Course outline for M.SC Statistics Semester 3rd

S. No.	Course Code	Course Title	Credits	L	T	P	S	Hours per week
1	STA601C	Statistical Inference	5	5	0	0	0	5
2	STA602C	Survival Analysis	4	4	0	0	0	4
4	STA603C	Multivariate Analysis	4	4	0	0	0	4
5	STA605E	Design of Experiments	4	4	0	0	0	4
6	MTH601E	Real Analysis	4	4	0	0	0	4
7	STA603E	Financial Mathematics	4	4	0	0	0	4
8	STA604E	Financial Reporting	4	4	0	0	0	4
9	STA606C	Python for Statistics	3	2	0	2	0	4

Course Title: Statistical Inference	**L**	**T**	**P**	**S**	**Semester:** 3rd
Course Code: STA601C	**5**	**x**	**x**	**x**	**Max Marks**: 100
Credits: 5					

Course Objectives: To make students aware of estimation (point, as well as interval) and testing (simple, as well as composite hypotheses) procedures.

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

1. Recognize basic concepts of statistical inference.
2. Recall various properties of estimators.
3. Understand Fisher Information, Lower bounds to variance of estimators, MVUE.
4. Apply various estimation and testing procedures to deal with real life problems.
5. Understand Neyman-Pearson theory of Hypothesis Testing.

Unit I: Concept of estimation; estimator and estimate; Criteria of a good estimator: unbiasedness, consistency, efficiency and sufficiency; Fisher-Neyman factorization theorem; Family of distributions admitting sufficient Statistic; Minimal sufficiency and completeness; Ancillary statistics and Basu's theorem; Pitman and exponential families of distributions.

Unit II: Uniformly minimum variance unbiased estimators; Fisher information matrix: illustration with one and two parameters; Chapman Robin's Inequality, Cramer-Rao inequality; Minimum variance bound estimator-examples; Rao-Blackwell and Lehmann-Scheffe theorems; Estimation Methods: Method of moments, Maximum likelihood estimation, Method of minimum Chi-Square, Method of least squares and Interval estimation.

Unit III: Testing of hypothesis: null and alternative, simple and composite; Type I and Type II errors; Test function; size and power function; Concept of p-value; Review of standard one and two-sample significance tests (Z, t, F tests); Most powerful tests; Neyman-Pearson lemma; UMP tests; MLR property.

Unit IV: Likelihood Ratio (LR) Test; Construction of LR tests for normal mean and variance-one and two sample problems; Asymptotic distribution of LR test statistic; Application to an r x c contingency table; Wald test and Rao's score test. Nonparametric tests: Sign test, Mann-Whitney-Wilcoxon test, Wilcoxon signed-rank test, median test, Kruskal-wallis H-test, Chi-square test of goodness of fit, Kolmogorov-Smirnov test, Wald-Wolfowitz run test, Test for randomness.

Textbooks/References:

1. Casella G, Berger R. L. (2001). Statistical Inference, 2/e, Cengage Learning Pvt. Ltd.
2. Rohatgi, V. K. and Saleh, A.K. Md. E. (2001). Introduction to Probability and Statistics, John Wiley & Sons, New York.

3. Kale, B.K. & Muralidharan, K. (2015) Parametric Inference: An Introduction, Alpha Science International Ltd.
4. Lehmann, E.L. and Casella, G. (1998). Theory of Point Estimation. Springer, New York.
5. Rao, C.R. (2002). Linear Statistical Inference and its Applications, 2/e, Wiley.

Course Title: Survival Analysis	**L**	**T**	**P**	**S**	**Semester:** 3rd
Course Code: STA602C	**4**	**x**	**x**	**x**	**Max Marks**: 100
Credits: 4					

Course Objective: The main objective of this paper is to introduce different concepts and applications of survival analysis.

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

1. Understand the need of life time distributions and their properties.
2. Identify the different type of censoring.
3. Estimate parameters in presence of censoring.
4. Implement different parametric and nonparametric estimators for estimating survival function.
5. Analyse the lifetime event data.

Unit I: Survival function, hazard rate, cumulative hazard function, and mean residual life. Parametric models for study of event time data: Exponential, Weibull, extreme value, gamma, Pareto, logistic, log-logistic, normal, log–normal and mixture models-their survival characteristics, Longitudinal studies, Censoring mechanisms- type I, type II and left right and interval censoring, Likelihood function under censoring, Fitting parametric models to survival data with right censoring, Large sample tests with censored data, The E–M algorithm.

Unit II: Actuarial and Kaplan–Meier estimators, Treatment of ties, Self consistency property and asymptotic properties of K–M estimator (statement), Pointwise confidence interval for S(t), Nelson-Aalen estimator of cumulative hazard function and estimation of S(t) based on it, Two–sample methods, Comparison of survival functions: Log rank and Tarone-Ware tests, Competing risks model; Kaplan-Meier estimator of survival function, Nelson-Aalen estimator.

Unit III: Explanatory variables- factors and variates, Cox proportional hazards model, The partial likelihood and estimation of regression coefficients and their standard errors, Breslow's estimator of the baseline hazard function; estimation of cumulative hazard rate and S(t), Statement of asymptotic properties of the estimator, Confidence interval for regression coefficients, Wald, Rao and likelihood tests for β, Accelerated life model, Model selection criteria and comparison of nested models (-2logL and AIC), Using information on prognostic variables in a competing risks model.

Unit IV: Parametric regression-Weibull and Gompertz models, Residuals and model checking under Cox and parametric models, Comparing two survival curves, Hazard plots, Survival plots, Comparing alternative models, AIC criterion, Comparing observed and fitted survival models, Testing proportional hazards hypothesis in the Weibull model of cumulative incidence function.

Textbooks/References:

1. Klien, J.P. and Moeschberger, M.L. (2003). Survival Analysis: Techniques for censored and Trun-cated Data. 2/e. Springer.
2. Kalbfleisch, J.D. and Prentice, R. L (2002), The Statistical Analysis of Failure Time Data, 2nd edition, J. Wiley, New York.
3. Miller, J (1980), Survival Analysis, J. Wiley, New York.
4. Elandt-Johnson, Regina C; Johnson, Norman L.(1999). Survival models and data analysis. Classics Library ed. – John Wiley & Sons.
5. Macdonald A S, An Actuarial Survey of Statistical Models for Decrement and Transition Data, British Actuarial Journal 2 (1996), (Research paper)

Course Title: Multivariate Analysis	**L**	**T**	**P**	**S**	**Semester:** 3rd
Course Code: STA603C	**4**	**x**	**x**	**x**	**Max Marks**: 100
Credits: 4					

Course Objective: The course aims to provide knowledge within multivariate statistics: theory, calculation techniques and applications.

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

1. Account for important theorems and concepts in multivariate analysis.
2. Account for the most common multivariate methods.
3. Understand structure of multivariate normal data.
4. Understand extensions of univariate techniques to multivariate frameworks and learn to apply dimension reduction techniques used in the data analysis.

Unit I: Introduction to Matrix Algebra, Multivariate Normal Distribution Theory: Marginal and conditional distribution, Joint distribution, Linear function of correlated normal variate. Characteristics function of multivariate normal distribution, Maximum likelihood estimation of the mean vector and co-variance matrix and their independence. Distribution of sample mean vector.

Unit II: Quadratic form and its distribution. Multiple linear equations, Multiple correlation, partial correlation in multiple setups. Partial and multiple correlation coefficients, their maximum likelihood estimators (MLE). Canonical Correlation and Canonical variables: Definition, uses, estimation and statistical inference. Distribution of characteristic roots and vectors: The distribution of canonical Correlation.

Unit III: Wishart matrix and its distribution. Some important properties of the Wishart distribution. Characteristic function of Wishart distribution, Hotelling's T^2statistic as a generalization of square of Student's statistic. Derivation of the generalized T^2 statistics and its distribution. Some important properties of T^2 statistics and its uses. Distance between two populations, Mahalnobis D^2 statistic and its relation with Hotelling's T^2 statistic.

Unit IV: Discrimination procedure for discrimination between two multivariate normal populations, Sample discriminant function, tests associated with discriminate function, Principal Component: Definition of principal components, uses, estimation and computation, Statistical inference on principal components. Factor Analysis: Definition of factor analysis and uses, linear factor models, estimation of factor loading, Factor rotation, estimation of factor scores.

Textbooks/References:

1. Alvin C. Rencher: Methods of Multivariate Analysis, 2nd Ed. John Wiely
2. Kshirsagar A. M. : Multivariate Analysis. Maral-Dekker.

3. Johnosn, R.A. and Wichern. D.W.:Applied multivariate Analysis. 5thAd.Prentice –Hall.
4. Anderson T. W.: An introduction to Multivariate statistical Analysis2nd Ed. John Wiely.
5. Morrison D.F.: Multivariate Statistical Methods McGraw-Hill.

Course Title: Design of Experiments	**L**	**T**	**P**	**S**	**Semester:** 3rd
Course Code: STA605E	**4**	**x**	**x**	**x**	**Max Marks**: 100
Credits: 4					

Course Objective: To equip students with the fundamental principles and techniques of designing, conducting and analyzing experiments using statistical methods, with emphasis on ANOVA and factorial designs.

Course Outcomes:

1. Understand the potential practical problems and applications of design of experiments in various fields.
2. Build a deeper understanding, and tools for analysis of experiments.
3. Describe how the analysis of the data from the experiment should be carried out.
4. Appreciate the advantages and disadvantages of a design for a particular experiment.

Unit I: Planning of experiment: Nomenclature, Introduction to basic designs and their analysis, Principles of Experimental design. Completely Randomized Design (CRD), Randomized Block Design (RBD), Latin Square Design (LSD), Missing plot techniques in RBD with one and two missing observations, Analysis of LSD with one missing observation.

Unit II: Introduction to Incomplete block design intra block analysis (estimability), estimates of estimable linear parametric function; Balanced Incomplete Block Design. Intra block analyses, recovery of inter block information;

Unit-III: Analysis of covariance. Practical situations where analysis of covariance is applicable. Model for analysis of covariance in CRD and RBD. Estimation of parameters. Preparation of analysis of covariance (ANOCOVA) table, test for $\beta = 0$, test for equality of treatment effects (computational technique only). Numerical illustrations.

Unit-IV: General description of factorial experiments, factorial effects, analysis of factorial experiment (2^n,3^n), main and interaction effects, advantages and disadvantages, total and partial confounding, split plot experiment.

Textbooks/References:

1. Goon, Gupta, Dasgupta: Fundamental of Statistics, Vol. I and II, The World Press Pvt. Ltd. Kolkata.
2. Montgomery, D.C.: Design and Analysis of Experiments, Wiley Eastern Ltd., New Delhi.
3. Cochran, W.G. and Cox, G.M.: Experimental Design, John Wiley and Sons, Inc., New York.
4. Gupta, S.C. and Kapoor, V.K.: Fundamentals of Applied Statistics, S. Chand & Sons, New Delhi.

5. Das, M.N. and Giri, N.C.: Design and Analysis of Experiments, Wiley Eastern Ltd., New Delhi.

Course Title: Real Analysis	**L**	**T**	**P**	**S**	**Semester:** 3rd
Course Code: MTH601E	**4**	**x**	**x**	**x**	**Max Marks**: 100
Credits: 4					

Course Objective: The main purpose is to provide mathematical foundation for statistics courses to enhance their knowledge in Real Analysis.

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

1. Understand the foundational concepts of real numbers, sequences and series and their convergence properties.
2. Apply the concepts of continuity, differentiability and integration to solve analytical problems in statistics.
3. Analyze and interpret the behavior of complex functions using algebraic and geometric approaches.
4. Employ Cauchy-Riemann equations and theorems of complex analysis to evaluate contour integrals.
5. Integrate real and complex analysis concepts to support advanced studies in probability and statistical theory.

Unit I: Real numbers, ordered sets; bounded sets**,** supremum and infimum, **o**rdered fields; completeness of the set of real numbers, Archimedean property, sequences and series, convergence Bolzano-Weistrass Theorem; Heine Borel Theorem.

Unit II: Continuity, uniform continuity; differentiability, Mean Value Theorem, sequences and series of functions, uniform convergence, Riemann sums, Riemann Integral, Improper integral.

Unit III: Algebra of complex numbers, the complex plane, power series, polynomials, transcendental functions, analytic functions, Cauchy-Riemann Equations.

Unit IV: Contour integrals, Cauchy's theorem and Cauchy's integral formula, Liouville's Theorem, Maximum Modulus Principle, Schwarz Lemma, `Taylor and Laurent series, calculus of residues.

Text Books/References:

1. Mathematical Analysis by Tom Apostol
2. Principles of Mathematical Analysis by Walter Rudin
3. A Course in Real Analysis by Shanti Narayan
4. Real Analysis by Terence Tao, Hindustan Book Agency (TRIM Series)
5. Complex Analysis by Lars V. Ahlfors

Course Title: Financial Mathematics	**L**	**T**	**P**	**S**	**Semester:** 3rd
Course Code: STA603E	**4**	**x**	**x**	**x**	**Max Marks**: 100
Credits: 4					

Course Objective: Introduces students the basic grounding in financial mathematics like simple interest, compound interest and their applications to calculate, accumulate value, present value, cash flows and loan calculation.

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

1. Understand the fundamental concepts of financial Mathematics.
2. Apply the learned techniques on real life financial decisions.
3. Learn about the types of derivatives and their characteristics.
4. Develop basic financial Decisions.

Unit I: Accumulation Function, Simple interest, compound interest, Generalized Cash- flow model, discounting. Nominal interest rates or discount rates in terms of different time periods. Force of interest.

Unit II: Definition of compound interest functions including annuities certain, Level payment annuities, Level payment perpetuities, Repayment mode, Non-level payment annuities and perpetuities: Geometric, Increasing and Decreasing, Continuous payment Cash flows.

Unit III: Inflations, arbitrage and hedging, the investment and risk characteristics of the different types of asset available for investment purposes. Variable interest rates, Investment and risk characteristics of various types of assets such as bonds, shares and other types of equity finance, options and derivatives.

Unit IV: Forwards and futures, hedging using futures, different types of futures, standardization, marking to market.

Textbooks/References:

1. Hull, J. C., (2003): Derivatives Options & Futures, Pearson Education.
2. Donald D.W.A. (1984): Compound Interest & Annuities Certain. Published for the Institute of Actuaries and the Faculty of Actuaries, London.
3. Mark Suresh Joshi,(2009): The Concept and Practice of Mathematical Finance, Cambridge University Press.
4. Dixit S. P., Modi C.S. and Joshi R.V. (2000): Mathematical Basis of Life Assurance, Insurance Institute of India, Bombay.
5. Kellison, Stephen G(1991): The Theory of Interest, Homewood, IL: Richard D. Irwin, 2/e

Course Title: Financial Reporting	**L**	**T**	**P**	**S**	**Semester:** 3rd
Course Code: STA604E	**4**	**x**	**x**	**x**	**Max Marks**: 100
Credits: 4					

Course Objective: Introduces students the basic grounding in financial reporting like investment and asset management. It will also introduce students to corporate finance and capital project appraisal.

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

1. Understand the fundamental concepts of financial Mathematics.
2. Apply the learned techniques on real life financial decisions.
3. Learn about the types of derivatives and their characteristics.
4. Develop basic financial Decisions.

Unit I: Investment and Asset management. Principles of finance, stakeholders in an organization, role and effects of capital markets, agency theory. Finances of joint stock, companies, types of companies, short, medium and long term financing of companies.

Unit II: Principles of personal and corporate taxation, principles of double taxation relief. Financial instrument: stocks, loan, debentures, bonds and shares. Types of shares and share issues. Capital structure, its effect of valuation of a company. Distribution of profits: dividend policies.

Unit III: Depreciation and its purpose, straight line basis, reducing balance method, generating accounts, the trial balance, goodwill on consolidation

Unit IV: Company's cost of capital, interaction with investment project, Evaluation of projects and risky investments. Annual reports and accounts, Construction of accounts, types of accounts. Construction of balance sheets. Various accounting ratios.

Unit V: Subsidiary and associated companies. Interpretation of accounts of a company or a group of companies. Assessment of capital investment projects.

Textbooks/References:

1. Chandra Prasana (2008).Financial Management, Theories and Practices.
2. Brealey, R.A., Meyers S.C. and Allen F. (2001). Principles of corporate Finance, 9th Edition, Mcgraw Hill Professional.
3. Financial statement analysis in Europe. (1999): Samuels, J M; Brayshaw, R E; Craner, J M. - Chapman & Hall.
4. Fundamentals of financial management.(2000): Brigham, Eugene F; Houston, Joel F., 9/e, Har-court Brace.
5. How to read the financial pages (2003): Brett, M., 2/e, Random House Business Books, 430 pages.

Course Title: Python for Statistics	**L**	**T**	**P**	**S**	**Semester:** 3rd
Course Code: STA606C	**2**	**x**	**2**	**x**	**Max Marks**: 100
Credits: 3					

Course Objectives: The course introduces students to Python programming for statistical applications. It focuses on data handling, visualization, probability distributions, inference, and modeling using Python libraries, enabling students to apply statistical methods to real-world data effectively.

Course Outcomes: By the end of the course, students will be able to:

1. Apply Python for data handling and statistical analysis.
2. Perform descriptive statistics, probability modeling, and hypothesis testing in Python.
3. Build and interpret regression models using real-world datasets.
4. Conduct applied statistical projects integrating theory and Python tools.

Unit I: Introduction to Python: Jupyter Notebook, basic syntax; Python data structures (lists, tuples, dictionaries, sets); NumPy arrays: creation, indexing, slicing, operations; Data input/output (CSV, Excel); Data handling with Pandas: importing, cleaning, transforming datasets.

Unit II: Descriptive statistics and visualization (Matplotlib, Seaborn); Probability distributions: Binomial, Poisson, Normal; Simulation of random variables and sampling distributions (t, chi-square, F); Hypothesis testing: one- and two-sample t-tests, chi-square test of independence; One-way ANOVA and confidence intervals.

Unit III: Correlation and regression analysis; Simple and multiple regression using Statsmodels/Scikit-learn; Logistic regression basics; Model selection (AIC, BIC, cross-validation).

List of Practicals

1. NumPy arrays and vectorized operations
2. Data import/export and cleaning using Pandas
3. Descriptive statistics and visualization using Matplotlib/Seaborn
4. Simulation of random variables from Binomial, Poisson, Normal distributions
5. Hypothesis testing: t-test and chi-square test with real datasets
6. One-way ANOVA using Python
7. Regression analysis (simple, multiple) with Statsmodels
8. Logistic regression on a real dataset
9. Mini-project: Statistical analysis of a dataset using Python

Textbooks/References

1. Wes McKinney, *Python for Data Analysis*, O'Reilly.

2. Allen B. Downey, *Think Stats: Probability and Statistics for Programmers*, O'Reilly.
3. Jake VanderPlas, *Python Data Science Handbook*, O'Reilly.
4. Thomas Haslwanter, *An Introduction to Statistics with Python*, Springer.

Online Resources

1. https://docs.python.org – Official Python documentation
2. https://numpy.org – NumPy for arrays and numerical computing
3. https://pandas.pydata.org – Pandas for data handling
4. https://matplotlib.org – Matplotlib for visualization
5. https://www.statsmodels.org – Statsmodels for statistical modeling