1) **Title of the course:** M. A. / M. Sc. in Statistics

2) **Pattern:** Semester and Credit system with pre-requisites for some courses

The M. A. / M. Sc. program in Statistics consists of 100 credits. Credits of a course are specified against the title of the course. A course with T in brackets indicates that it is a theory course whereas a course with P in brackets indicates that it is a practical course. Some of the practical courses are linked with a theory course and in such a case, both the courses will have the same number with T and P, indicating a theory and a practical course respectively. A student can enroll for a practical course if
(i) (s)he has enrolled for the corresponding theory course(as indicated) in the same term or
(ii) (s)he has passed the corresponding theory course in an earlier term or
(iii) terms for the corresponding theory course have been granted in an earlier term.

Some courses have pre-requisites as specified. Unless a student passes the pre-requisites of a course, (s)he will not be allowed to enroll in the said course. Pre-requisites of a course are indicated in curly brackets against the course. The Head of the Department, if necessary, may change the pre-requisites of a course on recommendations of the Departmental Committee.

3) **Eligibility**

For M. A. in Statistics following candidates are eligible.
(i) B.A. (Second class) with Statistics as principal and Mathematics at subsidiary level,
(ii) B.A. (Second class) with Mathematics as principal and Statistics at subsidiary level,
(iii) M.A. (First class) in Mathematics,
(iv) M.A. (First class) in Biometry.

For M. Sc. in Statistics following candidates are eligible.
(i) B.Sc. (Second class) with Statistics as principal and Mathematics at subsidiary level,
(ii) B.Sc. (Second class) with Mathematics as principal and Statistics at subsidiary level,
(iii) M.Sc. (First class) in Mathematics,
(iv) M.Sc. (First class) in Biometry.

4) **Examination**

A) (i) Pattern of examination: There would be continuous internal assessment (CIA) and an end of term examination (ETE) for each course. CIA includes examinations, assignments, viva-voce examinations and presentations.
(ii) Pattern of the question paper: For a compulsory course with 5 credits, a student is given an option to attempt 5 out of 8 questions.

B) Standard of passing: A student has to obtain 40% marks in the combined grading of ETE and CIA for passing the course, with a minimum passing of 30% in both CIA and ETE separately.

C) ATKT rules: A student can register for the third semester, if s/he completes 50% credits of the total credits expected to be completed within first two semesters, subject to the pre-requisite system mentioned above. Once registered, a student should complete M.A./M.Sc. within a period of 4 years (8 semesters).

D) Award of class: As per the University rules.

E) External students: Not applicable

F) Setting of question paper: The Department has been given autonomy for setting and grading of the papers. A teacher teaching a course sets the paper of the course and it is moderated by a committee of subject experts from other universities.

G) Verification or revaluation: As per the University rules
5) **Structure of the course**
   (i) a) Compulsory papers

   Compulsory courses in semester I and semester II are listed below.

   **Semester I**

   ST 1 Mathematical Analysis (5)
   ST 2 (T) Linear Algebra (3)
   ST 2 (P) Linear Algebra (2)
   ST 3 (T) Probability Distributions (5)
   ST 3 (P) Probability Distributions (1)
   ST 4 Multi-dimensional Calculus (3)
   ST 5 (T) Numerical Analysis and Programming (2)
   ST 5 (P) Numerical Analysis and programming (3)

   **Semester II**

   ST 6 Probability theory (5) {ST 1}
   ST 7 (T) Regression Analysis (4) {ST 2 (T), ST 3(T)}
   ST 7 (P) Regression Analysis (2)
   ST 8 Parametric Inference (5) {ST 3 (T)}
   ST 9 (T) Multivariate Analysis (4) {ST 3(T), ST 4, ST 2(T)}
   ST 9 (P) Multivariate Analysis (2)
   ST 10 Stochastic Processes I (4)

   In semesters III and IV some courses are compulsory and some are optional. List of optional
courses is given in 5(b).

   **Semester III**

   ST 11 (T) Asymptotic Inference (5) {ST 6 and ST 8}
   ST 11 (P) Asymptotic Inference (2)
   ST 12 (T) Design of Experiments and Analysis of Variance (4) {ST 7(T)}
   ST 12 (P) Design of Experiments and Analysis of Variance (2)
   ST 13 Stochastic Processes II (2)
   Optional course I (5)
   Optional course II (5)

   **Semester IV**

   ST 14 (T) Sampling Methods (4)
   ST 14 (P) Sampling Methods (2)
   ST 15 (P) Project (4)
   Optional course I (5)
   Optional course II (5)
   Optional course III (5)

   The compulsory courses in all the semesters are the core courses. These have 75 (55 (theory) + 20
   (practical) ) credits in all.

   A student may be allotted a module by the Department at the beginning of the second year of the
program. The allotment is based on the performance of students as indicated by her/his rank among all the
students given admission in the same year. A total of 15 credits should be obtained from the courses listed
as module specific courses. The remaining 10 credits need to be obtained from the optional courses listed in
5(b) or from courses from other departments, as allowed by the University.
The Head of the Department may not allow a student to continue with the allotted module and such a student can complete M.A. /M.Sc. in Statistics without having offered any module, subject to the requirement of at least 80 credits obtained from the courses offered by the Department, as described below.

A student can opt for a project in lieu of an optional course or two optional courses with the permission of the Head of the Department. A copy of rules and regulations regarding completion and submission of the project work by a student and assessment of the project work is available in the Department.

Below is given a list of 4 modules. Courses specific to a given module are also listed. Each of these courses, called as a module-specific course, has 5 credits.

A project can have either 5 credits or 10 credits. A project with 5 credits has to be completed within a semester. A project with 10 credits will be spread over two semesters; work in each semester is to be counted for 5 credits. A student can opt for two different projects also, each having 5 credits.

Modules

M1. Probability and Mathematical Statistics
M2. Industrial and Financial Statistics
M3. Bio-Statistics
M4. Computational Statistics

List of module specific courses

Module 1: Probability & Mathematical Statistics.

ST P1 Measure theory and Probability {ST 6}
ST P2 Advanced Stochastic Processes {ST 6, ST 10}
ST P3 Decision Theory and Bayesian Inference {ST 8}
ST P4 Advanced Multivariate Analysis {ST 9 (T)}
ST P5 Inference in Stochastic Processes {ST 10, ST 11, ST 13}

Module 2: Industrial and Financial Statistics

ST I1 Optimization Techniques {ST 2(T)}
ST I2 Statistical Methods for Quality and Reliability {ST 3 (T)}
ST I3 Survival Analysis {ST 11 (T)}
ST I4 Time Series {ST 7 (T)}
ST I5 Actuarial Statistics {ST 3 (T)}
ST I6 Stochastic models in Finance {ST 10 (T), ST 13}

Module 3: Bio-Statistics

ST B1 Statistical Ecology {ST 3 (T)}
ST B2 Statistical Genetics {ST 3 (T)}
ST B3 Survival Analysis {ST 11 (T)}
ST B4 Analysis of Clinical Trials {ST 7 (T), ST 12 (T)}
ST B5 Actuarial Statistics {ST 3 (T)}
ST B6 Medical and Health Statistics {ST 3 (T)}
Module 4: Computational Statistics

ST C1 Statistical Simulations {ST 3 (T), ST 5(T), ST 5(P)}
ST C2 Computer-Intensive Statistical Methods {ST 5 (T), ST 11 (T)}
ST C3 Advanced Computer Programming {ST 5 (T), ST 5 (P)}
ST C4 Knowledge Discovery and Data Mining {ST 6 (T), ST 9 (T)}
ST C5 Statistical Pattern Recognition {ST 6 (T), ST 9 (T)}

The Head of the Department may decide not to offer modules if sufficient faculty is not available.

b) Optional papers: A student may choose optional courses from the following list or from the courses which are listed as compulsory for a module other than the module allotted to her/him or from courses offered by other Departments subject to the approval of the Head of the Department. A student may opt for a project (or two projects) in lieu of an optional (or two optional courses), as described earlier.

List of Proposed Optional Courses

ST E01 Stochastic Models {ST 3 (T), ST 10 (T), ST 13}
ST E02 Queues and Inventories {ST 10 (T), ST 13}
ST E03 Empirical Processes {ST 10 (T), ST 6}
ST E04 Testing of hypotheses {ST 8}
ST E05 Sequential Analysis {ST 8}
ST E06 Nonparametric Inference {ST 8}
ST E07 Discrete Data Analysis {ST 11 (T)}
ST E08 Demography {ST 3 (T), ST 8}
ST E09 Quantitative Epidemiology {ST 3 (T), ST 10 (T), ST 13}
ST E10 Advanced Time Series {ST 14}
ST E11 Statistics in Micro-array data analysis {ST 9}
ST E12 Data Mining {ST 7, ST 9}
ST E13 Statistics in Natural Language Processing {ST 3(T)}
ST E14 Statistical methods for Bio-computing {ST 5, ST 10}
ST E15 Directional Data Analysis

The Head of the Department may introduce additional optional courses on recommendations of the Departmental Committee. The syllabus of the optional courses will be prepared by the concerned teacher and will be flexible to accommodate new developments in that area. Whenever such an optional course is floated, the concerned syllabus will be discussed and approved in the Departmental committee.

c) Question papers: In view of academic autonomy given to the Department, question papers are set by the teacher who teaches the course and these are moderated by a committee of experts, as indicated earlier.

ii) Medium of instructions: English

6) Equivalence subject/papers & transitory provision: The Head will take the decisions as and when the case arises.

7) University Terms: Time table for the beginning and end of the terms as announced by the University will be followed.

8 & 9) Subject wise detail syllabus and recommended books: Detailed syllabi, along with the list of recommended books of the compulsory courses and module specific courses are given below.
ST 1: Mathematical Analysis

Countability, supremum and infimum of sets of real numbers. Limit point of a set – open sets, closed sets etc. (will be developed through general metric space and \( \mathbb{R}^n \) will be considered as a special case), compactness, Bolzano-Weierstrass theorem, Heine-Borel Theorem. Continuous functions, uniform continuity, absolute continuity. Sequences and series of real numbers, limit superior, limit inferior and limit of a sequence. Cauchy sequences, convergence of series, tests for convergence of series, absolute convergence, Cauchy products.

Riemann and Riemann–Stieltjes integrals, integration by parts, mean value theorem. Uniform convergence of sequences and series. Term by term differentiation and integration, applications to power series. Improper Riemann–Stieltjes integrals: Improper integrals of first and second kind for one variable. Uniform convergence of improper integrals, differentiation under the sign of integral – Leibniz rule.

Books Recommended


ST 2(T): Linear Algebra

Matrix algebra, special types of matrices, determinants and their simple properties. Orthogonal and idempotent matrices: Linear dependence, basis of a vector space, an orthogonal basis, Gram-Schmidt orthogonalization, projection theorem, linear transformation, rank of matrix, linear equations, solution space and null space. Characteristic roots of real matrices, right and left characteristic vectors, independence of characteristic vectors corresponding to distinct characteristic roots, algebraic and geometric multiplicities; Generalized inverse, Spectral decomposition theorem.

Definiteness of a real quadratic form, reduction of quadratic forms, simultaneous reduction of two quadratic forms, maxima and minima of ratios of two quadratic forms.

Books Recommended


Additional Books for Reference


ST 2(P)

The following practicals are to be done by using the software R/SYSTAT/Matlab.

1. Calculation of determinant (4 x 4)
2. Calculation of eigen values, eigen vectors, g- inverse
3. Solution of simultaneous equations
4. Quadratic forms

ST 3(T): Probability Distributions

Random Experiment and its sample space, random variables, c.d.f., p.d.f., p.m.f., absolutely continuous and discrete distributions, mixtures of probability distributions. Some common distributions. Transformations, moments, m.g.f., p.g.f., quantiles and symmetry. Random vectors, joint distributions, joint m.g.f. mixed
moments, variance covariance matrix. Hazard rate and cumulative hazard rate, lack of memory property. IFR, IFRA, DFR and DFRA classes of distribution.

Independence, sums of independent random variables, convolutions, conditional expectation and variances, regression function and best linear regression function, multiple and partial correlation coefficients.

Sampling distributions of statistics from univariate normal random samples, such as linear and quadratic forms. Fisher Cochran theorem. Non-central chi-square, t and F distributions.

Order statistics: Distribution of r-th order statistic, joint distribution of several order statistics and their functions.

Probability Integral Transformation, Rank orders and their exact null distributions. One and two sample examples of rank statistic such as sign statistic, Wilcoxon signed rank statistic, Wilcoxon two sample statistic etc. Recurrence relations for the null distribution of the Wilcoxon two sample statistic.

Books Recommended

ST 3(P)

Using software packages: SYSTAT, R and MINITAB
1. Plotting of density functions, distribution functions and failure rates.
2. Wilcoxon test, Wilcoxon signed-rank test, Kruskal Wallis test., rank correlation
3. Model sampling from standard distributions and mixtures of distributions.

ST 4: Multi-dimensional Calculus

Review of calculus of one variable: differentiability, mean value theorem and Taylor series expansion. Functions of several variables: Continuity, directional derivatives, differential of functions of one variable, differentials of functions of several variables, the gradient vector, differentials of composite functions and the chain rule, the mean value theorem, a sufficient condition for the existence of the differential, partial derivatives of higher order and Taylor’s formula.

Multiple Integrals and evaluation of multiple integrals by repeated integration. Mean value theorem for multiple integrals.

Applications of partial differentiation: Jacobians, the inverse function theorem, the implicit function theorem, extremum problems.

Books Recommended

**ST 5(T): Numerical Analysis and Programming**

- Mixed congruential
- Multiplicative congruential
- Rejection
- Distribution specific methods

Testing for randomness of a sequence. Runs test, digit frequency test, gap test, serial correlation

Methods to compute integrals: quadrature formula, Monte Carlo Methods. Applications of Monte Carlo methods to compute expected values of random variables.

**Books Recommended:**

**ST 5(P)**

Elementary Statistics as listed below using software packages: SYSTAT, R and MINITAB.
- Calculation of summary statistics (mean, median, maximum, minimum, s.d.)
- Calculation of regression and correlation coefficients.
- ANOVA for one-way and two-way models
- Analysis of 2x2 contingency table.
- Computation of integrals by Riemann & RS sums.
- Calculation of p-value for standard normal distribution (for given Z value)
- Preparing frequency distribution of given data.
- Calculation of double integrals.
- Obtaining SRSWOR from a finite population,
- Calculation of order statistic
- Plotting of curves (standard & non-standard)
- Limits of functions
- Box Plot
- Computing integrals by statistical methods
- Computing expectations of complicated functions, mean and variance of estimates
- Calculation of empirical power & level of significance.
- Empirical confidence coefficient of a confidence interval
- Test for Randomness

**Books Recommended:**
ST 6: Probability Theory

Algebra of sets, fields and sigma-fields, limit of sequences of subsets, sigma-field generated by a class of subsets, Borel fields. Probability measure on a sigma-field, probability space, continuity of a probability measure, real and vector-valued random variables (r.v.s), distribution functions (d.f.), discrete r.v.s, r.v.s of the continuous type, decomposition of a d.f.

Expectation of a real r.v. and of a complex-valued r.v. Linear properties of expectations. Characteristic functions, their simple properties, uniqueness theorem.

Convergence of a sequence of r.v.s., convergence in distribution, convergence in probability, almost sure convergence and convergence in quadratic mean, their inter-relations. Cramer's theorem on composition of convergence in distribution and convergence in probability. Slutkey's theorem. Monotone convergence theorem and dominated convergence theorem.

Independence of two events and n (> 2) events, sequence of independent events, independent classes of events, \( \pi \)-system and \( \lambda \)-system of events, Dynkin’s theorem, independence of r.v.s, Borel zero-one law.

Khintchin's weak law of large numbers, Kolmogorov strong law of large numbers (without proof), continuity theorem for characteristic functions. Lindeberg's CLT and its particular cases,

Books Recommended

Additional Books for Reference

ST 7(T): Regression Analysis

Simple regression with one independent variable(X), assumptions, estimation of parameters, standard error of estimator, testing of hypothesis about regression parameters, standard error of prediction. Testing of hypotheses about parallelism, equality of intercepts, congruence. Extrapolation, optimal choice of X. Diagnostic checks and correction: graphical techniques, tests for normality, uncorrelatedness, homoscedasticity, lack of fit, modifications like polynomial regression, transformations on Y or X, WLS. Inverse regression X(Y).


Non Linear regression (NLS) : Linearization transforms, their use & limitations, examination of non linearity, initial estimates, iterative procedures for NLS, grid search, Newton- Raphson , steepest descent, Marquardt’s methods.

Generalized Linear model: link functions such as Poisson, binomial, inverse binomial, inverse Gaussian, gamma.

**Books Recommended:**


**ST 7(P)**

Simple regression, regression diagnostics. (2 practicals )
Multiple regression, forward method, backward method (2 practicals)
Non-linear regression – some non-standard models (2 practicals)
Logistic regression (2 practicals)
Following practicals are to be done with the help of a software package available in the Department.
1. Best subset selection, (2) GLM, (3) Linearization techniques in non-linear models.

**ST 8: Parametric Inference**

Sufficiency, completeness, Uniformly minimum variance unbiased estimators, C-R inequalities, exponential class of densities and its properties, some special classes of distributions admitting complete sufficient statistics, extensions of these results to multi-parameter situation.

Test function, Neyman- Pearson lemma for test functions. Uniformly most powerful tests for one sided alternative for one parameter exponential class of densities and extension to the distributions having monotone likelihood ratio property.

Confidence Intervals, shortest expected length confidence intervals, relations with testing of hypotheses, uniformly most accurate confidence intervals.

Bayesian estimation, prior distributions, posterior distribution, loss function, principle of minimum expected posterior loss, quadratic and other common loss functions, conjugate prior distributions. Common examples. Bayesian HPD confidence intervals.

**Books Recommended:**

ST 9(T): Multivariate Analysis

Multivariate normal distribution, pdf and mgf, singular and nonsingular normal distributions, distribution of a linear form and a quadratic form of normal variables, marginal and conditional distributions. Multiple regression and multiple and partial correlation coefficients, Definition and relationships.


Introduction to Principle Components , Canonical correlation coefficients, and canonical variables.

Cluster Analysis.

Classification problem. Discriminant analysis, Mahalanobis $D^2$-statistic.Methods and applications of MANOVA (without derivation of the distribution of Wilks' lambda).

Books Recommended:


ST 9 (P)

1. Model Sampling from multivariate normal distribution
2. Applications of Hotelling’s $T^2$ ( 2 practicals)
3. MANOVA
4. Discriminant Analysis
5. Principal Component Analysis
6. Cannonical Correlations
7. Cluster Analysis

ST 10: Stochastic Processes I

Markov chains with stationary transition probabilities, properties of transition functions, classification of states, Stationary distribution of a Markov chain, existence and uniqueness, convergence to the stationary distribution. Methods based on Markov chains for simulation of random vectors. MCMC algorithm.


Branching processes.

Introduction to Wiener Process and Brownian Motion.

Books Recommended


### Additional Books for Reference


### ST 11(T): Asymptotic Inference

Consistency and asymptotic normality (CAN) of real and vector parameters. Invariance of consistency under continuous transformation. Invariance of CAN estimators under differentiable transformations, generation of CAN estimators using central limit theorem. Method of moments, method of maximum likelihood. Special cases such as exponential class of densities and multinomial distribution, Cramer-Huzurbazar theorem, method of scoring.

Tests based on MLEs. Likelihood ratio tests, asymptotic distribution of log likelihood ratio, Wald Test, Score Test, locally most powerful tests. Applications to categorical data analysis, three dimensional contingency tables, Pearson’s chi-square test and LR test. Asymptotic comparison of tests. Asymptotic Relative Efficiency (Pitman’s), asymptotic normality of posterior distributions.

### Books Recommended:


### ST 11 (P)

1. Estimation of parameters of mixture distributions.
2. Functional estimation using kernels, bandwidth selection
3. Verification of consistency and asymptotic normality of estimates
4. Maximum likelihood estimation under various set up (includes censoring situations)
5. Comparing methods of estimation, MSE and sample size considerations
6. Power functions and comparison of tests (LR, WALD, RAO)
7. Analysis of higher dimensional contingency tables.

### ST – 12(T): Design of experiments and Analysis of Variance

Randomization, replication, local control, one way and two way classification with unequal and equal number of observations per cell (with / without interactions). Connectedness, balance, orthogonality, BIBD, ANOCOVA.

$2^k$ full factorial experiments: diagrammatic presentation of main effects and first order interactions, model, analysis of single as well as more than one replicates, using ANOVA.

Total confounding of $2^k$ design in $2^p$ blocks, $p \geq 2$. Partial confounding in $2^p$ blocks, $p =2, 3$. Fractional factorial experiments. Resolution of a design, (III, IV & V), aberration of a design.

Concept of rotatable design. Central composite designs, $3^2$ designs: contrasts for linear and quadratic effects, statistical analysis of $3^2$ design.
Response surface methodology (RSM): linear and quadratic model, stationary point, ridge systems, multiple responses, blocking in RSM.
Taguchi (orthogonal array) methods: concept of loss function, S/N ratio, linear graphs, ANOM, inner and outer arrays. ANOVA.

Random effect models for one way classification.

**Books Recommended:**

**ST 12(P): Design of experiments and Analysis of Variance**
1. One way classification. Multiple comparison tests.
2. Two way classification with equal / unequal number of observations per cell (model with interaction).
3. LSD and BIBD.
4. Analysis of covariance in one way and two way data.
5. $2^n$ Factorial Experiments, Analysis of Single Replicate of $2^n$.
6. Total and partial confounding for $n = 2,3,4,5$.
7. Fractional Replication.
8. $3^2$ Factorial Experiments.

**ST 13: Stochastic Processes II**
Markov pure jump processes, Poisson process, Birth and Death processes. Finite state continuous time Markov chains.
Renewal processes, Poisson process as a renewal process, elementary renewal theorem. Statement (without proof) of other renewal theorems.
Simple queueing systems.

**Books Recommended**

**ST 14(T): Sampling Methods**
Basic methods of sample selection, simple random sampling with replacement (SRSWR), simple random sampling without replacement (SRSWOR), probability proportional sampling with and without replacement, systematic sampling, estimation problems, Horwitz- Thompson estimator and its properties.
Stratification: Allocation problems and estimation problems, formation of strata and number of strata, method of collapsed strata.
Use of supplementary information for estimation, ratio and regression estimators with their properties and generalizations, Jackknife methods.
Cluster sampling, multistage-sampling. Double sampling procedures, Ratio and regression estimators, stratification.
Non-sampling errors, response and non-response errors and their treatments, randomized response.

**Books Recommended**


**Additional Books for Reference**


**ST 14(P): Sampling Methods**

SRSWOR, SRSWR, Stratified random sampling, various kinds of allocation, Post stratification, using auxiliary information.

Ratio and regression methods of estimation, pps sampling design.

Double sampling, two stage sampling, Systematic sampling, cluster sampling Randomized response technique.

Two practicals of consolidated nature each one of which would use theory of one or more of the above topics.

**ST 15(P) Practicals (Projects)**

The course consists of following two components - (i) Summary of research articles and (ii) Data Analysis

**Summary of Research Articles**

Students are expected to read some articles (number will be decided by the supervisor) on a specified topic or theme, summarize and write a comprehensive report and present the summary of the articles.

**Data Analysis**

Students are expected to choose her/his own project, wherein they are expected to analyze data pertaining to certain theme using a variety of statistical tools that they have studied so far.

This completes the detailed syllabi of 15 compulsory courses. Syllabi of module specific courses follow.

**Module 1: Probability and Mathematical Statistics**

**ST P1: Measure Theory and Probability**

Ring, σ-ring, Measure, Measure space, Caratheodory Extension theorem, Lebesgue measure.

Integral of a measurable function with respect to a measure, its properties. Hahn – Jordan decomposition, Lebesgue decomposition, Radon – Nikodym derivative.

Product measure, Fubini’s theorem. Convergence in measure, almost everywhere convergence, Kolmogorov Inequality. Kolmogorov three series criterion strong law of large numbers, conditional
Probability and conditional expectations, their simple properties. Martingales martingale convergence theorems (SLLN, CLT)

Books Recommended

5. Williams, P. ( ) Probability and Martingales

Additional Books for Reference


ST P2: Advanced Stochastic Processes

Markov Sequences: Definitions, Transition Densities, Stationary Distribution, Normal Markov Sequences [2] Ch. VI Section 11, ch. III Section 8 Markov Pure jump processes [4], [1]. Stochastic Processes (General Theory) : Probability spaces appropriate for stochastic processes, Kolmogorov’s extension theorem. (only sketch of the proof), separability, progressive measurability and strong Markov property of stochastic processes [3], Ch. 4 Sections 1 and 2.

Diffusion Processes : Definition, Elementary properties, infinitesimal parameters, standard process and Dynkin’s theorem, Continuity and non-differentiability of Diffusion processes. Modeling based on Diffusion processes, Standard Brownian Motion, Ornstein-Ulhenbeck process and other processes, Transformation of processes, Distribution of Hitting times and related problems, scale function and speed density, Kolmogorov’s backward differential equations, forward differential equations (without proof), transition density and stationary distribution of a diffusion process.

Books Recommended

5. Athreya & Lahiri ( )

ST P3: Decision Theory and Bayesian Analysis


Subjective probability and selection of prior distribution for Bayesian analysis. Bayesian analysis for statistical inference problems of estimation, testing hypotheses, confidence interval and prediction. Bayesian decision theory. Admissible and minimax decision rules. Complete class of decision rules.

Books Recommended

1. James O. Berger (1985), Statistical Decision Theory and Bayesian Analysis ,
   (Second Edn.) Springer -Verlag
   ( Specific material to be covered is in the following sections of the book 1, 2.1, 2.2, 2.4.1, 2.4.2, 3.1-3.3, 4.1-4.4.3, 4.8, 5.1-5.3, 8.1,8.2.)
Additional Books for Reference


ST P4 : Advanced Multivariate Analysis

Canonical variables, canonical correlation coefficients, and canonical vectors. Distributions of the sample canonical correlations coefficients and canonical vectors.

Wilk’s $\Lambda$ Criterion. The distribution of Wilk's $\Lambda$. Alternatives to Wilk's $\Lambda$ criterion.

Multivariate analysis of variance. Discrimination in case of several groups.

Application of Wilk's $\Lambda$ and canonical analysis in contingency tables. Likelihood ratio tests concerning mean vectors and covariance matrices of several populations, their properties. Principal components distributions associated with sample principal components (Null case). James Stein estimator of population mean vector.

Factor Analysis.

Books Recommended


ST P5 Inference in Stochastic Processes

Inference in Markov chains, estimation of transition probabilities, testing for order of a Markov chain, estimation of functions of transition probabilities, parametric models and their goodness of fit. Markov sequences, estimation of parameters based on likelihood and conditional least squares, autoregressive series. Models for higher order Markov chains. (Raftery)

Statement of martingale strong law of large numbers and CLT for martingales, CAN property of the mle from a general sequence of dependent random variables, Fisher information. Applications to Markov chains and sequences.

Likelihood of Poisson and other Pure Jump Markov processes from first principles, CAN property of mles , testing for a Poisson process, non-homogeneous processes, Analysis of parametric Pure Jump processes, Birth-Death-Immigration processes, testing goodness of fit of such models

Diffusion processes and their likelihood, properties of estimators (without proof)


Methods based on estimating functions, panel data, introduction to spatial models.

Books Recommended:

Module 2: Industrial and Financial Statistics

ST-I1 (Optimization Techniques)
Linear Programming: Convex sets, Supporting and Separating Hyper-planes, Standard linear Programming Problem, basic feasible solution, simplex algorithm and simplex method, graphical solution, two phase method. Duality in linear programming, duality theorems, dual simplex method with justification, sensitivity.

Transportation and assignment algorithms, Hungarian method of assignment, transshipment problems, duality theory of testing optimality of solution in transportation problem and transshipment problems, transportation problem and transshipment problems as network problems Balance and degeneracy in transportation problem.

Maximization, prohibitions and other variations of assignment problems. Integer linear Programming Problem, branch and bound method, Network flows, maximal flow in the network. Nonlinear Programming: Kuhn-Tucker conditions, Quadratic programming, Wolfe’s, Beale’s and Fletcher’s algorithms for solving quadratic programming problems.

Markovian and Non-Markovian queueing models, cost profit models of (M/M/1) and (M/M/s) queueing systems. Simulation, event type simulation, simulation of a queuing systems. Dynamic programming

Books Recommended:
4. Panneerselvam, R. Operations Research (Prentice hall of India)

ST I2: Statistical Methods for Quality


C_{p}, C_{pk}. Estimation & confidence intervals for estimators of C_{p}. Connection between proportion of defectives & C_{p}.

**Acceptance Sampling plans:** Single, double & multiple sampling plans for attributes. Curtailed double sampling plans. Operating characteristic functions & other properties of the sampling plan. Use of sampling plans for rectification. Dodge-Romig acceptance sampling plans.

**Acceptance sampling plan for variables.** Designing variable acceptance sampling plans. AQL based sampling plans. Continuous sampling plans CSP-I & CSP – II.

**Books Recommended:**

### ST I3 : Survival Analysis

Concepts of Time, Order and Random Censoring.

Life distributions - Exponential Gamma, Weibull, Lognormal, Pareto, Linear Failure rate. Parametric inference. Point estimation, Confidence Intervals, Scores, tests based on LR, MLE

Life tables, Failure rate, mean residual life and their elementary properties. Ageing classes - IFR, IFRA, NBU, NBUE, HNBUG and their duals, Bathtub Failure rate.

Estimation of survival function - Actuarial Estimator, Kaplan - Meier Estimator, Estimation under the assumption of IFR/DFR.

Tests of exponentiality against non-parametric classes - Total time on test, Deshpande test. Two sample problem - Gehan Test, Log rank test. Mantel - Haenszel Test, Tarone - Ware tests.

Semi-parametric regression for failure rate - Cox's proportional hazards model with one and several covariates.

**Books Recommended:**
5. Zacks, S. *Reliability*

### ST I4: Time Series Analysis

Time-series as discrete parameter stochastic process. Auto covariance and autocorrelation functions and their properties.

Exploratory time Series Analysis, Tests for trend and seasonality. Exponential and Moving average smoothing. Hot -Wiinters smoothing. Forecasting based on smoothing, adaptive smoothing.

Detailed study of the stationary processes: (1) moving average (MA), (2) Auto regressive (AR), (3) ARMA and (4) AR integrated MA (ARIMA) models. Box-Jenkins models. Discussion (without proof) of estimation of mean, auto covariance and autocorrelation functions under large sample theory. Choice of AR

Introduction to spectral analysis of weakly stationary process. Periodogram and correlogram analyses.

Nonstationary and Seasonal Time series Models: Unit-root nonstationarity, Unit-root tests, Integrated ARMA (ARIMA) models, Seasonal ARIMA (SARIMA) models
Conditional Heteroskedastic Models: Volatility models, ARCH and GARCH, Properties, Examples, Estimation & Forecasting,
Multivariate Time series model, VAR models, Vector ARMA models, Cointegration models.
Use of statistical software for time series analysis

Books Recommended:

Additional Books for Reference:

ST I5: Actuarial Statistics

Utility theory, insurance and utility theory, models for individual claims and their sums, survival function, curtate future lifetime, force of mortality.

Life table and its relation with survival function, examples, assumptions for fractional ages, some analytical laws of mortality, select and ultimate tables.

Principles of compound interest: Nominal and effective rates of interest and discount, force of interest and discount, compound interest, accumulation factor, continuous compounding.

Life insurance: Insurance payable at the moment’s of death and at the end of the year of death-level benefit insurance, endowment insurance, deferred insurance and varying benefit insurance.

Life annuities: Continuous life annuities, discrete life annuities, life annuities with monthly payments.

Net premiums: Continuous and discrete premiums, true monthly payment premiums.
Some practical considerations: Premiums that include expenses-general expenses types of expenses, per policy expenses.
Net premium reserves: Continuous and discrete net premium reserve, reserves on a semi continuous basis, reserves based on true monthly premiums

Multiple life functions, joint life and last survivor status, insurance and annuity benefits through multiple life functions evaluation for special mortality laws.

Distribution of aggregate claims, compound Poisson distribution and its applications. Distribution of aggregate claims, compound Poisson distribution and its applications.

**Books Recommended:**

Section I – Chapters: 1, 2, 3, 8, 9, and 11  
Section II – Chapters: 4, 5, 6, 7, 13, and 14  

**Books for Additional References:**


**ST I 6: Stochastic Models in Finance**


Options markets, properties of stock option prices. American and European options.  

Behaviour of stock prices : Conditional expectation, Martingale, Brownian Motion and  
Geometric Brownian motion, Markov property, Ito integral, Ito/diffusion and Mean-reverting processes  
Process, Ito Lemma.


**Books Recommended:**


**Module 3: Bio-statistics**

**ST B1: Statistical Ecology**

Population density estimation: Capture-recapture models, nearest neighbor models, line transect sampling.
Ecological Diversity: Simpson's index, Shannon–Weaver index, Diversity as average rarity.

Optimal Harvesting of Natural Resources, Maximum sustainable yield, tragedy of the commons.
Game theory in ecology: Concept of Evolutionarily stable strategy, its properties, simple cases such as Hawk-Dove game.
Foraging Theory: Diet choice problem, patch choice problem, mean variance tradeoff.

**Books Recommended:**


**ST B2: Statistical Genetics**

Gene frequency, random mating, Hardy-Weinberg Equilibrium, Matrix theory of random mating with applications.

Inbreeding, coefficients of inbreeding, inbreeding in randomly mating populations of finite size. Statistical problems in human genetics, blood group analysis.

Natural selection.
Quantitative genetics, study of inheritance of quantitative characters in random and nonrandom mating diploid populations. Detection and estimation of linkage.

Sequence similarity, homology and alignment. Algorithm for (a) pair wise sequence alignment, (b) multiple sequence alignment, construction of phylogenetic trees, UPGMA. Neighbor joining, maximum parsimony and maximum likelihood algorithms.

**Books Recommended:**

2. Ewens, W.J. (1979) Mathematical Population Genetics (Springer Verlag)

**Additional Books for Reference**


**ST B3: Survival Analysis**

Syllabus of this course is the same as that of ST I3.
**ST B4: Analysis of Clinical Trials**

Introduction to clinical trials: the need and ethics of clinical trials, bias and random error in clinical studies, conduct of clinical trials, overview of Phase I-IV trials, multi-center trials.

Data management: data definitions, case report forms, database design, data collection systems for good clinical practice.

Design of clinical trials: parallel vs. cross-over designs, cross-sectional vs. longitudinal designs, review of factorial designs, objectives and endpoints of clinical trials, design of Phase I trials, design of single-stage and multi-stage Phase II trials, design and monitoring of Phase III trials with sequential stopping, design of bio-equivalence trials.

Reporting and analysis: analysis of categorical outcomes from phase I – III trials, analysis of survival data from clinical trials.

Surrogate endpoints: selection and design of trials with surrogate endpoints, analysis of surrogate endpoint data.

Meta-analysis of clinical trials.

**Books Recommended:**


**ST B5: Actuarial Statistics**

This syllabus is the same as ST I 5.

**ST B6: Medical and Health Statistics**

Study designs in epidemiology. Measures of disease occurrence and association, variation and bias. Identifying non-causal association and confounding.

Defining and assessing heterogeneity of effects, interaction. Sensitivity and specificity of diagnostic test, Cohort Study designs, statistical power and sample size computations.

Log-linear models, 2xK and 2x2x2 contingency tables. Logistic model. Analysis of binary data.

Cross-control study designs, matched case-control studies.

Survival data: Proportional hazards model, multivariate survival data. Causal Inference, Longitudinal data.

Communicating results of epidemiological studies, ethical issues in epidemiology.

**Books Recommended:**

1. Selvin : Statistical analysis of epidemiological data.
2. Diggle, Liang and Zeger : Analysis of longitudinal data
3. Piantadosi : Clinical trials
4. Agresti : Categorical Data Analysis.
5. Clayton and Hills : Statistical methods in Epidemiology
8. Zhou, Obuchowski and McClish : Statistical Methods in Diagnostic Medicine
Module 4: Computational Statistics

ST C1: Statistical Simulations


Variance reduction techniques: importance sampling for integration, control variates and antithetic variables.

Simulating a non-homogeneous Poisson process.
Optimization using Monte Carlo methods, simulated annealing for optimization.
Solving differential equations by Monte Carlo methods.


ST C2: Computer Intensive Statistical Methods

Jackknife and Bootstrap.
Bootstrap methods: re-sampling paradigms, bias and standard errors. Bootstrapping for estimation of sampling distribution, confidence intervals, variance stabilizing transformation, bootstrapping in regression and sampling from finite populations.

Jackknife and cross-validation: jackknife in sample surveys, jack-knifing in regression with heteroscedasticity, cross-validation for tuning parameters.

EM algorithm: applications to missing and incomplete data problems, mixture models.
Applications to Bayesian analysis. Monte Carlo EM algorithm MCMC methods in missing data.

Smoothing with kernels: density estimation, simple nonparametric regression. Failure rate.
Permutation tests

Books Recommended:
ST C3: Advanced Computer Programming

This course is intended to introduce object-oriented computer programming. It assumes prior exposure to programming in languages such as C and/or Fortran. The language of choice here is C++.

Introduction to object-oriented programming concepts and design.
Programming in C++: data types and operations, functions and parameters, classes, constructors, input/output, control statements such as if-else, switch, for, while and do-while, pointers and references, dynamic allocation, processing of linked lists, arrays and character strings, libraries.

Introduction to program analysis: simple testing and debugging
Introduction to Web programming: simple examples in Java and the concept of byte code.

Note: If the previous exposure to C++ is observed then java can be substituted.

Books Recommended:

ST C4: Knowledge Discovery and Data Mining

Review of classification methods from multivariate analysis, classification and decision trees. Clustering methods from both statistical and data mining viewpoints, vector quantization.

Unsupervised learning from univariate and multivariate data, dimension reduction and feature selection.

Supervised learning from moderate to high dimensional input spaces, artificial neural networks and extensions of regression models, regression trees.

Introduction to databases, including simple relational databases, data warehouses and introduction to online analytical data processing.

Association rules and prediction, data attributes, applications to electronic commerce.

Books Recommended:

ST C5: Statistical Pattern Recognition

Linear classifiers: linear discriminant function (LDF) for minimum squared error, LDF for binary outputs, perception learning algorithm.

Nearest neighbor decision rules: description, convergence, finite sample considerations, use of branch and bound methods.
Probability of errors: two classes, normal distributions, equal covariance matrix assumptions, Chernoff bounds and Bhattacharya distance, estimation of probability of error.

Feature selection and extraction: interclass distance measures, discriminant analysis, probabilistic distance measures, principal components.

Books Recommended: