

WEEKLY PROGRESS REPORT

Name: Purva Jagdish Salgaonkar

Domain: Data Science & Machine Learning Internship

Date of submission: 24 Dec 2025

WEEK ENDING: 03

I. OVERVIEW:

During Week-3 of the Data Science & Machine Learning Internship, I focused on strengthening my statistical foundation, which is essential for understanding machine learning models and data analysis. The primary activity for this week was studying the book “An Introduction to Probability and Statistics (3rd Edition)” by Vijay K. Rohatgi. This week helped me build a clear understanding of probability theory, statistical distributions, estimation techniques, and hypothesis testing, all of which form the mathematical backbone of data science and machine learning.

II. ACHIEVEMENTS:

a. Book Study: Introduction to Probability and Statistics

I studied the prescribed book thoroughly and covered the following key concepts:

- **Fundamentals of Probability:**
Learned about random experiments, sample space, events, and basic probability laws. These concepts helped in understanding uncertainty and chance in real-life scenarios.
- **Conditional Probability and Bayes' Theorem:**
Understood how probability changes when prior information is available. Bayes' theorem was especially useful in understanding decision-making and predictive analysis.
- **Random Variables and Probability Distributions:**
Studied discrete and continuous random variables along with important distributions such as Binomial, Poisson, and Normal distributions, which are widely used in real-world modeling.

- **Mathematical Expectation and Moments:**
Learned about expectation, variance, standard deviation, moments, and moment generating functions, which help in measuring variability and distribution behavior.
- **Joint Distributions and Independence:**
Understood joint, marginal, and conditional distributions along with the concept of independent random variables.
- **Sampling Theory:**
Learned how samples are drawn from populations and how sampling distributions form the basis of statistical inference.
- **Estimation Theory:**
Studied point estimation, interval estimation, confidence intervals, and methods like maximum likelihood estimation.
- **Testing of Hypotheses:**
Gained clarity on null and alternative hypotheses, level of significance, p-values, and standard tests such as Z-test and t-test.
- **Correlation and Regression:**
Learned how relationships between variables are measured and how regression is used for prediction.
- **Practical Importance of Statistics:**
Understood how probability and statistics are applied in engineering, research, economics, and data science.

b. Conceptual Reinforcement Through Engineering Mathematics

Many concepts studied in this book were already familiar from Engineering Mathematics Semesters 1 to 4, such as probability distributions, hypothesis testing, correlation, and regression. This similarity helped in faster understanding and acted as a strong revision of previously learned topics while deepening conceptual clarity.

c. Learning Experience and Self-Assessment

Reading a 707-page book in four days was initially overwhelming. However, with focused reading and conceptual understanding, the content became manageable. Completing the book within a short time significantly boosted my confidence and strengthened my analytical thinking skills.

III. CHALLENGES:

a. Volume of Content

The book is extensive, covering a wide range of probability and statistics topics in great detail. Completing such a large 707-page book within a limited time

frame required strong time management and focused reading. It was necessary to prioritize core concepts, definitions, and important examples while avoiding excessive time on very detailed derivations.

b. Mathematical Depth

Several topics involved complex mathematical expressions, derivations, and formula-based explanations. Concepts related to probability distributions, expectation, variance, and hypothesis testing required multiple revisions to clearly understand their logic and practical significance. Additional effort was needed to connect formulas with their meanings rather than memorizing them.

c. Concept Integration

Although the theoretical concepts were clearly explained, relating probability and statistics topics to real-world machine learning applications required careful thinking and reflection. Understanding how statistical foundations support model building, prediction, and decision-making in machine learning took additional analysis and revision.

IV. LEARNING RESOURCES:

a. Primary Learning Resource

- Studied the book “An Introduction to Probability and Statistics (3rd Edition)” by Vijay K. Rohatgi
- Gained a strong theoretical foundation in probability and statistics
- Understood concepts through clear definitions and structured explanations
- Learned using detailed examples and mathematical reasoning
- Helped build conceptual clarity required for data science and machine learning

b. Supporting Academic Knowledge

- Applied prior knowledge from Engineering Mathematics Semesters 1 to 4
- Used familiarity with probability, distributions, hypothesis testing, correlation, and regression
- Related statistical concepts to engineering and data science applications
- Strengthened understanding by connecting theory with previously learned mathematics

V. NEXT WEEK'S GOALS

a. Introduction to Machine Learning

In the upcoming week, I will begin learning the fundamentals of Machine Learning. This will include understanding what machine learning is, how it works, and how it differs from traditional programming. I will focus on basic concepts such as types of learning—supervised, unsupervised, and reinforcement learning—and explore real-world applications where machine learning is commonly used.

b. E-Book Study: Machine Learning

I also plan to study the assigned Machine Learning e-book to gain a clear understanding of core algorithms, machine learning workflows, and model fundamentals. This will help me understand how data is prepared, how models are trained and tested, and how performance is evaluated. The e-book study will strengthen my theoretical foundation and prepare me for practical implementation.

c. Quiz Preparation

To assess my understanding, I will prepare for and attempt the “Test Your Knowledge – DS: Quiz-2”. Revising important machine learning concepts and attempting the quiz will help reinforce learning and identify areas that require further improvement.

d. Weekly Documentation

Along with learning activities, I will continue maintaining detailed weekly documentation. I will summarize all learning tasks, concepts covered, and progress made during the week and submit the Week-4 progress report in a clear and structured manner.

VI. ADDITIONAL COMMENTS

Week-3 played an important role in strengthening my statistical foundation for data science and machine learning. Understanding probability and statistics at a deeper level has prepared me to better grasp machine learning algorithms and data-driven decision-making in upcoming weeks.