SHIVAJI UNIVERSITY, KOLHAPUR

SYLLABUS

For

M.Sc. (Statistics) Part I

Choice Based Credit System
(CBCS) as per NEP-2020

To be implemented from

Academic Year 2022-23

M.Sc. Statistics Part I
Syllabus: w.e.f. academic year 2022-23

A. Ordinance and Regulations: - (as applicable to degree/ programme)

B. Shivaji University, Kolhapur, New/Revised Syllabus for Master of Science

1. Title of the Programme: **M. Sc. (Statistics)**
2. Faculty: Science and Technology
3. Year of Implementation: M.Sc. part-I: Academic year 2022-23
4. Preamble
5. Learning outcomes

- **Programme Outcomes**

  Post Graduates of the M.Sc. Statistics programme will be able to:

  1) Have sufficient knowledge of theoretical concepts in Statistics for (i) choosing and applying the most appropriate statistical methods/ techniques for collecting and analyzing data (ii) interpreting the results of analyses in relation to given real life situations.
  2) Have deep understanding and ability to explain the inter-connections between various sub disciplines and apt use of these inter-connections in modelling real life problems.
  3) Have ability to recognize the importance of statistical thinking and training, and to acquire the state-of-the-art developments in Statistics independently from available resources.
  4) Develop expertise in data management and analysis using widely used statistical software.

- **Programme Specific Outcomes (POs)**

  After completion of M.Sc. Statistics programme the student will be able to:

  1) Develop stochastic models for studying, analyzing, interpreting and forecasting real life phenomenon in diverse disciplines.
  2) Effectively use necessary statistical software and computing environment including R, MS-EXCEL among others.
  3) Have the versatility to work effectively in a broad range of establishments (including R&D sectors, analytics, scientific laboratories, government, financial, health, educational) or to continue for higher education, and exhibit ethical and professional behaviour in team work.

6. Duration: 2 Years
7. Pattern: CBCS as per NEP-2020
8. Fee Structure: As per university guidelines
9. Eligibility criteria for Admission: B. Sc. with Statistics as principal subject
10. Medium of Instruction: English

11. Structure of the Programme, Scheme of Teaching and Examination:

M.Sc. Part – I (Level-8)

SEMESTER-I (Duration- Six Month)												
	Sr. No.	Course Code	Teaching Scheme			Examination Scheme						
			Theory and Practical			University Assessment (UA)			Internal Assessment (IA)			
			Lectures (Per week)	Hours (Per week)	Credit	Maximum Marks	Minimum Marks	Exam. Hours	Maximum Marks	Minimum Marks	Exam. Hours	
CGPA	1	CC-101	4	4	4	80	32	3	20	8	1	
	2	CC-102	4	4	4	80	32	3	20	8	1	
	3	CC-103	4	4	4	80	32	3	20	8	1	
	4	CC-104	4	4	4	80	32	3	20	8	1	
	5	CC-105	4	4	4	80	32	3	20	8	1	
	6	CCPR-106	12	12	4	100	40	*	--	--	--	
Total (A)			**--**	**--**	**24**	**500**	**--**	**--**	**100**	**--**	**--**	
Non-CGPA	1	AEC-107	2	2	2	--	--	--	50	20	2	
SEMESTER-II (Duration- Six Month)												
CGPA	1	CC-201	4	4	4	80	32	3	20	8	1	
	2	CC-202	4	4	4	80	32	3	20	8	1	
	3	CC-203	4	4	4	80	32	3	20	8	1	
	4	CC-204	4	4	4	80	32	3	20	8	1	
	5	CC-205	4	4	4	80	32	3	20	8	1	
	6	CCPR-206	12	12	4	100	40	*	--	--	--	
Total (B)			**--**	**--**	**24**	**500**	**--**	**--**	**100**	**--**	**--**	
Non-CGPA	1	SEC-207	2	2	2	--	--	--	50	20	2	
Total (A+B)					**48**	**1000**	**--**	**--**	**200**	**--**	**--**	

• Student contact hours per week : **32** Hours (Min.)	• Total Marks for M.Sc.-I : **1200**
• Theory and Practical Lectures : **60** Minutes Each	• Total Credits for M.Sc.-I (Semester I & II) : **48**
• CC-Core Course • CCPR-Core Course Practical • AEC-Mandatory Non-CGPA compulsory Ability Enhancement Course • SEC- Mandatory Non-CGPA compulsory Skill Enhancement Course	• Practical Examination is annual. • Examination for CCPR-106 shall be based on Semester I Practical. • Examination for CCPR-206 shall be based on Semester II Practical. • *Duration of Practical Examination as per respective BOS guidelines • *Separate passing is mandatory for Theory, Internal and Practical Examination*
• **Requirement for Entry at Level 8:** **Completed all requirements of the relevant Bachelor's degree (Level 7) with principal / major subjects Statistics**	
• **Exit Option at Level 8:** Students can exit after Level 8 with **Post Graduate Diploma Statistics** if he/she completes the courses equivalent to minimum of 48 credits.	

M.Sc. Part – II(Level-9)

SEMESTER-III (Duration- Six Month)												
	Sr. No.	Course Code	Teaching Scheme			Examination Scheme					
			Theory and Practical			University Assessment (UA)			Internal Assessment (IA)		
			Lectures (Per week)	Hours (Per week)	Credit	Maximum Marks	Minimum Marks	Exam. Hours	Maximum Marks	Minimum Marks	Exam. Hours
CGPA	1	CC-301	4	4	4	80	32	3	20	8	1
	2	CCS -302	4	4	4	80	32	3	20	8	1
	3	CCS-303	4	4	4	80	32	3	20	8	1
	4	CCS-304	4	4	4	80	32	3	20	8	1
	5	DSE -305	4	4	4	80	32	3	20	8	1
	6	CCPR-306	12	12	4	100	40	*	--	--	-
Total (C)			--	--	**24**	**500**	--	--	**100**	--	
Non-CGPA	1	AEC-307	2	2	2	--	--	--	50	20	2
	2	EC-308	Number of lectures and credit shall be as specified on SWAYAM MOOC								
SEMESTER-IV (Duration- Six Month)											
CGPA	1	CC-401	4	4	4	80	32	3	20	8	1
	2	CCS -402	4	4	4	80	32	3	20	8	1
	3	CCS-403	4	4	4	80	32	3	20	8	1
	4	CCS-404	4	4	4	80	32	3	20	8	1
	5	DSE -405	4	4	4	80	32	3	20	8	1
	6	CCPR-406	12	12	4	100	40	*	--	--	--
Total (D)			--	--	**24**	**500**	--	--	**100**	--	--
Non-CGPA	1	SEC-407	2	2	2	--	--	--	50	20	2
	2	GE-408	2	2	2	--	--	--	50	20	2
Total (C+D)					**48**	**1000**	--	--	**200**	--	--

• Student contact hours per week : **32** Hours (Min.)	• Total Marks for M.Sc.-II : **1200**
• Theory and Practical Lectures : **60** Minutes Each	• Total Credits for M.Sc.-II (Semester III & IV) : **48**
• CC-Core Course	• Practical Examination is annual.

• CCS- Core Course Specialization • CCPR-Core Course Practical • DSE-Discipline Specific Elective • AEC-Mandatory Non-CGPA compulsory Ability Enhancement Course • SEC- Mandatory Non-CGPA compulsory Skill Enhancement Course • EC (SWM MOOC) - Non-CGPA Elective Course • GE-Generic Elective	• Examination for CCPR-306 shall be based on Semester III Practical. • Examination for CCPR-406 shall be based on Semester IV Practical. • *****Duration of Practical Examination as per respective BOS guidelines • ***Separate passing is mandatory for Theory, Internal and Practical Examination***
• **Requirement for Entry at Level 9:** **Completed all requirements of the relevant Post Graduate Diploma (Level 8) in Statistics**	
• **Exit at Level 9:** Students will exit after Level 9 with **Master's Degree in Statistics** if he/she completes the courses equivalent to minimum of 96 credits.	

	M.Sc.-I	M.Sc.-II	Total
Marks	1200	1200	2400
Credits	48	48	96

I. CGPA course:

1. There shall be 12 Core Courses (CC) of 48 credits per programme.
2. There shall be 06 Core Course Specialization (CCS) of 24 credits per programme.
3. There shall be 02 Discipline Specific Elective (DSE) courses of 08 credits per programme.
4. There shall be 4 Core Course Practical (CCPR) of 16 credits per programme.
5. Total credits for CGPA courses shall be of 96 credits per programme.

II. Mandatory Non-CGPA Courses:

1. There shall be 02 Mandatory Non-CGPA compulsory Ability Enhancement Courses (AEC) of 02 credits each per programme.
2. There shall be 01 Mandatory Non-CGPA compulsory Skill Enhancement Course (SEC) of 02 credits per programme.
3. There shall be one Elective Course (EC) (SWAYAM MOOC). The credits of this course shall be as specified on SWAYAM MOOC.
4. There shall be one Generic Elective (GE) course of 02 credits per programme. Each student has to take generic elective from the department other than parent department.
5. The total credits for Non-CGPA course shall be of 08 credits + 2-4 credits of EC as per availability.
6. The credits assigned to the courses and the programme are to be earned by the students and shall not have any relevance with the work load of the teacher.

12. Courses being offered

Semester I

	Course code	Title of course
CGPA	CC-101	Real Analysis
	CC-102	Linear Algebra
	CC-103	Distribution Theory
	CC-104	Estimation Theory
	CC-105	Statistical Computing
	CCPR-106	Practical I
Mandatory Non-CGPA	AEC-107	AEC: Communicative English-I

Semester II

	Course code	Title of course
CGPA	CC-201	Probability Theory
	CC-202	Theory of Testing of Hypothesis
	CC-203	Linear Models and Regression Analysis
	CC-204	Design and Analysis of Experiments
	CC-205	Sampling Theory and Official Statistics
	CCPR-206	Practical II
Mandatory Non-CGPA	SEC-207	SEC: Fundamentals of Information Technology-I

Semester III

	Course code	Title of course
CGPA	CC-301	Asymptotic Inference
	CCS-302	Multivariate Analysis
		Bayesian Inference
	CCS-303	Stochastic Processes
		Functional Data Analysis
	CCS-304	Data Mining
		Artificial Intelligence
	DSE-305	Time Series Analysis
		Statistical Ecology
	CCPR-306	Practical III
Mandatory Non-CGPA	AEC-307	AEC: Communicative English-II
	EC-308	EC: SWYAM/ MOOCS

Semester IV

	Course code	Title of course
CGPA	CC-401	Generalized Linear Models
	CCS-402	Survival Analysis
		Actuarial Statistics
	CCS-403	Biostatistics
		Econometrics
	CCS-404	Optimization Techniques
		Circular Data Analysis
	DSE-405	Spatial Data Analysis
		Statistical Quality Control
	CCPR-406	Practical VI and Project
Mandatory Non-CGPA	SEC-407	SEC: Fundamentals of Information Technology-II
	GE-408	GE: to be selected form the pool of courses offered by the other departments.

13. Standard of passing: 40% in each course. Separate passing for internal and semester examination.
14. Nature of Question paper and Scheme of marking

- **Nature of the theory question papers:-**
  a) There shall be 7 questions each carrying 16 marks.
  b) Question No.1 is compulsory. It consists of 8 questions for 2 marks each.
  c) Students have to attempt any 4 questions from question No. 2 to 7.
  d) Question No. 2 to 6 shall contain 2 to 4 sub-questions.
  e) Question No. 7 shall contain 4 short note type questions, each carrying 4 marks.

- **Practical Paper:-**
  a) Semester I, II, III "Practical CCPR-106, CCPR-206 and CCPR-306"
     1. There shall be 20 marks for day-to-day performance and journal.
     2. Examination (60): Practical Examinations of practical I and II will be conducted at the end the of respective year and practical III and IV will be conducted at the end of the respective year.
     3. Each exam will be of 3 hrs. duration carrying 60 marks. There shall be 8 questions each of 12 marks, of which a student has to attempt any 5 questions.
     4. Practical VIVA will be for 20 marks.

  b) Semester IV, Practical CCPR-406
     1. There shall be 10 marks for day-to-day performance and journal.
     2. Examination (30): Practical exam will be of 2 hrs. duration carrying 30 marks. There shall be 5 questions each of 10 marks, of which a student has to attempt any 3 questions.
     3. Practical VIVA will be for 20 marks.
     4. Project work carries 40 marks. Project work consists of understanding the domain of the problem, formulation of the problem, collection of the relevant data, Analysis of the data and report writing. They are expected to use software for which they are trained. 20 marks are reserved for project based VIVA. Project report will be evaluated for 20 marks. The project work should be preferably based on field work or problem in industry. Primary data collection using technically correct most feasible sampling scheme will carry 5 marks. If only secondary data is used, the student will loose these marks.

15. Each of the following courses have the same question paper in all examinations of M.Sc.(Statistics) and M.Sc.(Applied Statistics and Informatics)

Semester I

Course code	Title of course
CC-102	Linear Algebra
CC-103	Distribution Theory
CC-104	Estimation Theory
CC-105	Statistical Computing

Semester II

Course code	Title of course
CC-202	Theory of Testing of Hypothesis
CC-203	Linear Models and Regression Analysis
CC-204	Design and Analysis of Experiments
CC-205	Sampling Theory and Official Statistics

Semester III

Course code	Title of course
CCS-302	Multivariate Analysis
	Bayesian Inference
CCS-303	Stochastic Processes
	Functional Data Analysis
CCS-304	Data Mining
	Artificial Intelligence
DSE-305	Time Series Analysis
	Statistical Ecology

Semester IV

Course code	Title of course
CC-401	Generalized Linear Models
CCS-402	Survival Analysis
	Actuarial Statistics
CCS-403	Biostatistics
	Econometrics
DSE-405	Spatial Data Analysis
	Statistical Quality Control

16. Equivalence in accordance with titles and contents of courses

Semester-I			
Course Code		Title of the course	
Old	**New**	**Old**	**New**
CC-101	CC-101	Real Analysis	Real Analysis
CC-102	CC-102	Linear Algebra	Linear Algebra
CC-103	CC-103	Distribution Theory	Distribution Theory
CC-104	CC-104	Estimation Theory	Estimation Theory
CC-105	CC-105	Statistical Computing	Statistical Computing
CCPR-106	CCPR-106	Practical-I	Practical-I
AEC-107	AEC-107	Communicative English-I	Communicative English-I
Semester-II			
Course Code		Title of the course	
Old	**New**	**Old**	**New**
CC-201	CC-201	Probability Theory	Probability Theory
CC-202	CC-202	Theory of Testing of Hypothesis	Theory of Testing of Hypothesis
CC-203	CC-203	Regression Analysis	Linear Models and Regression Analysis
CC-204	CC-204	Design and Analysis of Experiments	Design and Analysis of Experiments
CC-205	CC-205	Sampling Theory and Official Statistics	Sampling Theory and Official Statistics
CCPR-206	CCPR-206	Practical - II	Practical II
SEC-207	SEC-207	Fundamentals of Information Technology-I	Fundamentals of Information Technology-I
Semester-III			
Course Code		Title of the course	
Old	**New**	**Old**	**New/Substitute**
CC-301	CC-301	Asymptotic Inference	Asymptotic Inference
CCS-302	CCS-302	Multivariate Analysis	Multivariate Analysis
		Bayesian Inference	Bayesian Inference
CCS-303	CCS-303	Stochastic Processes	Stochastic Processes
		Functional Data Analysis	Functional Data Analysis
CCS-304	CCS-304	Data Mining	Data Mining
		Artificial Intelligence	Artificial Intelligence
DSE-305	DSE-305	Time Series Analysis	Time Series Analysis
		Statistical Ecology	Statistical Ecology
CCPR-306	CCPR-306	Practical III	Practical III
AEC-307	AEC-307	AEC: Communicative English-II	AEC: Communicative English-II
EC -308	EC -308	EC: SWYAM/ MOOCS	EC: SWYAM/ MOOCS
Semester-IV			
Course Code		Title of the course	
Old	New	Old	New/Substitute
CC-401	CC-401	Generalized Linear Models	Generalized Linear Models
CCS-402	CCS-402	Survival Analysis	Survival Analysis
		Actuarial Statistics	Actuarial Statistics
CCS-403	CCS-403	Biostatistics	Biostatistics
		Econometrics	Econometrics
CCS-404	CCS-404	Optimization Techniques	Optimization Techniques
		Circular Data Analysis	Circular Data Analysis
DSE-405	DSE-405	Spatial Data Analysis	Spatial Data Analysis
		Statistical Quality Control	Statistical Quality Control
CCPR-406	CCPR-406	Practical VI and Project	Practical VI and Project
SEC-407	SEC-407	SEC: Fundamentals of Information Technology-II	SEC: Fundamentals of Information Technology-II

Semester I

CC-101: REAL ANALYSIS

Unit 1: Set of real numbers, countable and uncountable sets, countability of rational numbers, uncountability of the interval (0, 1) and other uncountable sets. Supremum and Infimum of bounded sets, limit point(s) of a set, closur of a set, open, closed, dense and compact sets and their properties. Bolzano-Weierstrass and Heine-Borel Theorems (Statements only). Applications of these theorems.

(12 L + 3 T)

Unit 2: Sequences of real numbers, convergence, divergence, monotone, bounded and unbounded sequences, Cauchy sequence, Convergence of bounded monotone sequence. Limit points, Limit inferior and limit superior of the sequences and their properties. Subsequences and properties associated with them. Series of numbers, tests for convergence (without proof) test for absolute convergence, convergence of series of non-negative terms.

(12 L + 3 T)

Unit 3: Real valued functions, continuous functions, Uniform continuity of functions and sequences of functions, Uniform convergence of series of functions with special emphasis on power series, radius of convergence. Riemann, Riemann-Steltjes Integrals and their common properties. Upper and lower integrals, integrability of functions, Integration by parts, Fundamental theorem on calculus, mean value theorem, their applications in finding functional of probability distributions.

(12 L + 3 T)

Unit 4: Vector and Matrix differentiation, Maxima, minima of functions of several variables. Constrained maxima, minima, Lagrange's method, Taylor's theorem (without proof), implicit function theorem and their applications. Multiple integrals, Change of variables, Improper integrals, Applications in multivariate probability distributions. Theorem on differentiation under integral sign and Leibnitz rule (statements only) with applications.

(12 L + 3 T)

References:

1. Malik S. C. &Arora S.(1991): Mathematical Analysis- Wiley Eastern Limited IInd edition.
2. Goldberg R. R. (1964): Methods of Real Analysis- Blaisdell Publishing company, New York, U.S.A.
3. Bartle G. R. (1976): Element of Real Analysis- Wiley, 2nd edition.
4. Bartle G.R. & Sherbert D. R. (2000): Introduction to Real Analysis- John Wiley & Son Inc.
5. Royden (1988): Principles of Real Analysis - Macmillian.
6. Shanti Narayan, Raisinghania M.D. (2013) Elements of Real Analysis, Fourteenth Revised Edition, S. Chand.
7. Apostol (1985): Mathematical Analysis - Narosa Publishing House, T.M.

CC-102: LINEAR ALGEBRA

Unit 1: Vector space, subspace, linear dependence and independence, basis, dimension of a vector space, example of vector spaces. Gram-Schmidt orthogonalisation process, Orthonormal basis, orthogonal projection of a vector, Linear transformations, algebra of matrices, types of matrices, row and column spaces of a matrix, elementary operations and elementary matrices, rank and inverse of a matrix, null space and nullity, partitioned matrices.

(12L+3T)

Unit 2: Permutation matrix, reducible/irreducible matrix, primitive/ imprimitive matrix, Kronecker product, Generalized inverse, Moore-Penrose generalized inverse, Solution of a system of homogenous and non-homogenous linear equations, theorem related to existence of solution and examples.

(12L+3T)

Unit 3: Characteristic roots and vectors of a matrix, algebraic and geometric multiplicities of a characteristic root, right and left characteristic vectors, orthogonal property of characteristic vectors, Cayley-Hamilton Theorem and its applications.

(12L+3T)

Unit 4: Spectral decomposition of a real symmetric matrix, singular value decomposition, Choleskey decomposition, real quadratic forms, reduction and classification, index and signature, extrema of a quadratic form, simultaneous reduction of two quadratic forms.

(12L+3T)

References:

1. Graybill, F.A (1961) An Introduction to Linear Statistical Models Vol 1, McGraw-Hill Book Company Inc.
2. Hadely G. (1962) Linear Algebra, Narosa Publishing House.
3. Harville D. (1997) Matrix Algebra From Statistics Perspective, Springer.
4. Rao A R. and Bhimasankaram P. (2000), Linear Algebra, Second edition, Hindustan Book Agency.
5. Rao C. R. (2001) Linear Statistical Inference and Its Applications, Second Edition, Wiley.
6. Schott J. (2016) Matrix Analysis for Statistics, Third edition Wiley.
7. Searl S. B.(2006) Matrix Algebra Useful for Statistics, Wiley.

CC-103: DISTRIBUTION THEORY

Unit 1: Review of Random experiment and its sample space, events, random variables, discrete, continuous and mixed type of random variables. Cumulative distribution function (CDF) and its properties, absolutely continuous and discrete distributions, mixtures of probability distributions, decomposition of mixture CDF into discrete and continuous CDFs, computation of probabilities of events using CDF, expectation and variance of mixture distributions. Quantiles of probability distributions and their computation for some well-known distributions.

(12L+3T)

Unit 2: Transformations of univariate random variables, probability integral transformation. Concepts of location, scale and shape parameters of distributions with examples. Symmetric distributions and their properties. Moment inequalities (with proof): Basic, Holder, Markov, Minkowski, Jensen, Chebyshev's inequality, and their applications.

(12L+3T)

Unit 3: Random vectors, joint distributions, Independence, variance-covariance matrix, joint MGF. Conditional expectation and variances, Transformations of bivariate random variables, Gumbel and Marshall-Olkin bivariate exponential distributions, Bivariate Poisson distribution. Convolutions, compound distributions.

(12L+3T)

Unit 4: Sampling distributions of statistics from univariate normal random samples: central and non-central chi-square, t and F distributions. Distributions of linear and quadratic forms involving normal random variables, Fisher Cochran and related theorems: statement and applications. Order Statistics: Distribution of an order statistics, joint distributions of two order statistics, distribution of spacings, normalized spacings with illustration to exponential case, distribution of sample median and sample range.

(12L+3T)

References:

1. Rohatagi V. K. & Saleh A. K. Md. E.(2001) : Introduction to Probability Theory and Mathematical Statistics- John Wiley and sons Inc.
2. Johnson N. L. & Kotz. S. (1996) : Distributions in Statistics Vol-I,II and III, John Wiley and Sons New York.
3. S. Kotz, N. Balakrishnan, N. L. Johnson: Continuous Multivariate Distributions - Second Edition, Wiley.
4. Casella & Berger (2002) : Statistical Inference - Duxbury advanced series. IInd edition.
5. C. R. Rao (1995) Linear Statistical Inference and Its Applications (Wiley Eastern) Second Edition.
6. Dasgupta, A. (2010) Fundamentals of Probability: A First Course (Springer).

CC-104: ESTIMATION THEORY

Unit 1: Principles of data reduction: sufficiency principle; sufficient statistics; factorization theorem; minimal sufficient statistic; minimal sufficient statistic for exponential family, power series family, curved exponential family, and Pitman family; completeness; bounded completeness; ancillary statistics, Basu's theorem and it's applications.

(12L + 3T)

Unit 2: Unbiased estimation: unbiased estimator; uniformly minimum variance unbiased estimator (UMVUE); A necessary and sufficient condition for an estimator to be UMVUE; Rao-Blackwell theorem and Lehmann-Scheffe theorem, and their applications in finding UMVUEs; Fisher information function and Fisher information matrix; concept of parameter orthogonality and its advantages for maximum likelihood estimation; Cramer-Rao lower bound; Chapman-Robbins-Kiefer lower bound.

(12L + 3T)

Unit 3: Methods of finding estimators: method of moments estimator; maximum likelihood estimator (MLE), properties of MLE, MLE in nonregular families; method of scoring; method of minimum chi-square, EM algorithm. Nonparametric estimation: degree of an estimable parameter, kernel, U-statistic and its properties.

(12L + 3T)

Unit 4: Bayesian estimation: The concepts of prior and posterior distributions, conjugate prior distribution, Jeffery priors and improper priors with examples, Bayes estimation under squared error and absolute error loss functions.

(12L + 3T)

References:

1. Rohatgi, V.K. and Saleh, A. K. MD. E. (2015). *Introduction to Probability Theory and Mathematical Statistics* - 3rd edition, John Wiley & sons.
2. Lehmann, E. L. (1983). *Theory of Point Estimation* - John Wiley & sons.
3. Rao, C. R.(1973). *Linear Statistical Inference and its Applications*, 2nd edition, Wiley.
4. Kale, B.K. and Muralidharan, K. (2015). *Parametric Inference: An Introduction*, Alpha Science International Ltd.
5. Mukhopadhyay, P. (2015). *Mathematical Statistics*, Books and Allied (p) Ltd.
6. Dudewicz, E. J. and Mishra, S. N. (1988). *Modern Mathematical Statistics*, John Wiley and Sons.
7. Casella, G., and Berger, R. L. (2001). *Statistical Inference*, 2nd edition, Duxbury press.
8. Bansal, A. K. (2007). Bayesian parametric inference. Alpha Science International Limited.
9. DasGupta, A. (2008). Asymptotic theory of statistics and probability. Springer Science & Business Media.

CC-105: STATISTICAL COMPUTING

Unit 1: MSEXCEL: Introduction to MSEXCEL. Cell formatting, conditional formatting, Data manipulation using EXCEL: sort and filter, find and replace, text to columns, remove duplicate, data validation, consolidate, what-if-analysis. Working with Multiple Worksheets and Workbooks. Built-in mathematical and statistical functions for obtaining descriptive statistic, computing PMF/PDF, CDF and quantiles of the well-known distributions, rand and randbetween function, Logical functions: if, and, or, not. Lookup functions: hlookup, vlookup, Formula Errors, Creating and Working with Charts, Database functions, Text functions, Date and time functions, Excel add-ins: analysis tool pack, Pivot tables and charts.

(12L+3T)

Unit 2: R-software: Introduction to R, data types and objects, operators, data input, data import and export, built in functions for descriptive statistics, random sampling and computation of pdf, cdf and quantiles of well known distribution. Strings and Dates in R. apply family of functions. Saving work in R. Matrix algebra, graphical procedures, frequencies and cross tabulation, built in functions: lm, t.test, prop.test, wilcox.test, ks.test, var.test, chisq.test, aov. Control statements. Programming, user defined functions, R-packages. R-studio.

(12L+3T)

Unit 3: Concept of simulation. Concept of random number generator, true random number and pseudo random number generators, requisites of a good random number generator. Tests for randomness. Congruential method of generating uniform random numbers. Algorithms for generating random numbers from well-known univariate discrete and continuous distributions, generating random vectors from multinomial, bivariate normal, and bivariate exponential distributions, generating random numbers from mixture of distributions (related results without proofs). Acceptance-Rejection Technique. Use of random numbers to evaluate integrals, to study the systems involving random variables, to estimate event probabilities and to find expected value of random variables. Use of random numbers in statistical inference.

(12L+3T)

Unit 4: Resampling techniques: Bootstrap methods, estimation of bias and standard errors, estimation of sampling distribution, confidence intervals. Jackknife method: estimation of bias and standard errors, bias reduction method. Solution to system of linear equations: Jacobi and Gauss-Seidel methods with convergence analysis. Finding roots of nonlinear equation: Newton-Raphson method, bisection method; Newton-Raphson for system of non-linear equations. Numerical integration: quadrature formula, trapezoidal rule and Simpson's rules for single integral.

(12L+3T)

References:

1. Atkinson K. E. (1989): An Introduction to Numerical Analysis. (Wiley).
2. Devroye L. (1986) : Non- Uniform Random Variate Generation. (Springer- Verlag New York).
3. Efron B. and Tibshirani. R. J. (1994): An Introduction to the Bootstrap. (Chapman and Hall).
4. Morgan B. J. T.(1984) : Elements of Simulation. (Chapman and Hall).
5. Robert C. P. and Casella G. (1999): Monte Carlo Statistical Methods. (Springer-verlag New York, Inc.).
6. Ross. S. M. (2006): Simulation. (Academic Press Inc).
7. Rubinstein, R. Y. (1998) Modern Simulation and Modeling. (Wiley Series in Probability and Statistics).
8. William J., Kennedy, James E. Gentle. (1980): Statistical Computing. (Marcel Dekker).

CCPR-106: PRACTICAL –I

1. Gram-Schmidt orthogonalization method
2. Computation of Inverse and G-inverse of a matrix
3. Applications of Cayley-Hamilton theorem
4. Characteristics roots and vectors and their applications
5. Classifications and reduction of quadratic forms
6. Sketching of CDF and computation of probabilities using CDF
7. Finding best possible probability distribution to observed data sets and allied inferences.
8. Probability plots for various univariate probability distributions and their interpretations
9. Model sampling from some univariate and bivariate discrete, continuous and mixture distribution
10. Sufficient, minimal sufficient, and complete sufficient statistics
11. UMVUE and lower bunds for variances of unbiased estimators
12. MME, MLE, and U-statistic
13. Methods of Scoring and method of minimum chi-square estimation
14. Estimation using EM Algorithm
15. Practical on MSEXCEL
16. Practical on R- Software
17. Random Number Generation from Discrete and Continuous distribution and testing uniformity of the generated data.
18. Applications of Simulation
19. Numerical Methods and Resampling Techniques

(Each practical should consist of problems to be solved using at least two of the following software: EXCEL/ R)

Semester II

CC-201: PROBABILITY THEORY

Unit 1: Classes of sets: Sequence of sets: limsup, liminf and limit of sequence of sets, field, σ-field, monotone field, their properties, minimal σ-field generated by a class of sets, Borel σ-field, various classes of sets which generate the Borel σ-field, measure, measurable space, finite measure, Lebesgue and Lebesgue - Steltjes measures on $\mathcal{R}$. Probability measure, Probability space, properties of a probability measure, continuity, conditional probability measure, mixture of probability measures. Computation of probabilities of arbitrary events using distribution function. Independence of events, class of independent events, independence of classes, independence of random variables, equivalent definitions of independence.

(12L+3T)

Unit 2: Measurable function, random variable, simple random variable, elementary random variable, liminf, limsup and limit of sequence of random variables. Method of obtaining a random variable as a limit of sequence of simple random variables. Integration of a measurable function with respect to a measure, expectation of simple random variable, non-negative random variable, arbitrary random variable, properties of expectation, expectation of the product of independent random variables.

(12L+3T)

Unit 3: Monotone convergence theorem, Fatous Lemma, Dominated Convergence theorem, Borel - Cantelli Lemma, (Statements only), and their applications. Convergence of sequence of random variables, Convergence in distribution, almost sure convergence and its characterizing property, convergence in probability, uniqueness of limit, Yule-Slutsky results and preservation under continuous transform. convergence in r^{th} mean, interrelationships (Statements only), their illustration with examples.

(12L+3T)

Unit 4: Characteristic function, simple properties. Inversion theorem, uniqueness property (Statement only) and their applications. Weak and Strong laws of large numbers, Kolmogorov's three series theorem for almost sure convergence (Statement only), Liaponove's, Lindeberg-Feller Theorems on CLT (Statement only). Applications of the above results.

(12L+3T)

References:

1. Bhat B. R.(1981) : Modern Probability Theory –IIIrd edition :New age international limited.
2. Alan Karr,(1993) : Probability Theory – Springer Verlag.
3. Billingsley P. (1986) : Probability & Measure –John Wiley and sons .
4. Athreya K. B. and Lahiri S. (2006). Probability Theory, (Hindustan Book Agency).
5. Feller, W. (1969). Introduction to Probability and its Applications vol. II (Wiley Eastern Ltd.) .
6. Loeve, M. (1978). Probability Theory (Springer Verlag). Fourth edition.
7. Rohatgi, V.K. and Saleh, A. K. MD. E. (2015). Introduction to Probability Theory and Mathematical Statistics -3rd Edition, John Wiley & sons.

CC-202: THEORY OF TESTING OF HYPOTHESES

Unit 1: Problem of testing of Hypothesis, null and alternative hypotheses, Simple and composite hypotheses, power function of a test, Randomized and non-randomized tests, Most powerful (MP) test, Neyman-Pearson Lemma, Monotone likelihood ratio property, Uniformly Most Powerful (UMP) test and its existence. Determination of minimum sample size to achieve the desired strength of a test. p-value of a test.

(12L+3T)

Unit 2: UMP tests for two sided alternatives, examples of their existence and non-existence. Unbiased test, Generalized Neyman Pearson lemma, UMPU tests and their existence in case of exponential families (Statements of the theorems only). Similar tests, test with Neyman structure.

(12L+3T)

Unit 3: Likelihood ratio test and its applications. Chi-square tests for goodness of fit and independence of attributes. Confidence interval, relation with testing of hypotheses problem, Uniformly Most Accurate (UMA) and Uniformly Most Accurate Unbiased (UMAU) confidence intervals, shortest length confidence intervals, Bayesian confidence intervals.

(12L+3T)

Unit 4: Non-parametric tests, one sample tests: Sign test, Median test, Wilcoxon Signed-Rank test, Kolmogorov Smirnov test. Two sample tests: Wald-Wolfowitz Runs test, Mann-Whiteny U test, Median test, Kolmogorov Smirnov test. Kruskal-Wallis Test. Spearman's Rank Correlation Test; Kendall's Rank Correlation Test; Tests for randomness.

(12L+3T)

References:

1. Rohatgi, V.K. and Saleh, A. K. MD. E. (2015).Introduction to Probability Theory and Mathematical Statistics -3rd Edition, John Wiley & sons.
2. Kale, B. K. and Muralidharan, K. (2015). Parametric Inference: An Introduction, Alpha Science International Ltd.
3. Dudewicz, E. J. and Mishra, S. N. (1988). Modern Mathematical Statistics, John Wiley and Sons.
4. Lehman, E. L. (1987). Theory of testing of hypotheses. Students Edition.
5. Ferguson, T. S. (1967). Mathematical Statistics: A decision theoretical approach. Academic Press.
6. Zacks, S. (1971) .Theory of Statistical Inference, John Wiley and Sons, NewYork.
7. Randles, R. H. and Wolfe, D. A. (1979).Introduction to theory of nonparametric Statistics, Wiley.
8. Gibbons J. D. and Chakraborty S.(2010) Nonparametric Statistical Inference, Fifth Edition, CRC Press.

CC-203: LINEAR MODELS AND REGRESSION ANALYSIS

Unit 1: General linear model: definition, assumptions, concept of estimability, least squares estimation, BLUE, estimation space, error space, Gauss Markov theorem, variances and covariances of BLUEs, Distribution of quadratic forms for normal variables: related theorems (without proof), Tests of hypotheses in general linear models. Description of the ANOVA and linear regression models as the particular cases of the general linear model.

(12L+3T)

Unit-2: Multiple regression model, Least squares estimate (LSE), Properties of LSE, Hypothesis testing, confidence and prediction intervals, General linear hypothesis testing. Dummy variables and their use in regression analysis. Model adequacy checking. Transformations to correct model inadequacies: VST and Box-Cox power transformation.

(12L+3T)

Unit-3: Multicollinearity: Consequences, detection and remedies, ridge regression. Autocorrelation: sources, consequences, detection (Durbin-Watson test) and remedies. Parameter estimation using Cochrane-Orcutt method. Variable Selection Procedures: R-square, adjusted R-square, Mallows' Cp, forward, backward and stepwise selection methods, AIC, BIC.

(12L+3T)

Unit-4: Robust Regression: need for robust regression, M-estimators, properties of robust estimators: breakdown and efficiency. Asymptotic distribution of M-estimator (Statement only). Nonlinear Regression Models: nonlinear least squares, transformation to a linear model, parameter estimation in a nonlinear system, linearization. Polynomial regression model, piecewise polynomial fitting. Introduction to quantile regression.

(12L+3T)

References:

1. Draper N.R. and Smith, H. (1998): Applied Regression Analysis. 3rd ed Wiley
2. Weisberg, S. (1985): Applied Linear Regression, Wiley.
3. Kutner, Neter, Nachtsheim and Wasserman (2003): Applied Linear Regression Models, 4th Edition, McGraw-Hill .
4. Montgomery, D.C., Peck, E. A., and Vining, G.(2012): Introduction to Linear Regression Analysis, 5th Ed . Wiley.
5. Cook R.D. & Weisberg.(1982): Residuals and Influence in Regression. Chapman and Hall.
6. Birkes, D and Dodge, Y. (1993). Alternative methods of regression, John Wiley & Sons.
7. Huber, P. J. and Ronchetti, E. M (2011) Robust Statistics, Wiley, 2nd Edition.
8. Seber, G. A., Wild, C. J. (2003). Non linear Regression, Wiley.

CC-204: DESIGN AND ANALYSIS OF EXPERIMENTS

Unit 1: Concept of design of experiments (DOE), applications of DOE; Basic principles of DOE; Analysis of completely randomized design using the fixed effect model and estimation of the model parameters; Contrasts, orthogonal contrasts, Scheffé's method for comparing contrasts; Comparing pairs of treatment means: Tukey's test, Fisher least significant difference method; Comparing treatment means with a control; Analyses of randomized complete block design, Latin square design, balanced incomplete block design using fixed effect models and estimation of the model parameters.

(12L + 3T)

Unit 2: Concepts of factorial designs, main effects, and interaction effects; The two-factor factorial design and its analysis using fixed effect model; The general factorial design; Analysis of replicated and unreplicated 2^k full factorial designs; Blocking and confounding in a 2^k factorial design. Construction and analysis of 2^{k-p} fractional factorial designs and their alias structures; Design resolution, resolution III and resolution IV designs; fold over designs; saturated designs.

(12L + 3T)

Unit 3: The 3^k full factorial design and its analysis using fixed effect model; Confounding in 3^k factorial designs; Construction and analysis of 3^{k-p} fractional factorial designs and their alias structures; Concept of random effects and mixed effects models, analysis of 2^k factorial designs using the random effect model, analysis of 2^k factorial designs using the mixed effect model, rules for expected mean squares, approximate F-tests.

(12L + 3T)

Unit 4: Response surface methodology: the method of steepest ascent, analysis of the response surface using first and second order models, characterizing the response surface, ridge systems, multiple responses, designs for fitting response surfaces: simplex design, central composite design (CCD), spherical CCD,; Robust parameter design: crossed array designs and their analyses, combined array designs and the response model approach; The concepts of nested and split-plot designs.

(12L + 3T)

References:

1. Montgomery D.C. (2017): *Design and Analysis of Experiments*, 9th edition, John Wiley & Sons, Inc.
2. Phadke, M. S.(1989). *Quality Engineering using Robust Design*, Prentice-Hall.
3. Voss, D., Dean, A., and Dean, A.(1999). *Design and Analysis of Experiments*, Springer verlag Gmbh.
4. Wu, C. F., Hamada M. S.(2000). *Experiments: Planning, Analysis and Parameter Design Optimization, 2nd edition*, John Wiley & Sons.

CC-205: SAMPLING THEORY AND OFFICIAL STATISTICS

Unit 1: Review of basic methods: probability and non-probability sampling, simple random sampling and stratified random sampling, Systematic sampling, Cluster sampling, multistage-sampling. Ratio and regression estimators with their properties. Double sampling procedures and its estimation.

(12L + 3T)

Unit 2: Varying probability sampling: PPS sampling, Cumulative total method, Lahiri's method, Hansen-Horwitz estimator and its properties. Horwitz- Thompson, Des Raj estimators for a general sample size and Murthy's estimator for a sample of size 2 and its properties. Midzuno sampling, Rao-Hartley-Cochran sampling Strategy.

(12L + 3T)

Unit 3: Non-sampling errors: Response and non- response errors. Hansen–Horwitz and Demings model for the effect of call-backs. Randomised response techniques, dichotomous population, Warners model, MLE in Warners model, unrelated question model.

(12L + 3T)

Unit 4: Official Statistics: Elements of Indian Official Statistics, National Accounts – different approaches, Indices for Development, GDP, HDI, Evaluation & Monitoring. Poverty measurement: Different issues related to poverty, Measures of incidence and intensity.

(12L + 3T)

References:

1. Parimal Mukhopadhyay (2008): Theory and methods of survey sampling – 2ndEdition, Prentice Hall of India private limited.
2. Sukhatme P. V., Sukhatme S. & Ashok C (1984): Sampling Theory of surveys and applications – Iowa university press and Indian society of agricultural statistics, New Delhi.
3. Chaudhuri and H. Stenger (2005): Survey Sampling: Theory and Methods, 2nd edition, chapman and hall/CRC.
4. Des Raj and Chandhok. P. (1998): Sample Survey Theory - Narosa publication.
5. William G. Cochran. (2008): Sampling Techniques- IIIrd edition –John and Wieley sons Inc.
6. Singh, D. and Chaudhary F.S (1986).Theory and Analysis of Sample Survey Designs, Wiley Eastern Limited.
7. CSO. National Accounts Statistics- Sources and Health.
8. Sen, A. (1997). Poverty and Inequality.
9. Datt R., Sundharam, K. P. M. (2016) Indian Economy, (Sultan Chand & company Ltd.).

CCPR-206: PRACTICAL –II

1. MP, UMP, and UMPU Tests
2. Likelihood ratio tests
3. Confidence Intervals
4. Non-parametric Tests
5. Linear Estimation: Estimation and Hypothesis testing
6. Multiple linear regression
7. Variable selection
8. Multicollinearity and Autocorrelation
9. Robust regression and nonlinear regression
10. Analysis of CRD, RBD, LSD, and BIBD
11. Analysis of full, confounded and fractional 2^k factorial designs
12. Analysis of full, confounded and fractional 3^k factorial designs
13. Response surface methodology and robust parameter designs
14. Simple random sampling, stratified sampling and Systematic Sampling
15. Cluster and two stage sampling
16. Application of Ratio and Regression method of estimation
17. Probability proportional to size sampling

(Each practical should consist of problems to be solved using at least two of the following software: EXCEL/ R/ MINITAB)