PUNYASHLOK AHILYADEVI HOLKAR
SOLAPUR UNIVERSITY, SOLAPUR

Name of the Faculty: Science & Technology

Syllabus

M.Sc. Statistics

Part-I (Semester-I and II)

As per NEP-2020

To be implemented from Academic Year 2026-27

PUNYASHLOK AHILYADEVI HOLKAR SOLAPUR UNIVERSITY

SCHOOL OF COMPUTATIONAL SCIENCES

DEPARTMENT OF STATISTICS

Syllabus of M.Sc. Statistics (NEP-2020)

(w.e.f.2026-27)

1) Title of the Programme: M.Sc. in Statistics.

2) Duration of the Programme: Two years.

3) Faculty: Science and Technology

4) Year of Implementation: M.Sc. Part-I : Academic year 2026-27

5) Fee Structure: As per university rules

6) Medium of Instructions: English

7) Eligibility criteria for Admission: For M. Sc. in Statistics following candidates are eligible.

(i) Three-year B.Sc. degree with Statistics as principal / major subject.

(ii) Three-year B.Sc. degree with Statistics as subsidiary / minor subject.

8) Learning Outcomes

- **Programme Outcomes (PO)**

On successful completion of M.Sc. Statistics programme the students will be able to:

i) Have a specialized knowledge and understanding of statistical theory at an advanced level which takes into account recent advances in the subject.

ii) Acquire the strong foundation of statistical concepts which will benefit them to become good academicians.

iii) Have deep understanding and ability to explain the inter-connections between various sub disciplines and apt use of these inter-connections in modelling and real life problems.

iv) Gain the knowledge of software which has wide range of opportunities in the planning and developments, IT sector, business, government and private sectors.

- **Programme Specific Outcomes (PSO)**

i) Understand, implement and develop statistical models.

ii) Effectively use necessary statistical software and computing environment including R, Python and MS-EXCEL among others.

iii) Get wide range of statistical skills in problem solving.

iv) Have the versatility to work effectively in a broad range of establishments (including R &D sectors, analytics, scientific laboratories, government, financial, health, education) or to continue higher education, and exhibit ethical and professional behavior in team work.

A Four Semester M.Sc. Statistics Course Structure as per NEP-2020

M. Sc. program in Statistics consists of 88 credits. Credits of a course are specified against the title of the course.

M.Sc. Statistics Semester-I			
Course Type	**Course code**	**Course Title**	**No. of Credits**
Major Mandatory	DSC-1	Distribution Theory	4
	DSC-2	Estimation Theory	4
	LAB-1	Statistics Practical-I	2
	LAB-2	Statistics Practical-II	2
Major Elective	DSE-1(A)	Statistical Mathematics	4
	DSE-1(B)	Demography	
	DSE-1(C)	Official Statistics	
	LAB-3	Statistics Practical-III	2
Research Methodology	RM	Research Methodology in Statistics	4
M.Sc. Statistics Semester-II			
Course Type	**Course code**	**Course Title**	**No. of Credits**
Major Mandatory	DSC-3	Stochastic Processes	4
	DSC-4	Theory of Testing of Hypotheses	4
	LAB-4	Statistics Practical-IV	2
	LAB-5	Statistics Practical-V	4
Major Elective	DSE-2 (A)	Probability Theory	4
	DSE-2 (B)	Nonparametric Methods	
	DSE-2 (C)	Modeling and Simulation	
	LAB-6	Statistics Practical-VI	
On Job Training/ Field Project	OJT/FP	-----	4

M.Sc. Statistics Semester-III			
Course Type	**Course code**	**Course Title**	**No. of Credits**
Major Mandatory	DSC-5	Regression Analysis	4
	DSC-6	Multivariate Analysis	4
	LAB-7	Statistics Practical-VII	2
	LAB-8	Statistics Practical-VIII	2
Major Elective	DSE-3 (A)	Design and Analysis of Experiments	4
	DSE-3 (B)	Time Series Analysis	
	DSE-3 (C)	Actuarial Statistics	
	LAB-9	Statistics Practical-IX	
Research Project	RP1	Research Project-I	4
M.Sc. Statistics Semester-IV			
Course Type	**Course code**	**Course Title**	**No. of Credits**
Major Mandatory	DSC-7	Data Mining	4
	DSC-8	Industrial Statistics	4
	LAB-10	Statistics Practical-X	4
Major Elective	DSE-4 (A)	Reliability and Survival Analysis	4
	DSE-4 (B)	Clinical Trials	
	DSE-4 (C)	Discrete Data Analysis	
	LAB-11	Statistics Practical-XI	
Research Project	RP2	Research Project-II	6

Punyashlok Ahilyadevi Holkar Solapur University Solapur

M. Sc. (Statistics) NEP-2020 Structure w.e.f. 2026-27

M.Sc. (Statistics) Semester –I											
Paper Code	**Title of the Paper**	**Credits**	**Contact hours/week**			**Distribution of Marks for Examination**					
			Th (L)	**Pr**	**Total**	**Internal**		**External**		**Total**	
						Th	**Pr**	**Th**	**Pr**	**Th**	**Pr**
DSC-1	Distribution Theory	4	4	---	4	40	---	60	---	100	---
DSC-2	Estimation Theory	4	4	---	4	40	---	60	---	100	---
DSE-1	(A) Mathematical Statistics (B) Demography (C) Official Statistics	4	4	---	4	40	---	60	---	100	---
RM	Research Methodology in Statistics	4	4	---	4	40	---	60	---	100	---
LAB-1	Statistics Practical-I	2	---	4	4	---	20	---	30	---	50
LAB-2	Statistics Practical-II	2	---	4	4	---	20	---	30	---	50
LAB-3	Statistics Practical-III	2	---	4	4	---	20	---	30	---	50
Total for Semester-I		**22**	**16**	**12**	**28**	**160**	**60**	**240**	**90**	**400**	**150**
M.Sc. (Statistics) Semester –II											
Code	**Title of the Paper**	**Credits**	**Contact hours / week**			**Distribution of Marks for Examination**					
			Th (L)	**Pr**	**Total**	**Internal**		**External**		**Total**	
						Th	**Pr**	**Th**	**Pr**	**Th**	**Pr**
DSC-3	Stochastic Processes	4	4	---	4	40	---	60	---	100	---
DSC-4	Theory of Testing of Hypotheses	4	4	---	4	40	---	60	---	100	---
DSE-2	(A) Probability Theory (B) Nonparametric Methods (C) Modeling and Simulation	4	4	---	4	40	---	60	---	100	---
LAB-4	Statistics Practical-IV	2	---	4	4	--	20	---	30	---	50
LAB-5	Statistics Practical-V	2	---	4	4	--	20	---	30	---	50
LAB-6	Statistics Practical-VI	2	---	4	4	--	20	---	30	---	50
OJT/FP	OJT/FP	4	---	8	8	---	40	---	60	---	100
Total for Semester-II		**22**	**16**	**20**	**32**	**120**	**100**	**180**	**150**	**300**	**250**

DSC: Discipline Specific Course **DSE:** Discipline Specific Elective **FP:** Field projects

OJT: On Job Training: Internship / Apprenticeship **RM**: Research Methodology

Evaluation Scheme

Each theory paper will have 100 marks out of which 60 marks will be for Term End examination and 40 marks for Internal Assessment. The candidate has to appear for internal evaluation of 40 marks and external evaluation (University Examination) of 60 marks for each theory paper.

Each practical paper will have 50 marks out of which 30 marks will be for Term End examination and 20 marks for Internal Assessment. The candidate has to appear for internal evaluation of 20 marks and external evaluation (University Examination) of 30 marks for each practical paper.

Internal Evaluation

- In case of theory papers internal examinations will be conducted by department / school.
- In case of practical papers, 5 marks shall be for day-to-day journal and 15 marks shall be for internal test, which will be conducted by the department / school.

External Evaluation (End of Term University Examination)

I) Nature of Theory question paper:

1) Each Theory paper is of 60 marks.
2) Each Theory paper will be of 2.5 hours.

II) Nature of Practical question paper: (End of Term Examination)

Sem-I and II: Practical examination will be conducted for 30 marks and is of two hours duration. There shall be 05 questions each of 08 marks, of which student has to attempt any 03 questions. VIVA will be for 06 marks.

M.Sc. (Statistics) Semester -I

Major Mandatory	DSC-1: DISTRIBUTION THEORY	Credits: 04

Course outcomes:

Upon successful completion of this course, the students will be able to:

i) Understand and explain the concept of univariate and bivariate random variables and related entities.
ii) Understand and explain the nature of various probability distributions and perform related computations.
iii) Understand probability models for univariate and bivariate 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: Brief review of basic distribution theory. Distribution function and its properties, continuous and discrete distributions, mixture of probability distributions, decomposition of mixture of distribution function into discrete and continuous distribution functions. Functions of random variables and their distributions in case of univariate random variables and its applications. Concept of location, scale and shape parameters of distributions with examples. Symmetric distributions and their properties. Probability integral transformation.

(15L)

Unit-2: Expectation, moments, probability generating function, moment generating function, convolution and examples. Moment inequalities: Basic, Markov, Chebychev, Holder, Minkowski and Jensen's inequalities with their applications.

(15L)

Unit-3: Bivariate random vector, joint distribution, independence, variance-covariance matrix, marginal and conditional distributions, conditional expectation and conditional variances. Transformation of bivariate random variables using Jacobian of transformation. Bivariate normal distribution and its properties, bivariate exponential (Marshall and Olkin), bivariate Poisson distributions. Multinomial distribution and its properties.

(15L)

Unit-4: Non-central Chi-square, t and F distributions and their properties. Order statistics and their distributions, joint and marginal distributions of order statistics, Distribution of sample median and sample range. Distribution of spacing, normalized spacing with illustration to exponential case.

(15L)

Reference Books

1. Rohtagi V.K. & Saleh A.K. Md. E (2001): Introduction to Probability Theory and Mathematical Statistics- John Wiley and Sons Inc.
2. Rao C.R. (1973): Linear Statistical Inference and Applications- John Wiley and Sons, Inc.

3. Johnson N.L. and. Kotz S. (1996): Distribution in Statistics, Vol-I, II and III- John Wiley and Sons, Inc.
4. Johnson N.L. and Kotz S.: Multivariate distributions- John Wiley and Sons, Inc.
5. Casella G. and Berger R.L. (2002): Statistical Inference-Duxbury Advanced Series, 2nd ed.

Major Mandatory	DSC-2: ESTIMATION THEORY	Credits: 04

Course outcomes:

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

Unit-1: Principles of data reduction: Sufficiency principle; factorization theorem; minimal sufficiency; minimal sufficient partition; construction of minimal sufficient statistic; minimal sufficient statistic for exponential families; power series family and Pitman family; completeness; bounded completeness; ancillary statistic, Basu's theorem and its applications.

(15L)

Unit-2: Unbiased estimation: unbiased estimators; 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; Chapmann-Robinson-Kiefer lower bound; Bhattacharya lower bound, and their applications.

(15 L)

Unit-3: Methods of Estimation: Method of maximum likelihood estimation (MLE); method of scoring and its application to estimation in multinomial distribution, method of moments, method of minimum Chi-square. Bayesian Estimation: Concept of prior and posterior distributions, various types of priors, loss functions, conjugate family and standard examples of such family. Bayes estimator under absolute and quadratic error loss functions.

(15L)

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. Best Asymptotic Normal (BAN) estimator; Consistent Asymptotic Normal (CAN) estimators;

invariance of CAN estimator under differentiable transformation; Methods for constructing CAN estimators for real and vector valued parameters. Cramer regularity conditions; Cramer-Huzurbazar theorem (Statement only).

(15L)

Reference Books:

1. Rohtagi V.K. & Saleh A.K. Md. E (2001): Introduction to Probability Theory and Mathematical Statistics- John Wiley and Sons Inc.
2. Rao C.R.(1973): Linear Statistical Inference and Applications- John Wiley and Sons, Inc.
3. Casella G. and Berger R.L. (2002): Statistical Inference-Duxbury Advanced Series, 2nd ed.
4. Kale B.K. (2005): First Course on Parametric Inference- 2nd ed. Narosa Publishing House.
5. Lehmann E.L. (1983): Theory of Point Estimation- John Wiley and Sons.
6. Ferguson T.S. (1967): Mathematical Statistics, Academic Press.
7. Statistical Inference : Theory Of Estimation : M. K. Srivastava, A. H. Khan, N. Srivastava
8. Anirban Das Gupta. (2008), Asymptotic Theory of Statistics and Probability, Springer Texts in Statistics.
9. Deshmukh Shailaja and Kulkarni Madhuri (2020). Asymptotic Statistical Inference: A Basic Course Using R, Kindle Edition.

Major Elective	DSE-1(A): STATISTICAL MATHEMATICS	Credits: 04

Course Outcomes:

After successful completion of this course, student will be able to-

1. Analyze sequences and series of real numbers.
2. Identify converging and diverging sequences and series.
3. Understand vector space and its properties
4. Understand the matrix algebra and its applications.

Unit 1: Set of real numbers, countable and uncountable sets, Sequence of real numbers, convergence, divergence, Cauchy sequence. Convergence of bounded monotone sequence. Limit inferior and limit superior of the sequences. Series of numbers, tests for convergence (without proof), test for absolute convergence, convergence of sequences of non-negative terms. Examples.

(15L)

Unit 2: Real valued function, Continuous function, Uniform continuity of sequence of functions, Uniform convergence of power series, Radius of convergence. Reimann, Reimann-Stieltjes, Integrals and their common properties. Integration by parts, Fundamental theorem on calculus, mean value theorem. Vector and Matrix differentiation, Maxima, minima of functions of variables. Constrained maxima, minima, Lagrange's method, Taylor's theorem (statement only) and its applications. Implicit function theorem and their applications. Multiple Integrals, Change of variables, Improper integrals, Applications in multivariate distributions. Theorem on differentiation under integral sign (statement only), Leibnitz rule (statement only) and applications. Examples.

(15L)

Unit 3: Vector space, subspace, linear dependence and independence, basis, dimension of a vector space, example of vector spaces. Null space, Matrices, elementary operations, rank of a matrix and related results. Orthonormal basis and orthogonal projection of a vector, Orthogonal matrices, Gram-Schmidt orthogonalisation. Idempotent matrices, Reduction of a matrix to echelon, diagonal, triangular forms, inverse of a matrix, their simple properties, partitioned matrices.

(15L)

Unit 4: G-inverse, Moore-Penrose inverse and its properties, Solution of a system of homogenous and non-homogenous linear equations and theorem related to existence of solution. Characteristic roots of a matrix, algebraic and geometric multiplicities, characteristics vectors and their orthogonal property. Cayley-Hamilton theorem and applications. Spectral decomposition, singular value decomposition and Choleskey decomposition. Quadratic forms: Definition and classification, reduction, simultaneous reduction of two quadratic forms, maxima and minima of ratio of quadratic forms.

(15L)

Reference Books:

1. Malik S. C. and Arora S. (1991). Mathematical Analysis- Wiley Eastern Limited 2nd Ed.
2. Goldberg R. R. (1964). Methods of Real Analysis-Blaisell Publishing company, New York, USA.
3. Bartle G.R. (1976). Element of Real Analysis-Wiley 2nd edition.
4. Royden (1988). Principles of Real Analysis-Macmillian.
5. Widder (1989). Advanced Calculus-Dover Publication.
6. Apostol T.M. (1985). Mathematical Analysis-Narosa Publishing House
7. Searle S. R. (1982): Matrix Algebra useful for Statistics, John Wiley and Sons, Inc.
8. Graybill F. A. (1983): Matrices with application in Statistics- 2nd ed. Wadsworth
9. Hadely C. (1987): Linear Algebra, Narosa Publishing House.

Major Elective	DSE-1(B): DEMOGRAPHY	Credits: 04

Course Outcomes:

After successful completion of this course, student will be able to

1. Understand the basic concepts of demography and vital statistics.
2. Understand the trends of mortality and compare and contrast among different age and sex group.
3. Identify the components of population change, including the effects of changing birth, death and migration rates, and demonstrate their influences on age structure.
4. Do population projection by different methods.

Unit-1: Demography and its interdisciplinary nature, sources of demographic data, Coverage and Content errors. The use of balancing equation, Chandrasekaran and Deming formula to check completeness of registration data. Use of Whipple's Myers's and UN Indices.

(15L)

Unit-2: Measures of Mortality: Various measures of mortality, infant mortality rate, cause specific death rates and standardized death rates. Measures of Fertility: Period and cohort fertility measures, use of birth order statistics, child-women ratio, Brass P/F ratio to estimate current level of fertility, Measures of reproduction and replacement. Sheps and Perrin stochastic human reproductive process.

(15L)

Unit-3: Life Tables: Types of life tables, inter-relationships between life table functions, construction of life tables using Reed- Meerel and Greville's Method. Probability distribution of life table functions and their optimum properties. Population estimation and Projection : Mathematical , Statistical and Demographic Methods, Component method.

(15L)

Unit-4: Stable and Quasi – stable population: Derivation of Lotka's stable population model and properties, Intrinsic growth rate and derivation, age structure and birth rate of a stable population, mean length of generation, momentum of population growth, Quasi-stable population under changing fertility and mortality situations.

(15L)

References:

1. Shryoch, Henry S, Jacob S, Siegel and Associates (1964). Methods and materials of demography (condensed edition) Academic press, London.
2. Barclay, George W. (1968). Techniques of population analysis, John Wiley and Sons, New York.
3. Keyfitz N. (1968). Introduction to Mathematics of Population. Addision-Wesley Publishing Co, Reading, Messachusetts.
4. R. Ramkumar (1986). Technical Demography, Wiley Eastern, New Delhi.
5. Sudhendu Biswas (1988). Stochastic processes in Demography and Applications, Wiley Eastern, New Delhi.

Major Elective	DSE-1(C): OFFICIAL STATISTICS	Credits: 04

- **Course Outcomes:**

Upon successful completion of this course, the student will be able to:

1. To know about Indian and International official statistical systems.
2. Prepare for qualifying Indian statistical service examination.

Unit-1: Introduction to Indian and International statistical systems. Role, function and activities of Central and State statistical organizations. Organization of large scale sample surveys. Role of National Sample Survey Organization. General and special data dissemination systems.

(15L)

Unit-2: Statistics for National Development-Economic Development: Growth in per capita income and distributive justice. Indices of development; Human development index. Estimation of national income-product approach, income approach and expenditure approach.

(15L)

Unit-3: Population growth in developed and developing countries, evaluation of performance of family welfare programmes, projections of labor force and manpower. Scope and content of population census of India.

(15L)

Unit-4: System of collection of Agricultural statistics. Crop forecasting and estimation, productivity, impact of irrigation projects. Statistics related to industries, foreign trade, balance of payment, cost of living inflation, educational and other social statistics.

(15L)

Reference Books:

1. Basic Statistics related to Indian Economy (CSO) 1990.
2. Guide to official Statistics (CSO) 1999.
3. Statistical System in India, (CSO) 1995.
4. Principle and accommodation of National Population Censuses, UNESCO.
5. Panse V.G. Estimation of Crop Yields, (FAO).
6. Family welfare, Yearbook, Annual publication of D/o Family Welfare.
7. Monthly Statistics of Foreign Trade in India, DGCIS, Calcutta and other Govt. Publication.

Research Methodology	RM: Research Methodology in Statistics	Credits: 04

Course outcomes:

After successful completion of this course, student will be able to

1. Understand research methodology and research ethics.
2. Understand the basic concepts involved in research.
3. Understand the basic principles underlying survey design and random sampling.
4. Apply the different sampling methods for designing and selecting a sample from a population.

Unit-1: Meaning of research, Objectives and motivation of research, Types of research. Research approaches, significance of research, Research methods versus research methodology, Importance of knowing how research is done, Research process, criteria of good research. Problems encountered by researchers in India.

(15L)

Unit-2: Meaning of literature survey and review. Understanding and selecting a research problem, steps in formulation of research problem and research objectives. Quality indices of research publication: citations, i10 index, impact factor, H-index, Science citation index, Using Web for literature review: Google scholar, Scopus, Web of Science, Math Sci net.

(15L)

Unit-3: Concept of population and sample, Need for sampling, Census and sample surveys, basic concepts in sampling and designing of large-scale survey design, sampling scheme and sampling strategy. Sampling and non-sampling errors, Response and non-response errors. Probability and non-probability sampling. Examples of non-probability sampling: Quota sampling, snowball sampling, judgment sampling. Basic methods of sample selection: SRSWR and SRSWOR, Stratified and systematic sampling.

(15L)

Unit-3. PPSWR Methods: Cumulative total method, Lahiri's method, related estimation problems, PPSWOR method and related estimation of a finite population mean (Horvitz Thompson and Des Raj estimator for general sample size and Murthy's estimator for sample of size 2, Midzuno sampling. Use of supplementary information for estimation: Ratio and Regression estimators and their properties. Cluster sampling, Double sampling, Two-stage sampling with equal number of second stage units.

(15L)

Reference Books:

1. Kothari C. R. and Garg G. (2014). Research Methodology: Methods and Techniques, 3rd Edition, New Age International Publisher.
2. Kumar R. (2014). Research Methodology-A step-by-step Guide for Beginners, 4th Edition, Sage Publication.
3. Sukhatme P.V., Sukhatme P.V., Sukhatme S. and Ashok C. (1997): Sampling Theory of Surveys and Applications- Piyush Publications.

4. Des Raj and Chandhok P. (1998): Sample Survey Theory. Narosa publishing House.
5. William G. Cochran (1977): Sampling Techniques, 3rd edition- John Wiley & Sons.
6. Parimal Mukhopadhyay (1988): Theory and methods of Survey sampling, Prentice Hall of India Pvt. Ltd.
7. Murthy M.N. (1977): Sampling Theory and Methods, Statistical publishing Society, Calcutta.
8. Latpate R., Kshirsagar J., Gupta, V. K. and Chandra G. (2021). Advanced Sampling Methods, Springer.
9. Chaudhuri, A., and Stenger, H. (2005). Survey sampling: theory and methods. CRC Press.
10. Cochran, W. G. (1977). Sampling techniques. John Wiley & Sons

LAB-1	**Statistics Practical-I** **(Based on DSC-1 and DSC-2)**	**Credits: 02**

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.

1	Sketching of PDF/PMF and CDF
2	Decomposition of Mixture Distribution
4	Construction of UMVUE
5	Methods of ML Estimation
6	Methods of Moments Estimation
7	Method of Scoring
8	Bayesian Estimation
9	Consistent Estimator and Comparison
10	Construction of CAN Estimators

LAB-2	**Statistics Practical-II** **(Based on MS-EXCEL and R)**	**Credits: 02**

Course outcomes:

After successful completion of this course, student will be able to

1. Perform data organization, data manipulation, statistical and mathematical computations, and data analysis using MS-EXCEL.
2. Perform data organization, data manipulation, statistical and mathematical computations, and data analysis using R.
3. Perform mathematical/statistical computations, statistical data analysis using built-in functions in MS-EXCEL and R and develop programs for various tasks.

Unit-1: MS-EXCEL: Introduction to MS-EXCEL, 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 statistics, computing pmf/pdf, cdf and quantiles of well-known distributions, rand and rand-between function, logical functions: if, and, or, not. Look-up 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.

MS-EXCEL Exercise

- Reading and creating data, certain computations using data. Descriptive statistics and construction of frequency distribution.
- Listing probabilities for standard distributions and plotting its probabilities and distributions.
- Plotting density functions and distribution functions for standard continuous distributions.
- Using macro programming for certain iterative computing.

Unit-2: R-software: Introduction to R, data type and objects, operators, data input, data import and export, built-in functions for descriptive statistics, random sampling and computation of pmf/pdf, cdf and quantiles of well known distributions. Strings and dates in R. apply family of functions, solving work in R. matrix algebra, graphical procedures, frequencies and cross-tabulations, built-in functions: lm, t.test, prop.test, wilcox.test, ks.test, vr.test, chisq.test, aov. Control statements, programming, user defined functions, R-packages, R-studio.

R-Programming

- Simple R exercise using –scan function, reading data from EXCEL and exercises, vectors, matrices, rbind and cbind.
- Exercises on Matrices
- Reading data from text file. Data frames, names etc. Exercises based on these data and graphics.
- Exercises on Functions in R
- Generating samples from standard discrete and continuous distributions.
- Verification of law of large numbers and central limit theorem.

LAB-3	Statistics Practical-III (Statistical Computing)	Credits: 02

Course outcomes:

Upon successful completion of this course, the student will be able to:

1. Understand the basic ideas of random number generation using different techniques.
2. Learn theoretical methods and practicable techniques of statistical simulations.
3. Understand how to apply Monte Carlo simulations and the EM algorithm.

Unit 1: Simulation: An introduction, need of simulation, random number generator, true random number and pseudo random number generators, requisites of a good random number generator. Statistical tests for random numbers. Congruential method of generating uniform random numbers, Inverse transforms method, Acceptance-rejection method. Algorithms for generating random numbers from well-known univariate discrete and continuous distributions.

- Generating samples from univariate discrete and continuous distributions.
- Generating random vectors from multinomial, bivariate normal, and bivariate exponential distributions.
- Generating random numbers from mixture of distributions.
- Use of random numbers to evaluate integrals and to study the systems involving random variables.
- Use of random numbers for performance evaluation of estimators and statistical tests.

Unit 2: Resampling methods

- Bootstrap method: 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 finding roots of nonlinear equation: Newton-Raphson method, bisection method; Newton-Raphson method for system of non- linear equations.
- Numerical integration: quadrature formula, trapezoidal rule and Simpson's rules for single integral.
- Introduction to EM algorithm.

M.Sc. (Statistics) Semester-II

Major Mandatory	DSC-3: STOCHASTIC PROCESSES	Credits: 04

Course outcomes:

Upon successful completion of this course, the student will be able to:

1. Study the fundamental concept of stochastic processes and its applications.
2. Understand Markov processes and Markov chains and their applications in real world.
3. Study the branching process and its properties.
4. Understand Poisson processes and its variations.

Unit-1: Definition of Stochastic process, Finite dimensional distributions. Examples of various stochastic processes. Markov chain and its Examples. Chapman-Kolmogorov equation, calculation of n-step transition probabilities. Classification of states of Markov chain, irreducible Markov chain, period of the state, First entrance theorem, First passage time distribution. Long run distribution of Markov chain, relation of mean recurrence time and stationary distribution.

(15L)

Unit-2: Random walk and gambler's ruin problem. Discrete state space continuous time Markov chain. Poisson process and related results. Birth and death processes and associated cases. M/M/1, M/M/S queuing models and related properties.

(15L)

Unit-3: Galton-Watson branching process. Probability of ultimate extinction, distribution of population size and associated results. Simulation of Markov chain, Poisson process, branching process (Algorithms).

(15 L)

Unit-4: Brownian motion, Hitting Time, Maximum, Reflection Principle Stopping Time and Strong Markov Property**,** Reflection principle.

(15 L)

Reference Books:

1. Medhi. J. (1982). Stochastic Process- Wiley Eastern.
2. Parzen E. (1962). Stochastic Process-Holden-Pay.
3. Karlin & Taylor (1975). A First Course in Stochastic Processes-Vol-I Academic Press.
4. Cinlar E. (1975). Introduction to Stochastic Process. Prentice Hall.
5. Srinivas and Mehta (1976). Stochastic Processes-Tata MgGraw-Hill.
6. Adke and Manjunath (1984). An introduction to finite Markov Processes- Wiley Eastern.
7. Bhat B. R. (2000). Stochastic Model: Analysis and Application, New Age International.
8. Sheldon Ross: Introduction to Probability Models, 11th Edition, Academic Press Publication

Major Mandatory	DSC-4: THEORY OF TESTING OF HYPOTHESES	Credits: 04

Course outcomes:

Upon successful completion of this course, the student will be able to:

1. Understand the concept of testing of hypothesis, test statistic, critical region, size and power of a test. Neyman-Pearson fundamental lemma, UMP test and interval estimation.
2. Understand the concept of critical regions, likelihood ratio test with its asymptotic distribution.

Unit-1: Problem of testing of hypothesis, null and alternative hypotheses, simple and composite hypotheses, randomized and non-randomized tests, power function of a test, Most powerful (MP) tests, Neyman-Pearson Lemma and its applications. Monotone likelihood ratio (MLR) property, uniformly most powerful (UMP) tests and their existence for one-sided alternatives.

(15 L)

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

(15L)

Unit-3: Interval estimation: confidence level, construction of confidence intervals using pivots, shortest length confidence interval, UMA confidence interval and its relation to UMP test, UMAU confidence interval and its relation with UMPU test, variance stabilizing transformations (VST) and their existence, asymptotic confidence intervals based on CAN estimators and VST.

(15L)

Unit-4: Likelihood ratio test (LRT) and its asymptotic distribution, application of LRT to standard distributions, Wald test, Rao's score test, Pearson Chi-square test for goodness of fit, application to contingency tables, Bartlett's test for homogeneity of variances, Kruskal-Wallis test.

(15L)

Reference Books:

1. Kale B. K. (1999). A first Course on Parametric Inference, Narosa
2. Rohatgi V.K. (1988), Introduction to Probability and Mathematical Statistics, Wiley Eastern Ltd. New Delhi. Student Edition.
3. Dudewicz E. J. and Mishra S. N. (1988), Modern Mathematical Statistics, Wiley Series
4. Lehman E. L. (1987). Theory of Testing of Hypotheses. Student Edition.
5. Srivastava M. and Srivastava N. (2009): Statistical Inference-Testing of Hypotheses, PHI Learning Pvt. Ltd.
6. Shanthakumaran A. (2001): Fundamentals of Testing of Hypotheses, Atlantic Publishers and Distributors.

Major Elective	DSE-2(A): PROBABILITY THEORY	Credits: 04

Course outcomes:

Upon successful completion of this course, the student will be able to:

1. Understand the concept of random variables, sigma-fields generated by random variables.
2. Learn probability distributions and independence of random variables related to measurable functions.
3. Gain the ability to understand the concepts of different types of generating function, sequence of random variables, convergence, modes of convergence.
4. Learn the concepts of weak, strong laws of large numbers and central limit theorem and their importance.

Unit-1: Classes of sets, Sequence of sets, limsup and liminf and limit of sequence of sets, field, σ- field, σ- field generated by a class, Borel σ-field. Probability measure, Probability space, properties of probability measure-continuity, mixture of probability measures. Lebesgue and Lebesgue-Stieltjes measures on R. Independence of events.

(15L)

Unit-2: Measurable function, random variable, distribution of a random variable, simple random variable, elementary random variable, liminf, limsup and limit of sequence of random variables. Method of obtaining a random variable as a limit of sequence of simple random variables. Integration of a measurable function with respect to a measure. Expectation of a random variable, independence. Characteristic function, simple properties, Inversion theorem and uniqueness property (Statement only).

(15L)

Unit-3: Monotone convergence theorem, Fatou's Lemma, Dominated Convergence theorem, Borel-Cantelli Lemma and their applications. Convergence of sequence of random variables, Convergence of distribution, Continuity theorem (Statement only), Almost sure convergence, a characterizing property, convergence in probability, uniqueness of limit. Yule-Slutsky results. convergence in r^{th} mean, interrelationships.

(15L)

Unit-4: Weak and Strong laws of large numbers, Kolmogorov's three series theorem for almost sure convergence (statement only), Liaponove's Lindeberg-Feller Theorems on CLT (statement only). Application of the above results.

(15L)

Reference Books:

1. Bhat B. R. (1999). Modern Probability Theory (3^{rd} Ed.), New Age International (P) Ltd.
2. Billingsley P. (1986). Probability and Measure-.John Wiley and Sons.
3. Chandra, T. and Gangopadhyay, S. (2017). Fundamentals of Probability Theory. Narosa Publishing House.

4. Gut, A. (2005). Probability: A Graduate Course. Springer.
5. Alan Karr (1993). Probability Theory-Springer Verlag.
6. Kingman, J F C and Taylor, S.J. (1966). Introduction to Measure and Probability-Cambridge University Press.

Major Elective	DSE-2(B): Nonparametric Methods	Credits: 04

Course outcomes:

Upon successful completion of this course, the student will be able to:

1. Learn the basic concepts of nonparametric techniques.
2. Learn about non-parametric methods for computing survival function.
3. Estimate nonparametric survival function of the data.

Unit-1: Definition of nonparametric test, Advantages and disadvantages of nonparametric tests. One sample tests: Sign test, Wilcoxon's signed rank test, Run test. Two sample tests: Wald-Wolfowitz run test, Mann-Whitney test, Median test, Kolmogorov-Smirnov test.

(15L)

Unit-2: Test for two-sample scale problem: Mood, Klotz, Ansari-Bradley, Siegel-Tukey and Sukhatme tests.

(15L)

Unit-3: Spearman's rank correlation test, Kendall's rank correlation test, Thiel's test for regression coefficients, Hollander's test for parallelism of two regression lines. Kruskal-Wallis test. Estimable parametric functions, kernel, symmetric kernel, Definition of U-statistics, U-statistics theorem for one sample and two samples (statements only).

(15L)

Unit-4: Concept of time order and random censoring, likelihood in these cases, survival function, hazard function, nonparametric estimation of survival function, Cox's proportional hazards model, the actuarial estimator, Kaplan-Meier estimator.

(15L)

Reference Books:

1. Gibbons J. D. and Chakraborti S. (2003). Nonparametric Statistical Inference, 2nd Edition, Marcel Dekker Inc.
2. Randles R. S. and Wolf D. A. (1979). Introduction to Theory of Nonparametric Statistics, John Wiley and Sons.
3. Rohatgi V. K. (1976). An Introduction to Probability Theory and Mathematical Statistics, John Wiley and Sons.
4. Daniel W. W. (2000). Applied Nonparametric Statistics, 2nd Edition, Duxbury.

Major Elective	DSE-2(C): MODELING AND SIMULATION	Credits: 04

- **Course Outcomes:**

Upon successful completion of this course, the student will be able to:

1. Construct models in discrete and continuous time based markov chains.
2. Describe and analyze basic markov queuing models and situations to which they may apply.
3. Construct simple network diagrams.
4. Analyze network that have deterministic time.
5. Learn different types of simulation techniques.
6. Use the methods of generating random numbers and understand the properties of random number generators.
7. Generate random values for variables with a specified probability distribution.

Unit-1: Stochastic Models: Introduction, Poisson Process, Markov chains and applications. Queuing Models: Introduction, Queuing System, Elements of queuing system, birth and death process model, Queuing Models M/ M/1, M/M/C.

(15L)

Unit-2: Network Analysis: Applications of PERT and CPM techniques, Network diagram representation, Rules for constructing the network diagram, Determination of critical path.

(15L)

Unit-3: Simulation: Introduction, Uses of simulation, Steps in simulation study, Advantages and disadvantages of simulation, Simulation models- continuous and discrete simulations. Simulation Models: (Flow chart and/or algorithms): Monte-Carlo simulation, Simulation of queuing system, Simulation of PERT problems.

(15L)

Unit-4: Random Number Generation: Introduction, Types of random numbers, Pseudo random number generator, Tests for random numbers, Techniques for generating random numbers, Inverse transformation technique, Generating random variates from Uniform, Bernoulli, Binomial, Exponential and Normal distributions.

(15L)

Reference Books:

1) Allen Arnold O. (1978). Probability, Statistics and Queuing with Computer Science Applications, Academic Press.
2) Kishore Trivedi. (1982). Probability and Statistics with Reliability, Queuing with computer science Applications, Prentice Hall.
3) Geofrey Gordon (1999). System Simulation, PHI, Second Ed.
4) NarsinghDeo (1979). System Simulation with Digital Computer, PHI.
5) Fred Maryanski (1987). Digital Computer, Simulation, CBSPD.
6) Medhi j. (1982): Stochastic Process, Wiley

LAB-4	ME24: STATISTICS PRACTICAL-IV (Based on DSC-3 and DSC-4)	Credits: 04

- **Course Outcomes:**

Upon successful completion of this course, the student will be able to perform:

- Classification of States
- Simulation from Markov Chain
- Stationary Distribution of Markov Chain
- Simulation from Counting Process
- Computation of Probabilities of Type-I and Type-II Errors
- Construction of MP , UMP and UMPU Tests
- Confidence Intervals
- Asymptotic Confidence Intervals based on CAN Estimators and VST.
- Likelihood Ratio Tests
- Chi-square Test of goodness of fit
- Bartlett's Test For Homogeneity of Variances
- Kruskal-Wallis Test.

LAB-5	STATISTICS PRACTICAL-V (Optimization Techniques-I)	Credits: 04

- **Course Outcomes:**

Upon successful completion of this course, the student will be able to:

1. Concept of Linear programming problem and its application to real life problems.
2. Importance of Transportation and Assignment problems.

Unit-1. Linear Programming Problem (LPP): Introduction to linear programming problems, Solution, basic solution, feasible solution, basic feasible solution and related theorem. Unbounded solution, optimal solutions, related theorems (statements only). Method for Solving LPP: Graphical method, simplex algorithm, artificial techniques: two-phase, big-M. Duality in linear programming.

Unit 2. The general transportation problem, duality in transportation problem, loops in transportation table, solution of transportation problem, test for optimality, degeneracy, transportation algorithm (MODI method), minimization transportation problem. Assignment problems: Mathematical formulation of assignment problem, Hungarian method for solving assignment problem.

LAB-6	Statistics Practical-VI (Data Analysis Using Python)	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.

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*.

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. Vander Plas, J. (2016). *Python data science handbook: Essential tools for working with data*. " O'Reilly Media, Inc.".
7. Mark Lutz . Programming Python, 4th Edition
8. 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/*

OJT/FP	On Job Training / Field Project	Credits: 04

On Job Training (OJT)/Field Project aims to provide necessary practical knowledge and hands-on experience in the application of Statistical tools to solve real-world problems. The objectives of OJT/FP in Statistics are as follows:

1. To understand the applications of statistical techniques in different fields.
2. To gain expertise in handling different statistical software.
3. To identify the potential use of statistical tools in different situation.
4. To enhance the revenue by using different statistical optimization techniques.
5. To explore the real-world data with the aim of getting meaningful insights from it.

Course Outcomes: Upon successful completion of this course, student will be able to:

1. Apply various statistical tools to different societal problems.
2. Analyze the real-world data using Statistical techniques.
3. Enhance the expertise on use of Statistical Software.
4. Interpret the output and convey it efficiently to the end user.
5. Explore himself / herself in competitive corporate environment.

On Job Training (OJT):

1. For On Job Training, students will be asked to work with various govt./semi-govt./private institutions/organizations/industries, where they can explore their statistical knowledge.
2. Students are expected to gain the practical knowledge under the mentoring of personnel available therein.
3. A student is expected to spend not less than 60 working hours for On Job Training and related activities.
4. On Job Training will be carried out in summer vacation after the students complete their second semester examination.
5. Students need to provide the confirmation letter from the organization or the Institute where they have joined for On Job Training.
6. Continuous evaluation of the students' performance in the On Job Training will be carried out with the assistance of the personnel of training institution/organization where the training is being given.
7. The proof of completion of On Job Training (work experience certificate / field report) should be submitted after the completion of On Job Training, duly issued and signed by the concerned training organization.

Evaluation: Students' performance will be evaluated for 100 marks (04 credits), which will be split into two parts. **Internal Evaluations:** 40 marks based on the certificate of assessment issued by organization, where the student completed his/her On Job Training. **External/End Semester Evaluation:** 60 marks, out of which 40 marks will be for submitted report of OJT and 20 marks for comprehensive viva-voce examination.

Field Project (FP):

The objective of Field Project is to train the students to undertake projects individually or in groups. The projects shall enable the students to take up their own statistical study and to understand the application of statistical methods that they learned during the course. For the project each student shall work under the supervision of a faculty member of the department. In consultation of the supervisor, students shall decide on a topic for their Field Project. The topic shall be presented before all faculty members for approval with details objective and methodology. Once approved, the student shall work on the project. There shall be a midterm evaluation of the project to appraise the continuous progress. Before the start of the semester (University) examination, students are required to submit the project report in hard copy in duplicate. During the end of semester examination, students shall present the same, whereby they shall be evaluated by an external examiner.

- Project report shall contain the Statement of Problem, Data collection, Methodology adopted, Statistical tools used for analysis, Findings, Conclusion, Suggestion and References.
- Project work will be assessed for 100 marks (04 credits), out of which 40 marks are reserved for internal evaluation based on primary preparation for the project like selection of topic, preparation of questionnaire, synopsis presentation and day-to-day project work reporting, mid-term project presentation etc.
- End of Term assessment of the project for 60 marks will be done on the basis of presentation, findings and report of the project, out of which 20 marks are reserved for VIVA.