ST. THOMAS' COLLEGE (AUTONOMOUS)
THRISSUR, KERALA – 680001

Affiliated to University of Calicut
Nationally reaccredited with 'A' Grade

CURRICULUM AND SYLLABUS
FOR
UNDERGRADUATE PROGRAMME IN STATISTICS

UNDER CHOICE BASED CREDIT AND SEMESTER SYSTEM
(w.e.f. 2020 Admission onwards)

ST. THOMAS COLLEGE (AUTONOMOUS),
THRISSUR
OUTCOME BASED EDUCATION

UG: Program Outcomes

At the end of an Undergraduate Program at St. Thomas College (Autonomous), a student would have obtained the following:

PO1:	**Critical Thinking:** Ability to take informed actions after identifying the assumptions that frame our thinking and actions, checking out the degree to which these assumptions are accurate and valid, and looking at our ideas and decisions (intellectual, organizational, and personal) from different perspectives.
PO2:	**Effective Communication:** Ability to speak, read, write and listen clearly in person and through electronic media in English and in one Indian language, and make meaning of the world by connecting people, ideas, books, media and technology.
PO3:	**Effective Citizenship:** Ability to demonstrate empathetic social concern and equity-centered national development, and the ability to act with an informed awareness of issues and participate in civic life through volunteering.
PO4:	**Environment and Sustainability:** Ability to understand the issues of environmental contexts and sustainable development.
PO5:	**Ethical Living:** Ability to recognize different value systems including your own, understand the moral dimensions of your decisions, and accept responsibility for them.
PO6:	**Social Interaction:** Ability to elicit views of others, mediate disagreements and help reach conclusions in group settings.
PO7:	**Problem Solving and Analytical Skills:** Ability to think rationally, analyze situations and solve problems adequately.

UG: Program Specific Outcomes

PSO 1	Demonstrate the ability in collection, presentation, analysis and interpretation of data.
PSO 2	Understand and solve problems in probability, statistical distributions, correlation and regression.
PSO 3	Understand and apply the theories of classical inference involving estimation of parameters and testing of hypotheses.
PSO 4	Understand and apply the techniques used in design of experiments, statistical quality control, time series and population studies

Cognitive Levels

A. Remember
B. Understand
C. Apply
D. Analyze
E. Evaluate
F. Create

B.Sc. DEGREE PROGRAMME COURSE STRUCTURE

Semester	Paper Code	Paper title	Instructional Hours per Week	Credit	Duration of Exam
I	STA1B01	OFFICIAL STATISTICS AND PROBABILITY	4	4	2.5
II	STA2B02	BIVARIATE RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS	4	4	2.5
III	STA3B03	STATISTICAL ESTIMATION	5	4	2.5
IV	STA4B04	TESTING OF HYPOTHESIS	5	4	2.5
V	STA5B05	MATHEMATICAL METHODS IN STATISTICS	5	4	2.5
	STA5B06	SAMPLE SURVEYS	5	5	2.5
	STA5B07	LINEAR REGRESSION ANALYSIS	5	4	2.5
	STA5B08	STATISTICAL COMPUTING	5	5	2.5
	STA5D01 STA5D 02 STA5D 03	ECONOMIC STATISTICS QUALITY CONTROL BASIC STATISTICS	3	3	2
		PROJECT WORK	2	--	--
VI	STA6B09	TIME SERIES AND INDEX NUMBERS	5	4	2.5
	STA6B10	DESIGN OF EXPERIMENTS	5	5	2.5
	STA6B11	POPULATION STUDIES, ACTUARIAL SCIENCE AND VITAL STATISTICS	5	4	2.5
	STA6B12	OPERATIONS RESEARCH AND STATISTICAL QUALITY CONTROL	5	4	2.5
	STA6B13	PROJECT WORK	2	2	--
	STA6B14(E)	PROBABILITY MODELS AND RISK THEORY	3	2	2

	STA6B15(E) STA6B16(E)	STOCHASTIC PROCESSES RELIABILITY THEORY			

Pattern of Question Papers

Question paper type 1 (for 80 marks)
Scheme of Examinations:
The external QP with 80 marks and Internal examination is of 20 marks. Duration of each externalexamination is 2.5 Hrs. The pattern of External Examination is as given below. The students cananswer all the questions in Sections A&B. But there shall be Ceiling in each section.

Section A
Short answer type carries 2 marks each - 15 questions — Ceiling - 25

Section B
Paragraph/ Problem type carries 5 marks each - 8 questions — Ceiling - 35

Section C
Essay type carries 10 marks (2 out of 4) — 2×10=20

Question paper type 1 (for 60 marks)

Scheme of Examinations:
The external QP with 60 marks and internal examination is of 15 marks. Duration of each external examination is 2 Hrs. The pattern of External Examination is as given below. The students cananswer all the questions in Sections A&B. But there shall be Ceiling in each section.

Section A
Short answer type carries 2 marks each - 12 questions — Ceiling - 20

Section B
Paragraph/ Problem type carries 5 marks each - 7 questions — Ceiling - 30

Section C
Essay type carries 10 marks (1 out of 2) — 1×10 =10

Questions in each part should be equally distributed among the various modules of the syllabus

SEMESTER I

STA1B01: OFFICIAL STATISTICS AND PROBABILITY (72 Hours)

CO Statement

I. CO1: Recognize important Statistical organizations in India and apply the measures of central tendency, measures of location, measures of dispersion and measures of shape.
II. CO2: Distinguish between correlation and regression in terms of explaining the relationship between two or more variables.
III. CO3: Apply the principle of least squares in fitting linear and non-linear curves.
IV. CO4: Understand the concepts of random experiments and definitions of probability.
V. CO5: Understand and apply the theorems and results to compute the probabilities of events.
VI. CO6: Explain discrete and continuous random variables, their probability functions and properties.

Module 1: Statistical organizations in India-MOSPI; CSO, NSSO, DES; Roles functions and activities of CSO, NSSO and DES; Measures of central tendency – Arithmetic Mean, Geometric Mean, Harmonic Mean, Median, Mode, Percentiles, Quartiles; Measures of dispersion- Variance, Standard deviation, Mean deviation, Quartile deviation, Coefficient of variation; moments, skewness, Kurtosis.
(I: A, B, C, D, E) 20 hours

Module 2: Fitting of straight line, parabola, exponential, polynomial, (least square method), correlation-Karl Pearson's Correlation coefficient, Rank Correlation-Spearman's rank correlation co-efficient, Partial Correlation, Multiple Correlation, Scatter diagram, regression, two regression lines, regression coefficients.
(II, III: A, B, C, D, E)
17 hours.

Module 3: Experiment, Non- random and Random experiments, Sample space, event, classical definition of probability, statistical regularity, field, sigma field, axiomatic definition of probability and simple properties, addition theorem (two and three events), conditional probability of two events, multiplication theorem, independence of events-pair wise and mutual, Bayes theorem.
(IV:A,B,C) 25 hours

Module 4: Random variable-discrete and continuous, probability mass function (pmf) and probability density function (pdf)-properties and examples, Cumulative Distribution function and its properties, change of variable (univariate case).

(V, VI: A, B, D, E)

10 hours

References

1. V. K. Rohatgi, An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern.
2. S.C.Gupta and V. K. Kapoor, Fundamentals of Mathematical Statistics, Sultan Chand and Sons
3. A.M. Mood, F.A. Graybill and D C Bose, Introduction to Theory of Statistics, McGraw Hill
4. John E Freund, Mathematical Statistics (6th edn), Pearson Edn, NewDelhi.
5. Statistical system in India (CSO), 1995.

SEMESTER II

STA2B02: BIVARIATE RANDOM VARIABLE AND PROBABILITY DISTRIBUTIONS (72 Hours)

CO Statement

I. CO1: Understand the idea of expectations and there by obtaining the moments
II. CO2: Understand and apply the concepts of bivariate random variables and their probability distributions
III. CO3: Describe the shape of frequency curve and compute the conditional mean and variance using mathematical expectation.
IV. CO4: Determine the nature of relationship and the independence of bivariate random variables using mathematical expectation.
V. CO5: Explain standard discrete probability distributions.
VI. CO6: Understand and apply the law of large numbers.

Module 1:Mathematical expectations-definition, raw and central moments (definition andrelationships), moment generating function and properties, characteristic function (definition and basic properties).

(I:A, B) 20 hours

Module 2:Bivariate random variable, joint pmf and joint pdf, marginal and conditionalprobability, independence of random variables.

(II: A, B, C) 15 hours

Module 3: Skewness and kurtosis using moments, bivariate case-conditional mean andvariance, covariance, Karl Pearson Correlation coefficient, independence of random variables based on expectation.

(III, IV:A, B, C) 12 hours

Module 4: Degenerate distribution, Standard discrete distributions-Bernoulli, Binomial,Poisson, Geometric, negative binomial, Hyper geometric (definition, properties and applications), Uniform. Limit Theorems: Chebyshev's inequality, Convergence in probability, Convergence in distribution (definition and example only), weak law of large numbers (iid case), Bernoulli's law of large numbers.

(V, VI: A, B, C) 25 hours

References

1. V. K. Rohatgi, An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern.
2. S.C.Gupta and V. K. kapoor Fundamentals of Mathematical Statistics, Sultan Chand and Sons
3. A.M. Mood, F.A. Graybill and D C Bose, Introduction to Theory of Statistics, McGraw Hill
4. John E Freund, Mathematical Statistics (6th edn), Pearson Edn, NewDelhi

SEMESTER III

STA3B03: STATISTICAL ESTIMATION (90 Hours)

CO Statement

I. CO1. Explain standard continuous distributions and their applications in real-life situations.
II. CO2. Establish and apply Lindberg- Levy central limit theorem for i.i.d case.
III. CO3. Distinguish between population and sample, and understand the concept of the sampling distribution.
IV. CO4. Establish and explain t, chi square and F distributions and their properties.
V. CO5. Explain the properties of point estimation and apply the methods of point estimation.
VI. CO6. Discuss the methods of interval estimation and construct confidence interval for mean and variance.

Module 1: Continuous type-Uniform, exponential, gamma, Beta, Normal (definition, properties and applications), Lognormal, Pareto and Cauchy (Definition only). Central limit theorem (Llindberg- Levy-iid case).

(I, II: A,B, C, D)
20 hours

Module 2: Sampling distributions: Parameter, Statistic, standard error, Sampling from normal distribution: distribution of sample mean, sample variance, chi-square, students t distribution, and F distribution (definition, derivation, property and relationships).

(III, IV: A, B, C)
15 hours

Module 3: Estimation of Parameter: Point Estimation. Desirable properties of a good estimator, unbiasedness, consistency, sufficiency, Fisher - Neyman factorization theorem (Statement and application only), efficiency, Cramer - Rao inequality; Methods of Estimation- method of maximum likelihood, method of moments, Bayesian estimation method. **(V: A, B, C)** 40 hours.

Module 4: Interval Estimation: Large sample confidence interval for mean, equality of means, equality of proportions. Derivation of exact confidence intervals for means, variance and ratio of variances based on Normal, t, chi square distribution and F distribution**.**

(VI: A, B, C)
15 hours

References

1. V. K. Rohatgi, An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern.
2. S.C.Gupta and V. K. Kapoor. Fundamentals of Mathematical Statistics, Sultan Chand and Sons.
3. A.M. Mood, F.A. Graybill and D C Bose, Introduction to Theory of Statistics, McGraw Hill.
4. John E Freund, Mathematical Statistics (6th edn), Pearson Edn, NewDelhi.

SEMESTER IV

STA4B04: TESTING OF HYPOTHESES (90 Hours)

CO Statement

I. CO1: Understand the concepts of testing of hypotheses and compute the probabilities of two types of errors.
II. CO2: Illustrate the concepts of most powerful tests and SPRT.
III. CO3: Understand and apply the statistical tests for means and proportions.
IV. CO4: Explain and apply tests based on F and Chi Square distribution
V. CO5: Discuss the concepts of Non parametric tests.
VI. CO6: Identify and apply non parametric tests for suitable situations.

Module 1: Testing of Hypotheses; concept of testing hypotheses, simple and composite hypotheses, null and alternative hypotheses, type I and type II errors, critical region, level of significance, power of test. Most powerful tests Uniformly most powerful test, Neyman Pearson Lemma (statement only). Sequential sampling and SPRT (Basic concepts only).

(I, II: A, B, C)
20 hours

Module 2: Large sample tests concerning mean, equality of means, proportions, equality of proportions. Small sample tests based on t distribution for mean, equality of means and paired t test, one-way ANOVA. (Include real life applications and practical problems).

(III:A, B, C)
30 hours

Module 3: Tests based on F distribution. Tests based on chi square distribution – Test for the significance of population variance, goodness of fit and for independence of attributes. Test for correlation coefficients. (Include real life applications and practical problems).

(IV: A, B, C)
20 hours.

Module 4: Non parametric tests - advantages, disadvantages; Kolmogorov - Smirnov test; one sample and two sample sign tests; Wilcoxon signed rank test; Median test; Mann Whitney test; Kruskal Wallis test and test for randomness (run test). (Include real life applications and practical problems).

(V: A, B, VI: A, B, C)
20 hours

References

1. V. K. Rohatgi, An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern.

2. S.C.Gupta and V. K. Kapoor, Fundamentals of Mathematical Statistics, Sultan Chand and Sons

3. A.M. Mood, F.A. Graybill and D C Bose, Introduction to Theory of Statistics, McGraw Hill

4. John E Freund, Mathematical Statistics (6th edn), Pearson Edn, NewDelhi

SEMESTER V

STA5B05: MATHEMATICAL METHODS IN STATISTICS (90 Hours)

CO Statement

I. CO1: Explain real number system and its properties.
II. CO2: Explain the concept of sequences and related theorems.
III. CO3: Describe infinite series and its convergence.
IV. CO4: Discuss continuity and uniform continuity of real valued functions and prove associated theorems.
V. CO5: Explain differentiation and supporting results.
VI. CO6: Explain Riemann integrability and fundamental theorems on integral calculus.

Module 1 : Real Number system: Mathematical induction, order properties of real number, Bernoulli, Cauchy, triangle inequality, absolute value, Completeness property-suprema&infima, Archimedean property, Density theorem, nested interval property.
(I: A, B)
20 hours

Module 2: Sequences: Limit, limit theorems, Squeeze theorem, convergence of sequence, root test and ratio test, monotone convergence theorem, subsequence and Bolzano-Weierstrass theorem, Cauchy criterion, limits of functions, limit theorems of functions. Infinite series and its convergence, Ratio test and Root test.
(II, III: A, B)
25 hours

Module 3: Continuous functions: Definition, Boundedness theorem, Maximum minimum theorem, Location of roots theorem, Intermediate value theorem, uniform continuity, Differentiation, Interior extremum theorem, Rolle's theorem, Mean value theorem, Taylor's theorem.
(IV, V: A, B) 25 hours

Module 4: Riemann Integration: Definition, Integrability criteria, integrability of continuous and monotone functions, properties of integrals, first and second fundamental theorems on integral calculus.
(VI: A, B) 20 hours

References

1. Malik S.C. and SavithaArora, Real Analysis, New Age International 2. Robert G Bartle, Real Analysis, Wiley
3. Shanti Narayanan, Elements of Real Analysis

STA5B06: SAMPLE SURVEYS (90 Hours)

CO Statements

I. CO1. Compare census and sampling and discuss the organization and execution of sample surveys and associated errors.
II. CO2. Design a questionnaire.
III. CO3. Explain the methods for simple random sampling and estimate the population mean, population total and their variances using simple random sampling methods.
IV. CO4. Explain stratified and systematic sampling methods and estimate the population mean, population total and their variances using these methods.
V. CO5. Describe cluster sampling and estimate the population mean, population total and their variances using cluster sampling methods.
VI. CO6. Compare stratified, systematic and cluster sampling with simple random sampling.

Module 1: Census and Sampling, principal steps in sample survey-probability sampling, judgment sampling, organizaion and execution of large sample surveys, sampling and non sampling errors, preparation of questionnaire.

(I, A, B, II,: A,B, F)
20 hours .

Module 2: Simple random sampling with and without replacement- methods of c collecting simple random samples, unbiased estimate of the population mean and population total-their variances and estimate of these variances-simple random sampling for proportions.

(III: A, B) 20 hours.

Module 3: Stratified random sampling: estimation of population mean and total, proportional and Neymann allocation of sample sizes-cost function-optimum allocation considering cost comparison with simple random sampling. Systematic Sampling: Linear and circular systematic sampling, comparison with simple random sampling.

(IV, VI: A, B) 30 hours.

Module 4: Cluster sampling: Clusters with equal sizes-estimation of the population mean and total, comparison with simple random sampling, two stage cluster sampling-estimate of variance of population mean.

(V,VI: A, B) 20 hours

References

1. Murthy M N, Sampling theory and methods, Statistical Publishing society, Calcutta
2. Daroja Singh and F S Chaudhary, Theory and Analysis of Sample Survey Designs, Wiely Eastern Limitted
3. Cochran W.G, Sampling Techniques, Wiely Eastern

STA5B07: LINEAR REGRESSION ANALYSIS (90 Hours)

CO Statement

I. CO1. Discuss the fundamentals of regression and model building.
II. CO2. Construct simple linear regression model, estimate its parameters and test their significance.
III. CO3. Describe and apply interval estimation of simple linear regression parameters and explain the method of maximum likelihood for estimating the parameters.
IV. CO4. Construct multiple linear regression model, estimate its parameters and test their significance.
V. CO5. Explain and apply residuals and residual plots for model adequacy checking.
VII. CO6. Discuss polynomial and logistic regression methods and estimate their parameters.

Module 1: Regression and Model building: Scatter diagram, regressor, response, error, uses of regression. Simple linear regression: Simple liner regression model, assumptions, Least square estimation of parameters, Properties of the Least-square estimators and the fitted Regression Model. Estimation of σ^2, Hypothesis testing of slope and intercept. Interval estimation of regression parameters (Slope, intercept and σ^2). Coefficient of determination.Estimation of regression parameters by the method of Maximum likelihood.
(I, II, III: A, B, C) 25 hours

Module 2: Multiple Linear Regression: Multiple linear regression model, assumptions, least square estimation of parameters, Properties of the Least-square estimators Hypothesis testing in Multiple linear regression (ANOVA), Test on individual regression coefficients, Interval estimation of coefficients, slope and intercept, co-efficient of determination.
(IV: A, B, C) 20 hours

Module 3: Model adequacy checking: Residual analysis, Methods of scaling residuals – standardized residuals, studentized residuals, PRESS residuals, R- Student. Residual plots – Normal probability plots, plot of residuals against fitted values, plot of residuals against the regressor, plot of residuals in time sequence. PRESS Statistic, $R^2$for prediction based on PRESS.
(V: A,B,C) 25 hours

Module 4: Polynomial and logistic regression: Polynomial models in one variable and two variables, Piece wise polynomial fitting (Basic concept only). Logistic regression- model with binary response variable, Estimation and Interpretation of the Parameters in logistic regression model.**(VI: A,B,C)** 20 hours

Book for Study

1. D C. Montegomerry, E A Peak and G G Vining, Introduction to Linear regression analysis, Wiley 2003.

References

1. Seber, Linear Regression Analysis, Wiley 1977.
2. D. D Joshi, Linear Estimation and Design of Experiments, Wiley 1987.
3. D N Gujarathi, D C Porter and G Sangeetha, Basic Econometrics, McGraw Hill, 2003

STA5B08: STATISTICAL COMPUTING (90 Hours)

Objectives: Statistical computing is a practical course. Its objectives are to develop scientific and experimental skills of the students and to correlate the theoretical principles with application based studies.

An introductory section is included in the course to familiarise the students with the basics of R package.

Introduction to R: R as a calculator, statistical software and a programming language, R preliminaries, getting help, data inputting methods(direct and importing from other spread sheet applications like Excel), data accessing, and indexing, Graphics in R, built in functions, saving, storing and retrieving work. Looping and Decision making – *for* loop, *while* loop, *if* command, *if else* command.

The practical is based on the following modules:

Module 1.	Diagrammatic representation of univariate and bivariate data - box plots, stem and leaf diagrams, bar plots, pie diagram, scatter plots.
Module 2.	Descriptive statistics - measures of central tendency (mean, median and mode), partition values, measures of dispersion (range, standard deviation, mean deviation and inter quartile range), summaries of a numerical data, skewness and kurtosis, random sampling with and without replacement.
Module 3.	Probability Distributions: Random number generation.
Module 4.	Statistical Inference: One- and two-sample tests, ztest, t-test,F-test, chi-square test of independence and goodness of fit, interval estimation for mean, difference of mean and variance, tests for normality (Shapiro-Wilks test, Wilcoxon's test and q-q plot), ANOVA (one- way)
Module 5.	Correlation and regression analysis(bivariate and multivariate data), polynomial regression, logistic regression.

Practical is to be done using R package. At least five statistical data oriented/supported problems should be done from each course. Practical record shall be maintained by each student and the same shall be submitted for verification at the time of external examination.

The Board of examiners (BoE) shall decide the pattern of question paper and the duration of the external examination. The external examination at each centre shall be conducted and evaluated on the same day jointly by two examiners – one external and one internal, appointed at the centre of the examination by the college on the recommendation of the Chairman, BoE. The question paper for the external examination at the centre will be set by

the external examiner in consultation with the Chairman, BoE and HoD of the centre. The questions are to be evenly distributed over the entire syllabus. Evaluation shall be done by assessing each candidate on the scientific and experimental skills, the efficiency of the algorithm/program implemented, the presentation and interpretation of the results.

References:

1. Michale J. Crawley, THE R BOOK, John Wiley & Sons, England (2009)

2. Sudha G. Purohitet.al., Statistics Using R, Narosa Publishing House, , India(2008)

3.JohnVerzani, simple R-Using R for Introductory Statistics, (http:// www.math.csi.cuny.edu/Statistics/R/SimpleR/Simple.)

4. W. N. Venables, D. M. Smith and the R Core Team, An Introduction to R , Notes on R: A Programming Environment for Data Analysis and Graphics, Version 2.15.2 (2012-10-26) (http://www.r-project.org)

SEMESTER VI

STA6B09: TIME SERIES AND INDEX NUMBERS (90 Hours)

CO Statement

I. CO1: Understand the components and models of time series
II. CO2: Determination of trends and construction of seasonal indices.
III. CO3: Discuss the income and allied distributions.
IV. CO4: Explain the definition and construction of index numbers.
V. CO5: Test the consistency for index numbers.
VI. CO6: Discuss the attitude measurement scale and issues associated with it.

Module 1: Time series analysis: Economic time series, different components, illustrations, additive and multiplicative models, determination of trends, growth curves, analysis of seasonal fluctuations, construction of seasonal indices.

(I: A, B, II: A, B, C) 25 hours

Module 2: Analysis of Income and allied distributions-Pareto distribution, graphical test, fitting of Pareto's law, illustrations, lognormal distribution and properties, Lorenz curve, Gini's coefficient.

(III: A, B) 20 hours

Module 3: Index numbers: Meaning and definition-uses and types-problems in the construction of index numbers-simple aggregate and weighted aggregate index numbers. Test for consistency of index numbers-factor reversal, time reversal and unit test, Chain base index numbers-Base shifting-splicing and deflating of index numbers. Consumer price index numbers-family budget enquiry-limitations of index numbers.

(IV, V: A, B, C) 30 hours

Module 4: Attitude Measurements and Scales: issues in attitude measurements scaling of attitude-Guttman scale, Semantic differential scale, Likert scale; selection of appropriate scale-limitations of scales.

(VI: A, B) 15 hours

Books for references

1. SC Gupta and V K Kapoor, Fundamentals of applied statistics, Sulthanchand and sons
2. Goon A M, Gupta M K and Das Gupta, Fundamentals of Statistics Vol. II, The World press, Calcutta
3. Box G E P and Jenkins G M, Time series analysis, Holden Day.
4. Meister David, Behavioural Analysis and Measurement methods, John Wiley New York
5. Luck et al. Marketing Research, Prentice Hall of India, New Delhi

STA6B10: DESIGN OF EXPERIMENTS (90 Hours)

CO Statement

I. CO1: Understand the idea of estimation and there by obtaining the best estimates.
II. CO2: Understand and apply analysis of variance and related post hoc tests.
III. CO3: Understand the concept of analysis of covariance with a single observation per cell.
IV. CO4: Understand the basic principles of experimentation
V. CO5: Compare and contrast complete block designs.
VI. CO6: Understand and apply factorial designs and incomplete block designs.

Module 1: Linear estimation, estimability of parametric functions and BLUE Gauss-Markovtheorem-Linear Hypothesis. **(I: A,B)** 15 hours

Module 2: Analysis of variance, one way and two way classification (with single observationper cell), Post Hoc Tests - Least Significant Difference (LSD) test, Duncan's multiple range test. Analysis of covariance with a single observation per cell (Concept and model only).

(II, III: A, B, C) 15 hours

Module 3: Principles of design-randomization- replication-local control, completelyrandomized design; Randomized block design; Latin square design. Missing plot technique; comparison of efficiency; Greco-Latin square design (Concept only).

(IV: A, B, C, D) 35 hours

Module 4: Basic concepts of factorial experiments, $2^2$and 2^3factorial experiments, Basicconcepts of Incomplete block design, Balanced incomplete block design and Partially Balanced incomplete block design.

(V, VI: A, B, C) 25 hours

References

1. S.C. Gupta and V K Kapoor, Fundamentals of applied Statistics, Sulthan Chand and Sons
2. M N Das and N Giri, Design of Experiments, New Age international,
4. D.D Joshy, linear Estimation and Design of Experiments, Wiley Eastern
5. Montgomery, D C, Design and Analysis of Experiments, John Wiley

STA6B11: POPULATION STUDIES, ACTUARIAL SCIENCE AND VITAL STATISTICS (90 Hours)

CO Statement

I. CO1: Recognize sources of vital statistics in India and its major functions
II. CO2: Understand and compute mortality rates, fertility and reproduction rates.
III. CO3: Understand and construct life tables.
IV. CO4: Understand the fundamentals of insurance.
V. CO5: Construct mortality tables to enhance the calculation of premiums of the life insurances.

Module 1: Sources of vital statistics in India-functions of vital statistics, Rates and ratios-mortality rates-crude, age specific and standard death rates -fertility and reproduction rates-crude birth rates-general and specific fertility rates-gross and net reproduction rates.

(I: A, II:A, B, C) 30 hours

Module 2: Life Tables-complete life tables and its characteristics-Abridged life tables and its characteristics, principle methods of construction of abridged life tables-Reed Merrel's method.

(III: A,B,C) 30 hours

Module 3: Fundamentals of insurance: Insurance defined meaning of loss, peril, hazard and proximate cause in insurance, Costs and benefits of insurance to society-branches of insurance. Insurable loss exposures-feature of loss that is deal of insurance, Construction of Mortality table-computation of premium of life insurance for fixed duration and for the whole life.

(IV: A, B, V: A, B, C) 30 hours

Books for reference

1. S.C. Gupta and V K Kapoor, Fundamentals of applied Statistics, Sulthan Chand and Sons
2. Benjamin B, Health and Vital Statistics, Allen and Unwin
3. Mark S Dorfman, Introduction to Risk Management and Insurance, Prentice Hall
4. C.D.Daykin, T. Pentikainen et al, Practical Risk Theory of Acturies, Chapman and Hill

STA6B12: OPERATIONS RESEARCH AND STATISTICAL QUALITY CONTROL

(90 Hours)

CO Statement

I. CO1: Understand and apply linear programming problem to solve real life problems.
II. CO2: Understand and apply the transportation and assignment problems to solve real life problems.
III. CO3: Explain the fundamental concept of control charts and causes of variations.
IV. CO4: Understand and apply control charts for variables and attributes.
V. CO5: Discuss the fundamentals of acceptance sampling and OC curve.
VI. CO6: Explain simple and double sampling plans and discuss the measures for the evaluation of performance of sampling plans.

Module 1: Linear programming: Mathematical formulation of LPP, Graphical and Simplex methods of solving LPP-duality in linear programming.

(**I**: A,**B,C**) 20 hours

Module 2 : Transportation and assignment problems, North-west corner rules, row column and least cost method-Vogel's approximation method, Assignment problem, Hungarian algorithm of solution.

(II: A, B, C) 20 hours

Module 3: General theory of control charts, causes of variations in quality, control limits, sub-grouping, summary of out-of-control criteria, charts of attributes, np chart, p chart, c chart, Charts of variables: X bar chart, R Chart and sigma chart, Revised control charts, applications and advantages.

(**III: A. B, IV: A, B, C**) 25 hours

Module 4: Principles of acceptance sampling-problems and lot acceptance, stipulation of good and bad lots-producer' and consumer' risk, simple and double sampling plans, their OC functions, concepts of AQL, LTPD,AOQL, Average amount of inspection and ASN function.

(V: A, B, VI: A, B) 25 hours

References

1. Gupta and Manmohan, Linear programming, Sulthan Chand and sons

2. Hardley G, Linear programming, Addison-Wesley

3. Taha, Operations Research, Macmillan,

4. V.K.Kapoor, Operations Research, Sultan Chand and Sons

5. S.C.Gupta and V.K.Kapoor Fundamentals of Applied Statistics, Sultan Chand and Sons

STA6B13: PROJECT WORK

Objectives:

1. The project work will helps the students to enhance their Research attitude.
2. It also helps in applying the theory of research in real life situations. Students get an exposure to study the working atmosphere of an enterprise or they can undertake research on any socially relevant area based on their various courses.

The following guidelines may be followed for project work.

1. The project is offered in the fifth and sixth semester of the degree course and the duration of the project may spread over the complete year.

2. A project may be undertaken by a group of students, the maximum number in a group shall not exceed 5. However the project report shall be submitted by each student.

3. There shall be a teacher from the department to supervise the project and the synopsis of the project should be approved by that teacher. The head of the department shall arrange teachers for supervision of the project work.

4. As far as possible, topics for the project may be selected from the applied branches of statistics, so that there is enough scope for applying and demonstrating statistical skills learnt in the degree course.

5. Field/Industrial/Organization visit is mandatory for the data collection.

ELECTIVE COURSES

STA6B14 (E): PROBABILITY MODELS AND RISK THEORY (54 Hours)

CO Statement

I. CO1: Discuss the model for individual claim random variables and sums of independent random variables of Individual risk model for a short time.
II. CO2: Explain the approximation for the distribution of sum of individual risk model for a short time and its application to insurance.
III. CO3: Describe the collective risk models for a single period and approximation to the distributions of aggregate claims.
IV. CO4: Discuss the collective risk models over an extended period and understand the maximal aggregate loss.
V. CO5: Establish the applications of risk theory and claim amount distributions approximating the individual model.
VI. CO6: Describe the stop-loss reinsurance and the effect of reinsurance on the probability of ruin.

Module 1: Individual risk model for a short time: Model for individual claim random variables-sums of independent random variables-Approximation for the distribution of sum-Application to insurance

(I, II, A, B) 10 hours

Module 2: Collective risk models for a single period: The distribution of aggregate claims-selection of basic distributions-properties of compound Poisson distribution-approximation to the distributions of aggregate claims (III A, B)
15 hours

Module 3: Collective risk models over an extended period: Claims process-The adjustment coefficients-Discrete time model-the first surplus below the initial level-The maximal aggregate loss

(IV A, B) 15 hours

Module 4: Application of risk theory: Claim amount distributions approximating the individual model-stop-loss re-insurance-the effect of reinsurance on the probability of ruin

(V, VI A, B, C) 14 hours

Books for reference

1. Institute of Actuaries, Act Ed. Study Materials

2. McCutcheon, JJ, Scott William (1986): An introduction to Mathematicsof Finance

3. Butcher M V, Nesbit, Cecil (1971) Mathematics of Compound Interest,Ulrich's book

4. Neil, Alistair, Heinemann (1977) Life contingencies

5. Bowers, Newton Let et al (1997) Actuarial mathematics, society ofActuaries, 2nd

STA6B15 (E): STOCHASTIC PROCESSES (54 Hours)

CO Statement:

I. CO1: Discuss Conditional probability and Baye's theorem.
II. CO2: Understand stochastic process and its classifications.
III. CO3: Understand the fundamentals of Markov process
IV. CO4: Explain Transition probability matrices
V. CO5: Discuss first passage probability and stationary distribution.
VI. CO6: Understand Ergodic theorems and periodicity.

Module 1: Conditional Probability, compound probability, Baye's theorem. Probability generating functions.

(I A, B) 6 hours

Module 2: Definition of stochastic process, Four classifications of Stochastic Processes, Markov Property, Markov process, Markov Chain, Graphical representations. Initial distributions (problems), Transition probabilities, Chapmann and Kolmogrov equations, Transition probabilitymatrices (examples and computation). Higher order transition probabilities (P^n)

(II, III, IV A, B) 30 hours

Module 3: Accessibility and communication of states. First passage probabilities, classification of states (recurrent, transient), mean recurrence. Periodicity, Ergodic theorem (Statement only), irreducible, class property.Stationary distribution.Limiting distribution (Definition and problems only).

(V, VI A, B) 18 hours

Books for reference

1. Medhi, J(1996). Stochastic Processes, Wiley Eastern Limited.
2. Karlin, S and Taylor, H.M.(1978). An Introduction to Stochastic Modeling (3rd edition), Academic Press
3. Karlin, S and Taylor, H.M.(1978). A first course in Stochastic Processes. Academic Press, New York.
4. Ross, S.M.(1983). Stochastic Processes, John Wiley and Sons.

STA6B16(E): RELIABILITY THEORY (54 Hours)

CO Statement

I. CO1: Understand the idea of notions of structural properties of binary system.
II. CO2: Explore the concepts of system reliability
III. CO3: Understand the type of ageing properties of lifetime distribution
IV. CO4: Determine the bounds of reliability.
V. CO5: Understand the applications of exponential and Poisson distributions in system reliability.

Module 1: Structural properties of coherent Systems: System of components series and parallel structure with example-dual structure function-coherent structure-preservation of coherent system in terms of paths and cuts representation of bridge structure - times to failure-relative importance of components - modules of coherent systems. (I A, B)

20 hours

Module 2: Reliability of Coherent systems: Reliability of a system of independent components-some basic properties of system reliability-computing exact system reliability-inclusion exclusion method-reliability importance of components (II A, B)

20 hours

Module 3: Parametric distributions in reliability: A notion of ageing (IFR and DFR only) with examples-exponential distribution-Poisson distribution. (III, IV, V, A, B)

14 hours

Books for references

1. R E Barlow and F Proschan (1975) Statistical Theory of Reliability andlife testing, Holt Rinhert, Winston

2. N Ravi Chandran, Reliability Theory, Wiley Eastern

OPEN COURSES

STA5D 01: ECONOMIC STATISTICS (54 Hours)

CO Statement

I. CO1: Understand components and models of time series data.
II. CO2: Determination of trends and construction of seasonal indices.
III. CO3: Explain the definition and construction of index numbers.
IV. CO4: Test the consistency for index numbers.

Module 1: Time series analysis: Economic time series, different components, illustrations, additive and multiplicative models, determination of trends, growth curves, analysis of seasonal fluctuations, construction of seasonal indices.

(I:A,B, II: A,B,C) 24 hours

Module 2: Index numbers: Meaning and definition-uses and types-problems in the construction of index numbers-simple aggregate and weighted aggregate index numbers. Test for consistency of index numbers-factor reversal , time reversal and unit test, Chain base index numbers-Base shifting-splicing and deflating of index numbers. Consumer price index numbers-family budget enquiry-limitations of index numbers.

(III:A, B, IV: A, B, C, D, E) 30 hours

Books for references

1. S C Gupta and V K Kapoor, Fundamentals of Applied Statistics, Sulthan Chand and sons

2. Goon A M, Gupta M K and Das Gupta, Fundamentals of Statistics Vol II, The World Press, Calcutta

STA5D 02: QUALITY CONTROL (54 Hours)

CO Statement

I. CO1: Explain the fundamental concept of control charts and causes of variations.
II. CO2: Understand and apply control charts for variables and attributes.
III. CO3: Discuss the fundamentals of acceptance sampling and OC curve.
IV. CO4: Explain simple and double sampling plans and discuss the measures for the evaluation of performance of sampling plans.

Module 1: General theory of control charts, causes of variations in quality, control limits, sub-grouping, summary of out-of-control criteria, charts of attributes, np chart, p chart, c

chart, Charts of variables: X bar chart, R Chart and sigma chart, Revised control charts, applications and advantages.

(I: A, B, II:A,B,C) 30 hours

Module 2: Principles of acceptance sampling-problems of lot acceptance, stipulation of good and bad lots-producer' and consumer' risk, simple and double sampling plans, their OC functions, concepts of AQL, LTPD,AOQL, Average amount of inspection and ASN function.

(III, IV: A, B) 24 hours

References

1. Grant E L, Statistical quality control, McGraw Hill
2. Duncan A J, Quality Control and Industrial Statistics, Taraporewala and sons
3. Montegomery D C, Introduction to Statistical Quality Control, John Wiley and son

STA5D03 : BASIC STATISTICS (54 Hours)

CO Statement

I. CO1: Compare census and sampling and discuss the principal steps in sample surveys and associated errors.
II. CO2: Understand and apply the measures of central tendency and measures of dispersion.
III. CO3: Discuss univariate and bivariate data and examine the linear correlation between two random variables.
IV. CO4: Apply the principle of least squares in fitting curves.
V. CO5: Understand the concepts of random experiments and definitions of probability.
VI. CO6: Understand and apply the theorems and results to compute the probabilities of events.

Module 1: Elements of Sample Survey: Census and Sampling, advantages, principal step in sample survey-sampling and non-sampling errors. Probability sampling, judgment sampling and simple random sampling.

(I: A,B) 10 hours

Module 2: Measures of Central tendency: Mean, median and mode and their empirical relationships ; Measures of Dispersion: absolute and relative measures, standard deviation and coefficient of variation.

(II: A, B) 12 hours

Module 3: Fundamental characteristics of bivariate data: univariate and bivariate data, scatter diagram, Pearson's correlation coefficient, limit of correlation coefficient. Curve fitting, principle of least squares, fitting of straight line.

(**III, IV: A,B,C**) 15 hours

Module 4 : Basic probability: Random experiment, sample space, event, algebra of events, Statistical regularity, frequency definition, classical definition and axiomatic definition of probability-addition theorem, conditional probability, multiplication theorem and independence of events (limited to three events).

(V,V**I: A,B,C**) 17 hours

References

1. V. K. Rohatgi, An Introduction to Probability Theory and Mathematical Statistics, Wiley Eastern.

2. S.C.Gupta and V. K. Kapoor, Fundamentals of Mathematical Statistics, Sultan Chand and Sons

3. A.M. Mood, F.A. Graybill and D C Bose, Introduction to Theory of Statistics, McGraw Hill

4. John E Freund, Mathematical Statistics (6th edn), Pearson Edn, New Delhi

STATISTICS: COMPLEMENTARY

SYLLABUS FOR B.Sc. (MATHEMATICS MAIN)

CBCSSUG 2019 (2019 admission onwards)

Sem No	Course Code	Course Title	Instructional Hours/week	Credit	Exam Hours	Ratio Ext: Int
1	STA 1C 01	INTRODUCTORY STATISTICS	4	3	2	4:1
2	STA 2C 02	PROBABILITY THEORY	4	3	2	4:1
3	STA 3C 03	PROBABILITY DISTRIBUTIONS AND SAMPLING THEORY	5	3	2	4:1
4	STA 4C 04	STATISTICAL INFERENCE AND QUALITY CONTROL	5	3	2	4:1

SEMSTER I
STA1C01: INTRODUCTORY STATISTICS (72 Hours)

CO Statement

I. CO1: Recognize important Statistical organizations in India.
II. CO2: Describe data and apply the measures of central tendency, measures of location, measures of dispersion and measures of shape.
III. CO3: Distinguish between correlation and regression in terms of explaining the relationship between two or more variables.
IV. CO4: Apply the principle of least squares in fitting linear and non-linear curves.
V. CO5: Understand the concept and components of time series and compare the methods for estimating trends and seasonal variations.

CO6: Explain the concept of index numbers and compare the methods for constructing index numbers

Module 1: *Official statistics*: The Statistical system in India: The Central and State Government organizations, functions of the Central Statistical Office (CSO), National Sample Survey Organization (NSSO) and the Department of Economics and Statistics.

(I, A)

7 hours

Module 2: *Introduction to Statistics*: Nature of Statistics, Uses of Statistics, Statistics in relation to other disciplines, Abuses of Statistics. Concept of primary and secondary data.Designing a questionnaire and a schedule.Concepts of statistical population and sample from a population, quantitative and qualitative data, Nominal, ordinal and time series data, discrete and continuous data.Presentation of data by table and by diagrams, frequency distributions by histogram and frequency polygon, cumulative frequency distributions (inclusive and exclusive methods) and ogives.Measures of central tendency (mean, median, mode, geometric mean and harmonic mean) with simple applications.Absolute and relative measures of dispersion (range, quartile deviation, mean deviation and standard deviation) with simple applications.Co-efficient of variation, Box Plot.Importance of moments, central and non-central moments, and their interrelationships. Measures of skewness based on quartiles and moments; kurtosis based on moments.

(II, A, B, C)

30 hours

Module 3: *Correlation and Regression:* Scatter Plot, Simple correlation, Simple regression, two regression lines, regression coefficients. Fitting of straight line, parabola, exponential, polynomial (least square method).**(III, IV, A, B, C)**

15 hours

Module 4:*Time series*: Introduction and examples of time series from various fields, Components of times series, Additive and Multiplicative models. Trend: Estimation of trend by free hand curve method, method of semi averages, method of moving averages and fitting various mathematical curves. Seasonal Component: Estimation of seasonal component by Method of simple averages, Ratio to Trend. *Index numbers:* Definition, construction of index numbers and problems thereof for weighted and unweighted index numbers including Laspeyre's, Paasche's, Edgeworth-Marshall and Fisher's.

(V, VI, A, B, C)

20hours

References:

1. S.C. Gupta and V.K. Kapoor. *Fundamentals of Mathematical Statistics*, Sultan Chand & Sons, New Delhi
2. Goon A.M., Gupta M.K. and Dasgupta B. (2002): Fundamentals of Statistics, Vol. I & II, 8th Edn. The World Press, Kolkata.
3. Mukhopadhyay P. (2011): Applied Statistics, 2nded. Revised reprint, Books and Allied
4. Hoel P.G. *Introduction to mathematical statistics*, Asia Publishing house.
5. Chatfield.C. *The Analysis of Time Series: An Introduction*, Chapman & Hall
6. *Guide to current Indian Official Statistics*, Central Statistical Office, GOI, New Delhi.
7. www.mospi.gov.in
8. www.ecostat.kerala.gov.in

SEMSTER II

STA2C02: PROBABILITY THEORY (72 Hours)

CO Statement

I. CO1:Understand the concepts of random experiments and definitions of probability.
II. CO2: Understand and apply the theorems and results to compute the probabilities of events.
III. CO3: Explain and apply discrete and continuous random variables, their probability functions and properties.
IV. CO4: Understand and apply the idea of expectations and there by obtaining the moments.
V. CO5: Understand and apply the concepts of bivariate random variables and their probability distributions.
VI. CO6: Determine the nature of relationship and the independence of bivariate random variables using mathematical expectation.

Module 1: *Introduction to Probability:* Random experiment, Sample space, events, classical definition of probability, statistical regularity, field, sigma field, axiomatic definition of probability and simple properties, addition theorem (two and three events), conditional probability of two events, multiplication theorem, independence of events-pair wise and mutual, Bayes theorem and its applications.

(I: A, B, II: A,B,C) 25 hour

Module 2: *Random variable*s: Discrete and continuous, probability mass function (pmf) and probability density function (pdf)-properties and examples, Cumulative distribution function and its properties, change of variables (univariate case only)

(III: A,B,C) 12 hours

Module 3: *Mathematical expectations (univaraite):* Definition, raw and central moments(definition and relationships), moment generation function and properties, characteristic function (definition and use only), Skewness and kurtosis using moments

(IV: A,B,C) 15 hours

Module 4: *Bivariate random variables*: Joint pmf and joint pdf, marginal and conditional probability, independence of random variables, function of random variable. Bivariate Expectations, conditional mean and variance, covariance, Karl Pearson Correlation coefficient, independence of random variables based on expectation.

(V: A, B,C VI: A,B,C) 20 hours

References :

1. Rohatgi V. K. and Saleh, A.K. Md. E. (2009): An Introduction to Probability and Statistics. 2nd Edn. (Reprint) John Wiley and Sons.
2. S.C.Gupta and V. K. Kapoor, *Fundamentals of Mathematical Statistics*, Sultan Chand and Sons.
3. Mood, A.M. Graybill, F.A. and Boes, D.C. (2007): Introduction to the Theory of Statistics, 3rd Edn., (Reprint), Tata McGraw-Hill Pub. Co. Ltd.
4. John E Freund, *Mathematical Statistics*, Pearson Edn, New Delhi
5. Hoel P.G. *Introduction to mathematical statistics*, Asia Publishing house.

SEMSTER III

STA3C03: PROBABILITY DISTRIBUTIONS AND SAMPLING THEORY.

(72 Hours)

CO Statement

I. CO1: Understand and derive discrete probability distributions and their properties.
II. CO2: Understand and derive continuous probability distributions and their properties.
III. CO3: Explain and apply Central Limit Theorems and laws of large numbers.
IV. CO4: Understand the methods of sampling and identify the suitable situations.
V. CO5: Understand the fundamentals of sampling distribution.
VI. CO6: Explain and derive Chi square, t, F distributions and their properties.

Module 1: *Standard distributions:* Discrete type-Bernoulli, Binomial, Poisson, Geometric, Negative Binomial (definition only), Uniform (mean, variance and mgf). Continuous type-Uniform, exponential and Normal (definition, properties and applications); Gamma (mean, variance, mgf); Lognormal, Beta, Pareto and Cauchy (Definition only).

(I: A,B,C, II: A,B,C) 30 hours

Module 2: *Limit theorems:* Chebyshev's inequality, Sequence of random variables, parameter and Statistic, Sample mean and variance, Convergence in probability (definition and example only), weak law of large numbers (iid case), Bernoulli law of large numbers, Convergence indistribution (definition and examples only), Central limit theorem (Lindberg levy-iid case).

(III: A,B,C) 25 hours

Module 3: Sampling methods: Simple random sampling with and without replacement, systematic sampling (Concept only), stratified sampling (Concept only), Cluster sampling (Concept only).

(IV: A,B,C) 10 hours

Module 4: Sampling distributions: Statistic, Standard error, Sampling from normal distribution, distribution of sample mean, sample variance, chi-square distribution, t-distribution, and F distribution (definition, derivations and relationships only).

(V: A,B, VI: A,B,C) 25 hours

References:

1. Rohatgi V. K. and Saleh, A.K. Md. E. (2009): An Introduction to Probability and Statistics. 2nd Edn. (Reprint) John Wiley and Sons.
2. S.C.Gupta and V. K. Kapoor, *Fundamentals of Mathematical Statistics*, Sultan Chand and Sons.

3. Mood, A.M. Graybill, F.A. and Boes, D.C. (2007): Introduction to the Theory of Statistics, 3rd Edn., (Reprint), Tata McGraw-Hill Pub. Co. Ltd.

4. John E Freund, *Mathematical Statistics*, Pearson Edn, NewDelhi

5. Cochran W.G. (1984):Sampling Techniques(3rd Ed.), Wiley Eastern.

SEMESTER IV

STA4C04: STATISTICAL INFERENCE AND QUALITY CONTROL. (72 Hours)

CO Statement

I. CO1: Understand and apply the theory of point estimation.
II. CO2: Understand and apply the theory of interval estimation.
III. CO3: Explore the concepts of testing of hypotheses.
IV. CO4: Identify and apply suitable parametric tests for varying situations.
V. CO5: Identify and apply suitable non parametric tests for varying situations.
VI. CO6: Explain and apply control charts for variables and attributes.

Module 1: *Estimation theory*: Parametric space, sample space, point estimation. Neyman Factorization criteria, Requirements of good estimator: Unbiasedness, Consistency, Efficiency, Sufficiency and completeness. Minimum variance unbiased (MVU) estimators.Cramer-Rao inequality (definition only). Minimum Variance Bound (MVB) estimators. Methods of estimation: Maximum likelihood estimation and Moment estimation methods (Detailed discussion with problems); Properties of maximum likelihood estimators (without proof); Least squares and minimum variance (concepts only). Interval estimation: Confidence interval (CI);CI for mean and variance of Normal distribution; Confidence interval for binomial proportion and population correlation coefficient when population is normal**.**

(I, II: A, B, C)

30 hours

Module 2: *Testing of Hypothesis*: Level of significance, Null and Alternative hypotheses, simple and composite hypothesis ,Types of Errors, Critical Region, Level of Significance, Power and p-values. Most powerful tests, Neyman-Pearson Lemma (without proof), Uniformly Most powerful tests. Large sample tests: Test for single mean, equality of two means, Test for single proportion, equality of two proportions. Small sample tests: t-test for single mean, unpaired and paired t-test. Chi-square test for equality of variances, goodness of fit, test of independence and association of attributes.Testing means of several populations: One Way ANOVA, Two Way ANOVA (assumptions, hypothesis, ANOVA table and problems).

(III, IV: A, B, C) 35 hours

Module 3: *Non-parametric methods***:** Advantages and drawbacks; Test for randomness, Median test, Sign test, Mann-Whiteny U test and Wilcoxon test; Kruskal Wallis test (Concept only).

(V: A, B, C) 10 hours

Module 4: *Quality Control:* General theory of control charts, causes of variations in quality, control limits, sub-grouping, summary of out-of-control criteria. Charts of variables - X bar chart, R Chart and sigma chart.Charts of attributes – c-charts, p-chart and np-chart.(Concepts and problems).

(VI: A, B, C) 15 hours

References:

1. Rohatgi V. K. and Saleh, A.K. Md. E. (2009): An Introduction to Probability and Statistics. 2nd Edn. (Reprint) John Wiley and Sons.
2. Gupta, S.P. *Statistical Methods*. Sultan Chand and Sons: New Delhi.
3. S.C.Gupta and V. K. Kapoor, *Fundamentals of Mathematical Statistics*, Sultan Chand and Sons

4. Mood, A.M. Graybill, F.A. and Boes, D.C. (2007): Introduction to the Theory of Statistics, 3rd Edn., (Reprint), Tata McGraw-Hill Pub. Co. Ltd.
5. John E Freund, *Mathematical Statistics*, Pearson Edn, NewDelhi
6. Grant E L, *Statistical quality control*, McGraw Hill

7. Montogomery, D. C. (2009): Introduction to Statistical Quality Control, 6th Edition, Wiley India Pvt. Ltd.

Model Question Papers

PATTERN OF QUESTION PAPER (B.Sc STATISTICS)

SEMESTER: I

STA1B01: OFFICIAL STATISTICS AND PROBABILITY

Contact Hours per Week : 4

Number of Credits : 4

Number of Contact Hours : 72

Course Evaluation : External 80 Marks + Internal 20 Marks

Duration of Exam : 2.5 hr

Module Blue Print For Question Paper Setting / Scrutiny						
Maximum Mark: 80						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE: 1	MODULE: 2	MODULE: 3	MODULE: 4
			Hour: 20	Hour: 17	Hour: 25	Hour: 10
			Marks: 32	Marks: 24	Marks: 28	Marks: 26
Expected Marks >>>>						
A	2	1.	2			
		2.	2			
		3.	2			
		4.	2			
		5.	2			
		6.	2			
		7.		2		
		8.		2		
		9.			2	
		10.			2	
		11.			2	
		12.			2	
		13.				2
		14.				2
		15.				2
B	5	16.	5			
		17.	5			
		18.		5		
		19.		5		
		20.			5	
		21.			5	
		22.				5
		23.				5
C	10	24.	10			
		25.		10		
		26.			10	
		27.				10

Total Marks >>>>	32	24	28	26

STA1B01: OFFICIAL STATISTICS AND PROBABILITY

Time: 2 ½ Hours Max.: 80 Marks

PART A

Each question carries 2 marks.

1. List out the important activities of Central Statistics Organization.
2. What are the main roles of state level statistical organizations?
3. The mean salary of 80 male employees in a firm is Rs. 5200 and that of 20 females in the same firm is Rs. 4200. What is the mean salary of all the employees in that firm?
4. Obtain the median for the following frequency distribution

X	1	2	3	4	5	6
Frequency	8	10	11	16	20	25

5. Compare range and quartile deviation.
6. The mean, median and standard deviation of a moderately asymmetrical distribution are 25, 23 and 4.5 respectively. Calculate Pearson's coefficient of skewness.
7. Write down the normal equation for fitting an exponential curve $y = ab^x$
8. Describe negative and positive correlation. Give example.
9. Define random experiment. Write an example.
10. State addition theorem of probability for three events.
11. Define independence of events.
12. Write the axiomatic definition of probability.
13. Distinguish between discrete and continuous random variables.
14. Define distribution function. Write any two properties of distribution function.
15. Let $f(x) = 2x+3$, $0 < x < 1$; 0 otherwise. Verify whether $f(x)$ is a probability density function or not.

Maximum Mark = 25

PART B

Each question carries 5 marks

16. For a group of 150 candidates, the mean and standard deviation of scores were found to be 38 and 16 respectively. Latter on it was found that the scores 45 ad 53 were misread as 54 and 35 respectively. Find the standard deviation of corrected figures.
17. Explain skewness and kurtosis.
18. Explain the least square method of fitting a parabola.
19. Consider the following set of rankings for a sample of 10 elements. Compute Spearman's rank correlation coefficient for the data.

Element	1	2	3	4	5	6	7	8	9	10
X	10	6	7	3	4	2	8	5	1	9
Y	8	4	10	2	5	7	6	3	1	9

20. State and prove Baye's theorem.

21. A problem in Statistics is given to 3 students A,B and C whose chances of solving it are ½, ¾ and ¼ respectively. What is the probability that the problem will be solved?
22. Let f(x) = 1, 0<x<1; 0 otherwise. Find the distribution of Y= -2log X
23. For a discrete r.v. X with probability distribution

x	-2	-1	0	1	2
P(X=x)	0.1	0.2	0.3	k	0.2

find the value of (i) k (ii) $p(-1 \leq X \leq 1)$

Maximum Mark = 35

PART C

Each question carries 10 marks(Answerany **TWO** Questions)

24. Police records show the following numbers of daily crime reports for a sample of days during the winter months and a sample of days during the summer months. Compare the variability of the two periods.

Winter	18	20	15	16	21	20	12	16	19	20
Summer	28	18	24	32	18	29	23	38	28	18

25. A box contains 3 blue and 2 red balls while another box contains 2 blue and 5 red balls. A ball drawn at random from one of the boxes turns out to be blue. What is the probability that it came from the first box
26. A committee of 4 people is appointed from 3 officers of the production department, 4 officers of the purchase department, 2 officers of the sales department and 1 charted accountant. Find the probability of forming the committee in the following manner.
    (1) There must be one from each category
    (2) It should have at least one from the purchase department
    (3) The charted accountant must be in the committee.
27. For n events $A_1, A_2, \ldots, A_n$ show that
    (1) $P(\bigcap_{i=1}^{n} A_i) \geq \sum_{i=1}^{n} P(A_i) - (n-1)$
    (2) $P(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$

$2 \times 10 = 20$

SEMESTER: II

STA2B02: BIVARIATE RANDOM VARIABLE AND PROBABILITY DISTRIBUTIONS

Contact Hours per Week : 4
Number of Credits : 4

Number of Contact Hours : 72

Course Evaluation : External 80 Marks + Internal 20 Marks

Duration of Exam : 2.5 hr

Module Blue Print For Question Paper Setting / Scrutiny

Maximum Mark: 80						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE: 1	MODULE: 2	MODULE: 3	MODULE: 4
			Hour: 20	Hour: 15	Hour: 12	Hour: 25
			Marks: 28	Marks: 24	Marks: 25	Marks: 33
Expected Marks >>>>						
A	2	1.	2			
		2.	2			
		3.	2			
		4.	2			
		5.		2		
		6.		2		
		7.			2	
		8.			2	
		9.			2	
		10.			2	
		11.			2	
		12.				2
		13.				2
		14.				2
		15.				2
B	5	16.	5			
		17.	5			
		18.		5		
		19.		5		
		20.			5	
		21.				5
		22.				5
		23.				5
C	10	24.	10			
		25.		10		
		26.			10	
		27.				10
Total Marks >>>>			28	24	25	33

STA2B02: BIVARIATE RANDOM VARIABLE AND PROBABILITY DISTRIBUTIONS

Time: 2 ½ Hours Max.: 80 Marks

PART A

Each question carries 2 marks

1. Define expectation of a random variable. Write any two properties of expectation.
2. Show that $V(aX + b) = a^2V(X)$.
3. Let X_1 and X_2 be two independent random variables. Prove that $M_{X1+X2}(t) = M_{X1}(t)M_{X2}(t)$.

4. Define characteristic function and write any two properties.
5. Define distribution function of a bivariate random variable and write any two properties.
6. When do you say that two random variables are independent?
7. State and prove the multiplication theorem of expectation
8. Define conditional expectation and conditional variance of a bivariate random variable.
9. For a discrete bivariate random variable (X, Y), prove that E{E(X|Y)} = E(X)
10. If X and Y are two independent random variables then show that Cov(X, Y) = 0.
11. Write the moment coefficients of skewness and kurtosis.
12. Describe degenerate probability distribution.
13. Define Binomial random variable.
14. Derive the mean of a Poisson random variable.
15. Define convergence in probability.

Maximum Mark = 25

PART B

Each question carries 5 marks

16. Let X be a random variable with the following probability distribution

X	-3	6	9
P(X = x)	1/6	1/2	1/3

Compute (i) V(X) and (ii) $E(2X+1)^2$

17. A random variable X has the density function f(x) = 1/2√x; 0 < x < 1 Obtain the m.g.f. of X and hence fid its mean and variance.
18. The joint pdf of a bivariate random variable ifs f(x y) = 2; 0 < x < 1; 0 < y < x. Check for independence f X ad Y
19. If X and Y have joint pdf $f(x,y) = x+y, 0 < x, y < 1$. Find $P\left(0 < X < \frac{1}{2}, \frac{1}{2} < Y < 1\right)$
20. Sow that V(X) = E[V(X|Y)] + V[E(X|Y)]
21. Let X and Y are two independent and identically distributed geometric random variables. Then show that the conditional distribution of X|X+Y = n is uniform
22. Establish the memory less property of geometric distribution.
23. State and prove Chebychev's inequality for continuous random variables.

Maximum Mark = 35

PART C

Each question carries 10 marks(Answerany **TWO** Questions)

24. (i) State and prove the addition theorem of expectation (ii) Derive the relationship between raw moments and central moments
25. If (X, Y) is a bivariate discrete random variable with joint probability mass function

Y \ X	1	2	3
1	$\frac{2}{21}$	$\frac{3}{21}$	$\frac{4}{21}$
2	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{5}{21}$

(i) Find the marginal distributions of X and Y (ii) Compute P(X|Y=1) and (ii) Compute P(Y|X=2)

26. If X and Y are two r.v.s having the joint pdf f(x y) = 2 – x – y ; 0 < x, y < 1. Find $\rho_{xy.}$
27. Show that Poisson distribution is a limiting case of Binomial distribution under certain conditions.

$2 \times 10 = 20$

SEMESTER: III

STA3B03:STATISTICAL ESTIMATION

Contact Hours per Week : 5
Number of Credits : 4

Number of Contact Hours : 90

Course Evaluation : External 80 Marks + Internal 20 Marks

Duration of Exam : 2.5 hr

Module Blue Print For Question Paper Setting / Scrutiny						
Maximum Mark: 80						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE: 1	MODULE: 2	MODULE: 3	MODULE: 4
			Hour: 20	Hour: 15	Hour: 40	Hour: 15
			Marks: 30	Marks: 26	Marks: 28	Marks: 26
Expected Marks >>>>						
A	2	1.	2			
		2.	2			
		3.	2			
		4.	2			
		5.	2			
		6.		2		
		7.		2		
		8.		2		
		9.			2	
		10.			2	
		11.			2	
		12.			2	
		13.				2
		14.				2
		15.				2
B	5	16.	5			
		17.	5			
		18.		5		
		19.		5		
		20.			5	
		21.			5	
		22.				5
		23.				5
C	10	24.	10			
		25.		10		
		26.			10	
		27.				10
Total Marks >>>>			30	26	28	26

STA3B03:STATISTICAL ESTIMATION

Time: 2 ½ Hours Max.: 80 Marks

PART A

Each question carries 2 marks.

1. If X is a random variable with a continuous distribution function F. Then show that F ~ Uniform[0, 1]
2. Write the relationship between normal and log normal distributions.
3. Define Pareto distribution.
4. Derive the moment generating function of Gamma distribution.
5. State Lindberg-Levy central limit theorem.
6. Distinguish between parameter and statistic.
7. State the additive property of Chi square distribution.
8. Establish the relationship between t and F distributions.
9. Define Cramer-Rao inequality
10. State Neyman –Pearson Factorisation theorem
11. Define unbiasedness and efficiency of an estimator.
12. What is the significance of Bayesian estimation method?
13. Describe interval estimation.
14. Define confidence interval and confidence coefficient.
15. Write the confidence for the mean of normal distribution when population standard deviation is unknown.

Maximum Mark = 25

PART B

Each question carries 5 marks

16. State and prove memory less property of exponential distribution.
17. If X follows Beta distribution of second kind then $(1 + X)^{-1}$ follows Beta distribution of first kind with same parameter.
18. If two independent random samples of sizes 15 and 20 are taken from $N(\mu,\sigma)$. What is the probability of $(S_1^2/ S_2^2) <2$.
19. Derive the moment generating function of X^2 distribution
20. Prove that in a Normal distribution sample mean is a consistent estimator of population mean.
21. Find the maximum likelihood estimator of λ for the Poisson distribution.
22. Establish confidence interval for the difference of proportions of two binomial populations
23. Derive exact confidence interval for the difference of population means based on t distribution.

Maximum Mark = 35

PART C

Each question carries 10 marks(Answerany **TWO** Questions)

24. X is a normal variate with mean 42 and standard deviation 4. Find the probability that a value taken by X is (i) less than 50, (ii) greater than 50, (iii) greater than 40, (iv) in between 40 and 50 and (v) equal to 45.
25. Derive Student's t distribution and state some of its applications.
26. (i) If T is a consistent estimator of θ, then prove that T^2 is a consistent estimator of θ^2. (ii) For a rectangular distribution over (a, b), a<b , find the Maximum likelihood estimates of a and b.
27. Derive the confidence interval for the mean of a Normal population $N(\mu, \sigma^2)$, when (a) σ is known (b) σ is unknown and the sample size is small

$2 \times 10 = 20$

SEMESTER: IV

STA4B04- TESTING OF HYPOTHESIS

Contact Hours per Week : 5
Number of Credits : 4

Number of Contact Hours : 90

Course Evaluation : External 80 Marks + Internal 20 Marks

Duration of Exam : 2.5 hr

Module Blue Print For Question Paper Setting / Scrutiny						
Maximum Mark: 80						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE: 1	MODULE: 2	MODULE: 3	MODULE: 4
			Hour: 20	Hour: 30	Hour: 20	Hour: 20
			Marks: 28	Marks: 30	Marks: 26	Marks: 26
Expected Marks >>>>						
A	2	1.	2			
		2.	2			
		3.	2			
		4.	2			
		5.		2		
		6.		2		
		7.		2		
		8.		2		
		9.		2		
		10.			2	
		11.			2	
		12.			2	
		13.				2
		14.				2
		15.				2
		16.	5			
		17.	5			

B	5	18.		5		
		19.		5		
		20.			5	
		21.			5	
		22.				5
		23.				5
C	10	24.	10			
		25.		10		
		26.			10	
		27.				10
Total Marks >>>>			28	30	26	26

STA4B04- TESTING OF HYPOTHESIS

Time: 2 ½ Hours Max.: 80 Marks

PART A

Each question carries 2 marks.

1. Define simple and composite hypotheses.
2. Distinguish between Type I error and Type II error.
3. Discuss sequential sampling.
4. What is the significance of SPRT?
5. Distinguish between small sample tests and large sample tests.
6. Write the critical regions of large sample test.
7. Compare parametric and non-parametric tests
8. Write a suitable situation where ANOVA test is applicable.
9. List the assumptions of one sample t test.
10. When do you use Yate's correction?
11. Identify a suitable test and its test statistic for testing the significance of a correlation coefficient.
12. What is meant by goodness of fit?
13. Write short note on Median Test
14. Write the test statistic of Wilcoxon-signed rank and identify its asymptotic distribution.
15. Briefly discuss run test.

Maximum Mark = 25

PART B

Each question carries 5 marks

16. The continuous random variable X has the frequency function

$$f(x,\theta) = \frac{1}{\theta}, 0 \le x \le \theta, \qquad 0\ otherwise$$

It is desired to test the hypothesis $H_0: \theta = 1\ against H_1: \theta = 2$ using a single observation X. X≥ 0.95 is used as the critical region. Evaluate Type I error and Type II error.

17. Explain pared t test. Give a practical situation where this test is suitable.

18. A sample of 25 boys who passed SSLC examination are found to have mean marks 50 with standard deviation 5 for English. The mean marks of 18 girls are found to be 48 with standard deviation 4 for the same subject. Does this indicate any significance difference between the marks of boys and girls assuming the population standard deviation are equal?
19. In a sample of 600 men from a certain city 400 are fund to be smokers In 900 from another city 450 are smokers Do the data indicate that the cities are significantly different as far as smoking habits of people are concerned.
20. Tests were carried out to assess the strength of single fibre yarn spun on two different machines A and B and the results are given below:

Machine A	4	4.4	3.9	3	4.2	4.4	5
Machine B	5.3	4.3	4.1	4.4	5.3	4.2	3.8

Assuming the samples have been taken from normal population, test the hypothesis that variability is same for both the machines.

21. Explain Chi square test for independence of attributes.
22. A sample of 10 men was used in a study to test the effects of a relaxant on the time required to fall asleep for male adults. Data for 10 subjects showing the number of minutes required to fall asleep with and without the relaxant follow. Use a 0.05 level of significance to determine whether the relaxant reduces the time required to fall asleep. Perform sign test and draw your conclusion.

Subject	1	2	3	4	5	6	7	8	9	10
Without Relaxant	15	12	22	8	10	7	8	10	14	9
With Relaxant	10	10	12	10	8	5	9	7	11	6

23. Explain Kolmogorov-Smirnov test. Suggest a situation where this test is useful.

Maximum Mark = 35

PART C

Each question carries 10 marks(Answerany **TWO** Questions)

24. State Neyman Pearson Lemma. Use the lemma to obtain the best critical region for testing $H_0: \mu = \mu_0$against $H_1: \mu = \mu_1$, in the case of a normal population with mean μ and variance σ^2. Find the power of the test.
25. List basic assumptions of ANOVA and explain the procedure or performing an ANOVA test.
26. Fit a Poisson distribution for the following data and test the goodness of fit.

X	0	1	2	3	4	5	6
frequency	275	72	30	7	5	2	1

27. Explain (i) Wilcoxon signed rank test and (ii) Mann Whitney test

$2 \times 10 = 20$

SEMESTER: V

STA5B05-MATHEMATICAL METHODS IN STATISTICS

Contact Hours per Week : 5
Number of Credits : 4

Number of Contact Hours : 90

Course Evaluation : External 80 Marks + Internal 20 Marks

Duration of Exam : 2.5 hr

Module Blue Print For Question Paper Setting / Scrutiny						
Maximum Mark: 80						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE: 1	MODULE: 2	MODULE: 3	MODULE: 4
			Hour: 20	Hour: 25	Hour: 25	Hour: 20
			Marks: 28	Marks: 28	Marks: 28	Marks: 26
Expected Marks >>>>						
A	2	1.	2			
		2.	2			
		3.	2			
		4.	2			
		5.		2		
		6.		2		
		7.		2		
		8.		2		
		9.			2	
		10.			2	
		11.			2	
		12.			2	
		13.				2
		14.				2
		15.				2
B	5	16.	5			
		17.	5			
		18.		5		
		19.		5		
		20.			5	
		21.			5	
		22.				5
		23.				5
		24.	10			
		25.		10		

C	10	26.			10	
		27.				10
Total Marks >>>>			28	28	28	26

STA5B05-MATHEMATICAL METHODS IN STATISTICS

Time: 2 ½ Hours Max.: 80 Marks

PART A

Each question carries 2 marks.

1. Write any two properties of real numbers?
2. Show that every convergent sequence is bounded
3. Find the region of x in $|x+1|<1$
4. State Taylor's theorem.
5. What is nested intervals?
6. Compute $\lim_{n\to\infty}(n!)^{1/n}$.
7. Why the set of rational numbers is not order-complete?
8. Write triangle inequality for two variables.
9. Identify the sixth term in the sequence {7, 26, 63, 124, 215, ….}.
10. Find $\frac{d}{dx}\left(e^{e^x}\right)$.
11. Discuss the relationship between continuability and differentiability of a function.
12. State a sufficient condition for uniform continuity.
13. Define Riemann integral.
14. Write the necessary and sufficient condition for a bounded function *f* is integrable.
15. Prove or disprove $\int_a^b f(x)dx = -\int_b^a f(x)dx$

Maximum Mark = 25

PART B

Each question carries 5 marks

16. If $\{x_n\}$ and $\{y_n\}$ are two sequences of real numbers converging to real numbers x and y respectively. Show that $\{x_n y_n\}$ converges to xy.
17. State and Prove Archmedian Property of Real numbers
18. Show every bounded sequence of real numbers has a convergent subsequence.
19. Show that every monotone sequence of bounded real numbers is convergent
20. State and prove mean value theorem
21. Establish Rolle's Theorem.
22. Prove that a bounded and continuous function f(.) over a closed interval (finite) is a sufficient condition for integrability.
23. If Q is a refinement of P, show that $L(P,f)\leq U(Q,f)$

Maximum Mark = 35

PART C

Each question carries 10 marks(Answerany **TWO** Questions)

24. /State and prove Cauchy convergence criteria
25. State and prove principle of Mathematical induction
26. State and prove intermediate value theorem.

27. If f and g are integrable then check whether the following areintegrable or not: a) $|f|$ b) f^2 c) $f.g2 \times 10 = 20$

SEMESTER: V

STA5B06: SAMPLE SURVEYS

Contact Hours per Week : 5
Number of Credits : 5

Number of Contact Hours : 90

Course Evaluation : External 80 Marks + Internal 20 Marks

Duration of Exam : 2.5 hr

Module Blue Print For Question Paper Setting / Scrutiny						
Maximum Mark: 80						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE: 1	MODULE: 2	MODULE: 3	MODULE: 4
			Hour: 20	Hour: 20	Hour: 40	Hour: 20
			Marks: 28	Marks: 26	Marks: 40	Marks: 16
Expected Marks >>>>						
A	2	1.	2			
		2.	2			
		3.	2			
		4.	2			
		5.		2		
		6.		2		
		7.		2		
		8.			2	
		9.			2	
		10.			2	
		11.			2	
		12.			2	
		13.				2
		14.				2
		15.				2
B	5	16.	5			
		17.	5			
		18.		5		
		19.		5		
		20.			5	
		21.			5	
		22.				5
		23.				5
		24.	10			
		25.		10		

C	10	26.			10	
		27.			10	
Total Marks >>>>			28	26	40	16

STA5B06: SAMPLE SURVEYS

Time: 2 ½ Hours Max.: 80 Marks

PART A

Each question carries 2 marks.

1. Distinguish between census and sampling.
2. What is meant by judgement sampling?
3. Define secondary data. State its major sources.
4. Compare sampling and non-sampling errors.
5. Write any four properties of a good questionnaire.
6. Discuss simple random sampling.
7. Compute the total number of samples of size n = 2 from a population of N = 6.
8. Write the unbiased estimate of population total and its variance.
9. Write Bowley's formula for proportional allocation.
10. In a systematic sampling N =40 and n = 4, find the value of k.
11. If ρis the intraclass correlation coefficient between the units of the same systematic sample, then when do you say that systematic sampling would be more efficient as compared with srswor?
12. Briefly explain cluster sampling.
13. Write a situation where two stage sampling is applicable.
14. Distinguish between a stratum and a cluster.
15. Discuss Neyman allocation of sample sizes.

Maximum Mark = 25

PART B

Each question carries 5 marks

16. What are the advantages of sampling over census?
17. Explain the probability sampling and non-probability sampling with the help of examples.
18. Explain the concept of stratified sampling.
19. Obtain an unbiased estimate of population mean in simple random sampling with replacement. Find the variance of the estimate.
20. Show that sample proportion, p is an unbiased estimate of population proportion, P. Also obtain the confidence interval for the population proportion
21. Show that sample mean is an unbiased estimate of population mean in stratified random sampling. Also find its variance.
22. What you mean by precision of an estimate? Show that mean of a systematic sample is more precise than mean of simple random sample.
23. Obtain an unbiased estimator of population mean in cluster sampling. Also find its variance.

Maximum Mark = 35

PART C

Each question carries 10 marks(Answerany **TWO** Questions)

24. Explain in detail the principal steps in a sample survey.
25. Show that Var $(\overline{y_{sys}}) = \frac{N-1}{Nn}$ (1+ (n-1)ρ)S^2, where ρ is the interclass correlation between the units of the same systematic sample.
26. If the population consists of a linear trend, then prove that
    Var $(\overline{y_{st}}) \leq$ Var $(\overline{y_{sys}}) \leq$ Var $(\overline{y_{ran}})$.
27. Compare the efficiencies of the Neyman and proportional allocations with that of an unstratified random sample of the same size.

$2 \times 10 = 20$

SEMESTER: V

STA5B07- LINEAR REGRESSION ANALYSIS

Contact Hours per Week : 5
Number of Credits : 4
Number of Contact Hours : 90
Course Evaluation : External 80 Marks + Internal 20 Marks
Duration of Exam : 2.5 hr

Module Blue Print For Question Paper Setting / Scrutiny						
Maximum Mark: 80						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE: 1	MODULE: 2	MODULE: 3	MODULE: 4
			Hour: 25	Hour: 20	Hour: 25	Hour: 20
			Marks: 33	Marks: 28	Marks: 21	Marks: 28
Expected Marks >>>>						
A	2	1.	2			
		2.	2			
		3.	2			
		4.	2			
		5.		2		
		6.		2		
		7.		2		
		8.		2		
		9.			2	
		10.			2	
		11.			2	
		12.				2
		13.				2
		14.				2
		15.				2

B	5	16.	5			
		17.	5			
		18.	5			
		19.		5		
		20.		5		
		21.			5	
		22.				5
		23.				5
C	10	24.	10			
		25.		10		
		26.			10	
		27.				10
Total Marks >>>>			33	28	21	28

STA5B07- LINEAR REGRESSION ANALYSIS

Time: 2 ½ Hours Max.: 80 Marks

PART A

Each question carries 2 marks.

1. Describe scatter diagram.
2. Write the confidence interval for the slope and intercept in simple linear regression model.
3. List some of the uses of regression.
4. Show that the least square estimate of multiple linear regression coefficient is unbiased.
5. Write the hypotheses and test statistic for testing the significance of slope coefficient in simple linear regression model.
6. Show that the residual mean square is an unbiased estimate of population variance of random error component in linear regression model.
7. Write the properties of least square estimates of simple linear regression coefficients.
8. What is the importance of ANOVA in multiple linear regression?
9. Define hat matrix.
10. Distinguish between R^2 and adjusted R^2.
11. What is the significance of coefficient of determination?
12. Write a situation where logistic regression is applicable.
13. List the assumptions of logistic regression.
14. Define Splines.
15. Define logit function.

Maximum Mark = 25

PART B

Each question carries 5 marks

16. Derive the variance of regression coefficients in simple linear regression
17. Estimate the variance of the response variable.

18. From a study conducted by the Department of transportation on driving speed and mileage for midsize automobiles, following results are obtained:

Driving speed (x)	30	50	40	55	30	25	60	25
Mileage (y)	28	25	25	23	30	32	21	35

Fit a linear regression model for the mileage and interpret the result.

19. Consider the simple linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$ with $E(\varepsilon) = 0, Var(\varepsilon) = \sigma^2 and \varepsilon$ uncorrelated
    a) Show that Cov$(\hat{\beta}_0, \hat{\beta}_1) = -\bar{x}\sigma^2/S_{xx}$
    b) Show that Cov$(\bar{y}, \hat{\beta}_1)$=0.
20. Describe multiple linear regression model.
21. Discuss the properties of least square estimators in multiple linear regression.
22. Explain polynomial regression models and list some of its applied areas.
23. Explain the logistic regression model with a Binary response variable.

Maximum Mark = 35

PART C

Each question carries 10 marks(Answerany **TWO** Questions)

24. Derive the least square estimates of simple linear regression coefficients and show that they are unbiased.
25. Explain the role of residual in model adequacy checking.
26. Explain test for significance of regression coefficients in multiple linear regression
27. Describe the estimation of regression coefficients in logistic regression model.

2 × 10 = 20

SEMESTER: VI

STA6B09: TIME SERIES AND INDEX NUMBERS

Contact Hours per Week : 5
Number of Credits : 4

Number of Contact Hours : 90

Course Evaluation : External 80 Marks + Internal 20 Marks

Duration of Exam : 2.5 hr

Module Blue Print For Question Paper Setting / Scrutiny						
Maximum Mark: 80						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE: 1	MODULE: 2	MODULE: 3	MODULE: 4
			Hour: 25	Hour: 20	Hour: 30	Hour: 15
			Marks: 30	Marks: 30	Marks: 31	Marks: 19
Expected Marks >>>>						
		1.	2			
		2.	2			
		3.	2			
		4.	2			

A	2	5.	2			
		6.		2		
		7.		2		
		8.		2		
		9.		2		
		10.		2		
		11.			2	
		12.			2	
		13.			2	
		14.				2
		15.				2
B	5	16.	5			
		17.	5			
		18.		5		
		19.		5		
		20.			5	
		21.			5	
		22.			5	
		23.				5
C	10	24.	10			
		25.		10		
		26.			10	
		27.				10
Total Marks >>>>			30	30	31	19

STA6B09: TIME SERIES AND INDEX NUMBERS

Time: 2 ½ Hours Max.: 80 Marks

PART A

Each question carries 2 marks.

1. Give example for seasonal and cyclic variations in time series.
2. Define Time series
3. What is meant by de-seasonalisation of data?
4. Write the steps for calculating the seasonal index using the method of simple averages.
5. Define moving averages.
6. What is family budget method?
7. Write the importance of lognormal distribution in income analysis.
8. Briefly describe the fitting of Pareto's law.
9. Describe additive models in time series.
10. What is Gini's coefficient?
11. Define index numbers.
12. Mention any three uses of index numbers.

13. Write Marshal-Edgeworth index number and Dorbish and Bowley's index number.
14. What are the scales of measurements?
15. Define Guttman scale.

Maximum Mark = 25

PART B

Each question carries 5 marks

16. How trend is measured using Moving Averages.
17. Explain periodic variations in Time series with suitable examples
18. Describe the Link Relative Method of measuring seasonal variation.
19. Explain Lorentz curve
20. Explain classification of index numbers
21. Give major limitations of index numbers
22. Explain the method of least squares.
23. How is a Likert scale measured? Is a Likert scale quantitative or qualitative?

Maximum Mark = 35

PART C

Each question carries 10 marks(Answerany **TWO** Questions)

24. What are the components of time series explain with example?
25. Explain the use of Pareto distribution and its applications.
26. Briefly explain Link relative methods also explain its merits and demerits
27. What is attitude scale? Explain the advantages and limitations of scales in attitude measurements?

2 × 10 = 20

SEMESTER: VI

STA6B10: DESIGN OF EXPERIMENTS

Contact Hours per Week : 5
Number of Credits : 5

Number of Contact Hours : 90

Course Evaluation : External 80 Marks + Internal 20 Marks

Duration of Exam : 2.5 hr

Module Blue Print For Question Paper Setting / Scrutiny						
Maximum Mark: 80						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE: 1	MODULE: 2	MODULE: 3	MODULE: 4
			Hour: 15	Hour: 15	Hour: 35	Hour: 25
			Marks: 26	Marks: 26	Marks: 28	Marks: 30
Expected Marks >>>>						
		1.	2			
		2.	2			
		3.	2			
		4.		2		

A	2	5.		2		
		6.		2		
		7.			2	
		8.			2	
		9.			2	
		10.			2	
		11.				2
		12.				2
		13.				2
		14.				2
		15.				2
B	5	16.	5			
		17.	5			
		18.		5		
		19.		5		
		20.			5	
		21.			5	
		22.				5
		23.				5
C	10	24.	10			
		25.		10		
		26.			10	
		27.				10
Total Marks >>>>			26	26	28	30

STA6B10: DESIGN OF EXPERIMENTS

Time: 2 ½ Hours Max.: 80 Marks

PART A

Each question carries 2 marks.

1. Explain Gauss Markov set up of linear model.
2. Define estimable parametric function.
3. What is meant by best linear unbiased estimator?
4. Discuss the concept of ANCOVA.
5. In a LSD with 4 treatments and error sum of squares is 16, find the Mean error sum of squares.
6. Write any situation where Graeco-Latin square design is suitable.
7. Distinguish between CRD and RBD.
8. Define treatment.
9. How to compare various designs of experiments?
10. In a 2^2 factorial design two factors A and B are given each at two levels, Write down the main effects and interaction effects of A and B
11. Distinguish between $2^2$and 2^3factorial designs.
12. Define incidence matrix

13. Discuss orthogonal design.
14. Describe BIBD.
15. What is the significance of PBIBD?

Maximum Mark = 25

PART B

Each question carries 5 marks

16. Consider three independent random variables y_1, y_2 and y_3 having common variance σ^2 and $E(Y_1) = \theta_1 - \theta_2, E(Y_2) = \theta_1 + \theta_2, E(Y_3) = 2\theta_1 - \theta_2$. Show that $3\theta_1 - 2\theta_2$ is an estimable parametric function
17. Find the least square estimate of the parameter vector θ in Gauss - Markov model and also find an unbiased estimator of $\sigma 2$.
18. What is meant by analysis of variance of experimental data? What are the assumptions used in it?
19. Explain the analysis for completely randomized design
20. Derive the expression for estimating one missing observation in RBD
21. Explain the efficiency of LSD compared to RBD
22. How can estimate the effects and calculate the sum of squares in factorial experiment?
23. Explain Duncan's multiple range test

Maximum Mark = 35

PART C

Each question carries 10 marks(Answerany **TWO** Questions)

24. State and prove Gauss-Markov theorem.
25. What is meant by missing plot technique and what are procedures used to obtain the missing observation. Write the expression for estimating two missing values in LSD and explain the ANOVA table in this case.
26. Explain the principles of design of experiments.
27. Define the main effects and interaction effects in a 2^3 factorial experiment. Also give its ANOVA table.

$2 \times 10 = 20$

SEMESTER: VI

STA6B11: POPULATION STUDIES, ACTUARIAL SCIENCE AND VITAL STATISTICS

Contact Hours per Week : 5
Number of Credits : 4

Number of Contact Hours : 90

Course Evaluation : External 80 Marks + Internal 20 Marks

Duration of Exam : 2.5 hr

Module Blue Print For Question Paper Setting / Scrutiny					
Maximum Mark: 80					
Question Paper			Syllabus		
Sections or Parts	Mark	Question Numbers	MODULE: 1	MODULE: 2	MODULE: 3
			Hour: 30	Hour: 30	Hour: 30
			Marks: 30	Marks: 30	Marks: 50
Expected Marks >>>>					
A	2	1.	2		
		2.	2		
		3.	2		
		4.	2		
		5.	2		
		6.		2	
		7.		2	
		8.		2	
		9.		2	
		10.		2	
		11.			2
		12.			2
		13.			2
		14.			2
		15.			2
B	5	16.	5		
		17.	5		
		18.		5	
		19.		5	
		20.			5
		21.			5
		22.			5
		23.			5
C	10	24.	10		
		25.		10	
		26.			10
		27.			10
Total Marks >>>>			30	30	50

STA6B11: POPULATION STUDIES, ACTUARIAL SCIENCE AND VITAL STATISTICS

Time: 2 ½ Hours Max.: 80 Marks

PART A

Each question carries 2 marks.

1. Define vital statistics. Give examples.
2. What is Crude birth rate?
3. Write the relation between N. R. R. And G. R. R.
4. Define mortality rate.
5. Write merits and demerits of general fertility rate.
6. Write any two characteristics of complete life tables.
7. What do you understand by an abridged life table?
8. What are the various uses of vital statistics?
9. What is expectation of life? Distinguish between 'curate expectation' and 'complete expectation' of life.
10. Discuss the costs and benefits of insurance to society.
11. Explain different kinds of policies.
12. Explain life insurance and fire insurance.
13. Define peril and hazard.
14. What is proximate cause in insurance?
15. List different branches of insurance.

Maximum Mark = 25

PART B

Each question carries 5 marks

16. Explain different methods of obtaining vital statistics.
17. Explain direct method of standardisation.
18. What are the assumptions in a life table?
19. State the merits and demerits of Crude Birth Rate.
20. Differentiate between Life insurance and General insurance.
21. Describe reinsurance and double insurance. Also explain the difference between them.
22. Explain the principles of insurance.
23. Discuss the costs and benefits of insurance to society.

Maximum Mark = 35

PART C

Each question carries 10 marks(Answerany **TWO** Questions)

24. Determine the standardised death rates for region A and B from the following.

Age Group	0-10	10-25	25-40	40-60	Above 60
Death per 1000, A (m_x^A)	2	8	15	32	41
Death per 1000, B(m_x^B))	4	10	18	28	38
Standard Population (in thousands)	3	12	12	30	50

25. Give brief account on sample registration system.
26. Explain the chief characteristics of an ideally insurable loss exposure.
27. Explain the method of calculating the premium of life insurance.2 × 10 = 20

STA6B12: OPERATIONS RESEARCH AND STATISTICAL QUALITY CONTROL

Contact Hours per Week : 5
Number of Credits : 4
Number of Contact Hours : 90
Course Evaluation : External 80 Marks + Internal 20 Marks
Duration of Exam : 2.5 hr

Module Blue Print For Question Paper Setting / Scrutiny						
Maximum Mark: 80						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE: 1	MODULE: 2	MODULE: 3	MODULE: 4
			Hour: 20	Hour: 20	Hour: 25	Hour: 25
			Marks: 26	Marks: 26	Marks: 30	Marks: 28
Expected Marks >>>>						
A	2	1.	2			
		2.	2			
		3.	2			
		4.		2		
		5.		2		
		6.		2		
		7.			2	
		8.			2	
		9.			2	
		10.			2	
		11.			2	
		12.				2
		13.				2
		14.				2
		15.				2
B	5	16.	5			
		17.	5			
		18.		5		
		19.		5		
		20.			5	
		21.			5	
		22.				5
		23.				5
C	10	24.	10			
		25.		10		
		26.			10	
		27.				10
Total Marks >>>>			26	26	30	28

STA6B12: OPERATIONS RESEARCH AND STATISTICAL QUALITY CONTROL

Time: 2 ½ Hours Max.: 80 Marks

PART A

Each question carries 2 marks.

1. Define slack and surplus variables.
2. Distinguish between feasible solution and optimal feasible solution
3. Define artificial variable? When do we use it?
4. Define degeneracy in transportation problem
5. What is meant by North West corner rule in lpp?
6. What are the disadvantages of BIG-M method over two phase method?
7. What are the causes of variation in quality control?
8. Explain the need for quality control techniques in production.
9. Describe C chart.
10. How to read a control chart?
11. What are the difference between defects & defectives?
12. What is the significance of OC curve?
13. Distinguish between consumer's risk and producer's risk.
14. Define LTPD.
15. What are ASN and ATI for the single sampling plan?

Maximum Mark = 25

PART B

Each question carries 5 marks

16. Describe linear programming problem. Write some of its applications.
17. Explain the graphical method for solving linear programming problem.
18. Find a geometrical interpretation and solution as well for the following linear programming problem.

Maximize $Z = 3x_1+5x_2$

Subject to

$x_1+2x_2 \leq 2000$

$x_1+x_2 \leq 1500$

$x_2 \leq 600$

$x_1 \geq 0 \quad x_2 \geq 0$

19. Explain assignment problem. Describe any method to solve it.
20. What is meant by a control charts? Explain the applications of these charts.
21. Describe procedure or drawing X bar and R charts.
22. Explain AQL and ASN.
23. Explain the main control charts for attributes and obtain their control limits.

Maximum Mark = 35

PART C

Each question carries 10 marks(Answerany **TWO** Questions)

24. Find the initial basic feasible solution of the following transportation problem. There are four origins three destinations. The availabilities are 9, 10, 8, 7 and the requirements are 17,10,7 respectively.

	A	**B**	**C**
D	2	3	2
E	1	3	4
F	2	3	1
G	2	4	3

25. What are the computational procedures of dual simplex method, explain with an example?
26. Distinguish between double and single sampling plans.
27. Draw the OC curve of the single sampling plan showing the consumers and producers risks.

$2 \times 10 = 20$

SEMESTER I

STA 1C 01- INTRODUCTORY STATISTICS

Contact Hours per week: 4
Number of credits: 3
Number of Contact Hours: 72
Course Evaluation: External 60 Marks+ Internal 15 Marks
Duration of Exam: 2 Hours

Blue Print for Question Paper Setting / Scrutiny						
Max. Marks: 60						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE 1	MODULE 2	MODULE 3	MODULE 4
			7 Hrs	30 Hrs	15 Hrs	20 Hrs
			7 Marks	30 Marks	19Marks	21 Marks
Expected mark >>>>						
A	2	1	2			
		2		2		
		3		2		
		4		2		
		5		2		
		6		2		
		7		2		
		8			2	
		9			2	
		10				2
		11				2
		12				2
B	5	13	5			
		14		5		
		15		5		
		16			5	
		17			5	
		18			5	
		19				5
C	10	20		10		
		21				10
Actual Mark >> >>			**7**	**32**	**19**	**21**

Question Paper setter has to give equal importance to both theory and problems in section B and C.

STA 1C 01 – INTRODUCTORY STATISTICS

Time: 2 Hours Max Marks: 60

SECTION-A

Each question carries 2 Marks.

Maximum Marks that can be scored in this section is 20.

1. What are the important functions of state statistical organizations?
2. Define data and observation. Give example.
3. What is meant by time series data? Describe it by using an example.
4. What are the main roles of state level statistical organizations?
5. The mean salary of 80 male employees in a firm is Rs. 5200 and that of 20 females in the same firm is Rs. 4200. What is the mean salary of all the employees in that firm?
6. Obtain the median for the following frequency distribution

X	1	2	3	4	5	6
Frequency	8	10	11	16	20	25

7. Compare range and quartile deviation.
8. Write down the normal equation for fitting an exponential curve $y = ab^x$
9. Describe negative and positive correlation. Give example.
10. Give example for seasonal and cyclic variations in time series.
11. Mention any three uses of index numbers.
12. Write Marshal-Edgeworth index number and Dorbish and Bowley's index number.

SECTION-B

Each question carries 5 Marks.

Maximum Marks that can be scored in this section is 30.

13. Discuss about National Sample Survey Organization.
14. For a group of 150 candidates, the mean and standard deviation of scores were found to be 38 and 16 respectively. Later on it was found that the scores 45 ad 53 were misread as 54 and 35 respectively. Find the standard deviation of corrected figures.
15. Explain skewness and kurtosis.
16. Explain the least square method of fitting a parabola.
17. Consider the following set of rankings for a sample of 10 elements. Compute Spearman's rank correlation coefficient for the data.

Element	1	2	3	4	5	6	7	8	9	10
X	10	6	7	3	4	2	8	5	1	9
Y	8	4	10	2	5	7	6	3	1	9

18. How trend is measured using Moving Averages?
19. Compute Marshal-Edgeworth's index number and Fisher's ideal index number for the following data.

Items	Price		Quantity	
	2001	2003	2003	2004
I	8	18	48	39
II	5	12	11	6
III	6	10	6	5

SECTION-C

Answer *any one* Question and carries 10 marks

20. The following are the figures of the production in a sugar factory. Fit a straightline trend for the data and estimate the value for the year 2009.

Year:	2001	2002	2003	2004	2005	2006	2007	2008
Production:	92	95	94	91	92	89	87	81

21. Police records show the following numbers of daily crime reports for a sample of days during the winter months and a sample of days during the summer months. Compare the variability of the two periods.

Winter	18	20	15	16	21	20	12	16	19	20
Summer	28	18	24	32	18	29	23	38	28	18

SEMESTER II

STA 2C 02- PROBABILITY THEORY

Contact Hours per week: 4
Number of credits: 3
Number of Contact Hours: 72
Course Evaluation: External 60 Marks+ Internal 15 Marks
Duration of Exam: 2 Hours

Blue Print for Question Paper Setting / Scrutiny						
Max. Marks: 60						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE 1	MODULE 2	MODULE 3	MODULE 4
			25Hrs	12Hrs	15 Hrs	20 Hrs
			28 Marks	16 Marks	16 Marks	19 Marks
Expected mark >>>>						
A	2	1	2			
		2	2			
		3	2			
		4	2			
		5		2		
		6		2		
		7		2		
		8			2	
		9			2	
		10			2	
		11				2
		12				2
B	5	13	5			
		14	5			
		15		5		
		16		5		
		17			5	
		18			5	
		19				5
C	10	20	10			
		21				10
Actual Mark >> >>			**28**	**16**	**16**	**19**

Question Paper setter has to give equal importance to both theory and problems in section B and C.

STA 2C 02 – PROBABILITY THEORY

Time: 2 Hours Max Marks: 60

SECTION-A

Each question carries 2 Marks.

Maximum Marks that can be scored in this section is 20.

1. Define random experiment. Write an example.
2. State addition theorem of probability for three events.
3. Define independence of events.
4. Write the axiomatic definition of probability.
5. Distinguish between discrete and continuous random variables.
6. Define distribution function. Write any two properties of distribution function.
7. Let $f(x) = 2x+3$, $0 < x < 1$; 0 otherwise. Verify whether $f(x)$ is a probability density function or not.
8. Define expectation of a random variable. Write any two properties of expectation.
9. Show that $V(aX + b) = a^2V(X)$.
10. Let X_1 and X_2 be two independent random variables. Prove that $M_{X1+X2}(t) = M_{X1}(t)M_{X2}(t)$.
11. When do you say that two random variables are independent?
12. State and prove the multiplication theorem of expectation

SECTION-B

Each question carries 5 Marks.

Maximum Marks that can be scored in this section is 30.

13. State and prove Baye's theorem.
14. A problem in Statistics is given to 3 students A,B and C whose chances of solving it are ½, ¾ and ¼ respectively. What is the probability that the problem will be solved?
15. Let f(x) = 1, 0<x<1; 0 otherwise. Find the distribution of Y= -2log X
16. For a discrete r.v. X with probability distribution

X	-2	-1	0	1	2
P(X=x)	0.1	0.2	0.3	k	0.2

find the value of (i) k (ii) $p(-1 \leq X \leq 1)$

17. Let f(x) = 1, 0<x<1; 0 otherwise. Find the distribution of Y= -2log X
18. For a discrete r.v. X with probability distribution

X	-2	-1	0	1	2
P(X=x)	0.1	0.2	0.3	k	0.2

find the value of (i) k (ii) $p(-1 \leq X \leq 1)$

19. Prove that V(X) = E[V(X|Y)] + V[E(X|Y)]

SECTION-C

Answer *any one* Question and carries 10 marks

20. A committee of 4 people is appointed from 3 officers of the productiondepartment, 4 officers of the purchase department, 2 officers of the

sales department and 1 charted accountant. Find the probability of forming the committee in the following manner.

(1) There must be one from each category

(2) It should have at least one from the purchase department

(3) The charted accountant must be in the committee.

21. If X and Y are two r.v.s having the joint pdf $f(x\ y) = 2 - x - y\ ;\ 0 < x, y < 1$. Find ρ_{xy}.

SEMESTER III

STA 3C 03- PROBABILITY DISTRIBUTIONS AND SAMPLING THEORY

Contact Hours per week: 4
Number of credits: 3
Number of Contact Hours: 90
Course Evaluation: External 60 Marks+ Internal 15 Marks
Duration of Exam: 2 Hours

Blue Print for Question Paper Setting / Scrutiny						
Max. Marks: 80						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE 1	MODULE 2	MODULE 3	MODULE 4
			30Hrs	25Hrs	10Hrs	25 Hrs
			28 Marks	21 Marks	9 Marks	21 Marks
Expected mark >>>>						
A	2	1	2			
		2	2			
		3	2			
		4	2			
		5		2		
		6		2		
		7		2		
		8			2	
		9			2	
		10				2
		11				2
		12				2
B	5	13	5			
		14	5			
		15		5		
		16		5		
		17		5		
		18			5	
		19				5
C	10	20	10			
		21				10
Actual Mark >> >>			**28**	**21**	**9**	**21**

Question Paper setter has to give equal importance to both theory and problems in section B and C.

STA 3C 03- PROBABILITY DISTRIBUTIONS AND SAMPLING THEORY

Time: 2 Hours Max Marks: 60

SECTION-A

Each question carries 2 Marks.

Maximum Marks that can be scored in this section is 20.

1. Define Binomial random variable.
2. Derive the mean of a Poisson random variable.
3. If X is a random variable with a continuous distribution function F. Then show that F ~ Uniform[0, 1]
4. Write the relationship between normal and log normal distributions.
5. Define convergence in probability.
6.
7. State Lindberg-Levy central limit theorem.
8. Discuss simple random sampling.
9. Distinguish between stratum and cluster.
10. Distinguish between parameter and statistic.
11. State the additive property of Chi square distribution.
12. Establish the relationship between t and F distributions.

SECTION-B

Each question carries 5 Marks.

Maximum Marks that can be scored in this section is 30.

13. Let X and Y are two independent and identically distributed geometric random variables. Then show that the conditional distribution of X|X+Y = n is uniform
14. Establish the memory less property of geometric distribution.
15. Suppose that the lifetime of an electronic device follows exponential distribution with mean 1. Determine the upper bound of $P(|X - 1| \geq 2)$ using Chebyshev's inequality.
16. Let X_n assumes the values $\frac{1}{\sqrt{n}}$ and $-\frac{1}{\sqrt{n}}$ with probabilities $\frac{2}{3}$ and $\frac{1}{3}$ respectively. Check whether the weak law of large numbers holds good for the sequence $\{X_n\}$ of independent random variables.
17. State and prove Bernoulli's law of large numbers.
18. Explain various methods for sampling.
19. Establish the sampling distribution of the sample variance of random sample drawn from Normal distribution.

SECTION-C

Answer *any one* Question and carries 10 marks

20. X is a normal variate with mean 42 and standard deviation 4. Find the probability that a value taken by X is (i) less than 50, (ii) greater than 50, (iii) greater than 40, (iv) in between 40 and 50 and (v) equal to 45.
21. Derive Student's t distribution and state some of its applications.

SEMESTER IV

STA 4C 04 - STATISTICAL INFERENCE AND QUALITY CONTROL

Contact Hours per week: 4
Number of credits: 3
Number of Contact Hours: 90
Course Evaluation: External 60 Marks+ Internal 15 Marks
Duration of Exam: 2 Hours

Blue Print for Question Paper Setting / Scrutiny						
Max. Marks: 60						
Question Paper			Syllabus			
Sections or Parts	Mark	Question Numbers	MODULE 1	MODULE 2	MODULE 3	MODULE 4
			30 Hrs	35 Hrs	10 Hrs	15Hrs
			26 Marks	30Marks	9 Marks	14 Marks
Expected mark >>>>						
A	2	1	2			
		2	2			
		3	2			
		4		2		
		5		2		
		6		2		
		7		2		
		8		2		
		9			2	
		10			2	
		11				2
		12				2
B	5	13	5			
		14	5			
		15		5		
		16		5		
		17			5	
		18				5
		19				5
C	10	20	10			
		21		10		
Actual Mark >> >>			**26**	**30**	**9**	**14**

Question Paper setter has to give equal importance to both theory and problems in section B and C.

STA 4C 04 - STATISTICAL INFERENCE AND QUALITY CONTROL

Time: 2 Hours Max Marks: 60

SECTION-A

Each question carries 2 Marks.

Maximum Marks that can be scored in this section is 20.

1. Define minimum variance unbiased estimator.
2. Write the confidence for the mean of normal distribution when population standard deviation is unknown.
3. State Cramer – Rao inequality.
4. Define simple and composite hypotheses.
5. Distinguish between Type I error and Type II error.
6. When do you use Yate's correction?
7. Write the critical regions of large sample test.
8. Write a suitable situation where ANOVA test is applicable.
9. Write short note on Median Test
10. Write the test statistic of Wilcoxon-signed rank and identify its asymptotic distribution
11. What are the causes of variation in quality control?
12. Distinguish between consumer's risk and producer's risk.

SECTION-B

Each question carries 5 Marks.

Maximum Marks that can be scored in this section is 30.

13. Find the maximum likelihood estimator of λ for the Poisson distribution.
14. Establish confidence interval for the difference of proportions of two binomial populations
15. The continuous random variable X has the frequency function

$$f(x,\theta) = \frac{1}{\theta}, 0 \le x \le \theta, \qquad 0\ otherwise$$

It is desired to test the hypothesis $H_0: \theta = 1\ against H_1: \theta = 2$ using a single observation X. X≥ 0.95 is used as the critical region. Evaluate Type I error and Type II error.

16. A sample of 25 boys who passed SSLC examination are found to have mean marks 50 with standard deviation 5 for English. The mean marks of 18 girls are found to be 48 with standard deviation 4 for the same subject. Does this indicate any significance difference between the marks of boys and girls assuming the population standard deviation are equal?
17. A sample of 10 men was used in a study to test the effects of a relaxant on the time required to fall asleep for male adults. Data for 10 subjects showing the number of minutes required to fall asleep with and without the relaxant follow. Use a 0.05 level of significance to determine whether the relaxant

reduces the time required to fall asleep. Perform sign test and draw your conclusion.

Subject	1	2	3	4	5	6	7	8	9	10
Without Relaxant	15	12	22	8	10	7	8	10	14	9
With Relaxant	10	10	12	10	8	5	9	7	11	6

18. During an examination of equal lengths of clothes, the following numbers of defects were observed 2, 3, 4, 0, 5, 6, 7, 4, 3, 2. Draw a control chart for the number of defects and comment whether the process is under control.
19. Construct control chart of mean and range for the following data and comment on the state of control.

Sample No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Sample values	42	46	66	36	57	77	87	45
	64	53	81	87	99	89	56	78
	44	75	34	60	46	56	39	34
	75	89	4	79	77	48	121	98
	86	44	75	66	44	40	56	65

($A_2 = 0.577$, $D_3 = 0$, $D_4 = 2.115$)

SECTION-C

Answer *any one* Question and carries 10 marks

20. Explain the desirable properties of a good estimator. Give examples.
21. The following table gives number of refrigerators sold by four different salesmen in three months. a) Test whether there is any significant difference in the average sales by the four salesmen b) Determine whether the sales significantly varying with respect to varying months.

Months	**Salesmen**			
	I	II	III	IV
May	50	40	48	39
June	46	48	50	45
July	39	44	40	39