)}$, $\mathbb{P}_{1,2}^{(5)}$, and $\mathbb{P}_{3,2}^{(5)}$.

Example: Ehrenfest Model, 4 Balls (Cont'd)

$$\mathbb{P}^5 = \mathbb{P}^2 \times \mathbb{P}^3$$

$$= \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1/4 & 0 & 3/4 & 0 & 0 \\ 0 & 5/8 & 0 & 3/8 & 0 \\ 1/8 & 0 & 3/4 & 0 & 1/8 \\ 0 & 3/8 & 0 & 5/8 & 0 \\ 0 & 0 & 3/4 & 0 & 1/4 \end{pmatrix} \end{array} \times \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0 & 5/8 & 0 & 3/8 & 0 \\ 5/32 & 0 & 3/4 & 0 & 3/32 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 3/32 & 0 & 3/4 & 0 & 5/32 \\ 0 & 3/8 & 0 & 5/8 & 0 \end{pmatrix} \end{array}$$

$$= \begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} & & 0 & & \\ & & 3/4 & & \\ & & 0 & & \\ & & 3/4 & & \\ 0 & 15/32 & 0 & 17/32 & 0 \end{pmatrix} \end{array}$$

So

$$\mathbb{P}^{(10)}_{4,2} = \mathbb{P}^{(5)}_{4,1}\mathbb{P}^{(5)}_{1,2} + \mathbb{P}^{(5)}_{4,3}\mathbb{P}^{(5)}_{3,2} = \frac{15}{32} \times \frac{3}{4} + \frac{17}{32} \times \frac{3}{4} = \frac{3}{4}.$$

Example: Ehrenfest Model, 4 Balls (Cont'd)

Q5: Given $\mathrm{P}(X_0 = i) = 1/5$ for $i = 0, 1, 2, 3, 4$, find $\mathrm{P}(X_4 = 2)$.

$$\pi_0 = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}).$$

$$\pi_4 = \pi_0 \mathbb{P}^4 = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}) \begin{pmatrix} 5/32 & 0 & 3/4 & 0 & 3/32 \\ 0 & 17/32 & 0 & 15/32 & 0 \\ 1/8 & 0 & 3/4 & 0 & 1/8 \\ 0 & 15/32 & 0 & 17/32 & 0 \\ 3/32 & 0 & 3/4 & 0 & 5/32 \end{pmatrix}$$

$$\pi_4(2) = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}) \begin{pmatrix} 3/4 \\ 0 \\ 3/4 \\ 0 \\ 3/4 \end{pmatrix}$$

$$= \frac{1}{5} \cdot \frac{3}{4} + \frac{1}{5} \cdot 0 + \frac{1}{5} \cdot \frac{3}{4} + \frac{1}{5} \cdot 0 + \frac{1}{5} \cdot \frac{3}{4} = \frac{9}{20}$$

Accessible State in Stochastic Processes

- In the context of stochastic processes, particularly in the study of Markov chains, a state j is said to be **accessible** from a state i if there exists a positive integer n such that the transition probability from i to j in n steps is greater than zero:

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

  where $P_{ij}^{(n)}$ is the probability of transitioning from state i to state j in n steps.
- If state j is accessible from state i, we denote this relationship as:

$$i \xrightarrow{\text{accessible}} j$$

Example: Weather States

States

- **State 1**: Sunny
- **State 2**: Cloudy
- **State 3**: Rainy

Transition Probabilities

The transition probabilities can be represented in matrix form as follows:

$$P = \begin{pmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix} = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.5 & 0.3 \\ 0.4 & 0.4 & 0.2 \end{pmatrix}$$

Accessibility of States

Accessibility from Sunny (State 1):

- $P_{12}^{(1)} = 0.3 > 0$ (to Cloudy)
- $P_{13}^{(1)} = 0.1 > 0$ (to Rainy)

Accessibility from Cloudy (State 2):

- $P_{21}^{(1)} = 0.2 > 0$ (to Sunny)
- $P_{23}^{(1)} = 0.3 > 0$ (to Rainy)

Accessibility from Rainy (State 3):

- $P_{31}^{(1)} = 0.4 > 0$ (to Sunny)
- $P_{32}^{(1)} = 0.4 > 0$ (to Cloudy)

Communicating States

In a Markov chain, two states i and j are said to **communicate** if:

- State j is accessible from state i:

$$P_{ij}^{(n)} > 0 \text{ for some } n$$

- State i is accessible from state j:

$$P_{ji}^{(m)} > 0 \text{ for some } m$$

We denote the communication relationship as:

$$i \leftrightarrow j$$

Communicating Classes are subsets of states where each state can reach every other state in that subset.

Transitivity of Communication

If we have three states i, j, and k in a Markov chain, the transitivity of communication states that:

- If state i communicates with state j (denoted as $i \leftrightarrow j$), and
- State j communicates with state k (denoted as $j \leftrightarrow k$),
- Then it follows that state i also communicates with state k (denoted as $i \leftrightarrow k$).

Communicating Class

A **communicating class** is a subset of states in a Markov chain that satisfies the following conditions:

- All states in the class communicate with each other.
- For any two states i and j in the communicating class, the following holds:

$$i \leftrightarrow j \quad (\text{i.e., } P_{ij}^{(n)} > 0 \text{ and } P_{ji}^{(m)} > 0 \text{ for some } n, m)$$

- The class is **maximal**, meaning that no state outside the class can be reached from any state within the class:

$$P_{ij}^{(n)} = 0 \quad \text{for all states } j \text{ not in the class.}$$

Imagine you have three states: A, B, and C.

If you can move from A to B, B to C, and C to A, then A, B, and C form a maximal class.
If you can't move from A, B, or C to any other state (like D), then D is not part of this maximal class.In simpler terms, a maximal class is like a closed group where all members can interact with each other, but they can't interact with anyone outside the group.

Example 1. Specify the classes of the following Markov chains.

$$\mathbb{P}_1 = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0.9 & 0.1 \end{pmatrix} \end{array} \quad \mathbb{P}_2 = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Example 2. How many classes does the Ehrenfest diffusion model with K balls have?

Example 1. Specify the classes of the following Markov chains.

$$\mathbb{P}_1 = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0.9 & 0.1 \end{pmatrix} \end{array} \quad \mathbb{P}_2 = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

For $\mathbb{P}_1$, $1 \leftrightarrow 2 \rightarrow 3 \leftrightarrow 4$. Classes: $\{1,2\}$, $\{3,4\}$.

Example 2. How many classes does the Ehrenfest diffusion model with K balls have?

Example 1. Specify the classes of the following Markov chains.

$$\mathbb{P}_1 = \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0.5 & 0.5 & 0 & 0 \\ 2 & 0.3 & 0.6 & 0.1 & 0 \\ 3 & 0 & 0 & 0.2 & 0.8 \\ 4 & 0 & 0 & 0.9 & 0.1 \end{array} \qquad \mathbb{P}_2 = \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 1/2 & 1/2 & 0 & 0 \\ 2 & 1/2 & 1/2 & 0 & 0 \\ 3 & 1/4 & 1/4 & 1/4 & 1/4 \\ 4 & 0 & 0 & 0 & 1 \end{array}$$

For $\mathbb{P}_1$, $1 \leftrightarrow 2 \rightarrow 3 \leftrightarrow 4$. Classes: $\{1,2\}$, $\{3,4\}$.

For $\mathbb{P}_2$, $\begin{array}{ccc} 1 & \longleftrightarrow & 2 \\ & \nwarrow & \uparrow \\ 4 & \leftarrow & 3 \\ \circlearrowleft & & \circlearrowleft \end{array}$. Classes: $\{1,2\}$, $\{3\}$, $\{4\}$.

Example 2. How many classes does the Ehrenfest diffusion model with K balls have?

Example 1. Specify the classes of the following Markov chains.

$$
\mathbb{P}_1 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \!\!\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0.9 & 0.1 \end{pmatrix} \end{array}
\qquad
\mathbb{P}_2 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \!\!\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}
$$

For $\mathbb{P}_1$, $1 \leftrightarrow 2 \to 3 \leftrightarrow 4$. Classes: $\{1,2\}$, $\{3,4\}$.

For $\mathbb{P}_2$,
$$
\begin{tikzcd}
1 \arrow[r, leftrightarrow] & 2 \\
4 \arrow[loop below] & 3 \arrow[l] \arrow[u] \arrow[ul] \arrow[loop below]
\end{tikzcd}
$$
. Classes: $\{1,2\}$, $\{3\}$, $\{4\}$.

Example 2. How many classes does the Ehrenfest diffusion model with K balls have?

All states communicate. Only one class.

Closed Classes

Definition. A class C is said to be closed if

$$P_{ij} = 0 \quad \text{for all } i \text{ in } C \text{ and } j \text{ not in } C.$$

Once the process gets into a closed class. It will never leave the class since the outgoing probabilities from the class are all 0.

Examples.

- For $\mathbb{P}_1$ in the previous slide, the class $\{1,2\}$ is not closed because it has a non-zero outgoing probability $P_{23} = 0.1 > 0$. The class $\{3, 4\}$ is closed.
- For $\mathbb{P}_2$ in the previous slide, the classes $\{1, 2\}$ and $\{4\}$ are closed, and $\{3\}$ is not closed.

A Markov Chain Restricted to a Closed Class is Also a Markov Chain

Example.

$$\mathbb{P}_1 = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0.9 & 0.1 \end{pmatrix} \end{array}$$

- For $\mathbb{P}_1$ above, the Markov chain restricted to the class $\{3,4\}$ is also a Markov chain, with transition matrix

$$\begin{array}{c c} & \begin{array}{cc} 3 & 4 \end{array} \\ \begin{array}{c} 3 \\ 4 \end{array} & \begin{pmatrix} 0.2 & 0.8 \\ 0.9 & 0.1 \end{pmatrix} \end{array}$$

- The Markov chain for $\mathbb{P}_1$ can not be restricted to $\{1,2\}$ as it may transit out of the state space from 2 to 3.

Irreducibility

- A Markov chain is said to be irreducible if all states communicate with each other.
- A Markov chain is said to be irreducible if it has only one communicating class.

Recurrent State

A recurrent state in a Markov chain is a state that, once visited, has a probability of returning to it equal to one.

- For any state i, we define the probability of recurrence to i is

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

  State i is **recurrent** if $f_{ii} = 1$, and it is **transient** if $f_{ii} < 1$.

- if i is a recurrent state, then,

$$\sum_{n=0}^{\infty} P_{ii}^{(n)} = \infty$$

- This means that if the process starts in state i, it will eventually return to i with certainty.

Mean Recurrent Time

The mean recurrent time, denoted as μ_i, is the expected time it takes for a Markov chain to return to a particular recurrent state i.
If T_i is the first return time to state i after leaving it, then we can define

$$T_i = \min\{n \geq 1 : X_n = i \mid X_0 = i\}$$

Hence, the mean recurrent time is

$$\begin{aligned} \mu_i &= E[T_i | X_0 = i] \\ &= \sum_{i=1}^{n} n f_{ii} \end{aligned}$$

Calculation of Mean Recurrent Time

To calculate the mean recurrent time, use the relationship:

$$\mu_i = \frac{1}{\pi_i}$$

where π_i is the stationary probability of state i.

1. Identify recurrent states.
2. Compute the stationary distribution π by solving $\pi P = \pi$ and $\sum \pi_i = 1$.
3. Calculate mean recurrent time: $\mu_i = \frac{1}{\pi_i}$.

Example

Consider a Markov chain with a stationary distribution:

$$\pi = (\pi_1, \pi_2, \pi_3) = (0.2, 0.5, 0.3)$$

The mean recurrent times are:

$$\mu_1 = \frac{1}{0.2} = 5$$
$$\mu_2 = \frac{1}{0.5} = 2$$
$$\mu_3 = \frac{1}{0.3} \approx 3.33$$

The mean recurrent time provides insights into how long, on average, the chain spends away from a recurrent state before returning. It is crucial for understanding the long-term behavior of Markov chains and their states.

Recurrent states can be further classified into two categories: **positive recurrent** and **null recurrent** states.

(i). Positive Recurrent State

- A state i is said to be positive recurrent if:

$$\mu_i = E[T_i | X_0 = i] < \infty$$

  where T_i is the time until the process returns to state i.
- This means the expected time to return to the state is finite.
- Example: In a simple random walk on a finite state space, all states are positive recurrent.

(ii). Null Recurrent State

- A state i is null recurrent if:

$$\mu_i = E[T_i | X_0 = i] = \infty$$

- This indicates that while the state will be revisited, the expected return time is infinite.

Theorem 1

If N is the number of visits to state i given $X_0 = i$, then:

$$E(N|X_0 = i) = \frac{1}{1 - f_i}$$

Proof: For simplicity, let $f_i - f_{ii}$ and then

$$\begin{aligned} E(N|X_0 = i) = & E[N|T_i = \infty, X_0 = i]P\{T_i = \infty|X_0 = i\} \\ & + E[N|T_i < \infty, X_0 = i]P\{T_i < \infty|X_0 = i\} \end{aligned}$$

This simplifies to:

$$E(N|X_0 = i) = 1 \cdot (1 - f_i) + f_i[1 + E(N|X_0 = i)]$$

- If $T_i = \infty$, except for $n = 0$ (where $X_0 = i$), there will never be a visit to i, i.e.:

$$E(N|T_i = \infty, X_0 = i) = 1$$

- If $T_i < \infty$, there is guaranteed to be one visit, say at X_k (where $X_k = i$). Therefore:

$$E(N|T_i < \infty, X_k = i) = E(N|T_i < \infty, X_0 = i)$$

by the Markov property, which gives us:

$$E[N|T_i < \infty, X_0 = i] = 1 + E[N|X_0 = i]$$

Substituting back, we have:

$$E[N|X_0 = i] = 1 \cdot (1 - f_i) + \{1 + E[N|X_0 = i]\} \cdot (f_i)$$

This leads to:

$$E[N|X_0 = i] = \frac{1}{1 - f_i}$$

Another expression for $E[N|X_0 = i]$ is:

$$E[N|X_0 = i] = \sum_{n=0}^{\infty} p_{ii}^{(n)}$$

Transient State

A **transient state** in a Markov chain is a state that, once left, may never be returned to with a probability greater than zero.

- For any state i, the probability of recurrence to i is less than one. That is,

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

- The probability of returning to state i after leaving is less than one:

$$\sum_{n=0}^{\infty} P_{ii}^{(n)} < \infty$$

- This implies that there exists a non-zero probability that the process may move to other states and never come back to i.
- In an irreducible chain, if one state is recurrent, all states must be recurrent. Similarly, if one state is transient, all states are transient.

Example: Consider a simple Markov chain with three states: A, B, and C with the following transition matrix:

$$\begin{pmatrix} 0.5 & 0.5 & 0.0 \\ 0.2 & 0.3 & 0.5 \\ 0.0 & 0.4 & 0.6 \end{pmatrix}$$

- State A is a transient state because once the process moves to states B or C, it may never return to A with probability 1.
- In contrast, states B and C are recurrent since there is a positive probability of returning to them.

Absorbing State

An **absorbing state** in a Markov chain is a state that, once entered, cannot be left. More formally, a state i is absorbing if:

- The probability of staying in state i is one:

$$P_{ii} = 1$$

- The probability of transitioning to any other state j (where $j \neq i$) is zero:

$$P_{ij} = 0 \quad \text{for all } j \neq i$$

Example: Consider a Markov chain with three states: A, B, and C, with the following transition matrix:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 0.5 & 0 \\ 0.2 & 0.8 & 0 \end{pmatrix}$$

- State A is an absorbing state because once the process enters A, it cannot leave.

- An absorbing state is a special case of a recurrent state.
- Every absorbing state is recurrent, but not every recurrent state is absorbing.

Every absorbing state is recurrent: Since you can't leave an absorbing state once you enter it, you are always in that state, making it recurrent.

Not every recurrent state is absorbing: A recurrent state allows for the possibility of leaving and returning, but it doesn't mean you can't leave it. You can leave a recurrent state and come back to it, whereas an absorbing state doesn't allow you to leave at all.

In simpler terms, think of an absorbing state as a one-way street with no exits, while a recurrent state is like a roundabout—you can leave, but you'll always come back around.

Classroom Exercise: Transition Probability Matrix

Dataset: Patient Health Events

Consider the following data from a study on patient health events over three visits (states):

Patient ID	Visit 1	Visit 2
1	Healthy	Sick
2	Sick	Healthy
3	Healthy	Healthy
4	Sick	Sick
5	Healthy	Sick
6	Sick	Healthy

States:

- Healthy
- Sick

Questions:

- Using the data provided, create a transition probability matrix that describes the probabilities of moving from one state to another between Visit 1 and Visit 2.
- Show your calculations and explain how you derived the probabilities.

Expected Output: A transition probability matrix of the form:

$$P = \begin{pmatrix} P(\text{Healthy} \to \text{Healthy}) & P(\text{Healthy} \to \text{Sick}) \\ P(\text{Sick} \to \text{Healthy}) & P(\text{Sick} \to \text{Sick}) \end{pmatrix}$$

Solution: Transition Probability Matrix

Data Overview:

- The data contains transitions from Visit 1 to Visit 2 for 6 patients.
- States: Healthy (H) and Sick (S).

Count Transitions:

- From Healthy to Healthy (H to H): 1 occurrence (Patient 3)
- From Healthy to Sick (H to S): 2 occurrences (Patients 1, 5)
- From Sick to Healthy (S to H): 2 occurrences (Patients 2, 6)
- From Sick to Sick (S to S): 1 occurrence (Patient 4)

Total Transitions:

- Total transitions from Healthy = 3 (1 to H + 2 to S)
- Total transitions from Sick = 3 (2 to H + 1 to S)

Constructing the Transition Probability Matrix: Transition probabilities are calculated as:

$$P(X \to Y) = \frac{\text{Number of transitions from } X \text{ to } Y}{\text{Total transitions from } X}$$

Transition Counts:

- From Healthy:
  - To Healthy: 1
  - To Sick: 2
- From Sick:
  - To Healthy: 2
  - To Sick: 1

Transition Probability Matrix:

$$P = \begin{pmatrix} P(H \to H) & P(H \to S) \\ P(S \to H) & P(S \to S) \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 0.33 & 0.67 \\ 0.67 & 0.33 \end{pmatrix}$$

Classroom Exercise: Transition Probability Matrix

Dataset: Patient Health Events

Consider the following data from a study on patient health events over three visits (states):

Patient ID	Visit 1	Visit 2
1	Healthy	Sick
2	Sick	Healthy
3	Healthy	Healthy
4	Sick	Sick
5	Healthy	Recovering
6	Sick	Recovering
7	Recovering	Healthy
8	Recovering	Sick
9	Healthy	Sick
10	Sick	Healthy

States:

- Healthy
- Sick
- Recovering

Questions:

- Using the data provided, create a transition probability matrix that describes the probabilities of moving from one state to another between Visit 1 and Visit 2.
- Show your calculations and explain how you derived the probabilities.

Transition Probability Matrix

Transition Probability Matrix:

The transition probability matrix P based on the dataset is given by:

$$P = \begin{pmatrix} 0.40 & 0.60 & 0.00 \\ 0.60 & 0.40 & 0.00 \\ 0.50 & 0.50 & 0.00 \end{pmatrix}$$

Where:

- The rows represent the states at Visit 1:
  - 1: Healthy
  - 2: Sick
  - 3: Recovering
- The columns represent the states at Visit 2:
  - 1: Healthy
  - 2: Sick
  - 3: Recovering

Analysis of States

Given Transition Probability Matrix:

$$P = \begin{pmatrix} 0.40 & 0.60 & 0.00 \\ 0.60 & 0.40 & 0.00 \\ 0.50 & 0.50 & 0.00 \end{pmatrix}$$

Classification of States:

- **Communicating States:**
  - States Healthy (H) and Sick (S) communicate with each other:

$$P(H \to S) > 0 \quad \text{and} \quad P(S \to H) > 0$$

  - States Sick (S) and Recovering (R) communicate with each other:

$$P(S \to R) > 0 \quad \text{and} \quad P(R \to S) > 0$$

  - States Healthy (H) and Recovering (R) do not communicate directly.

Analysis of States

- **State Healthy (H):**
  - Probability of returning to Healthy:

$$P_{HH}^{(n)} = P(H \to H) = 0.40$$

  - There is a positive probability of transitioning to Sick (S), which also has a probability of returning to H.
- **State Sick (S):**
  - Probability of returning to Sick:

$$P_{SS}^{(n)} = P(S \to S) = 0.40$$

  - Similar to H, it can return to H, indicating recurrent behavior.
- **State Recovering (R):**
  - Probability of returning to Recovering:

$$P_{RR}^{(n)} = P(R \to R) = 0.00$$

  - Since it can only transition to Healthy or Sick, R is classified as transient.

Hence,

- States Healthy (H) and Sick (S) are recurrent because there exists a positive probability of returning to these states.
- State Recovering (R) is transient as it does not return with certainty.

Absorbing States:

- An absorbing state is one where, once entered, it cannot be left (i.e., $P(i \to i) = 1$).
- In this matrix:

$$P(H \to H) = 0.40, \quad P(S \to S) = 0.40, \quad P(R \to R) = 0.00$$

- Since all states have non-zero probabilities to transition to other states, there are no absorbing states.

Periodic States

Imagine you can only return to a state every 2 steps. For instance, if you can only return to a state on steps 2, 4, 6, etc., that state is periodic (with a period of 2). Formally, a state i is periodic with period d if:

$$P(X_n = i|X_0 = i) > 0 \text{ only if } n \text{ is a multiple of } d.$$

That is,

$$d(i) = \text{GCD}\{n : P(X_n = i|X_0 = i) > 0\} > 1.$$

This means you can only return to state i at fixed intervals.

Example:

- Consider the transition matrix:

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

- From state A, you can only go to state B, and vice versa.
- Thus, you return to A every 2 steps.

Aperiodic States

If you can return to a state at any step, like steps 1, 2, 3, 4, etc., that state is aperiodic. You can go back to the state whenever you want, not just at fixed intervals. Formally, a state i is aperiodic if:

$$d(i) = \text{GCD}\{n : P(X_n = i | X_0 = i) > 0\} = 1.$$

This means you can return to state i at irregular intervals.

Example:

- Consider the transition matrix:

$$P = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$$

- From state C, you can go to C or D.
- From state D, you can also go back to C or stay at D.
- You can return to C in 1 step, 2 steps, etc., so C is aperiodic.

Irreducibility and Aperiodicity:

- The Markov chain is irreducible: all states communicate.
- It is aperiodic: no cycles restrict transitions.

Ergodic States

A state i in a Markov chain is said to be an **ergodic** state if state i is a

1. recurrent state
2. irreducible state
3. aperiodic state

Limiting Distribution

- If there are n states, transition matrix P is an $n \times n$ matrix.

$$P = \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{pmatrix}$$

- The limiting distribution describes the long-term behavior of the Markov chain. The limiting probability vector π satisfies:

$$\pi P = \pi$$

and

$$\sum_i \pi_i = 1$$

- The chain converges to a unique stationary distribution π. As $n \to \infty$:

$$\lim_{n\to\infty} P^n = \begin{pmatrix} \pi_0 & \pi_1 & \pi_2 & \pi_3 \end{pmatrix}$$

Example

Consider a transition matrix:

$$P = \begin{pmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{pmatrix}$$

- Set up the equations:

$$\pi_1 = 0.7\pi_1 + 0.4\pi_2 \quad (1)$$
$$\pi_2 = 0.3\pi_1 + 0.6\pi_2 \quad (2)$$

- Add normalization:

$$\pi_1 + \pi_2 = 1$$

Rearranging the Equations

Rearranging equation (1):

$$\pi_1 - 0.7\pi_1 = 0.4\pi_2 \implies 0.3\pi_1 = 0.4\pi_2 \implies \pi_1 = \frac{4}{3}\pi_2 \quad (3)$$

Rearranging equation (2):

$$\pi_2 - 0.6\pi_2 = 0.3\pi_1 \implies 0.4\pi_2 = 0.3\pi_1 \implies \pi_2 = \frac{3}{4}\pi_1 \quad (4)$$

The normalization condition is:

$$\pi_1 + \pi_2 = 1$$

Substituting (3) into the normalization:

$$\frac{4}{3}\pi_2 + \pi_2 = 1 \implies \frac{7}{3}\pi_2 = 1 \implies \pi_2 = \frac{3}{7}$$

Limiting Probabilities

Substituting π_2 back into (3):

$$\pi_1 = \frac{4}{3} \cdot \frac{3}{7} = \frac{4}{7}$$

The limiting probabilities are:

$$\pi_1 = \frac{4}{7}$$
$$\pi_2 = \frac{3}{7}$$

Limiting Probabilities of 2-state Markov Chain

The transition matrix is given by:

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

Find the limiting probabilities π_1 and π_2.

Setting Up the Limiting Distribution

Let $\pi = (\pi_1, \pi_2)$ be the limiting probabilities. We require:

$$\pi P = \pi$$

This results in the equations:

$$\pi_1 = (1-\alpha)\pi_1 + \beta\pi_2 \quad (1)$$
$$\pi_2 = \alpha\pi_1 + (1-\beta)\pi_2 \quad (2)$$

Rearranging (1):

$$\pi_1 - (1-\alpha)\pi_1 = \beta\pi_2 \implies \alpha\pi_1 = \beta\pi_2 \implies \pi_1 = \frac{\beta}{\alpha}\pi_2 \quad (3)$$

Rearranging (2):

$$\pi_2 - (1-\beta)\pi_2 = \alpha\pi_1 \implies \beta\pi_2 = \alpha\pi_1 \implies \pi_2 = \frac{\alpha}{\beta}\pi_1 \quad (4)$$

Normalization Condition

The normalization condition is:

$$\pi_1 + \pi_2 = 1 \quad (5)$$

Substituting (3) into (5):

$$\frac{\beta}{\alpha}\pi_2 + \pi_2 = 1.$$

Solving for Limiting Probabilities

Letting $\pi_2 = x$:

$$\left(\frac{\beta}{\alpha} + 1\right) x = 1 \implies x = \frac{\alpha}{\alpha + \beta}.$$

Thus,

$$\pi_2 = \frac{\alpha}{\alpha + \beta}.$$

Substitute back to find π_1:

$$\pi_1 = \frac{\beta}{\alpha + \beta}.$$

The limiting probabilities are:

$$\begin{aligned} \pi_1 &= \frac{\beta}{\alpha + \beta} \quad \text{(State 1)} \\ \pi_2 &= \frac{\alpha}{\alpha + \beta} \quad \text{(State 2)} \end{aligned}$$

Classroom Exercise

Given the transition matrix for weather states:

$$P = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.5 & 0.3 \\ 0.1 & 0.4 & 0.5 \end{pmatrix}$$

1. Determine whether the states (Sunny, Cloudy, Rainy) are ergodic.
2. If the system starts in the Sunny state, what is the long-term behavior of the weather states?

Solution: Part 1 - Determine Ergodicity

Step 1: Check Irreducibility

- From **Sunny (S)**:
  - To Sunny: $P(S \to S) = 0.6$
  - To Cloudy: $P(S \to C) = 0.3$
  - To Rainy: $P(S \to R) = 0.1$
- From **Cloudy (C)**:
  - To Sunny: $P(C \to S) = 0.2$
  - To Cloudy: $P(C \to C) = 0.5$
  - To Rainy: $P(C \to R) = 0.3$
- From **Rainy (R)**:
  - To Sunny: $P(R \to S) = 0.1$
  - To Cloudy: $P(R \to C) = 0.4$
  - To Rainy: $P(R \to R) = 0.5$

Conclusion: All states can reach each other, so the chain is irreducible.

Solution: Part 1 - Determine Ergodicity

Step 2: Check Recurrence Since the chain is irreducible, we can conclude that all states are recurrent. In irreducible chains, if at least one state is recurrent, then all states are recurrent.

Step 3: Check Aperiodicity

- A state is **aperiodic** if the GCD of return times is 1.
- Sunny can return in 1 step (stay) or through combinations.
- Cloudy and Rainy can also be revisited at various steps.

Conclusion: All states are **aperiodic**.

Hence, the transition matrix P is:

- **Irreducible**
- **Recurrent**
- **Aperiodic**

Therefore, all states (Sunny, Cloudy, Rainy) are **ergodic states**.

Solution: Part 2 - Long-Term Behavior

To find the stationary distribution $\pi = (\pi_1, \pi_2, \pi_3)$ such that:

$$\pi P = \pi$$

This leads to the equations:

$$\begin{aligned}
\pi_1 &= 0.6\pi_1 + 0.2\pi_2 + 0.1\pi_3 \\
\pi_2 &= 0.3\pi_1 + 0.5\pi_2 + 0.4\pi_3 \\
\pi_3 &= 0.1\pi_1 + 0.3\pi_2 + 0.5\pi_3
\end{aligned}$$

with the normalization condition:

$$\pi_1 + \pi_2 + \pi_3 = 1.$$

Solving the Equations

Rearranging the equations gives:

$$0.4\pi_1 - 0.2\pi_2 - 0.1\pi_3 = 0$$
$$0.3\pi_1 - 0.5\pi_2 + 0.4\pi_3 = 0$$
$$0.1\pi_1 + 0.3\pi_2 - 0.5\pi_3 = 0$$

Substituting π_2 and π_3 in terms of π_1:

$$\pi_2 = 2\pi_1 - 0.5\pi_3, \quad \pi_3 = \frac{14}{13}\pi_1$$

Normalization Condition

Using the normalization condition:

$$\pi_1 + \pi_2 + \pi_3 = 1$$

yields:

$$\pi_1 + (2\pi_1 - 0.5\pi_3) + \pi_3 = 1$$

After substituting π_3:

$$\pi_1\left(\frac{46}{13}\right) = 1 \implies \pi_1 = \frac{13}{46}$$

Final Stationary Distribution

Finally, we have:

$$\pi_1 = \frac{13}{46}, \quad \pi_2 = \frac{19}{46}, \quad \pi_3 = \frac{14}{46}$$

Long-term behavior:

- Sunny: $\frac{13}{46} \approx 28.26\%$
- Cloudy: $\frac{19}{46} \approx 41.3\%$
- Rainy: $\frac{14}{46} \approx 30.4\%$

Interpretation of Limiting Probabilities

The limiting probabilities represent the long-term behavior of weather conditions:

- $\pi_1 \approx 0.2826$: Approximately 28.26% of days will be sunny.
- $\pi_2 \approx 0.430$: Approximately 43.0% of days will be cloudy.
- $\pi_3 \approx 0.304$: Approximately 30.4% of days will be rainy.

This information helps in understanding the weather patterns over time, which can aid in planning outdoor activities and agricultural decisions.

Ehrenfest diffusion model

The Ehrenfest model simulates the behavior of particles (or "balls") moving between two containers (or "urns"). In the Ehrenfest model, if there are i balls in urn A (i.e., $X_n = i$), the transition probabilities are defined as follows:

- $P_{i,i+1} = P(X_{n+1} = i + 1 | X_n = i) = \frac{K-i}{K}$
  (probability of moving a ball from urn B to urn A)
- $P_{i,i-1} = P(X_{n+1} = i - 1 | X_n = i) = \frac{i}{K}$
  (probability of moving a ball from urn A to urn B)
- $P_{ii} = P(X_{n+1} = i | X_n = i) = 0$
  (no self-transition)

Transition Probability Matrix

For the Ehrenfest model with K total balls, the transition probability matrix P is defined as:

$$P = \begin{pmatrix} 0 & \frac{K}{K} & 0 & \cdots & 0 \\ \frac{1}{K} & 0 & \frac{K-1}{K} & \cdots & 0 \\ 0 & \frac{2}{K} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \frac{K-1}{K} & 0 & 1 \end{pmatrix}$$

For example, for $K = 3$:

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

Ehrenfest diffusion model

The Ehrenfest model simulates the behavior of particles (or "balls") moving between two containers (or "urns"). In the Ehrenfest model, if there are i balls in urn A (i.e., $X_n = i$), the transition probabilities are defined as follows:

- $P_{i,i+1} = P(X_{n+1} = i+1|X_n = i) = \frac{K-i}{K}$
  (probability of moving a ball from urn B to urn A)
- $P_{i,i-1} = P(X_{n+1} = i-1|X_n = i) = \frac{i}{K}$
  (probability of moving a ball from urn A to urn B)
- $P_{ii} = P(X_{n+1} = i|X_n = i) = 0$
  (no self-transition)

Transition Probability Matrix of Ehrenfest diffusion model

For the Ehrenfest model with K total balls, the transition probability matrix P is defined as:

$$P = \begin{pmatrix} 0 & \frac{K}{K} & 0 & \cdots & 0 \\ \frac{1}{K} & 0 & \frac{K-1}{K} & \cdots & 0 \\ 0 & \frac{2}{K} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \frac{K-1}{K} & 0 & 1 \end{pmatrix}$$

For example, for $K = 3$:

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

Finding Limiting Probabilities

The stationary distribution π satisfies the equation:

$$\pi P = \pi$$

This means:

$$\begin{pmatrix} \pi_0 & \pi_1 & \pi_2 & \cdots & \pi_K \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \frac{1}{K} & 0 & \frac{K-1}{K} & \cdots & 0 \\ 0 & \frac{2}{K} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \frac{K-1}{K} & 0 & 1 \end{pmatrix} = \begin{pmatrix} \pi_0 & \pi_1 & \pi_2 & \cdots & \pi_K \end{pmatrix}$$

Calculating Each Component

To find the i-th component of πP, we have:

$$\pi_i = \sum_j \pi_j P_{j,i}$$

For $i = 0$:

$$\pi_0 = \pi_1 \cdot P_{1,0} = \pi_1 \cdot \frac{1}{K}$$

For $i = 1$:

$$\pi_1 = \pi_0 \cdot P_{0,1} + \pi_2 \cdot P_{2,1} = \pi_0 \cdot 1 + \pi_2 \cdot \frac{2}{K}$$

For $i = 2$:

$$\pi_2 = \pi_1 \cdot P_{1,2} + \pi_3 \cdot P_{3,2} = \pi_1 \cdot \frac{K-1}{K} + \pi_3 \cdot 0$$

Balance Equation for States i and $i+1$

Outgoing Flow from State i:

$$\text{Flow out of } i = \pi_i \cdot P_{i,i+1}$$

Incoming Flow to State i:

$$\text{Flow into } i = \pi_{i+1} \cdot P_{i+1,i}$$

Setting these equal for equilibrium:

$$\pi_i \cdot P_{i,i+1} = \pi_{i+1} \cdot P_{i+1,i}$$

To find the limiting probabilities $\pi - i$, follow these steps:

1. **Write the Equilibrium Equations:**
   For each state i:
   $$\pi_i \cdot P_{i,i+1} = \pi_{i+1} \cdot P_{i+1,i}$$
2. **Substitute Transition Probabilities:**
   Using the Ehrenfest model transition probabilities:
   $$\pi_i \cdot \frac{K-i}{K} = \pi_{i+1} \cdot \frac{i+1}{K}$$
3. **Rearrange to Find a Recurrence Relation:**
   This leads to:
   $$\frac{\pi_{i+1}}{\pi_i} = \frac{K-i}{i+1}$$

1. **Express π_i) in Terms of π_0:**
   Start with π_0:

$$\pi_1 = K\pi_0, \quad \pi_2 = \frac{(K-1)}{2}K\pi_0, \ldots$$

2. **Find General Formula:**
   The general form for π_i:

$$\pi_i = \pi_0 \cdot \frac{(K-1)(K-2)\cdots(K-i+1)}{i!}K^i$$

3. **Normalize the Distribution:**
   Use the condition:

$$\sum_{i=0}^{K} \pi_i = 1$$

The limiting probabilities for the Ehrenfest model follow a binomial distribution:

$$\pi_i = \frac{K!}{i!(K-i)!} \cdot \frac{1}{2^K} \qquad ; \quad \text{for} \quad i = 0, 1, \ldots, K.$$

Limiting Probabilities for the Ehrenfest Model (K = 3)

For $K = 3$, the limiting probabilities are given by the binomial distribution:

$$\pi_i = \frac{3!}{i!(3-i)!} \cdot \frac{1}{2^3} = \frac{6}{i!(3-i)!} \cdot \frac{1}{8}$$

for $i = 0, 1, 2, 3$.

Calculating Limiting Probabilities:

$$\pi_0 = \frac{6}{0! \cdot 3!} \cdot \frac{1}{8} = \frac{1}{8},$$
$$\pi_1 = \frac{6}{1! \cdot 2!} \cdot \frac{1}{8} = \frac{3}{8},$$
$$\pi_2 = \frac{6}{2! \cdot 1!} \cdot \frac{1}{8} = \frac{3}{8},$$
$$\pi_3 = \frac{6}{3! \cdot 0!} \cdot \frac{1}{8} = \frac{1}{8}.$$

Ehrenfest Diffusion Model: Limiting Probabilities

For $K = 3$, the transition probability matrix P is:

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

Limiting Probabilities for the Ehrenfest Model

Given the transition matrix:

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

Discuss the limiting behavior of the Markov chain represented by the matrix P. What does it mean for the chain to converge to a limiting distribution?