Statistical Methods and Data Analysis

1.1   Course Number: MA 231

1.2   Contact Hours: 30 (L)   Credits: 08 [LTP: 2-1-0]

1.3   Semester-offered: Even (IVth)

1.4   Prerequisite: Real Analysis & Calculus

1.5   Syllabus Committee Member: Dr. C. Kundu (convener)

2.   **Objective:**

The objective of this course is to develop an understanding of the central statistical models and methods that are widely used in modern applications. A balance is struck between the presentation of the statistical techniques and data analysis including their appropriate use in a variety of practical contexts. The students are given insight into the theoretical foundation of the methods for analyzing data, and learn to perform such analyses in practice by using the software R. Intuitive developments and practical use of the techniques are emphasized rather than theorem/proof developments.

3.   **Course Content:**

Unit-wise distribution of content and number of lectures

Unit	Topics	Sub-topic	Lectures
1	**Random Variables and Some Special Distributions**	Various approaches of probability: classical, frequency and axiomatic, rules on probability, conditional probability, independence, Bayes' theorem. Univariate random variables: discrete and continuous, distribution functions and their properties, probability mass and density functions, expectation & moments, moment generating function & its properties. Multiple random variables: joint distributions, marginal and conditional distributions. Discrete probability distributions: Bernoulli, Binomial, Geometric, Negative Binomial, Hypergeometric and Poisson distribution. Continuous probability distributions: Uniform, Exponential, Gamma, Normal & Log-normal distribution.	10
2	**Descriptive Statistics**	Statistical Terminology: Inferential statistics, population, sample, parameter, statistic, random sample, sampling techniques. Summarizing and Exploring Data: Concept of frequency distribution, measures of central tendency, moments, measures of dispersion/variability, measures of skewness and kurtosis. Introduction to R-software, different aspects of probability distribution and analysis of data in R.	6

3	**Estimation**	Sampling distributions, basic concepts of inference (estimation & hypothesis testing), point estimation & interval estimation.	4
4	**Testing of Hypotheses**	Null and alternate hypothesis, simple & composite hypotheses, critical region, N-P lemma, tests for mean, variance and proportion in one and two sample problems. Chi-square goodness of fit test. Introduction to non-parametric test, Contingency table, test of independence.	5
5	**Regression & Correlation**	Simple linear regression, least squares fit and correlation analysis. Tests for slope & correlation, prediction problem, residual plots. Multiple linear regression. Analysis of Variance. Statistical models in R.	5
		Total	**30**

4. Readings

4.1 Textbooks:
- *Probability and Statistics in Engineering* by Hines, Montgomery, Goldsman & Borror. Wiley Student Edition.
- *An Introduction to Probability and Statistics* by Rohatgi and Saleh. Wiley Student Edition.
- *An Introduction to Mathematical Statistics and its Applications* by Larsen & Marx. Pearson.

4.2 Reference books:
- ✓ *Introduction to Probability Models* by Sheldon M. Ross, Academic Press.
- ✓ *Probability and Statistics* by Spiegel, Schiller and Srinivasan. Tata McGraw-Hill Pub. Co. Ltd.
- ✓ *Introduction to Probability and Statistics* by J. Susan Milton & J.C. Arnold, 4th Ed., Tata McGraw-Hill Pub. Co. Ltd.
- ✓ *Miller and Freund's Probability and Statistics for Engineers* by Johnson/Miller, Pearson Education India.
- ✓ *An Introduction to R*; manual from www.r-project.org

5 Outcome of the Course:

At the end of the course the students will be able to identify the statistical tools and to apply statistical/deterministic approaches to solve an engineering problem having stochastic in nature. Statistical software as introduced may help students to apply statistical techniques to practical problems in the workplace.