Project Report

On

AN OVERVIEW OF STOCHASTIC MARKOV CHAIN MODELS IN HEALTH APPLICATIONS

Submitted

in partial fulfilment of the requirements for the degree of

BACHELOR OF SCIENCE

in

MATHEMATICS

by

ANEETA ANTONY (Register No. AB23BMAT007)

Under the Supervision of

MS. PARVATHI T.S

DEPARTMENT OF MATHEMATICS AND STATISTICS

ST. TERESA'S COLLEGE (AUTONOMOUS)

ERNAKULAM, KOCHI - 682011

SEPETEMBER 2025

CERTIFICATE

This is to certify that the dissertation entitled, **AN OVERVIEW OF STOCHASTIC MARKOV CHAIN MODELS IN HEALTH APPLICATIONS** is a bonafide record of the work done by Ms. **ANEETA ANTONY** under my guidance as partial fulfillment of the award of the degree of **Bachelor of Science in Mathematics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

Date:30.09.2025
Place: Ernakulam

MS. PARVATHI T.S
Assistant Professor,
Department of Mathematics and Statistics,
St. Teresa's College(Autonomous),
Ernakulam.

ST. TERESA'S COLLEGE (AUTONOMOUS)
DEPARTMENT OF MATHEMATICS AND STATISTICS
ERNAKULAM

Dr.ELIZABETH RESHMA M.T
Assistant Professor and Head,
Department of Mathematics and Statistics,
St. Teresa's College(Autonomous),
Ernakulam.

External Examiners

1:........ 20/01/26 2: ....

ii

DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of Ms. Parvathi T.S, Assistant Professor, Department of Mathematics, St. Teresa's College(Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Ernakulam.

30.09.2025

ANEETA ANTONY

ACKNOWLEDGEMENT

It is with great respect and gratitude that we acknowledge the guidance and support received during the course of this project on Markov Chains and Stochastic Pro cesses.

We would like to express our deepest gratitude to our guide, Ms. PARVATHI T.S, Department of Mathematics and Statistics, St.Teresa's College, for their valuable guidance, encouragement, and constructive feedback throughout the development of this work. Their scholarly advice and constant supervision have been a source of inspiration and have greatly enriched our project.

We also extend our sincere thanks to the faculty members and staff of the Department of Mathematics and Statistics, St. Teresa's College, for their support and for providing the necessary academic resources that facilitated the completion of this project.

Lastly, we wish to convey our heartfelt appreciation to our parents, classmates, and friends for their constant encouragement, understanding, and moral support during the preparation of this project.

ERNAKULAM.
30.09.2025

ANEETAANTONY

Contents

Chapter 1

INTRODUCTION

1.1 Introduction

In the study of systems that evolve over time under the influence of chance, the idea of randomness plays a key role. Over the past century, especially during the 20th century, scientists and mathematicians began shifting from purely deter- ministic approaches to probabilistic models to better capture the complexity of real-world processes. One of the major developments in this shift was the evolu- tion of stochastic processes, which describe systems that change in unpredictable but measurable ways. Among the various classes of stochastic processes, Markov Chains hold a central position due to their simplicity and wide applicability. The concept was introduced by Andrei Andreyevich Markov, a Russian mathematician, in the early 1900s. He first applied this method to the analysis of patterns in Russian poetry to study sequences of vowels and consonants, thereby laying the foundation for what would become a major tool in both theoretical and applied probability. Since then, Markov Chains have become a powerful framework for modeling systems where the next step depends only on the current state and not the path taken to get there. This idea—known as the Markov property—has enabled researchers across disciplines to simplify complex systems into manageable, mathematical forms. Over the decades, Markov Chains have been extended and applied in fields as di- verse as physics, economics, operations research, computer science, and healthcare. In particular, they are widely used in the modeling of disease progression, where patients may move between various health states over time, such as in chronic conditions like hypertension, diabetes, or cancer.

1.2 Aim

The aim of this project is to explore the mathematical foundation and theoretical aspects of Markov Chains, with a focus on understanding their structure, properties, and classification. As an application, we will consider how these models can be used to study disease progression through synthetic patient data. The core of this study is to find, using mathematical and statistical methods, the relevance of Markov Chains in modeling the progression of health. The disease application will serve as an example to illustrate the practical relevance of Markov Chains, while our main focus remains on the mathematical theory. By the end of the project, we aim to not only understand the behavior of Markov Chains but also to demonstrate their usefulness in solving real-life problems where uncertainty and randomness are involved.

Chapter 2

BASIC CONCEPTS IN STOCHASTIC PROCESS

2.1 Introduction

Stochastic processes are used to study how systems evolve over time when uncertainty or randomness is involved. While a random variable gives us a single random outcome, a stochastic process extends this idea to multiple time points. That means instead of just asking ”What is the result of one random event?”, we ask ”How do random results change over time?” Imagine that we observe the value of a variable at different times (like a person’s heart rate every hour). At each moment, the value can change unpredictably; that is a stochastic process. It gives us a family of random variables, each tied to a specific point in time. Formally, we can write a stochastic process as:

$$\{X(t) : t \in T\}$$

- T is the set of time points (like minutes or days)
- Each X(t) is a random variable.

Example

Coin Toss Sequence:

Imagine tossing a coin every second.

Let Xt = 1 if the toss at time t is Heads, and 0 if Tails.

The sequence Xt is a stochastic process (discrete time, discrete state).

Gaussian stochastic processes

A Gaussian process isastochasticprocessforwhichE=RwhereEisthe State space & R is the set of Real numbers and all finite-dimensional distributions are Gaussian, that is, every finite-dimensional vector (Xt1, Xt2, . . . , Xt_k) is a multivariate normal random variable N (μk, Kk), for some vector μk and a non-negative symmetric definite matrix Kk, for all k = 1, 2, . . . and for all t1, t2, . . . , tk.

Examples

Daily Temperature with Gaussian Distribution

Suppose the noon temperature each day is modeled as (mean 30°C, variance 4).

Over days, the sequence of temperatures is a Gaussian process.

2.2 Classification of Stochastic Process.

Stochastic processes are classified on the basis of two main features:

1. Time parameter (T):This refers to the set over which the process is observed: If time is discrete (like 1, 2, 3,...), it is a discrete-time stochastic process. If time is continuous (like all real numbers), it is a continuous-time stochastic pro cess.

2. State Space (E):This refers to the set of possible values the process can take: If the state space is discrete, the process is called a discrete-state stochastic process.If the state space is continuous, it is called a continuous-state stochastic process.

2.3 Types of Stochastic Process

- Stationary Stochastic Processes

  A stationary stochastic process is one whose statistical characteristics, such as mean, variance, and autocorrelation, do not change over time. This implies that the behavior of the process remains consistent regard- less of when it is observed. Stationary simplifies analysis, especially when working with time series data, because the process behaves in a predictable manner over time.

  Types of stationary processes

  (a) Strictly stationary processes

  A stochastic process is called (strictly) stationary if all finite-dimensional distributions are invariant under time translation: for any integer k and times $\{t_i \in T\}$, the distribution of $(X(t_1), X(t_2), \ldots, X(t_k))$ is equal to that of $(X(s + t_1), X(s + t_2), \ldots, X(s + t_k))$ for any s such that $\{s+t_i \in T\}$ for all $i \in \{1, \ldots, k\}$.

  In other words,

  $$P(X_{t_1+s} \in A_1, X_{t_2+s} \in A_2, \ldots, X_{t_k+s} \in A_k) = P(X_{t_1} \in A_1, X_{t_2} \in A_2, \ldots, X_{t_k} \in A_k), \quad \forall s \in T.$$

  (b) Second order or Wide-Sense Stationary (WSS) or Weakly Stationary

  A stochastic process $X_t \in L^2$ (space of square integrable randome variables) is called second-order stationary, wide-sense stationary or weakly stationary if the first moment $E(X_t)$ is a constant and the covariance function $E[E(X_t - \mu)(X_s - \mu)]$ depends only on the difference $t - s$:

  $$E(X_t) = \mu,$$

  $$E[(X_t - \mu)(X_s - \mu)] = C(t - s)$$

- Ergodic Stochastic Process

  A stochastic process X_t is ergodic if the time averages of the process (computed along a single realization or sample path) are equal to the ensemble averages (statistical expectations of the probability distribution).

If you watch one system (one sequence of outcomes) for a long time, the average result you see will be the same as the average result across many systems observed at one moment.

Let X_t, $t \in T$ be a stationary process with mean $\mu = E[X_t]$

Ergodicity in Mean

The process is mean ergodic if:

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T X_t \, dt = \mu \qquad \text{(with probability 1)}$$

For a discrete-time process X_n:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} X_n = E[X_n]$$

Ergodicity in Covariance

A stationary process X_t is covariance ergodic if:

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T X_t X_{t+\tau} \, dt = E[X_t X_{t+\tau}]$$

for all lags τ.

2.4 QueingSystem

Queuing theory uses stochastic processes to model the random behavior of arrivals, waiting, and service in queues. Stochastic processes let us predict and analyze queue performance, since queue lengths and wait times change randomly over time due to unpredictable arrivals and service times

A queuing system is a mathematical model that represents the random process of entities (customers, patients, jobs, packets, etc.) arriving, waiting, and being served by one or more servers. Since both arrivals and services are random, queuing systems are studied using stochastic pro cesses.

A typical queuing model includes:

– Arrival process: pattern of arrivals (e.g., Poisson).

– Service process: distribution of service times (e.g., exponential).

– Number of servers: single or multiple.

– System capacity: finite or infinite.

– Queue discipline: FIFO, LIFO, priority, etc.

These models help analyze performance measures such as average waiting time, server utilization, queue length, and probability of congestion.

Notation (Kendall's Representation)
Queuing systems are often described using Kendall's notation:

$$A/S/c/K/N/SD$$

Where:

– A – interarrival time distribution (M = Poisson/Markovian, G = general, D = deterministic).

– S – service time distribution.

– c – number of servers.

– K – maximum system capacity (including those in service).

– N – size of customer population (often ∞).

– SD – service discipline (e.g., FIFO = first-in-first-out).

Examples:

– M/M/1: Poisson arrivals, exponential service, 1 server, infinite capacity.

– M/M/1/K: Single server with finite capacity K.

(a) The M/M/1 Queue (Single Server with infinite capacity)

Assumptions:

– Poisson arrivals with rate λ.

– Exponential service times with rate μ.

– FIFO discipline, infinite capacity.

Steady-State Distribution: For stability, $\rho = \frac{\lambda}{\mu} < 1$,

$$\pi_n = (1 - \rho)\rho^n, \quad n = 0, 1, 2, \ldots$$

(b) M/M/1/K Queues (single server with finite capacity)

If the system has finite capacity K (including the customer in service), only states 0, 1, . . . , K are possible.

The steady-state distribution is:

$$\pi_n = \frac{1 - \rho}{1 - \rho^{K+1}} \rho^n \quad n = 0, 1, \ldots, K$$

(for $\rho \neq 1$).

Remark: Blocking probability (probability that a new arrival is denied entry):

$$P_{block} = \pi_K$$

This model is useful for systems with limited space such as hospital wards, call centers, and network buffers.

Applications of Queuing Systems

Queuing models are widely used in real-life systems where demand and resources are random:

- Healthcare: Managing patient flow in hospitals and emergency departments.
- Telecommunications: Modeling packet traffic in routers and switches.
- Call Centers: Predicting waiting times and staffing (Erlang formulas).
- Manufacturing: Scheduling jobs and reducing machine idle times.
- Transportation: Traffic flow, ticket counters, airport check-ins.

By representing these systems as stochastic processes, decision-makers can estimate performance, optimize resource allocation, and reduce congestion.

2.5 Some Models Of Stochastic Process

i. Counting Process

A counting process is a stochastic process that represents the total number of events that have occurred up to a certain time.

Definition: $\{N(t), t \geq 0\}$ is a counting process if:

1. $N(t) \geq 0, \forall t \geq 0$,
2. $N(t) \in \{0, 1, 2, \ldots\}$,
3. $N(t)$ is non-decreasing in t,
4. $N(0) = 0$ (no events at time 0).

Example:

– Patients arriving at a hospital

– Cars entering a toll booth

– Calls received by a call center

ii. Martingales

A Martingale is a special type of stochastic process where the expected future value of the process, given all past information, is equal to the present value.

In simple words: A martingale is a "fair game" — there's no predictable gain or loss over time based on past events.

A process $X_t, t \geq 0$ is a martingale if:

A. $[E[|X_t|] < \infty]$ for all $t \geq 0]$

B. $[E[X_{t+1} \mid X_0, X_1, \ldots, X_t] = X_t]$ for all $t \geq 0$

C. X_t is adapted (value at time t depends only on information up to time t)

D. X_0 is known (often zero or a constant)

Example:

i. Fair coin tossing: Let S_n be the net winnings after n tosses, with [+1] for heads, [-1] for tails. S_n is a martingale (fair game).

ii. Random Walk of Drunk Person

(c) Independent Increment Process

An independent increment process is a stochastic process in which the increments (differences in values over non-overlapping intervals) are independent of each other.

Definition $\{X(t), t \geq 0\}$ is an independent increment process if:

i. $X(0) = 0$ (starts at zero)

ii. For any $0 \leq t1 < t2 < \cdots < tn$, the increments $X(t2) - X(t1), X(t3) - X(t2), \ldots, X(tn) - X(tn-1)$ are independent

iii. The process can have both positive and negative increments (unlike counting processes)

iv. Often used to model random movements such as the position of a particle over time

Key Idea:

The behavior of the process in one time interval does not affect its behavior in a different, non-overlapping time interval.

Examples of Independent Increment Processes

i. Random Walk with Independent Steps.

ii. Number of phone calls recieved in an hour.

Chapter 3

MARKOV CHAINS

3.1 INTODUCTION

A Markov chain is a type of stochastic process where the probability of moving to the next state depends only on the current state, not on the sequence of events that preceded it. This property is called "memorylessness" or the Markov property.

Definition:

A Markov chain is a sequence of random variables

$$\{X0, X1, X2, . .\}$$

with the property that

$$P(Xn_{+1} = x \mid Xn = xn, Xn_{-1} = x_{n-}1, \ldots, X0 = x0) = P(Xn_{+1} = x \mid Xn = x_n)$$

for all n. This means the future state depends only on the present state, not the past.

The outcomes are called the states of the Markov chain; if Xn, has the outcome j (i.e. Xn = j, the process is said to be at state j at nth trial.

Example:

Coin tossing type:

Suppose we have 2 states:

$$S = \{H, T\}$$

We write the chain as X_n where:

– $X_n = H$ means the n-th toss is a head.

– $X_n = T$ means the n-th toss is a tail.

Notation:

$P(X_{n+1} = H \mid X_n = H) = 0.5$, $P(X_{n+1} = T \mid X_n = H) = 0.5$

$P(X_{n+1} = H \mid X_n = T) = 0.5$, $P(X_{n+1} = T \mid X_n = T) = 0.5$

This shows that the next state depends only on the current state, not on earlier tosses.

3.2 Classification of Chains

i. Irreducible and Reducible Chains

Irreducible Chains: Every state can be reached from every other state, meaning the chain can move between any two states eventually.

Reducible Chains: Some states are not reachable from others; the system can split into isolated subgroups of states.

ii. Periodic and Aperiodic Chains

Periodic Chains: States are revisited at regular intervals only; for example, if a state can only be returned to in multiples of some integer greater than 1.

Aperiodic Chains: All states can be revisited at irregular or variable intervals; this is often needed for steady-state modeling in health contexts.

iii. Transient and Persistent (Recurrent) States

Transient States: These are states that may be left and never returned to, indicating temporary phases in the system (e.g., a temporary illness stage).

Persistent/Recurrent States: These are states the chain keeps returning to, typical of chronic health conditions or long-term outcomes.

iv. Absorbing States

Absorbing Chains: If a chain reaches an absorbing state, it stays there permanently (e.g., recovery, death, or a stable health outcome)

3.3 Transition Probabilities

Transition probabilities are the chances of moving from one state of a system to another in a Markov process.
Suppose the set of states is given by

$$S = \{1, 2, 3, \ldots, n\}.$$

The probability of moving from state i to state j in one step is denoted by

$$P_{ij} = P(X_{t+1} = j \mid X_t = i).$$

This represents that the given system is in state i at time t and the probability moves to state j at time t + 1.
Example: In the Bernoulli coin tossing experiment, the transition probabilitymatrixis $\begin{pmatrix} q & p \\ q & p \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}(q,p) = e(q,p).$

Higher Transistion Probabilities

Chapman-Kolmogorov equation In a Markov chain, the probability of moving from state i to state j in (m + n) steps is equal to the sum of probabilities of reaching an intermediate state k in m steps and then going from k to j in n steps:

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

Example 1:

Consider a 2-state Markov chain with transition probability matrix

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

We want to find the probability of going from state 1 to state 1 in two steps.

$$p_{11}^{(2)} = p_{11} \cdot p_{11} + p_{12} \cdot p_{21}$$

$$= (0.7)(0.7) + (0.3)(0.4) = 0.49 + 0.12 = 0.61.$$

Thus, the probability of returning to state 1 in two steps is 0.61.
In matrix notation:

$$P^{(2)} = P \times P = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} = \begin{bmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{bmatrix}.$$

Properties of Transition Probability Matrix

i. Non-negativity: Probability of Transition between states is always positive.

$$P_{ij} \geq 0 \quad \forall\, i, j$$

ii. Row sum property: For each state i, the probabilities of transitioning to all possible states must sum to 1:

$$\sum_j P_{ij} = 1.$$

iii. Transitionmatrix: The transition probabilities can be arranged in a matrix

$$P = [P_{ij}]_{n \times n},$$

which is called the transition probability matrix.

Example 2: Consider a simple weather model with two states: Sunny (S) and Rainy (R).

- If today is Sunny, then tomorrow will be Sunny with probability 0.8 and Rainy with probability 0.2.
- If today is Rainy, then tomorrow will be Sunny with probability 0.6 and Rainy with probability 0.4.

The transition matrix is

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

3.4 Properties of Markov Chains

Strong Markov Property

A stopping time for a sequence of random variables (X_n) is a random variable. Let N be a stopping time for a Markov chain $\{X_n, n \geq 0\}$ and let A and B be two events (relating to X_n and happening) prior and posterior respectively to N). Then

$$P(B \mid X_N = i, A) = P(B \mid X_N = i).$$

This is called the strong Markov property. It shows that if N is a stopping time for a Markov chain $\{X_n, n \geq 0\}$, then the evolution of the chain starts afresh from the state reached at time N.

The strong Markov property is implied by the Markov property; both properties are equivalent when N is constant (a degenerate random variable).

Every discrete-time Markov chain $\{X_n, n \geq 0\}$ possesses the strong Markov property.

Renewal Property in Markov Chains

For a discrete-time Markov chain $\{X_n\}$, fix a state i.

Let T_k be the time of the k-th return to state i.

The inter-return times $T_k - T_{(k-1)}$ are independent and identically distributed (i.i.d.).

3.5 MarkovChainsasGraphs

Markov chains can be described as graphs. The states of a Markov chain may be represented by the vertices (nodes) of the graph, and one-step transitions between states by directed arcs.

Ifi $\rightarrow$ j, then the vertices i and j are joined by a directed arc with an arrow towards j. The value of pij, which corresponds to the transition probability, can be indicated as the arc weight on the directed arc.

Directed Graph

IfS= {1,2,...,m} is the set of vertices corresponding to the state space of the chain and a is the set of directed arcs between these vertices, then the graph

$$G = (S, a)$$

is the directed graph (or digraph or transition graph) of the chain.

Digraph A digraph such that its arc weights are positive and the sum of the arc weights of all arcs from each node is unity is called a stochastic graph. The digraph (transition graph) of a Markov chain is therefore a stochastic graph.

Path

A path in a digraph is any sequence of arcs in which the final vertex of one arc is the initial vertex of the next arc.

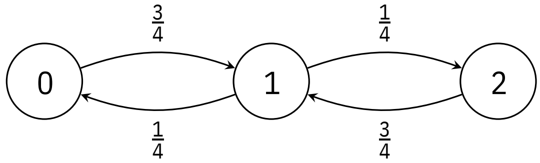

Fig3.1 Transition Graph of the Markov Chain of Example 2

Chapter 4

APPLICATIONS OF MARKOV CHAIN

4.1 DiseaseProgressionandHealthStateMod-eling

Disease progression modeling helps predict how a patient's illness may change over time and what health outcomes are likely. One powerful approach uses Markov chains, which split a disease into simple "states" (like 'early cancer', 'advanced cancer', 'remission', and 'death') and uses transition probabilities to describe how patients move between states from year to year.Markov chains are widely used in medicine because they:

– Only require knowledge of current state to predict future outcomes.

– Can easily incorporate empirical data about patient transitions.

STEP 1: Making Synthetic Data

Example : Transition Data

Current State	Early Cancer	Advanced Cancer	Remission	Death
Early Cancer	60%	30%	5%	5%
Advanced Cancer	0%	70%	10%	20%
Remission	10%	5%	80%	5%
Death	0%	0%	0%	100%

This means, for example, that in one year, 30% of early cancer patients progress to advanced, 5% recover, and 5% die.

STEP 2: Building the Markov Model

Defining States

Let's use four key states:

- – Early Cancer (EC)
- – Advanced Cancer (AC)
- – Remission (RM)
- – Death (D)

The Markov chain's transition matrix represents the above probabilities.

$$
\begin{array}{c} \\ EC \\ AC \\ RM \\ D \end{array}
\begin{array}{c}
\begin{array}{cccc} EC & AC & RM & D \end{array} \\
\begin{pmatrix}
0.60 & 0.30 & 0.05 & 0.05 \\
0.00 & 0.70 & 0.10 & 0.20 \\
0.10 & 0.05 & 0.80 & 0.05 \\
0.00 & 0.00 & 0.00 & 1.00
\end{pmatrix}
\end{array}
$$

STEP 3: Step by Step Stimulation

Year 0 – Initial distribution:

Early Cancer	Advanced Cancer	Remission	Death
1000	0	0	0

Year 1 – Apply matrix:

$$
\begin{aligned}
EC &= 1000 \times 0.60 = 600 \\
AC &= 1000 \times 0.30 = 300 \\
RM &= 1000 \times 0.05 = 50 \\
D &= 1000 \times 0.05 = 50
\end{aligned}
$$

Year 2 – Apply matrix again:

600 EC: 600 × 0.60 = 360 EC, 180 AC, 30 RM, 30 D

300 AC: 210 AC, 30 RM, 60 D

50 RM: 5 EC, 2.5 AC, 40 RM, 2.5 D

50 D: all stay D

Sum the results for each state.

Year 2 Table:

STEP 4: End Results and Interpretation

EC	AC	RM	D
365	392.5	100	142.5

– After a few years, see how patients are distributed across stages.
– For example, after 2 years:
  * About 36% remain in early cancer
  * About 39% progress to advanced cancer
  * About 10% are in remission
  * About 14% have died

  This helps doctors plan better treatments, support systems, or target interventions for those most at risk.

Conclusion

Markov models give doctors and researchers a powerful way to predict and understand disease progression using clear data tables and probabilities. You can easily extend this analysis by simulating more years, comparing treatments, or even switching to other diseases.

4.2 PatientFlow&QueuinginHospitals:Markov Model Analysis with Synthetic Data

Introduction

Managing patient flow in hospitals—especially in clinics, emergency rooms, and outpatient departments—is critical for reducing wait times, improving care, and using resources wisely. The Markovian queue- ing model (especially M/M/1 and M/M/k) helps us mathematically analyze these problems by predicting how patients move through the system, from “arriving” to “being served” to “leaving.”

Theoretical Model

M/M/1 Queue

– ”M” = No.of patients arriving
– ”M” = No.of services available
– ”1” = One doctor or server

Patients arrive at random times; the doctor can serve only one at a time.

States

– 0 patients waiting
– 1 patient waiting
– 2 patients waiting
– ... up to n patients

Creating Synthetic Data

Suppose a small clinic is open 8 hours per day with ONE doctor.

– Patient arrival rate (λ): 4 per hour
– Service rate (μ): 5 per hour

Let’s simulate one day (8 hours):

Synthetic Records

– Each hour: Random number of new arrivals (average = 4/hr)

– Doctor finishes each patient in 12 minutes (average = 5/hr)

Hourly Data Sample (first 4 hours):

Hour	Patients Arrived	Patients Served	Start Waiting	End Waiting
9-10	5	5	0	0
10-11	3	4	0	0
11-12	6	5	1	2
12-1	2	3	2	1

Total for 8 hours

– Arrivals: 32

– Served: 40

(doctor was fast some hours, so queue shrank)

Markov Queue Simulation (M/M/1)

The state at any time is ”number of patients in the system.”

Transitions happen when patients arrive or finish being served

Example Transition Matrix (for up to 3 in system):

Current	Next: 0	Next: 1	Next: 2	Next: 3
0	0.56	0.44	0	0
1	0.56	0	0.44	0
2	0	0.56	0	0.44
3	0	0	0.56	0.44

(Here, numbers are probabilities: Arrival rate/(Arrival+Service rate) and Service rate/(Arrival+Service rate)

Results

Queue Lengths and Waiting Times

Calculation:

– When patients come faster than they are served, the queue grows.

– When service is faster, the queue shrinks.

With the rates above, let's estimate average queue size and waiting time:

– Average queue = ρ2/(1-ρ)] where [ρ = λ/μ = 4/5 = 0.8

– Average queue = 0.82/(1-0.8) = 0.64/0.2 = 3.2 patients

– Average waiting time in system W = 1/(μ - λ) = 1/(5-4) = 1 hour

Visual Table

Hour	Patients Waiting	Avg. Wait Time (min)
9	0	0
10	0	0
11	1	15
12	2	24
...	3	36

Interpretation/Discussion

– Using queues and Markov models, hospital managers see when bottlenecks happen.

– If arrivals increase (more patients per hour) or service slows (doctor takes longer), waiting times increase fast!

– Hospitals can use these results to decide:

  * Should they add more doctors? (Switch to M/M/k with more servers)
  * Change appointment scheduling?
  * Implement fast-track systems for urgent patients?

Conclusion

Markovian queuing models are practical tools for hospital opera- tions. With arrival and service rate data—even synthetically gen- erated—you can forecast patient wait times, optimize staff, and im- prove care. In your simulation above, keeping arrival rate less than service rate keeps the queue short and waiting times low.You can show a table, matrix, and graphs of queue size over hours. You can also present this as proof of modeling with self-generated data—making the results and process fully transparent for any reviewer.Let me know if you'd like simple queue diagrams, further formulas, or additional hourly simulation!

4.3 Epidemiology Modeling and Infection Dynamics

A Markov Chain Approach Using Synthetic DataIntroductionEpidemiology uses mathematical models to understand and predict how diseases spread in a population. Markov chains and compartmental models like SIR (Susceptible-Infected-Recovered) are common frameworks. They divide the population into simple "states" and use probabilities or rates to show how people move from being healthy to sick to recovered. Such models help guide public health actions, forecast needs, and measure intervention effects.

i. Modeling Framework: SIR Model and Markov Chains

Let's use the simple SIR model:

- S (Susceptible): people who can catch the disease
- I (Infected): people who are currently sick and can spread it
- R (Recovered): people who have recovered and no longer spread the disease

Transitions:

- From S to I (catching the disease)
- From I to R (recovering)
- R to S (possible loss of immunity, not included here for simplicity)

ii. Building Synthetic Data

Imagine a small community of 1,000 people. The following is made-up week-by-week infection data:

Week	Susceptible	Infected	Recovered
1	990	10	0
2	946	46	10
3	870	110	20
4	760	190	50
5	650	260	90
6	590	250	160
7	550	200	250
8	520	100	380
9	515	50	435
10	512	15	473

iii. Markov Chain Transition Matrix

Suppose, for each week, the likely transitions (estimated from the above data) are:

– S to I: Probability of catching the disease (depends on infected)

– I to R: Probability of recovering each week

Let's use these synthetic rates:

– Each infected infects 3 new people per week ([$\beta = 0.003$] per S per I)

– Each infected recovers at rate [$\gamma = 0.2$] per week

A simple transition matrix for one week:

$$\begin{pmatrix} S \to S & S \to I & S \to R \\ & I \to I & I \to R \\ & R \to R & \end{pmatrix}$$

For this example:

– S stays S or becomes I – I stays I or becomes R – R stays R (no loss of immunity)

iv. Simulation Process Each week, update:

– Number who become infected: [β S I]

– Number who recover: [γ I]

Example for Week 2:

–=990,I=10,R=0 – New infections: 0.003 × 990 × 10 = 29.7 (30 become infected)

– New recoveries: 0.2 ×10=2

So for next week:

– S = [990 - 30 = 960]

–I=[10+30-2=38]

–R=[0+2=2]

v. Analysis and End Results

The synthetic data shows:

– How quickly infection spreads if parameters don't change

– When cases peak and decline

– How many people remain susceptible, infected, recovered over time

You can plot these results on a graph:

– Curve for S: starts high, drops fast, flattens

– Curve for I: rises sharply, peaks, drops as recoveries rise

– Curve for R: climbs steadily

Key Results:

– Peak infection around weeks 5-6 with nearly 260/1,000 infected

– By week 10, most of the population is either recovered or uninfected

– Simple changes to rates (β, γ), or adding vaccination, dramatically change results

vi. Extensions and Policy Implications

– Add vaccination: move some S straight to R, lower I peak

– Model quarantine: add a new state that reduces S-I contact

– Add birth/death for more realistic population change

Policy Insights:

– Early, rapid intervention lowers I peak

– Late or slow intervention means a bigger, longer outbreak

– Markov/SIR models guide hospital staff, bed planning, lockdown timing, and vaccination needs

Conclusion

Markov chains, when paired with even simple synthetic data, give powerful insight on how real infections behave in populations. Such models form the backbone of infectious disease planning, outbreak prediction, and public health decision-making. You can test ideas, create scenarios, and help leaders choose the best ac- tion—even before an epidemic starts.

4.4 ReliabilityandSurvival in Health Technology

Modeling with Synthetic Data

In healthcare, the reliability of medical devices (like ventilators, defibrillators, and dialysis machines) and the survival time of patients (like post-surgery or after diagnosis) are critical topics. Reliability means how long a device or system works before failing; survival analysis means predicting how long a patient will live or remain healthy after a treatment. Mathematical models—especially using Markov chains and related reliability engineering approaches—give hospitals and manufacturers the tools to:

– Predict device lifespan and replacement needs
– Improve patient outcomes by estimating survival under different treatments
– Manage risk and plan resources.

i. Theoretical Framework: Markov Reliability and Survival Models

Reliability Model

– Focus: Time until device failure, probability device is still functioning at time [t]
– Devices have “working” and “failed” states
– Failure can be exponential (constantly random risk per hour/day) or have stages (degrading performance)

Survival Model

– Tracks proportion of patients surviving after treatment, surgery, or diagnosis
– Commonly uses “survival curves” built from time-to-event (death, relapse, etc.) data
– Markov approach: states could be “healthy”, “sick”, “dead”; transitions based on rates

ii. Creating Synthetic Data

Reliability: Medical Device Example (e.g., Infusion Pumps)

Suppose a hospital uses 100 infusion pumps. Track monthly failures:

Month	NumberWorking	NumberFailed(thatmonth)
0	100	0
1	98	2
2	96	2
3	93	3
4	90	3
5	88	2
6	86	2
7	83	3
8	79	4
9	76	3
10	73	3
11	70	3
12	67	3

Survival: Post-Surgery Patient Example

Suppose we track 50 cancer patients after surgery, annually:

Year	SurvivingPatients	DeathsthatYear
0	50	0
1	47	3
2	43	4
3	38	5
4	34	4
5	30	4

iii. Building the model

Reliability analysis

– Assume exponential failure: Each month, pumps have a probability (p) of failing, estimated as:

$$p = \frac{\text{Number failing in period}}{\text{Number at start of period}}$$

– Survival function (R(t)) : Probability a pump survives to time t:

$$R(t) = e^{-\lambda t}, \quad \text{where } \lambda \text{ is the failure rate per month}$$

From the table above: average monthly failure rate is

$$\frac{3}{100} = 0.03 \text{ or } 3\%.$$

Patient Survival Analysis

– Kaplan-Meier estimator or Markov states (Healthy → Dead)

_ Survival at year t:

$$S(t) = \frac{\text{Number alive at t}}{\text{Number at start}}$$

$$\text{Foryear3: } S(3) = \frac{38}{50} = 0.76(76\%)$$

iv. Results and Curve Visualisations

Reliability Curve

Month	% DevicesSurviving
0	100%
3	93%
6	86%
12	67%

Devices decay gradually, planning for replacements by month 10-12 is smart.

Survival Curve

Year	Survival Percentage
0	100%
1	94%
3	76%
5	60%

Early sharp fall, then plateau—indicates risk is highest initially, later survivors tend to last longer.

5. Discussion and Interpretation

- Device reliability models tells hospitals when to buy spares, schedule replacements, and prevent unexpected failures—crucial for ICUs, surgical suites, dialysis clinics.
- Patient survival analysis informs doctors and families: sets realistic expectations, aids treatment choice, benchmarks hospital performance.
- Improvements (maintenance, new technology, better drugs) can be modeled by changing the failure or death rates and recomputing the survival curves.

Conclusion

Reliability and survival modeling using Markov chains and synthetic (or real) data equips healthcare teams with foresight: keeping equipment functional, improving patient care, and wisely allocating resources. Even simple models and tables reveal risks, budget needs, and life-saving intervention points.

CONCLUSION

This project presented a systematic study of stochastic processes, with special emphasis on Markov chains and their theoretical foundations. Basic concepts, classifications, and important properties of stochastic processes were discussed to build a strong mathematical background. The structure and behavior of Markov chains were explained through transition probabilities, classifications, and graphical representations.

The practical importance of Markov chains was demonstrated through applications in health state modeling, hospital patient flow, epidemiology, and reliability analysis. These applications show how probabilistic models can effectively represent real-life systems that evolve randomly over time. The study highlights that Markov chains provide a simple yet powerful framework for analyzing uncertainty and predicting long-term behavior.

Overall, this project emphasizes the relevance of stochastic modeling in real-world decision-making and establishes Markov chains as an essential tool in applied probability and mathematical modeling.

REFERENCES

[1] Stochastic Processes- J.Medhi

[2] Markov chains - J.R. Norris

[3] Basis of Applied Stochastic Processes- Richard Serfozo

[4] Stochastic Processes And Applications- G.A. Pavliotis

Reshma