M.Sc. in Statistics

with specialization in

Industrial Statistics/ Business Analytics / Biostatistics

(July 2022 onwards)

Programme Educational Objectives

PEO 1: To introduce students to the meaning, concepts and scope of the basic Statistical and Mathematical tools applicable for decision-making.

PEO 2: To familiarize students with the different tools and techniques used for collecting, handling, and managing and interpreting data.

PEO 3: To make the students acquainted with methods of data analysis and interpretation with focus on statistical tools.

PEO 4: To enable students to perform data visualization and analysis using modern statistical software packages such as Python, SPSS, R etc.

PEO 5: To enable students to get familiarized with different techniques of data driven statistical inference.

PEO 6: To encourage students towards various types of research and methodologies of conducting research.

Programme Outcomes

PO 1: Understanding the basic Statistical and Mathematical tools applicable for decision-making.

PO 2: Designing questionnaires or perform other survey methods to collect data from primary sources supporting independent research initiatives.

PO 3: Applying advanced quantitative and statistical tools envisaged in the course in collecting, analyzing and interpreting data.

PO 4: Using statistical software packages in the process of generation of graphs, charts, and other forms of data visualization, data analysis and data interpretation.

PO 5: Enabling to independently solve real life problems.

PO 6: Assisting the policy makers in framing plans on different aspects of national development.

PO 7: Understanding and predicting the movement different data series related to industry, business and healthcare.

Curriculum for M.Sc. Statistics

Total Credits: 96

Total Marks: 2500

Semester wise Details

Semester – I							
Number of Papers: 6							
Course code	**Course Title**	**Course Type**	**Credits in each course**				
			Theory	**Practical**	**Tutorial**	**Total Credits**	**Total Marks**
MSTR110T	Mathematical Analysis	Core	4	0	0	4	20(CIA) + 80(T)
MSTR120T	Probability Theory	Core	4	0	0	4	20(CIA) + 80(T)
MSTR130C	Statistical Inference-I (Estimation Theory / Hypothesis Testing)	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR140C	Linear Algebra & Linear Models	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR150C	Sampling Techniques	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR160P	R Programming	Core	0	4	0	4	20(CIA) + 80(P)
Total			**17**	**7**	**0**	**24**	**600**

***CIA: Continuous Internal Assessment, T: Theory, P: Practical and C: Combine**

Semester –II							
Number of Papers: 6							
Course code	**Course Title**	**Course Type**	**Credits in each course**				
			Theory	**Practical**	**Tutorial**	**Total Credits**	**Total Marks**
MSTR210T	Stochastic Processes	Core	4	0	0	4	20(CIA) + 80(T)
MSTR220C	Multivariate Analysis	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR230C	Statistical Inference II (Large sample Theory/Nonparametric Methods)	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR240C	Regression Analysis I	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR250C	Design of Experiments	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR260P	Python	Core	0	4	0	4	20(CIA) + 80(P)
Total Credits			**16**	**8**	**0**	**24**	**600**

***CIA: Continuous Internal Assessment, T: Theory, P: Practical and C: Combine**

***CIA: Continuous Internal Assessment, T: Theory, P: Practical and C: Combine**

Semester –III							
Number of Papers: 7							
Course code	**Course Title**	**Course Type**	**Credits in each course**				
			Theory	**Practical**	**Tutorial**	**Total Credits**	**Total Marks**
MSTR310C	Bayesian Inference	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR320C	Applied Multivariate Analysis	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR330C	Regression Analysis II	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR340C	Time Series Analysis and Development Statistics	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR350C	Advanced Data Analytic Techniques (Resampling Techniques/ Missing Data/Longitudinal Data)	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR361C	*Demography*	Elective	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR362C	*Optimization Techniques*						
MSTR363C	*Actuarial Statistics*						
MSTR370J	Internship	Project	0	0	2	2	[20(CIA) + 60(T)] + 20(P)
Total Credits			**18**	**6**	**2**	**26**	**700**

NOTE:

In this semester, students will be required to select one elective paper in Statistics from among from Demography, Actuarial Statistics and Optimization Techniques. However, all the papers may not be offered in a particular year, and it will be at the discretion of the Department to decide which papers to offer in the particular year.

Semester –IV							
Number of Papers: 6							
Course code	**Course Title**	**Course Type**	**Credits in each course**				
			Theory	**Practical**	**Tutorial**	**Credits**	**Total Marks**
MSTR410C	Statistical Analysis of Big Data	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
Specialization Bucket: 01 (Industrial Statistics)							
MSTR421C	*Operations Research*	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR431C	*Reliability Theory*	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR441C	*Statistical Quality Management*	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
Specialization Bucket: 02							
MSTR422C	*Econometrics*	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR432C	*Financial Time Series*	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR442C	*Machine Learning in Finance / Financial Econometrics*	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
Specialization Bucket: 03							
MSTR423C	*Survival Analysis*	Core	3	1	0	4	[20(CIA) + 60(T)]

							+ 20(P)
MSTR433C	*Clinical Trials*	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR443C	*Statistical Genetics / Epidemiology*	Core	3	1	0	4	[20(CIA) + 60(T)] + 20(P)
MSTR450J	Project	Core	0	0	6	6	40(CIA) + 120(T) +40(P)
Total Credits			**12**	**4**	**6**	**22**	**600**

***CIA: Continuous Internal Assessment, T: Theory, P: Practical and C: Combine**

NOTE:

In the Semester IV, students shall be required to select one specialization Bucket from the three specializations bucket as listed above. However, all of these may not be offered in a particular year, and it will be at the discretion of the Department to decide which paper or which Module to offer in a particular year.

St. Xavier's University, Kolkata
M.Sc. Statistics Syllabus
(July 2022 onwards)

SEMESTER: I

Mathematical Analysis

Credit: 4 (T) + 0 (P) **Paper Code: MSTR110T**
Marks: 100 (T) + 0 (P)

Real Number System: Axioms, Bounded and unbounded subsets of the real line, supremum, infimum, intervals, Countable and Uncountable Set, Archimedean property of R, Density property of R, Extended real line.

Sequence: subsequence, convergence, divergence, bounded sequences, limits superior and inferior, monotone sequences, Cauchy sequences, Bolzano-Weierstrass Theorem.

Series of real numbers: Different tests of convergence, Absolute convergence, conditional convergence, Rearrangement of series.

Topological properties: closed and open sets, limit points, closures, interiors. Compact sets, Heine-Borel Theorem (statement only).

Limits, continuity, uniform continuity, intermediate value theorem, differentiability, mean value theorem, Taylor's theorem (statement), extrema.

Riemann integral: Upper and lower integral, Fundamental theorem of calculus.

Sequences and series of functions: uniform convergence, power series, term-by-term differentiation and integration.

Multivariate calculus: partial, directional and total derivatives, mean value theorem.

References :

1. T.M.Apostol : Mathematical Analysis
2. W.Rudin: Principles of Mathematical Analysis
3. D. R. Sherbert & R.G. Bartle: Introduction to Real Analysis, 4ed
4. S. R. Ghorpade & B.V. Limaye: A Course in Calculus and Real Analysis
5. S. R. Ghorpade & B.V. Limaye: A Course in Multivariable Calculus and Analysis

---

Probability Theory

Credit: 4 (T) + 0 (P) **Paper Code: MSTR120T**
Marks: 100 (T) + 0 (P)

Fields, semi-fields, sigma-fields, measures, sigma-finite/finite/probability measures, properties, statement of Caratheodory extension theorem. Monotone class theorem, Dynkin's pi-lambda theorem.

Measurable functions and properties, Generated sigma-fields. Induced measures. Compositions. Random variables, probability distributions, distribution functions. Product measurable spaces; characterizations; random vectors, multivariate distributions.

Integration: simple, nonnegative, general measurable functions, integrability. Expectations, moments. Monotone Convergence Theorem, Fatou's lemma,Dominated Convergence Theorem.

Absolute continuity and singularity of measures. Radon-Nikodym Theorem (Statement). Discrete and absolutely continuous distributions, probability densities.

Product measures. Fubini's theorem. Lebesgue measure on R^k. Independent random variables. Asymptotics of independent random variables: tail sigma-field, Kolmogorov's 0-1 law.

Lp spaces, Lp-convergence of random variables, connections with other modes of convergence.

Convergence in distribution. Connections with other modes of convergence. Slutsky's theorem.

Characteristic functions. Convolutions. Inversion and Continuity theorems.

Weak and strong laws of large numbers. Central Limit Theorems.

References:

1. S. Resnick	:	A Probability Path
2. P. Billingsley	:	Probability and Measure
3. R.Ash& C. Doleans-Dade	:	Probability and Measure Theory
4. K. B. Athreya & S. N. Lahiri	:	Measure Theory and Probability Theory
5. A. K. Basu	:	Measure Theory and Probability

Statistical Inference-I (Estimation Theory / Hypothesis Testing)

Credit: 3 (T) + 1 (P) **Paper Code: MSTR130C**
Marks: 75 (T) + 25 (P)

Estimation Theory

Sufficiency and completeness. Rao-Blackwell and Lehmann-Scheffe Theorems, Minimum Variance Unbiased Estimators.Concepts of Likelihood function, Maximum likelihood Estimators, consistency.

Hypotheses Testing

Review of notions of nonrandomized and randomized tests, level, size, p-value, power function, Fundamental Neyman-Pearson lemma, UMP Tests. Monotone Likelihood Ratio.

Generalized Neyman-Pearson Lemma. UMPU Tests for Simple and Composite hypotheses.

Confidence sets, relation with hypothesis testing, UMA and UMAU confidence intervals.

***References* :**

1. E.L.Lehman : Testing Statistical Hypotheses
2. S.Zacks : The Theory of Statistical Inference
3. C.R.Rao : Linear Statistical Inference and its Applications
4. E.L.Lehmann : Theory of Point Estimation
5. T.S.Ferguson : Mathematical Statistics

Linear Algebra &Linear Models

Credit: 3 (T) + 1 (P) **Paper Code: MSTR140C**
Marks: 75 (T) + 25 (P)

Linear Algebra

Vector spaces over real fields, subspace, linear independence, basis and dimension, sum and intersection of subspaces, direct sum, complement and projection.

Homogeneous and non-homogeneous systems of linear equations, condition for consistency,

g-inverse and its elementary properties.

Characteristic roots, eigenvectors, spectral representation of real symmetric matrices, singular value decomposition.

Quadratic form, determinant criteria for n.n.d. and p.d. quadratic forms.

Linear Models

Gauss Markov Model: Estimable function, error function, BLUE, Gauss Markov theorem. Correlated set-up, least squares estimate with restriction on parameters.

Linear Set, General linear hypothesis – related sampling distribution, Multiple comparison techniques due to Scheffe and Tukey.

Analysis of variance: Balanced classification, Fixed Effects Model, Random Effects Model and Mixed Effects Model; Inference on Variance components. Analysis of covariance.

References :

1. C. R. Rao : Linear Statistical Inference and Its Applications.
2. A. Ramachandra Rao and P. Bhimasankaram : Linear Algebra.
3. Gilbert Strang : Introduction to Linear Algebra, Fifth Edition
4. R. B. Bapat : Linear Algebra and Linear Models
5. H.Scheffe : The Analysis of Variance
6. S.R.Searle : Linear Models
7. G.A.F.Seber : Linear Regression Analysis
8. N.C. Giri : Analysis of Variance
9. K. Hoffman and R. Kunze : Linear Algebra

Sampling Techniques

Credit:3 (T) + 1 (P) **Paper Code: MSTR150C**
Marks: 75 (T) + 25 (P)

Probability sampling from a finite population – Notions of sampling design, sampling scheme, inclusion probabilites, Horvitz-Thompson estimator of a population total.

Basic sampling schemes – Simple random sampling with and without replacement, stratified sampling, Unequal probability sampling with and without replacement, Systematic sampling. Related estimators of population total/mean, their variances and variance estimators – Mean per distinct unit in simple random with replacement sampling, Hansen-Hurwitz estimator in unequal probability sampling with replacement, Des Raj and Murthy's estimator (for sample size two) in unequal probability sampling without replacement.

Ratio, Product, Difference and Regression estimators.

Two-stage sampling with unequal number of second stage units and simple random sampling without replacement / unequal probability sampling with replacement at first stage, Ratio estimation in two-stage sampling.

Concept of sample size estimation. Determination of sample size for estimating mean and proportions. Design effect, sample size for comparisons of two means or proportions.

References:

1. W.G. Cochran : Sampling Techniques, 3rd ed.
2. Des Raj & Chandak : Sampling Theory.
3. P. Mukhopadhyay : Theory & Methods of Survey Sampling.
4. M.N. Murthy : Sampling Theory and Methods.
5. A.S. Hedayat and Bikas K. Sinha : Design and Inference in Finite Population Sampling

---

R Programming

Credit:0 (T) + 4 (P) **Paper Code: MSTR160P**
Marks: 0 (T) + 100 (P)

Introduction : History and Overview of R, the CRAN, Installing the R Software, The R-Console and the R-Script. Saving and Accessing Files. Libraries in R. Loading and Installing Packages in R. The *quit()* and the *history()* commands.

R as a calculator: Basic Mathematical Functions. The Base Library. Defining variables, calling variables, Unary and Binary Operators on Variables.

Modes of Data Storage : Vectors, Matrices, Data Frames, Lists. The *c(), edit()* and *scan()* commands. Defining Attributes. Creating Patterned Data – the *rep()* and *seq()* commands. Extracting rows and columns in data frames and lists. Assigning names to columns of data frames and matrices and rows of lists. The *$* operator. The *attach()/detach()* command. Conditional selections and subsetting of objects. The *length()* command. Merging multiple vectors or columns of different data frames into one - The *cbind(), rbind()* and *merge()* commands. Inter-Conversions of the various modes of storages.

Diagrammatic representations of Non-Frequency Data : the *plot()* command. Line Diagram, Bar (Horizontal and Vertical) diagrams, Multiple Bar diagrams, Multiple Line diagrams, Pie and Subdivided Charts. Adding

legends, Title, labels, limits on the axis. The *'graphics'* package and the *'ggplot2'* package. The *par()* parameter and its arguments.

Diagrammatic representations of Frequency Data : Frequency Distributions, the *table()* command. Column Diagrams and Histograms. Box Plots - the *summary()* command. Cumulative Frequency Diagrams. Juxtaposing frequency curves over histograms.

Univariate Statistics: Descriptive Measures of Central Tendency, Dispersion, Skewness and Kurtosis. The *'moments'* package and its functions.

Bivariate Statistics: Scatterplot, Various forms of correlations. Regression Theory – the *lm()* command, polynomial regression. Residual Plots.

Linear Algebra: Algebra of Matrices. The *'Matrix'* package. Obtaining Determinants, Trace, Rank and Inverse of a Matrix. Obtaining row reduced forms of matrices, obtaining an orthonormal basis. Eigen Values and Eigen Vectors. Solving a system of equations. Diagonalisation of Matrices.

Programming in R: Control Statements: if, if else. Loop Structures: for, while, repeat. User defined functions – Passing arguments, calling functions and returning values. Probability functions of random variables.

Statistical Simulations: Drawing Random Samples from different finite and infinite probability distributions – the *set.seed()* command. Illustrations through statistical problems (probability estimates by long-run relative frequencies, Bias and MSE's of estimates, coverage of Confidence Intervals, calculating empirical level and power of tests). Optimisation of Functions – the *optim()* function and its various arguments.

File Handling: Importing and Exporting Data from/to other softwares.

References:

1. Dalgaard, P : Introductory Statistics with R, Springer Pubications, 2nd edition, 2008.
2. Maindonald, J. & Braun, J. : Data Analysis and Graphics Using R , Cambridge University Press, Cambridge, 2nd edition, 2007.
3. Faraway, J. J. : Linear Models with R ,Chapman& Hall/CRC Texts in Statistical Science.

SEMESTER: II

Stochastic Processes

Credit: 4 (T) + 0 (P) **Paper Code: MSTR210T**
Marks: 100 (T) + 0 (P)

Stochastic Processes: Introduction, classifications of stochastic processes, counting processes.

Discrete-time Markov chains: time-homogeneity, one-step & multi-step transition probabilities, Chapman-Kolmogorov equations, Markov times, strong Markov property, classification of states, stationary distributions, periodicity, ergodicity, convergence, reversibility, elementary MCMC methods. Examples: birth-and-death processes, branching processes.

Poisson process: postulates, properties; compound PP, non-homogeneous PP.

Renewal Theory: renewal processes, renewal function, elementary renewal theorem, applications, inspection paradox, asymptotic normality of a renewal process.

Blackwell's and key renewal theorems (statements) in non-lattice and lattice cases, applications.

Conditional Expectation, Martingale.

Brownian Motion, Donsker Theorem (Statement) and its Applications.

References:

1. S.Karlin&H.M.Taylor : A First Course in Stochastic Processes
2. J. Medhi : Stochastic Process
3. D.R. Cox : Renewal Theory
4. S.Ross : Stochastic Process
5. A.K.Basu : Stochastic Process
6. P.G.Hoel, S.C.Port&C.J.Stone : An Introduction to Stochastic Process
7. R.N.Bhattacharyya& E. Waymire : Stochastic Processes and Applications

---

Multivariate Analysis

Credit: 3 (T) + 1 (P) **Paper Code: MSTR220C**
Marks: 75 (T) + 25 (P)

Non-central χ^2, t & F distributions – definitions and selected properties.

Distribution of quadratic forms – Cochran's theorem.

Sampling from Multivariate normal distribution – independence of sample mean vector and variance-covariance matrix. Wishart distribution.

Distributions of partial and multiple correlation coefficients and regression coefficients.

Hotelling T^2 and Mahalanobis D^2 applications in testing and confidence set construction.

Multivariate linear model: estimation of parameters, tests of linear hypotheses, Multivariate Analysis of variance of one and two way classified data (only LR test).

Canonical Correlation: Population and sample canonical variables and canonical correlations and their interpretations.

References :

1. C.R. Rao : Linear Statistical Inference and its Applications
2. T.W. Anderson : Introduction to Multivariate Analysis
3. A.M. Khirsagar : Multivariate Analysis
4. S.S. Wilks : Mathematical Statistics
5. M.S. Srivastava & C.G. Khatri: Introduction to Multivariate Statistics
6. R.J. Muirhead : Aspects of Multivariate Statistical Theory

---

Statistical Inference II

Credit:3 (T) + 1 (P) **Paper Code: MSTR230C**
Marks: 75 (T) + 25 (P)

Large Sample Theory:
Review of various modes of convergence of random variables and central limit theorems.Continuous mapping theorem. Cramer-Wold device and multivariate central limit theorem. Slutsky's theorem.

Asymptotic distribution of transformed statistics. Delta method. Derivation of the variance stabilizing formula. Asymptotic distribution of functions of sample moments like sample correlation coefficient, coefficient of variation.

Consistency and Asymptotic Efficiency of Estimators, Maximum Likelihood estimators and their large sample properties.

Asymptotic distributions and properties of Likelihood ratio tests, Rao's score test andWald's tests in the simple hypothesis case.

Nonparametric Methods:
Elementary concepts and properties of U-statistics and Linear Rank Statistics, Single sample location, location cum symmetry and goodness-of-fit problems.

Two-sample location, scale and homogeneity problems, Multi-sample location problem, Bivariate association problem.

Related nonparametric interval estimation;

Density Estimation

Concept of asymptotic relative efficiency.

References:

1. R.J. Serfling	:	Approximation Theorems of Mathematical Statistics
2. E.L. Lehmann	:	Large Sample Theory
3. A. W. van der Vaart	:	Asymptotic Statistics
4. J.D. Gibbons	:	Nonparametric Inference
5. T.P.Hettmansperger	:	Statistical Inference based on ranks
6. D.A.S.Fraser	:	Nonparametric methods in Statistics

Regression Analysis I

Credit:3 (T) + 1 (P) **Paper Code: MSTR240C**
Marks: 75 (T) + 25 (P)

Transformation of Variables – Box-Cox and Box-Tidwell models.
Stepwise regression.
Model selection (adjusted R2, Cp criteria, AIC).
Multicollinearity – detection and remedial measures.
Dummy variables, piecewise regression, splines and scatter plot smoothing.
Detection of outliers and influential observations: residuals and leverages, DFBETA, DFFIT, Cook's Distance.
Departures from the Gauss-Markov set-up: Heteroscedasticity and Autocorrelation – consequences, detection and remedies.
Checking for normality: Q-Q plots, Normal Probability plot, Shapiro-Wilks test.

References :

N.R. Draper & H. Smith	:	Applied Regression Analysis
D.W. Belsley, E. Kuh & R.E. Welsch	:	Regression Diagnostics – identifying Influential data & sources of collinearity
R.D. Cook & S. Weisberg	:	Residual and its Influence in Regression
S. Chatterjee & A.S. Hadi	:	Regression Analysis by Example
J. Johnston	:	Econometric Methods (3rd ed.)

G.G. Judge, W.E. Griffith, R.C. Hill, W. Lutkepohl& T.C. Lee	:	The Theory and Practice of Econometrics (2nd ed.)
T.P. Ryan	:	Modern Regression Methods (2nd ed.)

Design of Experiments

Credit: 3 (T) + 1 (P) **Paper Code: MSTR250C**
Marks: 75 (T) + 25 (P)

Block Designs: Connectedness, Orthogonality, Balance and Efficiency; Resolvable designs. Properties of BIB designs, Designs derived from BIB designs.
Intrablock analysis of BIB, Recovery of inter-block information in BIB designs; Row column and Youden Square designs, Missing plot technique. Elementary ideas of Lattice and PBIB designs.
Construction of mutually orthogonal Latin Squares (MOLS); Construction of BIBD through MOLS and other ways.
Factorial designs: Analysis, Confounding and balancing in Symmetric Factorials.
Response Surface Designs.

References :

M.C. Chakraborty	:	Mathematics of Design and Analysis of Experiments
A. Dey	:	Theory of Block Designs
D. Raghavarao	:	Constructions & Combinatorial Problems in DesignofExpts.
D.Raghavarao&L.V.Padgett	:	Block Design: Analysis, Combinatorics and Applications
D. G. Kabe and A. K. Gupta	:	Experimental Designs: Exercises and Solutions
G. Casella	:	Statistical Design
T. P. Ryan	:	Modern Experimental Design
C. F. J. Wu & M. S. Hamada	:	Experiments: Planning, Analysis and Optimization(2nd ed.)
D.C. Montogomery	:	Design and Analysis of Experiments

Python

Credit:0 (T) + 4 (P) **Paper Code: MSTR260P**
Marks: 0 (T) + 100 (P)

Introduction: Anaconda(individual edition) and Python Shell; Numbers, Variables, Comparisons and Logic, String, List and Tuple with examples.

Python Language I: Loops, Control Flow, File Input/Output and Functions with examples. Lambda, map, filter. Functions with List, Tuples, Dictionary.

Simple Plots: Basic plotting, Labels, Legends and Customization with examples.

Python Language II: Handling errors and exceptions, Dictionaries and Sets with examples; Some useful Python Idioms.

Spyder and Jupyter: How to edit, run and save Python codes using Spyder and Jupyter.

NumPy: Basic array methods; Reading and Writing an array to a File; Functions with Numpy arrays, Descriptive Statistics with NumPy; Polynomial fitting and root-finding; Basic matrix operations; Random number generation, Sampling from Uniform, Normal, Binomial and Poisson distributions, With and Without replacement sampling, Permutations.

Matplotlib: Line and bar diagram, Pie chart, Histogram, Box-Whisker plot, Error bar, Scatter plot; Plot customization and refinement; Multiple plots with Seaborn.

SciPy: Unconstrained optimization and root-finding of functions of a single variable, Weighted and unweighted least square fit; Discrete and Continuous probability distributions.

Pandas: Introduction to Series and Data Frame with Indexing and Slicing; Reading and Writing of Text, CSV and Excel File; Web Scraping; Data Cleaning and Data Grouping. Plotting on Pandas data frame, Statistical Function on Pandas data frames.

Statistics, Statsmodels, rp2: Simulating the sampling distributions of sample mean, sample proportion and sample variance with random samples drawn from Normal and Non-Normal (Uniform, Log-normal, Exponential) distributions; Critical Z score, Hypothesis testing, Z test, Left tailed, right tailed, two tailed, One-sample, Two-sample and Paired-sample tests and confidence intervals; Sign and Signed-rank tests; Evaluating power of a test; Multiple linear regression, regression diagnostics (residual analysis), prediction and confidence intervals; ANOVA (One way) with multiple comparisons; Logistic regression, LDA and ROC curve.

References

1. Learning Scientific Programming with Python (2020): Christian Hill, Cambridge University Press
2. Foundations of Statistics for Data Scientists with R & Python (2022): Alan Agresti, Maria Kateri; CRC Press
3. https://www.anaconda.com/products/distribution : installer and user-guide
4. https://docs.spyder-ide.org/current/index.html
5. http://stat4ds.rwth-aachen.de

SEMESTER: III

Bayesian Inference

Credit:2 (T) + 1 (P) **Paper Code: MSTR310C**
Marks:50 (T) + 25 (P)

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 – point estimation, interval estimation and prediction for some common models and common priors.
Bayesian Linear Regression with Conjugate priors, Bayesian Model Selection, Bayesian Information Criterion. Monte Carlo Method, Markov chains and MCMC, Gibbs Sampling with examples in R and WinBUGS, The Metropolis-Hastings Algorithm.
Bayesian Hypothesis Testing (One-sided and Two-sided Example), The Bayes Factor, A Test for Comparing Two Population Means. Hierarchical Bayes Examples.

References:

1. J.O. Berger : Statistical Decision Theory and Bayesian Analysis
2. J.K. Ghosh, M. Delampady & T. Samanta : Bayesian Inference
3. P. Lee : Bayesian Statistics –An Introduction
4. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari and D. B. Rubin : Bayesian Data Analysis (3rd Edition).
5. C. P. Robert : The Bayesian Choice (2nd Edition).

Applied Multivariate Analysis

Credit:3 (T) + 1 (P) **Paper Code: MSTR320C**
Marks: 75 (T) + 25 (P)

Clustering of Observations: Hierarchical clustering for continuous and categorical data - different choices of proximity measures, Agglomerative and Divisive algorithms, K-means clustering optimum choice of the number of clusters, Silhouette Index.
Classification and discrimination procedures: Discrimination between two known populations – Bayes, Minimax and Likelihood Ratio procedures. Discrimination between two multivariate normal populations. Sample discriminant function. Likelihood ratio rule. Tests associated with discriminant function, Probabilities of misclassification and their estimation. Classification of several populations. Fisher's method for discriminating among several populations.
Factor Analysis: The orthogonal factor model, Estimation of factor loading, Factor rotation, Estimation of Factor scores, Interpretation of Factor Analysis.
Multidimensional Scaling.
Correspondence Analysis.
Content Analysis: Steps in Content Analysis, Reliability of coded data, Interpretation of results.
Principal Component Analysis: Population and sample principal components and their uses. Biplot, Large sample inferences
Classification and Regression Tree. Random Forest.

References :

1. T.W. Anderson : An Introduction to Multivariate Statistical Analysis, (2^{nd} ed.)
2. N.C. Giri : Multivariate Statistical Inference
3. R.A. Johnson & D.W. Wichern : Applied Multivariate Statistical Analysis
4. A.M. Khirsagar : Multivariate Analysis
5. D.F. Morrison : Multivariate Statistical Methods
6. R.J. Muirhead : Aspects of Multivariate Statistical Theory
7. G.A.F. Seber : Multivariate Observations
8. S.C. Sharma : Applied Multivariate Techniques

Regression Analysis II

Credit: 3 (T) + 1 (P) **Paper Code: MSTR330C**
Marks: 75 (T) + 25 (P)

Measures of association for classified nominal and ordinal categorical data.
Generalized Linear Models: Introduction, Components of a GLM, Maximum Likelihood estimation, Deviance.
Binary data and Count data: ungrouped and grouped. Models with constant coefficient of variation. Polytomous data.
Overdispersion and fitting by quasi-likelihood.
Extensions of GLMs: Zero inflated Poisson models, Joint modeling of mean and variance, Concept of Generalized Linear Mixed Models (GLLM).

References :

1. A. Agresti : Analysis of Ordinal Categorical Data

2. A. Agresti : Categorical Data Analysis
3. P.McCullagh & A.J.Nelder : Generalized Linear Models
4. C.E.McCullough & S.R.Searle : Generalized, Linear and Mixed Models, 2nd ed.
5. T.Hastie&R.Tibshirani : Generalized Additive Models

Time Series Analysis and Development Statistics

Credit: 2 (T) + 1 (P) **Paper Code: MSTR340C**
Marks: 50 (T) + 25 (P)

Time Series Analysis: Classical Models, Smoothing Techniques – exponential and Holt-Winters methods. Evolutionary and Stationary time series. Autocorrelation and partial autocorrelation functions. Box-Jenkins Model. Tests for Unit Roots. Forecasting.
Volatility : ARCH, GARCH models – their variants.
Analysis in the frequency domain.
Development Statistics: Concept of economic development – role of statistics. National and international statistical systems. National accounts – estimation of national income.
Measurement of Poverty, Inequality and Unemployment.
Development indices – Human Development and Multi-dimensional Poverty Indices.

References :

1. C. Chatfield : The Analysis of Time Series – An Introduction
2. G.E.P. Box ,G.M. Jenkins & G.C.Reinsel : Time Series Analysis – Forecasting & Control
3. P.J. Brockwell & R.A. Davis : Introduction to Time Series Analysis and Forecasting
5. A.Pankratz :Forecasting with Univariate Box-Jenkins Model
6. G. Janacek and L. Swift :Time Series–Forecasting, Simulation, Applications
7. CSO (2007) : National Accounts Statistics – Sources and Methods
8. A.Sen : Poverty and Inequality
9. Y.P.Chaubey : Poverty Measurements : issues, approaches and indices
10. UNO : Yearly Human Development Reports

Advanced Data Analytic Techniques (Resampling Techniques/ Missing Data/Longitudinal Data)

Credit:3 (T) + 1 (P) **Paper Code: MSTR350C**
Marks: 75 (T) + 25 (P)

Resampling Techniques
Permutation tests
Introduction to Jackknife and Bootstrap-methods for estimating bias, standard error and distribution function based on iid random variables,
Standard examples
Bootstrap confidence intervals

Missing data analysis

Informative or non-informative missingness; MCAR, MAR and MNAR.
Complete case / Available case estimation
Mean imputation, Hot and cold deck imputation; MICE.
EM & MCEM algorithms and data augmentation techniques.

Longitudinal data analysis

Longitudinal regression : Cohort vs longitudinal effect, Bias and efficiency
Robust estimation -Weighted least-squares; Robust standard error estimation.
Parametric estimation: ML and REML.
Marginal, subject specific and transition models for continuous, binary and count outcomes.
Concept of GEE.

References :

1. J.J. Faraway : Linear Models with R
2. J.J. Faraway : Extending the Linear Model with R
3. D. Ruppert et al. : Semiparametric Regression
4. R.J.A. Little & D.B.Rubin : Statistical Analysis with Missing Data
5. C.K. Enders : Applied Missing Data Analysis
6. M.A. Tanner : Tools for Statistical Inference
7. G.J. McLachlan& T. Krishnan : The EM Algorithm and Extensions
8. B. Efron & R.J. Tibshirani : An introduction to bootstrap
9. B.Efron : The jackknife, the bootstrap, and other resampling plans
10. B. Efron : Bootstrap methods – another look at jackknife
11. J. Shao & D. Tu : The Jackknife and Bootstrap
12. P.J. Diggle et. al. : Analysis of Longitudinal Data (2nd ed).

Elective: 1 (anyone from the following three papers)

Demography

Credit: 3 (T) + 1 (P) **Paper Code: MSTR361C**
Marks: 75 (T) + 25 (P)

Use of balancing equation. Population composition-population pyramid.
Definition of concepts: Life and Death, Death Rates- age, space and cause specific. Adjusted death rates, Natality: Birth rates- age-sex adjusted, Quality Adjusted Life.
Migration-related measures.
Life tables: Distribution of life table functions and their estimates. Multiple Decrement tables.
Growth curve models, Population Estimation and Projection, Methods for Population projection. Stable and quasi stable population intrinsic growth rate.
Stochastic Models for Social and Occupational Mobility based on Markov Chains – closed and open systems, Estimation of Measures of Mobility. Manpower planning Models.

References :

1. D.J. Bartholomew : Stochastic Models for Social Processes (3rd ed.)
2. C.L. Chiang : Introduction to Stochastic Processes in Biostatistics

3. P.R. Cox : Demography
4. H.S. Shryock et.al. : The Methods and Materials of Demography
5. N.Keyfitz& N. Caswell : Applied Mathematical Demography

Optimization Techniques

Credit: 3 (T) + 1 (P) **Paper Code: MST R362C**
Marks: 75 (T) + 25 (P)

Search methods – different search methods for the single and several variable(s). Fibonacci, Golden Section, Steepest Descent methods.
Linear programming – Formulation of problems. Simplex algorithm. Artificial variables. Two-phase method. Duality. Dual simplex algorithm.
Integer programming – integer linear and mixed integer linear programming problems, Gomery's cutting plane method, Branch and Bound method. Balas algorithm for zero-one programming.
Goal programming – formulation of goal constraints, Partitioning algorithm.
Non-linear programming – multivariate optimization with inequality constraints. Kuhn-Tucker conditions. Convex programming, Quadratic Programming – Wolfe's algorithm.
Dynamic Programming.

References :

1. G. Hadley : Linear Programming
2. K.G. Murthy : Linear and Combinatorial Programming
3. S.N.S. Kambo : Mathematical Programming Techniques
4. S.S. Rao : Optimization – Theory and Applications
5. K.V. Mittal : Optimization Methods

Actuarial Statistics

Credit: 3 (T) + 1 (P) **Paper Code: MSTR363C**
Marks: 75 (T) + 25 (P)

Elements of decision theory and its actuarial applications.

Loss distributions: modelling of individual and aggregate losses, moments, fitting distributions to claims data, deductibles and retention limits, proportional and excess-of-loss reinsurance, share of claim amounts, parametric estimation with incomplete information.

Risk models: models for claim number and claim amount in short-term contracts, moments, compound distributions, moments of insurer's and reinsurer's share of aggregate claims.

Application of Bayesian techniques in credibility theory.

Experience rating: Rating methods in insurance and banking, claim probability calculation, stationary distribution of proportion of policyholders in various levels of discount.

Delay/run-off triangle: development factor, basic and inflation-adjusted chain-ladder method, alternative methods, average cost per claim and Bornhuetter-Ferguson methods for outstanding claim amounts, statistical models.

References :

1. N.L. Bowers, H.U. Gerber, J.C. Hickman, D.A. Jones & C.J. Nesbitt : Actuarial Mathematics, 2nd ed.
2. S.A. Klugman, H.H. Panjer, G.E. Willmot & G.G.Venter : Loss Models - From Data to Decisions.
3. C. D. Daykin, T. Pentikainen & M. Pesonen : Practical Risk Theory for Actuaries.

Internship

Credit: 2 **Paper Code: MSTR370J**
Marks: 75(T) + 25(P)

SEMESTER: IV

Statistical Analysis of Big Data

Credit: 3 (T) + 1 (P) **Paper Code: MSTR410C**
Marks: 75 (T) + 25 (P)

Machine learning: Supervised and unsupervised. The 5 v's of a big data problem. The curse of dimensionality. Prediction accuracy vs bias.
Measurement of model fit : Cross Validation and information theoretic criteria.
Generalization of Linear Regression – Shrinkage methods: Ridge Regression and LASSO.
Partial least squares.
Additive models.
Nonlinear models.
Classification using a separating hyperplane :The maximal margin classifier andseparability, Support Vector Machine, Data piling in high dimension, Case of multipleclasses.
Computer Architecture for Big Data

References:

1. T. Hastie, R. Tibshirani& J. Friedman : The Elements of Statistical Learning
2. B.L. Friedman, et al. : Classification and Regression Trees
3. R.A. Johnson & D.W. Wichern : Applied Multivariate Statistical Analysis

---

Elective Course (2)
Specialization Bucket: 01 (Industrial Statistics)

Operations Research

Credit: 3 (T) + 1 (P) **Paper Code: MSTR421C**
Marks: 75 (T) + 25 (P)

Definition and Scope of Operations Research.
Decision-making under uncertainty and risk, use of different criteria. Decision-making in the face of competition, two-person games, pure and mixed strategies, existence of solution and uniqueness of value in zero-sum games, finding solutions in 2x2, 2xm and mxn games.

Analytical structure of inventory problems, EOQ formula of Harris & Wilson, its sensitivity analysis and extensions allowing quantity discounts and shortages. Production inventory model. Models with random demand, the static risk model. P and Q- systems with constant and random lead times. Exact analysis of Q-system under Poisson demand. Airline problem. ABC analysis.
Queueing models – specification and effectiveness measures. Steady-state solutions of M/M/1, M/M/c M/M/c/N. M/G/1 queue and Pollazcek-Khinchine result. Machine interference problem, Little's formula.

Replacement problems – Deterministic models, Preventive replacement policies (cost and availability criteria), Staffing Problem

Traveling salesman problem - Branch and Bound method, Simulated annealing.

Project management and Network analysis - PERT and CPM.

References :

H.A. Taha	:	Operational Research
F.S. Hillier & G.J. Leiberman	:	Introduction to Operations Research
D.T. Philips, A. Ravindran & J. Solberg	:	Operations Research
C.W. Churchman, R.L. Ackoff & E.L. Arnoff	:	Introduction to Operations Research
T.M. Starr & D.W. Miller	:	Inventory Control – Theory & Practice
G. Hadley and TM Whitin	:	Analysis of Inventory Systems
L. Kleinrock	:	Queueing Systems
Sasieni, Yaspan & Friedman	:	Operations Research
Sasieni & Achoff	:	Operations Research

Reliability Theory

Credit: 3 (T) + 1 (P) **Paper Code: MSTR431C**
Marks: 75 (T) + 25 (P)

Reliability concepts and measures, components and systems, coherent systems, reliability of coherent systems. Life-distributions, reliability function, hazard rate, Mean residual life, common univariate life distributions, Mixture distribution, Convolution distribution, Bivariate exponential.
Notions of ageing – IFR, IFRA, NBU, DMRL and NBUE classes and their duals, preservation of such classes under reliability operations, Loss of memory property, Partial ordering of life distributions.

Censoring; Types of censoring. Reliability estimation based on failure times from variously censored life-tests data for parametric families.
Kaplan – Meier estimation of reliability curve, Greenwood formula, Non – parametric methods for comparison of several reliability curves, Log rank tests.

Regression models in reliability, Cox PH and Accelerated failure time models; Estimation of parameters and diagnostics.

Types of warranty; modelling; warranty cost analysis. Warranty prediction.

References:

R.E. Barlow & F. Proschan	:	Statistical Theory of Reliability and Life- Testing
J.F. Lawless	:	Statistical Models and Methods of Life-time data
L.J. Bain & M. Engelhardt	:	Statistical Analysis of Reliability and Life- testing Models
S. Zacks	:	Introduction to Reliability Analysis: Probability Models and Statistical Methods
J.D. Kalbfleisch &R.L. Prentice	:	The Statistical Analysis of Failure time data, 2nded.
P.J. Smith	:	Analysis of failure and survival data
C.D. Lai & M. Xie	:	Stochastic Ageing and Dependence for Reliability
I.B. Gertsbakh	:	Reliability Theory with Applications to Preventive Maintenance
W. Buschke and P.Murthy	:	Warranty cost analysis.

Statistical Quality Management

Credit: 3 (T) + 1 (P) **Paper Code: MSTR441C**
Marks: 75 (T) + 25 (P)

Group control chart. Extreme value chart. Moving average and exponentially weighted moving average charts. Cu-sum charts using V-masks and decision intervals. Economic design of control chart.

Multivariate Control Charts
Acceptance sampling plans for inspection by variables for two-sided specifications. Mil Std 105 plans. Continuous sampling plans of Dodge type and Wald-Wolfowitz type and their properties. Bayesian sampling plans
Process Capability
Taguchi method for quality improvement
QM System and ISO 9001 – brief exposition.

Basic concepts of 6 Sigma - DMAIC approach and the metrics used.

References :

D.C. Montogomery	:	Introduction to Statistical Quality Control
E.R. Ott	:	Process Quality Control
G.B. Wetherill	:	Sampling Inspection and Quality Control
G.B. Wetherill & D.W. Brown	:	Statistical Process Control – Theory and Practice
Thomas R. Ryan	:	Statistical Methods for Quality Improvement

Specialization Bucket: 02 (Business Analytics)

Econometrics

Credit: 3 (T) + 1 (P) **Paper Code: MSTR422C**
Marks: 75 (T) + 25 (P)

Single-equation linear model – censored data, errors-in-variables, lagged variables.

Simultaneous Equations models – identification & estimation. SUR models

Analysis of Panel Data

Nonparametric and semiparametric methods in econometrics

References :

J.Johnston	:	Econometric Methods
G.G.Judge, et.al.	:	The Theory and Practice of Econometrics (2^{nd} ed.)
W.Greene	:	Econometric Analysis
E.Malinvaud	:	Statistical Methods in Econometrics
A.Pagan&A.Ullah	:	Non-parametric Econometrics
B.M. Baltagi	:	Econometric Analysis of Panel Data

Financial Time Series

Credit: 3 (T) + 1 (P) **Paper Code: MSTR432C**
Marks: 75 (T) + 25 (P)

Granger Causality. Exogenity testing. Error Correction model.

Cointegration.

State Space Models and Kalman Filters.

Multivariate time series models : Crosscorrelation matrices. Portmanteau Test.

VAR and VARMA models. Cointegrated VAR models. Threshold Cointegration and Arbitrage. Multivariate Volatility models.

Nonlinear Models : Inference and applications.

Extreme Value, Quantile Estimation, Value at Risk.

References :

R.S. Tsay	:	Analysis of Financial Time Series
T.C. Mills and R.N. Merkellos	:	The Econometric Modelling of Financial Time Series
S.J. Taylor	:	Modelling Financial Time Series
T.G. Andersen, R.A. Davis, J.P. Kreib and T.V. Mikosch	:	Handbook of Financial time Series
G.S.Maddala & I.M. Kim	:	Unit Roots, Cointegration, and Structural Change

Machine Learning in Finance

Credit: 3 (T) + 1 (P) **Paper Code: MSTR442C**
Marks: 75 (T) + 25 (P)

Neural Networks: Perception, Feed-Forward networks, Back propagation, Regularization, Recurrent networks.
Binary and Multinomial Classifications.
Sequence Modeling.
Reinforcement Learning: Markov Decision Processes, Value and action value functions, Bellman optimality.
Inverse Reinforcement Learning and Imitation Learning
Case studies.

References :

M.F. Dixon, I. Halperin & P. Bilokon : Machine Learning in Finance – from Theory to Practice
T. Hastie, R. Tibshirani & J.H. Friedman : The elements of statistical learning – data mining, inference and prediction
C.E. Rasmussen & C.K.I Williams : Gaussian Processes for Machine Learning
R.S. Sutton & A.G. Barto : Introduction to Reinforcement Learning

Financial Econometrics

Credit: 3 (T) + 1 (P) **Paper Code: MSTR442C**
Marks: 75 (T) + 25 (P)

Risk-free and risky assets. Contracts and options. Continuous compounding, present valuation, risk, risk-neutral valuation.
Self-financing portfolios in finite markets. Replication, arbitrage, market completeness and existence. Hedging.
Harrison-Pliska arbitrage theorem. Capital assets pricing models.
Derivative pricing models. Option Pricing. Cox-Ross-Rubinstein Binomial and Black-Scholes models.

Portfolio management, Value-at-risk.

References :

A. N. Shiryaev : Essentials of stochastic finance: facts, models, theory
S. Ross : An elementary introduction to mathematical finance
J.Y.Campbell, A.W.Lo & C.MacKinlay : The Econometrics of Financial Markets
M.Baxter & A.Rennie : Financial Calculus: An Introduction to DerivativePricing
M. Ross : Intro. to Mathematical Finance: Options and OtherTopics
N.H.Bingham & R.Kiesel : Risk-Neutral Value Pricing & Hedging of Financial Derivatives
V.S.Bawa, S.J.Brown & R.W.Klein: Estimation Risk and Optimal Portfolio Choice

---

Elective Course (4)
Specialization Bucket: 03 (Biostatistics)

Survival Analysis

Credit: 3 (T) + 1 (P) **Paper Code: MSTR423C**
Marks: 75 (T) + 25 (P)

Introduction. Basic functions and Models,.Censoring and Truncation.
Parametric univariate estimation : Standard models – exponential, Weibull, log-logistic, log-normal and Gamma.

Nonparametric univariate estimation : Actuarial, Kaplan-Meier and Nelson-Aalen estimators.
Tests of equality of survival functions : Gehan's and Mantel-Haenszel tests.
Semiparametric regression models : Cox proportional hazard model – estimation, tests, Cox-Snell residuals.
Additive Models.
Accelerated Models – standard error distributions.
Competing Risk and Multivariate Survival models.
Frailty Models.

References :

J.P. Klein & M.L. Moeschberger	:	Survival Analysis : Techniques for Censored and Truncated Data
D.J. Kleinbaum & M. Klein	:	Survival Analysis – A Self-Learning Text
R.G. Miller	:	Survival Analysis
P.J. Smith	:	Analysis of Failure and Survival Data
J.D. Kalbfleisch & R.L. Prentiice	:	The Statistical Analysis of Failure Time Data, 2nd ed.

Clinical Trials

Credit: 3 (T) + 1 (P) **Paper Code: MSTR433C**
Marks: 75 (T) + 25 (P)

Introduction, Ethical issues in clinical trials, Types of clinical trials, Sample size determination, Group sequential monitoring.
Randomized clinical trials: Randomization for balancing treatment assignments (random allocation rule, truncated binomial design, biased coin designs), Incorporating covariate information.

Randomization to favor the better performing treatments for binary responses (play-the winner and randomized-play-the–winner rules).

References :

S. Piantadosi : Clinical Trials - A Methodologic Perspective

B.S. Everitt & A. Pickles : Statistical Aspects of Design & Analysis ofClinical Trials

S.J. Pocock : Clinical Trials

J. Whitehead : The Design and Analysis of Sequential Clinical Trials

W. F. Rosenberger & J.M. Lachin : Randomization in Clinical Trials- Theory and Practice

---

Statistical Genetics

Credit: 3 (T) + 1 (P) **Paper Code: MSTR443C**
Marks: 75 (T) + 25 (P)

Introduction to genetics

Gene mapping, sequence data, population genetics and coalescent theory

Phylogeny reconstruction

Pedigree analysis

Familial aggregation, segregation and linkage and association

Genetic epidemiology, role of genetic factors in human diseases

Analysis of complex and quantitative traits

References

B.H. Liu : Statistical Genomics Linkage, Mapping, and QTL Analysis

B. Neale, M. Ferreira, : Statistical Genetics: Gene Mapping Through Linkage

S. Medland & D. Posthuma (eds.) and Association

N.M. Laird & C. Lange : The Fundamentals of Modern Statistical Genetics

M. Lynch & B. Walsh : Genetics and Analysis of quantitative traits.

J. Felsenstein : Inferring Phylogenies

Z. Yang : Computational Molecular Evolution, Oxford University Press

---

Epidemiology

Credit: 3 (T) + 1 (P) **Paper Code: MSTR443C**
Marks: 75 (T) + 25 (P)

Definition of epidemiology. Case study on John Snow and the Lambeth cholera epidemic.

Study designs: Ecological, Cross-sectional, Cohort, Case-Control and its variants. Prospective and Retrospective studies.

Standard measures of Disease frequency and association based on rates and proportions.

Confounding and effect modification.

The design and analysis of cohort and case-control studies.

The design and analysis of matched studies.

Concept of causality and its measurement.

Case studies.

References :

K.J. Rothman, S. Greenland & T.L. Lash	:	Modern Epidemiology
S. Selvin	:	Statistical Analysis of Epidemiologic Data
D. McNeil	:	Epidemiological Research Methods
J.F. Jekel, J.G. Elmore & D.L. Katz	:	Epidemiology, Biostatistics and Preventive Medicine
N.E. Breslow & N.E. Day	:	Statistical Methods in cancer Research, Vol. 1, The Analysis of Case-Control Studies
N.E. Breslow & N.E. Day	:	Statistical Methods in cancer Research, Vol. 2, The Design and Analysis of Cohort Studies
S.J. Pocock	:	Clinical Trials
J. Whitehead	:	The Design and Analysis of Sequential Clinical Trials

---

Project

Credit: 6 Paper Code: MSTR450J

Marks:
