COURSE CONTENTS

for

M.Sc. in Applied Statistics and Informatics

(for 2016-17 Batch onwards)

Central University of Orissa

[A Central University Established by an Act of Parliament of India in 2009]

Central University of Orissa
Koraput (764021)
Odisha

(1) Course Name: M. Sc. in Applied Statistics and Informatics

(2) Preamble to the syllabus

M. Sc. in Applied Statistics and Informatics program is of minimum 80 credits spread over four semesters. This program emphasizes both theory and applications of statistics and is structured to provide knowledge and skills in depth necessary for the employability of students in industry, other organizations, as well as in academics. Accordingly, the program has important features such as projects, elective courses and courses on standard software packages such as MATLAB, SPSS, R. For every course, there shall be Continuous Internal Examination (CIE) conducted by department and End of the Semester Examination (ESE) shall be conducted by the department at the end of semester.

(3) Introduction

(a) M. Sc. in Applied Statistics and Informatics program will be conducted under credit system in four semesters. There shall be 20 credits in each semester for a total of 80 credits.

(b) Some courses are termed Open Courses (O). The open courses are those offered by other departments but relevant to M. Sc. in Applied Statistics and Informatics program.

(c) In addition, there are practical courses and Project course.

(d) For every course, there will be Continuous Internal Examination (CIE) conducted by department and End of the Semester Examination (ESE) will be conducted by the department at the end of semester.

(4) Structure of the program

Following is a structure of the two year M.Sc. in Applied Statistics and Informatics program.

T: Theory P: Practical O: Open E: Elective.

The courses will be offered with following provisions:

I. Elective courses to the students of Department of Statistics will be offered subject to the availability of the faculty members in the department.
II. All courses both core and elective shall be offered to the students of other departments subject to the availability of faculty members in the Department.

Course Structure

Semester	Course Code	Title of Paper	Core /Elective	L-T-P-D[C]
1st	STAT 101	Mathematical Analysis	Core	3-1-0-0[4]
	STAT 102	Probability Theory-I	Core	3-1-0-0[4]
	STAT 103	Statistical Methods	Core	3-1-0-0[4]
	STAT 104	C-Programming	Core	2-1-0-0[3]
	STAT 105	Linear Algebra	Core	2-1-0-0[3]
	STAT 106	Practical-I (Using C)	Core	0-0-2-0[2]
2nd	STAT 201	Statistical Inference-I	Core	3-1-0-0[4]
	STAT 202	Probability Theory-II	Core	3-1-0-0[4]
		Open Elective-I *	Elective	3-1-0-0[4]
	STAT 203	Sampling Theory	Core	2-1-0-0[3]
	STAT 204	Operations Research	Core	2-1-0-0[3]
	STAT 205	Practical-II (Using Excel / SPSS)	Core	0-0-2-0[2]
3rd	STAT 301	Statistical Inference-II	Core	3-1-0-0[4]
	STAT 302	Stochastic Processes	Core	3-1-0-0[4]
	STAT 303	Design and Analysis of Experiments	Core	3-1-0-0[4]
	STAT 304	Linear Models	Core	2-1-0-0[3]
		Department Elective-I (from Module-II)	Elective	2-1-0-0[3]
	STAT 305	Practical-III (Using R / Matlab)	Core	0-0-2-0[2]
4th	STAT 401	Applied Regression Analysis	Elective	2-1-0-0[3]
	STAT 402	Multivariate Analysis	Core	3-1-0-0[4]
		Department Elective-II (from Module-III)	Core	2-1-0-0[3]
		Open Elective-II**	Core	3-1-0-0[4]
	STAT 404	Project	Core	0-0-0-6[6]

* A Student of Department of Statistics is to choose open elective-I from Module-I/ Any course from other Department in CUO.

** A Student of Department of Statistics is to choose open elective-I from Module-IV /Any course from other Department in CUO.

LIST OF MODULE SPECIFIC ELECTIVE COURSES

MODULE I: APPLIED, INDUSTRIAL AND FINANCIAL STATISTICS

Sl. No.	Course Code	Title of Paper	L-T-P-D[C]
1	STAT 206	Demography	3-1-0-0[4]
2	STAT 207	Time Series Analysis	3-1-0-0[4]
3	STAT 208	Statistical Quality Assurance	3-1-0-0[4]
4	STAT 209	Applied Statistics for Nation Development	3-1-0-0[4]
5	STAT 210	Directional data analysis	3-1-0-0[4]
6	STAT 211	Statistical Finance	3-1-0-0[4]

MODULE II: PROBABILITY & MATHEMATICAL STATISTICS

Sl. No.	Course Code	Title of Paper	L-T-P-D[C]
1	STAT 212	Advanced Distribution Theory	2-1-0-0[3]
2	STAT 213	Advanced Probability Theory	2-1-0-0[3]
3	STAT 214	Statistical Decision Theory	2-1-0-0[3]
4	STAT 215	Reliability Theory	2-1-0-0[3]
5	STAT 216	Order Statistics	2-1-0-0[3]
6	STAT 217	Bayesian Inference	2-1-0-0[3]

MODULE III: BIO-STATISTICS

Sl. No.	Course Code	Title of Paper	L-T-P-D[C]
1	STAT 306	Statistical Genetics	2-1-0-0[3]
2	STAT 307	Survival Analysis	2-1-0-0[3]
3	STAT 308	Statistical Ecology	2-1-0-0[3]
4	STAT 309	Medical and Health Statistics	2-1-0-0[3]
5	STAT 310	Actuarial Statistics	2-1-0-0[3]
6	STAT 311	Statistics for Clinical Trials	2-1-0-0[3]

MODULE IV: COMPUTATIONAL STATISTICS

Sl. No.	Course Code	Title of Paper	L-T-P-D[C]
1	STAT 406	Computer intensive Statistical Methods	3-1-0-0[4]
2	STAT 407	Statistical Simulations	3-1-0-0[4]
3	STAT 408	Knowledge Discovery and Data Mining	3-1-0-0[4]
4	STAT 409	Statistical Pattern Recognizition	3-1-0-0[4]
5	STAT 410	Large Scale Data Analysis	3-1-0-0[4]
6	STAT 411	Advance data analytic techniques	3-1-0-0[4]

DETAILED SYLLABUS OF THE M.SC. IN APPLIED STATISTICS AND INFORMATICS PROGRAM (FOR 2016-17 BATCHES ONWARDS)

STAT 101	MATHEMATICAL ANALYSIS	4 Credits

Unit 1

Sequences, subsequences, convergence, divergence, bounded sequences, limits superior and inferior, monotone sequences, Cauchy sequences, completeness, Series of real numbers.

Unit 2

Bounded and unbounded subsets of the line, intervals, closed and open sets, characterizations, limit points, closures, interiors. Denseness. Compact sets. Heine- Borel Theorem, Bolzano-Weierstrass Theorem

Unit 3

Functions: limits, continuity, uniform continuity, intermediate value theorem, Differentiability; Chain rule, mean value theorem, Taylor's theorem (statement), extreme, Multivariate calculus: partial, directional and total derivatives mean value theorem, Gamma function and Beta function, Multiple integrals, change of variables, Jacobian formula.

Unit 4

Algebra of complex numbers, operations of absolute value and conjugate, standard inequalities for absolute value, concept of analytic functions via power series and differentiability methods, Exponential and logarithmic functions, trigonometric functions of a complex variable, complex line integral, Cauchy integral formula, Holomorphic functions, Laurent Series, Singularity, calculus of residues, evaluation of integration using contour integration.

References:

1. Bartle G.R. & Sherbert D. R. (2000): Introduction to Real Analysis- John Wiley & Son Inc.
2. Royden (1988): Principles of Real Analysis - Macmillian.
3. Widder (1989): Advanced Calculus - Dover Publication.
4. W, Rudin (2013): Real and Complex Analysis, Tata Mc-Graw Hill.
5. E. M. Strein, R, Shakarchi (2003): Complex Analysis, Princeton University Press.
6. W, Rudin. (2013): Real and Complex Analysis, Tata Mc-Graw Hill.

STAT 102	PROBABILITY THEORY -I	4 Credits

Unit 1

General probability space, various definitions of probability. Combinations of events: additive and multiplicative laws of probability. Conditional probability. Bayes' theorem and its applications. Concept of random variables, cumulative distribution function and probability density function, joint, marginal and conditional distribution. Brief review of joint, marginal and conditional probability density function, functions of random variables and their distributions using jacobian of transformation.

Unit 2

Mathematical expectation, moments, conditional expectation, moment generating functions, cumulative generating functions and their applications, Characteristic function, uniqueness theorem, Levy's continuity theorem (statement only). Probability inequalities and their applications: Chebyshev, Markov and-Johnson.

Unit 3

Convergence in probability and convergence in distribution, weak law of large numbers, Sequence of events and random variables: Zero one law of Borel and Kolmogorov, almost sure convergence in mean squares, Khintchin's weak law of large numbers, Kolmogorov inequality, and strong law of large numbers.

Unit 4

Central limit theorem for a sequence of independent random variables under Lindeberg's condition, central limit theorem for independent and identically distributed random variables with finite variance.

References:

1. P.G. Hoel, S.C. Port and C.J. Stone, Introduction to Probability, Universal Book Stall, New Delhi, 1998.
2. Sheldon Ross, A first course in Probability, 8th Edition, Prentice Hall, 2009.
3. Loeve (1996): Probability Theory Affiliated East –West Press Pvt. Ltd. New Delhi.
4. Bhatt, B.R. (2000): Probability, New Age International India.
5. Feller, W. (1971): Introduction to Probability Theory and its Applications, Vol. I and II. Wiley, Eastern-Ltd.
6. Rohatgi, V.K (1984): An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern.
7. Billingsley, P. (1986): Probability and Measure, John Wiley Publications.

STAT 103	**STATISTICAL METHODS**	**4 Credits**

Unit 1

Descriptive Statistics: Measures of central tendency, dispersion, skewness and kurtosis for the study of nature of data. Idea of correlation and regression for two and three variables; correlation coefficient, correlation ratio, multiple and partial correlations.

Unit 2

Some discrete statistical distributions: Binomial, Poisson, hypergeometric, negative binomial and multinomial distributions. Some continuous distributions (Normal, Uniform, Exponential, Cauchy, Pareto, Weibull, lognormal), Bivariate normal and bivariate exponential distributions and their properties.

Unit 3

Functions of random variables and their distributions using Jacobian and other tools, convolution and compound distributions, truncated and mixture distributions, Sampling distributions from normal population central and non-central Chi-square, t and F distributions.

Unit 4

Order statistics and their distributions and properties. Joint and marginal distributions of order statistics. Extreme values and their asymptotic distribution (statement only) with applications, Asymptotic distribution of median, distribution of quantiles.

References:

1. Hogg, R.V. and Craig, A.L. (1978). Introduction to Mathematical Statistics, MacMillan, New York.
2. Mood, A.M., Graybiil, F.A. and Boes, D.C. (2001). Introduction to Theory of Statistics, Tata McGraw Hill, New Delhi.
3. Ross, Sheldon M (2004) Introduction to Probability and Statistics for Engineers and Scientist, Third Edition, Elsevier Academic Press, USA.
4. Rohatgi V.K. and Saleh, A.K.Md. E. (2001). An Introduction to Probability and Statistics (Second Edition), John Wiley and Sons (Asia), Singapore.
5. Dudewicz, E.J. and Mishra, S.N. (1988). Modern Mathematical Statistics, John Wiley, New York.
6. Johnson, S and Kotz, S. (1970). Continuous univariate Distributions I and II John Wiley, New York.
7. David, W.S. (2003). Order Statistics. (Second Edition). John Wiley and Sons, New York.

STAT 104	**C-PROGRAMMING**	**3 Credits**

Unit 1

C language: Programmer's model of a computer, Algorithms, Flow Charts; Data Types, Arithmetic and input/output instructions.

Unit 2

Decisions control structures; Decision statements, Logical and Conditional operators; Loop, Case control structures, Functions.

Unit 3

Preprocesssors, Arrays, Puppettting of strings.

Unit 4

Structures, Pointers, File formatting.

References:

1. Henry Mullish & Hobert Looper, Spirit of C: An Introduction to Modern Programming, Jaico Publishers, Bombay.
2. Kernighan B.W. and Ritchie D.M., C Programming Language, Prentice Hall, Software Series.
3. Gottfried, B.S. (1996). Programming with C, Schaum's Series, Tata McGraw Hill.
4. E. Balaguruswamy. Programming in ANSI C, Tata McGraw Hill.
5. Kanitkar, Y., Let Us C, BPB.

STAT 105	LINEAR ALGEBRA	3 Credits

Unit 1

Vector spaces over fields of scalars, subspaces, linear independence of vectors, basis and dimension of a vector space, completion theorem, orthogonality of vectors and subspaces, Vector spaces with an inner product.

Unit 2

Gram-Schmidt ortghogonalization process, orthonormal basis, linear transformations and projections and their representation by matrices.

Unit 3

Non-singular matrices and their inversion, determinants, ranks, row and column rank of a matrix, Idempotent matrix, its properties, trace, invariance theorems, Sylvester and Frobenious inequalities, elementary matrices, partitioned matrices, G - inverse, Kronecker product.

Unit 4

Systems of homogeneous and non-homogeneous linear equations, their consistency and maximal linearly independent solutions, minimal and characteristic polynomials of a square matrix, Characteristic roots and vectors, Cayley – Hamilton theorem, similarity and diagonalization of square matrices. Real quadratic forms and their value classes, canonical reductions and simultaneous reducibility of quadratic forms.

References:

1. Hadley, G. (1987). Linear Algebra, Narosa Publishing House, New Delhi.
2. Lay, David C. (1997). Linear Algebra and its Applications, Addison-Wesley,
3. Searle, S.R. (1982). Matrix Algebra useful for Statistics, John Wiley & Sons, New York.
4. Hoffman, K. and Kunze, R. (1971). Linear Algebra, 2nd edition. Prentice Hall, New Jersey.
5. Rao, A.R. and Bhimasankaram, P. (1992). Linear Algebra, Tata McGraw-Hill, New Delhi
6. Rao, C.R. (1995). Linear Statistical Inference and its Applications, Wiley Eastern, New Delhi.
7. Strang, G. (2005). Linear Algebra and its application.

STAT 106	PRACTICAL-I (USING C)	2 Credits

Algorithms and programs to analyze statistical data and solve routine statistical problems. Measures of location and dispersion, sorting of data, solving systems of equations constructing inverse matrices and g-inverses. Numerical Algorithms using C.

STAT 201	STATISTICAL INFERENCE-I	4 Credits

Unit 1

Point estimation: Sufficiency and minimal sufficiency, Exponential family of distributions, Pitman family, Factorization criterion, Likelihood equivalence, Unbiased estimation, Completeness, Ancillary statistics and Basu's Theorem, UMVUE estimators and their characterizations, Rao-Blackwell Theorem, Lehmann-Scheffe Theorem, UMVUE estimation of parametric functions from standard distributions.

Unit 2

Fisher information measure and its properties, Fisher information matrix, Lower bound to the variance of an unbiased estimates, Cramer-Rao, Chapman-Robbin's and Bhattacharya bounds, BLUE of parametric functions, Efficiency, Consistency, Weak and strong consistency, Marginal and joint consistent estimators, Equivariance, Pitman estimators.

Unit 3

Methods of estimation: Methods of moments, Maximum likelihood, Minimum chi square and its modification, Least square estimation, Properties of maximum likelihood estimators, Cramer-Huzurbazar Theorem, Likelihood equation - multiple roots, Iterative methods, E.M Algorithm.

Unit 4

Basic elements of Bayesian Inference, Loss function, Prior distribution, Bayes Theorem, Posterior distributions, Bayes risk, Bayes principle, Bayes estimators, Minimax estimators, Metropolis-Hastings algorithm, Gibbs sampler, MCMC method.

References:

1. E.L.Lehmann (1998) Theory of Point Estimation, John Wiley and Sons.
2. V.K.Rohatgi and A.K.L. Saleh (2001) An Introduction to Probability and Mathematical Statistics, Wiley.
3. B.K. Kale (1999) A First Course in Parametric Inference, Narosa Publishing Company.
4. Robert C.P. and Casella, G (1999) Monte Carlo Statistical Methods, Springer Verlag.
5. Mukhopadhay, P. (1999) Mathematical Statistics, New Central Book Agency Pvt. Ltd.
6. Hogg, R.V. and Craig, A.T. (1971): Introduction to Mathematical Statistics, Princeton University Press.

STAT 202	**PROBABILITY THEORY-II**	**4 Credits**

Unit 1

Fields, sigma-fields and generators, semifields, Borel sigma-field on R. Monotone classes, monotone class theorem, pi-lambda theorem. Measures, finite, sigma-finite measures. Probability measures, properties. Independence of events, Borel-Cantelli lemmas. Measurable functions and properties, Generated sigma-fields. Induced measures. Compositions. Examples.

Unit 2

Product sigma-fields, Borel sigma-field on Euclidean spaces. Extensions of measures, Caratheodory's theorem (statement). Lebesgue measure on R and R_k: construction, properties. Random variables and vectors, probability distributions, distribution functions. Convergence in measure, almost everywhere and their connection.

Unit 3

Integration: simple, nonnegative, general measurable functions, integrability, Monotone Convergence Theorem, Dominated Convergence Theorem, Fatou's lemma. Change of variables. Lp spaces, Holder's and Minkowski's inequalities. Expectations, moments. Jensen's inequality. Generating functions.

Unit 4

Absolute continuity and singularity of measures. Radon-Nikodym Theorem (Statement). Discrete and absolutely continuous distributions. Lebesgue's differentiation theorem (statement), probability densities.

References:

1. Bhat, B.R. (1985) Modern Probability Theory, Second edition, Wiley Eastern.
2. Billingsley, P. (1986) Probability and Measure, Second Edition, John Wiley.
3. Feller, W. (1966) An Introduction to Probability Theory and Its Applications, Volume II, Wiley Eastern.
4. Rao, C.R. (1973) Linear Statistical Inference and Its Applications, Wiley.
5. Rohatgi, V.K. and A.K.E. Salah (2001) Introduction to Probability and Statistics, John Wiley and Sons.
6. Basu, A.K. (1999) Measure Theory and Probability, Prentice-Hall.

	OPEN ELECTIVE-I	**4 Credits**

To be chosen from Module-I / Any course from other Department in CUO.

STAT 203	**SAMPLING THEORY**	**3 Credits**

Unit 1

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

Unit 2

Basic sampling schemes - Simple random sampling with and without replacement, 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.

Unit 3

Stratified sampling – Allocation problem and construction of strata. Ratio, Product, Difference and Regression estimators. Unbiased Ratio estimators – Probability proportional to aggregate size sampling, Hartley – Ross estimator in simple random sampling.

Unit 4

Sampling and sub-sampling of clusters. Two-stage sampling with equal/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. Double sampling for stratification. Double sampling ratio and regression estimators. Sampling on successive occasions.

References:

1. Cochran, W.G. (1999) Sampling Techniques, Third edition, John Wiley & Sons.

2. Sampath, S. (2001) Sampling theory and methods, Alpha Science International Ltd., India.
3. Mukhopadhay, P (2009) Theory and methods of survey sampling, Second edition, PHI Learning Pvt Ltd., New Delhi
4. Des Raj (1976) Sampling Theory, McGraw Hill.
5. Murthy, M.N. (1977) Sampling Theory and Methods, Statistical Publishing Society, Calcutta.
6. Singh, D. and Chaudhary, F.S. (1986) Theory and Analysis of Sample Survey Designs, Wiley Eastern.

STAT 204	OPERATIONS RESEARCH	3 Credits

Unit 1

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.

Unit 2

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.

Unit 3

Two persons sum game, Network flows, maximal flow in the network.

Unit 4

Integer linear Programming Problem, branch and bound method, Job sequencing problem.

References:

1. Kambo, N.S. (1991) Mathematical Programming Techniques (Affiliated East-west press Pvt. Ltd.)
2. Hadley, G. (1987) Linear Programming.
3. Taha, H.A. (1992) Operations Research 5th ed. (Macmillan)
4. Panneerselvam, R. Operations Research (Prentice hall of India)
5. Medhi J. (1984) Stochastic Processes 2nd ed.(New Age International Pvt. Ltd.)

STAT 205	PRACTICAL-II (USING EXCEL/SPSS)	2 Credits

Practical based on STAT-101, STAT-103, STAT-201, STAT-203, STAT-301, STAT-303, STAT-304, STAT-401, STAT-402.

STAT 301	STATISTICAL INFERENCE-II	4 Credits

Unit 1

Tests of hypotheses, Formulation of problem, Null and alternative hypotheses, Size of a test, Two kinds of errors, Simple and composite hypotheses, Randomized and non-randomized tests, Power of a test, M.P test, Neyman-Pearson lemma and its generalization, Monotone likelihood ratio property, UMP tests, Unbiased tests and UMPU tests, Unbiased critical regions and similar regions, Neyman structure, UMPU tests in multi-parametric exponential families of distributions.

Unit 2

Confidence interval estimation, Relationship between confidence interval estimation and testing of hypothesis, UMA and UMAU confidence intervals, Shortest confidence intervals, Construction of confidence intervals using pivots, Large sample confidence interval based on maximum likelihood estimator, central limit theorem and Chebyshev's inequality, Bayesian credible regions.

Unit 3

Likelihood ratio tests and their properties, Testing mean and variance of a normal population, Testing equality of means and variances of two normal populations, Sequential probability ratio tests, Construction of sequential probability ratio tests, Wald's identity, OC and ASN functions, Properties of SPRT.

Unit 4

Non-parametric inference: Goodness of fit tests- Chi square test and Kolmogorov Smirnov test for one and two sample problems, Sign test, Signed rank test, Wald-Wolfowitz run test, Median test, Man-Whitney U-test, Non-parametric confidence intervals, Bootstrapping confidence intervals, P-P Plot and Q-Q plot, Kendall's tau, Tests for independence and homogeneity.

References:

1. Casella G. & Beregar R.L. (2002) Statistical Inference, 2nd Edition (Duxbury Advanced Series)
2. Dudewitz E.J. & Mishra S.N.(1988) Modern Mathematical Statistics (John Wiley)
3. Kale B.K. (1999) A First course on Parametric Inference (Narosa)
4. Lehman E.L (1988) Theory of point estimation (John Wiley)
5. Lehman E.L (1986) Testing of Statistical hypotheses (John Wiley)
6. Rohatagi V.K. (1976) Introduction to theory of probability & mathematical statistics (John Wiley & sons)
7. Wald, A. (1947) Sequential Analysis, Doves
8. Gibbons, J.K. (1971) Non-Parametric Statistical Inference, McGraw Hill

STAT 302	STOCHASTIC PROCESSES	4 Credits

Unit 1

Markov Chains: Definition, Examples and classification, Discrete renewal equation and basic limit theorem, Absorption probabilities, Criteria for recurrence.

Unit 2

Continuous time Markov chains, Examples, General pure birth process, Poisson process, Birth and death process, Finite state continuous time Markov chains.

Unit 3

Galton-Watson branching processes, Generating function, Extinction probabilities, Continuous time branching processes, Extinction probabilities, Branching processes with general variable life time.

Unit 4

Renewal equation, Renewal theorem, Applications, Generalizations and variations of renewal processes, Applications of renewal theory, Brownian motion.

References:

1. Karlin, S. and Taylor, H.M. (1975) A first Course in Stochastic Processes, second edition, Academic Press.
2. Bhat, B.R. (2002) Stochastic Processes, second edition, New Age Publication.
3. Cox, D.R. (1962) Renewal Theory, Methuen.
4. Ross, S. (1996) Stochastic Processes, Second edition, John Wiley.
5. Medhi, J. (1994) Stochastic Processes, Second edition, Wiley Eastern.
6. Basu, A.K. (2002) Elements of Stochastic Processes, Narosa Publications.

STAT 303	**DESIGN AND ANALYSIS OF EXPERIMENTS**	**4 Credits**

Unit 1

Randomization, Replication and local control, One way and two way classifications with equal and unequal number of observations per cell with and without interaction, Fixed effects and Random effects model. Model adequacy checking, CRD, RBD and Latin Square designs, Analysis of co-variance for completely randomized and randomized block designs. Analysis of experiments with missing observations.

Unit 2

Incomplete Block Designs: Balanced Incomplete Block designs, Construction of BIB Designs, Analysis with recovery of inter-block information and intra-block information. Partially balanced incomplete block designs, Analysis of partially balanced incomplete block designs with two associate classes, Lattice designs.

Unit 3

2^n Factorial experiments. Analysis of 2^n factorial experiments. Total confounding of 2^n designs in 2^p blocks. Partial confounding in 2^p blocks. Fractional factorial designs, Resolution of a design, 3^n factorial designs. Split plot and strip plot designs.

Unit 4

Response surface designs - orthogonality, rotatability blocking and analysis - Method of Steepest accent, Crossover designs, Models properties and Analysis.

References:

1. Chakrabarti, M.C. (1962). Mathematics of Design and Analysis of Experiments, Asia Publishing House, Bombay.

2. Das, M.N. and Giri, N.C. (1986). Design and Analysis of Experiments, Wiley Eastern Limited.
3. Dean, A. and Voss, D. (1999). Design and Analysis of Experiments, Springer. First Indian Reprint 2006.
4. Dey, A. (1986). Theory of Block Designs, John Wiley & Sons.
5. Hinkelmann, K. and Kempthorne, O. (2005). Design and Analysis of Experiments, Vol. 2: Advanced Experimental Design, John Wiley & Sons.
6. John, P.W.M. (1971). Statistical Design and Analysis of Experiments, Macmillan Co., New York.
7. Montgomery, D.C. (2005). Design and Analysis of Experiments, Sixth Edition, John Wiley & Sons.

STAT 304	LINEAR MODELS	3 Credits

Unit 1

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

Unit 2

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

Unit 3

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

Unit 4

Regression analysis, Analysis of covariance.

References:

1. Scheffe, H. (1959). The Analysis of Variance, John Wiley.
2. Searle, S. R. (1971). Linear Models, John Wiley Publication.
3. Seber, G.A.F. (2003). Linear Regression Analysis, Wiley Publication.
4. Giri, N.C. (1986). Analysis of Variance, South Asian Publisher.
5. Joshi, D. D. (1987). Linear Estimation & Design of Experiments. New Age International PVT. Ltd.

	DEPARTMENT ELECTIVE-I	3 Credits

To be chosen from Module-II.

STAT 305	PRACTICAL-III (USING R/MATLAB)	2 Credits

Practical based on STAT-101, STAT-103, STAT-201, STAT-203, STAT-301, STAT-303, STAT-304, STAT-401, STAT=402;

STAT 401	APPLIED REGRESSION ANALYSIS	3 Credits

Unit 1

Linear Regression Model, Least squares estimation, Gauss Markov Theorem, Properties of the estimates, Distribution Theory, Maximum likelihood estimation, Estimation with linear

restrictions, Generalised least squares; Hypothesis testing - likelihood ratio test, F-test; Confidence intervals.

Unit 2

Residual analysis, Departures from underlying assumptions, Effect of outliers, Collinearity, Non-constant variance and serial correlation, Departures from normality, Diagnostics and remedies.

Unit 3

Polynomial regression in one and several variables, orthogonal polynomials, Indicator variables, Subset selection of explanatory variables, stepwise regression and Mallows Cp-statistics, Introduction to non-parametric regression.

Unit 4

Introduction to nonlinear regression, Least squares in the nonlinear case and estimation of parameters, Generalized linear models, Models for binary response variables, estimation and diagnosis methods for logistic and Poisson regressions. Prediction and residual analysis.

References:

1. Seber, A.F. and Lee, A.J. (2003) Linear Regression Analysis, John Wiley
2. Montegomery, D.C., Peck, E.A. and Vining, G.G. (2001) Introduction to Regression Analysis, Third edition. Wiley.
3. Searle, S.R. (1971) Linear models, John Wiley & Sons, Inc.
4. N.Draper and H.Smith (1986) Applied Regression Analysis – John Wiley & Sons.
5. B.Abraham and Ledotter, J. (1983) Statistical Methods for Forecasting, John Wiley & Sons.
6. Fox, J. (1984) Linear Statistical Models and Related methods, John Wiley.
7. Christensen, R. (2001) Advanced Linear Modelling.

STAT 402	**MULTIVARIATE ANALYSIS**	**4 Credits**

Unit 1

Notion of multivariate distributions, Multivariate normal distribution, Marginal and conditional distributions, Characteristic function, Estimation of mean vector and covariance matrix.

Unit 2

Distribution of rectangular co-ordinates, Wishart distribution and its properties, Distribution of simple, partial and multiple correlations based on samples from normal population, Hotelling's T^2 and Mahalanobis D^2 statistics, Properties of T^2 and D^2, Multivariate Fisher Behren's problem.

Unit 3

Testing independence of sets of variates, Testing equality of covariance matrices and means, Sphericity tests, testing the hypothesis that a covariance matrix equal to given matrix, Mean and covariance equal to a given vector and given matrix.

Unit 4

Classification problem, Principal component analysis, Canonical variables, Canonical correlation, Basics of factor analysis.

References:

1. Anderson, T.W. (1984) An Introduction to Multivariate Statistical Analysis, John Wiley.
2. Kshirasagar, A.M. (1972) Multivariate Analysis, Marcel-Dekker.
3. Seber, G.A.F. (1977) Multivariate Observations, Wiley.
4. Morrison, D.F. (1976) Multivariate Statistical Methods, John Wiley.
5. Rancher, A.C. (1995) Methods of Multivariate Analysis, John Wiley.
6. Johnson, R.A. and Wichern, D.W. (1990) Applied Multivariate Statistical Analysis, Prentice Hall.

	DEPARTMENT ELECTIVE-II	**3 Credits**

To be chosen from Module-III.

	OPEN ELECTIVE-II	**4 Credits**

To be chosen from Module-IV / Any course from other Department in CUO.

STAT 404	**PROJECT WORK**	**6 Credits**

LIST OF MODULE SPECIFIC ELECTIVE COURSES

MODULE I: APPLIED, INDUSTRIAL AND FINANCIAL STATISTICS

STAT 206	DEMOGRAPHY	4 Credits

Unit 1

Sources of demographic Statistics, Basic demographic measures: Ratios, Proportions and percentages, Population Pyramids, Sex ratio Crude rates, Specific rates, Labour force participation rates, Density of population, Probability of dying.

Unit 2

Life tables: Construction of a life table, Graphs of lx, qx, dx, Functions Lx, Tx and Ex. Abridged life tables Mortality: Rates and Ratios, Infant mortality, Maternal mortality, Expected number of deaths, Direct and Indirect Standardization, Compound analysis, Morbidity.

Unit 3

Fertility: Measures of Fertility, Reproductivity formulae, Rates of natural increase, Fertility Schedules, Differential fertility, Stable populations, Calculation of the age distribution of a stable population, Model Stable Populations.

Unit 4

Population estimates, Population Projections: Component method, Mortality basis for projections, Fertility basis for projections, Migration basis for projections, Ageing of the population, Estimation of demographic measures from incomplete data.

References:

1. Pollard, A.H. Yusuf, F. and Pollard, G.N. (1990). Demographic Techniques, Pergamon Press.
2. Keyfitz, N. (1977) Applied Mathematical Demography A Wiley-Interscience Publication.
3. Keyfilz, N. (1968) Introduction to the Mathematic of Population Ready, Mass: Addition-Wesley.
4. Keyfilz, N. and Caswell, H. (2005) Applied Mathematical Demography, Third edition, Springer.

STAT 207	TIME SERIES ANALYSIS	4 Credits

Unit 1

Time series as a discrete parameter stochastic process, Auto - Covariance, Auto-correlation and their properties. Exploratory time series analysis, Test for trend and seasonality, Exponential and moving average smoothing, Holt – Winter smoothing, forecasting based on smoothing.

Unit 2

Wold representation of linear stationary processes, detailed study of the linear time series models: Autoregressive, Moving Average, Autoregressive Moving Average and Autoregressive Integrated Moving Average models.

Unit 3

Estimation of ARMA models: Yule-Walker estimation for AR Processes, Maximum likelihood and least squares estimation for ARMA Processes, Discussion (without proof) of estimation of mean, Auto-covariance and auto-correlation function under large samples theory, Residual analysis and diagnostic checking. Forecasting using ARIMA models.

Unit 4

Analysis of seasonal models: parsimonious models for seasonal time series, General multiplicative seasonal models, forecasting, identification, estimation and diagnosis methods for seasonal time series. Spectral analysis of weakly stationary process. Herglotzic Theorem. Periodogram and correlogram analysis.

References:

1. Box, G.E.P and Jenkins G.M. (1970) Time Series Analysis, Forecasting and Control, Holden-Day.
2. Brockwell, P.J and Davis R.A. (1987) Time Series: Theory and Methods, Springer-Verlag.
3. Abraham, B. and Ledolter, J.C. (1983) Statistical Methods for Forecasting, Wiley
4. Anderson, T.W (1971) Statistical Analysis of Time Series, Wiley.
5. Fuller, W.A. (1978) Introduction to Statistical Time Series, John Wiley.
6. Kendall, M.G. (1978) Time Series, Charler Graffin.
7. Chatfield, C. (2004) The Analysis of Time Series - An Introduction, Sixth edition, Chapman and Hall.

STAT 208	**STATISTICAL QUALITY ASSURANCE**	**4 Credits**

Unit 1

Quality and quality assurance, Methods of quality assurance, Introduction to TQM and ISO 9000 standards, statistical quality control: Acceptance sampling for attributes, Single sampling, Double sampling, Multiple and sequential sampling plans, Measuring the performance of these plans.

Unit 2

Control charts, Basic ideas, designing of control charts for the number of non-conformities and fraction non-conformities, mean charts, Median charts, Extreme value charts, R-charts, and S-charts, ARL, Economic design of Shewarts control charts.

Unit 3

Acceptance sampling by variables, Sampling plans for a single specification limit with known and unknown variance, Sampling plans with double specification limits, Comparison of sampling plans by variable and attributes, Continuous sampling plans I, II and III.

Unit 4

Process capability studies, Statistical aspect of six sigma philosophy, Control charts with memory - CUSUM charts, EWMA-mean charts, OC and ARL for control charts, The Taguchi Method: The Taguchi philosophy of Quality, Loss functions, SN ratios, Performance measures, Experimental design in Taguchi Methods: Orthogonal arrays and linear graph, Estimation of effects, Parameter Design.

References:

1. Montgomery, R.C. (1985). Introduction to Statistical Quality Control, Fourth edition, Wiley.
2. Mittag, H.J. & Rinne, H. (1993). Statistical Methods for Quality Assurance, Chapman & Hall.
3. Schilling, E.G. (1982). Acceptance Sampling in Quality Control, Marcel Dekker.
4. Amitava Mitra (2001). Fundamentals of Quality Control and Improvement – Pearson Education Asia.
5. Duncan, A.J. (1986). Quality control and Industrial Statistics.
6. Grant E.L. and Leaven Worth, R.S. (1980). Statistical Quality Control, McGraw Hill.
7. Chin-Knei Cho (1987). Quality Programming, John Wiley.

STAT 209	APPLIED STATISTICS FOR NATION DEVELOPMENT	4 Credits

Unit 1

Demographic methods:- Sources of demographic data - census, register, adhoc survey, hospital records, demographic profiles of Indian census; Measurement of mortality and life tables - crude, death rates, infant mortality rates, death date by cause, standardized death rate; Complete life tables – its main features, mortality rate and probability of dying, use of survival tables; Measurement of fertility - crude birth rate, general fertility rate, total fertility rate, gross reproduction rate, net reproduction rate; Population growth in developing and developed countries; Population projection using Leslie metric; Labour force projection.

Unit 2

Economic statistics: - Index number - its definition, price relatives and quantity or volume relatives, link and chain relatives, consumer price index; Demand analysis - static laws of demand and supply, price elasticity of demand, analysis of income and allied size distribution - Pareto distribution, graphical test, fitting of Pareto's law, log normal distribution and its properties, Lorenz curve and estimation of elasticity; Gini coefficient.

Unit 3

Economic development, growth in per capita income and distributive justice, indices of development; Human Development Index, Estimation of national income - product approach, income approach and expenditure approach; Measuring inequality in incomes, poverty measurement - measures of incidence and intensity combined; Time Series:- components of time series, determination of trend, analysis of seasonal fluctuations, construction of seasonal indices, measurement of cyclic movement, random component in time series, smoothing methods,

Unit 4

Introduction to Indian and International Statistical System - role, function activities of Central and State Statistical Organizations; Organization of large scale sample surveys; Role of National sample survey organization; General and special data dissemination systems; Principal publications containing such statistics on the topics - population, agriculture industry, trade, price, labour and employment transport and communications, and finance; Educational and Psychological statistics:- Scaling individual test items, scaling of scores on a test, different types of scores and scaling, scaling of ranking and rating in terms of normal curve, Reliability of test scores, Rulon and Kuder Richardson methods, Reliability of a test, validity, comparison between reliability and validity, Intelligence coefficient.

References:

1. Basic Statistics Relating to Indian Economy (CSO), 1990 - Current Indian Statistics
2. Cox PR (1957) Demography, Cambridge University Press
3. Croxton F E and Crowder D J (1967) Applied General statistics, Prentice - Hall India.
4. Guide to current Indian Official Statistics CSO, Govt. of India, New Delhi
5. Guide to official Statistics (CSO) -1990
6. Kendall, M.G. and Stuart, A. (1966). The Advanced Theory of Statistics, Charles Griffin
7. Keyfitz, N. (1977) Applied Mathematical Demography - Springer Verlag.
8. Mukhopadhyay, P. Applied Statistics, Books and Allied (P) Ltd
9. Pollard, A H, Yusuf , F and Pollard, G.N. (1998) Demographic Techniques
10. Saluja M.P, Indian Official Statistical Systems, Statistics Publishing Society, Calcutta
11. Sen, A. (1997) : Poverty and inequality
12. Statistical System in Indian (CSO) 1995

13. UNESCO: Principles for Vital Statistics system, Series M-12.

STAT 210	DIRECTIONAL DATA ANALYSIS	4 Credits

Unit 1

Graphical representation of data, Frequency distribution, Measures of location, circular variance and concentration, Correction for mean grouping, Measures of skewness and kurtosis.

Unit 2

Circular models, distribution theory, independence, convolution, moments, distributions of an arc, mixtures, lattice distributions, wrapped normal, Cauchy, Poisson distributions, Von Mises, Fisher distribution characteristic functions, Polar distributions, isotropic random walk on the circle.

Unit 3

Point estimation, Cramer Rao type bound, sufficiency, Methods of estimation, testing hypothesis from parametric models. Neyman-Pearson and likelihood ratio principles.

Unit 4

Non-parametric methods: Tests for randomness, goodness of fit, Rayleigh's test. Durand and Greenwood's test, Range test, Kuper's test, Watson's test, Uniform score tests, Runs test, Rank sum test, Tests for dispersion.

References:

1. Mardia, K.V. (1972). Statistics of Directional data, Academic Press.
2. Batschelet, E. (1981). Circular Statistics in Biology, Academic Press.
3. Watson, G.S. (1983). Statistics on Spheres, Wiley.

STAT 211	STATISTICAL FINANCE	4 Credits

Unit 1

Basic concepts of financial markets. Forward contracts, futures contracts, options-call and put options, European option and American options. Hedgers, speculators, arbitrageurs. Interest rates, compounding, present value analysis, risk free interest rates. Returns, gross returns and log returns. Portfolio theory – trading off expected return and risk, one risky asset and one risk free asset. Two risky assets, estimated expected return. Optimal mix of portfolio CAPM, capital market line, betas and security market line.

Unit 2

Options, pricing via arbitrage, law of one price. Risk neutral valuation. Binomial model-single and multiperiod binomial model, martingale measure. Modeling returns: lognormal model, random walk model, geometric Brownian motion process. Ito lemma (without proof). Arbitrage theorem. The Black-Scholes formula. Properties of the Black-Scholes option cost, the delta hedging arbitrage strategy. Some derivatives, their interpretations and applications.

Unit 3

Volatility and estimating the volatility parameter. Implied volatility. Pricing American options. Pricing of a European option using Monte-Carlo and pricing an American option using finite difference methods. Call options on dividend-paying securities. Pricing American put options, Modeling the prices by adding jumps to geometric Brownian motion. Valuing investments by expected utility. Modeling security market: Self-financing portfolio and no arbitrage, price process models, division rule, product rule

Unit 4

Financial Time Series – Special features of financial series, Linear time series models: AR (1), AR(p), ARMA(p,q) processes, the first and second order moments, estimation and forecasting methods. Models for Conditional heteroscedasticity: ARCH(1), ARCH (p), GARCH(p,q) models and their estimation. Comparison of ARMA and GARCH processes.

References:

1. Sheldon M. Ross (2003). "An elementary introduction to Mathematical Finance", Cambridge University Press.
2. David Ruppert (2004). "Statistics and Finance an Introduction" – Springer International Eddition.
3. Masaaki Kijima (2003). "Stochastic process with applications to finance", Chapman Hall.
4. Ruey S. Tsay (2005). "Analysis of Time Series III ed", John Wiley & Sons
5. John C. Hull (2008). "Options, Futures and other derivatives", Pearson Education India.
6. Christian Gourieroux and Joann Jasiak (2005). "Financial Econometrics", New Age International (P) Ltd.
7. Cuthbertson K and Nitzsche D (2001). "Financial Engineering - Derivatives and Risk Management", John Wiley & Sons Ltd.

MODULE 2: PROBABILITY & MATHEMATICAL STATISTICS

STAT 212	ADVANCED DISTRIBUTION THEORY	3 Credits

Unit 1

Stopped sum distributions: Poisson stopped sum, Neyman type A, Poisson- binomial, Poisson-negative binomial, Legrangian Poisson distributions, Distributions of order Poisson, negative binomial, Logarithmic series, Binomial.

Unit 2

Bivariate discrete distributions: bivariate power series distributions, bivariate Poisson, negative binomial and logarithmic series distributions, properties of these distributions, bivariate hypergeometric distribution and its properties.

Unit 3

Bivariate continuous models, bivariate Pearson system, Farlie Morgenstern distribution; distributions with specified conditionals, bivariate Pareto of I, II, III and IV kind, multivariate Liouville distributions.

Unit 4

Record values - definition, properties, distribution of *n*th record, record values from exponential, Weibull and logistic; Moments relationships, characterizations.

References:

1. Johnson, N.L., Kotz, S. and Kemp, A.W. (1992) Univariate discrete distributions, second edition, Wiley.
2. Kocherlakota, S. and Kocharlakota, K. (1992) Bivariate Discrete Distributions, Marcel-Dekker.
3. Johnson, N.L., Kotz, S. and Balakrishnan, N. (1997) Discrete multivariate distributions, second edition, Wiley.
4. Kotz, S. , Balakrishnan, N. and Johnson, N.L. (2000) Continuous multivariate distributions, Volume I, John Wiley and Sons.
5. Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1998) Records, John Wiley and Sons.

STAT 213	ADVANCED PROBABILITY THEORY	3 Credits

Unit 1

Review of Elementary Probability theory, Basic properties of expectations, Sequences of integrals, Lebesgue-Stieltjes integrals, Weak convergence - Theorems.

Unit 2

Complete convergence: Kolmogorov's three-series and two series theorems, Decomposition of normal distribution, Levy metric, Zolotarev and Lindeberg-Feller Theorems; Berry-Esseen Theorem.

Unit 3

More on Infinitely divisible distributions, Convergence under UAN, Convergence to special distributions, Cauchy functional equation, Stable distributions.

Unit 4

Conditional expectations (general case), Random-Nikodyn theorem, Martingales, Doob's decomposition, L_P-spaces Martingales, Martingale limit theorems, Exchangeability, Definite's theorem.

References:

1. Galambos J (1988) Advanced Probability Theory, Marcel Dekker, New York.
2. Ash R. B (2000) Probability and Measure Theory, Second edition. Academic Press.

3. Billingsley P (1985) Probability and Measure, Second edition, John Wiley and Sons, NewYork.
4. Laha R.G. and Rohatgi, V.K. (1979) Probability Theory, John Wiley and Sons, NewYork.

STAT 214	STATISTICAL DECISION THEORY	3 Credits

Unit 1

Basic elements of a decision problem, Randomized and non-randomized decision rules, Estimation and testing of hypothesis as decision problems, Bayes approach to inference and decision, Loss functions, Prior and posterior distributions, Prior - Posterior analysis for Bernoulli, Poisson and normal processes, Decision principles and Baye's risk.

Unit 2

Utility theory, axioms, construction of utility functions, sufficiency, equivalence of Classical and Bayesian sufficiency, complete and essentially complete classes of decision rules.

Unit 3

Minimax analysis, Basic elements of game theory, General techniques of solving games, Finite games, Supporting and separating hyperplane theorems, Minimax theorem, Minimax estimation for normal and Poisson means.

Unit 4

Admissibility of Bayes and minimax rules, General theorems on admissibility, Robustness of Bayes rules, Invariant decision rules, Location parameter problems, Confidence and credible sets.

References:

1. James O. Berger (1980) Statistical Decision Theory and Bayesian Analysis, Springer Verlag
2. M.H. DeGroot (1970) Optimal Statistical Decisions, John Wiley
3. H. Raiffa and R. Schlaifer (2000) Applied Statistical Decision Theory, Wiley Classics
4. Zellener (1971) An Introduction to Bayesian inference in Econometrics, Wiley
5. Hayes J, G and Winkler R I (1976) Probability, Statistics and Decision, Dower
6. Anthony O' Hagan (1994) Kendall's Advanced theory of Statistics vol. 2B, Bayesian Inference John Wiley.

STAT 215	RELIABILITY THEORY	3 Credits

Unit 1

Basic concepts in reliability: Failure rate, mean, variance and percentile residual life, identities connecting them; Notions of ageing - IFR, IFRA, NBU, NBUE, DMRL, HNBUE, NBUC etc and their mutual implications; TTT transforms and characterization of ageing classes.

Unit 2

Non monotonic failure rates and mean residual life functions, Study of life time models viz. exponential, Weibull, lognormal, generalized Pareto, gamma with reference to basic concepts and ageing characteristics; Bath tub and upside down bath tub failure rate distributions.

Unit 3

Discrete time failure models:- Definition of basic functions and their properties; Ageing classes and their mutual implications, Reliability systems with dependents components:- Parallel and series systems, k out of n systems, ageing properties with dependent and independents components, concepts and measures of dependence in reliability - RCSI, LCSD, PF2, WPQD.

Unit 4

Reliability estimation using MLE - exponential, Weibull and gamma distributions based on censored and non censored samples; UMVUE estimation of reliability function; Bayesian reliability estimation of exponential and Weibull models.

References:

1. Lai, C.D and Xie, M. (2006) Stochastic ageing and dependence in reliability (Relevant topics) Springer.
2. Sinha S K (1986) Reliability and Life Testing, Wiley Eastern.
3. Barlow, R.E. and Proschan, F. (1975) Statisical Theory of Reliability and Life Testing, Holt, Reinhart and Winston.
4. Marshall, A.W. and Olkin, I. (2007) Life Distributions, Springer
5. Galambos, J. and Kotz, S. (1978) Characterization of Probability distributions, Springer
6. Lawless, J.F. (2003) Statistical Models and Methods for Life Data, Wiley.

STAT 216	**ORDER STATISTICS**	**3 Credits**

Unit 1

Basic distribution theory. Order statistics for a discrete parent. Distribution-free confidence intervals for quantiles and distribution-free tolerance intervals.

Unit 2

Conditional distributions, Order Statistics as a Markov chain. Order statistics for independently and not identically distributed (i.ni.d.) variates. Moments of order statistics. Large sample approximations to mean and variance of order statistics. Asymptotic distributions of order statistics.

Unit 3

Recurrence relations and identities for moments of order statistics. Distibution-free bounds for moments of order statistics and of the range. Random division of an interval and its applications.

Unit 4

Concomitants. Order statistics from a sample containing a single outlier. Application to estimation and hypothesis testing.

References:

1. Arnold, B.C. and Balakrishnan, N. (1989). Relations, Bounds and Approximations for Order Statistics, Vol. 53, Springer-Verlag.
2. Arnold, B. C., Balakrishnan, N. and Nagaraja H. N. (1992). A First Course in Order Statistics, John Wiley & Sons.
3. David, H. A. and Nagaraja, H. N. (2003). Order Statistics, Third Edition, John Wiley & Sons.
4. Dwass, M. (1967). Simple random walk and rank order statistics. Ann. Math. Statist. 38, 1042-1053.
5. Gibbons, J.D. and Chakraborti, S. (1992). Nonparametric Statistical Inference, Third Edition, Marcel Dekker.
6. Takacs, L. (1967). Combinatorial Methods in the Theory of Stochastic Processes, John Wiley & Sons.

STAT 217	BAYESIAN INFERENCE	3 Credits

Unit 1

Basics on minimaxity: subjective and frequents probability, Bayesian inference, Bayesian estimation , prior distributions, posterior distribution, loss function, principle of minimum expected posterior loss, quadratic and other common loss functions, Advantages of being a Bayesian HPD confidence intervals, testing, credible intervals, prediction of a future observation.

Unit 2

Bayesian analysis with subjective prior, robustness and sensitivity, classes of priors, conjugate class, neighborhood class, density ratio class different methods of objective priors : Jeffrey's prior, probability matching prior, conjugate priors and mixtures, posterior robustness: measures and techniques.

Unit 3

Model selection and hypothesis testing based on objective probabilities and Bayes' factors, large sample methods: limit of posterior distribution, consistency of posterior distribution, asymptotic normality of posterior distribution.

Unit 4

Bayesian Computations : analytic approximation, E- M Algorithm, Monte Carlo sampling, Markov Chain Monte Carlo Methods, Metropolis – Hastings Algorithm, Gibbs sampling, examples, convergence issues

References:

1. Albert Jim (2009). Bayesian Computation with R, second edition, Springer, New York
2. Bolstad W. M. (2007). Introduction to Bayesian Statistics 2nd Ed. Wiley, New York
3. Christensen R. Johnson, W. Branscum A. and Hanson T.E. (2011) Bayesian Ideas and data analysis : A introduction for scientist and Statisticians, Chapman and Hall, London.
4. Ghosh, J.K. Delampady M. and T. Samantha (2006). An Introduction to Bayesian Analysis: Theory and Methods, Springer, New York.
5. Lee P.M. (2004) Bayesian Statistics: An Introduction, Hodder Arnold, New York.
6. Rao C.R. Day D. (2006) Bayesian Thinking, Modeling and Computation, Handbook of Statistics, Vol.25.
7. Berger, J.O. (1985). Statistical Decision Theory and Bayesian Analysis, Springer Verlag, New York.

MODULE 3: BIO-STATISTICS

STAT 306	STATISTICAL GENETICS	3 Credits

Unit 1

Introduction, Mendel's Laws, Linkage and Crossing cover, Linkage Maps, Statistical Analysis for Segregation and Linkage: Single Factor Segregation, Two factor segregation, Defection of Linkage, Estimation of Linkage.

Unit 2

Random mating: Hardy-Weinberg law of equilibrium. Single Locus, Sex-linked genes, Autopraploids, Forces affecting gene frequency, Fisher's fundamental theorem, inbreeding: Mutation and migration different approaches, concepts and definition, Path Coefficients, Stochastic Process of gene-frequency change, Diffusion approach, Transition matrix approach.

Unit 3

Genetic components of variance: Relationship between phenotype and genotype, Different approaches, Genetic components of covariance between Track; Linkage effects, Sex-linked genes, Maternal effect, Epistatic interaction, Genotype X Environment interaction.

Unit 4

Heritability, Estimation of Heritability, Precision of Heritability estimates, Repeatability, Estimates of Genetic correlation, Generalized Heritability Relation between phenotypic selection and genotypic selection, Intensity of selection correlated, Response to selection. Selection for improving several characters.

References:

1. Narain, P. (1990). Statistical Genetics, Wiley.
2. Liu, B.H. (1998). Statistical Genomics, CRC Press, New York.
3. Falconer, D.S. (1970). Introduction to Genetics, Oliver & Boyd.

STAT 307	SURVIVAL ANALYSIS	3 Credits

Unit 1

Basic concepts: Definition and properties, Lifetime models - Continuous and discrete models, General formulation, Different types of censoring, Truncation and other incomplete data, Likelihood inference with censored data, Inference procedures for exponential distributions. Methods based on large sample theory, Inference procedures for gamma and inverse Gaussian distribution.

Unit 2

The Product-Limit Estimate, Nelson Aalan Estimators, Interval estimation of survival probabilities and quantiles, Asymptotic properties of the estimators, Probability plots and hazard plots, Estimation of hazard rate or density functions, Methods for truncated and interval censored data.

Unit 3

Proportional hazards regression models, Methods for continuous multiplicative hazards models, Estimation of β, Comparison of two or more lifetime Distribution, Justification and properties of the likelihood function, Adjustments for tied life times, Estimation of baseline hazard function and baseline survivor function.

Unit 4

Regression diagnostics techniques: Cox-Snell residual method for assessing the fit, Graphical checks, Deviance residuals, Hypothesis testing: One sample tests for hazard function and

survivor function, Two sample tests for comparing hazard rates and survivor functions , Discrete time hazard based models and competing risks models.

References:

1. Lawless J.F (2003) Statistical Models and Methods for Lifetime Data, Second Editon, John Wiley & Sons, Relevant Sections of the Chapters 1, 2, 3, 4, 6,7 and 10.
2. Klein J.P. and Moeschberger M.L. (2003) Survival Analysis - Techniques for censored and truncated data, Second Edition, Springer-Verlag, New York, Chapters 7 and 11.
3. Kalbfleisch J.D and Prentice, R.L. (2002) The Statistical Analysis of Failure Time Data, Second Edition, John Wiley & Sons Inc.
4. Hosmer Jr. D.W and Lemeshow S (1999) Applied Survival Analysis - Regression Modelling of Time to event Data, John Wiley & Sons. Inc.
5. Nelson. W (1982) Applied Life Data Analysis.
6. Miller, R.G. (1981) Survival Analysis, John Wiley.

STAT 308	**STATISTICAL ECOLOGY**	**3 Credits**

Unit 1

Population Dynamics: One species - exponential, logistic and Gompertz models. Two species - competition, coexistence, predator - prey oscillation, Lotka - Volterra equations, isoclines. Leslie matrix model for age structured populations. Survivorship curves - constant hazard rate, monotone hazard rate and bath-tub shaped hazard rates.

Unit 2

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

Unit 3

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.

Unit 4

Foraging Theory: Diet choice problem, patch choice problem, mean variance tradeoff.

References:

1. Gore A.P. and Paranjpe S.A.(2000) A Course on Mathematical and Statistical Ecology, Kluwer Academic Publishers.
2. Pielou, E.C. (1977) An Introduction to Mathematical Ecology (Wiley)
3. Seber, G.A.F. (1982) The estimation of animal abundance and related parameters 2nd Ed. (C.Griffin)
4. Clark, C.W. (1976) Mathematical bio-economics: the optimal management of renewable resources (John Wiley)
5. Maynard Smith J. (1982) Evolution and the theory of games (Cambridge University Press)
6. Stephens D.W. & Krebs, J. R. (1986) Foraging Theory (Princeton University Press).

STAT 309	**MEDICAL AND HEALTH STATISTICS**	**3 Credits**

Unit 1

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

Unit 2

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

Unit 3

Log-linear models, 2xK and 2x2x2 contingency tables. Logistic model. Analysis of binary data. Cross-control study designs, matched case-control studies

Unit 4

Survival data: Proportional hazards model, multivariate survival data. Causal Inference, Longitudinal data. Communicating results of epidemiological studies, ethical issues in epidemiology.

References:

1. Selvin : Statistical analysis of epidemiological data.
2. Diggle, Liang and Zeger : Analysis of longitudinal data
3. Agresti : Categorical Data Analysis.
4. Clayton and Hills : Statistical methods in Epidemiology
5. McCullagh and Nelder : Generalized Linear Models.
6. Brookemeyer and Gail : AIDS Epidemiology : A Quantitative Approach
7. Zhou, Obuchowski and McClish : Statistical Methods in Diagnostic Medicine.

STAT 310	**ACTUARIAL STATISTICS**	**3 Credits**

Unit 1

Basic deterministic model: Cash flows, discount function, Interest and discount rates, balances and reserves, internal rate of return, The life table: Basic definitions, probabilities, construction of life tables, life expectancy, Life annuities: Introduction, calculating annuity premiums, interest and survivorship discount function, guaranteed payments, deferred annuities

Unit 2

Life insurance: Introduction, calculation of life insurance premiums, types of life insurance, combined benefits, insurances viewed as annuities, Insurance and annuity reserves: The general pattern of reserves, recursion, detailed analysis of an insurance, bases for reserves, non forfeiture values, policies involving a return of the reserve, premium difference and paid-up formula.

Unit 3

Fractional durations: Life annuities paid monthly, immediate annuities, fractional period premium and reserves, reserves at fractional durations, Continuous payments: Continuous annuities, force of discount, force of mortality, Insurance payable at the moment of death, premiums and reserves. The general insurance - annuity identity, Select mortality: Select an ultimate table, Changes in formulas.

Unit 4

Multiple life contracts: Joint life status, joint annuities and insurances, last survivor annuities and insurances, moment of death insurances. The general two life annuity and insurance contracts, contingent insurances Multiple decrement theory: Basic model, insurances, Determination of the models from the forces of decrement. Stochastic approach to insurance and annuities: Stochastic approach to insurance and annuity benefits, deferred contracts, Stochastic approach to reserves and premiums, variance formula.

References:

1. Promislow, S.D (2006) Fundamentals of Actuarial Mathematics, John Wiley, Chapters 2-11 & 14.
2. Neill, A. (1977) Life Contingencies, Heinemann, London.
3. Newton L.Bowers, Jr, Hans U.Gerber, James C.Hickman, Donald A. Jones and Cecil J. Nesbitt (1997) Actuarial Mathematics, The Society of Actuaries.
4. King, G. Institute of Actuaries Text Book. Part 11, Second edition, Charles and Edwin Layton, London.
5. Donald D.W.A. (1970) Compound Interest and Annuities, Heinemann, London.
6. Jordan, C.W. Jr. (1967) Life Contingencies, Second edition, Chicago Society of Actuaries.
7. Hooker, P.F. and Longley Cook, L.W. (1953) Life and Other Contingencies, Volume I and Volume II (1957) Cambridge University Press.
8. Spurgeen, E.T. Life Contingencies, Third edition, Cambridge University Press.
9. Benjamin, B. and Pollard, J.H. (1980) Analysis of Mortality and Other Actuarial Statistics, 2nd edition, Heinemann, London.
10. Freeman, H. (1960) Finite Differences for Actuarial Students, Cambridge University Press.
11. Biandt-Johnson, R.C. and Johnson, N.L. (1980) Survival Models and Data Analysis, John Wiley.

STAT 311	STATISTICS FOR CLINICAL TRIALS	3 Credits

Unit 1

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.

Unit 2

Design of clinical trials: parallel vs. cross-over designs, cross-sectional vs. longitudinal designs, review of factorial designs, objectives and endpoints of clinical trials.

Unit 3

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 bioequivalence trials.

Unit 4

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

References:

1. C. Jennison and B.W.Turnbul (1999). Group Sequential Methods with Applications to Clinical Trials, CRC Press.
2. E. Marubeni and M.G. Valsecchi (1994). Analyzing Survival Data from Clinical Trials and Observational Studies, Wiley and Sons.
3. J.L. Fleiss (1989). The Design and Analysis of Clinical Experiments. Wiley and Sons.
4. L.M.Friedman, C.Furburg, D.L.Demets (1998).Fundamentals of Clinical Trials, Springer Verlag.
5. S. Piantadosi (1997). Clinical Trials: A Methodological Perspective. Wiley and Sons.

MODULE 4: COMPUTATIONAL STATISTICS

STAT 406	COMPUTER INTENSIVE STATISTICAL METHODS	4 Credits

Unit 1

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.

Unit 2

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

Unit 3

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

Unit 4

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

References:

1. Fishman, G.S. (1996) Monte Carlo: Concepts, Algorithms, and Applications.(Springer).
2. Rubinstein, R.Y. (1981); Simulation and the Monte Carlo Method. (Wiley).
3. Efron, B. and Tibshirani. R.J. (1993); An Introduction to the Bootstrap.
4. Davison, A.C. and Hinkley, D.V. (1997) Bootstrap methods and their applications (Chapman and Hall).
5. Shao J. and Tu, D. (1995); The Jackknife and the Bootstrap. Springer Verlag.
6. McLachlan, G.J. and Krishnan, T. (1997) The EM Algorithms and Extensions. (Wiley.)
7. Kennedy W. J. & Gentle J. E. (1980) Statistical computing (Marcel Dekker)

STAT 407	STATISTICAL SIMULATIONS	4 Credits

Unit 1

Stochastic simulations: generating random variables, simulating normal, gamma and beta random variables. Comparison of algorithms to generate random variables. Generating random variables from failure rates.

Unit 2

Simulating multivariate distributions, MCMC methods and Gibbs sampler, Simulating random fields, simulating stochastic processes.

Unit 3

Variance reduction techniques: importance sampling for integration, control variates and antithetic variables. Simulating a non-homogeneous Poisson process.

Unit 4

Optimization using Monte Carlo methods, simulated annealing for optimization. Solving differential equations by Monte Carlo methods.

References:

1. Fishman, G.S. (1996) Monte Carlo: Concepts, Algorithms, and Applications. (Springer).
2. Rubinstein, R.Y. (1981) Simulation and the Monte Carlo Method. (Wiley).
3. Ripley B.D. (1987) Stochastic Simulations (Wiley)
4. Ross, S.M.(2002) Simulation (Third Edition) (Academic)

STAT 408	KNOWLEDGE DISCOVERY AND DATA MINING	4 Credits

Unit 1

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

Unit 2

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.

Unit 3

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

Unit 4

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

References:

1. Berson, A. and Smith, S.J. (1997). Data Warehousing, Data Mining, and OLAP. (McGraw-Hill.)
2. Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984). Classification and Regression Trees. (Wadsworth and Brooks/Cole).
3. Han,J. and Kamber, M. (2000). Data Mining; Concepts and Techniques. (Morgan Kaufmann.)
4. Mitchell, T.M. (1997). Machine Learning. (McGraw-Hill.)
5. Ripley, B.D. (1996). Pattern Recognition and Neural Networks. (Cambridge University Press).

STAT 409	STATISTICAL PATTERN RECOGNIZITION	4 Credits

Unit 1

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

Unit 2

Nearest neighbour decision rules: description, convergence, finite sample considerations, use of branch and bound methods.

Unit 3

Probability of errors: two classes, normal distributions, equal covariance matrix assumptions, Chernoff bounds and Bhattacharya distance, estimation of probability of error.

Unit 4

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

References:

1. Duda, R.O. and Hart, P.E. (1973). Pattern Recognition and Scene Analysis. (Wiley).
2. Fukunaga, K. (1990). Introduction to Statistical Pattern Recognition, 2nd Ed. (Academic Press).
3. McLachlan, G.J. (1992). Discriminant Analysis and Statistical Pattern Recognition. (Wiley).
4. Ripley, B.D. (1996). Pattern Recognition and Neural Networks. (Cambridge University Press).

STAT 410	LARGE SCALE DATA ANALYSIS	4 Credits

Unit 1

Generalization of Linear Regression- Ridge Regression, Partial least squares, LASSO and Least angle regression, Principal Components Regression.

Unit 2

Tree based methods- Classification and Regression Trees (CART), Patient rule induction method (PRIM), Multivariate Adaptive Regression Splines (MARS).

Unit 3

Generalization of Linear Discriminant Analysis- Flexible Discriminant Analysis Penalized Discriminant Analysis, Mixture Discriminant Analysis.

Unit 4

Generalization of Principal Component Analysis- Kernel Principal Components, Sparse Principal Component Analysis, Independent Component Analysis (ICA). Multidimensional Scaling. Applications of above methods in Astronomical 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. A.Hyvarinen, et al. : Independent Component Analysis
4. R.Stephen & E.Richard : Independent Component Analysis –Principles and Practice
5. R.A.Johnson & D.W.Wichern : Applied Multivariate Statistical Analysis

STAT 411	ADVANCE DATA ANALYTIC TECHNIQUES	4 Credits

Unit 1

Introduction to Jackknife and Bootstrap-methods for estimating bias, standard error and distribution function based on i.i.d. random variables, standard examples, Bootstrap confidence intervals.

Unit 2

Informative or non-informative missingness; complete case / available case estimation, Imputation, EM & MCEM algorithms and data augmentation techniques. Standard error estimation.

Unit 3

Longitudinal regression: Cohort vs. longitudinal effect, weighted least-squares, ML and REML techniques.

Unit 4

Marginal, subject specific and transition models, 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

****