MY LEARNING JOURNEY IN PROBABILITY & STOCHASTIC PROCESS (SEMESTER 1)

24pbd022 | Mauli Patel | SXCA | BDA(2024-2026)

What I Learned – Probability & Stochastic Process

Semester 1 | M.Sc. Big Data Analytics

24pbd022 Mauli Patel

St. Xavier's College, Ahmedabad

INTRODUCTION TO PROBABILITY

A. Core Concepts

- Experiments, Outcomes, Sample Space, Events
- Combinatorial Probability
- Birthday Paradox (brief illustration)
- Principle of Inclusion & Exclusion

B. Conditional Probability & Independence

- Definition and examples
- Understanding independent events

C. Bayes' Theorem

- Concept and formula
- Practical applications (e.g., medical testing, spam detection)

PROBABILITY DISTRIBUTIONS

A. Random Variables

Discrete vs. Continuous Random Variables

B. Key Discrete Probability Distributions

Binomial Distribution: Definition, parameters, applications (e.g., number of successes in fixed trials)
Poisson Distribution: Definition, parameters, applications (e.g., events in a fixed interval)
Practical Application: Visualizing Poisson process events over time (as demonstrated in R)
Geometric Distribution: Definition, applications (e.g., first success)
Hypergeometric Distribution: Definition, applications (e.g., sampling without replacement)

C. Key Continuous Probability Distributions

Normal Distribution: Definition, properties (mean, standard deviation, bell curve)
Practical Application: Histogram and summary statistics for normal distribution (as demonstrated in R)
Exponential Distribution: Definition, applications (e.g., time between events in Poisson process)

STOCHASTIC PROCESSES

A. Introduction to Stochastic Processes

Definition and examples (systems evolving over time with randomness)

B. Markov Chains

Definition and properties (memoryless property)
Classification of States (recurrent, transient, absorbing)
Stationary Distribution
Limit Theorems
Illustrations and applications (e.g., brand switching in Marketing Analytics)

C. Poisson Process

Definition and characteristics (constant rate, independent increments)
Relationship with Exponential distribution (inter-arrival times)
Practical Application: Visualizing Poisson process events over time (as demonstrated in R)

TIME SERIES MODELS

A. Introduction to Time Series Data

Concepts and characteristics of time series data

Importance in various fields

B. Autoregressive (AR) Models

Concept of dependence on past values

Order of AR models

C. Moving Average (MA) Models

Concept of dependence on past error terms

Order of MA models

D. Autoregressive Moving Average (ARMA) Models

Combining AR and MA components

E. Autoregressive Integrated Moving Average (ARIMA) Models

Handling non-stationary data (integration 'I' component)

Illustrations and practical considerations for model selection

Practical Application: Time series plot with trend and seasonality (as demonstrated in R)

PRACTICAL APPLICATIONS AND R IMPLEMENTATION

A. R as a Tool for Probability and Stochastic Processes

Demonstrations of key concepts using R

Generating random numbers from distributions (rnorm, rpois, rexp, etc.)

Calculating probabilities (dnorm, ppois, qexp, etc.)

B. Case Studies / Examples (Briefly touch upon if time permits)

Real-life scenarios where these concepts are applied (e.g., queuing theory, financial modeling, reliability analysis).

PRACTICAL PROJECTS AND R IMPLEMENTATION

- Visualizing Poisson Process Events:
- Normal Distribution Analysis
- Die Roll Frequency Distribution
- Time Series Decomposition

SUMMARY OF KEY LEARNINGS

- We have explored the foundational principles of probability, from basic concepts to advanced theorems like Bayes' Theorem.
- A deep dive into various discrete and continuous probability distributions has equipped us with tools to model diverse real-world phenomena.
- Understanding stochastic processes, particularly Markov Chains and Poisson Processes, has provided insights into systems evolving with randomness over time.
- Finally, we've gained a solid grasp of time series models (AR, MA, ARMA, ARIMA) essential for analyzing and forecasting sequential data.

Thank You

- Questions?
- Let’s connect: LinkedIn | GitHub | Email