College of Engineering Pune

(An Autonomous Institute of Government of Maharashtra)

Department of Mathematics

(MA- 21001) Probability and Statistics for Engineers

T.Y. B. Tech. Semester VI (Civil, Metallurgy, Manufacturing Engineering)

Academic Year 2023-24

Teaching Scheme: Lectures: 2hrs/week + Tutorial: 1 hr / week

Examination Scheme: Continuous evaluation: 40(20+20) marks, End Sem.Exam+oral : 60 marks

(Tb)Text book: Christian Heumann, Michael Schomaker, Shalabh, Introduction to Statistics and Data Analysis, ISBN: 978-3-319-46160-1, DOI: 10.1007/978-3-319-46162-5, Publisher: Springer,Year: 2016.

- Ronald E, Walpole, Sharon L. Myers, Keying Ye, Probability and Statistics for Engineersand Scientists (9^{th} Edition), Pearson Prentice Hall, 2007. (Use this book for some topics as stated in the lesson plan.)

Unit 1: Descriptive Statistics and Basic Probability Theory

Lesson no.	**Topic**	**Section nos. and books to be referred**
1	Introduction and framework: Population, sample, observations, variables: qualitative and quantitative variables, discrete and continuous variables, scales, Frequency measures: Absolute and relative frequencies	Tb Chapter 1, sections: 1.1, 1.2, 1.2.1, 1.2.2, 1.2.3, 1.2.4, 2.1
2	Graphical representation: Bar chart, pie chart and histogram, definitions of mean, median, mode, quantiles, variance, standard deviation with examples (Skip: weighted arithmetic mean and Pg no: 41,53,54,55)	Tb Chapter 2 and 3, sections 2.3.1, 2.3.2, 2.3.3, 3.1.1, 3.1.2, 3.1.4, 3.2.2
3	Basic concepts in Combinatorics : Permutation and combination examples	**Book by Walpole, Chapter 2 on Probability (section 3)**
4	Elements of probability theory: Axiomatic approach, Kolmogorv axioms, conditional probability, Bayes' rule and independence	Tb, Chapter 6, sections: all sections except section 6.2(revise 6.1)
5	Examples	-

Unit II: Some basic probability distributions

Lesson no.	**Topic**	**Section no. of text book**
1	Concept of a random variable (discrete and continuous): CDF of discrete and continuous random variables	Tb,Chapter 7, Section 7.1, 7.2, 7.2.1, 7.2.2
2	Mean or expectation and variance of a random variable, Calculation rules for expectation and variance	7.3.1, 7.3.2, 7.6
3	expectation and variance of arithmetic mean, Some discrete probability distributions: Bernoulli distribution	7.6.1, 8.1.3
4	Binomial distribution, Poisson distribution	8.1.4, 8.1.5
5	Continuous distributions : Normal and Exponential Distribution, Central limit theorem(**Walpole: Page no. 260**)	8.2.2, 8.2.3

Unit III :R software

R commands for possible calculations are covered along with respective theory topics.

Preliminary R software topics are as follows:

Lesson no.	**Topic**	**Students should See lecture no. 1-8 uploaded on R software on Moodle**
1	Overview of R and RStudio, installation,basicsyntax , mathematical operators and functions in R, plotting simple mathematical functions, vectors using sequence and length	-
2	Introduction to data frames in R, constructing data frame using vectors, saving data frame as CSV file, accessing built in data sets in R, creating and saving R script,indexing and slicing data frames	-
3	Performing data analysis: Prerequisites- setting working directory in R studio, merging and importing data, data types and factors, lists and its operations	-
4	Creating visualizations in R: Plotting histogram,Pie chart, bar chart, scatter plot, plots using ‘ggplot2’ package, data manipulation using ‘dplyr package’ and some useful functions in ‘dplyr’	-

Unit IV: Testing of Hypothesis and statistical inference

Lesson no.	Topic	Section no. of text book
1	Sampling distributions: Chi-squared distribution , t-distribution, F- distribution	8.3, 8.3.1, 8.3.2, 8.3.3
2	Inferences: Introduction, properties of estimators: unbiased nature	9.1, 9.2, 9.2.1
3	Hypothesis testing: Introduction, basic definitions, one and two sample problem, P-values	10, 10.1, 10.2, 10.2.1, 10.2.2, 10.2.3, 10.2.4
4	How to conduct statistical test, P-values approach using R, using confidence interval	10.2.5, 10.2.6, 10.2.7
5	Test for the Mean When the Variance is Known (One-Sample Gauss test) (Z-test)	10.3.1
6	Test for the Mean When the Variance is Unknown (One-Sample t-Test), Chi – squared goodness of fit test, F-test **(**SELF STUDY: Sign test and Wilcoxon signed rank test)**	10.3.2, 10.7, (Notes will be provided for F-test, sign test and Wilcoxon signed rank test)

Unit V: Regression Methods

Lesson no.	Topic	Section no. of text book
1	Linear Regression, linear regression model, method of least squares	11.1,11.2
2	Properties of linear regression line, inferences on regression coefficients (Goodness of fit)	11.2.1, 11.3
3	Multiple linear regression	11.6
4	Matrix notation and examples	11.6.1

Unit VI: Engineering applications of statistics

Lesson no.	Topic	Section no. of text book
1-4	Branch specific and teacher specific applications: Marcov Chains, Random Processes, Statistical Quality Control	NA

Topics marked with **are self study topics. Questions based on these topics will be asked in exams.

Reference Books:

- Ross S.M., Introduction to probability and statistics for Engineers and Scientists (8^{th} Edition), Elsevier Academic press, 2014.
- Ronald E, Walpole, Sharon L. Myers, Keying Ye, Probabilty and Statistics for Engineers and Scientists (9^{th} Edition), Pearson Prentice Hall, 2007.
- Tilman M. Davies, The book of R: A first course in Programming and Statistics (1^{st} Edition), No Starch Press, USA, 2016.
- S. P. Gupta, Statistical Methods, S. Chand & Sons, 37^{th} revised edition, 2008.
- Kishor S. Trivedi, Probability and Statistics with Reliability, Queuing and Computer Science Applications (2^{nd} Edition), Wiley Student edition, 2008.
- Stephens L.J., Schaum's outline of statistics for Engineers, Latest edition, 2019.
- The practice of Business Statistics by Manish Sharma and Amit Gupta, Khanna Publishing Company Private Limited, New Delhi, 2014.

References for R Software :

- Norman Matloff, The Art of R Programming - A Tour of Statistical Software Design, (1^{st} Edition), No Starch Press, USA, 2011.
- Sudha Purohit, Sharad Gore, Shailaja Deshmukh, Statistics using R (2^{nd} Edition), Narosa Publications, 2019.
- Randall Pruim, Foundations and Applications of Statistics - An introduction using R (2^{nd} Edition), American Mathematical Society, 2018.
- Hadley Wickham and Garrett Grolemund, R for Data Science: Import, Tidy, transform, Visualize and Model Data, (1^{st} Edition), O'Reilly Publications, 2017.

Outcomes : Students will be able to

1. **Recall and know** basics of probability theory, R software, probability distribution, statistical inference, linear regression.

2. **Understand** concepts of probability, probability distributions, estimation, regression, and use of R software.
3. **Evaluate** probability of compound events, **find** probabilities using standard distributions, **test for** basic statistical inference (t-test, z-test, F-test, χ^2 –test, confidence interval, non parametric tests), **Use** of statistical tables and data sets in R, **solve** problems on simple linear regression.
4. **Prove** theorems / statements**, run** standard programs on R, **solve** problems on multiple regression.
5. **apply** concepts of probability and statistics to various problems including real life problems.

Important Note :

- Two tests T1 and T2 (Each of 20 marks) and end semester examination (60marks) will be conducted .
- 100% attendance is compulsory.