SHIVAJI UNIVERSITY, KOLHAPUR - 416004, MAHARASHTRA
PHONE:EPABX-2609000, www.unishivaji.ac.in, bos@unishivaji.ac.in
शिवाजी विद्यापीठ, कोल्हापूर -४१६००४,महाराष्ट्र
दूरध्वनी-ईपीएबीएक्स -२६०९०००, अभ्यासमंडळे विभाग दुरध्वनी ०२३१—२६०९०९४
०२३१—२६०९४८७
Estd. 1962
"A++" Accredited by NAAC (2021) With CGPA 3.52

SU/BOS/Science/473 **Date: 19/08/2024.**

To,

The Principal, All Concerned Affiliated Colleges/Institutions Shivaji University, Kolhapur	The Head/Co-ordinator/Director All Concerned Department (Science) Shivaji University, Kolhapur.

Subject: Regarding Minor Change syllabi of M.Sc. Part-I & II as per NEP-2020 (2.0) degree programme under the Faculty of Science and Technology.

Ref: SU/BOS/Science/482 Date: 01/07/2023 & 09 Date: 02/01/2024

Sir/Madam,

With reference to the subject mentioned above, I am directed to inform you that the university authorities have accepted and granted approval to the Minor Change in syllabi, nature of question paper and equivalence of M.Sc. Part-I & II as per NEP-2020 (2.0) degree programme under the Faculty of Science and Technology.

M.Sc. Part-I & II as per NEP-2020 (2.0)			
1.	Statistics	2.	Applied Statistics and Informatics

This syllabus, nature of question shall be implemented from the academic year 2024-2025 onwards. A soft copy containing the syllabus is attached herewith and it is also available on university website www.unishivaji.ac.in,NEP-2020@suk(Online Syllabus).

The question papers on the pre-revised syllabi of above-mentioned course will be set for the examinations to be held in October /November 2024 & March/April 2025. These chances are available for repeater students, if any.

You are, therefore, requested to bring this to the notice of all students and teachers concerned.

Thanking you,

Dy Registrar
Dr. S. M. Kubal

Copy to:

1	The Dean, Faculty of Science & Technology	4	P.G Admission / Eligibility Section
2	The Chairman, Respective Board of Studies	5	Computer Centre/ Eligibility Section
3	B.Sc. Exam/ Appointment Section	6	Affiliation Section (U.G.) (P.G.)

SHIVAJI UNIVERSITY, KOLHAPUR

Established: 1962

A^{++} Accredited by NAAC (2021) with CGPA 3.52

Structure and Syllabus in Accordance with

National Education Policy – 2020 (NEP 2.0)

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)

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

Courses

M.Sc. Semester-III			
Sr. No.	**Course Category**	**Course Title**	**Course Code**
1	Major Mandatory	Stochastic Processes	MSU0325MML937I1
2		Theory and Practice of Machine Learning	MSU0325MML937I2
3		Multivariate Analysis	MSU0325MML937I3
4		Practical-III	MSU0325MMP937I4
5	Major Elective*	Design and Analysis of Experiments	MSU0325MEL937I1
6		Econometrics	MSU0325MEL937I2
7		Functional Data Analysis	MSU0325MEL937I3
8	Research Project	Research Project	MSU0325RPP937I
M.Sc. Semester-IV			
Sr. No.	**Course Category**	**Course Title**	**Course Code**
1	Major Mandatory	Biostatistics	MSU0325MML937J1
2		Time Series Analysis	MSU0325MML937J2
3		Practical-IV	MSU0325MMP937J3
4	Major Elective*	Optimization Techniques	MSU0325MEL937J1
5		Statistical Quality Control	MSU0325MEL937J2
6		Spatial Data Analysis	MSU0325MEL937J3
7		Actuarial Statistics	MSU0325MEL937J4
8	Research Project	Research Project	MSU0325RPP937J

*Only one course under this category is to be chosen by the student.

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: 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.

M. Sc. Statistics (Part II) (Level-6.5) (Semester III) (NEP-2020)

Title of Course: Stochastic Processes
Course Code: MSU0325MML937I1
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the students will be able to:

1. Identify appropriate stochastic process model for a given real life process.
2. Specify a given discrete time Markov chain in terms of a transition probability matrix and a transition diagram, and calculate higher step transition probabilities and limiting probabilities.
3. Understand and analyses discrete state space continuous time Markov chains and their practical applications
4. Explore the Galton-Watson Binary Branching process and understand the fundamental elements of Queuing models

Unit 1: Definition of stochastic process, classification of stochastic processes according to state space and time domain, finite dimensional distributions. Examples of various stochastic processes. Definition of Markov chain. Examples of Markov chains, Formulation of Markov chain models, initial distribution, transition probability matrix, Chapman-Kolmogorov equations, calculation of n-step transition probabilities. Simulation of Markov Chain.

(12L +3T)

Unit 2: Classification of states, irreducible Markov chain, period of the state, random walk and gambler's ruin problem, first entrance theorem, first passage time distribution. Long-Run proportions and limiting probabilities, relation with mean recurrence time, stationary distribution.

(12L +3T)

Unit 3: Discrete state space continuous time Markov chain, Poisson process and related results. Birth and death processes and associated cases. Renewal and delayed renewal processes, related theorems, key renewal theorem (Without proof) and its application. Simulation of Poisson process and discrete state space Markov processes.

(12L +3T)

Unit 4: Galton-Watson Binaymi Branching process. Generating functions and its properties, moments. Probability of ultimate extinction. Distribution of population size and association results. Simulation of branching process. Basic elements of Queuing model. Steady state probabilities and various average characteristics for the models: M/M/1, M/M/1 with balking, M/M/c and M/G/1.

(12 L+ 3T)

References:

1. Bhat B. R. (2000). Stochastic Models: Analysis and Applications, (New Age International)
2. Cinlar E. (2013): Introduction to Stochastic Process. (Courier Corporation)
3. Feller W.(2008): An Introduction to Probability Theory and Its Applications. (Wiley)
4. Hoel P. G., Port S. C. and Stone C. J. (1987): Introduction to Stochastic Processes. (Waveland Press)
5. Karlin S. and Taylor H. M. (1968): A First Course in Stochastic Process. (Academic Press)
6. Medhi J. (2009): Stochastic Process, (New Age International Publications)
7. Ross S. (1996): Stochastic Processes. (Wiley)
8. Ross S. (2014): Introduction to Probability Models. (Academic Press)
9. Taylor H. M. and Karlin S. (2014): An Introduction to Stochastic Modeling (Academic Press)

Title of Course: Theory and Practice of Machine Learning
Course Code: MSU0325MML937I2
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the students will be able to:

1. Explain supervised learning and construct classifiers namely, decision tree, k-nearest neighbour(s), logistic regression, naïve Bayes, Bayesian Belief Network
2. Compare different classifiers and employ techniques to improve their performance.
3. Use Artificial Neural Network and Support Vector Machine for classification and prediction.
4. Apply clustering techniques and generate association rules using apriori algorithm.

Unit 1: Data understanding and data cleaning, concept of supervised and unsupervised learning. Problem of classification, classification techniques: k-nearest neighbor, decision tree, Naïve Bayesian, classification based on logistic regression, Bayesian belief Network.

(12L+3T)

Unit 2: Model evaluation and selection: Metrics for Evaluating Classifier Performance, Holdout Method and Random Subsampling, Cross-Validation, Bootstrap, Model Selection Using Statistical Tests of Significance, Comparing Classifiers Based on Cost–Benefit and ROC Curves. Techniques to Improve Classification Accuracy: Introduction to Ensemble Methods, Bagging, Boosting and AdaBoost, Random Forests, Improving Classification Accuracy of Class-Imbalanced Data.

(12L+3T)

Unit 3: ANN and SVM: Artificial Neutral Network (ANN): Introduction to ANN, types of activation functions: identity, sigmoid, double sigmoid, tanh, softmax, loss functions: squared error, cross entropy; optimizers: gradient decent, stochastic and minibatch gradient decent; McCulloch-Pitts AN model, single layer network, multilayer feed forward network model, training methods, ANN & regression models. Support vector machine: Introduction to support vector machine, loss functions, soft margin, optimization hyperplane, support vector classification, support vector regression, linear programming support vector machine for classification and regression.

(12L+3T)

Unit 4: Unsupervised learning: Clustering: k-mediods, CLARA, DENCLUE, DBSCAN, Probabilistic model based clustering. Market Basket Analysis: Association rules and prediction, Apriori Algorithm, data attributes, applications to electronic commerce.

(12L+3T)

References:

1. Berson and Smith S.J. (1997) : Data warehousing, Data Mining, and OLAP, McGraw-Hill.
2. Breiman J.H Friedman, R.A. Olshen and stone C.J. (1984) : Classification and Regression Trees, Wadsworth and Brooks / Cole.
3. Han, J. and Kamber, M. and Pei, J. (2012) : Data Mining: Concepts and Techniques. MorganGaufmann.3rd Edition.
4. Mitchell T.M. (1997) : Machine Learning , McGraw-Hill.
5. Ripley B.D. (1996) : Pattern Recognition and Neural Networks. Cambridge University Press.
6. Vapnik V.N. The nature of Statistical learning theory, Springer.
7. Cristianini N. and Shawe-Taylor J. An Introduction to support vectormachines.
8. Data set source: http://www.ICS.uci.edu/~mlearn/MLRepository.html
9. Mehrika, K., Mohan, C., and Ranka (1997) Elements of Artificial neural networks. Penram international.
10. Hastie T, Tibshirani R, Friedmant J, (2009): The elements of statistical Learning, Springer.
11. Chattamvelli, R. (2015). Data mining methods. Alpha Science International.

Title of Course: Multivariate Analysis
Course Code: MSU0325MML937I3
Total Credits: 02
Course outcomes:
Upon successful completion of this course, the students will be able to:

1. Perform exploratory multivariate data analysis, understand and apply discrimination and classification techniques and apply hierarchical and non-hierarchical clustering techniques
2. Understand and apply the Canonical Correlation Analysis, principal component analysis and factor analysis.

Unit 1: Exploratory multivariate data analysis, sample mean vector, sample dispersion matrix, correlation matrix, graphical representation, Partial and multiple correlation coefficients. Discrimination and classification: Fisher's discriminant function and likelihood ratio procedure, minimum ECM rule, Rao's U statistics and its use in tests associated with discriminant function, classification with three populations. Cluster analysis, Hierarchical methods: Single, Complete, average linkage method and non-hierarchical clustering method: k-means clustering.

(12L+3T)

Unit 2: Canonical correlation analysis: Introduction, canonical variates and canonical correlation, interpreting population canonical variables, sample canonical variates and sample canonical correlations; Principal component analysis: Introduction, Population principal components, summarizing sample variation by principal components, Graphing the principal components, large sample inferences; Factor analysis: Introduction, Orthogonal factor model, methods of estimation, factor rotation and factor score

(12L+3T)

References:

1. Kshirsagar A. M.(1972) : Multivariate Analysis. Marcel-Dekker.
2. Johnson, R.A. and Wichern . D.W (2002) : Applied multivariate Analysis. 5th Ad.Prentice – Hall.
3. Anderson T. W. (1984) : An introduction to Multivariate statistical Analysis 2nd Ed. John Wiely.
4. Morrison D.F. (1976) : Multivariate Statistical Methods McGraw-Hill.
5. Bhuyan K. C. (2005): Multivariate Analysis and its applications, New central book agency Ltd. Kolkatta.

Title of Course: Design and Analysis of Experiments
Course Code: MSU0325MEL937I1
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the students will be able to:

1. Explain the fundamentals of experimental design.
2. Design and analyze a general full factorial experiment with more emphasis on two-level and three-level full factorial experiments.
3. Design and analyze two-level and three-level fractional factorial experiments.
4. Implement advanced experimental techniques such as Response Surface Methodology and Robust Parameter Design.

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; 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; The general factorial design; The two-factor factorial design and its analysis using fixed effect model: main effects, and interaction effects, 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.

(12L+3T)

Unit 3: The 3^k full factorial design and its analysis using fixed effect model; Confounding in3^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)

Unit4: Response surface methodology: the method of steepest ascent, analysis of the response surface using first and second order models, characterizing the response surface, 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.

(12L+3T)

References:

1. Dean, A., & Voss, D. (Eds.). (1999). *Design and analysis of experiments*. New York, NY: Springer New York.
2. Montgomery D.C. (2017): *Design and Analysis of Experiments*, 9thedition, John Wiley & Sons,Inc.
3. Phadke, M.S.(1995). *Quality Engineering using Robust Design*, Prentice-Hall.
4. Wu, C.J., and Hamada, M.S.(2000).*Experiments: Planning, Analysis and Parameter Design Optimization, 2ndedition*, John Wiley & Sons.

M. Sc. Statistics (Part II) (Level-6.5) (Semester IV) (NEP-2020)

Title of Course: Biostatistics
Course Code: MSU0325MML937J1
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the students will be able to:

1. Explain the need, ethics, and various concepts in clinical trials.
2. Design and perform various phases of clinical trials.
3. Analyze Phase I-III bio-equivalence trials, case-control and cohort designs.
4. Explain the concept of censoring, various types of censoring, and perform inference about survival function.

Unit 1: Introduction to clinical trials: aim, need and ethics of clinical trials, conduct of clinical trials, preclinical research, phase I-IV trials, multi-center trials, bias and random error in clinical studies, randomization, blinding/masking in clinical trials, data management: data definitions, case report forms, database design, data collection systems for good clinical practice.

(12L+3T)

Unit 2: Design of clinical trials: parallel vs. cross-over designs, cross-sectional vs. longitudinal designs, factorial designs, objectives and endpoints of clinical trials, design of Phase I trials, design of single-stage and multi-stage Phase II trials, design, and monitoring of Phase III trials with sequential stopping.

(12L+3T)

Unit 3: Bioavailability, pharmacokinetics, and pharmaco-dynamics. Design of bio-equivalence trials, Decision rules for bioequivalence, Inference for 2x2 crossover design: Classical methods of interval hypothesis testing for bioequivalence, Bayesian methods, nonparametric methods, assessment of Inter-and Intra-subject variability, drug interaction studies, reporting and analysis: analysis of categorical outcomes from Phase I - III trials. Epidemiological studies: case-control and cohort study designs. Measures of disease occurrence and association.

(12L+3T)

Unit 4: Introduction to survival analysis, examples of survival data/time to event data, measurement of survival time, concept of censoring, various types of censoring, type-I, type-II, random censoring, likelihood constructions and ML estimation. Non parametric estimation of survival function: Actuarial Estimator, Kaplan Meier product limit estimator, non parametric estimates of the mean, median and percentiles of the survival times; Nelson Aalen estimator, non parametric tests for two-sample problem: Gehen test, Log rank test, Mantel Haenszel test. Cox proportional hazard regression model.

(12L + 2T)

References:

1. Chow, S. C., & Liu, J. P. (2008). *Design and analysis of clinical trials: concepts and methodologies* (Vol. 507). John Wiley & Sons.
2. Clayton, D., & Hills, M. (2013). *Statistical models in epidemiology*. OUP Oxford.
3. Collett, D. (2023). *Modelling survival data in medical research*. CRC press.
4. Daniel, W. W., & Cross, C. L. (2018). *Biostatistics: a foundation for analysis in the health sciences*. Wiley.
5. Deshpande, J. V., & Purohit, S. G. (2015). *Lifetime Data: Statistical Models and Methods* (Vol. 16). World Scientific Publishing Company.
6. Fleiss, J. L. (2011). *Design and analysis of clinical experiments*. John Wiley & Sons.
7. Friedman, L. M., Furberg, C. D., DeMets, D. L., Reboussin, D. M., & Granger, C. B. (2015). *Fundamentals of clinical trials*. springer.
8. Hosmer Jr, D. W., Lemeshow, S., & May, S. (2011). *Applied survival analysis: regression modeling of time-to-event data*. John Wiley & Sons.

9. Jennison, C., & Turnbull, B. W. (1999). *Group sequential methods with applications to clinical trials*. CRC Press.
10. Marubeni .E. and Valsecchi M. G. (1994). *Analyzing Survival Data from Clinical Trials and Observational Studies*. Wiley.
11. Miller R. G. (1981). *Survival Analysis*, McFraw-Hill, New York.
12. Piantadosi, S. (2017). *Clinical trials: a methodologic perspective*. John Wiley & Sons.

Title of Course: Time Series Analysis
Course Code: MSU0325MML937J2
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the students will be able to:

1. Understand the time series; understand the concept of stationarity to the analysis of time series data in various contexts (such as actuarial studies, climatology, economics, finance, geography, meteorology, political science, and sociology).
2. Identify and isolate non deterministic components of observed time series; learn to translate an observed non-stationary series to stationary time series using an appropriate transformation.
3. Model, estimate, interpret and forecast observed time series through ARMA, ARIMA and SARIMA approach and Perform residual analysis for checking model adequacy.
4. Learn basics of time dependent volatility in time series, basics of ARCH and GARCH time series heteroscedastic models and basic of multivariate time series and its modeling

Unit 1: Exploratory time series analysis, Exponential, Double exponential and Holt-Winter smoothing and forecasting, auto-covariance, auto-correlation functions and their properties and characterization (without proof), partial auto covariance function, auto-covariance generating function. Weak and strong stationary time series, white noise process, Linear Process, estimates of mean, auto-covariance, auto-correlation and partial auto-covariance functions.

(12 L + 3 T)

Unit 2: Wold representation of linear stationary processes, linear time series models: autoregressive(AR), moving average(MA), autoregressive moving average models (ARMA). causality and invertibility of ARMA processes, computation of π -weights and ψ- weights, computation of ACVF, ACF and PACF for AR(1), AR(2), MA(1), MA(2), ARMA(1,1) processes and general procedure for ARMA(p,q) process. The need for differencing a time series, autoregressive integrated moving average models(ARIMA).

(12 L + 3 T)

Unit 3: Estimation of ARMA models: Yule-Walker estimation for AR Processes, Maximum likelihood and least squares estimation for ARMA Processes, Residual analysis and diagnostic checking. Minimum mean squared error forecasting for ARMA and ARIMA models, updating forecasts. Introduction to SARIMA models, Spectral Representation of the ACVF, Spectral density of an ARMA process, its computation for simple models.

(12 L + 3 T)

Unit 4: Introduction to ARCH and GARCH models. Properties and estimation under ARCH(1) and GARCH(1,1) model. Vector time-series models: Covariance and Correlation Matrix functions, MA and AR representation of vector processes, Covariance matrix function of the vector AR(1) and MA(1) models.

(12L + 3T)

Reference:

1. Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
2. Brockwell, P. J., & Davis, R. A. (2009). *Time series: theory and methods*. Springer science & business media.
3. Chatfield, C. (2004) *The Analysis of Time Series - An Introduction*, Sixth edition, Chapman and Hall.
4. Kendall, M.G. (1978) *Time Series*, Charler Graffin
5. Tsay, R. S. (2005). *Analysis of financial time series*. John wiley & sons.
6. WEI, W. W. (2006). *Time Series Analysis: Univariate and Multivariate Methods*.

Title of Course: Optimization Techniques
Course Code: MSU0325MEL937J1
Total Credits: 04
Course outcomes:
Upon successful completion of this course, the students will be able to:

1. Formulate a problem as an appropriate optimization problem (LPP, IPP, QPP).
2. Apply various methods to obtain optimum solution of a LPP, IPP and QPP.
3. Solve two person zero sum games with pure and mixed strategies using various methods.
4. Explain, formulate and solve dynamic Programming problem.

Unit 1: Convex Sets and Functions: Convex sets, supporting and separating hyperplanes, convex polyhedra and polytope, extreme points, convex functions. Linear programming problem (LPP): Definition and applications, methods of solving LPP: Graphical method, Simplex method, theorems related to the development of simplex algorithm, theorems related to a basic feasible solution, reduction of a feasible solution to a basic feasible solution, improvement of a basic feasible solution, existence of unbounded solution, optimality conditions and other related theorems (statements only), Examples. Artificial variable technique: Two phase method, Big M method, degeneracy.

(12L+3T)

Unit 2: Concept of Duality, related theorems, complementary slackness property and development of dual simplex algorithm. Sensitivity Analysis: Changes in the cost vector, requirement vector and non-basic activity vector; addition of new variables and addition of new constraints.

(12L+3T)

Unit 3: Integer Linear Programming Problem (ILPP): The concept of cutting plane, cutting plane method for all ILPP and mixed ILLP, Branch and Bound method. Quadratic programming: Kuhn-Tucker conditions, methods due to Beale, Wolfe.

(12L+3T)

Unit 4: Theory of games: two person zero sum games, minimax and maximin principles, Saddle point, mixed strategies; rules of dominance, solution of 2 x 2 game by algebraic method, Graphical method, Reduction of the game problem as LPP. Dynamic Programming: The Recursion Equation Approach, Computational Procedure, Characteristics of Dynamic Programming, Solution of L.P.P. by Dynamic Programming.

(12L+3T)

References:

1. Hadley G.(1969): Linear Programming, Addison Wesley
2. Taha H. A. (1971): Operation Research: An Introduction, Macmillan N.Y.
3. Kanti Swaroop & Gupta M. M.(1985): Operations Research, Sultan Chand & Co. ltd.
4. P.Gupta & D. S. Hira(2010): Operation Research, Sultan Chand & Co. ltd.
5. J. K. Sharma. (2003): Operation Research: Theory and Applications. Macmillan.