STATISTICS HONOURS SYLLABUS

Semester	Paper Code and No	Courses	Credits	Marks
I	STA-UG-E101	Descriptive Statistics and Probability Theory - I	4	100
II	STA-UG-E201	Statistical Methods-I and Probability Theory-II	4	100
III	STA-UG-E301	Statistical Methods – II and Computer Programming.	4	100
IV	STA-UG-C401	Mathematical Methods and Operation Research	4	100
	STA-UG-C402	Survey Sampling and Experimental Design	4	100
V	STA-UG-C501	Statistical Inference	4	100
	STA-UG-C502	Predictive Models: Regression and Time Series	4	100
VI	STA-UG-C601	Numerical Methods and Statistical Quality Control	4	100
	STA-UG-C602	Official and Economic Statistics	4	100

STA-UG-E101: Descriptive Statistics and Probability Theory – I

Unit I: Types of data

Concepts of a statistical population and sample from a population; qualitative and quantitative data; nominal and ordinal data; cross sectional and time series data; discrete and continuous data; frequency and non-frequency data, primary and Secondary data, concept of questionnaire and schedule. Presentation of Data and Statistical Graphics: Construction of tables with one or more factors of classification. Diagrammatic and graphical representation of non-frequency data. Frequency distributions, cumulative frequency distributions and their graphical and diagrammatic representation-column diagram, histogram, frequency polygon and ogives. Stem and leaf chart. Box plot, scatter diagram for bivariate data.

Unit II: Analysis of Quantitative Data

Univariate data. Concepts of central tendency or location, dispersion and relative dispersion, quartiles, skewness and kurtosis, Absolute moments, factorial moments, Sheppard's corrections for moments for grouped data.

Analysis of Qualitative data: Consistency of data, Independence and Association of attributes, measures of association and Contingency, Pearson and Yule's measures, Goodman – Kruskal's gamma coefficient.

Unit III: Probability Theory

Random experiment: trial, sample point and sample space, event, algebra of events, and concepts of mutually exclusive and exhaustive events.

Definition of probability: classical and relative frequency approach. Axiomatic approach, merits and demerits of these approaches.(only general ideas to be given)

Discrete probability space, Properties of probability, Independence of events, Conditional probability, total and compound probability rules, Baye's theorem and its applications, Paradoxes in Probability.

Unit IV: Laboratory Session

1. Construction of frequency distribution from raw data.
2. Graphical representation of frequency distribution: histogram and frequency polygon.
3. Box and Whisker plot.
4. Stem and leaf display.
5. Computation of mean, median, mode and quartiles.
6. Computation of variance, standard deviation and mean deviation.
7. Computation of skewness and kurtosis.
8. Computation of Pearson's product moment correlation coefficient.
9. Computation of Spearman's rank correlation.
10. Computation of Kendall's Tau.
11. Computation of Pearson- Yule's Phi.
12. Computation of Goodman-Krushkal's gamma.
13. Drawing Ogive curves and graphically determine the median.
14. Identifying the mode with help of frequency polygon.
15. Computation of measures of skewness and kurtosis for a frequency distribution.

Note: Some of the experiments may be done using Excel.

Reference books:

1. A M Goon , M K Gupta, B Dasgupta . Fundamental of Statistics, Vol 1 World Press
2. F E Croxton, D J Cowden, S Klein. Applied General Statistics, Prentice Hall.
3. M R Spiegel. Theory and Problems of Statistics
4. P Mukhopadhyay. Mathematical Statistics, New Central Book Agency.
5. K L Chung. Elementary Probability Theory with Stochastic Process, Springer
6. W Feller. An Introduction to Probability Theory and its Applications, Wiley
7. Bhattacharjee, D. and Das, K.K., Statistical Graphics for everyone, South Asian Publishers Pvt. Ltd., New Delhi.
8. Das, K.K., and Bhattacharjee, D. 101 Graphical Techniques, Asian Books Private Limited, New Delhi.
9. Gupta, S.C. and Kapoor, V.K., Fundamental of Mathematical Statistics, 11th Edn,(Reprint), Sultan chand and Sons.

STA-UG-E201: Statistical Methods-I and Probability Theory-II

Unit I: Random Variables

Discrete and continuous random variables, p.m.f., p.d.f., c.d.f, and their properties. Expectations of random variables and its properties. Moments and Cumulants, moment generating function, cumulant generating function and probability generating function. Chebychev's inequality, weak law of large numbers.

Unit II: Standard Univariate Distributions

Standard discrete probability distributions- Degenerate, Binomial, Poisson, Geometric, Negative Binomial, Hyper geometric and Multinomial. Standard continuous probability distribution: Normal, Uniform, Exponential, Beta, Gamma, Cauchy, Laplace.
Bivariate Distributions: Discrete and continuous type, c.d.f, p.d.f., p.m.f., marginal and conditional distribution, independence, product moments, conditional expectation, moment generating function and its properties. Bivariate normal distribution and its properties.

Unit III: Correlation and Regression

Karl Pearson's Coefficient of correlation, lines of regression, Multiple of partial Correlation Coefficients, Spearman's rank Correlation and Kandalls's Tau Coefficient.

Unit IV: Laboratory Sessions

1. Simulating a Binomial distribution using individual trials.
2. Simulating a Binomial distribution using the c.d.f.
3. Simulating a Negative Binomial distribution using individual trials.
4. Simulating a Negative Binomial distribution using the c.d.f.
5. Simulating a Poisson distribution using individual trials.
6. Simulating a Poisson distribution using the c.d.f.
7. Simulating a Discrete Uniform distribution using individual trials.
8. Simulating a Poisson distribution using the c.d.f.
9. Simulating a Exponential distribution using the c.d.f.
10. Simulating a Poisson distribution using Box – Muller and Polar Transformation method.
11. Simulating a Gamma distribution.
12. Simulating a Bivariate Normal distribution using marginal and Conditional Distribution.
13. Computation of a regression coefficient and intersect for a line of regression.
14. Computation Multiple Correlation Coefficient.
15. Computation of Partial Correlation Coefficient.

Note: Some of the laboratory sessions be conducted using R.

Reference books:

1. A M Goon , M K Gupta, B Dasgupta. Fundamental of Statistics, Vol 1 World Press
2. F E Croxton, D J Cowden, S Klein. Applied General Statistics, Prentice Hall.
3. M R Spiegel. Theory and Problems of Statistics
4. P Mukhopadhyay. Mathematical Statistics, New Central Book Agency.
5. K L Chung. Elementary Probability Theory with Stochastic Process, Springer
6. W Feller. An Introduction to Probability Theory and its Applications, Wiley
7. Bhattacharjee, D. and Das, K.K., Statistical Graphics for everyone, South Asian Publishers Pvt. Ltd., New Delhi.
8. Das, K.K., and Bhattacharjee, D. 101 Graphical Techniques, Asian Books Private

Limited, New Delhi.
9. Gupta, S.C. and Kapoor, V.K., Fundamental of Mathematical Statistics, 11th Edn,(Reprint), Sultan chand and Sons.

STA-UG-C301: Statistical Methods – II and Computer Programming

Unit I: Basic Concept of Testing of Hypothesis

Null and Alternative hypotheses, Types of Errors, Critical Region, Level of Significance, Power and p-values, Exact tests of hypotheses under Normal set-up for a single mean, the equality of two means, a single variance and the equality of two variances, Test of Significance of sample correlation coefficient (null case) and tests of hypotheses for the equality of means and equality of variances of a bivariate Normal distribution.

Unit II: Sampling Distribution

Random sample, Estimate, Parameter and Statistic, Sampling distribution of a statistic, Statement and applications of $\chi 2$, t, and F distribution, large sample tests, and confidence interval of sample mean and proportion, C.L.T., Standard Error of Sample mean and Sample Proprtion; Distribution of Sample mean and Sample variance. Distribution of Sample Correlation Coefficient (Null Case).

Unit III: Programming in C

Introduction to C language, Structure of a C – Programme. Data type, variable declaration, inputs/ output, arithmetic, relational and logical operators, Control Statement, Arrays, functions, pointers, recursion.

Unit IV: Laboratory Session

1. Writing a C program to carry out computations such as conversion between Celsius and Fahrenheit, calculation of interest, and similar other examples.
2. Computation the mean and variance of given data.
3. Sorting given numbers in ascending or descending order.
4. Writing program using conditional statements.
5. Writing program using " WHILE( )…." Loop.
6. Writing program using " DO ( ) WHLE" loop.
7. Writing program recursive procedure.
8. Writing program using a single array.
9. Writing program using double array.
10. Writing program define a function
11. -15. Special practice session to find errors and to debug programs. Also few more programs converting pointers, recursion etc. be covered.

Note: laboratory session should aim at giving students a hands-on experience of programming in C language.

Reference books

1. Goon A.M., Gupta M.K. & Dasgupta B. (1994): An Outline of Statistical Theory (Vol-1 World Press
2. Johnson, N.I. & Kotz S. (1970): Distributions in Statistics, John Wiley
3. Ross S.M. (1972): Introduction to Probability Models, Academic Press
4. Mood A.M., Graybill F. & Boes D.C. (1974): An Introduction to the Theory of Statistics (3rd ed), McGraw Hill
5. Rao C.R. (1952): Advanced Statistical Methods in Biometric Research, John Wiley
6. Hogg R.V. & Craig A.T. (1978): Introduction to Mathematical Statistics
7. Rohatgi V.K. (1984): An Introduction to Probability Theory & Mathematical Statistics, John Wiley
8. Stuart G & Ord J.K. (1991): Advanced Theory of Statistics (Vol 2), Charles Griffin
9. Goon A. M., Gupta M. K. and Dasgupta B. (1997): Fundamentals of Statistics (V-1), World Press
10. Bhattacharya GK & Johnson R. A. (1977): Concepts & Methods of Statistics, John Wiley
11. Balaguruswamy, E. (2008): Progrqamming in ANSI C, 4th Edn., Tata McGraw Hill.
12. Kanetkar, Y. (2008): Let Us C. BPB Publication.
13. Rajaraman, V. (2007): computer Programming in C, Prentice Hall of India.

STA-UG-C401: Mathematical Methods and O.R.

Unit I: Algebra

Sets, binary relation, Definition and examples of groups, rings and fields, vector space with illustrations, sub spaces, span, base and change of base, dimensions, orthogonal vectors, orthogonal basis, Gram – Schmidt orthognalization process.

Unit II: Real Analysis

Representation of real numbers as point on the line and the set of real numbers as complete order field. Bounded and unbounded sets, neighbourhood, limit and interior points, supremum and infimum, open and closed sets, sequences and convergences, limits of some special sequences such as $r^n, \left(1+\frac{1}{n}\right)^n, n^{\frac{1}{n}}$ and Cauchy's general principle of convergence, Cauchy's theorem on limits, monotonic sequences, limit superior and limit inferior of a bounded sequence.
Infinite series, positive term series and their convergences, comparison tests, D' Alembert's ratio test, Cauchy's n^{th} root test Raabe's test, Gauss test and Maclaurine's integral test. Leibnitz test for the convergence of alternating series, absolute convergence and conditional convergence of series.

Unit III: Introduction to OR

Phases of OR, model building and various types of OR problems.
Linear programming: Models, graphical solutions, simplex methods, Big M method. Concept of duality, dual simplex techniques.

Unit IV: The transportation problems

North West corner rule, Vogel's approximation method and MODI's method to find optimal solution. The assignment problems and its solution techniques.

Network: Idea of Network node, activities, dummy activity, construction of networkdiagram.
C.P.M: Network scheduling using C.P.M: determination of different types of floats and slacks, determination of Critical path.
PERT: Basic definition of PERT and its usefulness; brief idea of optimistic time, pessimistic time and most likely time (PERT calculations are not required)

Reference books

1. Hadley G. (1995): Linear Algebra, Addison Wesley/ Narosa
2. Rao A.R. & Bhimasankaran P. (1996): Linear Algebra
3. Narayan Shanti. Mathematical Analysis, S Chand and Co
4. Rudin W. Principles of Mathematical Analysis, McGraw Hill
5. Apostol T.M., Mathematical Analysis, Narosa Publishing House
6. Malik S.C., Mathematical Analysis
7. Deshpande J.V., Text Book of Mathematical Analysis, Tata McGraw Hill
8. H A Taha. Operations Research, Macmillan Publishing
9. F S Hiller , G J Libermann. Introduction to Operations Research, McGraw Hill

STA-UG-C402: Survey Sampling and Experimental Design

Unit I: Sample survey

Concepts of population and sample, complete enumeration vs sampling, principle and orthogonal aspects of sample survey, sampling errors and non-sampling errors and their sources. Probability sampling. Simple random sampling with and without replacement for the estimation of mean total and proportion.

Unit II: Stratified Random Sampling

Different allocations, proportional and Neyman allocation. Linear and circular systematic sampling, estimation of mean and variance, performance of systematic sampling in populations with linear trend.

Unit III: Experimental Designs

Role, historical perspective, terminology, basic principles. Basic designs-completely Randomized Design (CRD), Randomized Block Design (RBD), Latin Square Design (LSD)-Layout, model and statistical analysis, relative efficiency, analysis with missing observations. Factorial Design (2^2 $Case$).

Unit IV: Laboratory Session

1. Selecting a simple random sample with replacement using random numbers.
2. Selecting a simple random sample without replacement using random numbers.
3. Selecting a simple random sample without replacement using discrete uniform distribution.
4. Selecting a simple random sample with replacement using discrete uniform distribution.
5. Selecting a stratified random sample with replacement using proportional allocation.
6. Selecting a stratified random sample with replacement using Neyman allocation.

7. Estimating the population mean under SRSWOR, Stratified random sampling (proportional allocation) stratified random sampling (Neyman allocation).
8. Selecting a systematic sample and estimating the population mean. Computing the sampling variance of the estimation.
9. Analysis of CRD.
10. Analysis of RBD.
11. Analysis of LSD.
12. Relative efficiency of RBD with respect to CRD.
13. Relative efficiency of RBD with respect to CRD and RBD.
14. Analysis of RBD with one missing value.

Reference books

1. A M Goon, M K Gupta, B Dasgupta. Fundamental of Statistics, Vol & 2, World Press
2. A M Goon, M K Gupta, B Dasgupta. Outllines of Statistics, Vol 1& 2, World Press
3. P Mukhopadhyay. Applied Statistics, New Central Book Agency
5. Sukhatme et al. Sampling theory and methods
6. P Mukhopadhyay. Theory and Methods of Survey Sampling, New Central Book Agency
7. Gupta, S.C. and Kapoor, V.K., Fundamental of Applied Statistics, 11th Edn,(Reprint), Sultan Chand and Sons.

STA-UG-C501: Statistical Inference

Unit I: Point Estimation

Parameter space, sample space, requirements of good estimators-Unbiasedness, Consistency, Efficiency, Sufficiency and Completeness, minimum variance unbiased (MVU) estimators. Frechet – Rao – Cramer inequality and itts applications, Rao – Blackwell theorem. Method of estimation – maximum likelihood, moments, minimum variance, minimum Chi – Square, list squares. Properties of maximum likelihood estimators (without proof).

Unit II: Interval Estimation

Confidence intervals, pivotal quantity method of contracting confidence intervals, confidence intervals for parameters of one and two normal populations, confidence interval for proportions.

Unit III: Testing of Hypothesis

Simple and Composite hypotheses, most powerful test, randomized test, Neyman Pearson Lemma and applications, exact tests for parameters for normal populations, Likelihood Ratio test, properties of LR tests (without proof).

Unit IV: Non-Parametric test: tests for goodness of fit (one sample KS, chi-square), randomness (one sample run test) one sample location test (sign and signed rank test), two sample location tests (Mann Whitney, Wilcoxon rank sum, median), general two sample tests (run and two sample KS).

Reference books

1. A M Goon, M K Gupta, B Dasgupta. Fundamental of Statistics, Vol 2, World Press
2. P Mukhopadhyay. Mathematical Statistics, New Central Book Agency.
3. Gupta, S.C. and Kapoor, V.K., Fundamental of Mathematical Statistics, 11th Ed, (Reprint), Sultan Chand and Sons.

STA-UG-C502: Predictive Models: Regression and Time Series

Unit I: General linear Models

Gauss -Markov linear models, method of least squares, estimability and BLUE, Gauss-Markov theorem, estimation of error variance.

Analysis of Variance: Analysis of one-way classified data, two –way classified data with one observation per cell, two-way classified data with equal observation per cell, two-way classified data with unequal number of observations per cell for fixed effects model.

Unit II: Regression Models

Simple linear regression- fitting, tests and confidence intervals for parameters of the model, test for lack of fit. Multiple linear regression- fitting, tests and confidence intervals for parameters of the model.

Unit III: Time Series

Introduction, decomposition of a time series, different component with illustrations. Measurement of trend- Graphical method, method of semi- averages, method of fitting curves (straight line, polynomials, growth curves modified- exponential curves, Gompertz curve and Logistic curves). Method of moving averages, Measurement seasonal variation- method of simple averages, ratio to trend method, ratio to moving averages method and link relative method, residual method.

Unit IV: Laboratory Session

1. Analysis of two-way classified data with interaction
2. Pair wise comparisons of blocks and treatments. Tukey and Scheffe methods.
3. Fitting a simple linear regression. Estimation, testing and confidence intervals for the regression coefficients.
4. Test for lack of fit in simple linear regression.
5. Fitting multiple linear regression, estimation, testing and confidence intervals for the parameters of the models. Fitting may be done by shifting and origin methods.
6. Decomposition of time series in different components
7. Fitting curves to time series- 1
8. Fitting curves to time series- 2
9. Computation of moving average
10. Measurement of seasonal variation.
11. Residual analysis in time series
12. Multiple linear regression using R functions
13. Time series -1 using R packages.
14. Time series -2 using R packages.

Note: Some of the laboratory session be conducted using R functions and packages.

Reference books

1. Gupta, S.C. and Kapoor, V.K., Fundamental of Applied Statistics, 11th Edn, (Reprint), Sultan Chand and Sons.
2. P Mukhopadhyay. Applied Statistics, New Central Book Agency.
3. C Chatfield. The Analysis of Time Series – An Introduction , Chapman and Hall
4. Draper, N.R. and Smith, H. (1998): Applied Regression Analysis, 2nd Edn. John Wiley & Sons.
5. Montogomery, D.C., Peck, E.A. and Vining, G.G. (2006): Introduction to linear Regression Analysis, 4th Edn. John Wiley and Sons.

STA-UG-C601: Numerical Methods and Statistical Quality Control

Unit I: Numerical Analysis

Operators: Shift operators(E), forward difference(Δ), backward difference(∇), central difference(δ) and inter-relations between them. Finite difference of order n, divided differences of order n and interpolation. Newton's forward, backward and divided difference interpolation formula. Lagrange's interpolation formula. Central difference formula: Gauss, Everett and Stirling's formulae.

Unit II: Numerical Integration

Trapezoidal Rule, Simpson's 1/3 rd and 3/8 th rule, Weddle's rule, Euler-Maclaurin's formula. Solution of nonlinear equations- Bisection, Graphical method, Regula Falsi Method, Method of Iteration, Newton-Raphson methods.

Unit III: Statistical Quality Control

Introduction: Concepts of Quality and Quality Control, Process Control and Product Control. Process Control: Control Charts and their uses, Choice of Subgroup sizes, Construction of control charts by attributes (p, c, np) (including unequal subgroup size) and variables ($-X$, R). Interpretation of nonrandom patterns of points.

Product Control: Producer's Risk, Consumer's Risk, Acceptance Sampling Plan, Single and Double sampling plans by attributes, their OC, ASN (and ATI), LTPD and AOQL. Single sampling plan for inspection by variables (one-sided specification, known and unknown σ cases), Use of IS plans and tables.

Unit IV: Laboratory Session

1. Computation of finite differences.
2. Interpolation using Newton's – forward and Backward difference formulae.
3. Interpolation using Lagrange's interpolation formula.
4. Interpolation using central difference formula.
5. Numerical integration using Trapezoidal, Simpson's 1/3rd and Simpsons 3/8 rules.
6. Solving and algebraic and transcendental equations using bisection, regular – falsi and Newton – Raphson methods.

7. Drawing of X - charts.
8. Drawing of R – charts.
9. Drawing of np – charts.
10. Drawing of c – charts.
11. Drawing of OC curve for single and double sampling plans for attributes.
12. AOQ and ATI curves.
13. Writing C program for numerical integration and solving equations.

Note: Some of the laboratory session in numerical metods may be conducted using MATLAB.

Reference books

1. H C Saxena. Calculus of Finite Difference.
2. Bradie, B. (2006): A Friendly introduction to Numerical Analysis, Pearsons Education, India.
3. Sastry, S.S. (2000): Introductory Methods of Numerical Analysis, 3rd Edn, Prentice Hall of India Pvt. Ltd., New Delhi.
4. Goon A. M., Gupta M. K., Dasgupta B. (2001): Fundamentals of Statistics (V-2), World Press
5. Duncan A.J. (1953): Quality Control and Industrial Statistics, Richard D Irwin
6. Cowden D.J. (1957): Statistical Methods in Quality Control, Prentice Hall
7. Grant E.L. & Leavenworth (1964): Statistical Quality Control, McGraw Hill
8. Bowley A.H. & Goode H.P. (1952): Sampling Inspection by Variables, McGraw Hill
9. Ekambaram S. K. (1960): The Statistical Basis of Quality Cont. Charts, Asia Publishing House
10. Montgomery D.C. (1985): Introduction to Statistical Quality control, John Wiley
11. IS2500 Part I and Part II
12. Bureau of Indian Standards (1994): Handbook on Statistical quality Control.

STA-UG-C602: Official and Economic Statistics

Unit I: Study of population Census in India

(Highlights of last two censuses); India as a whole & the state of Sikkim. The Central and State Government organizations, the functions of the Central Statistical Organization (CSO), the National Sample Survey Organization (NSSO), Functions of National Statistical Commission (NSC) Simple Registration system.

National Income statistics: Income, expenditure and production approaches. Their applications in various sectors in India.

Unit II: Demography

Source of demographic data, vital rates, crude birth rates, general fertility rates, age specific fertility rates, total fertility rate, gross and net reproduction rates, crude death rate, age specific mortality rates, infant mortality rate, complete life table- concepts, different columns & their relationships; Standardised rates, idea of abridged life table without construction, Population Projection by Logistic curve and its fitting, Basic idea of Stable and Stationary population.

Unit III: Index Numbers

Definition, construction of index numbers by different methods, problems faced in their construction, criterion of a good index number, Tests for index numbers, time reversal, factor reversal. Errors in the construction of index numbers. Chain and fixed based index numbers. Base shifting, splicing and deflating index numbers. Cost of living index numbers- construction and uses.Wholesale price index and index of industrial production.

Unit IV: Demand Analysis

Theory of consumption and demand, demand function, elasticity of demand, determination of elasticity of demand by family budget method, Lorentz curve and Gini's coefficient, Engel's law and Engel's curve, Pareto's law of income distribution.

Reference books

1. C.S.O. (1984): Statistical System in India.
2. Goon A. M., Gupta M. K, and Dasgupta. B. (2001): Fundamentals of Statistics (V-2),World Press
3. Yule G.U. & Kendall M.G. (1953): An Introduction to the Theory of Statistics, C.Griffin
4. Kendall M.G. & Stuart A. (1966): Advanced Theory of Statistics (Vol 3), C.Griffin
5. Croxton F.E., Cowden D.J. & Klein (1969): Applied General Statistics, Prentice Hall
6. Mudgett B.D. (1951): Index Numbers, John Wiley
7. Allen R.G.D. (1975): Index Numbers in Theory and Practice, Macmillan
8. Mukhopadhyay P. (1999): Applied Statistics
9. Johnston J. & Dinardo J. (1997): Econometric Methods, McGraw Hill
10. Gupta, S.C. and Kapoor, V.K., Fundamental of Mathematical Statistics, 11^{th} Edn, (Reprint), Sultan Chand and Sons.