SOPHIA COLLEGE, (AUTONOMOUS)

Affiliated to

UNIVERSITY OF MUMBAI

Programme: STATISTICS

Programme Code: SBSSTT

F.Y.B.Sc.

2022-23

(Choice Based Credit System with effect from the year 2018-19)

Programme Outline: FYBsc (SEMESTER I)

Course Code	Unit No	Name of the Unit	Credits
SBSSTT101		DESCRIPTIVE STATISTICS-1	2
	1	Types of Data and Data Condensation	
	2	Types of Data and Data Condensation Classification of Data and Measures of central tendency	
	3	Measures of Dispersion, Skewness & Kurtosis	
SBSSTT102		STATISTICAL METHODS -1	2
	1	Elementary Probability Theory	
	2	Concept of Discrete random variable and properties of its probability distribution	
	3	Standard Discrete Distributions	
SBSSTTP1		STATISITCS PRACTICAL	2
	1.	Descriptive Statistics-1 based on the syllabus.	
	2.	Statistical Methods -1 based on the syllabus.	

Programme Outline: FYBsc (SEMESTER II)

Course Code	Unit No	Name of the Unit	Credits
SBSTT201		DESCRIPTIVE STATISTICS-2	2
	1	Correlation and regression analysis	
	2	Time Series	
	3	Index Numbers	
SBSTT202		Statistical Methods -2	2
	1	Continuous random variable and Standard Continuous Distributions.	
	2	Basic Concepts of Sampling and Estimation theory	
	3	Testing of Hypothesis	
SBSTTP2		STATISTICS PRACTICAL	2
	1.	Descriptive Statistics-2 based on the syllabus.	
	2.	Statistical Methods -2 based on the syllabus.	

Preamble:

In the current context, possessing a solid understanding of various statistical concepts has become crucial. Statisticians are constantly in demand in the software, research, industry, and education sectors. The various statistics course syllabi have been designed so that students can become competent in an extensive spectrum of statistical processes at the completion of each course. These techniques can be applied to further research while also applying statistical tools appropriately to a diversity of data sets in order to derive some reliable results.

Different environments require for different applications of statistics. Quantitative results in various areas of research are referred to as statistics. The study of statistics is an important domain of knowledge that focuses on various techniques of collecting, presenting, analyzing, and interpreting data. It is the science of data-driven learning. The subject provides tools to facilitate decision-making in uncertain situations. Decision-making can be rendered simpler by statistics, which measures uncertainties and chance. In addition to building the foundations for the development of essentially every contemporary field, its descriptive and inferential responsibilities offer a variety of unconventional career possibilities, from financial analysis to sports analysis. The main goal of the curriculum is to get students ready to enter into a promising professional life even after graduation.

Large volumes of data have been processed by computers over the past 20 years, and more complex methods of statistical analysis may be applied efficiently resulting in reliable results. Therefore, a number of fields, including agriculture, business, management, economics, finance, insurance, education, biotechnology, and medical science, among others, rely heavily on statistical techniques and procedures.

Statistics can be divided into three broad categories, (1) descriptive statistics, which summarizes and describes data; (2) inferential statistics, that arrives at decisions about the population based on sample; and (3) operations research, that utilizes statistics in the fields of industrial and management.

PROGRAMME OBJECTIVES

PO 1	To teach students methods for effective data collection, organization, and summarization skills as well as analysis and interpretation approaches.
PO 2	Introduce students to regression analysis to model relationships between variables and make predictions.
PO 3	To provide students with a understanding of fundamental concepts - probability, random variables, and distributions.
PO 4	Encourage students to use statistical techniques to solve practical issues and assess the reliability of statistical findings in order to develop their critical thinking abilities.

PROGRAMME SPECIFIC OUTCOMES

PSO 1	The learner will be able to understand the fundamentals of statistics, including the key concepts of probability theory, probability distributions, distribution theory, statistical inference, significance testing, and operations research.
PSO 2	The learner will be able apply the concepts taught in the practicals and will be able to analyse and evaluate data as well as come to reliable conclusions. This will prepare pupils for real-world situations.
PSO 3	Apply statistical, operations research, probability theory, time series, designs of experiments, and other principles to real-world issues.

PSO 4	Know how statistics are used in fields like finance, sociology, science, and economics, among others.

SEMESTER 1

NAME OF THE COURSE	DESCRIPTIVE STATISTICS-1	
CLASS	FYBSC	
COURSE CODE	SBSSTT101	
NUMBER OF CREDITS	2	
NUMBER OF LECTURES PER WEEK	3	
TOTAL NUMBER OF LECTURES PER SEMESTER	45	
EVALUATION METHOD	INTERNAL ASSESSMENT	SEMESTER END EXAMINATION
TOTAL MARKS	50	50
PASSING MARKS	20	20

COURSE OBJECTIVES:

CO 1.	To introduce the techniques of data collection and its presentation.
CO 2.	To emphasize the need for numerical summary measures for data analysis.
CO 3.	To learn to present the data graphically.
CO 4.	To understand and apply the descriptive techniques of statistical analysis to

COURSE LEARNING OUTCOMES:

CLO 1.	The learner will be able to distinguish between different types of scales of the characteristics.
CLO 2.	The learner will be able to compare the different types of data and describe various methods of data collection.
CLO 3.	The learner will be able to construct Univariate and Bivariate frequency distribution , Cumulative frequency distribution.
CLO 4.	The learner will be able to create appropriate graphical representation of the given data.
CLO 5.	The learner will be able to compute and interpret the relation between the qualitative characteristics in the data.
CLO 6.	The learner will be able to comprehend , compute and interpret the measures of central tendency and dispersion.
CLO 7.	The learner will be able to identify the nature of skewness and kurtosis of the data - mathematically & graphically.

UNIT 1	Types of Data and Data Condensation (15 LECTURES)
1.1	Concept of population and sample. Finite ,Infinite population ,Notion of SRS

	SRSWOR and SRSWR
1.2	Types of Characteristics, Different types of scales: nominal, ordinal, interval and ratio
1.3	Collection of Primary data: concept of a questionnaire and a schedule, Secondary data
1.4	Types of data: Qualitative and quantitative data; Time series data and cross section data, discrete and continuous data
1.5	Tabulation and Uni-variate frequency distribution of discrete and continuous variables. Cumulative frequency distribution, Bi-variate frequency distribution
1.6	Dichotomous classification- for two and three attributes, Verification for Consistency and Diagrams, Representation of data using bar diagrams(Simple, Multiple, Segmented and Percentage) , Pie diagram.
1.7	Association of attributes: Yule's coefficient of association Q. Yule's coefficient of Colligation.
UNIT 2	Graphical representation and Measures of central tendency (15 LECTURES)
2.1	Graphical representation of frequency distribution by Histogram, frequency polygon, Cumulative frequency curve. Stem and leaf diagram.
2.2	a)Concept of central tendency of data. Requirements of good measure b) Locational averages: Median, Mode, and Partition Values: Quartiles, Deciles, and Percentiles. c) Mathematical averages Arithmetic mean (Simple, weighted mean, combined mean), Geometric mean, Harmonic mean, d) Empirical relation between mean, median and mode. e) Merits and demerits of using different measures & their applicability.
UNIT 3	Measures of Dispersion, Skewness & Kurtosis (15 LECTURES)
3.1	a)Concept of dispersion. Requirements of good measure. b) Absolute and Relative measures of dispersion: Range, Quartile Deviation, Mean absolute deviation, Standard deviation. c) Variance and Combined variance, raw moments and central moments and relations between them and their properties d) Concept of Skewness and Kurtosis: Measures of Skewness: Karl Pearson's, Bowley's and Coefficient of skewness based on moments. Measure of Kurtosis, e) Box Plot.

References:

- Agarwal B.L.: Basic Statistics, New Age International Ltd.
- Spiegel M.R. : Tehory and Problems of Statistics, Schaum's Publications series, Tata Mc-Graw Hill
- Kothari C.R. : Research Methodology: Wiley Eastern Limited.
- Goon A.M. , Gupta M.K., Dasgupta B. : Fundamentals of Statistics, Volume II: The World Press Private Limited, Calcutta

NAME OF THE COURSE	STATISTICAL METHODS-1
CLASS	FYBSC
COURSE CODE	SBSSTT102

NUMBER OF CREDITS	2	
NUMBER OF LECTURES PER WEEK	3	
TOTAL NUMBER OF LECTURES PER SEMESTER	45	
EVALUATION METHOD	INTERNAL ASSESSMENT	SEMESTER END EXAMINATION
TOTAL MARKS	50	50
PASSING MARKS	20	20

COURSE OBJECTIVES:

CO 1.	To understand the basic concepts of probability and compute probability in various situations.
CO 2.	To learn the various concepts involved in creating the probability distribution of discrete random variables.
CO 3.	To learn the properties of the standard probability distributions of discrete random variables.
CO 4.	To fit appropriate distribution to the given data sets and interpret the results.

COURSE LEARNING OUTCOMES:

CLO 1.	The learner will be able to differentiate between random and non-random experiments.
CLO 2.	The learner will be able to compute the probabilities of various types of events.
CLO 3.	The learner will be able to understand the concept of a discrete random variable and its probability distribution.
CLO 4.	The learner will be able to compute different measures of the probability distribution using techniques involving discrete random variables.
CLO 5.	The learner will be able to apply standard discrete probability distributions to data based on real life situations.

UNIT 1	Elementary Probability Theory : (15 LECTURES)
1.1	Trial, random experiment, sample point and sample space.
1.2	Definition of an event. Operation of events, mutually exclusive and exhaustive events.
1.3	Classical (Mathematical) and Empirical definitions of Probability and their properties.
1.4	Theorems on Addition and Multiplication of probabilities.
1.5	Independence of events, pairwise and mutual independence for three event, Conditional probability
1.6	Bayes theorem and its applications.
UNIT 2	Discrete random variable and properties of probability distribution : (15 LECTURES)
2.1	Random variable. Definition and properties of probability distribution and cumulative distribution function of discrete random variable

2.2	Raw and Central moments (definition only) and their relationship.(upto order four).
2.3	Concepts of Skewness and Kurtosis and their uses.
2.4	Expectation and Variance of a random variable. Theorems on Expectation & Variance.
2.5	Concept of Generating function, Moment Generating function, Cumulant Generating function, Probability generating function - M.G.F. and C.G.F- Definition & Properties
2.6	Joint probability mass function of two discrete random variables, Marginal and conditional distributions. Covariance and Coefficient of Correlation. Independence of two random variables.
UNIT 3	Standard Discrete Probability Distributions (15 LECTURES)
3.1	Discrete Uniform Distribution– Definition, derivation of their mean and variance.
3.2	Bernoulli Distribution, Binomial distribution – Definition and properties, derivation of their mean and variance .
3.3	Poisson distribution – Definition and properties, derivation of their mean and variance. Poisson approximation to Binomial distribution(statement only).
3.4	Hyper geometric distribution- Derivation of their mean and variance, Binomial approximation to hyper geometric distribution(statement only).
3.5	Fitting of distribution.

REFERENCES

- Medhi J. : Statistical Methods, An Introductory Text, Second Edition, New Age International Ltd.
- Agarwal B.L. : Basic Statistics, New Age International Ltd.
- Spiegel M.R. : Theory and Problems of Statistics, Schaum' s Publications series. Tata McGraw-Hill.
- Kothari C.R. : Research Methodology, Wiley Eastern Limited.
- David S. : Elementary Probability, Cambridge University Press.
- Hoel P.G. : Introduction to Mathematical Statistics, Asia Publishing House.
- Hogg R.V. and Tannis E.P. : Probability and Statistical Inference, McMillan Publishing Co. Inc.
- Pitan Jim : Probability, Narosa Publishing House.
- Goon A.M., Gupta M.K., Dasgupta B. : Fundamentals of Statistics, Volume II : The World Press Private Limited, Calcutta.

NAME OF THE COURSE	STATISTICS PRACTICALS	
CLASS	FYBSC	
COURSE CODE	SBSSTTP1	
NUMBER OF CREDITS	2	
NUMBER OF LECTURES PER WEEK	3	
TOTAL NUMBER OF LECTURES PER SEMESTER	45	
EVALUATION METHOD	INTERNAL ASSESSMENT	SEMESTER END EXAMINATION
TOTAL MARKS	20	80
PASSING MARKS	-	40

UNIT 1	DESCRIPTIVE STATISTICS-1 (15 SESSIONS)

1.1	Tabulation
1.2	Attributes
1.3	Classification of Data
1.4.	Diagrammatic representation.
1.5	Measures of central tendency
1.6	Measures of dispersion
1.7	Practical using Excel and R i)Classification of Data and Diagrammatic representation. ii)Measures of central tendency iii)Measures of dispersion
UNIT 2	STATISTICAL METHODS-1 (15 SESSIONS)
2.1	Probability.
2.2	Discrete Random Variables
2.3	Bi-variate Probability Distributions.
2.4	Binomial distribution
2.5	Poisson distribution
2.6	Hyper geometric distribution
2.7	Practicals Using EXCEL and R- Binomial, Poisson, Hyper geometric distribution

ASSESSMENT DETAILS:

Internal Assessment (50 marks)

One written test of 25 mark and one project work of 25 mark.

Semester End Examination – External Assessment (50 marks)

At the end of the semester, Theory examination of 2 hours duration and 50 marks based on the three units shall be held for each course.

Pattern of **Theory question** paper at the end of the semester for **each course**.

Questions	Sub-questions	Maximum marks
Q1	Part A: two theory sub-questions each one is of 6 marks and attempt any one. Part B: Three sub-questions, each one is of 4 marks and attempt any two.	14 each
Q2		
Q3		
Q4	There shall be 3 sub-questions each one is of 4 marks and attempt any 2.	8
Total marks		50

Practical Assessment (for papers with practicals)

- Practical exam will be held on two days. Each session will be of two hours.
- The students are allowed to write the paper if the attendance for practicals is more than 75%
- To appear in the practical exam, students must bring a properly certified journal.

Programme Outline: FYBsc (SEMESTER II)

Course Code	Unit No	Name of the Unit	Credits
SBSSTT201		DESCRIPTIVE STATISTICS--2	2
	1	Correlation and regression analysis	
	2	Time Series	
	3	Index Numbers	
SBSSTT202		STATISTICAL METHODS--2	2
	1	Continuous random variable and Standard Continuous Distributions.	
	2	Basic Concepts of Sampling and Estimation theory	
	3	Testing of Hypothesis	
SBSSTTP2		STATISTICS PRACTICAL	2
	1.	DESCRIPTIVE STATISTICS—2 based on the syllabus.	
	2.	STATISTICAL METHODS -2 based on the syllabus.	

SEMESTER 2

NAME OF THE COURSE	DESCRIPTIVE STATISTICS-2	
CLASS	FYBSC	
COURSE CODE	SBSSTT201	
NUMBER OF CREDITS	2	
NUMBER OF LECTURES PER WEEK	3	
TOTAL NUMBER OF LECTURES PER SEMESTER	45	
EVALUATION METHOD	INTERNAL ASSESSMENT	SEMESTER END EXAMINATION
TOTAL MARKS	25	75
PASSING MARKS	10	30

COURSE OBJECTIVES:

CO 1.	To understand the nature and magnitude of relationship between the quantitative characteristics in the data.
CO 2.	To create suitable mathematical models that best represents the data given.
CO 3.	To enable the learners to understand forecasting techniques to predict trend and seasonal variation in the time series.
CO 4.	To enable the learners to understand the construction of index numbers & its applications in various field.

COURSE LEARNING OUTCOMES:

CLO 1.	The learner will be able to compute the numerical measures to identify the direction and strength of linear relationship between two variables.
CLO 2.	The learner will be able to build a simple linear regression model and interpret

	regression coefficients and coefficient of determination.
CLO 3.	The learner will be able to identify the relevant mathematical model which fits the data
CLO 4.	The learner will be able to identify various components of time series.
CLO 5.	The learner will be able to apply the appropriate methods to evaluate the impact of the different components of time series on the data.
CLO 6.	The learner will be able to comprehend the construction of different index numbers and to apply the methods in different situations.

UNIT 1	Correlation and regression analysis (15 LECTURES)
1.1	Scatter Diagram, Product moment correlation coefficient and its properties.
1.2	Spearman's Rank correlation.(With and without ties)
1.3	Concept of linear regression. Principle of least squares. Fitting a straight line by method of least squares.
1.4	Relation between regression coefficients and correlation coefficient. Concept
1.5	Fitting a quadratic curve by method of least squares.
1.6	Fitting of curves reducible to linear form by transformation.
UNIT 2	Time Series (15 LECTURES)
2.1	Definition of time series and its component. Models of time series.
2.2	Estimation of trend by: i) Freehand curve method ii) method of semi average iii)Method of Moving average iv) Method of least squares(linear trend only)
2.3.	Estimation of seasonal component by i) method of simple average ii) Ratio to moving average iii)Ratio to trend method (iv) Link Relative Method
UNIT 3	Index Numbers (15 LECTURES)
3.1	Index numbers as comparative tool. Stages in the construction of Price Index numbers.
3.2	Fixed base Index Numbers, Chain base Index Numbers. Base shifting, splicing
3.3	Composite & Weighted Index Numbers. Laspeyre's, Paasche's, Marshal-Edgeworth's, Dorbisch & Bowley's and Fisher's Index Numbers formula.
3.4	Quantity Index Numbers and Value Index Numbers Time reversal test, Factor reversal test, Circular test.
3.5	Cost of Living Index Number, Concept of Real Income based on Wholesale Price Index Number , deflating.
3.6	Index numbers as comparative tool. Stages in the construction of Price Index numbers.
3.7	Fixed base Index Numbers, Chain base Index Numbers. Base shifting, splicing

References:

- Agarwal B.L.: Basic Statistics, New Age International Ltd.
- Spiegel M.R. : Tehory and Problems of Statistics, Schaum's Publications series, Tata Mc-Graw Hill
- Kothari C.R. : Research Methodology: Wiley Eastern Limited.

- Goon A.M. , Gupta M.K., Dasgupta B. : Fundamentals of Statistics, Volume II: The World Press Private Limited, Calcutta

NAME OF THE COURSE	STATISTICAL METHODS-2	
CLASS	FYBSC	
COURSE CODE	SBSSTT202	
NUMBER OF CREDITS	2	
NUMBER OF LECTURES PER WEEK	6	
TOTAL NUMBER OF LECTURES PER SEMESTER	45	
EVALUATION METHOD	INTERNAL ASSESSMENT	SEMESTER END EXAMINATION
TOTAL MARKS	50	50
PASSING MARKS	20	20

COURSE OBJECTIVES:

CO 1.	To learn the various concepts involved in creating the probability distribution of continuous random variables..
CO 2.	To learn the properties of the standard probability distributions of continuous random variables.
CO 3.	To understand the significance of the normal distributions and its application in data analysis.
CO 4.	To introduce two branches of Statistical Inferential theory – Estimation theory and Testing of hypothesis.
CO 5.	To assess population characteristics on the basis of sample using estimation and hypothesis testing theory

COURSE LEARNING OUTCOMES:

CLO 1.	The learner will be able to create a probability density function and compute the cumulative distribution function for a continuous random variable.
CLO 2.	The learner will be able to apply the properties of standard continuous probability distributions to different data based on situations.
CLO 3.	The learner will be able to distinguish between point estimation and interval estimation of the population parameters.
CLO 4.	The learner will be able to frame a hypothesis and compute the probabilities of error that could arise while testing.
CLO 5.	The learner will be able to test the hypothesis by examining one or two random samples of the population.

UNIT 1	Continuous random variable and Standard Continuous Distributions (15 LECTURES)
1.1	Concept of Continuous random variable and properties of its probability distribution

1.2	Probability density function and cumulative distribution function – Properties and its graphical representation.
1.3	Expectation of a random variable and its properties. Measures of location, dispersion, skewness and kurtosis. Raw and central moments , M.G.F. and C.G.F- Definition and properties.
1.4	Rectangular Distribution- Derivations of mean, median and variance
1.5	Exponential (location scale parameter) - Derivations of mean, median and variance' memory less property of exponential distribution
UNIT 2	Normal Distribution , Sampling and Estimation theory: (15 LECTURES)
2.1	Gaussian (Normal) distribution- Properties of Normal distribution, Normal approximation to Binomial and Poisson distribution(statement only) , Use of normal tables.
2.2	Concept of Parameter and Statistic, estimator and estimate
2.3	Sampling distribution.
2.4	Concept of bias and standard error of an estimator.
2.5	Central Limit theorem (statement only).
2.6	Sampling distribution of sample mean and sample proportion. (For large sample only)
2.7	Standard errors of sample mean and sample proportion.
2.8	Point and Interval estimate (Confidence interval) of single mean, single proportion from sample of large size.
UNIT 3	Testing of hypothesis (15 LECTURES)
3.1	Concept of Statistical hypothesis
3.2	Null and alternate hypothesis
3.3	Simple and Composite Hypothesis
3.4	Types of errors, Critical region, Level of significance.
3.5	Large sample tests (using central limit theorem) • For testing specified value of population mean • For testing specified value in difference of two means • For testing specified value of population proportion • For testing specified value of difference of population proportion
3.6	Application of Chi-Square Distribution: • Test of Goodness of Fit • Contingency Table • Test of independence in a contingency table and Yates Correction • Derivation of a test statistic for a 2 x 2 contingency table.

References:

- Medhi J. : Statistical Methods, An Introductory Text, Second Edition, New Age International Ltd.
- Agarwal B.L. : Basic Statistics, New Age International Ltd.
- Spiegel M.R. : Theory and Problems of Statistics, Schaum' s Publications series.Tata McGraw-Hill.
- Kothari C.R. : Research Methodology, Wiley Eastern Limited.

- David S. : Elementary Probability, Cambridge University Press.
- Hoel P.G. : Introduction to Mathematical Statistics, Asia Publishing House.
- Hogg R.V. and Tannis E.P. : Probability and Statistical Inference, McMillan Publishing Co. Inc.
- Pitan Jim : Probability, Narosa Publishing House.
- Goon A.M., Gupta M.K., Dasgupta B. : Fundamentals of Statistics, Volume II : The World Press Private Limited, Calcutta.

NAME OF THE COURSE	STATISTICS PRACTICALS	
CLASS	FYBSC	
COURSE CODE	SBSSTTP2	
NUMBER OF CREDITS	2	
NUMBER OF LECTURES PER WEEK	6	
TOTAL NUMBER OF LECTURES PER SEMESTER	45	
EVALUATION METHOD	INTERNAL ASSESSMENT	SEMESTER END EXAMINATION
TOTAL MARKS	20	80
PASSING MARKS	-	40

UNIT 1	DESCRIPTIVE STATISTICS-2 (15 LECTURES)
1.1	Correlation analysis
1.2	Regression analysis
1.3	Fitting of curve
1.4.	Time series
1.5	Index number-I
1.6	Index number-II
1.7	Practical using Excel and R i) Correlation analysis ii) Regression analysis iii) Fitting of curve
UNIT 2	STATISTICAL METHODS-2 (15 LECTURES)
2.1	Continuous Random Variables
2.2	Uniform and Exponential distribution
2.3	Normal Distribution
2.4	Estimation and Sampling Theory
2.5	Testing of Hypothesis
2.6	Test of Significance
2.7	Practicals Using EXCEL and R

ASSESSMENT DETAILS:

Internal Assessment (50 marks)

The IA consists of test/project of 20 marks and class participation of 5 marks

Semester End Examination – External Assessment (50 marks)

At the end of the semester, Theory examination of 2 hours duration and 50 marks based on the three units shall be held for each course.

Pattern of **Theory question** paper at the end of the semester for **<u>each course</u>**.

Questions	Sub-questions	Maximum marks
Q1	Part A: two theory sub-questions each one is of 6 marks and attempt any one. Part B: Three sub-questions, each one is of 4 marks and attempt any two.	14 each
Q2		
Q3		
Q4	There shall be 3 sub-questions each one is of 4 marks and attempt any 2.	8
Total marks		50

Practical Assessment (for papers with practicals)

- Practical exam will be held on two days. Each session will be of two hours.
- The students are allowed to write the paper if the attendance for practicals is more than 75%
- To appear in the practical exam, students must bring a properly certified journal.