

WestBengal State University

Detailed Syllabus for Four-years Under Graduate Programme In STATISTICS

**Under Curriculum and Credit Framework for
Undergraduate Programmes**

**Based on
National Education Policy, 2020**

Syllabus for 4-year Undergraduate Programme (Honours / Honours with Research)

Major Papers Semester-I

STSDSC101T-Descriptive Statistics-I and Probability-I	
[Credit3]	[45LectureHours]
Unit1: Statistical Data	[12LectureHours]
<p>Statistics: Definition and scope. Concepts of statistical population and sample. Data: quantitative and qualitative, cross-sectional and time-series, discrete and continuous. Scales of measurement: nominal, ordinal, interval and ratio. Collection of data, concept of questionnaire. Presentation of data: tabular and graphical. Frequency distributions, cumulative frequency distributions and their graphical representations. Stem and leaf displays.</p>	
Unit2: Univariate Data Analysis	[15LectureHours]
<p>Measures of Central Tendency: Mean, Median, Mode. Measures of Dispersion: Range, Mean deviation, Standard deviation, Coefficient of variation, Gini's Coefficient, Lorenz Curve. Moments, skewness and kurtosis. Quantiles and measures based on them. Box Plot. Outliers and its detection using quantiles. Trimmed mean.</p>	
Unit3: Introduction to Probability	[18 Lecture Hours]
<p>Introduction, random experiments, sample space, events and algebra of events. Definitions of Probability: classical, statistical and axiomatic. Probability space and different properties of probability function. Conditional Probability, laws of addition and multiplication, independent events, theorem of total probability, Bayes' theorem and its applications.</p>	
STSDSC101P: List of Practical	
[Credit2]	[60 Lecture Hours]
<ul style="list-style-type: none"> • Graphical representation of data. • Problems based on construction of frequency distributions, cumulative frequency distributions and their graphical representations. • Problems based on measures of central tendency. • Problems based on measures of dispersion. • Problems based on combined mean and variance and coefficient of variation. • Problems based on moments, skewness and kurtosis. • Problems related to quantiles and measures based on them. • Problem of detection of outliers using quantiles, construction of boxplot. • Numerical sums using classical definition of Probability. • Numerical sums on conditional probability. 	
Reference Books	

- Goon, A.M., Gupta, M.K. and Dasgupta, B.(2002):Fundamentals of Statistics, Vol. I & II, 8thEdition, World Press, Kolkata.
- Miller, Irwin and Miller, Marylees John E. Freunds(2006): MathematicalStatisticswithApplications,7thEdition, Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes, D. C.(2007): Introduction to the Theory of Statistics,3rdEdition, Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J. W. (1977): Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Freedman, D., Pisani, R. and Purves, R. (2014): Statistics,4thEdition, W.W. Norton & Company.
- Chung, K. L. (1983): Elementary Probability Theory with Stochastic Process, Springer/ Narosa.
- Feller, W. (1968): An Introduction to Probability Theory & its Applications, John Wiley.
- Goon, A.M., Gupta, M.K. & Dasgupta, B. (2003): An Outline of Statistical Theory Vol- I, World Press.
- Parzen, E. (1972): Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V.(1937): Introduction to Mathematical Probability, McGraw Hill.
- Cacoullous, T.(1973): Exercises in Probability, Narosa.
- Ross, S. (2002): A First Course in Probability, Prentice Hall.
- Stirzaker, D. (2003): Elementary Probability, 2ndEdition, Cambridge University Press
- Rahman, N.A.(1983): Practical Exercises in Probability and Statistics, Griffin.
- Rohatgi, V. K. and Saleh, A.K. Md. E. (2009): An Introduction to Probability and Statistics. 2nd Edition, John Wiley and Sons.

STSHSE101T –Numerical Computations with C

[Credit2]

[30 Lecture Hours]

Unit 1: Numerical Analysis

[18 Lecture Hours]

Approximation of numbers and functions. Absolute and Relative errors. Interpolation: Polynomial approximation, Weierstrass Theorem (Statement). Difference Table, Newton's Forward and Backward interpolation formulae and Lagrange's general interpolation formula, Error terms. Numerical Differentiation and its Applications. Numerical Integration: Trapezoidal and Simpson's 1/3rd rules. Numerical solution of equations: method of bisection method of iteration and Newton-Raphson method in one unknown, conditions of convergence (statement only).

Unit2: Introduction to C

[12 Lecture Hours]

Components, basic structure of programming. Notion of header file. Concept of character and variable; allocation of memory. Declaration and assignment of variables and array variables. Input and output operations in C. Use of conditional operations and loops; if...else, for, while, do...while etc. dim arrays. User defined functions.

STSHSE101P: List of Practical**[Credit1]****[30 Lecture Hours]**

- Simple mathematical operations.
- Matrix addition, subtraction and multiplication.
- Finding values of a function $y=f(x)$ for given values of x
- Roots of a real quadratic equation.
- Finding output for a user-defined function.
- Sorting of an array and finding quantiles.
- Preparing a frequency table.
- Mean, median and mode of a grouped frequency Data.
- Variance and coefficient of variation of a grouped frequency data.
- Numerical methods: Interpolation by Lagrange's formula.
- Solving one variable equation using method of bisection.
- Solving one variable equation using Newton-Raphson method.
- Solving one variable equation using iteration method.
- Trapezoidal rule for numerical Integration.
- Simpson's 1/3rd rule for numerical integration.
- Storing the C-output in a file.

Reference Books

- Kernighan, B.W. and Ritchie, D. (1988): C Programming Language, 2nd Edition, Prentice
- Balagurusamy, E. (2011): Programming in ANSIC, 6th Edition Tata McGraw
- Gottfried, B.S. (1998): Schaums Outlines: Programming with C, 2nd Edition, Tata McGraw Hill.
- Jain, M. K., Iyengar, S. R. K. and Jain, R. K. (2003): Numerical methods for Scientific and Engineering computation, New age International Publisher, India.
- Mukherjee, Kr. Kalyan (1990): Numerical Analysis. New Central Book Agency.
- Sastry, S.S. (2000): Introductory Methods of Numerical Analysis, 3rd edition, Prentice Hall of India Pvt Ltd., New Del.
- Scarborough, J.B. (1966): Numerical Mathematical Analysis. Oxford and IBH Publishing.
- F.B. Hildebrand: Introduction to Numerical Analysis, Tata McGraw Hill

Semester II

STSDSC202T – Descriptive Statistics- II & Probability-II

[Credit3]

[45LectureHours]

Unit1: Bivariate Data Analysis

[12 Lecture Hours]

Bivariate data: Definition, scatterdiagram, simple correlation, linear regression, principle of least squares, fitting of polynomial and exponential curves, correlation ratio, correlation index, intra-class correlation. Rank correlation: Spearman's and Kendall's measures.

Unit2: Categorical Data Analysis

[15 Lecture Hours]

Analysis of Categorical Data: Contingency table, independence & association of attributes. Ideas of complete and absolute association. Yule's measures of association and colligation, Cramer's measure of association, extension to $k \times l$ contingency table: Pearson's chi-square, Kendall's r and r_b , Goodman Kruskal's γ . Difference of proportions, relative risk, odds-ratio, log-odds-ratio; types of observational studies.

Unit3: Random Variables

[18 Lecture Hours]

Random Variables: Definition of discrete and continuous random variables, cumulative distribution function (c.d.f.) and its properties (without proof), probability mass function (p.m.f.), and probability density function (p.d.f.). Expectation and Variance. Standard discrete probability distributions: Discrete Uniform, Binomial, Poisson, Geometric, Hypergeometric and Negative Binomial.

STSDSC202P: List of Practical

[Credit2]

[60 Lecture Hours]

- Correlation coefficient for a bivariate frequency distribution.
- Lines of regression, angle between lines and estimated values of variables.
- Fitting of polynomials, exponential curves.
- Spearman rank correlation with and with outliers.
- Computation of correlation ratio.
- Computation of intra class correlation coefficient.
- Fitting of binomial distribution for given n and p .
- Fitting of binomial distribution after computing mean and variance.
- Fitting of Poisson distribution for given value of λ .
- Fitting of Poisson distribution after computing mean.
- Fitting of negative binomial.
- Fitting of suitable distributions.
- Application problem based on binomial distribution.
- Application problem based on Poisson distribution.
- Application problem based on negative binomial distribution.

Reference Books

- Goon, A.M., Gupta, M.K. and Dasgupta, B. (2002): Fundamentals of Statistics, Vol. I & II, 8th Edition, World Press, Kolkata.

- Miller, Irwin and Miller, Marylees (2006): John E. Freunds Mathematical Statistics with Applications, 7th Edition, Pearson Education, Asia.
- Mood, A.M., Graybill, F.A. and Boes, D.C. (2007): Introduction to the Theory of Statistics, 3rd Edition, Tata McGraw-Hill Pub. Co. Ltd.
- Tukey, J.W. (1977): Exploratory Data Analysis, Addison-Wesley Publishing Co.
- Freedman, D., Pisani, R. and Purves, R. (2014): Statistics, 4th Edition, W. W. Norton & Company.
- Agresti, A. (2010): Analysis of Ordinal Categorical Data, 2nd Edition, Wiley.
- Chung, K.L. (1983): Elementary Probability Theory with Stochastic Process, Springer/ Narosa.
- Feller, W. (1968): An Introduction to Probability Theory & its Applications, John Wiley.
- Goon, A.M., Gupta, M.K. & Dasgupta, B. (2003): An Outline of Statistical Theory Vol- I, World Press.
- Parzen, E. (1972): Modern Probability Theory and its Applications, John Wiley.
- Uspensky, J.V. (1937): Introduction to Mathematical Probability, McGraw Hill.
- Cacoullos, T. (1973): Exercises in Probability, Narosa.
- Ross, S. (2002): A First Course in Probability, Prentice Hall.
- Stirzaker, D. (2003): Elementary Probability, 2nd Edition, Cambridge University Press
- Rahman, N.A. (1983): Practical Exercises in Probability and Statistics, Griffin.
- Rohatgi, V.K. and Saleh, A.K.Md.E. (2009): An Introduction to Probability and Statistics. 2nd Edition, John Wiley and Sons.

STSHSE202M-Computation using Software

[Credit 3]

[90 Lecture Hours]

Unit1: Introduction to Excel

[20 Lecture Hours]

Basic idea about software. Input and modification of data. Basic cell operations; operation of mathematical and inbuilt functions on cell. Display summary statistics for univariate and bivariate data. Regression and correlation computation. Use of Data Analysis' tool (only the applicable functions). Construction of data table and operations on it.

Graphical representation of data: Column diagram, pie diagram, line diagram, bar diagram, scatterplot, boxplot, stock diagram, surface plot and radar plot, histogram.

Unit2: Introduction to Minitab**[30 Lecture Hours]**

Basic idea about software. Input and modification of data. Basic cell operations; operation of mathematical and inbuilt functions on cell. Display summary statistics for univariate and bivariate data. Regression and correlation computation. Use of 'Stat' tab (only the applicable functions).

Graphical representation of data: Column diagram, pie diagram, line diagram, bar diagram, scatter plot, boxplot, stock diagram, surface plot and radar plot, histogram.

Unit3: Introduction to R**[40 Lecture Hours]**

Use of Rascal calculator. Operations within build mathematical functions. Input a vector, numeric and non-numeric vectors. Addition and deletion of data from a vector. Logical operations and use of different logical functions. Understanding the non-numeric outputs, like—NULL, NA and NaN. Array and Matrix with associated operations. Construction of new function in R. Use the help in R. Loading and installing packages in R.

ReferenceBooks

- Davies, T.M. (2016): The Book of R: A First Course in Programming and Statistics, 1st Edition, NoStarch Press, USA.

Semester III

STSDSC303T–Mathematical Analysis I	
[Credit5]	[75 Lecture Hours]
Unit1: Background on Real Analysis	[18 Lecture Hours]
Basic set theory; countability of a set and related results, countability of \mathbb{Q} ; supremum and infimum of a set; Concept of function; equivalent set. Basic theorem on real numbers; Density property, Completeness property, Archimedean property. Neighbourhood set, interior point, cluster point, isolated point and boundary point (definition and examples only). Open set, closed set and compact set (closed and bounded).	
Unit2: Sequence and Series of Real Numbers	[22 Lecture Hours]
Sequence of real numbers and their convergence, limits of sequences, Cauchy's general principle of convergence, Cauchy's first theorem on limits, monotonic sequences, monotone convergence theorem of sequence, Squeeze theorem. Concept of subsequences, Bolzano-Weierstrass theorem. Limit superior and limit inferior of a bounded sequence, Cauchy's second theorem on limits. Infinite series, positive-termed series and their convergence. Comparison tests, Limit comparison test, D'Alembert's ratio test, Cauchy's n^{th} root test, Integral test, Condensation test, Dirichlet's test. (Statements and examples only). Absolute convergence of real series, Leibnitz's test for the convergence of alternating series, conditional convergence.	
Unit3: Properties of Real-valued Functions	[25 Lecture Hours]
Limit, Continuity, Differentiability, Uniform continuity and Boundedness of functions, Intermediate value theorem, Indeterminate forms, L'Hospital's rule. Rolle's Theorem and Lagrange's Mean Value Theorems (without proof). Taylor's Theorem with Lagrange's and Cauchy's form of remainder (without proof). Taylor's series expansion. Functions on several variables, constrain optimization using Lagrange's multipliers. Double and multiple integrals, Fubini's theorem, change of variables, concept of Jacobian.	
Unit 4: Sequence and Series of Functions	[10 Lecture Hours]
Sequence and series of functions, pointwise and uniform convergences. Simple tests, Weierstrass M-test. Power series, radius and interval of convergence, basic properties of power series.	
Reference Books	
<ul style="list-style-type: none">• Bartle, R.G. and Sherbert, D.R. (2011): Introduction to Real Analysis, John Wiley & Sons, Inc.• Apostol, T. M. (1974): Mathematical Analysis, World Student Series Edition• Kumar, A. and Kumaresana, S. (2014): Basic Course in Real Analysis, Taylor & Francis Group, LLC• Thomson, B.S., Bruckner, J.B. and Bruckner, A.M. (2001): Elementary Real Analysis, Prentice-Hall, Inc.	

- Rădulescu, T-L.T., Rădulescu, V and Andreescu, T. (2009): Problems in Real Analysis, Springer Science+Business Media, LLC
- Trench, W.F. (2003): Introduction to Real Analysis, Prentice Hall
- Abbott, S. (2012). Understanding Analysis. Springer New York.

STSHSE303M -Computation using R

[Credit 3]

[90 Lecture Hours]

Unit 1: Statistical Measures and Graphical Representations of Statistical Data [40 Lecture Hours]

Calculation of descriptive measures for univariate and bivariate data. Graphical representation of statistical data; use of different graphical parameters. Linear and polynomial regression with the help of inbuilt functions, use of 'summary'. Input a data from a file stored in the computer. Use of 'slider' in R.

Unit 2: Mathematical and Statistical Operations [30 Lecture Hours]

Solution of polynomial, algebraic and transcendental equations. Representation of a function using 'curve'. Optimization of a function.
Generation of random numbers for known theoretical distributions, calculation of probability using inbuilt functions as well as using statistical definition.

Unit 3: Simulation with R [20 Lecture Hours]

Concept of loops and conditional functions. Notion of the reduction of complexity of a program in R. Use of 'apply' function.
Concept of simulation, generations of random numbers of functions of random variables and its graphical representation to visualized its probability distribution.

Reference Books

- Davies, T.M. (2016): The Book of R, No Starch Press, San Francisco
- Braun, W.J. and Murdoch, D.J (2021): A First Course in Statistical Programming with R, Cambridge University Press

Semester-IV

STSDSC404T-Vector and Matrix Algebra	
[Credit3]	[45 Lecture Hours]
Unit1: Vector & Matrix	[20 Lecture Hours]
<p>Vector and vector spaces: Vectors defined on Euclidean space, definition of null-vector, unit vector, sum and inner product of two vectors, linear dependence and independence of vectors, spanning set, basis and dimension of a subspace, norm of a vector, orthogonal and orthonormal vectors. Gram-Schmidt orthogonalization. Concept of vector space, vector subspace, sum and direct sum of subspaces, projection of a vector on subspace. Orthogonal complement and orthogonal projection.</p> <p>Matrix and matrix algebra: Basic matrix operations, Different types of matrices, orthogonal matrix and trace of a matrix. Elementary operations and elementary matrices. Matrix transformation, Determinant and its properties, evaluation of determinants. Non-singular matrix and inverse of a non-singular matrix. Partitioned matrix, its determinant and inverse.</p>	
Unit2: Rank of Matrix	[12 Lecture Hours]
<p>Rank of a matrix: Row-rank, column rank and rank of a matrix, echelon matrix and rank determination, rank inequalities, rank factorization. Null space and nullity of a matrix, Sylvester's and Frobenius inequality. Idempotent matrix and projector.</p> <p>Solution of equation and elementary concept of g-inverse of a matrix: Condition for consistency of a system of equations, left inverse, right inverse and generalized inverse of a matrix.</p>	
Unit3: Quadratic Form & Eigenvalues	[13 Lecture Hours]
<p>Eigenvalues and eigenvectors: Definition of eigenvalues and eigenvectors. Matrix polynomial, Caley Hamilton theorem and its application. Eigenspace, geometric and algebraic multiplicity of eigenvalues. Spectral decomposition theorem with application.</p> <p>Quadratic form: Quadratic form and its classification, important properties of positive definite and non-negative definite matrix. Diagonalization of quadratic form. Extrema of quadratic form.</p>	
STSDSC404P: List of Practical	
[Credit2]	[60 Lecture Hours]
<ul style="list-style-type: none">• Problems on independent and dependent set of vectors.• Problems on spanning set and basis.• Problems on Gram-Schmidt orthogonalization.• Problems on rank of a matrix.• Problems on inverse of a matrix.• Problems on g-inverse of a matrix.• Problems on system of linear equations.• Problems on quadratic forms.• Problems on eigenvalues and eigenvectors.• Other problems based on the above topics.	

Reference Books

- Hadley G. (2002): Linear Algebra. Narosa Publishing House (Reprint)
- Searle S.R. (1982): Matrix Algebra Useful for Statistics. John Wiley & Sons
- Fienberg, S. H., Insel, A. J. and Spence, L. E. (2017): Linear Algebra (4th De.), Pearson
- Rao, A. R. and Bhimasankaram, P. (2000): Linear Algebra. Hindustan Book Agency
- Narayan Shanti (2004): A Textbook of Matrices, S Chand & Co Ltd.
- Apostol, T. M. (2002): Mathematical Analysis, Narosa Publishing House

STSDSC405T–Sampling Distribution & Statistical Inference I

[Credit 3]

[45 Lecture Hours]

Unit 1: Sampling Distribution

[18 Lecture Hours]

Distribution of functions of random variables (using transformation), reproductive properties, illustration through standard discrete and continuous distributions. Definition and derivation of Chi-square distribution, t-distribution, and F-distribution; nature of p.d.f. curve for different degrees of freedom, mean, variance, m.g.f., property and limiting distribution of Chi-square and t-distribution. Student's, Fishers and paired t-distributions. Sampling distribution of sample correlation coefficient and sample regression coefficient while sampling from bivariate normal distribution. Sampling distribution of least square estimates of the coefficients of simple linear regression model under normality. Ordered Statistics.

Unit 2: Unbiasedness

[10 Lecture Hours]

Elementary concepts: Family of distributions, labelling parameter, parametric space, random sample, statistic, equivalent statistic, estimator and estimate, mean square error. Unbiased estimator and its properties, standard error. Exponential Family of distributions (one and multi-parameter), illustration through examples. Theory of partition, statistic and partition, illustrated through Bernoulli distribution. Notion of sufficiency through partition theory.

Unit 3: Best Choice of Estimator

[17 Lecture Hours]

Sufficiency, Factorization theorem, minimal sufficiency, finding of minimal sufficient statistic due to Lehmann & Scheffe (1950) and associate results. Completeness, complete-sufficient statistic, role of exponential family of distribution in finding complete-sufficient statistic. Ancillary statistic, examples, Basu's theorem. UMVUE, Cramer-Rao Inequality, equality case, Cramer-Rao Lower Bound. Fisher information contained in a statistic, efficient estimator, Bhattacharyya Bounds, relative efficiency of estimators and results on UMVUE. Rao-Blackwell theorem and Lehmann-Scheffé theorem. Maximum likelihood estimation, properties and examples.

STSDSC405P: List of Practical
[Credit 2] [60 Lecture Hours]

- Real life based practical problems on sampling distribution.
- Real life based practical problems on ordered statistics.
- Problems on finding of unbiased estimators of different parametric functions from known and unknown theoretical distributions.
- Problems on finding relative efficiencies of two unbiased estimators of a parametric function.
- Problems on finding of UMVUE of different parametric functions from known and unknown theoretical distributions.
- Problems on maximum likelihood estimation.

Reference Books

- Goon, A.M., Gupta, M.K. and Dasgupta, B. (2005): An Outline of Statistical Theory, Vol. I& II, World Press, Kolkata
- Goon, A.M., Gupta, M.K. and Dasgupta, B. (2001): Fundamental of Statistics, Vol. I, World Press, Kolkata
- Rohatgi, V.K. and Saleh, A.K.Md.E. (2009): An Introduction to Probability and Statistics. 2nd Edn. (Reprint), John Wiley and Sons
- Mood, A.M, Graybill, F.A. and Boes, D.C. (1974): Introduction to the Theory of Statistics, McGraw Hill
- Bhat, B.R, Srivenkatramana, T. and Rao, Madhava K. S. (1997): Statistics: A Beginner's Text, Vol. I, New Age International (P) Ltd.
- Casella, G. and Berger, R.L. (2002): Statistical Inference, 2nd Edn., Thomson Learning
- Lehman, E.L. and Casella, G. (1998): Theory of Point Estimation, 2nd Edn., Springer-Verlag New York, Inc.
- Ferguson, T.S. (1973). Mathematical Statistics. United States: Academic Press
- Li, B. and Babu, G.J. (2019): A Graduate Course on Statistical Inference, Springer Science + Business Media, LLC
- Mukhopadhyay, N. (2000): Probability and Statistical Inference, Marcel Dekker, Inc., New York
- Mukhopadhyay, P. (1996): Mathematical Statistics, New Central Book Agency, India

STSDSC406T -Probability Theory III

[Credit3]

[45 Lecture Hours]

Unit1: Moments and Inequalities

[8 Lecture Hours]

p.d.f. and c.d.f., illustrations and properties, univariate transformations through c.d.f. with illustrations. Derivation of moments. Probability Inequalities: Markov and Chebyshev.

Unit2: Generating Functions& Convergence	[10 Lecture Hours]
<p>Probability generating function, moment generating function, cumulant generating function and characteristic function. Uniqueness and inversion theorems (without proof) along with applications.</p> <p>Statement of convergence in distribution and probability, central limit theorem (iid case) and Lévy continuity theorem.</p>	
Unit3: Standard Continuous & Bivariate Distributions	[27 Lecture Hours]
<p>Uniform, normal, exponential, Cauchy, beta, gamma, lognormal, logistic, double exponential and Pareto along with their properties and limiting/approximation cases.</p> <p>Discrete two-dimensional random variable: joint, marginal and conditional, p.m.f. and c.d.f., statement of properties of c.d.f., independence of variables, trinomial distribution. Continuous two-dimensional random variable: joint, marginal and conditional, p.d.f., and c.d.f. Independence of two random variables, bivariate transformations with illustrations. Moments. Conditional expectation and Conditional variance. Correlation coefficient. Bivariate Normal Distribution and its mgf and properties.</p>	
STSDSC406P: List of Practical [Credit2]	[60 Lecture Hours]
<ul style="list-style-type: none"> • Problems based on the property of normal distribution. • To find the ordinate for a given area for normal distribution. • Application-based problems using normal distribution. • Fitting of normal distribution when parameters are given. • Fitting of normal distribution when parameters are not given. • Fitting of some other continuous distributions. 	
ReferenceBooks	
<ul style="list-style-type: none"> • Goon, A.M., Gupta, M.K. and Dasgupta, B. (2005): An Outline of Statistical Theory, Vol. I & II, World Press, Kolkata • Hogg, R.V., Tanis, E.A. and Rao J.M. (2009): Probability and Statistical Inference, Seventh Ed, Pearson Education, New Delhi • Hoel P.G., Port, S. and Stone, C. (1972): Introduction to Probability Theory, Houghton Mifflin • Miller, Irwin and Miller, Marylees (2006): John E. Freund's Mathematical Statistics with Applications, (7th Edn.), Pearson Education, Asia • S.M. Ross: A First Course in Probability • K.L. Chung: Elementary Probability Theory with Stochastic Process • Mukhopadhyay, P. (1996): Mathematical Statistics. New Central Book Agency 	

STSDSC407T –Survey Sampling**[Credit3]****[45 Lecture Hours]****Unit1: Simple Random Sample****[15 Lecture Hours]**

Concept of population and sample, complete enumeration versus sampling, sampling and non- sampling errors. Types of sampling: non-probability and probability sampling, basic principles of sample survey, simple random sampling with and without replacement, definition and procedure of selecting a sample, estimates of population mean, total and proportion, variances of these estimates, estimates of their variances and sample size determination.

Simple Random sampling using auxiliary information: Ratio and Regression methods of estimation, Hartley-Ross estimator.

Unit2: Stratified and Systematic Sampling**[18 Lecture Hours]**

Stratified random sampling, technique, estimates of population mean and total, variances of these estimates, proportional and optimum allocations and their comparison with SRS. Practical difficulties in allocation, estimation of gain in precision.

Systematic Sampling, Technique, estimates of population mean and total, variances of these estimates ($N=n \times k$ case). Comparison of systematic sampling with SRS and stratified sampling in the presence of linear trend and corrections.

Unit3: Other Sampling Methods**[12 Lecture Hours]**

Cluster sampling (equal-size clusters only) estimation of population mean and its variance, Concept of sub-sampling. Two-stage sampling, Estimation of Population mean and variance of the estimate, comparison between two-stage, cluster and uni-stage sampling. Randomized response technique: Warner's method.

STSDSC407P: List of Practical**[Credit2]****[60 Lecture Hours]**

- Select an SRS with and without replacement.
- For a population of size 5, estimate population mean, population mean square and population variance. Enumerate all possible samples of size 2 by WR and WOR and establish all properties relative to SRS.
- For SRSWOR, estimate mean, standard error, the sample size
- Ratio and Regression estimation: Calculate the population mean or total of the population. Calculate mean squares. Compare the efficiencies of ratio and regression estimators relative to SRS.
- Stratified Sampling: allocation of sample to strata by proportional and Neyman's methods. Compare the efficiencies of above two methods relative to SRS.
- Estimation of gain in precision in stratified sampling.
- Comparison of systematic with stratified sampling and SRS in the presence of a linear trend.
- Cluster sampling: estimation of mean or total, variance of the estimate, estimate of intra-class correlation coefficient, efficiency as compared to SRS.
- Two stage sampling.

ReferenceBooks

- Cochran, W.G. (1984): Sampling Techniques (3rd Ed.), Wiley Eastern
- Sukhatme, P.V., Sukhatme, B.V. Sukhatme, S. (1984). Sampling Theories of Survey with Application, IOWA State University Press and Indian Society of Agricultural Statistics
- Murthy, M.N. (1977): Sampling Theory & Statistical Methods, Statistical Pub. Society, Calcutta
- Des Raj and Chandhok P. (1998): Sample Survey Theory, Narosa PublishingHouse
- Goon A.M., Gupta M.K. and Dasgupta B. (2008): Fundamentals of Statistics, Vol-2, World Press

Semester V

STSDSC508T –Statistical Inference-II

[Credit 3]

[45 Lecture Hours]

Unit 1: Introduction to Hypothesis Testing

[15 Lecture Hours]

Elements of hypothesis testing: Null and alternative hypotheses, simple & composite hypotheses, critical region, type I and type II errors, level of significance, size, power, p -value. Exact tests: classical and p -value approaches, related confidence intervals. Tests of significance related to Binomial proportion(s), Poisson mean(s), Univariate Normal mean(s), standard deviation(s) and Bivariate normal parameters. Combination of probabilities in tests of significance.

Unit 2: Construction of Tests

[20 Lecture Hours]

Theory of hypothesis testing: Test function, randomized and non-randomized tests, most powerful (MP) test, uniformly most powerful (UMP) test, Fundamental Neyman-Pearson Lemma and its applications to construct MP and UMP tests. Monotone Likelihood Ratio property, Karlin-Rubin Theorem. Generalized Neyman-Pearson Lemma, UMPU tests: One parameter exponential family. Neyman structure, UMPU tests for composite hypotheses. Likelihood ratio tests, properties of likelihood ratio tests (without proof).

Unit 3: Confidence Interval and Sequential Analysis

[10 Lecture Hours]

Interval Estimation: Confidence intervals, Concepts of Uniformly Most Accurate (UMA) confidence sets, Uniformly Most Accurate Unbiased (UMAU) confidence sets, relationship with tests of hypotheses. Sequential Analysis: Stopping variables, Sequential Tests, Wald's equation for ASN, SPRT and its properties –fundamental identity, OC and ASN, Optimality of SPRT (under usual approximation).

STSDSC508P: List of Practical

[Credit 2] [60 Lecture Hours]

- Test of significance for single proportion and difference of two proportions.
- Test of significance for single Poisson mean and difference of two Poisson means.
- Test of significance and confidence intervals for single mean and difference of two means.
- Test of significance and confidence intervals for single variance and ratio of two variances.
- Test of parameters under bivariate normal distribution.
- Type I and Type II errors.
- Most powerful critical region.
- Uniformly most powerful critical region.
- Uniformly most powerful unbiased critical region.
- Power curves.
- Likelihood ratio tests for simple null hypothesis against simple alternative hypothesis.
- Likelihood ratio tests for simple null hypothesis against composite

alternative hypothesis.

- Confidence intervals, UMA confidence sets and UMAU confidence sets.
- Problems related to sequential testing procedure.

Reference Books

- Goon, A.M., Gupta, M.K. and Dasgupta, B. (2005): An Outline of Statistical Theory, Vol. I& II, World Press, Kolkata
- Goon, A.M., Gupta, M.K. and Dasgupta, B. (2001): Fundamental of Statistics, Vol. I, World Press, Kolkata
- Rohatgi, V.K. and Saleh, A.K. Md.E. (2009): An Introduction to Probability and Statistics. 2nd Edn. (Reprint), John Wiley and Sons
- Mood, A.M, Graybill, F.A. and Boes, D.C. (1974): Introduction to the Theory of Statistics, McGraw Hill
- Bhat, B.R, Srivenkatramana, T. and Rao, Madhava K. S. (1997): Statistics: A Beginner's Text, Vol. I, New Age International (P) Ltd.
- Casella, G. and Berger, R.L. (2002): Statistical Inference, 2nd Edn., Thomson Learning
- Lehman, E.L. and Romano, J.P. (2005): Testing Statistical Hypotheses, 3rd Edn., Springer Science + Business Media, LLC
- Ferguson, T.S. (1973). Mathematical Statistics. United States: Academic Press
- Li, B. and Babu, G.J. (2019): A Graduate Course on Statistical Inference, Springer Science + Business Media, LLC
- Mukhopadhyay, N. (2000): Probability and Statistical Inference, Marcel Dekker, Inc., New York
- Mukhopadhyay, P. (1996): Mathematical Statistics, New Central Book Agency, India

STSDSC509T -Linear Models

[Credit 3]

[45 Lecture Hours]

Unit 1: Basics of Linear Modeling[10 Lecture Hours]

General objectives of model building: inference and prediction, difference between parametric and nonparametric approaches. Review of simple linear regression and its properties, extension of linear regression: polynomial regression, multiple regression. General Linear Model: introduction and examples, use of dummy variables, different types of linear models and their parameter interpretations.

Unit 2: Gauss-Markov set-up [15 Lecture Hours]

Theory of linear estimation, estimability of linear parametric functions, method of least squares, Gauss-Markov Theorem, estimation space and error space, estimation of error variance. Fundamental theorems on least squares (statements only). Tests of General Linear Hypotheses (statements only). Classification of linear models. Generalized linear model: concept and examples only.

Unit 3: Analysis of variance, covariance and Regression [20 Lecture Hours]

Hypothesis testing in the case of simple and multiple regression models. Tests for parallelism and identity, linearity in regression models.

Analysis of Variance in one-way and two-way classified data (with equal number of observations per cell) for fixed, random, and mixed-effect models. Analysis of covariance for one-way and two-way classified data with one concomitant variable.

**STSDSC509P: List of Practical
[Credit 2] [60 Lecture Hours]**

- Estimability in Gauss Markov Model.
- Simple linear regression.
- Multiple regression using Gauss-Markov model.
- Tests for linear hypothesis.
- Analysis of variance of one way classified data.
- Analysis of variance of a two way classified data with one observation per cell.
- Analysis of variance of two way classified data with equal number of observations per cell.
- Analysis of covariance of a one way classified data with one concomitant variable.
- Analysis of covariance of a two way classified data with one concomitant variable.

Reference Books

- Goon, A.M., Gupta, M.K., and Dasgupta, B. (2002), Fundamental of Statistics, Volume 2, 8th Edn. The World Press, Kolkata.
- Scheffe, H, Linear Models
- Sengupta, D. and Jammalamadaka, S. R. (2020), Linear Models and Regression with R. World Scientific Publisher.
- Rao, C.R., Linear Statistical Inference.
- Mukhopadhyay, P. (2011): Applied Statistics, 2nd edition revised reprint, Books and Allied(P) Ltd.
- Weisburg, S (2005) Applied Linear Regression (Third edition), Wiley.
- Renchner, A.C. and Schaalje, G.B. (2008). Linear Models in Statistics (Second edition), John Wiley and Sons.

STSDSC510T – Stochastic Processes and Time Series Analysis**[Credit 3]****[45 Lecture Hours]****Unit 1: Time Series Data [15 Lecture Hours]**

Introduction to time series data and application of time series from various fields. Components of a time series (trend, cyclical and seasonal patterns, random error). Estimation of trend by free hand curve method, method of semi averages, method of moving averages, fitting various mathematical curves and growth curves. Effect of elimination of trend on other components of the time series. Estimation of seasonal component by method of simple averages, ratio to trend, ratio to moving average and link relative method. Harmonic Analysis. Variate difference method.

Unit 2: Stochastic Process [10 Lecture Hours]

Introduction of stochastic process and stationary process. Markov Chains: Definition of Markov Chain, examples including 2-state chain, random walk, etc. Transition probability matrix, higher order transition probabilities, order of a Markov chain, Markov chain as graphs, classification of states of a Markov Chain, stationary distribution, limiting distribution, period of a Markov Chain, convergence theorem.

Unit 3: Time Series Model, Estimation and Forecasting [20 Lecture Hours]

Time Series as a Stochastic Process. Stationary time series - weak stationarity, autocorrelation function and correlogram. Some special processes: Moving-average (MA) process and Autoregressive (AR) process of orders one and two. ACF, PACF and their graphical use in guessing the order of AR and MA processes. Estimation of the parameters of AR (1) and AR (2) – Yule-Walker equations. Introduction to ARMA and ARIMA models. Forecasting: Exponential smoothing methods, short term forecasting methods: Brown's discounted regression.

STSDSC510P: List of Practical

[Credit 2]

[60 Lecture Hours]

- Calculation of transition probability matrix, Identification of characteristics of reducible and irreducible chains, Identification of types of classes, Identification of ergodic transition probability matrix, Stationarity of Markov chain.
- Plotting a real time series, and detecting various features (trend, periodic behaviours etc.).
- Determination of trend by (i) Moving Average Method, ii) curve fitting and plotting de-trended series
- Determination of seasonal indices by (i) ratio to trend method, (ii) ratio to moving average method and (iii) method of link relatives
- Harmonic Analysis
- Correlogram Analysis
- Plotting ACF and PACF of a given time series.
- Fitting of AR 1 and AR 2 models
- Forecasting by short term forecasting methods.
- Forecasting by Exponential Smoothing

Reference Books

- Hoel, P.G.; Port, S.C. and Stone, C.J. (1986): Introduction to Stochastic Processes, Waveland Pr.
- Medhi, J. (2009): Stochastic Processes, New Age International Publishers.
- Ross, S.M. (2008): Stochastic Processes, Second Edition, Wiley Publisher.
- Bhat, B.R. (2000): Stochastic Models: Analysis and Applications, New Age International Publishers.
- R. N. Bhattacharya and Waymire, E.C. (2009): Stochastic Process and Applications, SIAM Classics Ed.
- Gun, Gupta and Dasgupta (2002): Fundamentals of Statistics, Vol. II, World Press.
- Cooray, TMJA (2008): Applied Time Series, Analysis and Forecasting, Narosa Publishing House.
- Chatfield, C. (2004): Analysis of Time Series, Chapman & Hall.
- Brockwell, P. and Davis, D. (2016): Introduction to Time Series and Forecasting, 3rd Ed., Springer.

STSDSC511T –Multivariate Distributions & Nonparametric Methods

[Credit 3]

[45 Lecture Hours]

Unit 1: Introduction to Random Vector

[12 Lecture Hours]

Random Vector: Probability mass/density functions, Distribution function, Mean vector and Dispersion matrix, Marginal and Conditional distributions.

Unit 2: Multivariate Distributions and Applications

[15 Lecture Hours]

Multinomial Distribution, Multivariate Normal distribution and its properties. Sampling distribution for mean vector and variance-covariance matrix (Statement only). Multiple and partial correlation coefficients and their properties. Multivariate t-distribution: properties and limiting case (statement).

Unit 3: Nonparametric Inference

[18 Lecture Hours]

Nonparametric Tests: Introduction and concept, nonparametric location problem and scale problem, simple linear rank statistics, score functions, test for randomness based on total number of runs, empirical distribution function, Kolmogorov Smirnov test for one sample and two samples, Sign tests one sample and two samples, Signed rank test, Wilcoxon rank sum test, Mann Whitney test, Median test, Bivariate Sign test, Kruskal Wallis test.

STSDSC511P: List of Practical

[Credit 2] [60 Lecture Hours]

- Problems on multiple correlation.
- Problems on partial correlation.
- Problems on multinomial distribution.
- Problems on multivariate normal distribution.
- Problems on discriminant analysis.
- Problems on principal components analysis.
- Problems on test for randomness based on total number of runs.
- Problems on Kolmogorov Smirnov test for one sample and two samples.
- Problems on Sign test: one sample, two samples, and large sample.
- Problems on Wilcoxon rank sum test, Mann Whitney test and median test.
- Problems on Kruskal Wallis test.

Reference Books

- Anderson, T.W. (2003): An Introduction to Multivariate Statistical Analysis, 3rd Edn., John Wiley
- Muirhead, R. J. (1982): Aspects of Multivariate Statistical Theory, John Wiley
- Kshirsagar, A.M. (1972): Multivariate Analysis, 1st Edn. Marcel Dekker
- Johnson, R.A. and Wichern, D.W. (2007): Applied Multivariate Analysis, 6th Edn., Pearson & Prentice Hall
- Mukhopadhyay, P. (1996): Mathematical Statistics
- Gibbons, J. D. and Chakraborty, S (2003): Nonparametric Statistical Inference. 4th Edn., Marcel Dekker, CRC
- Kotz, S., Nadarajah, S. (2004): Multivariate t-Distributions and Their Applications, Cambridge University Press, United Kingdom
- Mukhopadhyay, P. (2009). Multivariate Statistical Analysis, World Scientific, Singapore

Semester-VI

STSDSC612T –Design of Experiments

[Credit 3]

[45 Lecture Hours]

Unit 1: Experimental Designs and Models [15 Lecture Hours]

Experimental designs: role and historical perspective. Terminologies: experimental unit. Basic principles, uniformity trials, fertility contour maps, choice of size and shape of plots and blocks. Completely Randomized Design (CRD), Randomized Block Design (RBD), Latin Square Design (LSD) – layout, model, treatment contrast and analysis, relative efficiencies, analysis with one missing observation. Orthogonal Latin Square Design.

Unit 2: Incomplete Block Design [15 Lecture Hours]

Balanced Incomplete Block Design (BIBD) – parameters, relationships among its parameters, incidence matrix and its properties. Symmetric BIBD, resolvable BIBD, Affine Resolvable BIBD, Intra-block Analysis, Complementary BIBD, Residual BIBD, Dual BIBD, Derived BIBD.

Unit 3: Factorial Experiments

[15 Lecture Hours]

Advantages, notations and concepts of 2^n factorial experiments - their design and analyses. Total and partial confounding for 2^n ($n \leq 5$), factorial experiments. Construction of one-half and one-quarter fractions of 2^n ($n \leq 5$) factorial experiments, Alias structure, Resolution of a design. Response surface design.

STSDSC612P: List of Practical

[Credit 2] [60 Lecture Hours]

- Analysis of CRD
- Analysis of an RBD
- Analysis of an LSD
- Analysis of an RBD with one missing observation
- Analysis of an LSD with one missing observation
- Intra Block analysis of a BIBD
- Analysis of 2^2 and 2^3 factorial in CRD and RBD
- Analysis of 2^2 and 2^3 factorial in LSD
- Analysis of a completely confounded two level factorial design in 2 blocks
- Analysis of a completely confounded two level factorial design in 4 blocks
- Analysis of a partially confounded two level factorial design
- Analysis of a single replicate of a 2^n design
- Analysis of a fraction of 2^n factorial design

Reference Books

- Goon, A.M., Gupta, M.K. and Dasgupta, B. (2005): Fundamentals of Statistics, Vol. II, 8thEdn. World Press, Kolkata
- Mukhopadhyay, P. : Applied Statistics
- Cochran, W.G. and Cox, G.M. (1959): Experimental Design. Asia Publishing House
- Dey, A. (1986): Theory of Block Designs, Wiley Eastern Limited
- Montgomery, D. C. (2008): Design and Analysis of Experiments, John Wiley
- Das, M.N. and Giri, N.C. (1986): Design and Analysis of Experiments. Wiley Eastern Ltd
- Kempthorne, O. (1965): The Design and Analysis of Experiments. John Wiley

STSDSC613T - Large Sample Theory

[Credit 3]

[45 Lecture Hours]

Unit 1: Limit Theorems

[18 Lecture Hours]

Convergence in probability, convergence in distribution, Convergence in r^{th} mean, almost sure convergence, relation between these of convergences, associate theorems. Weak Laws of Large Numbers (WLLN), Strong Laws of Large Numbers (SLLN) and their applications. Khinchini's WLLN, Kolmogorov's SLLN (without proof). Slutsky's theorem. De Moivre-Laplace central limit theorem (CLT). Lindeberg-Lévy CLT (with proof), Lindeberg-Feller CLT (statement only), Lyapunov's CLT and their applications. CLT for uniformly bounded random variables.

Unit 2: Standard Errors of Statistics and Variance Stabilization

[12 Lecture Hours]

Delta method and Cramér-Wold device. Derivation and uses of large sample standard error of sample moments, standard deviation, coefficient of variation, g_1 & g_2 measures, correlation coefficient, regression coefficient, asymptotic distribution of sample quantiles. Variance stabilizing transformation of Statistic, Derivation of \sin^{-1} , square root, logarithmic & Fisher's Z transformations and associated applications.

Unit 3: Asymptotic Properties and Pearsonian χ^2 [15 Lecture Hours]

Consistency, asymptotic efficiency, ARE, CAN and BAN estimators. Asymptotic properties of MLE (statement only) and their uses in testing and confidence interval.

Large sample distribution of Pearsonian χ^2 statistic, its uses (goodness of fit, independence and homogeneity). Yates' correction in a 2×2 contingency table. Fisher's exact test in 2×2 contingency table.

STSDSC613P: List of Practical

[Credit 2]

[60 Lecture Hours]

- Tests of significance and confidence intervals for single proportion and difference of two Binomial proportions using CLT.
- Tests of significance and confidence intervals for single Poisson mean and difference of two Poisson means using CLT.

- Tests of significance and confidence intervals concerning sample standard deviation, coefficient of variation and correlation coefficient (both single sample, two sample cases).
- Tests of significance and confidence intervals using variance stabilizing transformations.
- Determination of the minimum sample size required to achieve normality by sample proportion, mean, standard deviation and correlation coefficient using CLT and variance stabilizing transformations.
- Tests for goodness of fit, independence and homogeneity using Pearsonian chi-square statistic.
- Test based on Fisher's exact test.

Reference Books

- Goon, A.M., Gupta, M.K. & Dasgupta, B. (2005): An Outline of Statistical Theory Vol- I & II, World Press
- Rohatgi, V. K. and Saleh, A.K. Md. E. (2009): An Introduction to Probability and Statistics. 2nd Edition, John Wiley and Sons
- Li, B. and Babu, G.J. (2019): A Graduate Course on Statistical Inference, Springer Science+Business Media, LLC
- Ferguson, T.S (1996): A Course in Large Sample Theory, Springer Science+Business Media Dordrecht
- Jiang, J. (2010): Large Sample Techniques for Statistics, Springer Science+Business Media, LLC
- DasGupta, A (2008): Asymptotic Theory of Statistics and Probability, Springer Science+Business Media, LLC
- Lehmann, E.L. (1999): Elements of Large-Sample Theory, Springer-Verlag New York, Inc.

STSDSC614T – Advance Regression & Applied Multivariate Analysis

[Credit 3]

[45 Lecture Hours]

Unit 1: Advance Regression Analysis

[20 Lecture Hours]

Model accuracy checking: Checking the linear relation between study and explanatory variables, residual analysis-standardized residual, Studentized residual, regression variable hull (RVH), PRESS residuals, R-student, residual plot. Normal probability plot. Partial regression and partial residual plot. Outlier detection based on R-student. Diagnostics of leverage and influence, leverage point, measure of influence- Cook's D-statistics, DEFFITS and DEBETAS, Transformation and weighting to correct model inadequacies, transformation to linearize the model, Box-Cox and Box-Tidwell transformation. Adjusted R square, Mallows' CP, AIC.

Unit 2: Applied Multivariate Analysis [15 Lecture Hours]

Inferential method under multivariate set-up. Multivariate normal likelihood, maximum likelihood estimator of mean vector μ and dispersion matrix Σ , distribution of sample mean and sample variance co-variance matrix (without derivation). Mahalanabis's Distance, Properties of Wishart distribution (without derivation), inference about mean vector, Hotelling's T^2 statistics and its distribution (without derivation). Likelihood ratio test for the mean vector of multivariate normal distribution. Confidence region and simultaneous comparison of component means.

Unit 3: Principal Component & Factor Analysis [10 Lecture Hours]

Population principal components, principal components obtained from standardized variables, principal components for covariance matrix with special structures. Summarizing sample variation by principal components, the number of principal components, interpretation of sample principal components.

Factor analysis: Orthogonal factor model, covariance structure model for orthogonal factors. Method of estimation – Principal component and principal factor method, maximum likelihood method. Discriminant Analysis.

STSDSC614P: List of Practical

[Credit 2]

[60 Lecture Hours]

- Fitting of ordinary linear regression equations with diagnostics.
- Multiple Correlation and Partial Correlation
- Multivariate Normal Distribution
- Principal Components Analysis
- Factor Analysis
- Discriminant Analysis

Reference Books

- Chatterjee, S. and Hadi, A. S. (2013): Regression Analysis by Example (5th Ed.). Wiley.
- Anderson, T.W. (2003): An Introduction to Multivariate Statistical Analysis, 3rdEdn., John Wiley.
- Muirhead, R.J. (1982): Aspects of Multivariate Statistical Theory, John Wiley.
- Kshirsagar, A.M. (1972): Multivariate Analysis, 1stEdn. Marcel Dekker.
- Johnson, R.A. and Wichern, D.W. (2007): Applied Multivariate Analysis, 6thEdn., Pearson & Prentice Hall
- Mukhopadhyay, P. (1996): Mathematical Statistics. New Central Book Agency.

STSDSC615T – Demography and SQC

[Credit 3]

[45 Lecture Hours]

Unit 1: Demographic Data and Measures [20 Lecture Hours]

Demographic events and processes. Sources of population data, Census, and registration. Errors in census and registration data. Rates and ratios of vital events. Crude death rate (CDR), specific death rate (SDR), infant mortality rate (IMR), and standardized death rates. Life Table: assumption, description, and uses. Force of mortality and graduation of mortality rates, Makeham's Graduation formula. Stable and stationary population. Crude birth rate (CBR), general fertility rate (GFR), specific fertility rate (SFR), and total Fertility rate (TFR). Measurement of population growth: crude rates of natural increase, Pearl's Vital index, gross reproduction rate (GRR) and net reproduction rate (NRR). Measurements of morbidity.

Unit 2: Population Size Estimation and Projection [8 Lecture Hours]

Population estimation, projection and forecasting: use of AP and GP methods for population estimates. Fitting of population curve for population forecasting using Rhode's method.

Unit 3: Statistical Quality Control[17 Lecture Hours]

Definition and perspective of SQC, different types of quality measure, chance and assignable causes of quality variation, rational sub-grouping. Process Control: statistical Control Charts- construction and statistical basis of 3- σ control charts. Control chart: X-bar, R-chart and s-chart. Control charts for attributes: np-chart, p-chart, c-chart and u-chart. Comparison between control charts for variables and control charts for attributes. Natural tolerance limits and specification limits. Product Control: principle of acceptance sampling plans, single and double sampling plans, their OC, AQL, LTPD, AOQ, AOQL, ASN, ATI functions with graphical interpretation, use and interpretation of Dodge and Romig's sampling inspection plan tables.

STSDSC615P: List of Practical

[Credit 2]

[60 Lecture Hours]

- Computation of Crude Birth Rate
- Computation of different Fertility Rates
- Computation of Vital Index, Reproduction Rates
- Fitting of population curve for population forecasting including Rhodes method
- Computation of different mortality rates and Makeham's Graduation formula
- Preparation of Life Table
- Construction and Interpretation of statistical control charts: X-bar, R chart, s-chart, np-chart, p-chart, c-chart
- Single sample inspection plan: Construction and interpretation of OC, AQL, LTPD, ASN, ATI, AOQ, AOQL curves

Reference Books

- Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol 2, 8th edition, The world Press, Kolkata
- Mukhopadhyay, P. (2011): Applied Statistics, 2nd edition revised reprint, Books and Allied (P) Ltd.
- Montgomery, D.C. (2009): Introduction to Statistical Quality control, 6th edition, Wiley India, Pvt. Ltd

Semester-VII

STSDSC716T-Mathematical Analysis II & Concept of Measure Theory	
[Credit5]	[75 Lecture Hours]
Unit1: Topology and Metric Space	[18 Lecture Hours]
<p>Topology: Neighbourhood set in $\mathbb{R}^d, d \geq 1$; interior point, cluster point, isolated point and boundary point — definition and relation between these points. Open set and closed set — definition, properties and associated results. Cantor set. Open cover, subcover, general definition of compact set, Heine-Borel Theorem. Preservation of compactness through continuity.</p> <p>Metrics: Definition of metric and metric space, different examples in $\mathbb{R}^d, d \geq 1$, associated results. Compactness in a metric space, continuous function on metric space.</p>	
Unit2: Riemann, Improper and Complex Integral	[15 Lecture Hours]
<p>Reimann Integration of Real-valued Functions, Darboux approach. Improperintegrals, convergence of integrals, simple tests, beta and gamma integral. Function of complex variable, Residue point, Residue Theorem and its application in complex integration with examples.</p>	
Unit3: Sigma-Field, Measures and Measurable Functions	[22 Lecture Hours]
<p>Fields, sigma-fields and generators, semi-fields, Borel sigma-field on \mathbb{R}. Monotone classes, monotone class theorem, π-λ theorem. Measures, finite and sigma-finite measures. Probability measures, properties. Independence of events, Borel-Cantelli lemma. Measurable functions and properties, non-measurable functions, Generated sigma-fields. Induced measures. Compositions. Extensions of measures, Carathéodory's theorem (statement). Lebesgue measure on \mathbb{R}, construction, properties. Random variables and vectors, probability distributions, distribution functions.</p>	
Unit4: Integration and Expectation	[20 Lecture Hours]
<p>Integration, simple, nonnegative, general measurable functions, integrability. Lebesgue Integration and Lebesgue-Stieltjes integration. Expectations, moments. Monotone Convergence Theorem, Dominated Convergence Theorem, Fatou's lemma. Absolute continuity and singularity of measures. Radon-Nikodym Theorem (Statement). Discrete and absolutely continuous distributions. Lebesgue's differentiation theorem (statement), probability densities.</p>	
ReferenceBooks	
<ul style="list-style-type: none"> • Bartle, R.G. and Sherbert, D.R. (2011): Introduction to Real Analysis, John Wiley & Sons, Inc. • Apostol, T. M. (1974): Mathematical Analysis, World Student Series Edition • Kumar, A. and Kumaresana, S. (2014): Basic Course in Real Analysis, Taylor & Francis Group, LLC • Thomson, B.S., Bruckner, J.B. and Bruckner, A.M. (2001): Elementary Real Analysis, Prentice-Hall, Inc. • Rădulescu, T-L.T., Rădulescu, V and Andreescu, T. (2009): Problems in Real 	

Analysis, Springer Science + Business Media, LLC

- Trench, W.F. (2003): Introduction to Real Analysis, Prentice Hall
- Abbott, S. (2012): Understanding Analysis. Springer New York
- Athreya, K. B., Lahiri, S. N. (2006): Measure Theory and Probability Theory, Springer
- Ash, R. B., Doleans-Dade, C. A. (2000): Probability and measure theory, Elsevier Science
- Billingsley, P. (2017): Probability and Measure, Wiley India
- Basu, a. K. (2012): Measure Theory and Probability, PHI learning, India
- Capinski, M., Kopp, P. E. (2004): Measure, Integral and Probability, Springer London, Germany

STSDSC717T - Statistical Computing

[Credit 3]

[45 Lecture Hours]

Unit 1: Simulation and Resampling Techniques

[20 Lecture Hours]

Simulation: Inversion method, Rejection Principle, Random sample generation from discrete distributions, convex polygon, mixture distributions, spherical distribution. Resampling techniques: Cross validation, Jackknife and Bootstrap: Parametric and nonparametric.

Unit 2: Robust Statistics and Nonparametric and Semi-parametric Regression [15 Lecture Hours]

Robust Statistics: Influence function, M-estimation, Different loss functions, robust regression, quantile regression.
Nonparametric and semi-parametric regression: kernel, local polynomial, regression tree.

Unit 3: EM and MM algorithm [10 Lecture Hours]

Missing data. Formulation of the EM and MM algorithms with illustrations.

STSDSC717P: List of Practical

[Credit 2]

[60 Lecture Hours]

- Drawing random samples using inversion method and rejection principle.
- Drawing random samples from discrete distributions, convex polygon, mixture distributions and spherical distributions.
- Estimation problems using cross validation, Jackknife and Bootstrap.
- Practical problems for depth functions, robust regression and quantile regression.
- Practical problems for nonparametric and semi-parametric regression
- Implementation of EM and MM algorithms.

Reference Books

- Voss, J. (2013). An Introduction to Statistical Computing: A Simulation based Approach. Wiley.
- Good, P. I. (2013). Introduction to Statistics through Resampling Methods and R, Second edition. Wiley.
- Huber, P. J. and Ronchetti, E. M. (2009): Robust Statistics, Second Edition, Wiley.
- Ruppert, D, Wand, M. D. and Carroll, R. J. (2003): Semiparametric Regression, Cambridge Series in Statistical and Probabilistic Mathematics.
- McLachlan, G. J. and Krishnan, T. (2008): The EM Algorithm and Extensions, 2nd Edition, Wiley Inter science.

Semester-VIII

STSDSC818T-Discrete Data Analysis*

[Credit3]

[45 Lecture Hours]

Unit 1: Analyzing Contingency Tables [17 Lecture Hours]

Measures of variability, Measures of association using variability, Association in three-way tables: Partial and Multiple association, Simpson's paradox, Inferences using odds ratios

Matched pairs data, Tests for marginal homogeneity for 2x2 and nxn tables, McNeymar Test, Cohen's Kappa.

Unit2: Generalized Linear Model

[18 Lecture Hours]

Generalized linear Model, Components of a generalized linear model, Random component, systematic component, Link function.

Generalized linear model for binary data, Logistic and probit regression model, Multiple logistic regression. Model fitting by using score function. Multicategory Logit models. Latent variable approach. Generalized linear model for count data, Poisson regression. Zero inflated Poisson (ZIP) model.

Unit3: Model for contingency table

[10 Lecture Hours]

Log linear model of independence for two-way table, Interpretation of the parameters in independence model, saturated model for two-way table. The log-linear-logistic connection.

STSDSC818P: List of Practical

[Credit2]

[60 Lecture Hours]

- Practical using measures of association using variability.
- Evaluating partial and multiple association for three-way tables.
- Testing problems using odds ratios.
- Fitting a logit model
- Fitting a probit model
- Fitting of multiple logistic regression.
- Fitting multi-category logistic regression.
- Fitting Poisson regression.
- Fitting ZIP model.
- Numerical problems form models for contingency table.

ReferenceBooks

- P.Mc Cullagh & J.A.Nelder.(1995):Generalized Linear Models. Chapman and Hall.
- J.F. Simonoff: Analyzing Categorical Data.
- S.E. Fienberg: The Analysis of Cross Classified data
- Michael S. Lewis Beck.: Basic Statistics.
- Agresti,A. (2007):An Introduction to Categorical data analysis. Wiley.

STSDSC819T – Financial Time Series Analysis*

[Credit 3]

[45 Lecture Hours]

Unit 1: Non-stationary Financial Time Series Data[16 Lecture Hours]

Recapitulation of Stationary Time Series Analysis: Strongly and weakly Stationary processes, ergodicity, stationarity of time series and its importance in analysis, Wald Decomposition, Correlogram, ACF and PACF, AR, MA and ARMA model, Stationarity and invertibility conditions, fitting of AR, MA and ARMA processes, diagnostics checking for stationary processes, forecasting with these models.

Non-Stationary Time Series: Asset data over time and their preliminary graphical analysis, return and log-return data of a financial asset. Non-stationarity: random walk and other stochastic models. Types of non-stationarity - integrated time series and time series with stochastic trend. Test of stationarity - graphical method, correlogram, unit root test, Dickey-Fuller (DF) test, and Augmented DF test. ARMA family of models - ARIMA and Seasonal ARIMA modelling using Box-Jenkins approach.

Unit 2: Volatility, its Characteristics & Modelling [16 Lecture Hours]

Concept of volatility, volatility clustering, conditional heteroscedasticity. Volatility modelling and forecasting - ARCH, GARCH, E-GARCH models and their illustrative applications in real data and associated forecasting. Option Pricing: importance, implied volatility, Black-Scholes Model and the use of volatility. Limitations of Black-Scholes Model, case studies with real-life financial data.

Unit 3: Asset Portfolio Risk Measure & CAPM [10 Lecture Hours]

Definition of risk, Value at Risk (VaR) as a quantile measure of risk, parametric and nonparametric approaches to VaR estimation. Expected Shortfall. Density forecasting of risk using ARCH models. Risk measure, real data applications. Capital asset pricing model (CAPM): share price, modern portfolio theory, CAPM modelling. Empirical evaluation of pricing data: data selection, modification of CAPM, statistical methods for testing CAPM such as two-pass regression tests. Case studies with real-life financial data.

STSDSC819P: List of Practical

[Credit 2]

[60 Lecture Hours]

- Basic time series analysis, decomposition, ACF, PACF, ARMA model identification, fitting
- Forecasting using fitted ARMA model(s) and comparing the results with the forecasting using the Exponential Smoothing technique
- Analysis of Non-stationary Time Series data: applications of unit root test, DF and ADF tests, fitting and forecasting of ARIMA and SARIMA models
- Identification of the presence of volatility in a given dataset, applications of ARCH, GARCH models in the suitable data
- Black-Scholes Model: Applications to real data

- VaR estimation, measuring risk from real data
- CAPM: tests of CAPM

Reference Books

- Tsay, R.S. (2010): Analysis of Financial Time Series (3rd Ed.), John Wiley.
- Box, George E. P., Gwilym M. Jenkins, Gregory C. Reinsel, Greta M. Ljung. (2015). Time Series Analysis, Forecasting and Control, Wiley
- Brockwell & Davis (2016): Introduction to Time Series and Forecasting (3rd Ed.), Springer.
- Brooks, C. (2019): Introductory Econometrics for Finance (4th Ed.), Cambridge University Press.
- Chan, N.H. (2002): Time Series: Applications to Finance, Wiley
- Mills, T. C. (1999): The Econometric Modelling of Financial Time Series (2nd Ed.), Cambridge University Press.

STSDSC820T- Introduction to Bayesian Inference*

[Credit 3]

[45 Lecture Hours]

Unit 1: Introduction [15 Lecture Hours]

Elements of decision theory - Preliminary ideas of decision rules, loss and risk. Overview and comparison of two paradigms – Classical statistical analysis and Bayesian analysis. Relative advantages and disadvantages, motivation for choice of different priors.

Bayesian Inference – estimation, testing, interval estimation and prediction for some common models and common priors, Hierarchical Bayes, brief discussions on Bayesian computational techniques and their applications.

Unit 2: Bayesian Analysis

[18 Lecture Hours]

Bayesian linear regression with conjugate priors, Bayesian model selection, Bayesian Information Criterion (BIC). Monte Carlo Method, Markov chains and MCMC, Gibbs sampling with examples in R or WinBUGS, The Metropolis-Hastings algorithm.

Generalized linear models and categorical data, longitudinal models, Bayesian multiple imputation. Bayesian hypothesis testing (One-sided and Two-sided Example), The Bayes factor, test for comparing two population means.

Unit 3: Applications

[12 Lecture Hours]

Hierarchical Bayes, Exchangeability, Hierarchical Bayesian shrinkage and Bayesian estimation, Empirical Bayes estimation (with examples), James-Stein estimator, Comparison of Hierarchical vs. Empirical Bayes.

Bayesian Probit and Logistic Regression (Multi-category Ordinal Response).

STSDSC820P: List of Practical**[Credit 2]****[60 Lecture Hours]**

- Problems on Bayesian estimation.
- Problems on Bayesian hypothesis testing.
- Problems on Bayesian interval estimation.
- Problems on Bayesian model selection.
- Problems on Bayesian Monte Carlo and MCMC
- Problems on Bayesian Probit and Logistic Regression.
- Related problems on other topics of Unit 1 – 3.

Reference Books

- Berger, J.O. (1985): Statistical Decision Theory and Bayesian Analysis, Springer New York
- Ghosh, J.K., Delampady, M. and Samanta, T. (2006): An Introduction to Bayesian Analysis: Theory and Methods, Springer New York
- Lee, P.M. (2012): Bayesian Statistics: An Introduction, Wiley, United Kingdom
- Robert, C. (2007): The Bayesian choice: From Decision-Theoretic Foundations to Computational Implementation, Springer, Ukraine

STSDSC821T- Advanced Sample Survey***[Credit 3]****[45 Lecture Hours]****Unit 1: Probability Sampling [10 Lecture Hours]**

Varying probability sampling: PPS sampling with replacement (WR), cumulative total method and Lahiri's method, estimation of population mean and variances, varying probability sampling without replacement, Midzuno scheme of sampling and its modification, Horvitz-Thompson estimator and its variance estimation, Yates-Grundy form of variance. Hijek estimator.

Unit 2: Unified Theory of Sampling [15 Lecture Hours]

Unified theory of sampling: Model unbiasedness and design unbiasedness, Godambe Joshi lower bound and optimality of estimators. Generalized regression estimator, QR-class of predictor, calibration technique in survey sampling.

Unit 3: Small Area Estimation [20 Lecture Hours]

Small area estimation: Direct, synthetic, and composite estimator. Explicit model approach- basic area level model of Fay and Herriot, basic unit level model and general linear model, E-BLUP approach, empirical Bayes and hierarchical Bayes approach of estimation under normality of error and prior distribution.

STSDSC821P: List of Practical**[Credit 2]****[60 Lecture Hours]**

- Practical problems based on unit 1.
- Assignment related to unit 2.
- Assignment related to unit 3.

Reference Books

- Cochran, W.G. (1984): Sampling Techniques (3rd Ed.), Wiley Eastern
- Sukhatme, P.V., Sukhatme, B.V. Sukhatme, S. (1984). Sampling Theories of Survey with Application, IOWA State University Press and Indian Society of Agricultural Statistics
- Murthy, M.N. (1977): Sampling Theory & Statistical Methods, Statistical Pub. Society, Calcutta
- Des Raj and Chandhok P. (1998): Sample Survey Theory, Narosa Publishing House
- Chaudhuri, A. and Vos, J. W. E. (1988): Unified Theory and Strategies of Survey Sampling, Elsevier Science Ltd.
- Rao, J. N. K. and Molina, I. (2015): Small Area Estimation (2nd Ed.), Wiley, Hoboken
- Morales, D., Esteban, M. D., Perez, A. and Hobza, T. (2021): A Course on Small Area Estimation and Mixed Models. Springer

STSDSC822T - Survival Analysis and Clinical Trials***[Credit 3]****[45 Lecture Hours]****Unit 1: Preliminary Topics and censoring [15 Lecture Hours]**

Survival Analysis: Functions of survival times, survival distributions and their applications: exponential, gamma, Weibull, lognormal, etc. Censoring Schemes: Type I, Type II and progressive or random censoring with biological examples. Estimation of mean survival time and variance of the estimator for Type I and Type II censored data with numerical examples. Survival function, Hazard Function and their relationship. Kaplan-Meier and Nelson-Aalen methods for estimating survival function and variance of the estimator.

Unit 2: Testing and model fitting [15 Lecture Hours]

Log rank test, Introduction to Cox Proportional Hazard (PH) regression, Model diagnostic checking, Accelerated failure time (AFT) model. Competing risk theory: Gray's test and Fine-Gray model.

Unit 3: Clinical Trials [15 Lecture Hours]

Clinical Trials: definition, ethics, masking and blinding, phases, Multicenter trials, objectives and endpoints of clinical trials, Randomized control trial. Dose-Response study, Basics of Adaptive Designs. Discussions of different historical clinical trials.

STSDSC822P: List of Practical**[Credit 2]****[60 Lecture Hours]**

- Fitting and plotting Kaplan-Meier estimator
- Performing log rank test
- Fitting Cox PH model and checking model diagnostics
- Fitting AFT model
- Performing Gray's test
- Fitting Fine-Gray model
- Problems for dose-response study

Reference Books

- Kalbfleisch J. D. and Prentice R. (1980): The Statistical Analysis of failure Time data, John Wiley.
- Kleinbaum, D.G. (1996): Survival Analysis, Springer
- Lee, Elisa, T. (1992). Statistical Methods for Survival Data Analysis, John Wiley & Sons.
- Miller, R.G. (1981). Survival Analysis, John Wiley & Sons.
- Piantadosi. S. (1997): Clinical Trials: A Methodologic Perspective. Wiley and Sons.
- Friedman, L. M. Furburg, C. Demets, D. L. (1998): Fundamentals of Clinical Trials. Springer Verlag.
- Marubeni. E. and Valsecchi. M. G. (1994): Analyzing Survival Data from Clinical Trials and Observational Studies, Wiley and Sons.

** The department offer the required number of course(s) to the students among the five courses in Semester VIII.*