Course outline for M.SC Statistics Semester 2nd

S. No.	Course Code	Course Title	Credits	L	T	P	S	Hours per week
1	STA551C	Probability distributions	5	5	0	0	0	5
2	STA552C	Stochastic Models	4	4	0	0	0	4
3	STA553C	Regression Theory & Linear Models	5	5	0	0	0	5
4	STA591C	Statistical Package	3	2	0	2	0	4
5	STA590C	Practical	2	0	0	4	0	4

Course Title: Probability Distributions	L	T	P	S	Semester: 2nd
Course Code: STA551C	**5**	**x**	**x**	**x**	**Max Marks**: 100
Credits: 5					

Course Objective: The main objective is to introduce standard discrete and continuous distributions along with some generalized distributions and order statistics.

Course Outcomes: After completion of this course student will able to:

1. Recognize and apply standard discrete and continuous probability distributions and their properties.
2. Analyze and determine distributions of functions of random variables.
3. Understand the concept of compound, truncated, mixture and sampling distributions.
4. Understand the concept of order statistics, their distributions, properties and applications in real-life situations.
5. Use various probability distributions including generalized and bivariate distributions, to solve practical problems in diverse contexts.

Unit I: Discrete distributions: Binomial, Poisson, multinomial, Geometric, hypergeometric, negative binomial, Logarithmic series distribution; properties and applications.

Unit II: Continuous Distributions: uniform, normal, exponential, gamma, Weibull, Pareto, beta, Burr, lognormal, Laplace, Cauchy distributions; properties and applications.

Unit III: Functions of random variables and their distributions using Jacobian of transformation and other tools; Concept of a sampling distribution; Sampling distributions of t, $\chi 2$ and F (both central and non-central): properties and applications.

Unit IV: Compound, truncated and mixture distributions; Convolutions of two distributions; Order statistics: their distributions and properties; Joint, marginal and conditional distribution of order statistics; Distribution of range and median; Extreme values and their asymptotic distribution (statement only) with applications.

Unit V: Generalized Poisson, generalized negative binomial, generalized geometric and generalized logarithmic series distributions; Generalized power series distributions; Generalized exponential distributions; Bivariate normal and bivariate exponential distributions; Logistic and Log-logistic distributions; Rayleigh distribution: properties and applications.

Textbooks/References:

1. Rohatgi, V. K. (1990): An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern Ltd.
2. Johnson, N. L. and Kotz, S. (1969): Distributions in Statistics; Discrete distributions. John Wiley and Sons, New York.

3. Hogg, R. V. and Craig, A. T. (1989): Introduction to Mathematical Statistics, Macmillan Publishing Company.
4. Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994): Continuous Univariate Distributions-1, 2nd Edition John Wiley and Sons, New York.
5. Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995): Continuous Univariate Distributions-2, 2nd Edition, John Wiley and Sons, New York.

Course Title: Stochastic Models	L	T	P	S	Semester: 2^{nd}
Course Code: STA552C	4	x	x	x	**Max Marks**: 100
Credits: 4					

Course objective: The main objective of this course is to apprise the students about the existence of several stochastic processes in real life situations and to equip them with the techniques to study their statistical behavior as a sequence of dependent random variables.

Course Outcomes: After completion of this course student will able to:

1. Describe and classify various stochastic processes and their key characteristics, including random walks and Markov chains.
2. Apply Markov process concepts, including transition matrices, stationary distributions, and birth-death processes, to model real-life phenomena.
3. Understand and analyze renewal processes, branching processes and Brownian motion, including their properties and applications.
4. Examine and solve queuing models, including M/M/1 and M/M/C systems and compute performance measures such as waiting time distributions and server utilization.
5. Utilize stochastic process techniques to model, interpret and solve practical problems in areas such as reliability, population growth, competing risks and operational systems.

Unit I: Definition and examples of stochastic processes; classification of stochastic processes; random walk and gambler's ruin problems, including elementary problems. Markov chains**:** definition and examples; transition probability matrix; stationary distribution and its interpretation; Chapman–Kolmogorov equation; classification of states; communicating classes; recurrence and non-recurrence; irreducibility.

Unit II: Stationary probability distribution and its applications; computation of n-step transition probability matrix by spectral representation; absorption probability and mean time to absorption. Markov processes: Kolmogorov forward and backward equations; Poisson process; pure birth process (arrival distribution); inter-arrival distribution; death process (departure distribution); birth-death processes; Yule–Furry process (linear growth process); differential-difference and steady-state equations for processes; examples based on these concepts; sickness and marriage models in terms of Markov processes..

Unit III: Renewal processes; renewal function; elementary renewal theorem and its applications. Galton–Watson branching processes: definition and examples of discrete-time branching processes; probability generating function and its properties; offspring mean and probability of extinction; Introduction to Brownian motion process and its basic properties.

Unit IV: Stochastic processes in queuing systems: structure, elements, operating characteristics, classifications and applications. Queuing models: single-server M/M/1 system with infinite (M/M/1/∞) and finite (M/M/1/N) capacity and multi-server (M/M/C) system; Difference equations and steady-state solutions; probabilities of busy and idle periods; measures of effectiveness; waiting time distribution for M/M/1 model. Special cases of M/M/1 queue and numerical illustrations based on the concepts.

Textbooks/References:

1. Bhat, B. R. (2000). Stochastic Models: Analysis and Applications, 2/e, New Age International, India.
2. Medhi, Jyotiprasad (1994): Stochastic Processes, Wiley Eastern Limited, 2/e.
3. Adke, S. R. and Manjunath S. M. (1985). Finite Markov Processes. Wiley Eastern (New Age Publishing).
4. Feller, W. (1972) An Introduction to Probability Theory and its Applications,Vol. 3/e Wiley Eastern Ltd.
5. Kulkarni, Vidyadhar (1995): Modeling and Analysis of Stochastic systems, G. Thomson Science and Professional.

Course Title: Regression Theory and Linear Models	L	T	P	S	Semester: 2nd
Course Code: STA553C	**5**	**x**	**x**	**x**	**Max Marks**: 100
Credits: 5					

Course Objective: The main objective of this course is to provide students with a solid theoretical foundation in Linear Estimation Theory and Regression Analysis.

Course Outcomes: Upon successful completion of this course, the student will be able to:

1. Employ multiple linear regression models for analyzing real-life data sets.
2. Perform statistical tests and construct statistical intervals in a multiple linear regression framework.
3. Validate regression models using graphical and numerical diagnostic procedures.
4. Estimate regression parameters in the presence of multicollinearity.
1. Select and apply appropriate link functions for building regression models.

Unit I: Correlation Analysis – conceptual framework; methods of studying correlation – scatter diagram, Karl Pearson's correlation coefficient, Spearman's rank correlation coefficient and concurrent deviation methods; probable error (ungrouped data); coefficient of determination.

Unit II: Simple regression model with one independent variable (X) – assumptions; estimation of parameters using least squares theory; standard error of estimator; testing of hypotheses about parameters; coefficient of determination and its use to measure the goodness of fit of a linear regression model; prediction of response with confidence limits.

Unit III: Diagnostic checks for suitability and validity of a linear regression model – graphical techniques; tests for normality, un-correlatedness and lack of fit; multiple regression model; standard Gauss–Markov setup; least squares estimation.

Unit IV: Fundamental concepts of Generalized Linear Model (GLM); exponential family of distributions; link functions such as Poisson, binomial, normal, exponential and Gamma; logistic regression.

Unit V: Concepts of deviancy; estimation of parameters of a GLM; suitability of a model by using analysis of deviancy and by examining the significance of parameters; Pearson and deviancy residuals; statistical tests for acceptability of a fitted model – Pearson's chi-square test and likelihood ratio test.

Textbooks/References:

1. Montgomery, Douglas C., Peck, Elizabeth A., Vining, G. Geoffrey (2003) Introduction to Linear Regression Analysis.
2. McCullagh, P. and Nelder, J. A. (1989) Generalized Linear Models (Chapman & Hall).
3. Draper, N. R. and Smith, H. (1998) Applied Regression Analysis, 3rd Ed. (JohnWiley).

4. Ratkowsky, D. A. (1983) Nonlinear Regression Modelling (Marcel Dekker).
5. Neter, J., Wasserman, W., Kutner, M. H. (1985) Applied Linear Statistical Models. (Richard D. Irwin).

Course Title: Statistical Package	**L**	**T**	**P**	**S**	**Semester:** 2nd
Course Code: STA591C	**2**	**x**	**2**	**x**	**Max Marks**: 100
Credits: 3					

Course Objective: To provide students with practical skills in using R and SPSS software for data analysis, including data manipulation, visualization, statistical computation and simulation.

Course Outcomes: After the completion of paper student will able to:

1. Use R as a calculator and as a tool for performing statistical analysis.
2. Apply R and SPSS software for effective data visualization and exploratory data analysis.
3. Perform simulation and generate random data using R software.
4. Conduct descriptive and inferential statistical analyses, including correlation, regression, t-tests and ANOVA using R and SPSS.
5. Prepare, manage and manipulate datasets efficiently in R and SPSS for various types of analyses.

Unit I: Introduction to R – R as a calculator; R data structures, help functions in R, assignment operator, vectors, operations on vectors, setting working directories, importing different data formats (.csv, .xlsx, .sav) into R, sub-setting and writing output; Handling matrices, data frames and lists in R; Introduction to creating functions, calling functions, plots and graphics in R.

Unit II: Descriptive statistics in R; various summary statistics commands; correlation and regression. Apply family of functions in R – apply(), lapply(), sapply() and tapply(); Random data generation in R; creating frequency tables, proportion tables and crosstabs for categorical variables.

Unit III: Introduction to SPSS – variable view and data view; working with data files, SPSS windows, menus and dialogue boxes. Preparing the data file: creating data files, entering data, defining variables, modifying data files and importing files; Descriptive statistics for categorical and continuous variables; checking normality and outliers; Running correlation, simple linear regression and multiple linear regression analyses; Conducting one-sample and two independent-sample t-tests, paired-sample t-test and one-way analysis of variance in SPSS; Graphics and plots in SPSS.

Textbooks/References:

1. Jones, O., Maillardet. R. and Robinson, A. (2014). Introduction to Scientific Programming and Simulation Using R. Chapman & Hall/CRC, The R Series.
2. Matloff, N. (2016). The art of R programming: A tour of statistical software design. No Starch Press

Course Title: Practical	L	T	P	S	Semester: 2^{nd}
Course Code: STA590C	**x**	**x**	**4**	**x**	**Max Marks**: 50
Credits: 2					

Objective:

The main objective is to enhance the practical knowledge of the students in the courses of Regression theory and Probability theory.

Course Outcomes:

1. Learning to perform different regression techniques using R.
2. Learning to perform fitting of probability models on real life data.

Practical Exercises: Practical work will be based on courses STA551C, STA553C and STA591C.