M. C. E. Society's
Abeda Inamdar Senior College
Of Arts, Science and Commerce, Camp, Pune-1
(Autonomous) Affiliated to SavitribaiPhule Pune University
NAAC accredited 'A' Grade

Faculty of Science &Technology

Choice Based Credit System Syllabus
To be implemented from the academic year 2024-2025

S.Y.B.Sc.(Computer Science) Statistics
Semester III (CBCS – Autonomy 2023Pattern)

CourseTitle	Statistics for Data Science -II	
Course Code: 23SBCS31MNB		**No. of Credits:2**
Course Type: Minor		**Total Teaching Hours:30**

Course Objectives	
1.	A student should be able to recall basic concepts and terminology in Statistics and covers basic tools and methods required for data analysis from their studies.
2.	A student able to recall basic concepts in probability, conditional probability and independent events, random variable, mathematical expectation, and different types of distributions.
3.	A student able to design a statistical hypothesis about the real world problem and conduct appropriate test for drawing valid inference.
4.	A student able to recall basic arithmetic operations on vectors and matrices, including inverse, determinants, eigenvalues and eigenvectors of a matrix

Course Outcome	
1.	Calculate the simple linear regression equation for a set of data, principles of linear regression and correlation, including least square method, predicting a particular value of Y for a given value of X and significance of the correlation coefficient
2.	Use discrete and continuous probability distributions, mean and variance, and making decisions, define binomial outcomes and compute the probability of getting X successes in N trials, use Poisson, exponential distributions to solve statistical problems
3.	Calculate inverse of a square matrix, the determinant, transpose of a square matrix, and of eigenvalues and corresponding eigenvectors of a given matrix

Syllabus

	Syllabus	
Unit I	**Regression (for ungrouped data)**	**07**
	1. Concept of linear and nonlinear regression.	01
	2. Principle of least square, fitting straight line by method of least square.	02
	3. Concept of regression coefficients and correlation coefficient.	01
	4. Concept of multiple regressions, Yule's Notations and Fitting of multiple regression planes.	02
	5. Concept of partial regression coefficients, interpretations.	01
Unit II	**Theory of Probability**	**10**
	1. Counting Principles, Permutation, and Combination.	02
	2. Deterministic and non-determination models.	01
	3. Random Experiment, Sample Spaces (Discrete and continuous).	01
	4. Concept of Event.	01
	5. Concept of Probability: classical definition, probability models, axioms of probability, probability of an event, theorems of probability (without proof).	02
	6. Independence of two events, Conditional probability, multiplication theorem $P(A \cap B) = P(A).P(B\|A)$	02
	7. Bayes' theorem (without proof). True positive, false positive and sensitivity	01
Unit III	**Random Variable and Mathematical Expectation**	**07**
	1. Random variable (r.v.), discrete random variable and Continuous random Variable	01
	2. Probability mass function (P.m.f.) of discrete r.v. and cumulative distribution function (c.d.f.) of discrete r.v.	02
	3. Mathematical expectation and variance of discrete r.v.	01
	4. Probability mass function (P.m.f.) of continuous r.v. and cumulative distribution function (c.d.f.) of continuous r.v.	02
	5. Mathematical expectation and variance of continuous r.v.	01
Unit IV	**Standard Discrete Distribution**	**06**
	1. Discrete Uniform Distribution.	01
	2. Binomial Distribution.	02
	3. Geometric Distribution.	01
	4. Poisson Distribution	02

	Suggested Readings
1.	Introduction to Linear Regression Analysis: Douglas C. Montgomery, Elizabeth A. peck, G. Geoffrey Vining, Wiley
2.	Fundamentals of Statistics. Vol I, A.M.Goon, M.K. Gupta, B.Das Gupta, World press

Web Reference:
1. https://open.umn.edu/opentextbooks/textbooks/459
2. https://openstax.org/books/introductory-business-statistics/pages/5-1-properties-of-continuous-probability-density-functions

Course Title	Statistics Practical-II	
Course Code: 23SBCS32MNB		No. of Credits:2
Course Type: Minor		Total Teaching Session:10

Sr. No	Title of the Practical	No. of Practical
1	**Introduction to Pivot Table and Pivot Charts**	1
2	**Lookup and Logical Functions**	1
3	**Regression Using MS-EXCEL**	1
4	**Probability using MS-EXCEL**	1
5	**Random Variables and Distributions Using MS-EXCEL**	2
6	**Problems Based on Unit I :23SBCS31MNB**	1
7	**Problems Based on Unit II : 23SBCS31MNB**	1
8	**Problems Based on Unit III : 23SBCS31MNB**	1
9	**Problems Based on Unit IV : 23SBCS31MNB**	1

Course Title	Statistics for Data Science -III	
Course Code: 23SBCS41MNB		**No. of Credits:2**
Course Type: Minor		**Total Teaching Hours:30**

Syllabus

	Syllabus	
Unit I	**Matrix Algebra**	**08**
	1. Vectors in IRn	02
	2. Matrices and Operations on Matrices	02
	3. Determinant	02
	4. Inverse of a matrix	02
Unit II	**Introduction to Testing of Hypothesis**	**02**
	1. Concepts of population and sample.	01
	2. Definitions: random sample from a probability distribution, parameter, statistic, standard error of estimator.	
	3. Concept of the null hypothesis and alternative hypothesis (Research hypothesis), critical region.	01
	4. Concept level of significance, type I and type II error, one-sided and two-sided tests, a test of hypothesis, p-value.	
Unit III	**Parametric Tests**	**10**
	1. Large Sample Tests.	02
	2. $H_0: \mu = \mu_0$ Vs $H_1: \mu \neq \mu_0$ or $H_1: \mu < \mu_0$ or $H_1: \mu > \mu_0$ (One sided and two-sided tests).	02
	3. $H_0: \mu_1 = \mu_2$ Vs $H_1: \mu_1 \neq \mu_2$ or $H_1: \mu_1 < \mu_2$ or $H_1: \mu_1 > \mu_2$ (One-sided and two-sided tests).	02
	4. $H_0: P = P_0$ Vs $H_1: P \neq P_0$ or $H_1: P < P_0$ or $H_1: P > P_0$ (One sided and two-sided tests).	02
	5. $H_0: P_1 = P_2$ Vs $H_1: P_1 \neq P_2$ or $H_1: P_1 < P_2$ or $H_1: P_1 > P_2$ (One sided and two-sided tests).	02
Unit IV	**Test Based on Chi-square, t distribution & F distribution**	**10**
	1. Tests based on Chi-square distribution, Chi-square test for the goodness of fit, Test for the independence of attributes (mxn and 2x2)	03
	2. Tests based on t –distribution, $H_0: \mu_1 = \mu_2$ Vs $H_1: \mu_1 \neq \mu_2$ or $H_1: \mu_1 < \mu_2$ or $H_1: \mu_1 > \mu_2$ (One-sided and two-sided tests).	03
	3. Paired t-test.	02
	4. Test based on F- distribution, F-test for testing significance of equality of two population variances.	02

	Suggested Readings
1.	Fundamentals of Statistics, Sixth Revised and Enlarged Edition, S.C. Gupta, HimalayaPublishing House.
2.	Linear Algebra and its Applications, David C Lay, Steven R. Lay, Judi J. MacDonald Pearson Publication, 2016, Fifth Edition.

Web Reference:

1. http://math.mit.edu/~gs/linearalgebra/
2. https://openstax.org/books/introductory-business-statistics/pages/5-1-properties-of-continuous-probability-density-functions

Course Title	Statistics Practical-III	
Course Code: 23SBCS42MNB		No. of Credits:2
Course Type: Minor		Total Teaching Session:10

Sr. No	Title of the Practical	No. of Practical
1	**Introduction to R-software**	1
2	**Charts and Plots using R-Software**	1
3	**Basics of Statistics Using R-software**	1
4	**Probability and Distributions Using R-software**	1
5	**Testing of Hypothesis Using R-software**	1
6	**Matrix and Matrix Operations Using R-software**	1
7	**Problems Based on Unit I :23SBCS42MNB**	1
8	**Problems Based on Unit III : 23SBCS42MNB**	2
9	**Problems Based on Unit IV :23SBCS42MNB**	1

Suggested Readings	
1.	Introduction to Statistics and Data Analysis With Exercises, Solutions and Applications in R, Christian Heumann, Michael Schomaker Shalabh.
2.	Using R for Data Analysis and Graphics Introduction, Code and Commentary J H Maindonald Centre for Mathematics and Its Applications, Australian National University.