PMDS502L: Probability And Distribution Models

Course Coordinator: Dr. Jisha Francis

Contents

1 Preface to the course: Probability And Distribution Models : PMDS502L

1.1 Contact of Course Coordinator

Dr. Jisha Francis
PRP 215 E
Assistant Professor,
Department of Mathematics,
School of Advanced Sciences
Vellore Institute of Technology, Vellore.
Email : jishafrancis@vit.ac.in

1.2 Course Details with L-T-P-C

Session   : August - December 2024
Semester   : I
Subject Code   : PMDS502L
Subject Name   : Probability And Distribution Models
Subject Type   : Departmental core subject
Lectures (L)   : 3 hours/week
Credits (C)   : 3

1.3 Time Table

Day	Slot (Time)	Activity
Monday	TB_2 (17:00 - 17:50)	Lecture
Tuesday	B_2 (14:00 - 14:50)	Lecture
Thursday	B_2 (15:00 - 15:50)	Lecture
Office Hours	PRP 215E (Through prior appointment)	Doubts clarification

1.4 Classroom Details

Mode   : Physical (Regular)
Theory Class : GDN121

1.5 Evaluation Criteria

Final Assessment Test   : 40%
Continuous Assessment Test 1   : 15%
Continuous Assessment Test 2   : 15%
Digital Assignment 1   : 10%
Digital Assignment 2   : 10%
Quiz 1   : 10%
Attendance   : Minimum 75% (including medical and duty-leaves) is mandatory
                                   to attendant CAT 1, CAT 2, & FAT exams.

What is the role of the Probability and Distribution Models in Data Science?
Imagine you are navigating a sea of data ? tons of information, patterns, and randomness. Probability and Distribution Models act as our compass, our guide, helping us make sense of this chaotic data landscape. Probability theory provides the foundation for making predictions, while distribution models offer/ a lens through which/we can understand and describe patterns in data.

Course Objectives

1. To incorporate the concepts of probability theory and its applications as the core material in building theoretical ideas along with the practical notion.

2. To integrate the intrinsic ideas of preliminary and advanced distributions to correlate with real-world scenarios.

Course Outcomes

At the end of the course, students will be able to:

1. Develop the problem-solving techniques needed to calculate probability and conditional probability.

2. Describe and construct the probability distribution functions and illustrate the mathematical expectation.

3. Demonstrate the various types of generating functions used in statistics.

4. Apply the commonly used univariate discrete and continuous probability distributions.

5. Illustrate the sampling distributions and their importance in inferential statistics.

Module Breakdown

Module 1: Probability (4 hours)

- Introduction-Random Experiments, Empirical basis of probability, Algebra of events, laws of probability

- Conditional Probability, Independence, Bayes? law

- Application of probability to business and economics

Module 2: Random Variables (7 hours)

- One-dimensional Random variable- Discrete and Continuous

- Distribution functions and their properties

- Bivariate Random Variables- Joint Probability functions, marginal distributions, conditional distribution functions

- Notion of Independence of Random variables

- Functions of random variables: introduction, distribution function technique, transformation technique: one variable, transformation technique: several variables, theory and applications

Module 3: Mathematical Expectation (9 hours)

- Expectation, Variance, and Co-variance of random variables

- Conditional expectation and conditional variance

- Markov, Holder, Jensen, and Chebyshev?s Inequality

- Weak Law of Large numbers, Strong law of large numbers, and Kolmogorov theorem

- Central Limit Theorem

Module 4: Generating Functions (4 hours)

- Moment Generating Function, Characteristic Function, and Probability Generating Function - Properties and Applications

Module 5: Discrete Distributions (8 hours)

- Bernoulli, Binomial, Poisson, Geometric, Hyper-geometric, Negative Binomial, Multinomial, distributions and Discrete Uniform distribution - definition, properties and applications with numerical problems

Module 6: Continuous Distributions (8 hours)

- Uniform, Normal distribution function, Exponential, Gamma, Beta distributions (First and Second kind), Weibull, Cauchy, and Laplace distribution functions - definition, properties and applications

- Concept of truncated distributions

Module 7: Sampling Distributions (3 hours)

- Introduction, The sampling distribution of the Mean: Finite Populations, Sampling distribution of the proportion: Finite Populations, distribution of sample variance

- Chi-square distribution, t- distribution, F distribution, order statistics: properties and applications

Module 8: Contemporary Issues (2 hours)

Total Lecture hours: 45 hours

Text Book(s)

1. Sheldon M. Ross, *A First Course in Probability*, 2020, 10th Edition, Pearson.

2. R.V. Hogg, J. W. McKean, and Allen T. Craig, *An Introduction to Mathematical Statistics*, 2019, 8th Edition, Pearson Education.

Reference Book(s)

1. Rohatgi, V.K. and Ebsanes Saleh, A.K. Md., *An Introduction to Probability and Statistics*, 2002, 2nd Edition, John Wiley & Sons.

2. Krishnamoorthy, K., *Handbook of Statistical Distributions with Applications*, 2006, Chapman & Hall/CRC.

3. Gupta, S.C. and Kapoor V.K., *Fundamentals of Mathematical Statistics*, 2020, Sultan Chand & sons.

4. Maurits Kaptein, ?Edwin van den Heuvel, *Statistics for Data Scientists: An Introduction to Probability and Statistics and Data Analysis*, 2022, Springer.

1.6   FAQs

1. I found Probability And Distribution Models difficult. How should I plan Probability And Distribution Models ?

    (a) **Review Probability And Distribution Models Concepts:** Before starting Probability And Distribution Models , take some time to review the concepts covered in previous semesters. Make sure you have a solid understanding of the fundamentals as they will form the basis for the advanced topics in Probability And Distribution Models .

    (b) **Create a Study Schedule:** Break down the topics in Decision Modelling Techniques into manageable chunks and create a study schedule. Allocate sufficient time for each topic, and set aside regular study sessions to stay on track. This course is of 3 credits with 3 hours of classroom sessions (3 lectures). Thus **every student is expected to do 6 hours** $(3 \times 2 = 6 \; hours)$ **of self study per week for the course**.

    (c) **Use Multiple Learning Resources:** Explore various learning resources, such as textbooks, lecture notes, video tutorials, and online courses. Different resources may provide different perspectives and examples, making it easier to grasp complex concepts.

(d) **Practice with Problems:** Practice solving problems related to each topic in Probability And Distribution Models . Solving numerical and practical problems will help you gain confidence in applying the concepts to real-world scenarios.

(e) **Participate in Group Study:** Join study groups or engage in discussions with classmates who are also studying Probability And Distribution Models . Collaboration can lead to better understanding and help clarify doubts.

(f) **Seek Help Early:** If you encounter difficulties during your study, don't hesitate to seek help from faculty (Dr. Jisha Francis)

(g) **Stay Consistent:** Consistency is key to success. Dedicate regular study time, and avoid last-minute cramming to allow for better comprehension and retention of the material.

(h) **Stay Positive and Persistent:** Believe in your abilities and keep pushing forward. Take care of your physical and mental health, as it directly impacts your learning ability.

2. How the understanding of the subject helps me in my future studies or profession? Understanding probability and distribution is crucial for various aspects of future studies and professional development. Here's how it helps:

- **Data Analysis and Interpretation:** Mastery of probability and statistical distributions enables precise analysis and interpretation of data, a fundamental skill in research, economics, social sciences, and more.

- **Decision Making:** In business and economics, probabilistic models aid in making informed decisions under uncertainty, optimizing strategies, and improving outcomes.

- **Advanced Studies:** A solid foundation in probability theory is essential for advanced studies in fields like statistics, machine learning, artificial intelligence, finance, and actuarial science.

- **Research and Development:** Understanding statistical distributions and their properties facilitates rigorous research design, hypothesis testing, and data-driven innovation.

- **Professional Skills:** Skills in probability and statistics are highly valued in industries such as finance, healthcare, technology, and engineering, enhancing employability and career advancement opportunities.