HSNC UNIVERSITY, MUMBAI

(2024-2025)

Ordinances and Regulations With

Respect to

Choice Based Credit System (CBCS)

For the Programmes Under

The Faculty of Science and Technology

Framed According to the National Education Policy (NEP 2020)

To be implemented from Academic Year: 2024-2025

For the Course

Statistics

Curriculum – Second Year Undergraduate

Semester-III and Semester -IV

2024-2025

HSNC UNIVERSITY, MUMBAI

Board of Faculty of Science & Technology

Board of Studies in the Subjects of Statistics

1) **Name of Chairperson/Co-Chairperson/Coordinator:-**

   a) **Dr Asha Jindal,** Professor and (UG: Head &PG: Coordinator), Department of Statistics, K. C. college, HSNC University Churchgate, Mumbai –400 020. Email ID- asha.jindal@kccollege.edu.in Mobile no- 9821235627

2) **Two to five teachers each having minimum five years teaching experience amongst the full time teachers of the Departments, in the relevant subject.**

   a) **Dr. S. B. Muley,** Assistant Professor, Department of Statistics, K. C. college, HSNC University Churchgate, Mumbai – 400 020. Email ID sakharam.muley@kccollege.edu.in , Mobile No- 9323817918

   b) **Mrs. Pratiksha Kadam,** Assistant Professor, Department of Statistics, K. C. college, HSNC University Churchgate, Mumbai – 400 020. Email ID pratiksha.kadam@kccollege.edu.in , Mobile No- 7507162816

   c) **Ms. Shailaja Rane,** Assistant Professor, Department of Statistics, K. C. college, HSNC University Churchgate, Mumbai – 400 020. Email ID shailaja.rane@kccollege.edu.in, Mobile No- 7506986359

3) **One Professor / Associate Professor from other Universities or professor / Associate Professor from colleges managed by Parent Body;**

   a) **Dr Anjum Ara Ahmed;** Professor and I/C Principal, Rizvi College, Mumbai. Email ID **anjumahmed8@gmail.com, Mobile No-** 8451046220

4) **Four external experts from Industry / Research / eminent scholar in the field relevant to the subject nominated by the Parent Body;**

   a. **Prof. Suresh Kumar Sharma,** Senior Professor, Department of Statistics, Panjab University, Chandigarh.
   Email ID ssharma643@yahoo.co.in, **Mobile No-**9815911381

   b. **Mr Mukesh Jain,** Vice President and Chief Technological Officer, Capgemini. Email ID mdjain@hotmail.com, **Mobile No-**7972637347.

   c. **Dr Santosh Gite,** Professor, Dept. of Statistics, University of Mumbai, Mumbai. Email ID santgite@yahoo.com, **Mobile No-** 9167157717.

   d. Mr Prashant Kumar Nair, Director, Geo Spatial Analytics Global Lead, Intelligent Analytics, Nielsen Connect, Email ID prashantkumar.nair@nielsen.com , **Mobile No-**9833747057.

5. Top rankers of the Final Year Graduate and Final Year Post Graduate examination of previous year of the concerned subject as invitee members for discussions on framing or revision of syllabus of that subject or group of subjects for one year.

   a) **Mr. Chinmay Mokal** (Postgraduate student 22-23) Email ID- chinmay30mokal@gmail.com ; Mobile no- 9372323901

   b) **Ms. Aanchal Goyal** (undergraduate student 22-23) Email ID- aanchalgoyal2703@gmail.com ; Mobile no- 7738886488

Statistics

Part 1- Preamble

B. Sc. Statistics program is of minimum 120 credits cover six semesters. Statistics is the language of the uncertainties riddled modern information age. Statistics facilitates the choice making process by quantifying the element of chance or uncertainties. It's descriptive and inferential roles not only formulate the basis of the expansion of almost all the disciplines of the modern world, but also provide an array of non-traditional employment avenues starting from that of sport analysts to business analysts to actuaries. The thrust of the course is to prepare students to enter into a promising career even after graduation, as also provide to them a platform for pursuing higher studies resulting in post-graduate or doctorate degrees. The program has some unique features like number of elective courses and extensive computer training of statistical computations including standard software packages like IBM SPSS, Excel, MINITAB, R and PYTHON. The integration of continuous assessment, project work, and applied learning ensures that students develop not only academic knowledge but also critical thinking, ethical decision-making, and effective communication skills, essential for careers in statistics, data science, finance, research, and beyond. The curriculum supports the objectives of the National Education Policy (NEP 2020) by focusing on interdisciplinary collaboration, sustainable practices and lifelong learning.

1. Programme Outcomes

1) **Understanding Core Scientific Concepts**: Students will gain clear insight and understanding to recall key scientific principles across various fields. A well-established foundational knowledge of the subject will play a crucial role for deeper learning and future studies.
2) **Commit to Lifelong Scientific Learning**: Students will cultivate a habit of continuous learning and shall learn to stay updated with the latest scientific and technological advancements. This mindset will ensure that they remain relevant, engaged and informed throughout their future academic journey.
3) **Abilities to Analyse and Evaluate**: Students will learn to classify and scrutinize complex problems into manageable parts, critically analyse data, and evaluate potential solutions to scientific problems.
4) **Assessing Ethical Implications**: Students will be trained to evaluate the ethical dimensions of research and technological innovations, ensuring that their decisions consider societal impacts and they adhere to ethical standards. This is vital for responsible and sustainable practices.

5) **Design Experiments and Innovate**: Students will learn to design and conduct experiments, developing innovative solutions to challenges through Research Projects. They will also learn to evaluate their results and refine their experimental approaches over time.
6) **Application of Scientific and Technical Knowledge to Real-World Problems:** Students will use their scientific and Technical knowledge and expertise to identify and solve real-world problems. This would involve applying theoretical concepts to practical situations, bridging the gap between classroom learning, Industry-academia and real-life applications.
7) **Communicating Scientific Findings Effectively**: Students will develop the ability to communicate scientific information clearly and effectively, both in writing and verbally. Whether presenting research findings or writing technical reports, clear communication is key to knowledge sharing and collaboration.
8) **Foster an Interdisciplinary Approach**: Students will cultivate leadership and teamwork skills, enabling them to collaborate effectively in diverse, interdisciplinary teams. Leadership qualities such as decision-making and delegation will help them achieve successful outcomes in various projects.
9) **Promote Environmental Sustainability**: Students will understand the environmental impact of scientific activities and advocate for sustainable practices. By considering environmental factors in their work, they will contribute to the long-term health of the planet.
10) **Enhanced Skills for Entrepreneurship and Employability**: Students will be equipped with essential skills for entrepreneurship and employability, focusing on job readiness, soft skills, and practical business knowledge. Hands-on experience through internships and mentorship will further enhance their readiness for the job market and entrepreneurial ventures.

2. Program Objective: The main objectives of the course are-

1) Graduates will build a strong foundation in core statistical principles and methodologies, allowing them to apply statistical techniques to real-world problems.
2) Graduates will commit to lifelong learning and remain updated on emerging statistical methods, computational tools, and technological advancements.
3) Graduates will develop problem-solving and analytical skills, enabling them to apply statistical models and methods in various industries, including finance, healthcare, and research.

4) Graduates will demonstrate professionalism, ethical responsibility, and effective communication in multidisciplinary and collaborative environments.

5) Graduates will possess the skills necessary for employability and entrepreneurship, and they will be well-prepared to pursue higher education or careers in data analysis, statistical consulting, or related fields.

3. Process adopted for curriculum designing.

The members of Department of Statistics initially drafted the syllabus. The draft syllabus was shown to Industry Partners, Academic Partners and Research Institute Partners through mail and in person invited to college. They suggested some changes. These changes were incorporated.

4. Salient features, how it has been made more relevant.

Statistics deals with collection, organization, analysis and interpretation of data. Statistical knowledge is very important as it helps to use appropriate methodologies for collecting data, tools for employing analysis and interpretation of results. It also provides us with techniques which are important in designing and planning of experiments.

A lot of data is generated at each and every moment. Data literacy has become crucial and indispensable to the society. Statistics has the quality of quantifying and measuring uncertainty which helps in assessing risk. It helps in extracting the meaningful information from the data, making predictions and taking decisions. Study of data has become an integral part of education, business, and overall human progress. This has put Statistics on the center stage of teaching, research, policy making and development all over the globe.

The S.Y.B.Sc Statistics syllabus is a Choice based credit system comprising of two papers having three units each in both the semesters. The current course is designed to enhance the knowledge of the subject. While designing of the syllabus care has been taken to balance the fundamental techniques of Statistics with soft skills like analysis using Statistical Software.

Several radical changes have been made in the syllabi. Firstly, the concept of having seperate papers in practicals has been abondoned. All the numerical / practical work has been integrated with the teaching of theory courses. Secondly, majorly the numerical /practical work be carried out on computers.

The course would give the students option to develop skills in areas which have direct relevance to employability in insurance and finance industries, banks, econometrics, quality control, pharmaceutical, medical statistics, agricultural statistics, weather forecasting, civil services, stock market, machine learning and artificial intelligence related job opportunities in Statistics.

5. Program-Specific Outcomes (PSOs)

1. **Comprehension of Statistical Concepts**: Develop a strong foundational understanding of probability, statistical distributions, and inference.
2. **Analytical and Critical Thinking**: Apply statistical techniques to analyze data, solve problems, and make informed decisions.
3. **Technological Proficiency**: Gain hands-on expertise in statistical software such as R, Python, Excel, and Power BI for data analysis and visualization.
4. **Data Interpretation and Experimentation**: Design experiments, interpret data meaningfully, and apply statistical reasoning in real-world applications.
5. **Interdisciplinary Integration**: Collaborate with other disciplines and integrate statistical methods for diverse applications like business, health, and environment.
6. **Communication and Collaboration**: Communicate statistical concepts effectively in both written and verbal formats, fostering teamwork and interdisciplinary collaboration.
7. **Ethics and Sustainability**: Ensure ethical data handling, emphasizing sustainable practices and societal responsibility.

6. Learning Outcomes:

Semester III

1. **Probability Distributions (STA201B)**
    - **Unit I**: Understanding and computing Bivariate probability distributions, covariance, and correlation.
    - **Unit II**: Understanding and computing Moment generating functions, cumulant generating functions and Probability generating functions.
    - **Unit III**: Properties and applications of standard distributions (Binomial, Poisson etc) through MGF, CGF and PGF.
2. **Theory of Sampling (STA202B)**
    - **Unit I**: Understand sampling techniques and their applications.
    - **Unit II**: Implement stratified sampling with proportional and optimal allocation.
    - **Unit III**: Learn advanced methods like ratio and regression estimation.
3. **STA203B – Statistical Methods**:
    - **Unit I**: Understand estimation, related theory of point estimation, theory of testing and confidence sets, together with their applications.
    - **Unit II**: Develop test procedure and to learn applications of Z, t, F, Chi-Square. Perform hypothesis testing for various datasets.
    - **Unit III**: Conduct ANOVA and design quality control charts for Variables and Attributes(X-bar, R, and p-charts) and learn the application of Quality Control in Industry.

4. **Statistical Data Analysis using R Programming (STA209B)**
    - **Unit I: Data Wrangling**: Students will learn basics of R like Data Types, Variable Type and Data Manipulation (Pre-processing, cleaning, and transforming data using R) of datasets for analysis, ensuring readiness for modeling and visualization.

    Unit II: Data Visualization: Students will create visualizations to explore and present data effectively using ggplot2.

    - **Unit III: Modeling Techniques**: Students will compute basic statistics, analyze relationships using correlation and build, validate & learn interpretations of statistical models like linear regression for data-driven decision-making using real datasets.
5. **Python for Statistical Computing (STA210C/B)**
    - **Unit I: Computing**: Students will learn basics of Python like Numpy arrays, Pandas data frames, Reading and writing data, Control statements etc.
    - **Unit II: Data Visualization**: Students will create visualizations and interactive dashboards to explore and present data effectively using libraries like Matplotlib, Seaborn, and Plotly.
    - **Unit III: Modeling Techniques**: Students will compute basics of Hypothesis Testing, analyze relationships using correlation and build, validate & learn interpretations of statistical models like linear regression for data-driven decision-making using real datasets.

Semester IV

1. **Probability and Sampling Distributions (STA204B)**
    - **Unit I**: To obtain the central location, dispersion Moment Generating Distribution, Cumulant Generating Function and Characteristics Function of the Continuous distributions like Exponential, Uniform, Triangular and Gamma distributions.
    - **Unit II**: To obtain the central location, dispersion Moment Generating Distribution, Cumulant Generating Function and Characteristics Function of the Continuous distributions like Beta and Normal distributions.
    - **Unit III**: Learn derivations of pdf and other characteristics of t, chi-square, and F distributions with applications.
2. **ANOVA and Design of Experiments (STA205B)**
    - **Unit I**: Learn to derive various estimators of parameters and Sum of Squares Analysis of variance (one-way and two-way) with application.

- **Unit II**: Understanding concepts and learn to derive various estimators of parameters and Sum of Squares of Completely randomized and randomized block designs with application.
- **Unit III**: Understand Factorial experiments and its application in real life situation /datasets.

3. **STA206B – Operations Research Techniques**:

   - **Unit I**: Learn mathematical formulation of real-life situations using LPP and study methods to solve the formulated problems using graphical and simplex methods.
   - **Unit II**: Learn mathematical formulation of real-life situations using Transportation., Assignment Problems and Apply transportation and assignment problem-solving techniques.
   - **Unit III**: Learn mathematical formulation of real-life situations using Game Theory, Decision Theory and to study methods to solve the formulated problems manually and using TORA software.

4. **Statistical Modeling and Data Analysis in SPSS (STA211B)**
   - **Unit I**: **Managing Data in SPSS**: Students will learn Data entry, manage datasets in SPSS, including manipulation, transformation, and handling incomplete data.
   - **Unit II**: **Charts in SPSS**: Students will create and interpret a variety of visualizations for univariate, bivariate, and multivariate data.
   - **Unit III**: **Advanced Analysis**: Students will apply advanced techniques like Multi-Dimensional Scaling (MDS) and Correspondence Analysis and integrate R scripts within SPSS.
5. **Open Elective: Introduction to Six Sigma with MS Excel (STA212 B)**
   - **Unit I: Introduction to Six Sigma**

     CO1: Explain the principles of Six Sigma, Lean methodology, and the components of the DMAIC framework, including the concept of quality and waste reduction.

     CO2: Develop SIPOC diagrams, problem statements, and project charters while identifying key terms like VOC, VOP, CTQ, and COPQ.
   - **Unit II: Measure and Analyse**

     CO3: Perform Measurement System Analysis (MSA) and Process Capability analysis to assess and improve process performance.

     CO4: Utilize statistical tools and data visualization techniques to analyze process

data and identify root causes.

- **Unit III: Improve and Control**

  CO5: Apply improvement techniques such as Kaizen, Poka Yoke, and Multi-Voting to optimize processes effectively.

  CO6: Design and interpret control charts to monitor process stability and implement Six Sigma principles in real-world scenarios.

Part 2- The Scheme of Teaching and Examination is as under:Semester – III Summary

Sr. No.	Choice Based Credit System			Subject Code	Remarks
1	Core Course (**Statistics**)			STA201B STA201D STA202B STA202D STA203B STA203D	Nil
2	Elective Course	Discipline Specific Elective (DSE) Course			
		2.1	Interdisciplinary Specific Elective (IDSE) Course		
		2.2	Dissertation/Project		
		2.3	Generic Elective (GE) Course	STA210B	
3	Ability Enhancement Courses (AEC)				
	Skill Enhancement Courses (SEC)			STA201C STA209D	

First Year Semester III Internal and External Detailed Evaluation Scheme

Sr. No.	Semester	Subject Code	Subject Title	NEP Course Type	Hours Per Week						Seasonal Evaluation Scheme (Internal + External)			Total Marks
					Units	S.L.E.	L	T	P	Credit	S.L.E	PA/AT	SEE	
1	III	STA201B	Probability Distributions	Major	3	20%	3	0	0	3	10	5	60	100
		STA201D	**Computer Applications & Practical** Based on Probability Distributions		1				2	1			25	
2	III	STA202B	Theory of Sampling	Major	3	20%	3	0	0	3	10	5	60	100
		STA202D	**Computer Applications & Practical** Based on Theory of Sampling		1				2	1			25	
3	III	STA203B	Statistical Methods	Minor	3	20%	3	0	0	3	10	5	60	100
		STA203D	**Computer Applications & Practical** Based on Statistical		1				2	1			25	

			Methods											
4	III	STA201C	**Statistical Data Analysis using R Programming**	SEC	2		2	0	0	2	0	0	50	75
		STA209D	**Computer Applications & Practical** Based on Statistical Data Analysis using R Programming		1				2	1			25	
5	III	STA210B	**Wanna Learn Python: Statistical Computing with Python**	OE	3	20 %	3	0	0	3	10	15	50	75

***One to two lectures to be taken for CONTINUOUS self -learning Evaluation.**

Second Year Semester III - Units – Topics – Teaching Hours

S. No.	Subject Code	Subject Unit Title		Hours	Total No. of hours	Credit	Total Marks
1	STA201B	I	Bivariate Probability Distributions	15	45	3	100 (60+40)
		II	Moment Generating Function, Cumulant generating Function, Probability Generating Function	15			
		III	M.G.F. & C.G.F. of Standard Discrete Distributions	15			
	STA201D	IV	Practical based on STA201B	30	30	1	
2	STA202B	I	Concepts of Sampling and Simple Random Sampling	15	45	3	100 (60+40)
		II	Stratified Random Sampling	15			
		III	Ratio and Regression Estimation Method	15			
	STA202D	IV	Practical based on STA202B	30	30	1	
3	STA202B	I	Elementary topics on Estimation and Testing of Hypothesis	15	45	3	100 (60+40)
		II	Applications of Z, T, Chi-square and F	15			
		III	ANOVA and Statistical Quality Control	15			
	STA202D	IV	Practical based on STA203B	30	30	1	
4	STA201C	I	Introduction to R programming	15	30	2	**75** (50+25)
		II	Data Visualization using GGPLOT2	15			
	STA209D	III	Practical based on STA201C	30	30	1	
5	STA210B	I	Introduction to Python	15	45	3	**75** (50+25)
		II	Data Handling and Visualisation	15			
		III	Statistical Computing and Modeling	15			

- **Lecture Duration – One hour**
- **One Credit =15 class room teaching hours.**

L: Lecture: Tutorials P: Practical Ct-Core Theory, Cp-Core Practical, SLE- Self learning evaluation CT-Commutative Test, SEE- Semester End Examination , PA-Project Assessment, AT- Attendance

Part -3 Detailed Scheme Theory

Curriculum Topics along with Self-Learning topics - to be covered, through self-learning mode along with the respective Unit. Evaluation of self-learning topics to be undertaken before the concluding lecture instructions of the respective UNIT

SY BSc SEM III Syllabus

Course Code: STA201B
Course Title: Probability Distributions (Major-Paper I)

Unit I	**Bivariate Probability Distributions**	**15**
1.1	Two dimensional Discrete random variables: Joint Probability mass function and its properties, Distribution function of (X,Y) and its properties,-Definition of raw and central moments, covariance, correlation coefficient, Independence and correlation between two variables,-Marginal and conditional probability distributions. Conditional expectation, conditional variance.	**05**
1.2	Continuous bivariate random variables: Joint Probability density function and its properties, Distribution function of (X,Y) and its properties, Definition of raw and central moments, covariance, correlation coefficient, Independence and correlation between two variables.	**05**
1.3	Marginal and conditional probability distributions -Conditional expectation, conditional variance - Regression Function. Transformation of Random Variables and Jacobian of transformation with illustrations.	**05**
Unit II	**Moment Generating Function, Cumulant generating Function, Probability Generating Function**	**15**
2.1	Moment Generating Function Definition and Properties: Effect of change of origin and scale. M.G.F of sum of two independent random variables X and Y, Extension of this property for n independent random variables and for n i.i.d. random variables. All above properties with proof, Uniqueness Property without proof. Raw moments using M.G.F: using expansion method and using derivative method.	**05**
2.2	Cumulant generating Function: Definition, Properties: Effect of change and origin and scale, Additive Property of C.G.F. and cumulants Both properties with proof. Obtaining Cumulants using C.G.F. Derivation of relationship between moments and cumulants upto order four. Characteristic Function: Definition and properties (without Proof), Examples of obtaining raw moments and central moments up to order four using M.G.F. and C.G.F. for continuous and discrete distributions.	**05**
2.3	Definitions of generating function and probability generating function. Expression for mean and variance in terms of generating functions. Definition of a convolution of two or more sequences. Generating function of a convolution.	**05**

	Degenerate distribution (One point distribution) P(X=c), Mean, Variance, Use of Degenerate distribution. Discrete Uniform distribution. Mean, Variance, coefficient of skewness using m.g.f., Bernoulli distribution. Mean, Variance, coefficient of skewness using m.g.f.	
Unit III	**M.G.F. & C.G.F. of Standard Discrete Distributions**	**15**
3.1	Binomial distribution: Mean, Variance, Measures of skewness and Kurtosis based on moments using M.G.F. and C.G.F. , Nature of probability curve, Mode, Additive property ,If X follows Binomial, then to find distribution of n-X. Recurrence relation for moments with proof, Relation between Bernoulli and Binomial using m.g.f. Transformation of random Variable (Univariate) : examples based on it. Real life examples of Binomial distribution. Poisson Distribution: Mean, Variance, Measures of skewness and Kurtosis based on moments using M.G.F. and C.G.F. Nature of probability distribution with change in the values of parameters, Mode, Additive property. Recurrence relation for moments with proof for μ'_{r+1}, μ_{r+1} . If X and Y are two independent Poisson variables Conditional distribution of X given X+Y with proof, Poisson distribution as limiting distribution of Binomial (with proof),Real life examples of Poisson distribution.	**05**
3.2	Geometric Distribution: Definition in terms of No. of failures and No. of trials. Mean, Variance, M.G.F., Mean and Variance using M.G.F.,C.G.F., Mean and Variance,μ_3, μ_4 using C.G.F., Coefficients of skewness and kurtosis and nature of the probability distribution. Lack of memory property with proof. If X and Y are two i.i.d. Geometric variables; Conditional distribution of X given X+Y with proof. Distribution of sum of k i.i.d. Geometric variables. Negative Binomial Distribution: Definition, Mean, Variance, M.G.F., Mean and Variance using M.G.F., C.G.F., Recurrence relation for central moments, Mean, Variance,μ_3, μ_4 using C.G.F., Coefficients of skewness and Kurtosis and nature of probability distribution. Lack of Memory property with proof. Recurrence relation for probabilities, Fitting of distribution. Limiting distribution of Negative Binomial distribution (with proof).	**05**
3.3	Hyper geometric Distribution: Definition, Mean, Variance, Limiting distribution of Hyper geometric distribution(with proof),If X and Y are two independent Binomial variables Conditional distribution of X given X+Y (with proof) Truncated distribution: Definition, Truncated Binomial and Truncated Poisson Distribution:(truncated at 0),Probability mass function, mean and variance. Real life situations of Geometric, Negative Binomial and Hypergeometric distributions.	**05**

Self-Learning topics (Unit wise)

Unit	Topics
2.1	Binomial Distribution: Additive Property
2.2	Poisson Distribution: Additive Property
2.3	**Geometric Distribution:** Distribution of k i.i.d. Geometric variables
2.3	**Negative Binomial Distribution:**Definition, Mean, Variance

1.1	Joint Probability mass function and its properties, Distribution function of (X, Y)and its properties
3.2	Joint Probability density function and its properties, Distribution function of (X, Y) and its properties
3.3	**Transformation of discrete & continuous random variables**

Online **Resources**

"Probability and Statistics" by Prof. Somesh Kumar, Department of Mathematics, IIT Kharagpur Source : https://nptel.ac.in/courses/111/105/111105090/ on unit 3.3

Course Code: STA202B
Course Title: Theory of Sampling (Major-Paper II)

Unit 1	**Concepts of Sampling and Simple Random Sampling**	15
1.1	Population, Population unit, Sample, Sample unit, Parameter, Statistic, Estimator, Bias, Unbiased Mean square error and Standard error. Census survey, Sample Survey. Steps in conducting sample survey with examples on designing appropriate Questionnaire.	**05**
1.2	Concepts of Sampling and Non-sampling errors. NSSO, CSO and their functions. Concepts and methods of Probability and Non-Probability Sampling. **Simple Random Sampling: (SRS):**Definition, Sampling with and without replacement (WR/WOR).Lottery method and use of Random numbers to select simple random sample. Estimation of population mean and population total.	**05**
1.3	Expectation and Variance of the estimators, Unbiased estimator of variance of these estimators. (WR/WOR). Estimation of population proportion. Estimation of sample size based on a desired accuracy in case of SRS for variables and attributes. (WR/WOR).	**05**
Unit 2	**Stratified Random Sampling**	**15 hrs**
2.1	Need for Stratification of population with suitable examples. Stratified Random Sampling: Definition and its advantages.	5 hrs
2.2	Stratified Random Sampling method, Estimation of population mean and population total in case of Stratified Random Sampling (WOR within each stratum). Expectation and Variance of the unbiased estimators, Unbiased estimators of variances of these estimators	**05**
2.3	Proportional allocation, Optimum allocation with and without varying costs. Comparison of Simple Random Sampling, Stratified Random Sampling using Proportional allocation and Neyman allocation	**05**
Unit 3	**Ratio and Regression Estimation Method**	**15**
3.1	**Ratio and Regression Estimation assuming SRSWOR:** Ratio Estimators for population Ratio, Mean and Total. Expectation and Mean Square Error of the Estimators. Estimators of Mean Square Error. Uses of Ratio Estimator	**05**

3.2	Regression Estimators for population mean and population total. Expectation and Variance of the Estimators assuming known value of regression coefficient 'b'. Estimation of 'b'. Resulting variance of the estimators. Uses of regression Estimator. Comparison of Ratio, Regression and mean per Unit estimators.	**05**
3.3	Systematic sampling, Cluster sampling and Two Stage Sampling, Double Sampling: Introduction its application, Differences and Case Studies.	**05**

Self-Learning topics (Unit wise)

Unit	**Topics**
1.1	Census Survey, Sample Survey. Steps in conducting sample survey with examples on designing appropriate Questionnaire.
1.4	Estimation of sample size based on a desired accuracy in case of SRS for variables
2.1	Need for Stratification of population with suitable examples. Definition of Stratified Sample. Advantages of Stratified Sampling.
3.1	Ratio Estimators for population Ratio, Mean and Total. Expectation & MSE of the Estimators
3.3	Introduction to Systematic sampling, Cluster sampling with suitable illustrations. Comparing various Sampling Methods.

Online Resources

'Business Statistics' by PROF. M. K. BARUA, Department of Management Studies, IIT Roorkee, available on the NPTEL portal **https://nptel.ac.in/courses/110/107/110107114/# for 2.1**
"Sampling Theory" **Swayam Prabha Course, Channel 16, MOE, GOI by Prof. Shalabh, IIT Kanpur** Source: https://www.youtube.com/playlist?list=PLqMl6r3x6BUTP4XPysDab-RrLAt4_PP6E for units 1.4, 2.1, 3.1, 3.3

Course Code: STA203 B
Course Title: Statistical Methods**(Minor Course)**

Unit	Content	No. of hours
I	**Elementary topics on Estimation and Testing of Hypothesis** 1.1 Sample from a distribution : Concept of Population and sample, Concept of Parameter, statistic, estimator and estimate. **1.2** Properties of good estimator (Only names), unbiasedness and standard error of an estimator. **1.3** Central Limit theorem (statement only). **1.4** Sampling distribution of sample mean and sample proportion (For large sample only). **1.5** Standard errors of sample mean and sample proportion. **1.6** Interval estimate of single mean, single proportion, Difference between two population mean and Difference between two population proportions from sample of large size. 1.7 **Testing of Hypothesis:** Concept of hypothesis Simple Hypothesis and composite hypothesis Null and alternate hypothesis, Types of errors, Critical region, Level of significance and Power of test. Concept of p-value, One tail and two tailed tests.	15
II	**Applications of Z, T , Chi-square and F:** 2.1 Large Sample Tests for Mean and Proportions**:** (Development of critical region is not expected.) (i) For testing specified value of population mean (ii) For testing specified value of population proportion (iii) Test for Difference between Two Population Means (iv) Test for Difference between Two Population Proportions. 2.2 Applications of Chi-square tests: goodness of fit, independence of two attributes and variance of normal distribution.Yate's correction for 2x2 contingency table. 2.3 Applications of Student's t-test for one and two population means and for correlation coefficient. 2.4 Z-transformation and its uses. 2.5 F-test for equality of variances.	15

III	**ANOVA and Statistical Quality Control** **ANOVA** 3.1 One-way classification as an extension of t-test and Two Way Classification: Model, Layout and Calculation of various sum of squares, Hypothesis, ANOVA Table. **Statistical Quality Control** : 3.2 The Meaning of Quality and Quality Improvement, Chance and Assignable Causes of Quality Variation, Statistical Basis of the Control Chart, Control Limits, Specification Limits, and Natural Tolerance Limits. The choice between Attributes and Variables Control Charts. Elementary ideas, Assignable and unassignable Causes, Control Charts, Various Patterns and its interpretation **3.3** Control Chart for Variables: X bar and R charts, Statistical Basis of the Charts, Construction of Control Chart, interpretation and Use of X bar and R Charts, when standards are known and unknown./at least one of them is unknown. **3.4** Control Chart for attributes: p-chart, np chart, c-chart (with constant/ variable sample size), Construction of Control Chart and their interpretation.	15

Self-Learning topics (Unit wise)

Unit	Topics
3	Central Limit theorem (statement only).
3	Sampling distribution of sample mean and sample proportion (For large sample only).
3	Standard errors of sample mean and sample proportion.
3	Point and Interval estimate of single mean, single proportion
3	Concept of hypothesis
3	Simple Hypothesis and composite hypothesis Null and alternate hypothesis
3	Types of errors, Critical region, Level of significance.
3	Test for Difference between Two Population Proportions.

Online Resources

'Business Statistics' by Dr Mukesh Kumar Barua from IIT Roorkee available on the Swayam portal, https://nptel.ac.in/courses/110/107/110107114/ for US-FST-202 for unit III.

Course Code: STA201C
Course Title: Statistical Data Analysis using R Programming (Skill Enhancement Course-Paper I)

Unit I	**Introduction to R programming**	**15**
1.1	**Introduction to R programming:** What is R? - Installing R and R Studio – R Studio Overview - Working in the Console - Starting and ending R session,	**05**

	getting help in R , Installing and loading packages. Arithmetic Operators – Logical Operations Using Functions - log 10, log , sort, max, unique, range, length, mean, var, prod, sum, summary, dim, sort, five num etc	
1.2	**Data structures, variables, and data types in R:** Data Types : Vector, list, matrices, array and data frame Variable Type : logical, numeric, integer, complex, character and factor Data Manipulation : Selecting random N rows, removing duplicate row(s), dropping a variable(s),Renaming variable(s), sub setting data, creating a new variable(s), selecting of random fraction of row(s), appending of row(s) and column(s), simulation of variables.	**05**
1.3	Data Processing : Data import and export, setting working directory, checking structure of Data: Str(), Class(), Changing type of variable (for eg as.factor, as.numeric)	**05**
Unit 2	**Data Visualization using GGPLOT2**	**15**
2.1	Introduction to ggplot2 & Grammar of Graphics: Overview of ggplot2, core structure (ggplot() + aes() + geom_*())	1
2.2	Basic Components of a ggplot: Aesthetics, Layers, Geoms, Stats, Facets, Scales	2
2.3	Working with Aesthetics: Mapping vs setting, color, fill, size, shape	1
2.4	Common Chart Types: Histogram, Bar Chart, Boxplot, Violin, Scatter, Line, Jitter, Frequency polygon, Pie (coord_polar)	2
2.5	Titles, Labels, Annotation & Themes: labs(), annotate(), geom_text(), theme_*() and theme() customization	1
2.6	Faceting & Plot Layout: facet_wrap(), facet_grid(), Arranging with patchwork, gridExtra	2
2.7	Scales: setting same scale to different charts, Zoom in option, Zoom out option, Reverse scale on x-axis, y-axis	1
2.8	Interactivity with esquisse: Hands-on with GUI plot builder, exporting code	1
2.9	ggcorrplot, ggpol, ggwordcloud, ggthemes	2
2.10	Additional packages to GGplot2: patchwork- Plots positioned side by side ,directlabels- Labels to points, add2ggplot - For additional themes, ggforce, gghighlight	2

Course Code: STA210B
Course Title: Wanna Learn Python (Open Electives-Paper I)

Unit	Content	No. of Hours
I	**Introduction to Python:** 1.1. Python Setup, Python Arithmetic, Basic Data Types, Variables, Lists, Tuples and Strings, Dictionaries and sets. 1.2. Numpy arrays: Creating arrays crating n-dimensional arrays using np.array and array operations(indexing and slicing, transpose, mathematical operations) 1.3. Pandas data frames: Creating series and data frames and Operations on series and data frames 1.4. Reading and writing data: From and to Excel and CSV files 1.5. Control statements: if, if-else, if-else-if, while loop, for loop 1.6. Defining functions: def statement 1.7. Text data operations: len, upper, lower, slice, replace	15

II	**Data Handling and Visualisation:** 2.1. **Data Manipulation**: Selecting random N rows, removing duplicate row(s), dropping a variable(s), Renaming variable(s), sub-setting data, creating a new variable(s), selecting of random fraction of row(s), appending of row(s) and column(s), simulation of variables. 2.2. **Data Processing**: Data import and export, setting working directory, checking structure of Data, Changing type of variable, Data split into training and Test 2.3. **Data Visualisation**: Simple bar diagram, subdivided bar diagram, multiple bar diagram, pie diagram, Box plot for one and more variables, histogram, frequency polygon, scatter plot, correlation plot Time series, Relationship maps, Heat maps, Geo Maps,3-D Plots, Higher-Dimensional Plots, Word clouds using Matplotlib, Plotly, Seaborn, **Story telling** on Data Sets **Iris /Pigeons/Car Design Data Set** for 2.1 to 2.3	15
III	**Statistical Computing and Modeling** : 3.1 **Creating Dashboard, Comparison of Data Sets and Storytelling with** using Panda Profiling, Sweetviz, Autoviz. **Story telling** on Data Sets **Iris /Pigeons/Car Design Data Sets** 3.2 **Statistical Computing**: Descriptive Statistics: Mean, Median, Mode, Standard Deviation, Variance, Coefficient of Variation, Skewness, Kurtosis, Correlation 3.3 **Linear Regression**: Model, splitting of Data set, Interpretation of output , R^2 ,adjusted R^2, Overall significance of model, Significance of Individual Coefficient, Confidence intervals for the regression coefficients, Validity Tests, Cross Validation with interpretations Story telling on dataset **Auto MPG dataset** **Story telling** on dataset **Titanic dataset/Iris dataset**	15

Self-Learning topics (Unit wise)

Unit	Topics
II	Simple bar diagram, subdivided bar diagram, multiple bar diagram, pie diagram, Box plot for one and more variables, histogram, frequency polygon, scatter plot
III	Descriptive Statistics: Mean, Median, Mode, Standard Deviation, Variance, Coefficient of Variation, Skewness, Kurtosis, Correlation

Online Resources

https://nptel.ac.in/courses/106/107/106107220/ https://nptel.ac.in/courses/106/106/106106212/

Sr No	**List of practical experiments**
Probability Distributions	
1	Bivariate Probability Distributions, Marginal & Conditional distributions, Conditional Mean, Conditional Variance, Correlation.
2	Moment Generating Function
3	Cumulant generating Function, Characteristic function and PGF.
4	Discrete Uniform, Bernoulli, Binomial and Poisson Distribution.
5	Geometric, Negative Binomial and Hypergeometric Distribution.
6	Fitting of Binomial and Poisson Distribution and Negative Binomial Distribution
7	Transformation of discrete & continuous random variables.
8	Practical Using R: Standard Discrete distribution -I
9	Practical Using R: Standard Discrete distribution -II
Theory of Sampling	
1	Designing of Questionnaire
2	Simple Random Sampling for Variables.
3	Simple Random Sampling for Attributes.
4	Estimation of Sample Size in Simple Random Sampling.
5	Stratified Random sampling and Problem based on proportional allocation
6	Stratified Random Sampling and Problem based on optimum allocation
7	Ratio Estimation.
8	Regression Estimation.
9	Practical Using R: Simple Random Sampling and Stratified Random sampling
Statistical Method	
1	Sampling distribution and estimation
2	Testing of Hypothesis
3	Large Sample Tests for Attributes
4	Large Sample Tests for Variables
5	Applications of Chi-square test for Goodness of Fit
6	Applications of Chi-square test for Independence
7	Applications of Student's t-test
8	F- Tests
9	Practical Using Megastat Add on to Excel on abo
10	Anova
11	Control Chart for Variables
12	Control Chart for Attributes
Statistical Data Analysis using R Programming	
1	Vectors, Arithmetic & Logical Operations
2	Working with Data Frames
3	Working with Matrices
4	Data Import, Structure, and Type Conversion
5	Univariate Analysis with ggplot2
6	Bivariate Analysis: Comparing Two Variables
7	Multivariate Analysis: Adding Context
8	Creating Correlation Heatmaps, ggpol and Word Clouds
9	Dashboard-like Plot Arrangement

References:

Course Code: STA201B

Probability Distributions

1. Introduction to the theory of statistics: A. M. Mood, F.A. Graybill, D. C. Boyes, Third Edition; McGraw-Hill Book Company.(1957)
2. Introduction to Mathematical Statistics: R.V. Hogg, A.T. Craig; Fourth Edition; Collier , McMillan Publishers.(2005)
3. Probability and Statistical Inference: R.V. Hogg, E. A. Tannis, Forth Edition; Collier McMillan Publishers.(2001)
4. John E. Freund's Mathematical Statistics: I. Miller, M. Miller; Eighth Edition; Pearson Education Inc.(2014)
5. Introduction to Mathematical Statistics: P.G. Hoel; Fourth Edition; John Wiley & Sons Inc.(2016)
6. Fundamentals of Mathematical Statistics: S.C. Gupta, V.K. Kapoor; Eighth Edition; Sultan Chand &Sons, Tenth edition (2000)
7. An Outline of Statistical Theory Vol. 1: A.M. Goon, M.K. Gupta, B. Das Gupta; Third Edition; The World Press Pvt. Ltd.(2013)

Course Code: STA202B

Theory of Sampling

1. Sampling Techniques: W.G. Cochran; 3rd Edition; Wiley. (1978)
2. Sampling Theory and methods: M.N. Murthy; Statistical Publishing Society. (1967)
3. Sampling Theory: Des Raj; McGraw Hill Series in Probability and Statistics. (1968)
4. Sampling Theory of Surveys with Applications: P.V. Sukhatme and B.V.Sukhatme; 3rd Edition; Iowa State University Press. (1984)
5. Fundamentals of Applied Statistics: S. C. Gupta and V.K. Kapoor; 3rd Edition; Sultan Chand and Sons. (2001)
6. Sampling Techniques: W.G. Cochran; 3rd Edition; Wiley. (1978)
7. Sampling Theory and methods: M.N. Murthy; Statistical Publishing Society. (1967)
8. Sampling Theory: Des Raj; McGraw Hill Series in Probability and Statistics. (1968)
9. Sampling Theory of Surveys with Applications: P.V. Sukhatme and B.V.Sukhatme; 3rd Edition; Iowa State University Press. (1984)
10. Fundamentals of Applied Statistics: S. C. Gupta and V.K. Kapoor; 3rd Edition; Sultan Chand and Sons. (2001)

Course Code: STA203B

Statistical Methods

1. Medhi J.: Statistical Methods, An Introductory Text, Second Edition, New Age International Ltd.
2. Spiegel M.R.: Theory and Problems of Statistics, Schaum's Publications series. Tata McGraw-Hill.
3. Kothari C.R. : Research Methodology, Wiley Eastern Limited.
4. David S.: Elementary Probability, Cambridge University Press.
5. Hoel P.G.: Introduction to Mathematical Statistics, Asia Publishing House.
6. Hogg R.V. and Tannis E.P.: Probability and Statistical Inference. McMillan Publishing Co. Inc.
7. Pitan Jim: Probability, Narosa Publishing House.

Course Code: STA209B

1. R Cookbook, Paul Teetor, Oreilly: R Cookbook [R CKBK] [Paperback] R Cookbook [RKBK] [Paperback].

2. ggplot2: Elegant Graphics for Data Analysis, Hadley Wickham, Danielle Navarro, and Thomas Lin Pedersen, 3rd edition, Springer, https://ggplot2-book.org/.
3. Torsten Hothorn, Brian S. Everitt: A Handbook of Statistical Analyses using R, Chapman and Hall/CRC Press, 3rd Edition
4. Pierre-Andre Cornillon, Arnaud Guyader, Francois Husson, Nicolas Jegou, Julie Josse, Maela Kloareg, Eric Matzner-Lober, Laurent Rouvière: R for Statistics, Chapman and Hall/CRC .
5. Asha Jindal(Ed): Analyzing and Visualizing Data with R Software, Shailja Prakashan and K.C.College, 2018.

Course Code: STA210B

1. Mark Lutz: Programming Python, O'Reilly Media, 4th Edition.
2. Wes McKinney: Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, O'Reilly Media, 2nd Edition.
3. Kenneth A. Lambert: The Fundamentals of Python: First Programs, 2011, Cengage Learning.
4. Asha Jindal(Ed): Analyzing and Visualizing Data using Free Open Source Software: Python Programming with Case Studies, Shailja Prakashan and K. C. College, 2020.

Part 5- The Scheme of Teaching and Examination is as under:Second Year Semester – IV Summary

Sr. No.	Choice Based Credit System			Subject Code	Remarks
1	Core Course (**Statistics**)			STA204B STA204D STA205B STA205D STA206B STA206D	Nil
2	Elective Course	Discipline Specific Elective (DSE) Course			
		2.1	Interdisciplinary Specific Elective (IDSE) Course		
		2.2	Dissertation/Project		
		2.3	Generic Elective (GE) Course	STA212B	
3	Ability Enhancement Courses (AEC)				
	Skill Enhancement Courses (SEC)			STA211B	

Second Year Semester -IV Internal and External Detailed Evaluation Scheme

Sr. No.	Semester	Subject Code	Subject Title	NEP Course Type	Hours Per Week						Seasonal Evaluation Scheme (Internal+ External)			Total Marks
					Units	S. L. E.	L	T	P	Credit	S. L. E	PA/AT	SEE	
1	IV	STA204B	Probability and Sampling Distributions	Major	3	20 %	3	0	0	3	10	5	60	100
		STA204D	**Computer Applications & Practical** Based on STA204B						2	1			25*	
2	IV	STA205B	ANOVA and Design of Experiments	Major	3	20 %	3	0	0	3	10	5	60	100
		STA205D	**Computer Applications & Practical** Based on STA205B						2	1			25	
3	IV	STA206B	**Operation Research Techniques**	Minor	3	20 %	3	0	0	3	10	5	60	100
		STA206D	**Computer Applications & Practical** Based on STA206B						2	1			25	
4	IV	STA203C	**Statistical Data Analysis using SPSS**	VSEC	2		2	0	0	2			50	75

		STA211D	**Computer Applications & Practical** Based on STA203C		1				2	1			25	
5	IV	STA212B	**Introduction to Six Sigma with MS-Excel**	OE	3	20 %	3	0	0	3	10	15	50	75

***One to two lectures to be taken for CONTINUOUS self -learning Evaluation.**

Second Year Semester – IV Units – Topics – Teaching Hours

S. No.	**Subject Code**	**Subject Unit Title**		**Hours**	**Total No. of hours**	**Credit**	**Total Marks**
1	STA204B	I	Standard Continuous Probability Distributions-I	15	45	3	100 (60+40)
		II	Standard Continuous Probability Distributions-II	15			
		III	Exact Sampling Distributions	15			
	STA204D	IV	Practical based on STA204B	30	30	1	
2	STA205B	I	Analysis of Variance	15	45	3	100 (60+40)
		II	Design of Experiments	15			
		III	Latin square design and Factorial Experiments	15			
	STA205D	IV	Practical based on STA205B	30	30	1	
3	STA206B	I	Linear Programming Problem(L.P.P.)	15	45	3	100 (60+40)
		II	Transportation Problem and Assignment problem	15			
		III	Decision theory and Game Theory	15			
	STA206D	IV	Practical based on STA206B	30	30	1	
4	STA203C	I	Managing Data in SPSS	15	30	2	**75** (50+25)
		II	Charts in SPSS, Multi-Dimensional Scaling (MDS) & Correspondence Analysis	15			
	STA211D	III	Practical based on STA203C	30	30	1	
5	STA212B	I	Introduction to Six Sigma	15	45	3	**75** (50+25)
		II	Measure and Analyse	15			
		III	Improve and Control	15			

- **One Credit =15 Classroom hours**

L: Lecture: Tutorials P: Practical Ct-Core Theory, Cp-Core Practical, SLE- Self learning evaluation CT-Commutative Test, SEE- Semester End Examination, PA- Project Assessment, AT- Attendance

Part -6 - Detailed Scheme Theory

Curriculum Topics along with Self-Learning topics - to be covered, through self-learning mode along with the respective Unit. Evaluation of self-learning topics to be undertaken before the concluding lecture instructions of the respective UNIT

Course Code: STA204B

Major-Paper I (Probability and Sampling Distributions)

Unit 1	**Standard Continuous Probability Distributions-I**	15
1.1	**Rectangular or Continuous Uniform distribution over (a, b)** Mean, Median Standard deviation, C.D.F., M.G.F., Mean, variance, μ3 using M.G.F., skewness of distribution For X following U (0,1), distribution of i) X/(1+X) , ii) X/(1-X) **Triangular distribution:** Symmetric and asymmetric over (a, b) with peak at c , M.G.F. Mean ,Variance, c.d.f. Median.	05
1.2	**Exponential Distribution:** Definition, M.G.F.,C.G.F. raw moments and central moments up to order four using M.G.F..and C.G.F. Measures of Skewness and Kurtosis, Nature of Probability curve, Median and Quartiles and Percentiles, Forgetfulness Property with proof and examples based on it. Distribution of X(1) ,first order statistic, Distribution of ratio of two i.i.d. Exponential random variables. Distribution of $1/\lambda \ln(1-x)$ if follows Uniform (0,1). Distribution of X+Y and x/(x+y), for two independent Exponential variables X and Y with mean1. (All with proof.)	
1.3	**Gamma Distribution.(with Scale and shape parameter)** Expression for rth raw moment, Mean, variance, Mode & Standard deviation. M.G.F., Additive property, C.G.F.. raw moments and central moments up to order four using M.G.F.. and C.G.F., Coefficients of skewness and Kurtosis and nature of probability curve, Distribution of sum of independent Exponential random variables.	5 hrs.
Unit 2	**Standard Continuous Probability Distributions-II**	**15**
2.1	**Beta Distribution: Type I & Type II** Expression for rth raw moment, Mean, Mode and Standard deviation, H.M., If a r.v. X follows Beta of type 1, distribution of 1- X, If a r.v. X follows Beta of type 2, distribution of i) 1/(1+X) ii)X/(X+Y) With proof. For two independent Gamma variables X and Y with parameters m and n respectively, distribution of U = X/Y and V=X/(X+Y) with proof.	**05**
2.2	**Normal Distribution:** Definition, Derivation of Mean, Median, Mode, Standard deviation, M.G.F. , C,G,F., Moments & Cumulants (up to fourth order). skewness & kurtosis, Nature of Normal curve, Mean absolute deviation. Properties of Normal Distribution.	**05**
2.3	Expression for even order central moments and to show that odd order central moments are zero. Percentile, Distribution of Standard normal variable, Percentiles. Distribution of linear function of independent Normal variables. Fitting of Normal Distribution. Central Limit theorem for i.i.d. random variables. (without proof). Application of	**05**
		15

Unit 3	Exact Sampling Distributions	
3.1	**Chi-Square Distribution:** Derivation of p.d.f., Concept of degrees of freedom. Mean, Mode &Standard, deviation. M.G.F., C.G.F., Measures of skewness and Kurtosis, Additive property, Distribution of ratio of two independent Chi-square variables, Distribution of if X and Y are two independent Chi-square variables (All with proof). Distribution of the sum of squares of independent Standard Normal variables. Sampling distributions of sample mean and sample variance and their independence for a sample drawn from Normal distribution (with proof).	5 hrs
3.2	**Testing of Hypothesis :** Review of Concept of hypothesis Simple Hypothesis and composite hypothesis Null and alternate hypothesis, Types of errors, Critical region, Level of significance and Power of test. Concept of p-value, One tail and two tailed tests. **Applications of Chi-Square:** Development of decision criterion with test procedures of (i) Test of significance for specified value of variance of a Normal population (ii) Test for goodness of fit, Test Procedure for independence of attributes. (i) r × c contingency table, (ii) 2 × 2 contingency table, Derivation of test statistic, Yates' correction with proof Derivation of Confidence interval for the variance of a Normal population when (i) mean is known, , (ii) mean is unknown. **Student's t-distribution:** Derivation of p.d.f., Mean, Median, Mean Deviation & Standard deviation. M.G.F., C.G.F. , Measures of skewness and Kurtosis and Additive property, Limiting distribution of t with proof.	5 hrs
3.3	**Applications of t:** Development of decision criterion with test procedure of Test of significance for specified value of mean of Normal population. Test procedure of test of significance for difference between means of (i) two independent Normal populations with equal variances (ii) Dependent samples (Paired t test) Derivation of Confidence intervals for (i) Mean of Normal population, (ii) difference between means of two independent Normal populations having the same variance **Snedecor's F-distribution:** Derivation of p.d.f., Expression for r^{th} raw moment, Mean, variance, Mode & Standard deviation ,Distribution of Reciprocal of F variable with proof. **Applications of F:** Test procedure for testing equality of variances of two independent Normal populations i. Mean is known ii. Mean isunknown Derivation of confidence interval for ratio of variances of two independent Normal populations.	5 hours

Self-Learning topics (Unit wise)

Unit	Topics
1.3	Exponential Distribution: Forgetfulness Property with proof and examples based on it.
2.1	Expression for even order central moments and to show that odd order central moments zero for Normal distribution
2.2	**Distribution of Standard Normal Variable**
3.1	Chi-Square Distribution: Derivation of p.d.f. Mean, Mode & Standard deviation
3.5	**Student t-distribution** :Derivation of p.d.f. , Mean, variance, rth order raw moment, Mean Deviation, Measures of skewness and Kurtosis
3.9	F Distribution: Derivation of p.d.f. , Expression for r th raw moment, Mean, Mode & Standard Deviation, Distribution of Reciprocal of F variable with proof.

Online Resources

"Probability and Statistics" by Prof. Somesh Kumar, Department of Mathematics, IIT Kharagpur Source : https://nptel.ac.in/courses/111/105/111105090/ on unit 3.3.

Course Code: STA205B
Course Title: Analysis of Variance and Design of Experiments (Major-Paper II)

Unit1	**Analysis of Variance**	**15 Hours**
1.1	Introduction, Uses, Cochran's Theorem (Statement only). **One-Way ANOVA:** One way classification with equal & unequal observations per class, Mathematical Model, Assumptions Least square estimators of the parameters, Variance of the estimators,	**05**
1.2	**Two- Way ANOVA:** Two-way classification with one and K observations per cell. Mathematical Model, Assumptions, Least square estimators of the parameters, Variance of the estimators,	**05**
1.3	Expectation of various sums of squares, F-test, Analysis of variance table. Estimation of treatment contrasts, Standard Error and Confidence limits for elementary treatment contrasts for one Way and Two way.	05
Unit2	**Design of Experiments**	**15**
2.1	Concepts of Experiments, Experimental unit, Treatment, Yield, Block, Replicate, Experimental Error, Precision. Principles of Design of Experiments: Replication, Randomization & Local Control. Efficiency of design D1 with respect to design D2 .	**05**
2.2	Choice of size, shape of plots & blocks in agricultural & non-agricultural experiments. **Completely Randomized Design (CRD) & Randomized Block Design (RBD):** Mathematical Model, Assumptions, Expectation of various sums of squares, F-test and Analysis of variance table.	**05**
2.3	Least square estimators of the parameters, Variance of the estimators, Estimation of treatment contrasts, Standard error and Confidence limits for elementary treatment contrasts. Efficiency of RBD relative to a CRD.	**05**

Unit3	**Latin square design and Factorial Experiments**	**15**
3.1	**Latin square design:** Mathematical Model, Assumptions, Expectation of various sums of squares, F-test, Analysis of variance table. Least square estimators of the parameters, Variance of the estimators,	**05**
3.2	Estimation of treatment contrasts, Standard error and Confidence limits for elementary treatment contrasts. Efficiency of the design relative to RBD, CRD. Missing plot technique for one missing observation in case of CRD, RBD & LSD.	**05**
3.3	**Factorial Experiments:** Definition, Purpose & Advantages. 2^2, 2^3 Experiments. Calculation of Main & interaction Effects. Definition of contrast and orthogonal contrast, Yates' method. Analysis of 2^2& 2^3 factorial Experiments.	**05**

Self-Learning topics (Unit wise)

Unit	Topics
1	Introduction to Multiple Comparison Tests.
2	Concepts of Experiments, Experimental unit, Treatment, Yield, Block, Replicate, Experimental Error, Precision. Principles of Design of Experiments: Replication, Randomization & Local Control. Efficiency of design D1 with respect to design D2. Choice of size, shape of plots and blocks in agricultural & nonagricultural experiments
3	Basics in Latin Square Design , Factorial Experiments: Definition, Purpose & Advantages. 2^3 Experiments. Calculation of Main & interaction Effects. Definition of contrast and orthogonal contrast, Yates' method. Analysis of 2^3 factorial Experiments.

Online Resources

Analysis of Variance and Design of Experiments, Swayam Prabha Course, MOE, GOI by Prof. Shalabh, IIT Kanpur **http://home.iitk.ac.in/~shalab/spanova.htm?fbclid=IwAR3mmXTpm6P6BSnoaAX25qkyrLx9LGy5SXLj3CodHFYWwHrnL-5IKI5f6SI** for Unit 2 and 3.

Course Code: STA206B

Course Title: **Operation Research Techniques(Minor Course)**

Unit	Content	No. of Lectures
I	**Linear Programming Problem (L.P.P.)** 1. Introduction to OR model, convex set 2. Definition, Mathematical Formulation (Maximization and Minimization), Concepts of Solution, Feasible Solution, Basic Feasible Solution, Optimal solution, Slack, Surplus & Artificial variable, Standard form, Canonical form 3. Graphical Method & Simplex Algorithm to obtain the solution to an L.P.P. Problems involving Unique Solution, Multiple Solution, Unbounded Solution and Infeasible Solution. 4. Big M method. 5. Primal and Dual conversion	15
II	**Transportation Problem and Assignment problems** 1. Definition, Basic concepts of Transportation Problem. 2. Initial Basic Feasible Solution using (i) North-West Corner rule.(ii) Matrix Minima Method.	15

	(iii)Vogel's Approximation Method. 3. Optimum Solution using MODI Method. 4. Problems involving unique solution, multiple solutions, degeneracy, maximization, prohibited route(s) and production costs. Unbalanced Transportation problem. 5. Introduction to Assignment problem, Optimum solution using Hungarian method 6. Special cases like Unbalanced and Maximization type Assignment problem	
III	**Decision Theory and Game Theory:** **Decision Theory:** 1. Introduction and basics of Decision theory, Decision making under uncertainty: Laplace criterion, Maximax (Minimin) criterion, Maximin (Minimax) criterion, Hurwitz criterion, Minimax Regret criterion. 2. Decision making under risk: Expected Monetary Value criterion (EMV), Expected Opportunity Loss (EOL) criterion, EPPI, EVPI. **3.** Decision tree analysis **Game Theory:** **1. Introduction and Basics of game theory,** Definitions of Two persons Zero Sum Game, Saddle Point, Value of the Game, Pure and Mixed strategy, Optimal solution of two person zero sum games. Dominance property, Solution of Mixed Strategy Games 2 X 2. **2.** Graphical solution of (2xn) and (mx2) games.	15

Self-Learning topics (Unit wise)

Unit	Topics
1	Introduction to OR model, convex set, Graphical Method, Big M method, Primal and Dual conversion
II	Definition, Basic concepts of Transportation Problem, Initial Basic Feasible Solution using Matrix Minima Method. Introduction to Assignment problem, Optimum solution using Hungarian method, Special cases like Unbalanced and Maximization type Assignment problem
III	**Introduction and Basics of game theory,** Definitions of Two persons Zero Sum Game, Saddle Point, Value of the Game, Solution of Mixed Strategy Games 2 X 2.

Online Resources

'Operations Research' by PROF.KUSUMDEEP, Department of Mathematics, IIT Roorkee available on the NPTEL portal, https://nptel.ac.in/courses/111/107/111107128/# for US-FAST-401 for unit I, II and III.

Course Code: STA203C
Course Title: **Statistical Data Analysis using SPSS (Skill Enhancement Course-Paper II)**

Unit I	**Managing Data in SPSS**	**15**
1.1	Introduction to SPSS, Understanding Environment, Creating and Editing Data File Data Manipulation: Sorting Data, Merging and Appending Data/files, Reshaping Data, Recording Variables, Sub setting Data,	**05**
1.2	Data Type Conversions, Sampling, Renaming-formatting data, Handling duplicates/Missing values, Computing new variables	**04**
1.3	Split file, Merge file, Selecting cases, OLAP Cubes	**04**
1.4	Data Preprocessing: 1. Outlier handling 2. Missing frequency handling 3. Handling Improper codes 4. Category merging	**02**
Unit II	**Charts in SPSS, Multi-Dimensional Scaling (MDS) & Correspondence Analysis**	**15**
2.1	Introduction, Error bar, Box plot, Clustered Chart, Chart from the output table, Chart Editor	**03**
2.2	Pie Chart ,Area Chart, Scatter and dot charts, 2Dand 3D charts	**02**
2.3	Visualization for Univariate, Bivariate and Multivariate Data, **Story telling** on dataset **Titanic dataset** (http://biostat.mc.vanderbilt.edu/wiki/pub/Main/DataSets/titanic3.csv)	**02**
2.4	**Multi-Dimensional Scaling (MDS) & Correspondence Analysis:** **Multi-Dimensional Scaling (MDS) Analysis** • Objective and Concept of MDS • Comparison with Other Interdependence Techniques • Research Design and Assumptions • Deriving MDS & Assessing Overall Fit • Validating MDS Models **Correspondence Analysis –** • Objective and Concept of Correspondence Analysis (CA) • Research Design and Assumptions • Deriving CA & Assessing Fit • Applications of CA in Market Research and Social Science **Story telling** using **grades-mds.sav** And experiential learning on other datasets	**06**
2.5	Programmability Extension-Using R inside SPSS Using R inside SPSS (Programmability Extension) and syntax option in SPSS	**02**

Course Code: STA212B
Course Title: **Introduction to Six Sigma with MS-Excel (Open Electives-Paper II)**

Course Code	**Title**	**Credits**
	Introduction to Six Sigma with MS-Excel	**3 Credits (45 Hours)**
Unit I	**Introduction to Six Sigma** 1.1 Quality concept and definition, Six – Sigma: Definition of six-sigma, Six Sigma as business strategy, Principles and Meaning of Lean- six – sigma, Introduction to Lean: Definition of Lean , 5 S in Lean, The 5 Principles of Lean,7	**15 Hours**

	Wastes in Lean . Over view of DMAIC. 1.2 Define Phase: Introduction, Method and its importance, Definitions of VOC,VOP,VOB,CTQ,COPQ, 1.3 Process map, Problem statement , Goal statement, SIPOC, process components in SIPOC, Mapping the process, Steps to draw SIPOC 1.4 Project charter , Brain storming, Negative Brain storming	
Unit II	**Measure and Analyse** 2.1 Measure phase: Measurement System Fundamentals, Accuracy , Precision, Repeatability, Reproducibility, Overview of Sampling Techniques 2.2 Measurement System Analysis: Gage R & R ,Process performance. DPMO, DPU, Yield 2.3 Process Capability Analysis: Introduction to process capability, concept, Specification limits natural tolerance limits and their comparisons, estimate of percent defectives 2.4 Capability ratio and Capability indices (Cp), Capability performance indices Cpk with respect to machine and process interpretation, relationship between (i) Cp and Cpk (ii)Defective parts per million and Cp.(iii)Pp and Ppk , comparison with Cp , Cpk 2.5 Analyse Phase : Basic statistics : Types of Data , Descriptive statistics correlation and simple linear regression, Residuals, R^2, adjusted R^2 and its interpretation , probability distributions. 2.6 Process Door and Data Door, Root cause analysis, 2.7 Graphs: Histogram, Box – plot, scatter plot , Pareto chart, Run chart,	**15 Hours**
Unit III	**Improve and Control** 3.1 Improve Phase : Multi Voting , Delphi Technique , Nominal Group Technique , Kaizen, SCAMPER. 3.2 Poka Yoke , Introduction, Chance and assignable causes, Statistical basis of the control chart: Basic principles of control chart (Shewhart control charts), Concept of Rational subgroups, Analysis of patterns on control charts, Sensitizing rules for control chart. 3.3 Control chart for variables: X-Bar, R, and IMR :Construction, charts based on standard values, Interpretation, Applications of variable control charts, *problems involving setting up standards for future use is also expected.* 3.4 Introduction to the concept of attribute, Defect and Defective. P, np, c and U charts:. Construction, charts based on standard values, Interpretation, Applications of variable control charts. *problems involving setting up standards for future use is also expected.* 3.5 Six Sigma Case Study	**15 Hours** **(13+02)**

Self-Learning topics (Unit wise)

Unit	Topics
I	The 5 Principles of Lean, Negative brain storming
II	Process Capability Analysis, correlation and simple linear regression, Process Door and Data Door, : Histogram, SCAMPER
III	Statistical Process Control : Introduction

Online Resources

The 5 Principles of Lean ⟶ https://kissflow.com/project/agile/5-principles-of-lean/
Negative Brain storming ⟶ https://www.slideshare.net/guestf761cb/reverse-brainstorming-2154911
Process Capability Analysis ⟶ https://sixsigmastudyguide.com/process-capability-pp-ppk-cp-cpk/
Correlation and simple linear ⟶ https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression Regression
Process Door and Data Door ⟶ https://www.benchmarksixsigma.com/forum/topic/36217-process-door-vs-data-door/
Histogram ⟶ **https://asq.org/quality-resources/histogram**
SCAMPER ⟶ https://youtu.be/G8w0rJhztJ4 https://youtu.be/vqnIEtlp9d8
"Quality Control and Improvement With Minitab: by Prof. Indrajit Mukherjee, Management, IIT Bombay, Source:**https://nptel.ac.in/courses/110/101/110101150/** for unit 2(PCA) and unit 3

Sr No	**List of practical experiments**
	Probability & Sampling Distribution
1	Continuous Uniform distribution
2	Triangular distribution
3	Exponential distribution
4	Normal Distribution.
5	Central Limit Theorem.
6	Chi Square distribution.
7	t distribution.
8	F distribution.
9	**Practical Based On R :** Exponential and Normal distribution
10	**Practical Based On R:** Chi square test, t-test and F test
	Analysis of Variance & Design of Experiment
1	Analysis of Variance- One Way
2	Analysis of Variance- Two Way.
3	Critical Difference
4	Completely Randomized Design.
5	Randomized Block Design.

6	Latin Square Design.
7	Missing Observations in CRD, RBD & LSD.
8	Factorial Experiments.
9	Questionnaire Designing for Factorial Experiment
10	**Practical Using SPSS:** One way ANOVA , Two Way ANOVA and Factorial Experiments
Operation Research Techniques	
1	Formulation and Graphical
2	Simplex Method
3	Big-M Method
4	Transportation
5	Assignment
6	Practical based on 2, 3 and 4 using LPP Solver
7	Decision Theory-1
8	Decision Theory-2
9	Game Theory
Statistical Data Analysis using SPSS Statistical Data Analysis using SPSS	
1	Introduction to SPSS Interface & Basic Data Entry
2	Data Manipulation – Sorting, Merging, and Appending Files
3	Data Cleaning, Data Preprocessing and Transformation
4	Sampling and Subsetting Data
5	OLAP Cubes and Summary Statistics
6	Charting – Boxplots, Error Bars, and Clustered Charts
7	Advanced Charts – Scatter, Pie, Area & 3D Charts
8	Multi-Dimensional Scaling (MDS)
9	Correspondence Analysis
10	R Integration and Syntax Options in SPSS

References:

Course Code: STA204B

1. Introduction to the theory of statistics: A M Mood, F.A. Graybill, D C Boyes; Third Edition; McGraw-Hill Book Company.(1973)
2. Introduction to Mathematical Statistics: R.V.Hogg, A.T. Craig; Sixth Edition; Collier McMillan Publishers.(2005)
3. Probability and Statistical Inference: R.V.Hogg, E. A.Tannis, Third Edition; Collier McMillan Publishers.(2001)
4. John E. Freund's Mathematical Statistics: I. Miller, M. Miller; Eighth Edition; Pearson Education Inc.(2014)
5. Introduction to Mathematical Statistics: P.G. Hoel; Fourth Edition; John Wiley & Sons Inc.
6. Fundamentals of Mathematical Statistics: S.C. Gupta, V.K. Kapoor; Tenth Edition; Sultan Chand &Sons.(2000)
7. An Outline of Statistical Theory Vol. 1: A.M. Goon, M.K. Gupta, B. Das Gupta; eighth Edition; The World Press Pvt. Ltd.(2002)

Course Code: STA205B

1. Experimental Designs: W.G. Cochran and G.M. Cox; Second Edition; John Wiley and Sons (1957).
2. The Design and Analysis of Experiments: Oscar Kempthorne, John Wiley and Sons, Second Edition (2007)
3. Design and Analysis of Experiments: Douglas C Montgomery; 6thEdition;John Wiley & Sons (2008).

4. Design and Analysis of Experiments: M.N. Das and N.C. Giri, 2nd Edition; New Age International (P) Limited;(1986).
5. Fundamentals of Applied Statistics: S.C. Gupta and V.K. Kapoor; 3rd Edition; Sultan Chand and Sons (2001).

Course Code: STA206B

1. Vora N. D. : Quantitative Techniques in Management, Third edition, GcGraw Hill Companies.
2. Kantiswarup, P. K. Gupta, Manmohan : Operations Research, Twelfth edition, Sultan Chand & sons.
3. Sharma S. D. : Operations Research, Eighth edition, Kedarnath Ramnath& Co.
4. TahaHamdyA. : Operations Research : Eighth edition, Prentice Hall of India Pvt. Ltd.
5. Vora N. D. ; Quantitative Techniques in Management, Third edition, McGraw Hill Companies.

Course Code: STA211B

1. SPSS for Windows Step by Step A Simple Guide and Reference by Darren George and Paul Mallery, Pearson
2. Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). SAGE Publications.
3. Brian C. Cronk, How to Use SPSS:A Step-By-Step Guide to Analysis and Interpretation Paperback
4. Naresh K. Malhotra(2021), Marketing Research: An Applied Orientation, Pearson, 7th edition.
5. Applied Multivariate Statistical Analysis, Richard A. Johnson and Dean W. Wichern , sixth edition, Pearson Prentice Hall

Course Code: STA212B

Main References:

1. The Six Sigma Way: How GE, Motorola, And Other Top Companies Are Honing Their Performance, (2000), Peter S. Pande, Robert P. Neuman, Roland R. Cavanagh, McGraw-Hill
2. Six Sigma For Beginners: Pocket Book (2018), Rajiv Tiwari Kindle Edition
3. Six Sigma for Organizational Excellence: A Statistical Approach (2015), K. Muralidharan, springer

Additional Resources

1. Fundamental of Mathematical Statistics, Gupta and Kapoor.
2. Probability and Random process by T. Veerarajan.
3. Six Sigma For Business Excellence, (2005), Penelope Przekop, McGraw-Hill Six Sigma Handbook, by Pyzdek, McGraw Hill Education; 4 edition (1 July 2017).
4. The Certified Six Sigma Green Belt Handbook, Roderick A. Munro and Govindarajan Ramu , American Society for Quality (ASQ),
5. What Is Design For Six Sigma,(2005), Roland Cavanagh, Robert Neuman, Peter Pande, Tata McGraw-Hill
6. What Is Lean Six Sigma,(2004), Mike George, Dave Rowlands, Bill Kastle, McGraw-Hill
7. Six Sigma Deployment,(2003), Cary W. Adams, Charles E Wilson Jrs, Praveen Gupta, Elsevier Science.
8. Introduction to Statistical Quality Control(2009), Montgomery, Douglas, C ,Sixth Edition, John Wiley & Sons.Inc.:.
9. Statistical Quality Control: E.L.Grant. 2nd edition, McGraw Hill, 1988.
10. Quality Control and Industrial Statistics: Duncan. 3rd edition, D.Taraporewala sons & company

The Scheme of Teaching and Examination:
The performance of the learners shall be evaluated in two components: Internal Assessment with 40% marks by way of continuous evaluation and by Semester End Examination with 60% marks by conducting the theory examination.

Examination Pattern for First Year Degree as per NEP 2020 Academic Year 2023-2024

1) Evaluation of Major and Minor Subjects

Subject	Formative Assessment (Marks)	Summative Assessment (Marks)
Major Subject	40	60
Minor Subject	40	60
Major (Practical based Subject)	-	25
Minor (Practical based Subject)	-	25
General Elective (GE/OE)	20	30

.

FORMATIVE ASSESSMENT:- It is defined as the assessment of the learners on the basis of continuous evaluation as envisaged in the credit based system by way of participation of learners in various academic and correlated activities in the given semester of the programme.

A). Formative Assessment – 40
B). Practical's (internal Components of the Practical Course) 40 marks

1. For Theory Courses

Sr.No.	Particulars	Marks
1	Assessment of Practical exercise / case study / presentation, assignment presentation / MCQ designed with experiential learning and hands on learning experience **OR**	20 Marks
2	Self-Learning Evaluation	10 Marks
3	Active participation in routine class instructional deliveries	5 Marks

2. For Courses with Practical
Each practical course can be conducted out of 50 marks with 10 marks for internal **component of the Practical** and 40 marks for formative assessment which will be converted to 25 marks.

Practical's (Internal component of the Practical Course)

Sr. No	Evaluation type	Marks
1	Journal	5
2	Viva	5

C). SUMMATIVE ASSESSMENT =SEMESTER END EXAMINATION :-
It is defined as the examination of the learners on the basis of performance in the semester end theory / written examinations.

The semester end examination (external component) of 60 % for each course will be as follows:

a. For Theory Courses

i) **Duration – 2 Hours**

Theory Question Paper pattern for Main Papers of 60 Marks:

Sr. No.	All questions are Compulsory.	Marks
Q. 1	Attempt either a & b or p & q based on unit 1.	20
Q. 2	Attempt either a & b or p & q based on unit 2.	20
Q. 3	Attempt either a & b or p & q based on unit 3.	20

Theory Question Paper pattern for GE/ OE Papers of 50 Marks:

Sr. No.	All questions are Compulsory.	Marks
Q. 1	a) Attempt any 10 out of 15 based on MCQ based on all units. b) Attempt any 5 out of 7 based on True/False with Justification) based on all units.	20 (10+10)
Q. 2	Attempt any two out of four based on unit 1.	10
Q. 3	Attempt any two out of four based on unit 2.	10
Q. 4	Attempt any two out of four based on unit 3.	10

Theory Question Paper pattern for SEC Papers of 50 Marks:

Sr. No.	All questions are Compulsory.	Marks
Q. 1	a) Attempt any 10 out of 15 based on MCQ based on all units. b) Attempt any 5 out of 7 based on True/False with Justification) based on all units.	20
Q. 2	Attempt any three out of five based on unit 1.	15
Q. 3	Attempt any three out of five based on unit 2.	15

b. For Practical Courses
Duration – 2 Hours

Practical Question Paper Pattern:-

Sr. No.	All questions are Compulsory.	Marks
Q. 1	Attempt any two out of three based on unit 1.	10
Q. 2	Attempt any two out of three based on unit 2.	10
Q. 3	Attempt any two out of three based on unit 3.	10
Q. 4	Attempt any two out of three based on all three units using Statistical Software.	10

For SEC Practical Practical Examination Paper Pattern:

Q. No.	**Question Type**	**Marks**
Q1	Any 2 out of 3 based on Unit 1 (Each question carries 10 Marks)	20
Q2	Any 2 out of 3 based on Unit 2 (Each question carries 10 Marks)	20