SHIVAJI UNIVERSITY, KOLHAPUR

Established: 1962

A++ Accredited by NAAC (2021) with CGPA 3.52

Structure and Syllabus in Accordance with

National Education Policy - 2020

with Multiple Entry and Multiple Exit

of

Master of Science (Statistics)

under
Faculty of Science and Technology

(To Be Implemented From Academic Year 2023-24)

INDEX

1. Preamble

The M.Sc. (Statistics) programme at Shivaji University, Kolhapur, is a comprehensive postgraduate programme designed to equip students with a profound understanding of statistical theories, methodologies, and applications. This programme fosters a strong foundation in probability, data analysis, and inferential techniques, preparing students to excel in diverse fields such as research, industry, and academia. Through a balanced blend of theoretical insights and practical exposure, students develop skills in data manipulation, modeling, and interpretation. The curriculum promotes critical thinking, problem-solving, and advanced statistical software usage. By offering a rigorous academic environment, the programme cultivates statisticians capable of contributing effectively to data-driven decision-making processes across various sectors.

2. Duration

- 2 Years

3. Eligibility for Admission

- Three-year B. Sc. Degree with Statistics as principal / major subject

4. Medium of Instruction

- English

5. Programme Structure

Structure in Accordance with National Education Policy - 2020
With Multiple Entry and Multiple Exit Options
M.Sc. (Statistics) Part – I (Level-6.0)

	Course Code	Teaching Scheme			Examination Scheme					
		Theory and Practical			University Assessment (UA)			Internal Assessment (IA)		
		Lectures + Tutorial/ (Hours/ week)	Practical (Hours/ week)	Credit	Maximum Marks	Minimum Marks	Exam. Hours	Maximum Marks	Minimum Marks	Exam. Hours
Semester-I										
Major Mandatory	MMT11	4	--	4	80	32	3	20	8	0.5
	MMT12	4	--	4	80	32	3	20	8	0.5
	MMT13	2	--	2	40	16	1.5	10	4	0.25
	MMP14	--	12	4	--	--	--	100	40	3
Major Elective	MET1	4	--	4	80	32	3	20	8	0.5
Research Methodology	RM	4	--	4	80	32	3	20	8	0.5
Total				22	360			190		
Semester-II										
Major Mandatory	MMT21	4	--	4	80	32	3	20	8	0.5
	MMT22	4	--	4	80	32	3	20	8	0.5
	MMT23	2	-	2	40	16	1.5	10	4	0.25
	MMP24	--	12	4	--	--	--	100	40	3
Major Elective	MET2	4	--	4	80	32	3	20	8	0.5
OJT/FP	OJT/FP	4	--	4	*					
Total				22						
Total (Sem I + Sem II)				44						

• MMT–Major Mandatory Theory • MMP–Major Mandatory Practical • MET–Major Elective Theory • RM - Research Methodology • OJT/FP- On Job Training/ Field Project	• Total Marks for M.Sc.-I : 1100 • Total Credits for M.Sc.-I (Semester I & II) : 44 • Separate passing is mandatory for University and Internal Examinations
*Evaluation scheme for OJT/FP shall be decided by concerned BOS	
• Requirement for Entry at Level 6.0: B. Sc. degree of 3 years with Statistics as Principle/Major Subject.	
• Requirement for Exit after Level 6.0: Student can exit after completion of Level 6.0 with Post Graduate Diploma in Statistics	
• Requirement for Entry at Level 6.5: Completion of Level 6.0 with Statistics as Principle/Major Subject.	

Structure in Accordance with National Education Policy - 2020
With Multiple Entry and Multiple Exit Options
M.Sc. (Statistics) Part – II (Level-6.5)

	Course Code	Teaching Scheme			Examination Scheme					
		Theory and Practical			University Assessment (UA)			Internal Assessment (IA)		
		Lectures + Tutorial/ (Hours/ week)	Practical (Hours/ week)	Credit	Maximum Marks	Minimum Marks	Exam. Hours	Maximum Marks	Minimum Marks	Exam. Hours
Semester-III										
Major Mandatory	MMT31	4	--	4	80	32	3	20	8	0.5
	MMT32	4	--	4	80	32	3	20	8	0.5
	MMT33	2	-	2	40	16	1.5	10	4	0.25
	MMP34	--	12	4	--	--	--	100	40	3
Major Elective	MET3	4	--	4	80	32	3	20	8	0.5
Research Project	RP			4	#					
Total				22						
Semester-IV										
Major Mandatory	MMT41	4	--	4	80	32	3	20	8	0.5
	MMT42	4	--	4	80	32	3	20	8	0.5
	MMP43	--	12	4	--	--	--	100	40	3
Major Elective	MET4	4	--	4	80	32	3	20	8	0.5
Research Project	RP			6	##					
Total				22						
Total (Sem I + Sem II)				44						

• MMT–Major Mandatory Theory • MMP–Major Mandatory Practical • MET–Major Elective • RP- Research Project	• Total Marks for M.Sc.-II : 1100 • Total Credits for M.Sc.-II (Semester III & IV) : 44 • Separate passing is mandatory for University and Internal Examinations
# Evaluation scheme for Research Project shall be decided by concerned BOS	
## Evaluation scheme for Research Project shall be decided by concerned BOS	

6. Programme Outcomes (POs) and Programme Specific Outcomes (PSOs)

- **Programme Outcomes (POs)**

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 fromavailable resources.
4) Develop expertise in data management and analysis using widely used statistical software.

- **Programme Specific Outcomes (PSOs)**

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, Python and 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.

7. Course Codes

Semester-I		
Sr. No.	**Major Mandatory**	
1	Distribution Theory	MSU0325MML937G1
2	Estimation Theory	MSU0325MML937G2
3	Statistical Computing	MSU0325MML937G3
4	Practical I	MSU0325MMP937G4
5	Research Methodology in Statistics	MSU0325RML937G
	Major Elective	
6	Statistical Mathematics	MSU0325MEL937G1
	Real Analysis	MSU0325MEL937G2
	Linear Algebra	MSU0325MEL937G3
Semester-II		
Sr. No.	**Major Mandatory**	
1	Theory of Testing of Hypothesis	MSU0325MML937H1
2	Linear Models and Regression Analysis	MSU0325MML937H2
3	Statistical Programming using Python	MSU0325MML937H3
4	Practical-II	MSU0325MMP937H4
5	On Job Training/Field Project	MSU0325OJP937H/ MSU0325FPP937H
	Major Elective	
6	Probability Theory	MSU0325MEL937H1
	Bayesian Inference	MSU0325MEL937H2
	Reliability Theory	MSU0325MEL937H3
	Reliability Theory	MSU0325MEL937H4

8. Syllabus

M. Sc. Statistics (Part I) (Level-6.0) (Semester I)
(NEP-2020)
(Introduced from Academic Year 2023-24)

Title of Course: DISTRIBUTION THEORY
Course Code: MSU0325MML937G1
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the students will be able to:

i) Understand and explain the concept of univatiate and multivariate random variables and related entities
ii) Understand and explain the nature of various probability distributions and perform related computations.
iii) Understand probability models for multivariate data and perform related computations
iv) Understand non-central sampling distribution and able to perform their applications, able to perform computations related to order statistics

Unit 1: Random variable, Cumulative distribution function (CDF) and its properties, 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. Transformations of univariate random variables, probability integral transformation. (12L+3T)

Unit 2: Concepts of location, scale and shape parameters of distributions with examples. Symmetric distributions and their properties. Moment inequalities: Basic, Holder, Markov, Minkowski, Jensen, Chebyshev's inequalities, and their applications. Random vectors, joint distributions, Independence, variance-covariance matrix, joint MGF. Conditional expectation and variances. Transformations of bivariate random variables, Convolutions, compound distributions. (12L+3T)

Unit 3: Multivariate normal distribution, two definitions and their equivalence, singular and nonsingular normal distribution, characteristic function, moments, marginal and conditional distributions. Maximum likelihood estimators of the parameters of the multivariate normal distribution and their sampling distributions. Wishart matrix and its distribution (statement only), properties of Wishart distribution, distribution of generalized variance(statement only). Marshall-Olkin bivariate exponential distributions. (12L+3T)

Unit 4: Non-central chi-square, t and F distributions, distributions of linear and quadratic forms involving normal random variables, Fisher Cochran and related theorems (statements only) and their 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. Casella, G., & Berger, R. L. (2021). Statistical inference. Cengage Learning.
2. DasGupta, A. (2010). Fundamentals of probability: A first course. Springer Science & Business Media.
3. Johnson N. L. & Kotz. S. (1996). Distributions in Statistics Vol-I, II and III. JohnWiley and Sons New York.
4. Kotz, S., Balakrishnan, N., & Johnson, N. L. (2004). Continuous multivariate distributions, Volume 1: Models and applications (Vol. 1). John Wiley & Sons.
5. Rao C. R. (1995). Linear Statistical Inference and Its Applications. John Wiley & Sons.
6. Rohatagi V. K. & Saleh A. K. Md. E.(2001). Introduction to Probability Theory andMathematical Statistics. John Wiley and sons Inc.

Title of Course: ESTIMATION THEORY
Course Code: MSU0325MML937G2
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the students will be able to:

i) Explain the principles of data reduction and obtain sufficient, minimal sufficient, and complete statistics for various families of distributions.
ii) Obtain UMVUE of parameters of various distributions and determine Cramer-Rao and Chapman-Robbins-Kiefer lower bounds for the variances of unbiased estimators.
iii) Apply parametric and nonparametric methods to obtain estimators.
iv) Obtain CAN and BAN estimators.

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; 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: Consistency of an estimator, weak and strong consistency, joint and marginal consistency, invariance property under continuous transformations, methods of constructing consistent estimators, Comparison of consistent estimators, asymptotic relative efficiency, minimum sample size required by the estimator to attain certain level of accuracy. Consistent Asymptotic Normal (CAN) Estimators: Definition of CAN estimator for real and vector valued parameters, invariance of CAN property under non-vanishing differentiable transformation (delta method). Methods of constructing CAN estimators: Method of Moments, method of percentiles, comparison of CAN estimators. BAN estimators, CAN and BAN estimators in one parameter and multi-parameter exponential family of distributions, Crammer family of distributions, Cramer – Huzurbazar theorem (Statement only).

(12L + 3T)

References:

1. Casella, G., & Berger, R. L. (2021). Statistical inference. Cengage Learning.
2. Deshmukh S., Kulkarni M. (2022). Asymptotic Statistical Inference: A Basic Course Using R. Springer Verlag, Singapor.
3. Dudewicz, E. J., & Mishra, S. (1988). Modern mathematical statistics. John Wiley & Sons, Inc.
4. Kale, B. K., & Muralidharan, K. (2015). Parametric inference: An introduction. Alpha Science International Limited.
5. Lehmann, E. L. (1983). Theory of Point Estimation. John Wiley & sons.
6. Mukhopadhyay, P. (2015). Mathematical Statistics, Books and Allied (p) Ltd.
7. Rao C. R. (1995). Linear Statistical Inference and Its Applications. John Wiley & Sons.
8. Rohatgi, V. K., & Saleh, A. M. E. (2015). An introduction to probability and statistics. John Wiley & Sons.

Title of Course: STATISTICAL COMPUTING
Course Code: MSU0325MML937G3
Total Credits: 02
Course outcomes:
Upon successful completion of this course, the students will be able to:

i) Perform data organization, data manipulation, statistical and mathematical computations, and data analysis using MSEXCEL.
ii) Perform data organization, data manipulation, statistical and mathematical computations, and data analysis using R.

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)

References:

1. Gardener, M. (2012). Beginning R: the statistical programming language. John Wiley & Sons.
2. Held, B., Moriarty, B., & Richardson, T. (2019). Microsoft Excel Functions and Formulas with Excel 2019/Office 365. Mercury Learning and Information.
3. Herkenhoff, L., & Fogli, J. (2013). Applied statistics for business and management using Microsoft Excel. New York: Springer.
4. Purohit, S. G., Gore, S. D., & Deshmukh, S. R. (2015). Statistics using R. Alpha Science International.
5. Thulin, M. (2021). Modern Statistics with R: From wrangling and exploring data to inference and predictive modelling. BoD-Books on Demand.
6. Weblinks:
   https://support.microsoft.com/en-us/excel
   https://cran.r-project.org/manuals.html

Title of Course: STATISTICAL MATHEMATICS
Course Code: MSU0325MEL937G1
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the students will be able to:

i) Explain the vector Space, its dimension, and linear dependence/independence of vectors, and able to perform relation operations
ii) Understand matrix theory and perform matrix operations
iii) Understand sequences and series of real numbers and their convergence
iv) Understand the concept of real valued function, continuity of functions, convergence of series of functions, integration of functions, and obtain optima of a function.

Unit 1: Vectors, linear dependence and independence of vectors, vector space, subspace, 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: Generalized inverse, Vector and Matrix differentiation, 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.

(12 L + 3 T)

Unit 3: 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 4: 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. integrability of functions, Fundamental theorem on calculus, mean value theorem, their applications in finding functional of probability distributions. Maxima, minima of functions of several variables. Constrained maxima, minima, Lagrange’s method, Taylor’s theorem (without proof), Multiple and Improper integrals, their applications in multivariate probability distributions. Theorem on differentiation under integral sign and Leibnitz rule (statements only) with applications.

(12 L + 3 T)

References:

1. Apostol (1985). Mathematical Analysis. Narosa Publishing House, T.M.
2. Bartle, R. G., & Sherbert, D. R. (2000). Introduction to real analysis (Vol. 2, p. 2). New York: Wiley.
3. Goldberg, R. R. (1970). Methods of real analysis. Oxford and IBH Publishing.
4. Hadely G. (1962). Linear Algebra. Narosa Publishing House.
5. Malik, S. C., & Arora, S. (1992). Mathematical analysis. New Age International.
6. Narayan, S., Raisinghania M. D. (2013). Elements of Real Analysis. S. Chand.
7. Rao C. R. (1995). Linear Statistical Inference and Its Applications. John Wiley & Sons
8. Rao, A. R., & Bhimasankaram, P. (2000). Linear algebra (Vol. 19). Springer.
9. Royden (1988). Principles of Real Analysis. Macmillian
10. Searl S. B.(2006). Matrix Algebra Useful for Statistics. Wiley.

Title of Course: REAL ANALYSIS
Course Code: MSU0325MEL937G2
Total Credits: 04
Course outcomes:
Upon successful completion of the requirements for this course, students will be able to:

i) Understand set theory, obtain the supremum and infimum of bounded sets, and limit point of a set
ii) Understand sequences and series, the convergence of sequence and series, and compute the limit of sequences.
iii) Examine continuity and uniform continuity of functions, sequences, and series of functions, and evaluate Riemann and Riemann-Steltjes Integrals.
iv) Differentiate vector and matrix valued functions, obtain optima of functions of several variables, determine the Taylor series expansion of functions, evaluate multiple integrals, and use in connection with multivariate distributions

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. Apostol (1985). Mathematical Analysis. Narosa Publishing House, T.M.
2. Bartle G. R. (1976). Element of Real Analysis. Wiley, 2nd edition.
3. Bartle, R. G., & Sherbert, D. R. (2000). Introduction to real analysis (Vol. 2, p. 2). New York: Wiley.
4. Goldberg, R. R. (1970). Methods of real analysis. Oxford and IBH Publishing.
5. Malik, S. C., & Arora, S. (1992). Mathematical analysis. New Age International.
6. Narayan, S., Raisinghania M. D. (2013). Elements of Real Analysis. Fourteenth Revised Edition, S. Chand.
7. Royden (1988). Principles of Real Analysis. Macmillian

Title of Course: LINEAR ALGEBRA
Course Code: MSU0325MEL937G3
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the student will be able to:

i) Understand the vector Space, its dimension, and linear dependence/independence of vectors.
ii) Understand the matrix theory and related computations.
iii) Solve systems of linear equations.
iv) Understand the quadratic form and its classification

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 ofsolution 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 (No. 04; HA29, G73 V. 1.).
2. Hadely G. (1962). Linear Algebra. Narosa Publishing House.
3. Harville D. (1997). Matrix Algebra From Statistics Perspective. Springer.
4. Rao C. R. (1995). Linear Statistical Inference and Its Applications. John Wiley & Sons
5. Schott, J. R. (2016). Matrix analysis for statistics. John Wiley & Sons.
6. Searl S. B.(2006). Matrix Algebra Useful for Statistics. Wiley.
7. Rao, A. R., & Bhimasankaram, P. (2000). Linear algebra (Vol. 19). Springer.

Title of Course: RESEARCH METHODOLOGY IN STATISTICS
Course Code: MSU0325RML937G
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the students will be able to:

i) Understand the concept of research, research process, and research ethics.
ii) Understand and apply various sampling methods for data collection and estimate the parameters.
iii) Understand the concept of simulation and able to simulate real life processes
iv) Estimate bias and standard error of an estimator using resampling techniques, apply, numerical methods to solve systems of linear equations, to obtain the roots of a nonlinear equation, and to solve definite integrals.

Unit 1: Meaning of research, objectives of research, motivation in research, types of research, research approaches, significance of research, research methods vs. methodology, research and Scientific method, research process, criteria of good research, defining research problem, research design, Research Ethics, publication of research, Plagiarism, Intellectual property rights, Patents and its filing procedures.

(12L+3T)

Unit 2: Sampling techniques: review of simple random sampling stratified random sampling, systematic random sampling, cluster sampling, two phase sampling, ratio and regression method of estimation. Probability proportional to size sampling: Cumulative total method, Lahiri's method, Hansen-Horwitz estimator and its properties, Horwitz-Thompson estimator, Des Raj estimators for a general sample size. Non-sampling errors, techniques for handling non-response: 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 3: Concept and need of simulation, 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 for performance evaluation of estimators and statistical tests.

(12L+3T)

Unit 4: Resampling methods: 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. Numerical methods for solution to system of linear equations: Jacobi and Gauss-Seidel methods with convergence analysis. Numerical methods for 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, John Wiley and Sons.
2. Chaudhuri, A., & Stenger, H. (2005). Survey sampling: theory and methods. CRC Press.
3. Cochran, W. G. (1977). Sampling techniques. John Wiley & Sons.
4. Devroye L. (1986). Non- Uniform Random Variate Generation. Springer- Verlag New York.

5. Efron, B., & Tibshirani, R. J. (1994). An introduction to the bootstrap. CRC press.
6. Kennedy, W. J., & Gentle, J. E. (2021). Statistical computing. Routledge.
7. Kothari, C. R. (2004). Research methodology: Methods and techniques. New Age International.
8. Morgan, B. J. (1984). Elements of simulation (Vol. 4). CRC Press.
9. Mukhopadhyay, P. (2008). Theory and methods of survey sampling. PHI Learning Pvt. Ltd..
10. Robert, C. P., Casella, G., & Casella, G. (1999). Monte Carlo statistical methods (Vol. 2). New York: Springer.
11. Ross, S. M. (2022). Simulation. Academic Press.
12. Rubinstein, R. Y., & Melamed, B. (1998). Modern simulation and modeling (Vol. 7). New York: Wiley.
13. Singh, D., & Chaudhary, F. S. (1986). Theory and analysis of sample survey designs. John Wiley & Sons.
14. 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.

Title of Course: PRACTICAL-I
Course Code: MSU0325MMP937G4
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the student will be able to:

1. Sketching of various distribution functions and finding possible probability distribution to observed Data.
2. Compute UMVUE, MME and MLE using various methods.
3. Perform mathematical/statistical computations, statistical data analysis using built-in functions in MS-EXCEL and R and develop programs for various tasks.
4. Perform/Solve various statistical problems though simulation, numerical and re-sampling techniques.

Practical List:

1. Sketching of PDF/PMF and CDF
2. Fitting probability distribution and related inference.
3. Probability plots for various univariate probability distributions and their interpretations.
4. Applications of multivariate normal distribution.
5. UMVUE, MME and MLE
6. Methods of Scoring and method of minimum chi-square estimation
7. Estimation using EM Algorithm
8. Verification of consistency and CAN property of estimators.
9. Computations using MSEXCEL
10. Statistical analysis using MSEXCEL
11. Computations using R
12. R-Programming
13. Formulations of research problem and its design
14. Sampling techniques - I
15. Sampling techniques - II
16. Applications of Simulation techniques
17. Numerical Methods and Resampling Techniques

18-20. Based on elective course

M. Sc. Statistics (Part I) (Level-6.0) (Semester II)
(NEP-2020)
(Introduced from Academic Year 2023-24)

Title of Course: THEORY OF TESTING OF HYPOTHESES
Course Code: MSU0325MML937H1
Total Credits: 04
Course outcomes:
Upon successful completion of this course, a student will be able to:

i) Understand the concept of testing of hypothesis, test statistic, critical regions, size and power of a test.
ii) Develop and apply MP test, UMP test, UMPU test, similar tests and a test with Neyman structure.
iii) Obtain and interpret different interval estimates of parameters.
iv) Develop and apply large sample tests.

Unit 1: Problem of testing of hypothesis, null and alternative hypotheses, Simple and composite hypotheses, test function, Randomized and non-randomized tests, power function of a test, Most powerful (MP) test, Neyman-PearsonLemma, 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. Concept of p-value.

(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. Tests for means of multivariate normal distributions based on Hotelling's T^2 Statistic.

(12L+3T)

Unit 3: Interval estimation: Confidence interval, relation with testing of hypotheses problem, Uniformly Most Accurate(UMA) and Uniformly Most Accurate Unbiased(UMAU) confidence intervals, shortest length confidence intervals, Asymptotic Confidence Intervals based on CAN estimators, Variance stabilizing transformations (VST), confidence interval based on VST, Asymptotic Confidence regions.

(12L+3T)

Unit 4: Likelihood ratio test (LRT) and its asymptotic distribution, Wald test, Rao's Score test, Pearson Chi-square test of goodness of fit, Bartlett's test for homogeneity of variances. Large sample tests based on VST. Consistent test, comparison of tests: asymptotic relative efficiency of tests (Pitman and Bahadur efficiency). Performance evaluation (based on simulation) of asymptotic tests and confidence intervals.

(12L+3T)

References:

1. Kale, B. K., &Muralidharan, K. (2015). Parametric inference: An introduction. Alpha Science International Limited.
2. Dudewicz, E. J. and Mishra, S. N. (1988). Modern Mathematical Statistics, John Wiley & Sons.
3. Ferguson, T. S. (2014). Mathematical statistics: A decision theoretic approach. Academic press.
4. Gibbons, J.D., & Chakraborti, S. (2010). Nonparametric Statistical Inference (5th ed.). Chapman and Hall/CRC.
5. Lehman, E. L. (1987). Theory of testing of hypotheses. Students Edition.
6. Randles, R. H., & Wolfe, D. A. (1979). Introduction to the theory of nonparametric statistics. John Wiley.
7. Rohatgi, V. K., & Saleh, A. M. E. (2015). An introduction to probability and statistics. John Wiley & Sons.
8. Zacks, S. (1971). Theory of Statistical Inference, John Wiley & Sons, New York.

Title of Course: LINEAR MODELS AND REGRESSION ANALYSIS
Course Code: MSU0325MML937H2
Total Credits: 04
Course outcomes:
Upon successful completion of this course, a student will be able to:

i) Understand the concept of general linear model and associated inferential procedures.
ii) Understand and develop multiple linear regression models
iii) Identify the problems in developing multiple linear regression models and apply remedies.
iv) Understand generalized linear models and apply them for analyzing real data.

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 linear regression model, Least squares estimates (LSE) of parameters, Properties of LSE, Hypothesis testing, confidence and prediction intervals, General linear hypothesis testing, Model adequacy checking, Dummy variables and their use in regression analysis. 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. Robust Regression: need for robust regression, M-estimators, properties of robust estimators: breakdown and efficiency. Asymptotic distribution of M-estimator (Statement only).

(12L+3T)

Unit 4: Generalized linear models: concept of generalized linear model, Link function, ML estimation, large sample tests about parameters, goodness of fit, analysis of deviance. Residual analysis, types of residuals: raw, Pearson, deviance, Anscombe, quantile; residual plots. Variable selection: AIC and BIC. Logistic regression: logit, probit and cloglog models for dichotomous data, ML estimation, Odds ratio and its interpretation, hypothesis tests about model parameters. Hosmer-Lemeshow test, multilevel logistic regression, Logistic regression for Nominal response. Poisson regression.

(12L+3T)

References:

1. Birkes, D., & Dodge, Y. (2011). Alternative methods of regression. John Wiley & Sons.
2. Cook, R. D., & Weisberg, S. (1982). Residuals and influence in regression. New York: Chapman and Hall.
3. Draper, N. R., & Smith, H. (1998). Applied regression analysis. John Wiley & Sons.
4. Huber,P.J.andRonchetti,E.M (2011)Robust Statistics,Wiley,2ndEdition.
5. Kutner, M. H., Nachtsheim, C. J., Neter, J., & Wasserman, W. (2004). Applied linear regression models. New York: McGraw-Hill/Irwin.
6. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). Introduction to linear regression analysis. 5th Ed. John Wiley & Sons.
7. Seber,G.A.,Wild,C.J.(2003).Non linear Regression,Wiley.
8. Weisberg, S. (1985). Applied Linear Regression, John Wiley & Sons. New York.

Title of Course: STATISTICAL PROGRAMMING USING PYTHON
Course Code: MSU0325MML937H3
Total Credits: 02
Course outcomes:
Upon successful completion of this course, a student will be able to:

i) Develop programs in Python.
ii) Perform data organization, data manipulation, statistical and mathematical computations, and data analysis using Python.

Unit 1: Introduction, installation, keywords, identifiers: variables, constants, literals; comments, Operators, statements and expressions, data types with methods: numbers, string, lists, tuple, dictionary, set; indexing and slicing of each data type, data type conversion, built-in functions, control statements and loops, list comprehensions, user defined functions, anonymous/lambda function, local and global variables, modules: math, stat, random;creating own modules.

(12L+3T)

Unit 2: Concept of library and its working, Data storage, manipulation, visualization and analysis using the libraries: Numpy, Pandas, Scipy, statsmodels, Matplotlip, Seaborn, Regular Expressions (RegEx), Ski-kit learn.

(12L+3T)

Reference:

1. Gowrishankar, S., &Veena, A. (2018). Introduction to Python programming. CRC Press.
2. Guttag, J. V. (2021). Introduction to Computation and Programming Using Python: With Application to Computational Modeling and Understanding Data. Mit Press.
3. Haslwanter, T. (2016). An Introduction to Statistics with Python. With Applications in the Life Sciences. Switzerland: Springer International Publishing.
4. Nelli, F. (2018). Python data analytics with Pandas, NumPy, and Matplotlib.
5. Unpingco, J. (2016). Python for probability, statistics, and machine learning (Vol. 1). Springer International Publishing.
6. VanderPlas, J. (2016). Python data science handbook: Essential tools for working with data. " O'Reilly Media, Inc.".
7. URLs:
   - https://scikit-learn.org/stable/
   - https://numpy.org/
   - https://scipy.org/
   - https://www.statsmodels.org/stable/index.html
   - https://matplotlib.org/
   - https://pandas.pydata.org/

Title of Course: PROBABILITY THEORY
Course Code: MSU0325MEL937H1
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the student will be able to:

i) Understand the concept of sets and events, field, measurable functions, measure space, random variable.
ii) Understand the concept of probability space, probability measure, distribution function, moments and expectation, convergence of random variables
iii) Understand the concept of independence of events and 0-1 laws.
iv) Understand and apply characteristic function, law of large numbers and central limit theorems.

Unit 1: Sets and classes of events: the event, algebra of sets, classes of events, limsup, liminf and limit of sequence of sets, Monotone Class, fields, sigma fields and properties. Minimal σ-field generated by a class of sets, Borel σ-field. Tail events and tail Sigma field. Measurable Functions: functions and inverse functions, measure, measurable space and measure space, finite measure, sigma finite measure. Random variable: Definition of random variable, Indicator function and its properties, Simple and elementary random variable, limits of random variables and properties. (12L+3T)

Unit 2: Probability space: definition of probability, its properties, discrete probability space, general probability space, induced probability space, conditional probability measure, Dirac Measure, Counting measure, Lebesgue and Lebesgue - Steltjes measures, linear combinations of probability measures. Distribution Function and its properties. Expectations and moments: Definition and properties of expectation, conditional expectation, moments, moment generating functions. Convergence of random variables: Types of convergence, Cauchy convergence, Convergences in probability, Criterion for Convergences in probability, Slutsky's theorems, Almost sure convergence, Convergence in distribution, Convergence in r^{th} mean, Interrelations (12L+3T)

Unit 3: Convergence Theorem for expectation: Monotone convergence theorem, Fatous Lemma, Dominated Convergence theorem. Independence: Independence of events, class of independent events, independence of classes, independence of random variables, expectation of product of independent random variables, equivalent definitions of independence, Zero-One laws: Kolmogorov 0-1 law, Borel 0-1 criterion, Borel-Cantelli Lemma
(12L+3T)

Unit 4: Characteristic function: definition and properties of characteristic function, inversion formula (without proof), characteristic function and moments. Laws of large numbers: Convergence of distribution functions, convergence of series of independent random variables, Kolmogorov inequalities and almost sure convergence, weak law of large numbers (without Proof) for iid and non-iid random variables, Strong law of large numbers (without Proof). Central limit Theorem(CLT) (without proof): Lindeberg-Levy, Liaponove's and Lindeberg-Feller forms and applications. (12L+3T)

References:

1. Alan Karr. (1993): Probability Theory- Springer Verlag.
2. Athreya, K. B., &Lahiri, S. N. (2006). Probability theory. Hindustan Book Agency.
3. Bhat, B. R. (2007). Modern probability theory. New Age International.
4. Billingsley, P. (1986). Probability and measure.,(John Wiley and Sons: New York). Billingsley
5. Loeve, M. (1978). Probability Theory (Springer Verlag). Fourth edition.
6. Rohatgi, V. K., & Saleh, A. M. E. (2015). An introduction to probability and statistics. John Wiley & Sons.

Title of Course: BAYESIAN INFERENCE
Course Code: MSU0325MEL937H2
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the student will be able to:

i) Describe the role of the prior and posterior distribution in Bayesian inference about a parameter and provide various prior distributions
ii) Carry out the statistical inference (point estimate, testing of hypothesis and interval estimate) of the parameter using Bayesian approach
iii) Perform Bayesian analysis with various prior distribution
iv) Generate random samples from distributions for which direct sampling is difficult and use iterative methods for estimation

Unit 1: Basic elements of Statistical Decision Problem. Expected loss, decision rules (non-randomized and randomized). Overview of Classical and Bayesian Estimation. Advantage of Bayesian inference, Prior distribution, Posterior distribution, Subjective probability and its uses for determination of prior distribution. Importance of non-informative priors, improper priors, invariant priors. Conjugate priors, construction of conjugate families using sufficient statistics, hierarchical priors. Admissible and minimax rules and Bayes rules.

(12L + 3T)

Unit 2: Point estimation, Concept of Loss functions, Bayes estimation under symmetric loss functions, Bayes credible intervals, highest posterior density intervals, testing of hypotheses. Comparison with classical procedures. Predictive inference. One- and two-sample predictive problems.

(12L + 3 T)

Unit 3: Bayesian analysis with subjective prior robustness and sensitivity, classes of priors, conjugate class different methods of construction of objective priors: Jeffrey’s prior, probability matching prior, conjugate priors and mixtures, posterior robustness: measures and techniques. Bayes factors large sample methods: Limit of posterior distribution, consistency of posterior distribution, asymptotic normality of posterior distribution.

(12L + 3 T)

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.

(12L + 3T)

References:

1. Bolstad, W. M. (2007). Introduction to Bayesian Statistics Second Edition. Amerika: A John Wiley & Sons.
2. Christensen, R., Johnson, W., Branscum, A., & Hanson, T. E. (2011). Bayesian ideas and data analysis: an introduction for scientists and statisticians. CRC press.
3. Congdon, P. (2007). Bayesian statistical modelling. John Wiley & Sons.
4. Ghosh, J. K., Delampady, M., & Samanta, T. (2006). An introduction to Bayesian analysis: theory and methods (Vol. 725). New York: Springer.
5. Albert, J. (Ed.). (2009). Bayesian computation with R. New York, NY: Springer New York.
6. Rao. C.R. and Day. D. (2006). Bayesian Thinking, Modeling & Computation, Handbook of Statistics, Vol. 25. Elsevier

Title of Course: PRACTICAL II
Course Code: MSU0325MMP937H4
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the student will be able to:

i) Test hypothesis about the parameters and provide interval estimates involved in random experiments based on random sample.
ii) Perform statistical analysis, such as estimation, hypothesis testing, and analysis of variance, under generalized linear models
iii) Fit linear regression model or Generalized Linear Regression Models to the data, perform diagnostic analysis and apply rectifying measures to overcome the problem of Multicollinearity, auto-correlation, outliers and non-linearity.
iv) Perform mathematical/statistical computations, statistical data analysis using built-in functions in Python and develop programs for various tasks.

Practical List:

1. MP, UMP, and UMPU Tests
2. Tests based on Hotelling's T^2 statistic
3. Confidence Intervals
4. Likelihood ratio tests
5. Linear Estimation: Estimation and Hypothesis testing
6. Performance evaluations of confidence intervals and tests through simulation.
7. Multiple linear regression
8. Variable selection in regression
9. Dealing with multicollinearity, autocorrelation and outliers
10. Logistic Regression
11. Poison regression
12. Monte Carlo Simulation of Regression Models
13. Python programming-I
14. Python programming-II
15. Python Programming-III

14-20. Based on elective paper

9. Scheme of Teaching

- Theory courses: One hour per week per credit
- Practical courses: Three hours per week per credit
- Project: Three hours per week per credit

10. Examination Pattern

- Theory: 20% marks for internal examination and 80% marks for university examination
- Practical: 100% marks for internal evaluation
- On Job Training: To be decided by BoS
- Field Project: To be decided by BoS
- Research Methodology: 20% marks for internal examination and 80% marks for university examination

11. Nature of Question Paper and Scheme of Marking

Theory Examination:

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

Practical Examination:

Component	Max marks
Practical examination: Examination will be of 3 hour duration. There shall be 8 questions each of 12 marks, of which a student has to attempt any 5 questions.	60
Day-to-day practical performance and journal	20
Viva: Viva will be based on all practical's	20

11. Equivalence of courses

M. Sc. Part I (Semester I and II)

Old Course				Equivalent Course		
Sem No.	Course Code	Title of Old Course	Credit	Course Code	Title of New Course	Credit
I	CC-101	Real Analysis	4	ME12	Real Analysis	4
I	CC-102	Linear Algebra	4	ME13	Linear Algebra	4
I	CC-103	Distribution Theory	4	MM11	Distribution Theory	4
I	CC-104	Estimation Theory	4	MM12	Estimation Theory	4
I	CC-105	Statistical Computing	4	RM	Research Methodology in Statistics	4
II	CC-201	Probability Theory	4	ME21	Probability Theory	4
II	CC-202	Theory of Testing of Hypothesis	4	MM21	Theory of Testing of Hypothesis	4
II	CC-203	Linear Models and Regression Analysis	4	MM22	Linear Models and Regression Analysis	4
II	CC-204	Design and Analysis of Experiments	4	ME31	Design and Analysis of Experiments	4
II	CC-205	Sampling Theory and Official Statistics	4	No equivalence can be given for these courses in the new syllabus as per NEP 2020.		