Probability and Stochastic Processes
Lecture 0: Introduction

Prof. Washim Uddin Mondal
Department of Electrical Engineering
Indian Institute of Technology Kanpur

The Goal of the Course

Main Objective

- To introduce the formal description of probability and random processes, and discuss some of its potential applications.

Questions

- What do you intuitively understand by the terms *probability* and *random outcomes*?
- Why do you think we need to formalize the concept of probability, i.e., why is the intuitive description not sufficient?
- What are some modern applications of probability theory?

A Brief History of Probability

The Birth of Probability: From Gambling to Mathematics

Chevalier De Méré	Blaise Pascal	Pierre De Fermat
		

A Brief History of Probability

How Gambling Inspired Probability Theory

- In the 17th century, gambling was a popular pastime in Europe.
- One famous French gambler, Chevalier de Méré, became curious about the chances of winning certain games of dice.
- Since he was unsure how to calculate these chances, he sought help from his mathematician friend, Blaise Pascal.
- Pascal discussed the problem with another renowned mathematician, Pierre de Fermat. Together, they developed systematic methods for analyzing uncertainty.

The Concept of Randomness

The Source of Randomness

- Until the late 19th century, our worldview was dominated by Newtonian physics, which *deterministically* predicts the future given the current state of the universe.
- The early 20th century saw the emergence of quantum physics, which describes reality as inherently *non-deterministic*.
- Despite this, the quantum effect is predominantly a microscopic phenomenon, and our day-to-day macroscopic world remains more or less deterministic.

Question

- What is the source of randomness in our familiar macroscopic world?

The Concept of Randomness

Randomness in the Macroscopic World

- Which of the following tasks is easier and why?
  - (a) Predicting the occurrences of solar eclipses at Kanpur for the next 1000 years.
  - (b) Predicting the occurrences of rainfalls at Kanpur for the next 5 years.
- Solar eclipses can be predicted based on a few unknowns (the masses and the current locations of the planets and the sun), all of which can be precisely measured.
- Weather forecasting depends on thousands of unknowns (urbanization, solar activity, geopolitical events, to name a few), many of which cannot be accurately quantified.
- Randomness can be a useful tool to capture the effects of these unknowns.

Applications of Randomness

Example 1: Finance/Economics

- Economic growth is modeled via the stochastic Solow equation, where stochasticity captures unpredictable shocks to productivity, population growth, etc.
- Stock prices are often modeled using stochastic differential equations driven by Brownian motion.

Example 2: Wireless Communications

- Data transmitted over a wireless channel often gets unpredictably corrupted. This is modeled as an additive white Gaussian noise (AWGN).

Applications of Randomness

Example 3: Cryptography

- Commonly used public key cryptography (such as the RSA system) uses large prime numbers. The fastest prime-number-generation algorithms (such as Miller-Rabin test-based algorithms) are probabilistic.

Example 4: Weather Forecasting

- Using momentum conservation (Navier-Stokes equations), mass conservation (continuity equation), and other relevant fluid dynamics principles, weather evolution can be modeled as a stochastic process.

Applications of Randomness

Example 5: Machine Learning

- The goal of most machine learning algorithms is to minimize a stochastic loss function. This is accomplished via a stochastic gradient descent (SGD) algorithm.
- Most large language models (LLMs) can be viewed as a stochastic sequential predictor, which predicts the most likely next word based on the current context.
- Generative AI models use noise as a part of their core architecture.

Example 6: Epidemiology

- Disease propagation models are inherently probabilistic.

Towards Quantifying Randomness: A Few Definitions

Random Experiment

- An experiment where the possible outcomes are known, but which of these possibilities will be realized when the experiment is conducted is unknown.
  - Example 1: Tossing a coin; the outcomes are $\{H, T\}$.
  - Example 2: Rolling a 6-faced die; the outcomes are $\{1, 2, 3, 4, 5, 6\}$.
  - Example 3: The daily closing price of a stock; the outcomes are $[0, \infty)$.

Towards Quantifying Randomness: A Few Definitions

Sample Space

The collection of *all* possible outcomes of a random experiment is called a sample space.

- For example, $\{H, T\}$ is the sample space for the coin tossing example, and $\{1, \cdots, 6\}$ is the sample space for the die throwing example.

Events

The collection of *some* of the outcomes of a random experiment is called an event.

- For example, $\{2, 4, 6\}$ indicates the event that an even number appears after throwing a 6-sided die. Similarly, $[0, 10]$ denotes the event that the daily closing price of the stock does not exceed 10 units.

The Concept of Probability

Quantifying Randomness

- In a random experiment, not all possible events are equally likely.
  - For example, in a gaseous mixture enclosed in a container, it is possible for the component gases to become locally segregated at some point.
  - However, this is highly unlikely and therefore, never observed in practice.
- Probability assigns values to each event, thereby quantifying their likelihood.

Question

- How to quantitatively define the probability of an event?

Quantifying Probability: Classical Approaches

Frequentist Approach

- In this approach, the probability of the event A is defined to be the long-run relative frequency of the event A. In particular, if $n(A)$ is the number of times the event A occurs when the experiment is run n times, then, assuming the existence of the limit,

$$P(A) = \lim_{n\to\infty} \frac{n(A)}{n}$$

Example

- If a fair coin is tossed 1000 times, approximately 500 heads and 500 tails will occur.

Quantifying Probability: Classical Approaches

Combinatorial Approach

- If the number of outcomes is *finite*, and all outcomes are *equally likely*, then

$$P(A) = \frac{\text{Number of favorable outcomes}}{\text{Total number of equally likely outcomes}}$$

Example

- If a 6-faced *fair* die is thrown, the probability of getting an even number as an outcome is $3/6 = 0.5$.

Drawbacks of Classical Probability

Drawbacks

- The combinatorial approach does not extend to infinite sample spaces.
  - If a number is randomly generated between 0 and 1, what is the probability that it is rational[1]? The combinatorial approach leads to the answer ∞/∞.
- The Frequentist approach cannot be extended to experiments that cannot be repeated.
  - What is the probability that life exists on Mars?
  - What is the probability that a specific political party wins an election?
- Without formalization, ambiguity may seep into apparently concrete questions.

[1] Rational numbers can be expressed as a ratio of two integers.

Bertrand's Paradox

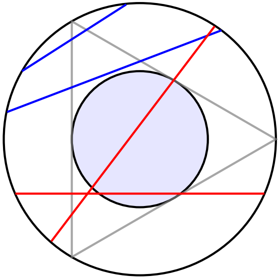
The Problem

Consider two concentric circles with radii $r = 1/2$ and $R = 1$. What is the probability of the event A that a *randomly drawn* chord of the outer circle intersects the inner circle?

Remarks

- Joseph Bertrand posed this problem in his work *Calcul des probabilités* in 1899, which showed that this apparently straightforward question is deeply ambiguous.

Bertrand's Paradox

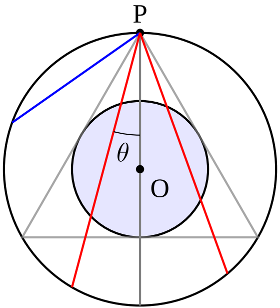


Method 1: Random Midpoint

- **Assumption:** We choose the midpoint of the chord uniformly at random over the area of the outer circle.
- The chord intersects the inner circle if and only if its midpoint lies within the inner circle. Therefore,

$$P(A) = \frac{\text{Area of inner circle}}{\text{Area of outer circle}} = \frac{\pi r^2}{\pi R^2} = \frac{\mathbf{1}}{\mathbf{4}}$$

Bertrand's Paradox

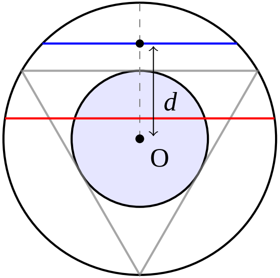


Method 2: Random Angle

- Fix one endpoint of the chord on the outer circle and define θ to be the angle between the chord and the diameter of the outer circle passing through the same endpoint.
- **Assumption:** We choose θ uniformly in $[-\pi/2, \pi/2]$. Clearly, the chord intersects the inner circle if $\theta \in [-\pi/6, \pi/6]$. Hence,

$$P(A) = \frac{\text{Range of Favorable Angles}}{\text{Range of Possible Angles}} = \frac{\pi/3}{\pi} = \mathbf{\frac{1}{3}}$$

Bertrand's Paradox



Method 3: Random Radius

- Fix a radius of the outer circle and pick a point on this radius whose distance from the center O is d. The chord is constructed perpendicular to the radius passing through the chosen point.
- **Assumption:** We choose the distance d uniformly in $[0, R]$. The chord intersects the inner circle if $d \in [0, r]$. Hence,

$$P(A) = \frac{\text{Favorable Length}}{\text{Total Length}} = \frac{r}{R} = \frac{\mathbf{1}}{\mathbf{2}}$$

The Paradox and its Resolution

The Contradiction

Depending on how we define a *random chord*, we obtain three different probabilities for the exact same physical event: **1/4**, **1/3**, and **1/2**.

The Source of Ambiguity and its Resolution

- The term *random chord* does not unambiguously specify the sample space. Are we distributing probability uniformly over an *area*, an *angle* or a *distance*?
- Formalization is needed because two different persons may interpret the probability of the same event in completely different ways.
- Modern probability theory requires a *measurable space* over which the probability values are defined. This avoids ambiguities.

Modern Approach to Probability

Modern Approaches

- Bayesian philosophy models probability as a *degree of belief*, which can be updated if new evidence comes into light (Bayes' Theorem).
- Kolmogorov modeled probability as a *finite measure* defined on a *measurable space*.

What shall we learn in this course?

- We will discuss both Bayesian and Kolmogorov's approaches in detail, with a few lectures dedicated to the combinatorial approach.
- We will learn, among many other things, that not all events are measurable, i.e., cannot be assigned a probability to, and not all infinities are the same.

Interesting Facts You'll Learn

Interesting Facts

- Some infinite sets can be counted, while others are too big to do so.
- Zero probability events may not be impossible, while unit probability events may not be absolutely certain.
- Mathematicians have an insider joke about unconscious statisticians.
- Gambling is a bad idea for all practical purposes.

Course Contents (1/2)

- **Module 1: Foundations of Probability**
  - Set theory, functions, pre-images, and countability
  - Axiomatic definition of probability; σ-algebra, Generated σ-algebra, Borel σ-algebra
  - Conditional probability and independence

- **Module 2: Random Variables**
  - Random variables as measurable functions, induced measure
  - Discrete and continuous random variables
  - Distribution functions (CDF, PDF, PMF), special random variables
  - Functions of random variables, multiple random variables
  - Expectation, variance, correlation, and conditional expectation

Course Contents (2/2)

- **Module 3: Advanced Probability Tools**
  - Moment Generating Functions (MGF)
  - Convergence of a sequence of random variables
  - Law of Large Numbers (LLN) & Central Limit Theorem (CLT)
  - Concentration inequalities: Markov, Chebyshev, and Chernoff
  - Simulation of random variables

- **Module 4: Random Processes**
  - Introduction to random processes; statistical descriptors
  - Stationarity, ergodicity
  - Random walk, Poisson processes
  - Discrete Time Markov Chains (DTMC)