Sant Gadge Baba Amravati University, Amravati

Faculty of Science and Technology
Programme: M. Sc. Statistics

Program Outcomes (POs)

At the end of the programme, students would be able to
PO1: **Critical Thinking**: Think critically for data driven solution with advanced Methodology. Students may be able to think critically everything around him/ her. They will be able to take better action in any critical situation in their life.

PO2: Nurture their curious minds toward translation and application & Statistical knowledge to Find solutions to real world problems.

PO3: **Employability**: get knowledge and skills in depth necessary for employability of students in industry, Govt Sector and organizations as well as in academics.

PO4: **Empowerment**: be empowered to investigate, solve questions for which answer lies beyond the boundaries of conventional thinking.

PO5: **Effective decision making**: take proper decision in critical and complicated situations. Reduce risk factor and be able to maximize project.

PO6: **Effective citizenship**: get acquainted with the needs of the industry and society, and they become the assets for the society.

PO7: **Social interaction**: Develop sensitivity for social issues and become productive citizen of the nation.

Program Specific outcomes (PSOs)
Upon completion of the programme successfully, students would be able to

PSO1: understand basic theoretical and applied principles of Statistics needed to enter the job force. They will be able to communicate key statistical concepts to non-statisticians. They also gain proficiency in using statistical ideas/ principles for data analysis.

PSO2: groom as the next generation statisticians ready for scientific decision making, aided with advanced statistical software translating into sharp and extensive analytics pertinent to various domains.

PSO3: start consultancy for statistical analysis and can be helpful in interdisciplinary research.

PSO4: perform Statistical quality control, which is the most important sector of any industry where students can work as statisticians to approve the quality.

Employability Potential of the programme:

The programme offers many opportunities, where statistician can work.

By completing this programme, students are able to:

- Analyse the things
- Understand patterns in them by asking different questions to it
- Compete with the current demand of field

- To solve a specific problem
-

This “skill” is a key requirement for many analysis type jobs like,

1. Statisticians
2. Business Analyst
3. Mathematician
4. Professor
5. Risk Analyst
6. Data Analyst
7. Content Analyst
8. Statistic Trainer

Besides all these students can work in various banking sector.

Students can also work in government sector :

- Indian statistical services (ISS)
- Staff selection services(SSC)
- Reserve Bank of India (Junior statistical officer)

By taking the programme, students are able to:

- Analyse numbers
- Understand patterns in them by asking different questions to it
- Go about it in a systematic fashion
- To solve a specific problem

This “skill” is a key requirement for many analysis type jobs.

Career Options::

Data Analytics #1: Students could get into any Analytics firm, and can assist customers in getting patterns out of data.

Data Analytics #2: For Data Analytics in banks, there can be algorithms developed for fraud deduction using the digital imprints. This requires analysing large amounts of data. That could a career choice - Digital Forensics.

Market Research: For doing a survey for customer expectations and behaviours, the data pours in, from online and offline channels - how students draw meaningful, actionable conclusions? Students need to use the statistical methods learnt. So, Market Research in a MR firm or a corporate entity can be a large area of focus.

Software Programmer: With analytical bent of mind, Students could take up a software programming job. It might not leverage learning but will leverage the “bent of mind” cultivated out of the education.

Students could focus on areas like: Visual Representation of Data (Tableau, Quilk, PowerBI), Data Reporting (Crystal Reports) - that are aligned to the core skills.

Government Statistician: Our country requires a lot of econometric and statistical data for its running. Acreage, Yields, Health Statistics and the like. Bright young idealistic people are required to run our country too.

Students could be a District Statistical Officer, who are in charge of collecting information from the district, analysing it and sharing with the State Authorities.

M. Sc. (Statistics): Name of the papers
Semester – I Total credits: 22

S. N.	Type	Name of Paper	Credits
1	DSC -I	Elementary Probability & Distribution Theory	3
2	AEC on DSC II	Data analysis & modeling Based on DSC - I	1
3	DSC-III	Estimation Theory	4
4	DSC -IV	Sampling Theory	4
5	DSE-1	1. Industrial Statistics 2. Time Series 3. Demography 4. Real Analysis and Measure Theory	4
6	Lab	Practical based on DSC- I & DSC-II	3
7	Lab	Practical based on DSC- III & DSE-1	3

M. Sc. (Statistics)
Semester – II Total credits: 22

S. N.	Type	Name of Paper	Credits
1	DSC IV	Testing of Hypothesis	3
2	AEC	Test of significance	1
3	DSCV	Advanced Probability theory	4
4	DSC VI	Design of experiment	4
5	DSE II	1. Survival Analysis 2. Advanced Sampling theory 3. Data Mining 4. Statistical Genetics and Bio-informatics	4
6	Lab	Practical based on DSC- IV & DSC-V	3
7	Lab	Practical based on DSC- VI & DSE-II	3

Duration of examination: 3 Hrs. each.
Theory + MCQ = 80 marks & Internal marks: 20 = 100 marks.
For AEC: Internal marks = 25
Total marks for Sem-I = 625
Total marks for Sem-II = 625

M. Sc. (Statistics)
Semester – III Total credits: 22

S. N.	Type	Name of Paper	Credits
1	DSC -VI1	Advanced Statistical Inference	3
2	DSC -VI1I	Mathematical Programming	1
3	DSC -IX	Linear and Non linear Modelling	4
4	DSE-1I	1. Advanced Operation Research 2. Bioassay 3. Econometrics I 4. Actuarial Statistics	4
5	Lab	Practical based on DSC- 1 & DSC-2	3
6	Lab	Practical based on DSC- 3 & DSE-2	3

M. Sc. (Statistics)
Semester – IV Total credits: 22

S. N.	Type	Name of Paper	Credits
1	DSC -X	Computational Statistics	3
2	AEC	Applications of Monte Carlo Simulation	1
3	DSC -XI	Multivariate Analysis	4
4	DSC -XII	Stochastic Processes	4
5	DSE-1I	1. Bayesian Inference 2. Clinical Trial 3. Econometrics II 4. Regression Analysis	4
6	Lab	Practical based on DSC-X & DSC-XI	3
7	Lab	Practical based on DSC- XII & DSE-1	3

M.Sc. (Statistics)

Semester – I

DSC – 1

ELEMENTARY PROBABILITY & DISTRIBUTION THEORY

SYLLABUS

Unit I: Definition of probability: Classical, relative, frequency approach, axiomatic approach. Conditional probability, independence of events, Bayes theorem and its applications, random variable, expectation of random variable, brief review of basic distribution theory, concept of pmf, pdf, cdf, joint, marginal and conditional expectation.

Unit II: Discrete distributions: binomial, Poisson, negative binomial, geometric, uniform, hypergeometric distributions.

Unit III: Continuous distributions: Uniform, normal, exponential, beta, gamma, Cauchy, Log-normal, Weibull, Laplace.

Unit IV: Chi-square, t, f distributions and their properties, Markov, Jensen, Lyapunov inequalities, approximating distributions of sample moments.

Unit V: Compound, truncated and mixture distributions (only concepts) ordered statistics, their distribution and properties. Joint and marginal distributions of order statistics, extreme value and their asymptotic distributions (Statement only) with applications.

References:

1) Johnson S. and Kotz (1972): Distributions in Statistics Vol.-I, II, and III.

2) Rao C.R. (1973): Linear Statistical Inference and its applications, Second Edition, Wiley.

3) Rohatgi V.K. (1984): An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern.

The course explores the basic concepts of Modern Probability Theory and its applications for decision making in various fields of social sciences in our everyday lives as well as economic and business activities which are full of uncertainties. The course is heavily oriented towards the formulation of mathematical concepts of probability and probability distributions with practical applications.

Course Outcome

At the successful completion of the course, students will be able to

1. develop problem solving techniques needed to accurately calculate probabilities and selected probability distributions
2. apply problem solving techniques to solve real world events.
3. be in a position to calculate statistic such as mean and variance of commonly used probability distributions.

M.Sc. (Statistics)

Semester – I

AEC on DSC – 1

DATA ANALYSIS & MODELING BASED ON DSC-1

SYLLABUS

Study of different distributions (covered in DSC 1) using real life data sets, testing goodness of fit for different distributions.

Course Outcome

After completing this Ability Enhancement Course, student will be able to;

1. Propose a suitable probability distribution as a model for a real-life data.
2. be in a position to check the validity of proposed model.

M.Sc. (Statistics)

Semester – I

DSC – 2

ESTIMATION THEORY

SYLLABUS

Unit I: Parametric Models, Probability of point estimation consistent estimators, unbiased estimator, minimum variance unbiased estimator (MVUE)/ Likelihood function Likelihood equivalence.

Unit II: Methods of Estimation: Maximum likelihood, minimum Chi-square, Method of moments, method of scoring, Cramer family. Cramer –Huzurbazar theorem, properties of maximum likelihood estimator.

Unit III: Fisher information and information matrix, Cramer-Rao Inequality, sufficiency, Factorization theorem, minimal Sufficiency, minimal sufficient statistics for exponential family & pitman family.

Unit IV: Rao-Blackwell theorem, completeness, bounded completeness, Bounded completeness, Lehman-Scheffe theorem and their use.

Unit V: Interval Estimation, confidence level, construction of confidence interval using pivots, shortest length C.I. shortest expected length C.I.

References:

1) E.L. Lechman (1986): Theory of Point Estimation.

2) B.K. Kale (199): First Course on parametric inference.

3) Rao C.R. (1973): Linear Statistical inference and its applications.

4) V.K. Rohatri (1988): An Introduction to Probability and Mathematical Statistics.

Course outcomes:

After completing this course successfully, students would be able to

1. be Aware of estimation (point, as well as, interval)
2. Apply various estimation and testing procedures to deal with real life problems.
3. Use Various methods for Estimation.
4. Understand Fisher Information, Lower bounds to variance of estimators, MVUE.
5. Get Knowledge of the concept of completeness, Basu's theorem, Rao Blackwell theorem, Lehmann Scheff theorem etc.

M.Sc. (Statistics)

Semester – I

DSC – 3

SAMPLING THEORY

SYLLABUS

Unit I: Review of basic finite population sampling techniques (SRS, WR/WOR, Stratified Systematic) and related results on estimation of Population mean, total and variance, Allocation problem in Stratified Sampling, Comparison of systematic sampling with SRS and Stratified random sampling.

Unit II: Unequal probability sampling: PPS, WR/WOR methods (including Lahiri's Scheme) and related estimators of a finite population mean (Hansen-Hovitz and Desraj estimators for a general sample size and Murthy's estimator for a sample of size two), Horvitz Thompson Estimator.

Unit III: Use of supplementary information for estimation: Ratio and regression estimators based on SRSWOR, method of sampling.

Unit IV: Cluster sampling with equal and unequal cluster sizes, two stage sampling with equal number of second stage Units.

Unit V: Double sampling estimating Strata sizes in ratio and regression method of estimation, Randomized response technique. (Warner's Model: Related and unrelated questionnaire methods.)

References:

1) Sukhatme et al: Sampling theory of surveys with applications.

2) Singh, D. and Chaudhari, F.S.: Theory and Analysis of Sample Survey designs.

3) Murthy, M.N.: Sampling theory and methods.

4) Des Raj and Chandak: Sampling Theory.

5) Cochram, W.G.: Sampling Techniques.

Course Outcome

After completing this course successfully, students would be able to

1. understand basic concepts of Sample Survey theory, sampling, and non-sampling error and randomized response technique.
2. Apply various sampling methods such as SRS, stratified and systematic sampling, sampling with varying probability, methods of drawing sample by PPS sampling method.
3. Use the method of using supplementary information in survey sampling by using ratio and regression method. They will be able to compare between ratio and regression method.

4. be equipped with the concepts of cluster sampling with equal and unequal cluster size and variances of the estimators. Concept of two stage sampling with equal first stage unit will be introduced to them.
5. use double sampling in stratification and ratio and regression method of estimation.

M.Sc. (Statistics)

Semester – I

DSE – 1

1. INDUSTRIAL STATISTICS

SYLLABUS

Unit I : Basic concept of process monitoring and control, process capability and process optimization. Review of control charts for attributes and variable data. O.C. and ARL of control charts. Moving average charts. Cusum & V-masks charts.

Unit II : Acceptance sampling plan for attributes inspection, single, double and sequential sampling plans and their properties. Plan for Inspection by variable (one & two sided specifications) Mil Std & I.S. plans. Continuous sampling plans of Dodge types, Wald-Wolfiwitz type and their properties. Bayesian sampling plans.

Unit III : Use of experiments is SPC, factorial experiments, fractional factorial designs, construction of such designs and anlaysis data.

Unit IV : Capability indices Cp, Cpk and Cpm, estimation, confidence intervals and tests of hypothesis relating to capability indices for normally distributed characteristics. Multivariate quality control.

Unit V : Quality Systems : ISO 9000 standards, QS 9000 standards, concept of six sigma and define-measure-analyze-Improve-Control Approach. Precision and accuracy in measurement systems. Estimation of measurement uncertainty. Total Quality management.

References:

1) Montgomery D.C (1985): Introduction to Statistical Quality Control, Wiley.
2) Montgomery D.C. (1985): Design and Analysis of Experiments, Wiley.
3) Wetherill G. B. (1977): Sampling Inspection & Quality Control, Halsted Press.
4) Wetherill G. B. & Brown D.W. (…..): Statistical Process Control, Theory and Practices, Chapman & Hall.
5) Logothetis N. (1992): Managing Total Quality, Prentice Hall of India
6) Oakland J.S. (1989): Total Quality Management; Butterworth Heinemann.
7) Mittog HJ, and Rinne H. (1993): Statistical Methods of Quality Assurance.

Course Outcome

After completing this course successfully, students would be able to

1. Use Quality systems ISO 9000 and QS 9000, Total Quality Management, PDSA cycle
2. know how to use various statistical tools such Design of Experiment for quality improvement
3. perform Process Capability Analysis, various Capability indices
4. use Taguchi philosophy – system, parameter and tolerance design

5. compute Signal to Noise Ratio, FMECA

M.Sc. (Statistics)
Semester – I
DSE – 1
2. TIME SERIES
SYLLABUS

Unit 1: Exploratory time series analysis, tests for trend and seasonality. Exponential and Moving average smoothing. Holt –Winters smoothing. Forecasting based on smoothing, adaptive smoothing. Time – series as a discrete parameter stochastic process. Auto covariance and autocorrelation functions and their properties.
Unit 2: Stationary processes: General linear processes, moving average (MA), auto regressive (AR), and autoregressive moving average (ARMA). Auto regressive integrated moving average (ARIMA) models, Box –Jenkins models Stationarity and inevitability conditions.
Unit 3: -Nonstationary and seasonal time series models:
ARIMA models, Identification techniques, Unit roots in time series, Forecasting ARIMA models, Seasonal ARIMA models, Regression with ARMA errors, Transfer function models (Time series regression)
Unit 4: Forecasting in time series models, Durbin-Levinson algorithm, innovation algorithm (without proof). Estimation of mean, auto covariance and autocorrelation functions, Yule-Walker estimation, Estimation of ARIMA model parameters, maximum likelihood method, large sample theory (without proofs). Choice of AR and MA periods, FPE, AIC, BIC, residual analysis and diagnostic checking.
Unit 5: Multivariate Time series model, VAR models, Vector ARMA models. Conditional heteroschedastic models, ARCH and GARCH, properties, examples, estimation & forecasting, extensions of ARCH & GARCH.

References:
1. Brockwell, P.J. and Davis, R. A. (2003). Introduction to Time Series Analysis, Springer
2. Chatfield, C. (2001). Time Series Forecasting, Chapmann & Hall, London
3. Fuller, W. A. (1996). Introduction to Statistical Time Series, 2nd Ed. Wiley.
4. Hamilton N. Y. (1994). Time Series Analysis. Princeton University press.
5. Box, G.E.P & Jenkins G.M (1976): Time Series Analysis – Forecasting & Control, Holden-Day,San Francisco .
6. Lutkepohl, H. and Kratzing, M. (Ed.) (2004). Applied Time Series Econometrics, Cambridge University Press.
7. Shumway, R. H.and Stoffer D. S. (2010). Time Series Analysis & Its Applications, Springer.
8. Tsay, R. S. (2010). Analysis of Financial Time Series, Wiley.
9. Montgomery, D.C & Johnson, L.A (1977): Forecasting and Time Series Analysis, McGraw Hill

Course Outcome

After completing this course successfully, students would be able to

1. describe and verify mathematical considerations for analyzing time series, including concepts of white noise, stationarity, autocovariance, autocorrelation

2. apply various techniques of time series models, including the seasonal autoregressive moving average (SARIMA) models, regression with ARMA models

3. apply various techniques for the modeling: including parameter estimation, assumption verification, and residual sequence diagnosis

4. verify the properties of linear predictor operator, and apply various linear forecasting techniques

5. describe and apply techniques of selected additional topics, such as spectral analysis, state space models, ARCH and GARCH, multivariate time series, principle component analysis, process control, and other topics.

6. Use R to construct time series models and conduct analysis.

M.Sc. (Statistics)
SE M E S T E R – I
D S E – 1
3. DEMOGRAPHY
S Y L L A B U S

Unit-I: Definition and Scope: Development of demography as a interdisciplinary discipline, Basic demographic concept and components of population dynamics. Coverage and content errors in demographic data, use of balancing equations and Chandrasekharan Deming formula to check completeness of registration data. Adjustment of age data. Use of whipple, Myer and UN indices. Population composition, dependency ratio.
Unit-II: Measure of Fertility: Stochastic models for reproduction, distribution of time to first birth, inter live birth intervals and of number of births (for both homogeneous and non-homogeneous groups of women), estimation of parameters, estimation of parity progression ratios from open birth interval data.
Unit-III: Measure of Mortality: Various measures of mortality, infant mortality rate, cause specific death rate and standardised death rates. Construction of a bridge life table. Distribution of life table functions and their estimation.
Unit-IV: Migration: Migration Rates and Ratios : Indirect measures of net-internal migration. National growth rate method. Stochastic models for migration and for Social and occupational mobility based on Markov chains. Estimation of Measures of Mobility.
Unit-V: Measurement of Population Change: Linear, Geometric, 2 exponential, Gompretz, Logistic Population growth models. Methods of population projection, use of Leslie matrix. Stable and Quasi Stable populations, intrinsic growth rate. Models for population growth and their fitting to population data. Stochastic models for population growth.

References:

1) Benjamin, B. (1969): Demographic Analysis, George, Allen & Unwin.
2) Cox P.R. (1970): Demography, Cambridge University Press.
3) Keyfitz N. (1977): Applied Mathematical Demography, Springer Verlag.
4) Spiegelman M. (1969): Introduction to Demographic Analysis, Harvard University Press.
5) Bartholomew D.J. (1982): Stochastic Models for Social Processes, John Wiley.

COURSE OUTCOME
After completing this course successfully, students would be able to

1. Understand the core social demographic variables, and how these variables influence population growth, composition, and structure
2. Use demographic tools in understanding public health issues Knowledge attitude and practices.
3. Discuss global demographic regimes and impact on public health.
4. Identify appropriate sources of data, perform basic demographic analyses using various techniques and ensure their comparability across populations.

M.Sc. (Statistics)
Semester – I
DSE – 1
4. REAL ANALYSIS & MEASURE THEORY
SYLLABUS

Unit – I: Real valued functions. Riemann and Riemann Stieltjes integral, Integration by Parts, mean value theorem. Elements of complex integration, Analytic function, definition of line integral, Cauchy integration formula, Residue theorem.

Unit – II: Matrix algebra: characteristic roots of real matrices, right and left characteristic vectors. Independence of characteristic vectors their multiplicities. Generalized inverse Definiteness of a real quadratic form Reduction to quadratic form.

Unit – III: Sets : Classes of sets, Sequences of sets, lim sup and lim inf of sequences of sets, field, s-field, s field generated by a class, Borel s- field. Set functions, additive set functions & their properties. Measure, Measure spaces, Measurable function, simple function, Integral of measure function w.r.t. measure.

Unit – IV: Sequences of measurable functions, convergence a.e. and in measure. Monotone convergence theorem, Fatous lemma, Dominated convergence theorem and their application.

Unit V:- − Construction of measures: Caratheodory construction, product measures, Fubini's theorem; definition of Lebesgue measure in R^n and its determination by translational invariance, Lebesgue integral on the line and Riemann integral, calculation of Lebesgue integral in R^n including the use of the substitution theorem but not its proof.

References :

1) T Apostol : Mathematical analysis

2) Churchill : Functions of a complex variable

3) M E Munroe : Introduction to measure and Integration

4) R Ash : Real analysis and Probability theory

5) A. K. Basu : Measure Theory and Probability

Course Outcome:

After completing this course successfully, students would be able to

1. work comfortably with classes of sets.

2. recognise the concept of Reimain Stieljes integral.

3. Enhance the knowledge regarding convergence its application.

4. understand matrix algebra.

5. acquire knowledge of complex integral.

M.Sc. (Statistics)

Semester – I

DSE – 1

Lab: Practical based on DSC-1 & DSC-2

COs:

After completing this course successfully, students would be able to perform/demonstrate/accomplish the following

1. Fitting of Binomial distribution and testing the goodness of fit.

2. Fitting of Poisson distribution and testing the goodness of fit.

3. Fitting of Negative Binomial distribution and testing the goodness of fit.

4. Fitting of Normal distribution and testing the goodness of fit.

5. Fitting of Truncated Binomial distribution and testing the goodness of fit.

6. Fitting of Truncated Poisson distribution and testing the goodness of fit.

7. Application of Chi Square distribution.

8. Application of t distribution.

9. Application of F distribution.

10. Computation of MLE by the method of scoring

M.Sc. (Statistics)

Semester – I

DSE – 1

Lab: Practical based on DSC-3 & DSE-1

COs:

After completing this course successfully, students would be able to perform/demonstrate/accomplish the following

1. Estimation of mean and variance in SRSWR and SRSWOR.
2. Comparison of systematic sapling with SRS and Stratified random sampling.
3. Selection of sample with unequal probabilities and unbiased estimation of mean and variance.
4. Des Raj estimator.
5. Horvitz Thompson estimator.
6. Ratio and Regression estimate of population mean and s.e.
7. Cluster sampling.
8. Two stage sampling.
9. Double sampling for estimating strata sizes in Ration and regression method of estimation.
10. Practical based on chosen DSE

M.Sc. (Statistics)
Semester – II
DSC – IV
TESTING OF HYPOTHESIS
SYLLABUS

Unit I: Test of Hypothesis, concept of critical region, test function, two kinds of errors, size function, power function, level, p-value concept, MP and UMP test in the class of size level test, Neyman-Pearson lemma, MP test for simple hypothesis against simple alternative hypothesis.

Unit II: UMP test for simple null hypothesis against one sided alternative and for one sided null and one-sided alternative in one parameter exponential family. Non-existence of UMP test for simple null against two-sided alternative in one parameter exponential family.

Unit III: Likelihood ratio test. Asymptotic distribution of LR test statistic (Without proof) Wald's test, Rao's score test, Pearson's chi-square test for goodness of fit. Bartlett's test for homogeneity of variances (without proof)

Unit IV: Sequential Testing, sequential probability ratio test. Relation among parameters. Application of SPRT for Binomial, Poisson, Normal distribution.

Unit V: Generalized Neyman-Pearson lemma (Statement only), unbiased test, UMPU test and their existence in case of exponential family, similar test and test with Neyman structure.

References:

1) Ferguson T.S.: Mathematical Statistics.
2) Goon, A.M., Gupta, Dasgupta: An Outline of Statistical Inference.

3) H.C. Saxena, Surendran : Statistical Inference : S. Chand.
4) Kale B.K. (1999): A First Course in Parametric Inference, Narosa Publishing House.
5) Lehman E.L. (1986): Testing Statistical Hypothesis, Student Edition.
6) Rao C.R. (1973): Linear Statistical Inference.
7) Rohatgi V. (1998): An Introduction to Probability and Mathematical Statistics, Wiley Eastern Ltd. New York.
8) Zack. (1971): Theory of Statistical Inference, John Wily.

Course Outcome:

COs:

After completing this course successfully, students would be able to

1. Apply Concept of hypothesis, be able to state appropriate null and alternative hypothesis.
2. calculate p – value and interpret it.
3. Understand Neyman-Pearson fundamental lemma, generalized NP lemma, MP and UMP test.
4. Understand MLR property, its applications, Karlin Rubin theorem
5. use Likelihood Ratio test, Wald test, Rao's score test, Bartlett's test.
6. Understand SPRT, OC and ASN function for SPRT.

M.Sc. (Statistics)

Semester – II

AEC

TEST OF SIGNIFICANCE

SYLLABUS

Test of significance for mean and Variance for normal distribution of real-life examples.

Test of significance based on Chi Square, t, F, inference drawing for real-life problems.

Course Outcome

COs:

After completing this course successfully, students would be able to

1. work comfortably with test of significance of mean.

2. recognise the concept of formulating hypothesis for different test of significance problems.

3. Enhance the knowledge regarding application of exact sample test.

4. understand variance.

5. acquire knowledge of testing goodness of fit.

M.Sc. (Statistics)
Semester – II
DSC – V
ADVANCED PROBABILITY THEORY
SYLLABUS

Unit I:
Definition of Probability: classical, relative, frequency approach, Axiomatic definition of probability. Probability measure on a 6 field. Probability space. Properties of Probability measure, conditional probability. Independence of events, mutual and pairwise independence, Bayes theorem and its applications, Boreal Connelly lemma (Zero-one law)

Unit II:
Distribution function and its properties, convergence of sequence of random variables. Convergence in probability, convergence almost sure, convergence in distribution, convergence in r th mean. Inter relations between different types of convergences. Cramer's theorem. Concept of Martingles.

Unit III:
Characteristic function, properties, necessary and sufficient condition for a function ϕ (t) to be characteristic function, Multivariate characteristic function. Inversion theorem, Uniqueness theorem. Necessary and sufficient condition for independence of random variables in terms of characteristic function. Hally Bray theorem continuity theorem for characteristic function.

Unit IV:
Weak Law of large numbers strong law of large numbers chabyshevs weak law of large numbers Khinchins weak law of large numbers, Necessary and sufficient condition condition for any sequence {Xn} to satisfy the WLLN. Kolmogorov's strong law of large no. (Statement only) Kolmogorov's inequality comparison between WLLN and SLLN.

Unit V:
Central Limit theorem – Demovr's Laplace Central limit theorem. Lindeberg Levy and Liapounoff's central limit theorem. Application of central limit theorem Multivariate central limit theorem and Relation between CLT and WLLN.

References:

1. V.K. Rohatgi: Introduction to probability theory.
2. Bhat B.R. : Modern probability theory
3. Basu A.K.: Measure theory and probability
4. Fisz Mi: Probability Theory and Mathematical Statistics.

Course outcomes
COs:

After completing this course successfully, students would be able to

1. Get a deeper understanding of the foundations of Probability Theory.
2. Study and use important theorems such as Borell - Cantelli Lemma, General Central Limit Theorems, Inversion Theorem, and also Characteristics Function
3. deal with 'Convergence In Distribution', 'Convergence Almost Sure', 'Convergence In Probability' and their inter-relationships.
4. Apply modes of convergence of sequence of r.v., the weak law of large number and strong law of large number and their utility in day-to-day life.

M.Sc. (Statistics)
Semester – II
DSC – VI
DESIGN OF EXPERIMENT
SYLLABUS

Unit I

Linear estimation, estimable function and conditions for estimability, Best linear unbiased estimator (BLUE), Linear models - Principle of least squares, Gauss – Markov Theorem, Analysis of variance: one way classification with equal and unequal number of observations per cell, two way classification with equal number of observations per cell (with and without interaction), Two way classification with unequal number of observations per cell without interaction model.

Unit II

Introduction to designed experiments, General block design and its information matrix, concepts of connectedness, balance and orthogonality. Review of elementary Designs (CRD, RBD and LSD), Missing plot technique in RBD and LSD with one and two missing values.

Unit III

Incomplete Block Designs, BIBD: Definition, properties and intra block analysis, Definitions and parametric relations of SBIBD, RBIBD, ARBIBD, PBIBD, Definition and analysis of Youden square design.

Unit IV

Analysis of Covariance of one way and two-way classified data. Definition and analysis of Split plot and strip plot designs.

Unit V

General factorial experiments, factorial effects, best estimates and testing the significance of factorial effects, Study of 2^3 and 2^4 factorial experiments in RBD. Confounding in factorial experiments, Complete and partial Confounding, concept of generalized interaction.

References

1) Alok Dey (1986): Theory of Block Designs – Wiley Eastern.
2) Das, M.N. and Giri, N. (1979): Design and Analysis of Experiments- Wiley Eastern.
3) Joshi,D.D. (1987) : Linear Estimation and Design of Experiments, Wiley Eastern.
4) Montgomery, D.C. (2017). Design and analysis of Experiments, 9th Edition. John Wiley & Sons.
5) Cochran, W. G. and Cox, G. M. (1957). Experimental Designs, John Wiley & Sons, NewYork.
6) Federer, W.T. (1955). Experimental Design: Theory and Applications. Oxford & IBH Publishing Company, Calcutta, Bombay and New Delhi.

Course Outcome

COs:

After completing this course successfully, students would be able to

1) study linear estimation, BLUE, one way and two-way classification with equal and unequal number of observations per cell.

2) use General block design and its information matrix, Missing plot technique in RBD and LSD with one and two missing values.

3) learn Analysis of Covariance of one way and two-way classified data. Definition and analysis of Split plot and strip plot designs.

4) learn the basic principles in the design of simple experiments. To learn different tests for comparing pairs of treatment means, factorial experiments, confounding, concepts of complete and partial confounding, generalized interaction, BIBD, PBIBD and Youden Square design.

M.Sc. (Statistics)

Semester – II

DSE – II

1. SURVIVAL ANALYSIS

SYLLABUS

UNIT I: Concept of censoring, order and random censoring (left and right), progressive censoring, survival function, density function, Life distribution, Exponential, Gamma, Weiball, Pareto, parametric inference, point estimation, confidence interval, Scores, test based on LR, MLE

UNIT II: Life tables, Failure rate, mean residual life and their elementary properties, Ageing classes IFR, IFRA, NBU, NBUE, HNBUE and their duals, Bathtube failure rate.

UNIT III: Estimation of survival function, Actuarial estimator, Kaplan, Meier estimator, Estimator under the assumption of IFR / DFR, test of exponentiality against non-parametric classes, Total time on test.

UNIT IV: Two sample problem: Gahen test, Long rank test, Mentel Haenszed test, Tarone Wane test, competing risk models, parametric and non-parametric inference for this models.

UNIT V: Semi parametric regression for failure rate – Cox's proportional hazards model with one and several covariates, related estimation and test procedure. Introduction to accelerated time models: Linear rank test Least squares

References:

1. Cox DlRl and Oakes D. (1984): Analysis of Survival Data Chapman and Hall, New York.
2. Elandt-Johnson R.E., Johnson N.L.: Survival Models and Data Analysis, John Wiley and Sons
3. Miller R.G (1981): Survival Analysis, John Wiley.

Course Outcome

After completing this course successfully, students would be able to

1. describe survival data and format it appropriately for analysis and understanding.
2. graph data, specify and fit proportional hazard models, check assumptions and compute hazard ratios.
3. Describe survival data, and the roles played by censoring, and survival and hazard functions
4. Format data appropriately for analysis, and understanding
5. Graph survival data, and the Kaplan – Meier curve
6. Specify and fit the Cox Proportional Hazards model
7. Describe the fully-extended Cox model

M.Sc. (Statistics)

Semester – II

DSE – II

2. ADVANCED SAMPLING THEORY

SYLLABUS

Unit-I: Two stage sampling with unequal number of second stage units. Issues in stratified sampling: Allocation problems involving several study variables, stratum boundary, determination problems. Double sampling.

Unit II: Introduction to Unified theory of finite population sampling, population, sample (ordered and unordered). Sample space, sampling design, parameters of sampling design & their properties sampling algorithm. Hanurao existence theorem. Parametric function, Estimability of parametric function. Classes of estimators (as suggested by Horvitz-Thompson)

Unis-III: Horvitz-Thompson estimator (HTE) of a finite population total/mean. Expression for variance of H-T estimator and its unbiased estimator. Concept of sufficiency in survey sampling and use of Rao-Blackwell theorem. Likelihood function. Generalization of H-T estimator, difference estimator.

Unit-IV: IPPS schemes of sampling due to Midzuno-sen, Brewer, Durbin and JNK Rao (Sample size 2 only). Rao-Hartley Cochran sampling. Schemes for sample size n with random grouping.

Unit-V: Super population models, design unbiasedness & model unbiasedness. Criteria for comparison prediction approach.

Reference:

(1) P.Mukhopadhyay: Theory and Methods of Survey Sampling.

(2) Daroga Singh and Choudhary: Theory and analysis of sample survey designs.

(3) Murthy M.N.: Sampling Theory and Methods. (4) Cochran W.G: Sampling Techniques.

(5) Des-Raj and Chandak: Sampling Theory

(6) Sukhatma P.V. & B.V.: Sampling Theory of Surveys with Applications.

(7) Cassel, Saranda! & Wretman: Foundation course in survey sampling for finite population.

Course Outcomes

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

1. Understand the non –existence of uniform estimators and repetitive surveys.
2. Apply the re-sampling techniques for variance estimation – independent and dependent random groups.
3. Understand the design-based estimation procedures and double sampling technique for stratification.

4. Understand the response and non- response techniques; Randomized Response Technique and a technique to predict non observed residue under design and model-based model.
5. Understand the model assisted sampling strategies; super population model.

M.Sc. (Statistics)

Semester – II

DSE – II

3. DATA MINING

SYLLABUS

Unit-I: Introduction to object-oriented programming concepts and design. Introduction to Web Programming (Simple example in Java) and the concept of byte codes.

Una-II: Review of classification methods from multivariate analysis; classification and decision trees. Clustering methods from data mining viewpoints.

Unit-III: Vector quantization, unsupervised learning from univariate and multivariate data, dimension reduction and feature selection.

Unit IV: Supervised learning from moderate to high dimensional input spaces, artificial neural networks and extensions of regression models, regression trees.

Unit-V: Introduction to database including simple relational databases, data ware houses, introduction to online analytical data processing. Data attributes, applications to electronic commerce.

References:

1) A.Borson and S.J.Smith (1997): Data Ware Housing, Data Mining and OLAP, McGraw Hill.
2) Breiman, J.H.Friedman, R.A.Olsher and C.J.Stone (1984): Classification and Regression Trees, Wordsworth & Brooks/Cole.
3) J.Han and M.Kamber (2000): Data Mining Concepts and Techniques, Morgan Gaufmann.
4) T.M. Mitchell (1997): Machine Learning McGraw Hill.
5) P.Naughton (1996): The Java Handbook, Tata McGraw Hill.
6) W.J.Savich (2001): Problem Solving with C++: The Object of Programming (3" Edn.) Addison Wesley, Longman.

Course outcome

After completing this course successfully, students would be able to

1. Evaluate and implement a wide range of emerging and newly-adopted methodologies and technologies to facilitate the knowledge discovery
2. Assess raw input data, and process it to provide suitable input for a range of data mining algorithms.
3. Discover and measure interesting patterns from different kinds of databases
4. Characterize and discriminate data summarization forms and determine data mining functionalities.

M.Sc. (Statistics)

Semester – II

DSE – II

4. Statistical Genetics and Bio-informatics

SYLLABUS

Unit-I: Basic biological concepts in genetics Mendel's law, Hardy Weinberg equilibrium. Mating tables, estimation of allele frequency (dominant / co dominant cases).

Unit-II: Approach to equilibrium for X-linked gene, Natural selection, mutation, genetic drift, equilibrium when both natural selection and mutation an operative.

Unit-III: Non-random mating, inbreeding, phenotypic assortative mating

Unit-IV: Analysis of family data, (a) Relative pair data, I. T. O matrices, identity of descent, (b) family data - estimation of segregation ratio under ascertainment bias, (c) Pedigree data Elston-Estewart algorithm for calculation of likelihoods. Linkage, Estimation of recombination fraction, inheritance of quantitative traits, models and estimation of parameters.

Unit-V: Sequence similarity, homology and alignment. Algorithms for (a) Pair wise sequence alignment, (b) multiple sequence alignment. Construction of Phylogenetic trees, UPGMA, Neighbour joining. maximum parsimony and maximum likelihood algorithms. Bayesian approach to Bio-informatics.

References:

(1) C.C.LI. (1976): First Course on Population genetics, Boxwood Press, California.

(2) W. J. Ewens (1979): Mathematical Population Genetics, Springer Verlag.

(3) T. Nagylaki (1992): Introduction to Theoretical Population genetics, Springer Verlag.

(4) R Durbin, S. R. Eddy, A. Krogh, G. Mitchinson (1998): Biological sequence Analysis Probabilistic Models of Proteins and Nucleic Acids.

Course Outcome

After completing this course successfully, students would be able to

1. be familiar with biological concepts in genetics which will be helpful for analyzing biological data.
2. gain expertise in analysis in analysis of family data.
3. draw inference of estimation of parameter and use of maximum likelihood of estimator for biological data.

PRACTICAL BASED ON DSC IV AND DSC V

COs:

After completing this course successfully, students would be able to perform/demonstrate/accomplish/solve the following

List of Problems

(1) Construction of MP test and power curve.

(2) MP test for families with MLR property

(3) Construction of UMP test and power curve.

(4) Sequential probability ratio test for normal distribution.

(5) Sequential probability ratio test for binomial distribution.

(6) Sequential probability ratio test for Poisson distribution

(7) Problems on likelihood ratio test.

(8) Construction UMPU test

PRACTICAL BASED ON DSC VI AND DSE II

COs:

After completing this course successfully, students would be able to perform/demonstrate/accomplish/solve the following

List of Practicals

(1) Elementary Designs (CRD, RBD, LSD)

(2) Missing Plot Techniques in RBD and LSD with one and w missing values.

(3) BIBD

(4) Youden Square Design

(5) Analysis of Covariance (one way, two way classification)

(6) Split plot design.

(7) 2^3 and 2^4 factorial experiments in RBD (without confounding)

(8) 2^3 and 2^4 factorial experiments in RBD (with confounding)

(9) Practicals on chosen DSE