Simulation & Modeling

Course Code: BIT534CO
Year/ Semester: III/VI

Compiled by Shishir Ghimire

Credit Hours: 3hrs

Unit 2: Monte Carlo Method

Class Load : 4 hrs

TABLE OF CONTENTS

2. Monte Carlo Method [4 Hrs]

2.1. Monte Carlo Method

2.2. Normally distributed random number

2.3. Monte Carlo Method V/S Stochastic Simulation

▶ What is Monte Carlo ?

Monte Carlo Method > History :

- **1930's:** ***Enrico Fermi*** uses **Monte Carlo** in the calculation of **neutron diffusion**.
- **1940's:** ***Stan Ulam*** while playing solitaire tries to calculate the ***likelihood of winning*** based on the initial layout of the cards.
  - After exhaustive combinatorial calculations, he decided to go for ***practical approach***.
  - He tries many different layouts and observing the number of successful games.
  - He realized that **computers** could be used to solve such problems.
  - ***Stan Ulam*** worked with ***John Von Neumann*** to develop algorithms including ***importance sampling and rejection sampling***.
  - ***Ulam*** and ***Von Neumann*** suggested that aspects of research into **nuclear fission (building Atom Bomb)** at Los Alamos could be aided by use of computer experiments based on chance.

Monte Carlo Method > History :

- The project was top secret so Von Neumann chose the name Monte Carlo in reference to the Casino in Monaco.
- **1950's:** Many papers on Monte Carlo simulation appeared in physics literature. The first major MCMC paper was published by Metropolis and Ulam titled "The Monte Carlo Method" in 1953.
- **1970:** Generalization of the Metropolis algorithm by Hastings which led to development of MCMC (Markov Chain Monte Carlo).
- **1980's:** Important MCMC papers appeared in the fields of computer vision and artificial intelligence but there were few significant publications in the field of statistics
- **1990:** MCMC made the first significant impact in statistics in the work of Gelfand and Smith.

Monte Carlo Simulation:

- The Monte Carlo method is a **numerical method** for **statistical simulation** which utilizes sequences of **random numbers** to perform the simulation.
- Monte Carlo Simulation is a method used to **predict** and **understand** the behaviour of systems involving **uncertainty**. By running multiple simulations with **random inputs**, this technique helps **estimate** possible outcomes and their probabilities.
- Instead of relying on a **single solution**, it generates many **different scenarios** to give a clearer picture of the **range** of possible results.
- Monte Carlo simulations are used to model the **probability of different outcomes** in a process that **cannot easily be predicted** due to the **intervention of random variables**. It is a technique used to understand the impact of risk and uncertainty in prediction and forecasting models.
- **Underlying concept** is to use randomness, **use repeated random sampling.**

Probability and Monte Carlo Simulation:

- The name "Monte Carlo" is inspired by the **Monte Carlo Casino** in Monaco where **randomness** and **chance** are central to the games.
- Similarly, it uses **random numbers** to predict a **range of possible outcomes**.

- Three important characteristics of Monte-Carlo method:
  - Its **input distribution** must be **known.**
  - Its **output** must generate **random samples.**
  - Its **result must be known** while performing an experiment.

What is Monte Carlo Simulation?

- **Monte Carlo methods** are a widely used class of **computational algorithms** for simulating the behavior of various physical and mathematical systems, and for other computations.

- **Monte Carlo algorithm** is often a **numerical Monte Carlo method** used to find solutions to mathematical problems (which may have many variables) that cannot easily be solved, (e.g. integral calculus, or other numerical methods).

- **Monte Carlo methods**, or Monte Carlo experiments, are a broad class of computational algorithms that **rely on repeated random sampling** to obtain numerical results.

Monte Carlo Simulation > Applications:

- <u>**The applications of the Monte Carlo**</u> method include:
  - Simulation of Natural Phenomena
  - Simulation of Experimental Apparatus
  - Numerical Analysis
  - **Integration:** the integration of a complex, discontinuous, high dimension function is hard to give the analytical result.
  - The Simulation of Systems with linear equations
  - Sophisticated Statistical Modeling
  - Probabilistic Graphical Models
  - **Physical Phenomena -** radiation transport in the earth's atmosphere
  - **Communications -** bit error rates

General Algorithm of Monte Carlo Simulation:

- In general Monte Carlo Simulation is roughly **composed of five steps**:

  1. **Set up probability distributions:** Identify the probability distribution that will be used in the simulation.
  2. Build cumulative probability distributions.
  3. Establish an interval of random numbers for each variable.
  4. **Generate random numbers:** Only accept numbers that satisfy a given condition.
  5. Simulate Trials and Analyze

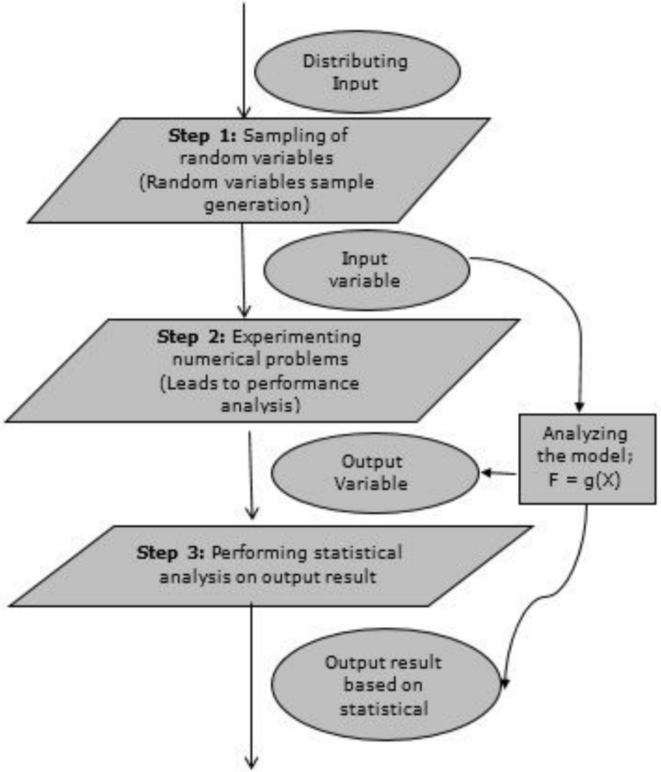
General Algorithm of Monte Carlo Simulation:

1. **Define Problem and Identify Random Variables**
   - Understand the system.
   - Determine **uncertain** or **random variables** (e.g., demand, lead time, prices).
   - Select appropriate **probability distributions** (e.g., normal, uniform, exponential).
2. **Build Cumulative Probability Distributions**
   - **Convert** each probability distribution into a cumulative form.
   - This helps in **linking** random numbers to possible outcomes.

General Algorithm of Monte Carlo Simulation:

3. **Create Random Number Intervals**
   - Divide the **cumulative probability range [0,1]** into **subintervals** corresponding to possible values of the variable.
   - Each interval is associated with a specific outcome.
4. **Generate Random Numbers**
   - Use a **random number generator** to generate a sequence of random numbers.
   - Assign outcomes based on which interval the number falls into.
5. **Simulate Trials and Analyze**
   - **Repeat the process** for many trials.
   - Collect and **analyze** results (mean, variance, etc.) to make data-driven decisions.

General Algorithm of Monte Carlo Simulation:

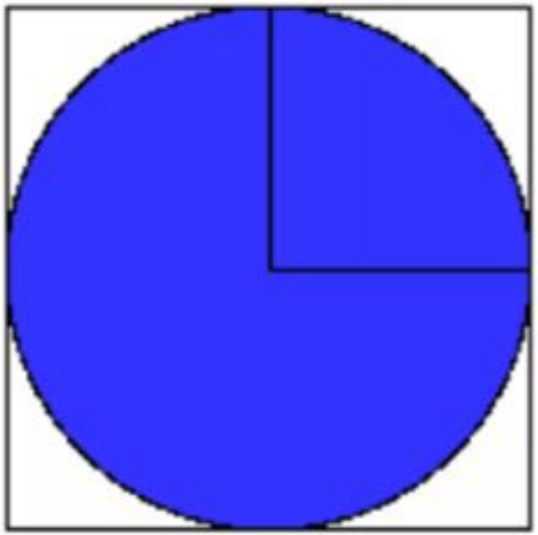


Mathematical Statement of Monte Carlo Method:

Derivation Done in Class

Monte Carlo Simulation > Example :

Example 1: Estimating Value of PI



$$\frac{\#\text{ darts hitting shaded area}}{\#\text{ darts hitting inside square}} = \frac{\text{area of shaded area}}{\text{area of square}}$$

$$\frac{\#\text{ darts hitting shaded area}}{\#\text{ darts hitting inside square}} = \frac{\frac{1}{4}\pi r^2}{r^2} = \frac{1}{4}\pi$$

or

$$\pi = 4\,\frac{\#\text{ darts hitting shaded area}}{\#\text{ darts hitting inside square}}$$

▶ Monte Carlo Simulation > Example :

❖ **Calculation of π:** Consider a quadrant (Circular Sector) inscribed in a unit square. Given that the ratio of their areas is π/4, the value of π can be approximated using a Monte Carlo Method:

1. Draw a square, then inscribe a quadrant within it.
2. Uniformly scatter a given number of points over the square.
3. For each point, we calculate its distance from the origin (0, 0) using the formula: distance = $\sqrt{(x^2 + y^2)}$
4. Count the number of points inside the quadrant (points that satisfy the condition $\sqrt{(x^2 + y^2)} \leq 1$).
5. The ratio of the **inside-count** to the **total-sample-count** is an estimate of the ratio of the two areas, π/4. Multiply the result by 4 to estimate π.

Monte Carlo Simulation > π :

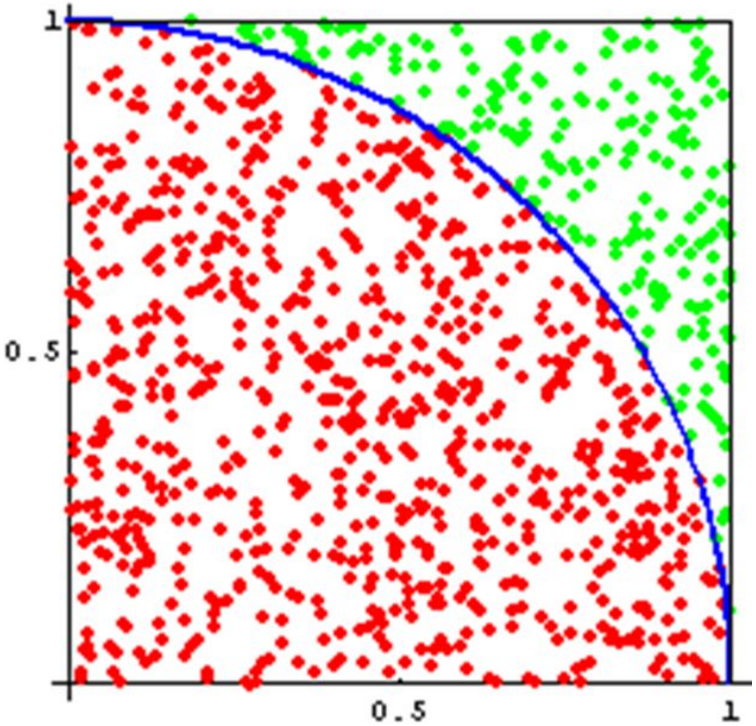
Pseudocode to estimate value of π

1. Start
2. Set spoint = cpoint = 0
3. set value of n
4. Loop for range(n)

    Generate x

    Generate y

5. $d = \sqrt{(x^2 + y^2)}$
6. If $d \leq 1$, then increment cpoint
7. Increment spoint
8. Calculate $\pi = 4 * cpoint/spoint$



Calculating Value of π

Calculation Done in Class

Monte Carlo Simulation > Applications:

- **Optimization of Hyperparameters:** It is used to explore various **hyperparameter configurations** in **machine learning models** such as neural networks, to identify the optimal settings for improved model performance.
- **Reinforcement Learning:** It estimate the value of actions within **reinforcement learning algorithms** helping AI systems learn optimal policies by **simulating and evaluating** different decision-making strategies.
- **Uncertainty Estimation:** It quantifies the impact of **input uncertainties** on AI model predictions. This helps **improve** the model's robustness and reliability by providing insight into how variations in data affect outcomes.
- **Risk Assessment in Financial Models:** AI models predicting financial outcomes benefit from it by evaluating the **probability of various financial scenarios**, enabling better risk assessment and decision-making in uncertain environments.
- **Monte Carlo Tree Search (MCTS):** MCTS is a Monte Carlo-based algorithm is used in **decision-making** for AI in games and other sequential problems, **simulating future states** to guide decision processes and optimize strategies.

Monte Carlo Simulation > Advantages :

- Easy to implement.
- **Risk Analysis:** It helps to visualize the range of **possible outcomes and assess** the probability of different risks.
- **Better Decision-Making:** Provides a clearer picture of potential outcomes, providing more **informed decisions**.
- **Flexibility Across Fields:** Applicable across industries from **finance to engineering to AI**, for various problem types.
- **Handling Complex Systems:** Models systems with **multiple uncertain variables** that traditional methods **can't handle**.
- **Visualization of Uncertainty:** Offers a **distribution of outcomes**, making it easier to understand and plan for uncertainty.
- Can be used for both **stochastic** and **deterministic** problems.

Monte Carlo Simulation > Disadvantages :

- **Time consuming** as there is a need to **generate large number of sampling** to get the desired output.
- The **results** of this method are only the **approximation of true values**, not the **exact**.
- **Computational Intensity:** Requires significant time and resources, especially with **complex models** and many simulations.
- **Quality of Input Data:** The results depend on the **accuracy of the input data**, poor data leads to unreliable outcomes.
- **Convergence Rate:** High accuracy often needs a **large number of simulations**, especially for rare events.
- **Over-Simplification:** May miss complex system details or relationships, leading to **potential oversights**.
- **Interpretation of Results:** Results are **probabilistic** and **misinterpretation** can lead to poor decision-making.

Scenario for Using Monte Carlo Method Over Other Simulation Methods:

The Monte Carlo Method is **preferred** when:

1. **Complex or Unsolvable Mathematical Models:** Analytical solutions are **difficult** or **impossible** (e.g., multi-dimensional integrals, nonlinear systems).
2. **High Uncertainty in Inputs:** Problems involve **random variables** with **known** probability distributions (finance, risk analysis).
3. **Large-Scale Probabilistic Problems:** Systems with many interacting **random factors** (e.g., weather prediction, nuclear physics).
4. **Statistical Estimation:** When estimating probabilities, expected values, or variances where **direct measurement is impractical**.
5. **Optimization in Uncertain Conditions:** Decision-making under **uncertainty**, especially in operations research and engineering design.
6. **Feasibility of Repeated Random Sampling:** When computational power allows millions of trials to approximate accurate results.

Normally Distributed Random Number:

- Normal Distribution is the **most common** or **normal form** of distribution of **Random Variables**, hence the name "normal distribution."
- It is also called the **Gaussian Distribution** in Statistics or Probability. We use this distribution to represent a large number of random variables as it serves as a **foundation** for statistics and probability theory.
- Normal distribution is a **continuous probability distribution** that is symmetric about the mean, depicting that data near the mean are **more frequent** in occurrence than data far from the mean.
- We can draw a Normal Distribution for various types of data that include:
  1. Distribution of Height of People.
  2. Distribution of Errors in any Measurement.
  3. Distribution of Blood Pressure of any Patient, etc.

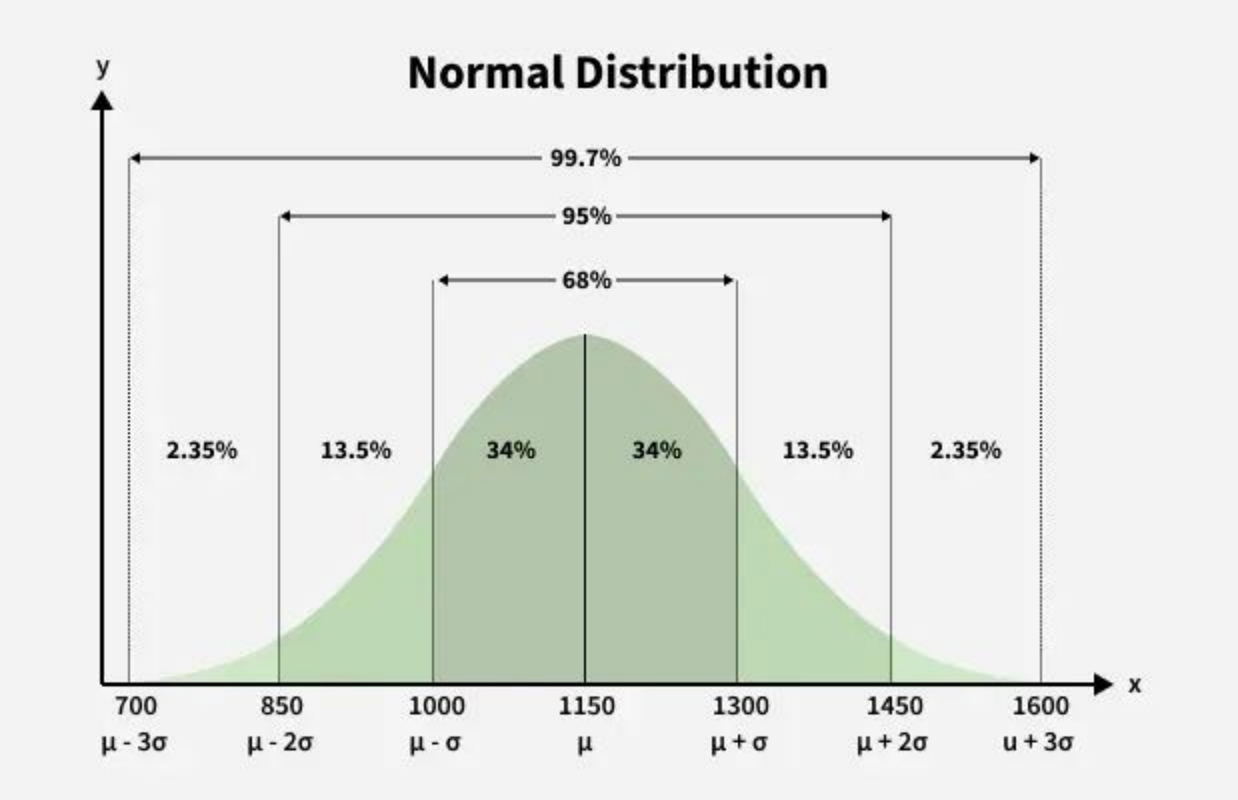
Example: If the average height of men is 175 cm and σ = 7 cm,

- About 68% of men are between 168 cm and 182 cm.
- About 95% are between 161 cm and 189 cm.
- Almost everyone (99.7%) is between 154 cm and 196 cm.

Normally Distributed Random Number:

- It is defined by **two parameters**:
    1. **Mean (μ)** → Center of the distribution.
    2. **Standard Deviation (σ)** → Spread or dispersion.
- This distribution is characterized by a **bell-shaped curve** where values are **clustered around** a central **mean**, with **probabilities decreasing** as you move away from the mean in either direction.
- We define Normal Distribution as the **probability density function** of any continuous random variable for any given system. Now for defining Normal Distribution suppose we take f(x) as the probability density function for any random variable X.
- The area under the curve **(total probability)** of a normal distribution is **always 1**.
- The curve traced by the upper values of the Normal Distribution is in the shape of a Bell, hence Normal Distribution is also called the "**Bell Curve**".

Normally Distributed Random Number:



where,

- The curve is bell-shaped and **symmetric** around μ
- Maximum probability density occurs at **x=μ**.
- As x moves away from μ, the density **decreases** exponentially.

▶ Probability Density Function (PDF):

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

where,

- ❖ x is Random Variable
- ❖ μ is Mean
- ❖ σ is Standard Deviation
- ❖ e is Euler’s number (~2.71828)
- ❖ Π is constant (~3.1416).

Normally Distributed Random Number:

- **Normal Distribution Characteristics:**

1. **Symmetry:** The normal distribution is symmetric **around its mean**. This means the left side of the distribution **mirrors** the right side.
2. **Mean, Median, and Mode:** In a normal distribution, the mean, median, and mode are all **equal** and located at the center of the distribution.
3. **Bell-shaped Curve:** The curve is bell-shaped, indicating that most of the observations **cluster** around the **central peak** and the probabilities for values further away from the mean **taper off equally in both directions**.
4. **Standard Deviation:** The spread of the distribution is determined by the **standard deviation**.
   - About 68% of the data falls within one standard deviation of the mean, ($\mu \pm 1\sigma$)
   - 95% within two standard deviations ($\mu \pm 2\sigma$), and
   - 99.7% within three standard deviations ($\mu \pm 3\sigma$).

Normally Distributed Random Number:

❖ **Methods to Generate Normally Distributed Random Number:**

1. **Box–Muller Transform :**
   - Uses two independent uniform random numbers U_2, U_1 to produce two independent normal random variables.
2. **Central Limit Theorem (CLT) Method:**
   - The sum of many independent uniform random variables tends toward a normal distribution.

Normally Distributed Random Number:

- **Why Generate Normally Distributed Random Numbers?**
  - **Many natural inputs follow** a normal distribution (e.g., customer service times, product dimensions).
  - Randomness in engineering problems often has **Gaussian characteristics.**
  - Statistical quality control often **models defects/errors** as normally distributed.
- **Advantages:**
  - **Matches** many real-world processes.
  - **Well-studied** mathematical properties.
  - **Basis** for parametric statistical analysis.
- **Limitations:**
  - Not suitable for skewed or **heavy-tailed distributions**.
  - **Assumes** symmetrical distribution.
  - Sensitive to **outliers**.

Q 1: Find the probability density function of the normal distribution of the following data. x = 2, μ = 3 and σ = 4.

Solution:

Given,

- *Variable (x) = 2*
- *Mean = 3*
- *Standard Deviation = 4*

Using formula of probability density of normal distribution

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

Simplifying,

f(2, 3, 4) = 0.09666703

Monte Carlo Method vs Stochastic Simulation

- **Stochastic Simulation**
  - Stochastic simulation is a method of **modeling and analyzing systems** whose behavior involves **random variables** or uncertainty.
  - It uses **probability distributions** to represent uncertain elements and simulates how the system evolves over time.
  - **Example:** Simulating a hospital emergency room to predict **patient wait times**, considering random arrivals and treatment durations.
- **Applications:**
  - Queueing systems (banks, call centers).
  - Inventory control and supply chains.
  - Reliability and risk analysis.
  - Manufacturing processes.

🤔 **How is it different from Monte Carlo?**

✅**Monte Carlo Method** is about using randomness **to solve** numerical problems, while **Stochastic Simulation** is about using randomness **to simulate** real-world systems over time.

Monte Carlo Method vs Stochastic Simulation

- **Key Features:**
  - **Randomness Involved:** At least one input or process is probabilistic.
  - **Time-dependent:** Often simulates events in chronological order.
  - **Multiple Runs:** Repeated simulations (replications) are used to obtain reliable results.
  - **Probability Distributions**: Inputs are generated using random numbers and statistical distributions.
- **Advantages:**
  - Captures real-world uncertainty.
  - Can handle complex systems **without** analytical solutions.
  - Provides performance measures like averages, variances, and probabilities.
- **Limitations:**
  - Computationally intensive.
  - Results are approximate and depend on the quality of the model.
  - Requires knowledge of probability and statistics.

Monte Carlo Method vs Stochastic Simulation

Aspect	Monte Carlo Method	Stochastic Simulation
Definition	A simulation technique that uses **random sampling** to estimate numerical results for deterministic or probabilistic problems.	A broader simulation technique that models **systems influenced by random variables or processes** over time.
Purpose	To **approximate mathematical results** (integrals, probabilities, optimization) using randomness.	To **simulate the dynamic behavior of real-world systems** that have inherent randomness.
Focus	Primarily focuses on **numerical estimation** and **probability evaluation**.	Focuses on **modeling system behavior** under uncertainty.
System Type	Often used for **static problems** (no time evolution), but can also be extended to dynamic problems.	Mostly used for **dynamic systems** that evolve over time (e.g., queues, inventory, population growth).
Time Dependency	Usually **time-independent** (sampling does not require a time clock).	**Time-dependent**, simulating events chronologically.
Input	Probability distributions + random number generators.	Probability distributions + event scheduling + system rules.
Output	Estimated numerical value(s) with associated statistical error.	Detailed system performance measures (e.g., waiting times, stock levels, throughput).
Examples	• Estimating the value of π using random points • Risk analysis in finance. • Integration in physics problems.	• Simulating a bank's customer queue. • Modeling traffic flow. • Predicting inventory shortages.

PYQs:

1. Describe Monte Carlo Simulation and state the scenario of using Monte Carlo Method over other simulation methods. **2024 PU (8)**
2. Estimate the value of Pi using Monte carlo Method. Consider the radius of circle as 1 unit. **2025 PU (8)**

References:

1. Jerry Banks, John S. Carson II, Barry L. Nelson, David M. Nicol – *Discrete-Event System Simulation* – Pearson (2013)
2. Averill M. Law – *Simulation Modeling and Analysis* – McGraw-Hill Education (2014)
3. Geoffrey Gordon – *System Simulation*

THANKS!

Do you have any questions?

theciceerguy@p m.me

9841893035

ghimires.com.np