**P. V. K. N. GOVT. COLLEGE (AUTONOMOUS): CHITTOOR**

**Accredited NAAC by ‘A’- Grade**

**DEPARTMENT OF STATISTICS (**w.e.f. AY 2024-25**)**

**Programme: B.Sc., Honors in Applied Computer Statistics (Major)**

**COURSE STRUCTURE**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Year**  | **Semester**  | **Course**  | **Title of the Course**  | **No. of Hrs /Week**  | **No. of Credits**  |
| I  | I  | 1  | Essentials and Applications of Mathematical, Physical and Chemical Sciences  | 3+2  | 4  |
| 2 | Advances in Mathematical, Physical and Chemical Sciences | 3+2  | 4  |
| II | 3  | Descriptive Statistics, and Mathematical Expectations | 3  | 3  |
| Descriptive Statistics, and Mathematical Expectations Practical Course | 2 | 1 |
| 4 | Fundamentals of Computer Science and Information Technology | 3 | 3  |
| Fundamentals of Computer Science and Information Technology Practical course | 2 | 1 |
|  | III  | 5  | Theoretical Discrete and Continuous Distributions  | 3  | 3  |
| Theoretical Discrete and Continuous Distributions Practical Course | 2 | 1 |
| 6 | Statistical Methods and Inferential Statistics | 3 | 3  |
| Statistical Methods and Inferential Statistics Practical Course  | 2 | 1 |
| 7 | Statistical Analysis and Reporting Using MS Word, MS Excel, and PowerPoint | 3 | 3  |
| Statistical Analysis and Reporting Using MS Word, MS Excel, and Power Point Practical Course | 2 | 1 |
| 8 | Statistical Programming with C - Language and Data Structures | 3 | 3  |
| Statistical Programming with C - Language and Data Structures Practical Course | 2 | 1 |
| IV  | 9  | Sampling Techniques  | 3  | 3  |
| Sampling Techniques Practical Course | 2 | 1 |
| 10 | Design and Analysis of Experiments | 3 | 3  |
| Design and Analysis of Experiments Practical Course | 2 | 1 |
| 11 | Foundations of Statistical Data Analysis Using ‘R’ | 3 | 3  |
| Foundations of Statistical Data Analysis Using ‘R’ Practical Course | 2 | 1 |
| III  | V  | 12  | Operations Research – I  | 3  | 3  |
| Operations Research – I Practical Course  | 2 | 1 |
| 13 | Statistical Quality Control | 3  | 3 |
| Statistical Quality Control Practical Course | 2 | 1 |
|  | 14A | Applied Statistics  | 3  | 3  |
| Applied Statistics Practical Course  | 2  | 1  |
| **OR** |
|  |  | 14 B  | **Introduction to Core Java and Its Applications in Statistics** | 3  | 3  |
| **Introduction to Core Java and Its Applications in Statistics** Practical Course  | 2  | 1  |
| 15 A  | Operations Research – II  | 3  | 3  |
| Operations Research – II Practical Course  | 2  | 1  |
| **OR** |
|  |  | 15 B  | **Big Data Analysis through Apache Hadoop** | 3  | 3  |
| **Big Data Analysis through Apache Hadoop** Practical Course | 2  | 1  |
|  | VI | **INTERNSHIP** |

**P. V. K. N. GOVT. COLLEGE (AUTONOMOUS): CHITTOOR**

**Accredited NAAC by ‘A’- Grade**

**DEPARTMENT OF STATISTICS**

**BLUE PRINT FOR THE QUESTION PAPER SETTING UNIT WISE**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Chapter Name** | **Very Short Questions2 Marks** | **Short Questions4 Marks** | **Essay Question10 Marks** | **Marks allotted to the Chapter** |
| UNIT - I | 1 | 1 | 2 | 16 |
| UNIT - II | 1 | 2 | 1 | 20 |
| UNIT - III | 1 | 2 | 1 | 20 |
| UNIT – IV | 1 | 2 | 1 | 20 |
| UNIT – V | 1 | 1 | 1 | 16 |
| **Total No. of Questions** | **5** | **8** | **6** | **102** |

**Blue Print for Question Paper for all SEM-END examinations**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sl.****No.** | **Type of Questions** | **No. of Questions given** | **No. of Questions to be answered** | **Marks Allotted to each Question** | **Total Mark s** |
| 1 | Section – A Short Answer Questions | 5 (Knowledge/application, Analysis, comprehension and Synthesis, Understanding and Evaluation) | 5 | 2 | 10 |
| 2 | Section – B Short Answer Questions | 8 (Knowledge/application, Analysis, comprehension and Synthesis, Understanding and Evaluation) | 5 | 5 | 25 |
| 3 | Section –C Essay Questions | 6 (Knowledge/application, Analysis, comprehension and Synthesis, Understanding and Evaluation) | 4 | 10 | 40 |
|  |  |  | **Total** |  | **75** |

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –II (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 3: DESCRIPTIVE STATISTICS AND MATHEMATICAL EXPECTATIONS**

**No. of** Hours/week: 03 **Course Code: 24 – STA – 2C3 C**redits **–** 3

**Learning Objectives:**

1. To understand and apply measures of central tendency, dispersion, moments, skewness, and kurtosis.
2. To learn the fundamental concepts of probability, including various probability theorems and their applications.
3. To develop the ability to analyze random variables, including their types, probability distributions, and functions.
4. To gain proficiency in calculating mathematical expectations, covariance, and understanding their properties.
5. To comprehend generating functions and inequalities in probability theory, including their properties and applications.

**Learning Outcomes:**

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

1. Calculate and interpret measures of central tendency, dispersion, and moments for different data sets.
2. Apply probability theorems such as the addition and multiplication theorems, and understand their real-life applications.
3. Understand and apply the concepts of random variables, including the ability to calculate and interpret probability distributions.
4. Solve problems related to mathematical expectation, covariance, and understand their practical significance.
5. Utilize generating functions and inequalities in probability theory to solve complex problems and analyze data.

This table connects **Course Outcomes (COs)** with **Program Outcomes (POs)** and **Program Specific Outcomes (PSOs)**, outlining how each course outcome contributes to achieving broader educational objectives.

| **Course Outcomes (COs)** | **Program Outcomes (POs)** | **Program Specific Outcomes (PSOs)** |
| --- | --- | --- |
| **CO1**: Apply measures of central tendency (mean, median, mode) and dispersion (range, standard deviation) to summarize and interpret data. | **PO1**: Apply knowledge of mathematics to solve complex problems in the field of statistics and probability. | **PSO1**: Demonstrate the ability to apply advanced statistical methods and techniques to analyze and interpret large datasets. |
| **CO2**: Use graphical methods to derive and interpret median and mode from data. | **PO2**: Develop critical thinking and analytical skills to evaluate data and make informed decisions based on statistical methods. | **PSO2**: Utilize probability models and random variables to make data-driven predictions and decisions in real-world applications. |
| **CO3**: Understand the concept of moments, including central and non-central moments, and apply Sheppard's correction. | **PO3**: Demonstrate an understanding of probabilistic models and random variables in both theoretical and practical contexts. | **PSO3**: Develop problem-solving skills using generating functions, inequalities, and other advanced statistical tools to address practical challenges in fields such as finance, healthcare, and engineering. |
| **CO4**: Understand and apply the concept of probability, including conditional probability, and various probability theorems (addition, multiplication, Bayes’ Theorem). | **PO4**: Gain proficiency in using statistical software (e.g., Excel, R, or Python) to perform calculations and statistical analysis. |  |
| **CO5**: Develop proficiency in calculating and analyzing random variables, their distributions, and their functions, including the use of joint and marginal distributions. | **PO5**: Work effectively both individually and as part of a team to apply statistical concepts to real-world problems. |  |
|  | **PO6**: Understand the role of mathematical expectations, variance, and covariance in interpreting and solving statistical problems. |  |
|  | **PO7**: Effectively use generating functions, inequalities, and probability distributions to model complex systems. |  |
|  | **PO8**: Apply probability theory to various fields, including economics, engineering, and healthcare, to analyze risk and uncertainty. |  |

**Mapping Explanation:**

* **CO1** and **CO2** contribute to **PO1**, as students learn how to apply basic statistical measures and graphical techniques to solve problems in data analysis.
* **CO3** and **CO4** contribute to **PO3**, equipping students with an understanding of advanced statistical concepts such as moments, Sheppard's correction, and probability theory.
* **CO5** aligns with **PO4** and **PO6**, ensuring that students gain the necessary proficiency in random variable analysis and statistical software tools.
* **COs** are also aligned with **PSOs**, particularly **PSO1** and **PSO2**, by enabling students to apply advanced methods to interpret data, model real-world phenomena, and solve problems using statistical and probabilistic models.

**Mapping with Bloom's Taxonomy:**

| **Unit** | **Bloom's Taxonomy Level** |
| --- | --- |
| **UNIT – I: Descriptive Statistics** | Understanding, Applying, Analyzing |
| **UNIT – II: Probability Concepts** | Remembering, Understanding, Applying |
| **UNIT – III: Random Variables** | Remembering, Understanding, Applying |
| **UNIT – IV: Mathematical Expectation** | Understanding, Analyzing, Applying |
| **UNIT – V: Generating Functions and Inequalities** | Remembering, Understanding, Applying |

This mapping reflects the cognitive skills students will develop, progressing from remembering and understanding basic concepts to applying and analyzing more advanced topics.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –II (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 3: DESCRIPTIVE STATISTICS AND MATHEMATICAL EXPECTATIONS**

**No. of** Hours/week: 03 **Course Code: 24 – STA – 2C3 C**redits **–** 3

**II. Syllabus**

**UNIT – I**

**Descriptive Statistics:**

**Measures of Central Tendency**: Mean, Median, Mode, Geometric Mean and Harmonic Mean, Graphical Derivation of Median and Mode. **Measures of Dispersion:** Range Quartile Deviation, Mean Deviation, Standard Deviation and Variance. **Moments:** Central and Non-Central Moments, Interrelationships Between Moments, Sheppard's Correction for Moments. **Skewness and Kurtosis:** Definitions and Measures, Interpretation and Applications

**UNIT – II**

**Probability Concepts:**

**Basic Concepts of Probability:** Random Experiments, Trials, and Outcomes, Sample Spaces and Events, Mutually Exclusive and Exhaustive Events, Equally Likely and Favorable Outcomes. **Definitions of Probability:** Mathematical, Statistical, and Axiomatic Definitions. **Conditional Probability:** Concept and Calculation, Independence of Events. **Theorems in Probability:** Addition Theorem for Two and Multiple Events, Multiplication Theorem for Two and Multiple Events, **Advanced Topics:** Boole's Inequality, Bayes' Theorem and Its Applications in Real-Life Scenarios

**UNIT – III**

**Random Variables**

**Introduction to Random Variables:** Definition of Random Variables, Discrete and Continuous Random Variables, Functions of Random Variables ,**Probability Distributions:** Probability Mass Functions (PMF), Probability Density Functions (PDF), Distribution Functions and Their Properties, **Practical Applications**

**UNIT – IV**

**Mathematical Expectation**

**Bi-variate Random Variables:** Joint, Marginal, and Conditional Distributions, Independence of Random Variables, **Practical Applications, Mathematical Expectation of a Random Variable:** Definition and Concept of Expectation, Properties of Expectation. **Moments and Covariance:** Definition of Moments (Central and Non-Central), Covariance and Its Interpretation. **Theorems of Expectation:** Addition Theorem of Expectation, Multiplication Theorem of Expectation.

**UNIT – V**

**Generating Functions and Inequalities**

**Moment-Generating Function (MGF):** Definition and Properties of MGF. **Cumulant-Generating Function (CGF):** Definition and Properties of CGF. **Probability-Generating Function (PGF):** Definition and Properties of PGF. **Characteristic Function (CF):** Definition and Properties of CF, **Inequalities in Probability Theory:** Chebyshev's Inequality, Cauchy-Schwarz Inequality

**Textbooks**

1. V.K. Kapoor and S.C. Gupta: *Fundamentals of Mathematical Statistics*, Sultan Chand & Sons, New Delhi.
2. BA/BSc I Year Statistics: Descriptive Statistics, Probability Distribution – Telugu Academy, Dr. M. Jaganmohan Rao, Dr. N. Srinivasa Rao, Dr. P. Tirupathi Rao, Smt. D. Vijayalakshmi.
3. K.V.S. Sarma: *Statistics Made Simple: Do It Yourself on PC*, PHI.

**Reference Books**

1. William Feller: *Introduction to Probability Theory and Its Applications*, Volume I, Wiley.
2. Goon, A.M., Gupta, M.K., Das Gupta, B.: *Fundamentals of Statistics*, Vol. I, The World Press Pvt. Ltd., Kolkata.
3. Hoel, P.G.: *Introduction to Mathematical Statistics*, Asia Publishing House.
4. M. Jaganmohan Rao and Papa Rao: *A Textbook of Statistics*, Paper I.
5. Sanjay Arora and Bansi Lal: *New Mathematical Statistics*, Satya Prakashan, New Delhi.

**Practicals - Paper III Credits: 2**

1. Creation of Subdivided and Percentage Bar Diagrams
2. Construction of Pie or Circular Diagrams (for two graphs)
3. Development of Histograms and Frequency Polygons
4. Construction of histograms and frequency polygons to represent the distribution
5. Construction of Ogive Curves
6. Calculation of Measures of Central Tendency for Grouped Data
7. Calculation of Geometric Mean and Harmonic Mean for Grouped Data
8. Computation of Quartile Deviation and Range for Grouped Data
9. Calculation of Mean Deviation, Standard Deviation, and Coefficient of Variation for Grouped Data
10. Consistency Determination for Grouped Data
11. Calculation of Karl Pearson’s and Bowley’s Coefficients of Skewness
12. Computation of Non-central and Central Moments, β₁, β₂, γ₁, γ₂ for Grouped Data
13. Calculation of Non-central and Central Moments with Sheppard’s Corrections for Grouped Data

**Note:** The training will involve using Excel to implement formulas and derive results. The output will then be exported to MS Word for writing inferences and analysis.

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER: II (CBCS, w.e.f 2024-25)**

**COURSE TITLE - 3: DESCRIPTIVE STATISTICS AND MATHEMATICAL EXPECTATIONS**

**Time**: 03Hrs **Course Code: 24 – STA – 2C3 Max. Marks: 75**

**SECTION-A**

Answer any **FIVE** questions. Each question carries **TWO** marks **5 X 2 = 10M**

1. Define Measures of Central Tendency.
2. Define Random experiment and Sample Space.
3. Define Random variable.
4. Define Mathematical Expectation.
5. Define Probability Generating function.

**SECTION – B**

Answer any **FIVE** questions and each question carries **FIVE** marks **5 X 5 = 25M**

1. Define and explain the concept of Kurtosis.
2. Calculate Mean and Median for the following data.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class | 0 – 20 | 20 – 40  | 40 – 60  | 60 – 80  | 80 - 100 |
| fi | 15 | 23 | 50 | 25 | 10 |

1. Explain the three definitions of Probability.
2. State and prove Baye’s theorem of Probability.
3. Explain different types Random variables.
4. A Random variable ‘X’ has the following probability distribution then find

i) Constant ‘K’ ii) P(X<6) ii) P(0<X<5) iii) P(X<4)

1. The Probability distribution of a Random variable is given below, and then find its Mean and Variance.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| X | -2 | -1 | 0 | 1 | 2 | 3 |
| P(X) | 0.1 | K | 0.2 | 2K | 0.3 | K |

1. What are the properties of Moment Generating function?

**SECTION – C**

Answer any **FOUR** questions and each question carries **TEN** marks**. 4 X 10 = 40M**

1. What do you understand by measures of Dispersion and explain various measures of Dispersion?
2. Define and explain the concept of skewness and measure of Skewness.
3. State and prove Boole’s inequality in Probability.
4. Define and explain Distribution function and its properties.
5. State and prove Chebyshev’s Inequality.
6. Define Characteristic function and state its properties.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –II (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE – 4**: **FUNDAMENTALS of COMPUTER SCIENCE AND INFORMATION TECHNOLOGY**

**No. of Hours/week: 03 Course Code: 24 – STA – 2C4 C**redits **–** 3

**Learning Objectives (LOs):**

1. **Understand the foundational concepts of computer science**, including computer systems, hardware, software, and their evolution.
2. **Develop proficiency in using various computer hardware components**, such as input/output devices, storage devices, and CPUs, and understand their functionality.
3. **Gain knowledge of operating systems**, their types, functions, and popular operating systems like Windows, macOS, and Linux.
4. **Learn the fundamentals of office software** such as word processing, spreadsheets, and presentation tools, and their practical applications.
5. **Understand internet technologies and networking**, including internet protocols, web technologies, security measures, and online communication tools.

**Learning Outcomes (LOs):**

Upon completion of the course, students should be able to:

1. **LO1:** Describe the components of computer systems, their evolution, and their classification, including different generations of computers and software types.
2. **LO2:** Identify and explain the various input/output devices, storage devices, and peripherals, and understand their function in a computer system.
3. **LO3:** Compare and contrast different operating systems and their functionalities, including file management and system utilities.
4. **LO4:** Apply word processing, spreadsheet, and presentation software tools to create and manage documents, presentations, and data analysis tasks.
5. **LO5:** Understand and apply networking concepts, web technologies, and internet security, and use internet tools for communication, collaboration, and browsing.

**Course Outcomes (COs):**

1. **CO1:** Understand and explain the key concepts of computers, hardware components, and software types.
2. **CO2:** Identify, compare, and utilize various computer hardware and peripheral devices in practical scenarios.
3. **CO3:** Use operating systems effectively, manage files, and understand their structure and security mechanisms.
4. **CO4:** Proficiently use word processing, spreadsheet, and presentation software for document creation, data manipulation, and presentation.
5. **CO5:** Demonstrate knowledge of internet technologies, networking, and security, and apply them to real-world online communication and web development.

**Program Outcomes (POs):**

1. **PO1:** Demonstrate a foundational understanding of computer systems, their components, and functionalities.
2. **PO2:** Apply technical skills in using computer hardware and software for various tasks in computing.
3. **PO3:** Utilize operating systems efficiently for managing files, processes, and resources.
4. **PO4:** Apply office productivity software to manage data, create documents, and deliver presentations.
5. **PO5:** Understand and apply internet and networking concepts for effective online communication and web browsing.
6. **PO6:** Demonstrate proficiency in using common communication tools (email, video conferencing) for effective collaboration.
7. **PO7:** Understand internet security practices and use them to ensure safe and secure online activity.
8. **PO8:** Apply networking concepts and protocols to set up and troubleshoot basic network configurations.

**Program Specific Outcomes (PSOs):**

1. **PSO1:** Demonstrate the ability to work with different computer hardware and software, including operating systems, to solve problems effectively.
2. **PSO2:** Apply office software (word processors, spreadsheets, presentations) to solve practical business and educational tasks.
3. **PSO3:** Demonstrate an understanding of internet technologies and security measures to create secure and functional web applications and online communications.

**Mapping Summary Table:**

| **Learning Outcomes (LOs)** | **Program Outcomes (POs)** | **Program Specific Outcomes (PSOs)** | **Course Outcomes (COs)** | **Bloom's Taxonomy Level** |
| --- | --- | --- | --- | --- |
| **LO1:** Describe the components of computer systems, their evolution, and their classification. | PO1, PO2 | PSO1 | CO1 | Knowledge (Remembering) |
| **LO2:** Identify and explain the various input/output devices, storage devices, and peripherals. | PO2, PO3 | PSO1, PSO2 | CO2 | Comprehension (Understanding) |
| **LO3:** Compare and contrast different operating systems and their functionalities. | PO3, PO4 | PSO1 | CO3 | Analysis (Analyzing) |
| **LO4:** Apply word processing, spreadsheet, and presentation software tools to create and manage documents. | PO4, PO5 | PSO2 | CO4 | Application (Applying) |
| **LO5:** Understand and apply networking concepts, web technologies, and internet security. | PO5, PO7 | PSO3 | CO5 | Synthesis (Creating) |

**Bloom's Taxonomy Mapping Summary:**

1. **Remembering (Knowledge)**: Students will recall key components of computer systems, hardware types, and software classifications. This is reflected in **CO1** and **LO1**.
2. **Understanding (Comprehension)**: Students will interpret the functions and use of different input/output devices, storage devices, and peripherals, aligning with **CO2** and **LO2**.
3. **Analyzing (Analysis)**: Students will compare and contrast various operating systems, identify their features, and analyze their structures. This corresponds to **CO3** and **LO3**.
4. **Applying (Application)**: Students will apply their knowledge of word processing, spreadsheets, and presentations for practical tasks like document creation and data management. This is represented by **CO4** and **LO4**.
5. **Creating (Synthesis)**: Students will be expected to understand internet technologies, web design, and networking protocols to build secure and functional applications. This aligns with **CO5** and **LO5**.

**Mapping of COs to POs and PSOs**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **COs / POs & PSOs** | **PO1** | **PO2** | **PO3** | **PO4** | **PO5** | **PO6** | **PO7** | **PO8** | **PSO1** | **PSO2** | **PSO3** |
| **CO1** | 3 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 |
| **CO2** | 3 | 2 | 1 | 2 | 3 | 2 | 1 | 1 | 2 | 2 | 1 |
| **CO3** | 3 | 3 | 2 | 2 | 3 | 2 | 2 | 1 | 2 | 2 | 2 |
| **CO4** | 2 | 2 | 3 | 3 | 3 | 1 | 2 | 2 | 1 | 1 | 3 |
| **CO5** | 2 | 3 | 2 | 3 | 3 | 3 | 2 | 2 | 2 | 3 | 1 |

*(3: High, 2: Medium, 1: Low)*

**Bloom's Taxonomy Levels Mapping**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| COs | Remember | Understand | Apply | Analyze | Evaluate | Create |
| **CO1** | 3 | 3 | 2 | 1 | 1 | - |
| **CO2** | 3 | 3 | 2 | 2 | 1 | - |
| **CO3** | 3 | 3 | 2 | 2 | 1 | - |
| **CO4** | 2 | 2 | 3 | 3 | 2 | 2 |
| **CO5** | 2 | 3 | 3 | 3 | 3 | - |

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –II (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE – 4**: **FUNDAMENTALS OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY**

**No. of Hours/week: 03 Course Code: 24 – STA – 2C4**  **C**redits **–** 3

**II. Syllabus**

**UNIT - I**

**Introduction to Computers**

**Computer:** Definition and evolution of computers, Components of a computer system (Hardware, Software, Data),Classification of computers (Analog, Digital, Hybrid), Generation of computers (First to Fifth Generation) .**Computer System Organization**: Overview of hardware and software components, Functions of the central processing unit (CPU) and memory (RAM), Input/output (I/O) systems, Concept of a bus system in computers. **Types of Software**: **System Software**: Operating systems, device drivers, and utility programs, **Application Software**: Word processors, spreadsheets, media players, etc, **Firmware**: Software embedded in hardware. **Data Representation**: Binary and hexadecimal number systems, Computers process and store data, Conversion between different numeral systems, Encoding schemes like ASCII, Unicode, and Extended ASCII. **Basic Computer Applications**: **Education**: E-learning tools, simulations, digital libraries, **Business**: Inventory management, accounting software, CRM systems, **Entertainment**: Video games, multimedia content, and digital art.

**UNIT – II**

**Computer Hardware**

**Input Devices**: **Basic Input Devices**: Keyboard, mouse, track pad, touch screen, **Advanced Input Devices**: Joystick, scanner, biometric devices, stylus, **Speech Recognition**: Microphone and voice-to-text technology. **Output Devices**: **Basic Output Devices**: Monitor, printer, speakers, **Advanced Output Devices**: Projectors, VR headsets, 3D printers. **Display Technologies**: CRT, LCD, LED, OLED, and touch screens. **Storage Devices**: **Primary Storage**: RAM, ROM, Cache memory. **Secondary Storage**: HDD, SSD, Optical Discs, USB Flash Drives, Magnetic Tapes. **Tertiary and Cloud Storage**: Network-attached storage (NAS), cloud storage solutions (Google Drive, Drop box). **Motherboard and CPU**: **Motherboard Architecture**: Chipsets, buses, power connectors, PCI slots. **CPU Components**: Registers, Clock Speed, Cache Memory. **Multicore and Hyper threading Technology**: Improving processing power. **Peripheral Devices**: **USB Devices**: USB hubs, external storage devices. **Network Interface Cards (NIC)**: Wired and wireless connections. **Other Peripherals**: External hard drives, webcams, printers, and scanners.

**UNIT – III**

**Operating Systems**

**Definition and Functionality of Operating Systems**: The role of OS in controlling hardware and managing software resources, Processes, threads, and multitasking in OS, User Interface: Command-Line Interface (CLI) vs. Graphical User Interface (GUI), **Memory Management**: Virtual memory, paging, segmentation. **Types of Operating Systems**: **Batch Processing Systems**: Early systems used for large-scale data processing, **Multiprogramming and Time-sharing Systems**: Running multiple applications concurrently, **Real-time Operating Systems (RTOS)**: For embedded systems with stringent timing constraints, **Distributed Operating Systems**: Managing resources across multiple machines in a network .**Popular Operating Systems**: **Windows**: Versions (XP, 7, 8, 10, 11), file system (NTFS, FAT). **macOS**: Unique features of the Apple OS (Finder, Spotlight, Time Machine). **Linux**: Open-source OS, distribution variants (Ubuntu, CentOS, Fedora). **Mobile Operating Systems**: Android, iOS, and their app ecosystems. **File Management**: Files and file systems: FAT32, NTFS, ext3/ext4, File and folder creation, deletion, copy, move, and organization. **File Permissions and Security**: Read, write, execute permissions. **Backup and Restore**: Methods to secure data. **Basic OS Operations**: **Boot Process**: BIOS/UEFI, POST, and boot loaders. **Process Scheduling**: Scheduling algorithms (Round Robin, Shortest Job First, etc.). **System Utilities**: Disk cleanup, task manager, and control panel settings.

**UNIT – IV**

**Word Processing and Office Software**

**Introduction to Word Processing**: **Basic Word Operations**: Typing, undo, redo, search, and replace. **Formatting Text**: Fonts, size, styles, color, highlighting, and special characters. **Page Layout**: Margins, orientation, headers, footers, and page numbering. **Tables and Lists**: Creating tables, bulleted and numbered lists, multi-column text. **Document Review**: Spelling and grammar check, comments, track changes. **Spreadsheets**: **Basic Spreadsheet Operations**: Creating cells, rows, columns, and ranges. **Formulas and Functions**: SUM, AVERAGE, IF, VLOOKUP, and more. **Charts and Graphs**: Creating bar, pie, line, and scatter plots. **Data Management**: Sorting, filtering, and pivot tables. **Presentation Software**: **Creating Slides**: Adding text, images, videos, and charts. **Animations and Transitions**: Slide transitions, object animations. **Presentation Design**: Choosing themes, color schemes, and layouts. **Collaborative Features**: Cloud-based tools like Google Slides for group work. **Email and Communication Tools**: **Email Basics**: How E Mail works, Compose, reply, forward, attach files, and CC/BCC. **Email Management**: Organizing inbox with folders, labels, and filters. **Instant Messaging & Video Calling**: Using platforms like Skype, Slack, Zoom, and Microsoft Teams. **Calendar and Task Management**: Scheduling meetings, reminders, and managing deadlines.

**UNIT – V**

**Internet and Networking**

**Introduction to the Internet**: IP addresses, DNS, TCP/IP model. **Internet Services**: Web browsing, email, FTP, VoIP, and cloud computing. **Browsers and Search Engines**: Google Chrome, Mozilla Firefox, Bing, and search engine optimization (SEO). **Web 2.0 and 3.0**: Evolution of the web from static to dynamic content. **Web Technologies**: **HTML Basics**: Structure of a webpage, tags, elements, and attributes. **CSS Basics**: Styling HTML elements, layouts, colors, fonts, and responsive design. **Web Development Frameworks**: Introduction to tools like Bootstrap and Word Press. **Internet Security**: **Cyber Threats**: Viruses, Trojans, worms, spyware, and ransom ware. **Encryption and Authentication**: SSL/TLS, passwords, two-factor authentication. **Safe Browsing Practices**: Phishing, scam emails, secure websites (HTTPS). **Firewalls and Antivirus Software**: Preventive security measures. **Networking Fundamentals**: **Types of Networks**: LAN, WAN, PAN, MAN, 4G/5G NETWORKS Introduction – 4G/5G vision – 4G /5G features and challenges – Applications of 4G/5G. **Networking Devices**: Routers, switches, modems, and hubs. **Internet Protocols**: IP, HTTP, FTP, SMTP, DNS, and DHCP. **Wireless Networks**: Wi-Fi, Bluetooth, and cellular networks. **Social Media and Communication Tools**: **Social Media Platforms**: Face-book, Instagram, LinkedIn and Twitter, and Tik-Tok. **Collaboration Tools**: Google Drive, Dropbox, and Microsoft One-Drive. **Online Communication**: Forums, discussion boards, Cisco Webx and web conferencing tools (Zoom, Google Meet).

**Practical Sessions: Fundamentals of Computer Science and Information Technology**

**Session 1: Computer Basics**

* Identify and describe the components of a computer system.
* Questions:
	1. What are the primary components of a computer system?
	2. How do hardware and software interact in a computer system?
	3. Explain the role of the CPU in a computer.

**Session 2: Data Representation**

* Convert binary numbers to decimal and hexadecimal.
* Questions:
	1. Convert the binary number 10101 to decimal and hexadecimal.
	2. Explain the significance of ASCII and Unicode in data representation.
	3. How is a character stored in memory?

**Session 3: Input and Output Devices**

* Analyze the functioning of various input and output devices.
* Questions:
	1. Describe the working mechanism of a scanner.
	2. Compare and contrast LED and OLED display technologies.
	3. How does a barcode scanner work?

**Session 4: Operating Systems**

* Perform basic file management operations.
* Questions:
	1. How do you create, rename, and delete a folder in Windows?
	2. What are the differences between FAT32 and NTFS file systems?
	3. Explain the process of setting file permissions in Linux.

**Session 5: Word Processing**

* Create and format a document using MS Word.
* Questions:
	1. How can you insert a table into a Word document?
	2. Explain the process of adding a header and footer.
	3. Describe the steps to apply a style to a paragraph.

**Session 6: Spreadsheets**

* Perform data analysis using spreadsheet formulas.
* Questions:
	1. Write a formula to calculate the average of marks in a column.
	2. Create a pie chart to represent sales data.
	3. What is the difference between relative and absolute cell references?

**Session 7: Networking Fundamentals**

* Explore IP configuration and basic troubleshooting.
* Questions:
	1. What is the purpose of an IP address?
	2. How can you troubleshoot connectivity issues using the ping command?
	3. Explain the difference between a private and a public IP address.

**Session 8: Internet Security**

* Demonstrate safe browsing practices.
* Questions:
	1. What are the key features of a secure website?
	2. Explain the role of two-factor authentication in cyber security.
	3. How does encryption enhance internet security?

**Session 9: Database Basics**

* Perform basic operations in a database management system.
* Questions:
	1. How do you create a new table in MS Access?
	2. What is the difference between a primary key and a foreign key?
	3. Explain the process of running a query to filter data.

**Session 10: Presentation Tools**

* Create a slideshow presentation.
* Questions:
	1. How can you add transitions to a presentation in PowerPoint?
	2. What are the different views available in PowerPoint?
	3. Describe the steps to embed a video in a slide.

**Textbooks for Reference**

1. **Peter Norton**: *Introduction to Computers*, McGraw Hill.
2. **Silberschatz, Galvin, and Gagne**: *Operating System Concepts*, Wiley.
3. **Perry and Morton**: *Microsoft Office 365 & Office 2019 Introductory*, Cengage Learning.

**Reference Books**

1. **Stallings, W.**: *Computer Organization and Architecture*, Pearson.
2. **Tanenbaum, A. S.**: *Structured Computer Organization*, Prentice Hall.
3. **Forouzan, B. A.**: *Data Communications and Networking*, McGraw Hill.
4. **Rainer, R. K., Prince, B., and Watson, H. J.**: *Introduction to Information Systems*, Wiley.
5. **Godfrey, B. and Chappell, D.**: *Introduction to Networking: How the Internet Works*, Wiley.

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER: II (CBCS, w.e.f 2024-25)**

**COURSE TITLE - 4: FUNDAMENTALS OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY**

**Time**: 03Hrs **Course Code: 24 – STA – 2C**4 **Max. Marks: 75**

**SECTION-A**

Answer any **FIVE** questions. Each question carries **TWO** marks **5 X 2 = 10M**

1. What are the main components of a computer system?
2. Explain the concept of a bus system in a computer.
3. List the different types of storage devices and give one example of each.
4. Explain the function of the Arithmetic Logic Unit (ALU) in the CPU.
5. What is the purpose of an operating system?

**SECTION – B**

Answer any **FIVE** questions and each question carries **FIVE** marks **5 X 5 = 25M**

1. What is the role of the Central Processing Unit (CPU) in a computer system?
2. Describe the different types of operating systems and their functions.
3. What are the various types of memory used in a computer system? Explain each one briefly.
4. Explain the working and features of a real-time operating system (RTOS).
5. Explain the process of backing up and restoring data in an operating system.
6. What are input and output devices? Provide examples of each, including advanced input/output devices.
7. Discuss the use of cloud storage and its advantages over traditional storage devices.
8. Explain the concept of file permissions and security in an operating system.

**SECTION – C**

Answer any **FOUR** questions and each question carries **TEN** marks**. 4 X 10 = 40M**

1. Describe the different display technologies (CRT, LCD, LED, OLED). Discuss the advantages of each.
2. Explain the role of the motherboard in a computer. Describe the architecture of a motherboard and its components.
3. Write a detailed explanation of the process of email communication, including the features of email management tools and protocols.
4. Describe the various networking devices (routers, switches, modems, hubs) and their roles in network communication.
5. What is web development? Explain the basic structure of an HTML webpage and its key elements.
6. Discuss the importance of internet security, including common cyber threats, encryption, and preventive security measures.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –III (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 5: THEORETICAL DISCRETE AND CONTINUOUS DISTRIBUTIONS**

**No. of** Hours/week: 03 **Course Code: 24 – STA – 3C5 C**redits **–** 3

**Learning Objectives and Outcomes for Probability Distributions**

**Learning Objectives (LOs):**

1. Understand the fundamental concepts of various probability distributions, including their definitions, properties, and applications.
2. Learn how to compute moments, moment-generating functions (MGF), cumulant-generating functions (CGF), and probability-generating functions (PGF) for different distributions.
3. Study the concepts of skewness, kurtosis, and recurrence relations in probability distributions.
4. Apply limiting cases of specific distributions, such as the binomial distribution to the normal distribution, and the Poisson distribution in various scenarios.
5. Develop skills to solve real-world statistical problems using different probability distributions and related functions.

**Learning Outcomes (LOs):**

1. Demonstrate a clear understanding of the definitions, properties, and applications of various probability distributions.
2. Calculate moments, MGF, CGF, PGF, and solve problems involving skewness and kurtosis.
3. Analyze the limiting cases of distributions like binomial and Poisson to the normal distribution.
4. Apply the various probability distributions to solve practical problems in real-world contexts.
5. Interpret and analyze the behavior of probability distributions using different generating functions and recurrence relations.

**Course Outcomes (COs):**

| **Course Outcomes (COs)** | **Bloom's Taxonomy Level** |
| --- | --- |
| **CO1**: Define and explain the properties of Binomial, Poisson, Negative Binomial, Hypergeometric, Normal, Uniform, Exponential, and Gamma distributions. | **Remembering** |
| **CO2**: Derive and apply the Moment-Generating Function (MGF), Cumulant-Generating Function (CGF), and Probability-Generating Function (PGF) for different distributions. | **Applying** |
| **CO3**: Calculate and interpret skewness, kurtosis, and moments for various distributions. | **Analyzing** |
| **CO4**: Use the limiting behavior of distributions (e.g., binomial to normal, Poisson to normal) to solve practical statistical problems. | **Evaluating** |
| **CO5**: Solve complex problems using the properties and applications of different probability distributions. | **Creating** |

**Program Outcomes (POs):**

1. **PO1**: Apply mathematical and statistical knowledge to solve real-world problems in various fields.
2. **PO2**: Develop critical thinking and analytical abilities in understanding and solving complex problems.
3. **PO3**: Use modern statistical techniques and tools to analyze data and interpret results.
4. **PO4**: Apply probability theory and statistical methods to analyze and interpret the behavior of various distributions.
5. **PO5**: Communicate complex statistical results effectively through appropriate methods.
6. **PO6**: Demonstrate a thorough understanding of statistical distributions and their applications in decision-making.
7. **PO7**: Use computational tools and software for solving statistical problems.
8. **PO8**: Contribute to multidisciplinary teams to apply statistical knowledge in diverse problem-solving contexts.

**Program-Specific Outcomes (PSOs):**

1. **PSO1**: Apply advanced statistical methods to analyze large datasets and make data-driven decisions.
2. **PSO2**: Utilize probability models and distributions to analyze trends and make predictions in practical scenarios.
3. **PSO3**: Solve real-world problems by selecting the appropriate probability distribution and applying related techniques.

**Mapping Table: COs, POs, PSOs with Bloom’s Taxonomy Connectivity**

| **Course Outcomes (COs)** | **Program Outcomes (POs)** | **Program Specific Outcomes (PSOs)** | **Bloom’s Taxonomy** |
| --- | --- | --- | --- |
| **CO1**: Define and explain the properties of various distributions. | **PO1, PO3, PO6**: Apply statistical knowledge to understand distributions and solve problems. | **PSO1, PSO2**: Use distributions to analyze and interpret data. | **Remembering** |
| **CO2**: Derive and apply MGF, CGF, and PGF for different distributions. | **PO2, PO4, PO7**: Develop analytical skills and use computational tools to solve problems. | **PSO2**: Use probability distributions for real-world applications. | **Applying** |
| **CO3**: Calculate and interpret skewness, kurtosis, and moments for various distributions. | **PO2, PO4, PO5**: Analyze statistical data and communicate findings effectively. | **PSO1**: Apply statistical methods to interpret data. | **Analyzing** |
| **CO4**: Use limiting behavior of distributions to solve problems. | **PO3, PO6**: Analyze and interpret the limiting behavior in statistical contexts. | **PSO3**: Solve real-world problems using distribution theory. | **Evaluating** |
| **CO5**: Solve problems using the properties and applications of distributions. | **PO1, PO5, PO8**: Apply statistical methods to solve complex problems in diverse contexts. | **PSO2**: Analyze and interpret practical problems using statistical distributions. | **Creating** |

**Summary of Bloom's Taxonomy Mapping:**

* **Remembering (CO1)**: This outcome involves recalling the basic definitions and properties of probability distributions.
* **Applying (CO2)**: This outcome focuses on applying generating functions (MGF, CGF, PGF) to solve problems involving probability distributions.
* **Analyzing (CO3)**: In this outcome, students calculate and analyze skewness, kurtosis, and moments for different distributions.
* **Evaluating (CO4)**: This outcome involves using the limiting case behavior of distributions to evaluate statistical problems.
* **Creating (CO5)**: This outcome focuses on solving complex problems using all the concepts learned in the course, demonstrating creative application of distribution theory.

This approach ensures that students not only grasp theoretical concepts but also develop the necessary skills to apply them in real-world statistical analysis, while progressing through various cognitive levels of Bloom’s Taxonomy.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –III (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 5: THEORETICAL DISCRETE AND CONTINUOUS DISTRIBUTIONS**

**No. of** Hours/week: 03 **Course Code: 24 – STA – 3C5 C**redits **–** 3

**II. Syllabus**

**UNIT – I**

**Binomial Distribution**

Definition and Properties, Moments, Moment-Generating Function (M.G.F), Cumulant-Generating Function (C.G.F), Probability-Generating Function (P.G.F), Additive Property (if exists), Skewness and Kurtosis, Problems and Applications. **Additional Concepts:** Moments Obtained Through M.G.F, Recurrence Relation for Probabilities, Limiting Case of Binomial Distribution to Normal Distribution.

UNIT - II

**Poisson Distribution:** Definition and Properties, Moments, Moment-Generating Function (M.G.F), Cumulant-Generating Function (C.G.F), Probability-Generating Function (P.G.F), Additive Property (if exists),Skewness and Kurtosis, Problems and Applications. **Additional Concepts:** First Two Moments Obtained Through M.G.F, Recurrence Relation for Probabilities, Limiting Case of Poisson Distribution

**UNIT – III**

Negative Binomial Distribution and Hyper Geometric Distribution

**Negative Binomial Distribution:** Definition and Properties, Moments, Moment-Generating Function (M.G.F), Cumulant-Generating Function (C.G.F), Additive Property (if exists),Recurrence Relation for Probabilities, Limiting Case of Negative Binomial Distribution to Normal Distribution. **Geometric Distribution (G.D.):** Definition, mean and variance, Lack of Memory Property of Geometric Distribution: **Hyper geometric Distribution:** Definition and Properties, Limiting Case of Hyper geometric Distribution to Binomial Distribution

**UNIT – IV**

**Normal Distribution**

Definition and Properties, Importance in Statistics’ **Moment-Generating Functions (M.G.F), Cumulant-Generating Functions (C.G.F), and Characteristic Functions (C.F),** Additive Property, **Skewness and Kurtosis,** Even and Odd Order Moments About the Mean, **Linear Combination of Normal Variates,** Identification and Calculation of Points of Inflexion of the Normal Probability Curve, **Problems and Applications**

**UNIT – V**

**Uniform, Exponential and Gamma distributions**

**Uniform Distribution:** Definition and Properties, Moments, Moment-Generating Function (M.G.F), Cumulant-Generating Function (C.G.F), Characteristic Function (C.F), Distribution Function. **Exponential Distribution:** Definition and Properties, Moments, Moment-Generating Function (M.G.F), Cumulant-Generating Function (C.G.F), Characteristic Function (C.F), Memory less Property of Exponential Distribution. **Gamma Distribution:** Definition and Properties, Moments, Moment-Generating Function (M.G.F), Cumulant-Generating Function (C.G.F), Characteristic Function (C.F), Additive Property, Limiting Form of Gamma Distribution

**Practical Syllabus Practical Credits: 1 2 hrs/week**

1. Fitting of Binomial distribution – Direct method.

2. Fitting of Binomial distribution – Recurrence relation Method.

3. Fitting of Poisson distribution – Direct method.

4. Fitting of Poisson distribution - Recurrence relation Method.

5. Fitting of Negative Binomial distribution – Direct method.

6. Fitting of Negative Binomial distribution – Recurrence relation Method.

7. Fitting of Geometric distribution – Direct method.

8. Fitting of Geometric distribution – Recurrence relation Method.

9. Fitting of Hyper Geometric distribution.

10. Fitting of Exponential distribution.

11. Fitting of Normal distribution – Areas method.

12. Fitting of Normal distribution – Ordinates method.

**Note:** Training shall be on establishing formulae in Excel cells and derive the results. The excel output shall be exported to MS word for writing inference.

**III. Text Books/References**

1. S. C. Gupta & V. K. Kapoor: Fundamentals of Mathematical Statistics, Sultan Chand & Sons, New Delhi.

2. O. P. Gupta: Mathematical Statistics, Kedar nath Ram nath& Co.

3. P. N. Arora & S. Arora: Quantitative Aptitude Statistics – Vol II, S. Chand & Company Ltd.

4. K. Rohatgi &Ehsanes Saleh: An Introduction to Probability and Statistics, John Wiley & Sons.

**IV. Suggested Co-curricular Activities:**

1. Training of students by related industrial experts

2. Assignments including technical assignments if any.

3. Seminars, Group Discussions, Quiz, Debates etc on related topics.

4. Preparation of audio and videos on tools of diagrammatic and graphical representations.

5. Collection of material/figures/photos/author photoes of related topics.

6. Invited lectures and presentations of stalwarts to those topics.

7. Visits/field trips of firms, research organizations etc.

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER: III (CBCS, w.e.f 2024-25)**

**COURSE TITLE - 5: THEORETICAL DISCRETE AND CONTINUOUS DISTRIBUTIONS**

**Time**: 03Hrs **Course Code: 24 – STA – 3C5 Max. Marks:75**

**SECTION-A**

Answer any **FIVE** questions. Each question carries **TWO** marks **5X 2 = 10M**

1. Define Binomial distribution.
2. What are the Mean and variance of Poisson Distribution..
3. What are the MGF and CF of Negative Binomial Distribution.
4. Define Exponential Distribution.
5. Define Normal Distribution.

**SECTION – B**

Answer any **FIVE** questions and each question carries **FIVE** marks **5 X5 = 25M**

1. Find its Mean and Variance through MGF of Binomial Distribution.
2. Discuss the recurrence relation for probabilities in the context of the Poisson distribution.
3. Define HGD. Derive its Mean and Variance.
4. State and prove Lack of memory property of Geometric distribution
5. Derive the Mean Deviation about mean of Exponential Distribution
6. Deduce mean and variance of Gamma Distribution.
7. Derive m.g.f. of Normal Distribution
8. Prove that in Normal distribution odd ordered moments does not exists but even order moments exists.

**SECTION – C**

Answer any **FOUR** questions and each question carries **TEN** marks**. 4 X 10 = 40M**

1. How can the first two moments of the Poisson distribution be obtained using the C.F.?
2. Discuss the limiting case of the Binomial distribution to the Poisson distribution.
3. Explain the first two moments of the Geometric distribution obtained through the M.G.F.
4. Discuss the Mean Deviation about the mean for the Uniform distribution.
5. Define the Gamma distribution and discuss its moments using M.G.F.
6. How can you obtain the mean, median, and mode of the Normal distribution? Discuss the relationship between them.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –III (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 6: STATISTICAL METHODS AND INFERENTIAL STATISTICS**

**No. of Hours/week: 03 Course Code: 24 – STA – 3C6 Credits – 3**

**Learning Objectives:**

1. Understand and apply the concepts of correlation, regression analysis, curve fitting, and hypothesis testing.
2. Develop proficiency in estimation techniques such as Maximum Likelihood Estimation (MLE) and Moment Estimation.
3. Gain a thorough understanding of large and small sample tests, including their assumptions and applications.
4. Analyze data using various statistical tools and interpret the results effectively.
5. Enhance critical thinking by comparing and contrasting correlation with regression, and understanding the different types of errors in hypothesis testing.

**Learning Outcomes:**

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

1. Define and calculate different types of correlation coefficients and interpret them in the context of bi-variate data.
2. Understand the process of curve fitting, apply the least squares principle, and interpret regression lines for bi-variate data.
3. Apply the method of moments and Maximum Likelihood Estimation (MLE) for parameter estimation in different distributions.
4. Perform hypothesis testing for different distributions, including binomial, Poisson, and normal, and calculate p-values and test statistics.
5. Implement large and small sample tests effectively, including t-tests, chi-square tests, and confidence intervals for mean, proportions, and variances.

**Mapping Summary Table**

| **Course Outcomes (COs)** | **Program Outcomes (POs)** | **Program Specific Outcomes (PSOs)** |
| --- | --- | --- |
| **CO1**: Define and calculate different types of correlation coefficients (positive, negative, zero) and interpret their meaning in bi-variate data. | **PO1**: Apply knowledge of mathematics to solve complex problems in statistics and probability. | **PSO1**: Demonstrate the ability to apply advanced statistical methods and techniques to analyze and interpret large datasets. |
| **CO2**: Understand and apply curve fitting techniques using the principle of least squares, including fitting different types of curves and regression analysis. | **PO2**: Develop critical thinking and analytical skills to evaluate data and make informed decisions based on statistical methods. | **PSO2**: Utilize probability models and random variables to make data-driven predictions and decisions in real-world applications. |
| **CO3**: Apply the method of moments and Maximum Likelihood Estimation (MLE) to estimate population parameters in different distributions. | **PO3**: Demonstrate an understanding of probabilistic models and random variables in both theoretical and practical contexts. | **PSO3**: Develop problem-solving skills using generating functions, inequalities, and other advanced statistical tools to address practical challenges in fields such as finance, healthcare, and engineering. |
| **CO4**: Perform hypothesis testing using different distributions, and understand key concepts like errors, significance levels, and power of a test. | **PO4**: Gain proficiency in using statistical software (e.g., Excel, R, or Python) to perform calculations and statistical analysis. |  |
| **CO5**: Implement large and small sample tests, including t-tests, chi-square tests, and calculate confidence intervals for various population parameters. | **PO5**: Work effectively both individually and as part of a team to apply statistical concepts to real-world problems. |  |
|  | **PO6**: Understand the role of mathematical expectations, variance, and covariance in interpreting and solving statistical problems. |  |
|  | **PO7**: Effectively use generating functions, inequalities, and probability distributions to model complex systems. |  |
|  | **PO8**: Apply probability theory to various fields, including economics, engineering, and healthcare, to analyze risk and uncertainty. |  |

**Bloom's Taxonomy Mapping:**

| **Learning Objective** | **Bloom’s Level** | **Course Outcome (CO)** | **Program Outcome (PO)** | **Program Specific Outcome (PSO)** |
| --- | --- | --- | --- | --- |
| Understand and apply the concepts of correlation, regression analysis, curve fitting, and hypothesis testing. | **Understanding, Applying** | **CO1**: Define and calculate different types of correlation coefficients and interpret their meaning in bi-variate data. | **PO1**, **PO2** | **PSO1** |
| Develop proficiency in estimation techniques such as Maximum Likelihood Estimation (MLE) and Moment Estimation. | **Applying** | **CO3**: Apply the method of moments and Maximum Likelihood Estimation (MLE) to estimate population parameters in different distributions. | **PO3** | **PSO2** |
| Gain a thorough understanding of large and small sample tests, including their assumptions and applications. | **Understanding, Applying** | **CO5**: Implement large and small sample tests, including t-tests, chi-square tests, and calculate confidence intervals for various population parameters. | **PO5**, **PO6** | **PSO3** |
| Analyze data using various statistical tools and interpret the results effectively. | **Analyzing** | **CO4**: Perform hypothesis testing using different distributions, and understand key concepts like errors, significance levels, and power of a test. | **PO4**, **PO5** |  |
| Enhance critical thinking by comparing and contrasting correlation with regression, and understanding the different types of errors in hypothesis testing. | **Evaluating** | **CO2**: Understand and apply curve fitting techniques using the principle of least squares, including fitting different types of curves and regression analysis. | **PO2**, **PO6** | **PSO1**, **PSO2** |

**Mapping Explanation:**

* **CO1** and **CO2** align with **PO1** and **PO2**, ensuring students can apply correlation and regression techniques to real-world data and develop analytical skills to interpret relationships.
* **CO3** and **CO5** link with **PO3**, **PO5**, and **PO6**, focusing on estimation methods like MLE and hypothesis testing, which students will use to make informed decisions and solve complex statistical problems.
* **CO4** contributes to **PO4** and **PO5**, developing students’ abilities to perform hypothesis tests using various distributions, with a focus on error analysis and understanding power tests.
* **PSO1** and **PSO2** are closely mapped to **CO1**, **CO2**, and **CO3**, ensuring students apply statistical tools for interpreting large datasets, fitting curves, and making predictions in real-world scenarios.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –III (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 6: STATISTICAL METHODS AND INFERENTIAL STATISTICS**

**No. of Hours/week: 03 Course Code: 24 – STA – 3C6 Credits – 3**

**II. Syllabus**

**UNIT – I**

**Introduction to Correlation:** Meaning and Importance of Correlation, Types of Correlation: Positive, Negative, and Zero Correlation, **Measures of Correlation:** Scatter Diagram: Interpretation and Construction, Karl Pearson’s Coefficient of Correlation: Definition and Calculation, Rank Correlation Coefficient: With Ties and Without Ties, Properties of Correlation Coefficients. **Bi-Variate Frequency Distribution:** Concept and Construction of Bi-Variate Frequency Distribution, **Correlation Coefficient for Bi-Variate Data,** Calculation and Interpretation of Correlation Coefficient for Bi-Variate Data.

**UNIT – II**

**Curve Fitting and Regression Analysis**

**Curve Fitting for Bi-Variate Data:** Concept and Importance of Curve Fitting, Principle of Least Squares, Fitting of k-th Degree Polynomial, **Fitting of Different Curves:** Fitting of Straight Line, Fitting of Second Degree Polynomial (Parabola), Fitting of Family of Exponential Curves, Fitting of Power Curve. **Regression Analysis:** Concept of Regression, Linear and Non-Linear Regression, **Regression Lines for Bi-Variate Data:** Calculation and Interpretation of Regression Lines, Simple Problems on Regression, **Correlation vs Regression differences.**

**UNIT*–* III**

**Theory of estimation**: Estimation of a Parameter**,** Concept of Parameter Estimation, Criteria for a Good Estimator: Unbiasedness , Consistency, Efficiency and Sufficiency, Statement of **Neymann's Factorization Theorem. Methods of Estimation:** Estimation by the Method of Moments, Estimation by Maximum Likelihood Estimation (M.L.E) and Properties of Maximum Likelihood Estimators (MLEs), **Estimation of Population Parameters Using MLE:** Estimation of Binomial, Poisson, and Normal Population Parameters by MLE Method. **Confidence Intervals**

***UNIT – IV***

**Testing of Hypothesis:** Concepts of Statistical Hypotheses**,** Definition and Importance of Statistical Hypotheses, Null and Alternative Hypotheses. **Key Components of Hypothesis Testing:** Critical Region ,Types of Errors: Type I and Type II Errors , Level of Significance, Power of a Test. **One and Two-Tailed tests. Neymann-Pearson’s Lemma. Examples of Hypothesis Testing:** Hypothesis Testing for Binomial, Poisson, Exponential, and Normal Distributions.

***UNIT – V***

**Large Sample Tests: Test for Single Mean**: Concept, Formula, and Application, **Test for Difference of Two Means**: Procedure, Formula, and Application, **Confidence Intervals for Mean(s)**: Calculation and Interpretation, **Test for Single Proportion**: Concept, Formula, and Application, **Test for Difference of Proportions**: Procedure and Application, **Tests for Standard Deviation(s)**: Hypothesis Testing for Population Standard Deviation, **Test for Correlation Coefficient(s)**: Testing for Linear Relationship

**Small Sample Tests :** Assumptions Underlying t-tests: **t-Test for Single Mean**: Definition, Procedure, and Application, **t-Test for Difference of Means**: Formula, Assumptions, and Application, **Paired t-Test**: Concept, Formula, and Application, **Chi-Square (𝜒²) Test for Goodness of Fit**: Procedure and Application , **Chi-Square (𝜒²) Test for Independence of Attributes**: Testing Association Between Two Categorical Variables, **Chi-Square (𝜒²) Test for Single Variance**: Hypothesis Testing for Population Variance

**Note:** Training shall be on establishing formulae in Excel cells and derive the results. The excel output shall be exported to MS word for writing inference.

**III. References**

1. S. C. Gupta & V. K. Kapoor: Fundamentals of Mathematical Statistics, Sultan Chand & Sons, New Delhi.

2. O. P. Gupta: Mathematical Statistics, Kedar nath Ram nath& Co.

3. P. N. Arora & S. Arora: Quantitative Aptitude Statistics – Vol II, S. Chand & Company Ltd.

4. K. Rohatgi &Ehsanes Saleh: An Introduction to Probability and Statistics, John Wiley & Sons.

**IV. Suggested Co-curricular Activities:**

1. Training of students by related industrial experts

2. Assignments including technical assignments if any.

3. Seminars, Group Discussions, Quiz, Debates etc on related topics.

4. Preparation of audio and videos on tools of diagrammatic and graphical representations.

5. Collection of material/figures/photos/author photoes of related topics.

6. Invited lectures and presentations of stalwarts to those topics.

7. Visits/field trips of firms, research organizations etc.

**24-STA-3C6P Practicals for Course – II**  **Credits 2**

1. Fitting of straight line by the method of least-squares.
2. Fitting of parabola by the method of least-squares.
3. Fitting of exponential curve of two types by the method of least-squares.
4. Fitting of power curve of the type by the method of least-squares.
5. Computation of correlation coefficient and regression lines for ungrouped data.
6. Computation of correlation coefficient for Bi-Variate frequency distribution.
7. Computation of correlation coefficient, forming regression lines for grouped data.
8. Computation of Yule's coefficient of association and colligation.
9. Computation of Pearson's, Tschuprow’s coefficient of contingency.
10. Large sample test for single mean and for difference of two means
11. Large sample test for single proportion and for difference of two proportions
12. Large sample test for difference of standard deviations
13. Large sample test for correlation coefficient (Fisher Z – transformation).
14. Small sample test for single mean and for difference of two means
15. Paired t-test (paired samples).
16. Small sample test for difference of variances (F - test)
17. 𝜒2 test for goodness of fit and independence of attributes.
18. Nonparametric tests for single sample(run test, sign test and Wilcoxon signed rank test)
19. Nonparametric tests for related samples (sign test and Wilcoxon signed rank test)
20. Nonparametric tests for two independent samples (Median test, Wilcoxon –Mann- Whitney - U test, Wald - Wolfowitz' s runs test)

**Note:** Training shall be on establishing formulae in Excel cells and derive the results. The excel output shall be exported to MS word for writing inference.

Suggested Co-curricular Activities:

1. Training of students by related industrial experts

2. Assignments including technical assignments if any.

3. Seminars, Group Discussions, Quiz, Debates etc on related topics.

4. Preparation of audio and videos on tools of diagrammatic and graphical representations.

5. Collection of material/figures/photos/author photos of related topics.

6. Invited lectures and presentations of stalwarts to those topics.

7. Visits/field trips of firms, research organizations etc.

***Reference books:***

1. V.K. Kapoor and S.C. Gupta: Fundamentals of Mathematical Statistics, Sultan Chand & Sons, New Delhi.

2. BA/B. Sc I year statistics - descriptive statistics, probability distribution - Telugu Academy

1. Dr M. JaganmohanRao, Dr N. Srinivasa Rao, Dr P. Tirupathi Rao, Smt. D. Vijayalakshmi.
2. K.V.S. Sarma: Statistics Made Simple: Do it yourself on PC. PHI.
3. Willam Feller: Introduction to Probability theory and its applications. Volume –I,Wiley
4. Goon AM, Gupta MK, Das Gupta B : Fundamentals of Statistics , Vol-I, the World Press Pvt. Ltd., Kolakota.
5. Hoel P.G: Introduction to mathematical statistics, Asia Publishing house.
6. M. Jagan Mohan Rao and Papa Rao: A Text book of Statistics Paper-I.
7. Sanjay Arora and Bansi Lal: New Mathematical Statistics: Satya Prakashan , New-Delhi
8. Hogg Tanis Rao: Probability and Statistical Inference. 7thedition. Pearson.

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER: III (CBCS, w.e.f 2024-25)**

**COURSE TITLE 6: STATISTICAL METHODSAND INFERENTIAL STATISTICS**

**Time: 3 Hours Subject Code: 24-STA-3C6 Max. Marks: 75**

**SECTION – A**

Answer any **ALL** questions. Each question carry **TWO** marks **5 X 2 = 10M**

1. Define Correlation.
2. What is Curve fitting?
3. What is an unbiased estimator?
4. Define Statistical Hypothesis.
5. What is the paired t-test?

**SECTION – B**

Answer any **FIVE** questions. Each question carry **FIVE** marks **5 X 5 = 25M**

1. Obtain the Rank Correlation Coefficient to the following data

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| x | 11.1 | 10.3 | 12 | 15.1 | 13.7 | 18.5 | 17.3 | 14.2 | 14.8 | 15.3 |
| y | 10.9 | 14.2 | 13.8 | 21.5 | 13.2 | 21.1 | 16.4 | 19.3 | 17.4 | 19.0 |

1. Calculate karl Pearson Correlation Coefficient for the given data.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X | 78 | 25 | 56 | 94 | 23 | 45 | 76 | 31 | 13 | 65 |
| Y | 28 | 76 | 47 | 24 | 80 | 55 | 35 | 69 | 98 | 40 |

1. Fit a straight line of the type y = a + bx for the given data.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| X | 1960 | 1970 | 1980 | 1990 | 2000 | 2010 | 2020 |
| Y | 27 | 22 | 45 | 86 | 65 | 74 | 98 |

1. Define Regression Coefficients and give its properties.
2. What are the criteria for a good estimator?
3. Define i) Null Hypothesis ii) Alternative Hypothesis
4. Explain student t-test to test equality between two sample means
5. Explain the χ2 test for goodness of fit?

**SECTION – C**

 Answer any **FOUR** questions and each question carries **10** marks **4 X 10 = 40M**

1. Explain principle of least squares to fit a Parabola of the form y = a + bx+cx2.
2. Calculate Coefficient of Correlation for the give bi-variate frequency data.

|  |  |
| --- | --- |
| Hours studied | Test Scores |
|  | 20 – 30  | 30 – 40  | 40 – 50  | 50 – 60  | 60 – 70  |
| 0 – 2 | 15 | 26 | 12 |  |  |
| 2 – 4  |  | 20 | 35 | 13 |  |
| 4 – 6  |  | 12 | 25 | 10 | 6 |
| 6 – 8 |  |  | 15 | 24 | 8 |
| 8 – 10 |  |  | 05 | 07 | 01 |

1. Derive Regression lines of y on x and x on y.
2. Define and explain the concept of consistency.
3. State and prove Neymann Pearson Lemma.
4. Explain the i) Paired t - test ii) F – test.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –III (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 7: STATISTICAL ANALYSIS AND REPORTING USING MS WORD, MS EXCEL, AND POWERPOINT**

**No. of Hours/week: 03 Course Code: 24 – STA – 3C7 Credits – 3**

**Learning Objectives (LOs):**

This course aims to provide students with practical knowledge and skills in using MS Word, MS Excel, and MS PowerPoint for statistical data analysis, reporting, and presentations. Students will learn to use these tools effectively for data entry, statistical functions, data visualization, and presentation of statistical results.

1. Understand and apply MS Word features for effective statistical report creation.
2. Develop proficiency in using MS Excel for statistical data analysis and visualization.
3. Use advanced MS Excel functions for performing statistical modeling and forecasting.
4. Master the tools in MS PowerPoint for presenting statistical data effectively.
5. Integrate MS Word, Excel, and PowerPoint for a cohesive statistical analysis project.

**Course Outcomes (COs):**

1. **CO1**: Demonstrate the ability to use MS Word for structuring and formatting professional statistical reports.
2. **CO2**: Apply MS Excel for performing statistical analysis, including data entry, analysis, and visualization.
3. **CO3**: Conduct advanced statistical analysis such as regression, correlation, and ANOVA using MS Excel.
4. **CO4**: Create effective and engaging presentations of statistical results using MS PowerPoint.
5. **CO5**: Integrate MS Word, Excel, and PowerPoint to complete a statistical project with data analysis, written reports, and presentations.

**Program Outcomes (POs):**

1. **PO1**: Apply critical thinking skills to solve data-driven problems using MS Office tools.
2. **PO2**: Demonstrate proficiency in using MS Word, Excel, and PowerPoint for statistical reporting and analysis.
3. **PO3**: Analyze and interpret data using MS Excel’s statistical functions and visualization tools.
4. **PO4**: Effectively present complex statistical results to a non-technical audience using MS PowerPoint.
5. **PO5**: Integrate and combine multiple MS Office tools for a unified approach to statistical analysis and reporting.
6. **PO6**: Develop advanced skills in statistical analysis techniques including regression, correlation, and ANOVA.
7. **PO7**: Apply ethical considerations in presenting statistical results and analysis.
8. **PO8**: Work collaboratively using MS Office tools for group projects and statistical analyses.

**Program Specific Outcomes (PSOs):**

1. **PSO1**: Ability to effectively use MS Word, MS Excel, and MS PowerPoint for statistical reporting, analysis, and presentations.
2. **PSO2**: Demonstrate proficiency in advanced statistical analysis, including forecasting, regression, and hypothesis testing, using MS Excel.
3. **PSO3**: Present statistical results effectively through integrated MS Office tools (Word, Excel, PowerPoint) for professional use.

**Mapping Summary Table**

| **Unit** | **Core Topics** | **COs** | **POs** | **PSOs** | **Bloom’s Taxonomy Levels** |
| --- | --- | --- | --- | --- | --- |
| **Unit 1** | Introduction to MS Word for Statistical Reporting | CO1 | PO1, PO2 | PSO1 | Remember, Understand, Apply |
| **Unit 2** | Introduction to MS Excel for Statistical Data Analysis | CO2 | PO2, PO3 | PSO1, PSO2 | Understand, Apply, Analyze |
| **Unit 3** | Using MS Excel for Statistical Modeling and Forecasting | CO3 | PO3, PO6 | PSO2 | Apply, Analyze, Evaluate |
| **Unit 4** | MS PowerPoint for Presenting Statistical Results | CO4 | PO4, PO5 | PSO1 | Create, Evaluate, Apply |
| **Unit 5** | Integrating MS Word, MS Excel, and PowerPoint for Statistical Projects | CO5 | PO5, PO7, PO8 | PSO1, PSO3 | Apply, Analyze, Create |

**Bloom’s Taxonomy Connectivity:**

* **Remembering**: Recall basic features of MS Word, Excel, and PowerPoint (Unit 1, 2).
* **Understanding**: Grasp the purpose and utility of statistical analysis and visualization in Excel (Unit 2, 3).
* **Applying**: Use statistical tools in Excel and reporting features in Word and PowerPoint (Unit 2, 4).
* **Analyzing**: Evaluate statistical results, create forecasts, and conduct regression/correlation analysis (Unit 3, 5).
* **Evaluating**: Interpret statistical results to decide the best model or visualization (Unit 3, 4).
* **Creating**: Develop integrated reports and presentations combining MS Word, Excel, and PowerPoint (Unit 5).

This mapping connects the core topics to the course outcomes, program outcomes, and program-specific outcomes, while ensuring each unit aligns with specific levels of Bloom's Taxonomy for cognitive development. It promotes a comprehensive understanding and application of MS Office tools in statistical analysis and reporting.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –III (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 7: STATISTICAL ANALYSIS AND REPORTING USING MS WORD, MS EXCEL, AND POWERPOINT**

**No. of Hours/week: 03 Course Code: 24 – STA – 3C7 Credits – 3**

**II. Syllabus:**

**UNIT *–* I**

**Introduction to MS Word for Statistical Reporting**

**Overview of MS Word**: Interface, document formatting, and tools. **Data Entry and Text Formatting**: Inserting tables, text styles, headings, and bullet points. **Inserting Statistical Elements**: Charts, graphs, equations, and symbols. **Report Writing**: Structuring a statistical report (introduction, methods, results, discussion). **Referencing and Citations**: Using citation tools and generating bibliographies. **Statistical Templates**: Utilizing templates for report formatting and style consistency.

**UNIT *–* II**

**Introduction to MS Excel for Statistical Data Analysis**
**Overview of MS Excel**: Understanding the spreadsheet interface, cells, rows, columns. **Data Entry and Management**: Entering and organizing data, sorting, filtering, and data validation. **Descriptive Statistics in Excel**: Using functions like MEAN, MEDIAN, MODE, STDEV, etc. **Data Visualization**: Creating histograms, pie charts, bar charts, and scatter plots. **Advanced Excel Functions**: VLOOKUP, INDEX MATCH, IF statements, and pivot tables. **Statistical Analysis Tools**: Using Data Analysis Toolpak for regression, correlation, and hypothesis testing.

**UNIT *–* III**

**Using MS Excel for Statistical Modeling and Forecasting**
**Regression Analysis**: Simple linear regression, multiple regression analysis using Excel. **Correlation Analysis**: Understanding and calculating Pearson correlation coefficient. **Time Series Analysis**: Creating and interpreting trend lines and forecasting. **ANOVA**: One-way and two-way ANOVA in Excel. **Chi-Square Tests**: Conducting chi-square tests for independence and goodness of fit. **Excel Solver**: Using Solver for optimization problems and finding statistical solutions. Project: Perform regression analysis and create statistical models using Excel.

**UNIT *–* IV**

**MS PowerPoint for Presenting Statistical Results**
**Overview of PowerPoint**: Interface, slide layout, and design basics. **Creating Statistical Presentations**: How to present data effectively using PowerPoint. **Inserting Statistical Graphs and Charts**: Adding graphs from Excel, using data-driven visuals. **Using SmartArt and Diagrams**: Presenting statistical processes and flowcharts. **Statistical Tables and Figures**: Formatting tables and figures for clear data presentation. **Effective Presentation Tips**: Tips for presenting complex statistical results clearly and engagingly.

**UNIT *–* V**

**Integrating MS Word, MS Excel, and PowerPoint for Statistical Projects**
**Integrating Data from Excel to Word**: Importing charts, tables, and graphs into MS Word for reports. **Creating Interactive Presentations**: Combining Excel data and PowerPoint charts for dynamic presentations. **Data Analysis Workflow**: Managing data in Excel, generating results, and presenting findings in Word and PowerPoint. **Report and Presentation Finalization**: Formatting documents and presentations for consistency and clarity. **Collaboration Tools**: Using MS Office collaboration tools (comments, track changes, sharing).

**24-STA-3C7P Practical Course – VII**  **Credits 2**

**UNIT *–* I**

***Introduction to MS Word for Statistical Reporting***

1. **Document Formatting**: Create a statistical report template in MS Word with a title, headings, and a table of contents. Format the document using different styles for headings and body text.
2. **Tables and Text**: Insert a table into an MS Word document, populate it with sample data, and format it with a professional style.
3. **Statistical Elements**: Add a bar graph and a scatter plot to an MS Word document. Include equations and symbols relevant to statistical analysis.
4. **Report Structure**: Create a sample statistical report with sections for *Introduction, Methods, Results,* and *Discussion.* Ensure consistent formatting throughout.
5. **Referencing**: Use the citation tool in MS Word to add references in APA style. Generate a bibliography at the end of the report.

**UNIT – II**

 **Introduction to MS Excel for Statistical Data Analysis**

1. **Data Entry**: Enter sample data into an Excel spreadsheet. Sort and filter the data to display records meeting specific criteria.
2. **Descriptive Statistics**: Calculate the mean, median, mode, standard deviation, and variance for a given dataset using Excel functions.
3. **Data Visualization**: Create a pie chart and a histogram using the data provided. Format the charts for clear presentation.
4. **Advanced Functions**: Use the VLOOKUP function to find data from a lookup table. Combine INDEX and MATCH functions to retrieve data based on specific conditions.
5. **Statistical Tools**: Use the Data Analysis Toolpak to perform a regression analysis and interpret the results.

**UNIT – III**

**Using MS Excel for Statistical Modeling and Forecasting**

1. **Regression Analysis**: Conduct a simple linear regression analysis on a dataset. Include the regression equation and R² value in your results.
2. **Correlation Analysis**: Calculate the Pearson correlation coefficient for a dataset and interpret its value.
3. **Time Series Analysis**: Create a time series plot for a dataset. Add a trend line and forecast the next three data points.
4. **ANOVA**: Perform a one-way ANOVA using a dataset in Excel. Interpret the F-statistic and p-value.
5. **Chi-Square Test**: Conduct a chi-square test of independence on a contingency table and interpret the results.

**UNIT *–* IV**

***MS PowerPoint for Presenting Statistical Results***

1. **Statistical Presentation**: Design a PowerPoint presentation with slides summarizing statistical results, including charts, tables, and key findings.
2. **SmartArt**: Use SmartArt to create a flowchart explaining a statistical process.
3. **Graph Integration**: Import a graph from Excel into a PowerPoint slide and customize its appearance for better presentation.
4. **Effective Slides**: Create a slide with key statistical results and format it to be visually appealing and easy to understand.
5. **Final Presentation**: Compile a full presentation with a title slide, introduction, methods, results, and conclusions, using best practices for clarity and engagement.

**UNIT *–* V**

***Integrating MS Word, MS Excel, and PowerPoint for Statistical Projects***

1. **Data Integration**: Import charts and tables from an Excel analysis into an MS Word report. Maintain consistency in formatting.
2. **Interactive Presentation**: Link Excel data to PowerPoint charts so that updates in the Excel file automatically reflect in the presentation.
3. **Collaboration**: Use the "Track Changes" feature in MS Word to review and comment on a statistical report collaboratively.
4. **Workflow**: Demonstrate the process of managing data in Excel, analyzing results, and presenting findings in a PowerPoint presentation.
5. **Final Report**: Combine elements from Word, Excel, and PowerPoint to create a polished statistical report and presentation.

**Reference Books and Textbooks**

1. **"Microsoft Word for Beginners: Master the Essentials"** by M.L. Humphrey
	* Covers fundamental concepts of MS Word, including document formatting, tools, and citation techniques.
2. **"Microsoft Word in Easy Steps"** by Scott Basham
	* A beginner-friendly guide to using MS Word for creating professional documents, reports, and templates.
3. **"Excel 2019 for Statistical Analysis: A Guide to Solving Practical Problems"** by Thomas J. Quirk
	* Focuses on using Excel for statistical analysis, including descriptive statistics, regression, and ANOVA.
4. **"Statistical Analysis with Excel for Dummies"** by Joseph Schmuller
	* Explains statistical tools and concepts, data visualization, and analysis using MS Excel in a simplified way.
5. **"Microsoft Excel Data Analysis and Business Modeling"** by Wayne Winston
	* A detailed resource on advanced Excel techniques, including data analysis tools, Solver, and forecasting.
6. **"Presentation Zen: Simple Ideas on Presentation Design and Delivery"** by Garr Reynolds
	* Offers techniques for creating impactful and visually engaging PowerPoint presentations.
7. **"Microsoft PowerPoint: Tips and Tricks for Effective Presentations"** by Vickie L. Milazzo
	* Provides practical advice on designing professional slides, integrating visuals, and presenting data.
8. **"Integrating Microsoft Office for Statistical Analysis and Reporting"** by Ray J. Ouellette
	* Guides users in integrating Word, Excel, and PowerPoint for comprehensive statistical projects and reports.

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER: III (CBCS, w.e.f 2024-25)**

**COURSE TITLE 7: STATISTICAL ANALYSIS AND REPORTING USING MS WORD, MS EXCEL, AND POWERPOINT**

**Time: 3 Hours Subject Code: 24-STA-3C7 Max. Marks: 75**

**SECTION – A**

Answer any **ALL** questions. Each question carry **TWO** marks **5 X 2 = 10M**

1. List any four key features of MS Word used in statistical report writing.
2. Define the purpose of the Data Analysis Toolpak in MS Excel.
3. What is the significance of regression analysis in statistical modeling?
4. Mention two tips for presenting statistical results effectively in PowerPoint.
5. What are collaboration tools in MS Office used for?

**SECTION – B**

Answer any **FIVE** questions. Each question carry **FIVE** marks **5 X 5 = 25M**

1. Explain the steps to insert and format statistical tables in MS Word.
2. Describe how descriptive statistics functions (e.g., MEAN, MEDIAN, MODE) are used in MS Excel.
3. How can time series analysis be performed using Excel?
4. Illustrate the process of integrating Excel charts into MS Word documents.
5. How can you import Excel charts and tables into Word for a report?
6. Explain the role of Smart-Art and diagrams in presenting statistical results in PowerPoint.
7. What are the key steps in performing a Chi-Square Test in Excel?
8. How do collaboration tools help in finalizing statistical reports?

**SECTION – C**

Answer any **FOUR** questions and each question carries **10** marks **4 X 10 = 40M**

1. Explain the process of structuring a statistical report in MS Word, including referencing and citations.
2. Describe the steps for performing multiple regression analysis using MS Excel with an example.
3. Discuss the workflow of data analysis using Excel and presenting the results in Word and PowerPoint.
4. How can interactive presentations be created by integrating data from Excel into PowerPoint?
5. Write a detailed note on ANOVA analysis using MS Excel with an example.
6. Explain the importance of consistency and clarity in finalizing statistical reports and presentations.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER: III (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 8: STATISTICAL PROGRAMMING WITH ‘C’ LANGUAGE AND DATA STRUCTURES**

**Time: 03**Hrs **Course Code: 24 – STA – 3C**8 **Max. marks: 75**

**Program Outcomes (POs):**

1. **PO1**: Apply mathematical, statistical, and computational skills to analyze and solve real-world problems.
2. **PO2**: Develop efficient programming solutions using C language for statistical problems.
3. **PO3**: Understand and implement data structures to organize and manage statistical data.
4. **PO4**: Perform statistical computations using algorithms in C.
5. **PO5**: Communicate technical information and computational results effectively.
6. **PO6**: Adapt to various programming paradigms and debugging techniques.
7. **PO7**: Exhibit problem-solving skills with an emphasis on precision and efficiency.
8. **PO8**: Develop teamwork and collaboration skills in computational projects.

**Course Outcomes (COs):**

1. **CO1**: Understand the basics of C programming for statistical applications.
2. **CO2**: Develop programs in C for data manipulation and statistical calculations.
3. **CO3**: Implement fundamental data structures and algorithms.
4. **CO4**: Apply C programming to solve statistical problems.
5. **CO5**: Demonstrate effective data handling and computational efficiency.

## ****Mapping of COs, POs, and PSOs with Bloom’s Taxonomy****

### ****Course Outcomes (COs)****

| **CO No.** | **Course Outcome** | **Bloom's Taxonomy Level** |
| --- | --- | --- |
| **CO1** | Understand the syntax, variables, data types, and control structures in C programming. | Understand (Level 2) |
| **CO2** | Apply functions and arrays for statistical computations. | Apply (Level 3) |
| **CO3** | Analyze data structures like linked lists, stacks, and queues for statistical data management. | Analyze (Level 4) |
| **CO4** | Evaluate statistical algorithms such as regression and correlation using C programming. | Evaluate (Level 5) |
| **CO5** | Create optimized C programs for statistical analysis and effective data presentation. | Create (Level 6) |

### ****Program Outcomes (POs)****

| **PO.No.** | **Program Outcome** | **Bloom's Taxonomy Level** |
| --- | --- | --- |
| **PO1** | Apply knowledge of programming concepts for solving statistical problems. | Apply (Level 3) |
| **PO2** | Identify and define statistical problems and solve them using C programming. | Understand (Level 2) |
| **PO3** | Design algorithms for statistical data analysis. | Create (Level 6) |
| **PO4** | Implement statistical algorithms using C for real-time problems. | Apply (Level 3) |
| **PO5** | Analyze the performance and optimization of statistical programs. | Analyze (Level 4) |
| **PO6** | Develop solutions for statistical challenges using programming tools. | Create (Level 6) |
| **PO7** | Evaluate statistical program outputs for accuracy and efficiency. | Evaluate (Level 5) |
| **PO8** | Communicate statistical findings effectively using structured program outputs. | Apply (Level 3) |

### ****Program-Specific Outcomes (PSOs)****

| **PSO No.** | **Program-Specific Outcome** | **Bloom's Taxonomy Level** |
| --- | --- | --- |
| **PSO1** | Develop efficient statistical models using programming concepts. | Create (Level 6) |
| **PSO2** | Apply statistical algorithms for data analysis and problem-solving. | Apply (Level 3) |
| **PSO3** | Analyze statistical data outputs and interpret results effectively. | Analyze (Level 4) |

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –III (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 8: STATISTICAL PROGRAMMING WITH ‘C’ LANGUAGE AND DATA STRUCTURES**

**No. of Hours/week: 03 Course Code: 24 – STA – 3C**8 **Credits – 3**

**II. Syllabus**

**UNIT *–* I**

**Introduction to C Programming for Statistical Applications**

**Basics of C Programming**: Understanding C syntax, variables, data types, operators, and control structures (if-else, loops). **Input and Output**: Use of scanf, printf for data input and output, formatted I/O for statistical data. **Memory Allocation**: Understanding variables, memory allocation, pointers, and their applications in data manipulation. **Debugging and Errors**: Common errors in C programming and debugging techniques.

**UNIT *–* II**

**Arrays, Functions, and Statistical Computations**

**Arrays and Matrices**: Creating and manipulating arrays, matrix representation, and operations using arrays. **Functions in C**: Creating user-defined functions, passing arguments, return values, and using functions to modularize code. **Statistical Functions**: Writing ‘C’ functions for mean, median, mode, variance, and standard deviation. **Use of Libraries**: Introduction to C basic libraries and math for advanced computations –Incorporate Unit I instead of Unit –II .Without knowing ‘C’ Basic Library functions the student cannot develop the ‘C’ program. In the same unit we can include advanced library functions. The student can understand in a better way.

**UNIT *–* III**

**Data Structures for Statistical Data Management**

**Introduction to Data Structures**: Understanding the need for data structures in managing statistical data, Types of Data Structures, Operations performed on Data Structures, Arrays. **Linked Lists**: Implementing singly and doubly linked lists, use cases in statistical data storage. **Stacks and Queues**: Concepts, implementation, and applications in sequential data handling. **Trees and Binary Trees**: Basics of binary trees, applications in data indexing, and hierarchical data organization.

**UNIT *–* I**V

**Statistical Algorithms and Sorting Techniques**

**Sorting Algorithms**: Implementation of bubble sort, selection sort, insertion sort, and their relevance in data organization. **Searching Algorithms**: Linear and binary search for data retrieval, applications in large datasets. **Random Number Generation**: Generating random samples in C, applying randomization in statistical sampling. **Statistical Algorithms**: Implementing regression analysis and correlation computations in C.

**UNIT *–* V**

**Application of C Programming in Statistical Analysis**

**Data Handling and Optimization**: Techniques for managing large datasets in C, optimizing code for speed and memory efficiency. **Collaboration in Programming**: Version control practices and documentation. **Effective Presentation of Results**: Structuring output for clarity and precision, presenting statistical findings through program outputs. Practical Applications: Writing C programs to demonstrate statistical concepts such as regression, correlation, and sampling.

**24-STA-3C8P Practicals for Course – VIII**  **Credits 2**

**UNIT *–* I**

**Introduction to C Programming for Statistical Applications**

1. **Basic Syntax and Control Structures**: Write a C program that uses if-else statements to classify a given dataset of numbers as "positive," "negative," or "zero."
2. **Variable Declaration and Data Types**: Create a C program that takes input from the user for a dataset of integers and stores them in an array. Display the sum and average of the dataset.
3. **Input and Output Operations**: Develop a C program that uses scanf and printf to take input of 10 integers from the user, and then display them in a formatted table (including column headers).
4. **Memory Allocation**: Write a C program that dynamically allocates memory for an array of 100 integers, initializes the array with random values, and prints the array.
5. **Debugging Techniques**: Implement a C program with a deliberate error (e.g., division by zero) and debug it to handle errors gracefully (e.g., by displaying an error message).

**UNIT *–* II**

**Arrays, Functions, and Statistical Computations**

1. **Array Manipulation**: Write a C program that stores a dataset of numbers in an array and finds the highest and lowest values using array manipulation techniques.
2. **Matrix Operations**: Develop a C program that accepts two matrices and performs matrix addition, subtraction, or multiplication based on user input.
3. **User-defined Functions**: Create a C function to compute the mean and standard deviation for a given dataset. Implement the function and call it from the main program.
4. **Statistical Functions**: Write a C program with functions for calculating the median, mode, and variance of a dataset of numbers. Ensure that all functions are modular and reusable.
5. **Using Libraries**: Implement a C program that uses the math library (math.h) to calculate the square root, power, and logarithms of given statistical values.

**UNIT *–* III**

**Data Structures for Statistical Data Management**

1. **Linked List Operations**: Write a C program that implements a singly linked list to store statistical data (e.g., a list of student grades). Implement insert and delete functions for manipulating the list.
2. **Stack Implementation**: Develop a C program that implements a stack data structure to manage a sequence of statistical calculations, allowing for "undo" functionality.
3. **Queue Implementation**: Create a C program that implements a queue to store datasets in a FIFO (First In, First Out) manner and performs simple operations like enqueue and dequeue.
4. **Binary Tree Implementation**: Write a C program to build a binary search tree (BST) to store statistical data (e.g., ordered values) and implement traversal methods (in-order, pre-order, post-order).
5. **Linked List for Sorting**: Implement a C program using a doubly linked list to sort a dataset of integers in ascending order using the insertion sort algorithm.

**UNIT *–* IV**

**Statistical Algorithms and Sorting Techniques**

1. **Bubble Sort Algorithm**: Write a C program to implement the bubble sort algorithm on a dataset and display the sorted dataset.
2. **Selection Sort Algorithm**: Implement the selection sort algorithm in C to arrange a set of numbers in descending order and print the result.
3. **Binary Search Algorithm**: Write a C program to perform a binary search on a sorted dataset and find the index of a given value.
4. **Regression Analysis**: Develop a C program to implement simple linear regression and calculate the regression line equation for a dataset.
5. **Random Sampling**: Create a C program that generates a specified number of random values from a population and calculates basic descriptive statistics (mean, variance).

 **Reference books and text books**

**Introduction to C Programming for Statistical Applications**

1. **"Programming in C"** by **E. Balagurusamy**
	* A comprehensive textbook that covers the basics of C programming, including syntax, operators, control structures, and data types.
2. **"Let Us C"** by **Yashavant Kanetkar**
	* An excellent introductory book for C programming, explaining the fundamentals with simple examples and exercises.

**Arrays, Functions, and Statistical Computations**

1. **"C Programming: A Problem-Solving Approach"** by **R. S. Salaria**
	* Focuses on problem-solving using C programming, covering arrays, functions, and statistics-related computations in C.
2. **"Programming with C"** by **B. A. Forouzan and Richard F. Gilberg**, translated by **S. S. R. Anjaneyulu**
	* A widely recommended book on C programming, providing clear explanations of arrays, functions, and advanced topics like memory management.

**Data Structures for Statistical Data Management**

1. **"Data Structures Using C"** by **Reema Thareja**
	* A book focusing on data structures in C, explaining linked lists, stacks, queues, and trees with practical examples.
2. **"Data Structures and Algorithms in C"** by **A. K. Sharma**
	* A practical guide that covers various data structures like arrays, linked lists, stacks, and queues, with emphasis on statistical data management.

**Statistical Algorithms and Sorting Techniques**

1. **"Data Structures and Algorithms in C"** by **Y. Langsam, M. J. Augenstein, A. M. Tenenbaum**, translated by **V. Rajaraman**
	* This book provides a solid foundation in C programming, focusing on algorithms, sorting techniques, and their applications in statistical analysis.
2. **"Algorithms in C"** by **Robert Sedgewick**, adapted for Indian curriculum by **P. B. S. J. M. Rajan**
	* An in-depth reference on implementing various algorithms, including sorting, searching, and statistical algorithms, specifically in C programming.

**Project and Application of C Programming in Statistical Analysis**

1. **"Practical C Programming"** by **K. N. King**, adapted by **M. D. Chidambaram**
	* Provides practical approaches to C programming, with a focus on project development and implementing statistical analysis in real-world scenarios.
2. **"C Programming and Data Structures"** by **B. K. P. Subrahmanyam**
* A practical guide to C programming that also emphasizes implementing data structures and solving real-world statistical problems with C.

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER: III (CBCS, w.e.f 2024-25)**

**COURSE TITLE 8: STATISTICAL PROGRAMMING WITH ‘C’ LANGUAGE AND DATA STRUCTURES**

**Time: 3 Hours Subject Code: 24-STA-3C8 Max. Marks: 75**

**SECTION – A**

Answer any **ALL** questions. Each question carry **TWO** marks **5 X 2 = 10M**

1. Define pointers in C and mention their importance in data manipulation.
2. What is the purpose of scanf and printf in C programming?
3. Write syntax for creating a user-defined function in C.
4. What is the difference between linear and binary search?
5. Mention two key optimization techniques for handling large datasets in C.

**SECTION – B**

Answer any **FIVE** questions. Each question carry **FIVE** marks **5 X 5 = 25M**

1. Explain control structures in C with examples.
2. Write a program to calculate the mean of an array of numbers in C.
3. Describe the concept of singly linked lists with an example.
4. Differentiate between stacks and queues with their applications in statistical data handling.
5. Write a short note on bubble sort with an example.
6. Explain random number generation in C and its role in statistical sampling.
7. Describe the role of collaboration tools in C programming.
8. Write a C program to demonstrate correlation computation.

**SECTION – C**

Answer any **FOUR** questions and each question carries **10** marks **4 X 10 = 40M**

1. Explain in detail the process of memory allocation in C with suitable examples.
2. Write a C program to calculate mean, median, and mode of a dataset using functions.
3. Discuss the implementation of doubly linked lists and their applications in statistical data storage.
4. Write a C program for implementing insertion sort and explain its working.
5. Explain regression analysis in statistical computations and write a C program to demonstrate it.
6. Describe data handling and optimization techniques used for managing large statistical datasets in C.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –IV (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 9: SAMPLING TECHNIQUES**

**No. of Hours/week: 03 Course Code: 24 – STA – 4C9 Credits – 3**

**Learning Objectives and Outcomes for Sampling Theory and Methods**

**Learning Objectives (LOs):**

1. Understand the basic concepts of sampling theory, including the distinction between parameters and statistics.
2. Learn about various sampling methods, including probability and non-probability sampling, and their applications.
3. Understand the importance of sampling distribution and its role in survey design and statistical analysis.
4. Gain knowledge of the different sampling techniques such as Simple Random Sampling (SRS), Stratified Sampling, and Systematic Sampling.
5. Learn how to calculate population parameters (mean, variance, total, proportion) using sampling methods.
6. Understand the structure and role of national statistical organizations and their importance in data collection and reporting.

**Learning Outcomes (LOs):**

1. Demonstrate a clear understanding of the differences between parameters and statistics, and explain their significance in sampling.
2. Apply various sampling techniques in real-world survey situations and analyze their advantages and disadvantages.
3. Calculate population parameters such as mean, variance, and total using different sampling methods.
4. Evaluate the effectiveness and efficiency of different sampling methods and make decisions on the appropriate sampling technique.
5. Interpret the role of national statistical organizations and their importance in the collection, analysis, and reporting of statistical data.

**Course Outcomes (COs):**

| **Course Outcomes (COs)** | **Bloom’s Taxonomy Level** |
| --- | --- |
| **CO1**: Understand the concepts of sampling distributions, parameters, statistics, and their differences. | **Remembering** |
| **CO2**: Apply different sampling methods such as simple random, stratified, and systematic sampling in real-world scenarios. | **Applying** |
| **CO3**: Estimate population parameters such as mean, variance, and total using various sampling techniques. | **Applying** |
| **CO4**: Evaluate the efficiency of different sampling methods and their suitability for different survey situations. | **Evaluating** |
| **CO5**: Understand the role of national statistical organizations in conducting surveys, data analysis, and policy-making. | **Understanding** |

**Program Outcomes (POs):**

1. **PO1**: Develop a strong foundation in statistical theory and apply it to solve real-world problems.
2. **PO2**: Understand and apply statistical methods to draw valid conclusions from data.
3. **PO3**: Demonstrate analytical and critical thinking in applying various statistical techniques.
4. **PO4**: Use modern statistical tools and software to analyze data effectively.
5. **PO5**: Communicate statistical findings and analyses clearly and effectively.
6. **PO6**: Work collaboratively in multidisciplinary teams to apply statistical knowledge in real-world contexts.
7. **PO7**: Interpret and analyze complex statistical data to support decision-making processes.
8. **PO8**: Engage in continuous learning and development in statistical methods and techniques.

**Program-Specific Outcomes (PSOs):**

1. **PSO1**: Apply advanced sampling and statistical techniques to interpret large datasets and solve complex problems.
2. **PSO2**: Design and conduct surveys using various sampling methods to collect accurate and reliable data.
3. **PSO3**: Contribute to the development and implementation of statistical surveys and data analysis projects in both academic and professional settings.

**Mapping Table: COs, POs, PSOs with Bloom’s Taxonomy Connectivity**

| **Course Outcomes (COs)** | **Program Outcomes (POs)** | **Program Specific Outcomes (PSOs)** | **Bloom’s Taxonomy** |
| --- | --- | --- | --- |
| **CO1**: Understand the concepts of sampling distributions, parameters, statistics, and their differences. | **PO1, PO2**: Apply statistical knowledge to understand concepts and solve problems. | **PSO1**: Use sampling methods to interpret data and solve real-world problems. | **Remembering** |
| **CO2**: Apply different sampling methods such as simple random, stratified, and systematic sampling. | **PO3, PO4**: Develop analytical skills and apply statistical techniques to design surveys and solve problems. | **PSO2**: Design surveys and apply appropriate sampling methods in research. | **Applying** |
| **CO3**: Estimate population parameters using various sampling techniques. | **PO1, PO7**: Use statistical methods to make inferences and estimate population parameters. | **PSO3**: Use sampling methods to estimate parameters in practical scenarios. | **Applying** |
| **CO4**: Evaluate the efficiency of different sampling methods. | **PO5, PO8**: Evaluate the suitability of different statistical methods and communicate the findings. | **PSO2**: Evaluate the application of various sampling methods in surveys. | **Evaluating** |
| **CO5**: Understand the role of national statistical organizations in surveys and data analysis. | **PO6, PO7**: Understand the role of statistical organizations and communicate their impact. | **PSO1, PSO3**: Contribute to statistical organizations through data collection and analysis. | **Understanding** |

**Summary of Bloom’s Taxonomy Mapping:**

* **Remembering (CO1)**: The outcome involves recalling fundamental concepts such as parameters, statistics, and their differences.
* **Applying (CO2, CO3)**: These outcomes focus on applying sampling methods (SRS, Stratified, Systematic) and estimating population parameters in practical scenarios.
* **Evaluating (CO4)**: In this outcome, students assess the effectiveness of different sampling methods, evaluating their efficiency and suitability for specific problems.
* **Understanding (CO5)**: This outcome focuses on understanding the role of national statistical organizations and their impact on statistical data collection and reporting.

This approach helps students progress through various levels of Bloom's Taxonomy, from basic recall to the application of knowledge and evaluation of sampling methods, while fostering a comprehensive understanding of the subject.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –IV (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**Course - 9: SAMPLING TECHNIQUES**

**No. of Hours/week: 03 Course Code: 24 – STA – 4C9 Credits – 3**

**II. Syllabus**

**UNIT –** I

**Sampling Theory and Methods**

**Review of Parameter and Statistic:** Definition and Differences Between Parameter and Statistic, Sampling Distribution: Concept and Importance. **Sampling in Surveys:** Principal Steps and Principles in a Sample Survey, Sampling vs Non-Sampling Errors: Definition and Examples, Advantages of Sampling Over Census: Time, Cost, and Efficiency, Limitations of Sampling: Errors and Biases, **Types of Sampling: Subjective Sampling**: Definition and Applications, **Probability Sampling**: Concept and Types (Simple Random Sampling, Stratified Sampling, Systematic Sampling, etc.), **Mixed Sampling**: Combination of Probability and Non-Probability Sampling Methods.

**UNIT – II**

**Simple Random Sampling and Estimation**

**Notations and Terminology:** Key Terms in Sampling: Population, Sample, Sampling Frame, Sample Size, etc., Various Probabilities of Selection: Definition and Calculation, **Random Numbers and Their Uses:** Concept and Use of Random Numbers, Random Number Tables: Construction and Application, **Methods of Selecting Simple Random Samples: Lottery Method**: Explanation and Procedure, **Method Based on Random Numbers**: Step-by-step Process and Applications, **Estimation of Population Parameters: Estimates of Population Total**: Formula and Calculation, **Estimates of Population Mean**: Formula and Calculation, **Variance and Standard Errors**: Estimation Methods for Variance and Standard Error. **Determination of Sample Size:** Calculation of Sample Size for Estimation of Mean and Total, **Simple Random Sampling of Attributes:** Definition and Methodology of Sampling for Attributes, Estimation of Population Proportions and Variance for Attributes.

**UNIT – III**

**Stratified Random Sampling**

**Stratified Random Sampling:** Concept and Procedure of Stratified Random Sampling, Importance of Stratification in Sampling. **Advantages and Disadvantages of Stratified Random Sampling:** Advantages: Improved Precision, Reduced Variance, Disadvantages: Complexity in Implementation, Need for Detailed Information. **Estimation in Stratified Random Sampling: Estimation of Population Mean**: Formula and Calculation for Stratified Sampling: **Estimation of Population Variance**: Formula and Calculation for Stratified Sampling: **Allocations in Stratified Random Sampling :Proportional Allocation**: Concept, Formula, and Application: **Optimum Allocation**: Definition, Formula, and Application: **Comparison Between Proportional and Optimum Allocations:** Comparison with Simple Random Sampling Without Replacement (SRSWOR), Advantages and Disadvantages of Each Allocation Method, Practical Examples of Stratified Sampling with Different Allocations

**UNIT – IV**

**Systematic sampling**

**Systematic Sampling:** Definition and Procedure of Systematic Sampling when N=nk : Merits of Systematic Sampling: Simplicity, Cost-Effectiveness, Demerits of Systematic Sampling: Potential Bias, Requires Ordered Population, **Estimation in Systematic Sampling: Estimation of Population Mean**: Formula and Calculation for Systematic Sampling, **Estimation of Population Variance**: Formula and Calculation, **Comparison of Systematic Sampling with Other Sampling Methods: Comparison with Stratified Sampling**: Differences in Precision and Cost, **Comparison with Simple Random Sampling Without Replacement (SRSWOR)**: Efficiency and Application, **Variance Comparison for Different Sampling Methods, Comparison of Variance for Simple Random Sampling (SRS), Stratified Random Sampling (St.R.S), and Systematic Sampling (SYS):** Analysis for a Linear Trend in the Population

**UNIT – V**

**National Statistical Organizations**

**National Statistical Organization: Vision and Mission**: Overview and Goals of the National Statistical Organization, **National Sample Survey Office (NSSO)**: Structure, Roles, and Responsibilities, **Central Statistical Organization (CSO)**: Structure, Roles, and Responsibilities, **Important Activities**: Surveys, Data Collection, Analysis, and Reporting, **Publications**: Key Statistical Reports and Publications of NSSO and CSO. **National Statistical Commission (NSC): Need for NSC**: Importance of an Independent Statistical Commission, **Constitution of NSC**: Structure, Members, and Appointment Process, **Role and Functions of NSC**: Advisory, Regulatory, and Policy Functions, **Important Acts**: Relevant Legal Frameworks for Statistical Activities in India

**Practical Syllabus**  Practical Credits: 1 2 hrs/week

1. Show the sample mean is unbiased estimator of population mean in SRSWOR and also find variance of sample mean.

2. Show the sample mean square is unbiased estimator of population mean square in SRSWOR.

3. Show the sample mean is unbiased estimator of population mean in SRSWR and also find variance of sample mean.

4. Compare means and variances between SRSWR and SRSWOR.

5. Allocation of sample sizes to various strata in proportional and in optimum allocations to draw a Stratified random sample.

6. Compare precision in proportional and optimum allocations with SRSWOR and gain in efficiency due to proportional and optimum allocations.

7. Systematic sampling with N = nk and Compare the precision of an estimate in systematic sampling with that of in Stratified and in SRSWOR.

**Note:** Training shall be on establishing formulae in Excel cells and derive the results. The excel output shall be exported to MS word for writing inference.

**III. References**

1. S. C. Gupta & V. K. Kapoor: Fundamentals of Applied Statistics, Sultan Chand& Sons, New Delhi.

2. K.V.S. Sarma: Statistics Made Simple: Do it yourself on PC. PHI.

3. M. R. Saluja: Indian Official Statistics. ISI publications.

**IV. Suggested Co-curricular Activities:**

1. Training of students by related industrial experts

2. Assignments including technical assignments if any.

3. Seminars, Group Discussions, Quiz, Debates etc on related topics.

4. Preparation of audio and videos on tools of diagrammatic and graphical representations.

5. Collection of material/figures/photos/author photoes of related topics.

6. Invited lectures and presentations of stalwarts to those topics.

7. Visits/field trips of firms, research organizations etc.

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER: IV (CBCS, w.e.f 2024-25)**

**COURSE TITLE: SAMPLING TECHNIQUES**

**Time: 3 Hours Subject Code: 24-STA-4C9 Max. Marks: 75**

**SECTION – A**

Answer any **ALL** questions. Each question carry **TWO** marks **5 X 2 = 10M**

1. Define parameter and a statistic.
2. What is Simple Random Sampling (SRS)?
3. Define Stratified Random Sampling?
4. What is Systematic Sampling?
5. Define NSSO and CSO?

**SECTION – B**

Answer any **FIVE** questions. Each question carry **FIVE** marks **5 X 5 = 25M**

1. What are the principal steps in a sample survey?
2. Explain about Sampling and Non-Sampling errors?
3. Define and explain Simple Random Sampling (SRS) with and without replacement?
4. Describe the lottery method of selecting a sample in SRS**.**
5. Find the Variance of sample mean in Stratified random sampling.
6. What are the advantages and disadvantages of Stratified Random Sampling?
7. What are the merits and demerits of Systematic Sampling?
8. Explain the functions of the National Statistical Commission (NSC).

**SECTION – C**

 Answer any **FOUR** questions and each question carries **10** marks **4 X 10 = 40M**

1. What are the advantages and limitations of sampling over census?
2. How do you estimate the population mean and variance using SRS?
3. In SRS show that sample mean is unbiased estimator of population mean
4. What are the differences between proportional and optimum allocation in Stratified Random Sampling?
5. How do you estimate the population mean and variance using Systematic Sampling?
6. What are the functions of NSSO and CSO? Explain in detail.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –IV (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 10: DESIGN AND ANALYSIS OF EXPERIMENTS**

**No. of** Hours/week: 03 **Course Code: 24 – STA – 4C10 C**redits **–** 3

**Learning Objectives and Outcomes for Experimental Designs**

**Learning Objectives (LOs):**

1. Understand the basic concepts of analysis of variance (ANOVA) and experimental designs.
2. Learn the formulation, assumptions, and applications of One-Way and Two-Way ANOVA.
3. Understand the principles, advantages, and limitations of various experimental designs such as CRD, RBD, LSD, and factorial experiments.
4. Gain practical experience in designing experiments and analyzing data using ANOVA techniques.
5. Develop skills in statistical analysis for different experimental designs, including solving practical problems.

**Learning Outcomes (LOs):**

1. Demonstrate an understanding of the key concepts of ANOVA and experimental design, including the purpose, assumptions, and steps involved.
2. Apply One-Way and Two-Way ANOVA techniques to analyze data and interpret results.
3. Design and implement Completely Randomized Designs (CRD), Randomized Block Designs (RBD), and Latin Square Designs (LSD) in experiments.
4. Analyze and interpret experimental data using ANOVA for CRD, RBD, and LSD designs.
5. Conduct factorial experiments, including 2² and 2³ designs, and interpret main and interaction effects.

**Course Outcomes (COs):**

| **Course Outcomes (COs)** | **Bloom’s Taxonomy Level** |
| --- | --- |
| **CO1**: Understand and apply the concept of ANOVA, including One-Way and Two-Way classifications. | **Understanding** |
| **CO2**: Design and implement experimental designs such as CRD, RBD, LSD, and factorial experiments. | **Applying** |
| **CO3**: Perform statistical analysis of experimental data using ANOVA for various designs. | **Applying** |
| **CO4**: Analyze the efficiency and advantages of different experimental designs in specific research contexts. | **Evaluating** |
| **CO5**: Solve problems related to missing values in experimental designs and interpret results. | **Applying** |

**Program Outcomes (POs):**

1. **PO1**: Develop a strong foundation in statistical methods for experimental design and data analysis.
2. **PO2**: Apply statistical techniques to design and conduct experiments and analyze data effectively.
3. **PO3**: Analyze and interpret data using advanced statistical tools and methods.
4. **PO4**: Use modern software tools to implement and analyze experimental designs.
5. **PO5**: Communicate statistical findings clearly, both in written reports and oral presentations.
6. **PO6**: Work effectively in multidisciplinary teams to design and conduct experiments.
7. **PO7**: Apply experimental design techniques to solve real-world research problems.
8. **PO8**: Engage in lifelong learning and professional development in statistical analysis and experimental design.

**Program-Specific Outcomes (PSOs):**

1. **PSO1**: Design and conduct statistical experiments to collect and analyze data using various experimental designs.
2. **PSO2**: Interpret and evaluate results from ANOVA and other statistical methods to make informed decisions.
3. **PSO3**: Apply factorial designs and other complex statistical models to analyze interactions and effects in research experiments.

**Mapping Table: COs, POs, PSOs with Bloom’s Taxonomy Connectivity**

| **Course Outcomes (COs)** | **Program Outcomes (POs)** | **Program-Specific Outcomes (PSOs)** | **Bloom’s Taxonomy** |
| --- | --- | --- | --- |
| **CO1**: Understand and apply the concept of ANOVA, including One-Way and Two-Way classifications. | **PO1, PO2**: Apply statistical methods and techniques to analyze data using ANOVA. | **PSO1**: Apply ANOVA techniques in experimental research to interpret data. | **Understanding** |
| **CO2**: Design and implement experimental designs such as CRD, RBD, LSD, and factorial experiments. | **PO3, PO4**: Use statistical tools to design and conduct experiments and analyze data. | **PSO2**: Design and apply experimental designs in research projects. | **Applying** |
| **CO3**: Perform statistical analysis of experimental data using ANOVA for various designs. | **PO5, PO6**: Communicate findings from statistical analyses and collaborate in teams for design and analysis. | **PSO3**: Analyze experimental data using ANOVA and interpret interaction effects. | **Applying** |
| **CO4**: Analyze the efficiency and advantages of different experimental designs in specific research contexts. | **PO7, PO8**: Evaluate the suitability of different experimental designs for specific contexts. | **PSO2**: Evaluate experimental designs for their efficiency and effectiveness in specific research scenarios. | **Evaluating** |
| **CO5**: Solve problems related to missing values in experimental designs and interpret results. | **PO3, PO8**: Solve real-world problems by applying appropriate statistical techniques. | **PSO3**: Address missing values and analyze their impact on experimental results. | **Applying** |

**Summary of Bloom’s Taxonomy Mapping:**

* **Understanding (CO1)**: The first outcome involves grasping the basic concepts of ANOVA, including One-Way and Two-Way classifications, and their assumptions and applications.
* **Applying (CO2, CO3, CO5)**: These outcomes focus on applying experimental design principles (CRD, RBD, LSD, factorial) and performing statistical analyses on experimental data, including handling missing values.
* **Evaluating (CO4)**: The fourth outcome encourages students to evaluate the efficiency and appropriateness of different experimental designs in various research contexts.

This mapping demonstrates the alignment of the course content with both Bloom's Taxonomy and the program’s educational objectives, helping students progress from foundational knowledge to practical application and critical evaluation of statistical methods.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –IV (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 10: DESIGN AND ANALYSIS OF EXPERIMENTS**

**No. of** Hours/week: 03 **Course Code: 24 – STA – 4C10 C**redits **–** 3

**II. Syllabus**

**UNIT – I**

**Analysis of Variance (ANOVA)**

**Concept and Definition of ANOVA:** Overview of ANOVA: Purpose and Application, Assumptions of ANOVA: Normality, Homogeneity of Variance, Independence, **One-Way Classification ANOVA: Mathematical Model**: Formulation for One-Way ANOVA, **Analysis of Variance**: Steps and Procedure for Equal and Unequal Classifications, **Problems**: Practical Applications and Problem-Solving. **Two-Way Classification ANOVA: Mathematical Model**: Formulation for Two-Way ANOVA with and without Interaction, **Analysis of Variance**: Steps for Two-Way Classification and Interpretation of Results, **Problems**: Practical Examples and Application of Two-Way ANOVA

**UNIT – II**

**Completely Randomized Design (CRD)**

**Introduction to Experimental Designs: Definition and Terminology**: Key Concepts in Design of Experiments, **Principles of Design of Experiments**: Randomization, Replication, and Control, **Completely Randomized Design (CRD): Concept of CRD**: Overview and Application of CRD in Experiments, **Advantages of CRD**: Simplicity, Flexibility, and Ease of Implementation, **Disadvantages of CRD**: Assumption of Homogeneity, Limited to Simple Experimental Settings, **Applications of CRD**: Use Cases in Agricultural, Industrial, and Social Sciences, **Layout of CRD: Design Layout**: Structure and Setup of CRD Experiments, **Statistical Analysis**: Steps in Analyzing CRD Data, Analysis of Variance (ANOVA).

**UNIT – III**

**Randomized Block Design (RBD)**

**Introduction to Randomized Block Design (RBD): Concept of RBD**: Overview and Explanation of Randomized Block Design, **Advantages of RBD**: Control of Variability, Increased Precision, and Improved Estimates, **Disadvantages of RBD**: Complexity in Layout, Limited to Homogeneous Blocks, **Applications of RBD**: Use in Agricultural, Industrial, and Social Research, **Layout of RBD, Design Layout**: Structure and Arrangement of Experimental Units in Blocks, **Statistical Analysis**: Steps for Analyzing RBD Data using Analysis of Variance (ANOVA), **Efficiency of RBD Relative to CRD, RBD with One Missing Value: Analysis of RBD with One Missing Value**: Treatment of Missing Data and Statistical Adjustments, **Practical Problems**: Solving Problems Involving Missing Values in RBD

**UNIT – IV**

**Latin Square Design (LSD)**

**Introduction to Latin Square Design:** Concept of Latin Square Design, Advantages and disadvantages, Applications of LSD, **Layout of Latin Square Design:** Structure and arrangement of treatments in LSD. **Statistical Analysis of LSD:** Steps for statistical analysis of LSD, Calculation of critical differences, **Efficiency of LSD:** Efficiency of LSD compared to Randomized Block Design (RBD), Efficiency of LSD compared to Completely Randomized Design (CRD). **Estimation of Missing Value:** Estimation of one missing value in LSD, Analysis of LSD with a missing value

**UNIT – V**

**Factorial Experiments**

**Introduction to Factorial Experiments:** Concept of factorial experiments, Importance of studying main effects and interaction effects. **Analysis of 2 2 Factorial Experiments**, Main effects in **2 2**factorial experiments, Interaction effects in **2 2**factorial experiments, Statistical analysis of **2 2**factorial experiments. **Analysis of 2 3 Factorial Experiments:** Main effects in **2 3** factorial experiments, Interaction effects in **2 3** factorial experiments, Statistical analysis of **2 3** factorial experiments. **Yates' Procedure:** Step-by-step process to compute factorial effect totals using Yates' procedure, Application of Yates' procedure in factorial designs. **Problem-Solving:** Examples and practice problems for **2 2**and **2 3**factorial experiments, Interpretation of results from statistical analyses

 **Practical Syllabus**  Practical Credits: 1 2 hrs/week

1. ANOVA - one - way classification with equal number of observations.

2. ANOVA - one - way classification with unequal number of observations.

3. ANOVA Two-way classification.

4. Analysis of CRD and critical differences.

5. Analysis of RBD and critical differences. Relative efficiency of CRD with RBD.

6. Estimation of single missing observation in RBD and its analysis.

7. Analysis of LSD and efficiency of LSD over CRD and RBD.

8. Estimation of single missing observation in LSD and its analysis.

9. Analysis of 22 with RBD layout.

10. Analysis of 23 with RBD layout.

**Note:** Training shall be on establishing formulae in Excel cells and derive the results. The excel output shall be exported to MS word for writing inference.

**I. References**

1. S. C. Gupta & V. K. Kapoor: Fundamentals of Applied Statistics, Sultan Chand& Sons, New Delhi.

2. K.V.S. Sarma: Statistics Made Simple: Do it yourself on PC. PHI.

3. M. R. Saluja: Indian Official Statistics. ISI publications.

**II. Suggested Co-curricular Activities:**

1. Training of students by related industrial experts

2. Assignments including technical assignments if any.

3. Seminars, Group Discussions, Quiz, Debates etc on related topics.

4. Preparation of audio and videos on tools of diagrammatic and graphical representations.

5. Collection of material/figures/photos/author photoes of related topics.

6. Invited lectures and presentations of stalwarts to those topics.

7. Visits/field trips of firms, research organizations etc.

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (WM)**

**SEMESTER: IV (CBCS, w.e.f 2024-25)**

**Course - 10: DESIGN AND ANALYSIS OF EXPERIMENTS**

**Time: 3 hrs Course Code: 24 – STA – 4C10 Max. Marks: 75**

**SECTION – A**

Answer any **FIVE** questions. Each question carries ‘**2’** marks. **5 X 2 = 10M**

* + - 1. Define ANOVA.
			2. What are the main advantages of CRD?
			3. What is a Randomized Block Design (RBD)?
			4. What is a Latin Square Design (LSD)?
			5. What is a factorial design?

**SECTION – B**

Answer any **FIVE** questions. Each question carries ‘**5’** marks. **5 X 5 = 25 M**

1. What is “Analysis of Variance” and where it is used? And also give its assumptions.
2. Explain the statistical analysis of CRD.
3. Derive the expression to estimate the single missing plot in R.B.D.
4. Explain the Layout of RBD in brief.
5. Explain the statistical analysis of LSD.
6. Explain the Layout of LSD in brief.
7. Define i) Main Effects ii) Interaction Effects.
8. Explain the Yates procedure to find factorial effect totals.

**SECTION – C**

Answer any **FOUR** questions. Each question carries **10** marks. **4 X 10 = 40 M**

1. Explain the ANOVA of two way classification.
2. Explain the statistical analysis of RBD.
3. Explain the missing plot technique of Latin square design.
4. Derive the efficiency of LSD over RBD.
5. Explain statistical analysis of 22 - factorial design using RBD.
6. Explain Statistical analysis of 23 factorial design.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –IV (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**Course - 11: FOUNDATIONS OF STATISTICAL DATA ANALYSIS USING ‘R’**

**No. of** Hours/week: 03 **Course Code: 24 – STA – 4C1**1 **C**redits **–** 3

**Learning Objectives (LOs):**

1. **Understand the fundamentals of statistical data analysis** using R programming.
2. **Develop proficiency in importing, cleaning, and transforming data** for statistical analysis in R.
3. **Apply descriptive and inferential statistical techniques** using R for analysis and interpretation of real-world data.
4. **Gain expertise in data visualization** using R, including the use of advanced visualization libraries like ggplot2.
5. **Understand and apply machine learning algorithms** for statistical modeling and prediction in R.

**Learning Outcomes (LOs):**

Upon completion of the course, students should be able to:

1. **LO1:** Describe key concepts in statistical data analysis and demonstrate proficiency in handling data using R.
2. **LO2:** Import, clean, and manipulate data from various sources (CSV, Excel, etc.) using R.
3. **LO3:** Perform descriptive and inferential statistical analyses in R, such as calculating measures of central tendency, dispersion, hypothesis testing, and regression analysis.
4. **LO4:** Visualize data effectively using ggplot2 and interpret graphical outputs to derive meaningful insights.
5. **LO5:** Develop and implement basic machine learning algorithms for predictive modeling and evaluation in R.

**Course Outcomes (COs):**

1. **CO1:** Understand the concepts and applications of statistical data analysis and R programming.
2. **CO2:** Import and manipulate data from different formats (CSV, Excel) and perform data cleaning and transformation tasks.
3. **CO3:** Analyze datasets using descriptive statistics and inferential tests (t-tests, ANOVA, regression) in R.
4. **CO4:** Use ggplot2 to create professional-quality data visualizations and interpret statistical results graphically.
5. **CO5:** Apply machine learning algorithms in R for classification, clustering, and regression tasks, including model evaluation.

**Program Outcomes (POs):**

1. **PO1:** Demonstrate knowledge of statistical methods and their application in real-world problems.
2. **PO2:** Use software tools (such as R) for data analysis, statistical computations, and visualization.
3. **PO3:** Work effectively with datasets of different formats and perform cleaning, transformation, and analysis.
4. **PO4:** Critically analyze the results of statistical methods and provide insightful interpretations.
5. **PO5:** Apply descriptive, inferential, and predictive analysis to solve problems in various domains (e.g., business, health, education).
6. **PO6:** Communicate data insights through visualizations, reports, and presentations.
7. **PO7:** Work collaboratively in teams to solve data analysis problems using R.
8. **PO8:** Demonstrate the ability to apply machine learning techniques for predictive analytics and model evaluation.

**Program Specific Outcomes (PSOs):**

1. **PSO1:** Apply statistical and machine learning techniques to solve complex problems in data science and analytics.
2. **PSO2:** Demonstrate proficiency in using R for data wrangling, visualization, and analysis in various domains such as business, healthcare, and research.
3. **PSO3:** Use R to build and evaluate predictive models for real-world applications, ensuring model accuracy and reliability.

**Mapping Summary Table:**

| **Learning Outcomes (LOs)** | **Program Outcomes (POs)** | **Program Specific Outcomes (PSOs)** | **Course Outcomes (COs)** | **Bloom's Taxonomy Level** |
| --- | --- | --- | --- | --- |
| **LO1:** Describe key concepts in statistical data analysis and demonstrate proficiency in handling data using R. | PO1, PO2, PO3 | PSO1, PSO2 | CO1 | Knowledge (Remembering) |
| **LO2:** Import, clean, and manipulate data from various sources (CSV, Excel, etc.) using R. | PO2, PO3, PO5 | PSO1, PSO3 | CO2 | Application (Applying) |
| **LO3:** Perform descriptive and inferential statistical analyses in R. | PO1, PO4, PO5 | PSO1, PSO2 | CO3 | Analysis (Analyzing) |
| **LO4:** Visualize data effectively using ggplot2 and interpret graphical outputs to derive meaningful insights. | PO4, PO6 | PSO2, PSO3 | CO4 | Synthesis (Creating) |
| **LO5:** Develop and implement basic machine learning algorithms for predictive modeling and evaluation in R. | PO7, PO8 | PSO1, PSO3 | CO5 | Evaluation (Evaluating) |

**Bloom's Taxonomy Mapping Summary:**

1. **Remembering (Knowledge)**: Students will recall key concepts in statistical analysis, such as types of data, basic statistical methods, and R functions. This is reflected in **CO1** and **LO1**.
2. **Understanding (Comprehension)**: Students will interpret basic concepts of data manipulation, statistical techniques, and the use of R for these tasks, reflected in **CO2** and **LO2**.
3. **Applying (Application)**: The ability to apply statistical methods using R (e.g., performing t-tests, regression) will be developed. These maps to **CO3** and **LO3**.
4. **Analyzing (Analysis)**: Students will learn to analyze datasets for trends, correlations, and relationships, applying inferential statistics like ANOVA, regression, and hypothesis testing. This corresponds to **CO3** and **LO3**.
5. **Creating (Synthesis)**: Students will be expected to create meaningful visualizations using ggplot2 and analyze complex data structures, aligning with **CO4** and **LO4**.
6. **Evaluating (Evaluation)**: Students will assess the performance of machine learning models, evaluate statistical results, and make decisions based on their analysis. This is reflected in **CO5** and **LO5**.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –IV (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 11: FOUNDATIONS OF STATISTICAL DATA ANALYSIS USING ‘R’**

**No. of** Hours/week: 03 **Course Code: 24 – STA – 4C1**1 **C**redits **–** 3

**Syllabus:**

**UNIT – I**

***Introduction to R and R Studio***

**Overview of R**: Introduction to the R programming language, its applications in statistical data analysis, and its benefits. **Setting up R**: Installation and setup of R and R-Studio. **R-Studio Interface**: Understanding the environment, console, script editor, and workspace in R-Studio. **Basic R Commands**: Arithmetic operations, data types (numeric, character, logical), creating and manipulating variables. **R Functions**: Built-in functions and how to define custom functions in R. **R Help System**: Using help files, documentation, and online resources for R.

**UNIT – II**

***Data Structures in R***

**Vectors**: Creation, indexing, and manipulation of vectors. Operations on vectors (addition, subtraction, etc.). **Matrices**: Creating and managing matrices, matrix operations (addition, multiplication, transposition). **Data Frames**: Introduction to data frames, accessing and modifying data frames, adding/removing columns and rows. **Lists**: Creating and using lists, accessing list elements, lists of different data types. **Factors**: Introduction to factors, levels, and their application in categorical data analysis. **Handling Missing Data**: Techniques for dealing with missing data in R.

**UNIT – III**

***Statistical Data Analysis with R***

**Descriptive Statistics**: Measures of central tendency (mean, median, mode), measures of variability (variance, standard deviation), and shape of the distribution (skewness, kurtosis). **Data Visualization**: Creating and customizing plots in R (bar charts, histograms, pie charts, scatter plots). **Exploratory Data Analysis (EDA)**: Techniques for summarizing and visualizing data to detect patterns and anomalies. **Correlation and Covariance**: Calculating and interpreting correlation coefficients and covariance matrices. **Statistical Functions in R**: Using R’s built-in functions to compute summary statistics and visualize distributions.

**UNIT – I*V***

***Inferential Statistics Using R***

**Hypothesis Testing**: Conducting one-sample and two-sample t-tests, chi-square tests, and z-tests using R. **Confidence Intervals**: Computing and interpreting confidence intervals for population parameters. **Analysis of Variance (ANOVA)**: One-way and two-way ANOVA using R, assumptions, and interpretation of results. **Regression Analysis**: Simple linear regression, multiple regression, and model diagnostics in R. **Chi-Square Tests**: Goodness-of-fit test and test for independence using R. **Non-Parametric Tests**: ruskal-Wallis and Mann-Whitney tests in R.

**UNIT – V**

***Advanced Statistical Techniques and Applications***

**Time Series Analysis**: Introduction to time series data, trend analysis, and forecasting using R. **Principal Component Analysis (PCA)**: Understanding PCA, dimensionality reduction, and application in multivariate data. **Cluster Analysis**: K-means clustering, hierarchical clustering, and visualizing clusters in R. **Data Mining Techniques**: Association rules, decision trees, and random forests in R. **Statistical Modeling**: Generalized Linear Models (GLM), logistic regression, and model interpretation. **Advanced Data Visualizations**: Interactive plots using libraries like ggplot2, 3D plots, and advanced chart customizations.

**Practical Sessions**

**UNIT – I**

***Introduction to R and R Studio***

**Practical Sessions:**

1. **Setting Up R and RStudio**: Installing R and RStudio on local systems and getting familiar with the RStudio interface.
2. **Basic R Commands**: Using R for basic arithmetic, variable assignments, and exploring data types (numeric, character, logical).
3. **Creating Variables and Using Functions**: Writing simple functions in R, using in-built functions, and performing basic data manipulations.
4. **Exploring R Help System**: Using R's help system to get documentation on functions and packages.
5. **Basic Data Visualization**: Creating simple plots (scatter plots, histograms, and bar charts) using R commands.

**Reference Books**:

* *"R for Data Science"* by Hadley Wickham and Garrett Grolemund.
* *"The Art of R Programming"* by Norman Matloff.

**Text Books**:

* *"R Programming for Data Science"* by Roger D. Peng.
* *"Hands-On Programming with R"* by Garrett Grolemund.

**UNIT – II**

***Data Structures in R***

**Practical Sessions:**

1. **Working with Vectors**: Creating, indexing, and manipulating vectors in R. Operations such as element-wise arithmetic, sorting, and sub setting.
2. **Matrix Operations**: Creating matrices, performing matrix operations (transpose, multiplication), and indexing matrices.
3. **Data Frames**: Creating and manipulating data frames, adding/removing rows and columns, and accessing specific elements.
4. **Lists and Factors**: Creating lists, accessing list elements, and working with factors for categorical data.
5. **Handling Missing Data**: Identifying and handling missing values using functions like na.omit(), na.rm(), and replacing missing values.

**Reference Books**:

* *"R for Data Science"* by Hadley Wickham and Garrett Grolemund.
* *"R Programming for Data Science"* by Roger D. Peng.

**Text Books**:

* *"Data Science from Scratch: First Principles with Python"* by Joel Grus (This book has conceptual overlap with R, useful for understanding the core concepts).
* *"The R Book"* by Michael J. Crawley.

**UNIT – III**

***Statistical Data Analysis with R***

**Practical Sessions:**

1. **Descriptive Statistics**: Calculating mean, median, variance, standard deviation, skewness, and kurtosis for a dataset.
2. **Creating and Customizing Plots**: Using ggplot2 to create advanced visualizations (boxplots, scatter plots, histograms).
3. **Exploratory Data Analysis (EDA)**: Summarizing data using summary(), detecting outliers, and visualizing distributions.
4. **Correlation and Covariance**: Computing correlation coefficients and covariance matrices, visualizing relationships between variables.
5. **Statistical Functions**: Using R functions to compute statistical summaries such as quantiles, summary tables, and visualize results.

**Reference Books**:

* *"R for Data Science"* by Hadley Wickham and Garrett Grolemund.
* *"Data Science for Business"* by Foster Provost and Tom Fawcett (Conceptual application of statistical methods).

**Text Books**:

* *"Statistics with R: A Beginner’s Guide"* by John Verzani.
* *"Modern Applied Statistics with S"* by W.N. Venables and B.D. Ripley.

**UNIT – IV**

***Inferential Statistics Using ‘R’***

**Practical Sessions:**

1. **Hypothesis Testing**: Conducting one-sample t-tests, two-sample t-tests, and chi-square tests using R functions.
2. **Confidence Intervals**: Computing and interpreting confidence intervals for population means and proportions.
3. **ANOVA**: Performing one-way and two-way ANOVA using R, interpreting the results.
4. **Regression Analysis**: Fitting simple and multiple linear regression models, evaluating model fit, and interpreting regression outputs.
5. **Non-Parametric Tests**: Conducting Kruskal-Wallis and Mann-Whitney U tests in R.

**Reference Books**:

* *"Practical Statistics for Data Scientists"* by Peter Bruce and Andrew Bruce.
* *"Statistics for Business and Economics"* by Paul Newbold, William L. Carroll, and Betty Thorne.

**Text Books**:

* *"Applied Multivariate Statistical Analysis"* by Richard A. Johnson and Dean W. Wichern.
* *"The Essence of Multivariate Thinking"* by Lisa L. Harlow.

**UNIT – V**

***Advanced Statistical Techniques and Applications***

**Practical Sessions:**

1. **Time Series Analysis**: Analyzing and forecasting time series data using R’s forecast package.
2. **Principal Component Analysis (PCA)**: Performing PCA for dimensionality reduction and visualizing the results.
3. **Cluster Analysis**: Implementing k-means clustering and hierarchical clustering for segmenting data.
4. **Statistical Modeling**: Fitting generalized linear models (GLM) and logistic regression models, interpreting model coefficients.
5. **Advanced Data Visualizations**: Creating interactive plots and 3D visualizations using plotly and ggplot2.

**Reference Books**:

* *"The Elements of Statistical Learning"* by Trevor Hastie, Robert Tibshirani, and Jerome Friedman.
* *"Practical Data Science with R"* by Nina Zumel and John Mount.

**Text Books**:

* *"Data Mining with R: Learning with Case Studies"* by Luis Torgo.
* *"Advanced R"* by Hadley Wickham.

**Evaluation and References:**

* **Assignments**: Hands-on programming tasks and application of R in data analysis.
* **Project**: A mini-project involving real datasets, statistical analysis, and visualization using R.
* **Exams**: Short tests based on the practical implementation of concepts.

**Recommended Reading:**

1. *"R for Data Science"* by Hadley Wickham and Garrett Grolemund.
2. *"The R Book"* by Michael J. Crawley.
3. *"Practical Data Science with R"* by Nina Zumel and John Mount.
4. *"Applied Multivariate Statistical Analysis"* by Richard A. Johnson and Dean W. Wichern.

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER: IV (CBCS, w.e.f 2024-25)**

**COURSE TITLE: FOUNDATIONS OF STATISTICAL DATA ANALYSIS USING ‘R**’

**Time: 3 Hours Subject Code: 24-STA-4C11 Max. Marks: 75**

**SECTION – A**

Answer any **ALL** questions. Each question carry **TWO** marks **5 X 2 = 10M**

1. Define the term “statistical data analysis” and explain its significance in research.
2. Describe the difference between mean () and median() in R.
3. Explain the concept of missing data in R. How can you handle it?
4. Briefly explain the ggplot2 package in R and its importance for data visualization.
5. How would you create a data frame in R? Give a simple example.

**SECTION – B**

Answer any **FIVE** questions. Each question carries ‘**5’** marks. **5 X 5 = 25 M**

1. Describe the process of importing data into R. How would you import a CSV file?
2. Explain the concept and use of data frames in R. Provide an example of creating and modifying a data frame.
3. Explain the concept of vectorization in R. How does it improve performance in data analysis?
4. What is a scatter plot? How would you create one in R using the ggplot2 package?
5. Describe the steps to perform a one-way ANOVA test in R. What is its significance?
6. What are the various types of visualizations you can create in R? Explain a basic visualization example using ggplot2.
7. How would you perform linear regression analysis in R? Explain the process and interpretation of output.
8. What is PCA (Principal Component Analysis)? Explain how you would perform PCA in R for dimensionality reduction.

**SECTION – B**

**Answer any FOUR Questions. Each Question Carries 10 marks. 4X 10 = 40M**

1. Explain in detail the concept of Exploratory Data Analysis (EDA). How can you perform EDA using R? Provide examples of various functions used in EDA.
2. How would you handle categorical data in R? Explain how to work with factors and tables in R with suitable examples.
3. Discuss the concept of time series analysis. How would you analyze and forecast a time series dataset using R? Provide an example.
4. Discuss the process of handling missing data in R. Explain the various strategies and functions available for dealing with missing values.
5. Provide a detailed explanation of the ggplot2 package in R. Discuss its components and how it can be used to create complex visualizations.
6. Discuss the importance of normalization in data analysis. How can you normalize data in R? Provide an example where normalization is necessary.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –V (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 12: OPERATIONS RESEARCH – I**

**No. of** Hours/week: 03 **Course Code: 24 – STA – 5C12** **C**redits **–** 3

**Objective:** The Objective of the Course is to introduce the basic concepts of Operational Research and linear programming to the students.

**Learning Objectives and Outcomes for Operations Research**

**Learning Objectives (LOs):**

1. Understand the concepts, origins, and applications of Operations Research (OR) in various fields.
2. Learn the methods used in solving Linear Programming Problems (LPP) and their graphical and algebraic solutions.
3. Gain knowledge of the simplex method and artificial variable techniques for solving LPPs.
4. Understand duality theory in LPP and its application in solving problems.
5. Develop skills in interpreting the solutions of optimization problems and handling exceptional cases.

**Learning Outcomes (LOs):**

1. Describe the origin, features, and applications of Operations Research.
2. Formulate and solve Linear Programming Problems using graphical and simplex methods.
3. Apply the artificial variable technique and handle degeneracy in Linear Programming Problems.
4. Understand the concept of duality and solve primal problems using duality concepts.
5. Solve optimization problems and interpret results, including handling exceptional cases like unbounded and degenerate solutions.

**Course Outcomes (COs):**

| **Course Outcomes (COs)** | **Bloom’s Taxonomy Level** |
| --- | --- |
| **CO1**: Understand and explain the nature, origin, and applications of Operations Research. | **Understanding** |
| **CO2**: Formulate Linear Programming Problems and solve them using graphical methods and simplex methods. | **Applying** |
| **CO3**: Apply artificial variable techniques, including Big-M and two-phase methods, to solve LPPs. | **Applying** |
| **CO4**: Solve optimization problems involving duality, including the dual simplex method. | **Applying** |
| **CO5**: Interpret and analyze the solutions of LPPs, including handling exceptional cases. | **Analyzing** |

**Program Outcomes (POs):**

1. **PO1**: Demonstrate the ability to understand and solve optimization problems using Operations Research techniques.
2. **PO2**: Develop problem-solving skills through the application of mathematical modeling in real-world scenarios.
3. **PO3**: Apply linear programming and related techniques to formulate and solve practical optimization problems.
4. **PO4**: Use computational tools and techniques, including simplex and dual simplex methods, to solve complex problems.
5. **PO5**: Communicate effectively the results and solutions of optimization problems in both written and oral forms.
6. **PO6**: Work collaboratively in teams to solve operational research problems and apply interdisciplinary knowledge.
7. **PO7**: Critically evaluate the efficiency and feasibility of different optimization techniques.
8. **PO8**: Engage in lifelong learning to stay updated with advancements in Operations Research and its applications.

**Program-Specific Outcomes (PSOs):**

1. **PSO1**: Apply Operations Research techniques to solve real-life problems in industries like manufacturing, logistics, and service sectors.
2. **PSO2**: Formulate optimization models for decision-making processes and apply them using computational methods.
3. **PSO3**: Analyze the behavior of optimization problems under different constraints and solve them using simplex and duality methods.

**Mapping Table: COs, POs, PSOs with Bloom’s Taxonomy Connectivity**

| **Course Outcomes (COs)** | **Program Outcomes (POs)** | **Program-Specific Outcomes (PSOs)** | **Bloom’s Taxonomy** |
| --- | --- | --- | --- |
| **CO1**: Understand and explain the nature, origin, and applications of Operations Research. | **PO1, PO2**: Understand and apply the methods of Operations Research to real-world problems. | **PSO1**: Apply OR techniques to solve industrial and service sector problems. | **Understanding** |
| **CO2**: Formulate Linear Programming Problems and solve them using graphical methods and simplex methods. | **PO3, PO4**: Formulate and solve optimization problems using appropriate methods and computational tools. | **PSO2**: Formulate and solve optimization models using simplex methods. | **Applying** |
| **CO3**: Apply artificial variable techniques, including Big-M and two-phase methods, to solve LPPs. | **PO3, PO4**: Use computational tools and methods like Big-M and two-phase simplex for optimization problems. | **PSO2**: Apply artificial variable techniques to solve optimization problems. | **Applying** |
| **CO4**: Solve optimization problems involving duality, including the dual simplex method. | **PO4, PO7**: Use duality theory and the dual simplex method to solve optimization problems and evaluate solutions. | **PSO3**: Analyze and solve problems using duality theory and the dual simplex method. | **Applying** |
| **CO5**: Interpret and analyze the solutions of LPPs, including handling exceptional cases. | **PO5, PO8**: Analyze the solutions to optimization problems and interpret exceptional cases such as unbounded and degenerate solutions. | **PSO3**: Analyze and handle exceptional cases in linear programming problems. | **Analyzing** |

**Summary of Bloom’s Taxonomy Mapping:**

* **Understanding (CO1)**: The first outcome introduces the foundational concepts of Operations Research, focusing on its origin, features, and applications.
* **Applying (CO2, CO3, CO4)**: These outcomes encourage students to apply linear programming and related techniques like the simplex method, artificial variable techniques, and duality methods to solve real-world optimization problems.
* **Analyzing (CO5)**: The final outcome is about analyzing and interpreting the results of linear programming problems, including exceptional cases like unbounded and degenerate solutions.

This mapping reflects the connection between the course content and the program's educational objectives, advancing students from understanding foundational concepts to applying and analyzing complex optimization techniques.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –V (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE - 12: OPERATIONS RESEARCH – I**

**No. of** Hours/week: 03 **Course Code: 24 – STA – 5C12** **C**redits **–** 3

**II. Syllabus**

**UNIT – I**

**Operations Research**

**Introduction to Operations Research:** Origin and development of Operations Research, Nature and features of Operations Research, Scientific method in Operations Research. **Modeling in Operations Research:** Role of models in OR, Advantages and limitations of OR models, General solution methods for OR models. **Applications of Operations Research:** Practical applications across various fields. **Linear Programming Problem (LPP):** Introduction to Linear Programming Problem, Mathematical formulation of Linear Programming Problems, Illustrations and examples of LPP formulation, **Problem-Solving and Applications :**Real-life examples and problem-solving exercises for LPP, Interpretation of LPP solutions in practical scenarios.

**UNIT – II**

**Graphical Solution of Linear Programming Problems (LPP)**

**Introduction to Graphical Method:** Concept and importance of the graphical solution for LPP, Overview of maximizing and minimizing objective functions. **Graphical Solution for LPP:** Graphical representation of constraints, Solving LPP with up to 3 variables using graphical methods. **Convex and Non-Convex Hulls in LPP:** Identification and determination of convex hulls, Understanding and finding non-convex hulls. **Exceptional Cases in LPP:** Alternative solutions and their graphical representation, Identification of unbounded solutions, Non-existing feasible solutions and interpretation. **Problem-Solving:** Practical examples and exercises to illustrate each concept, Analysis of exceptional cases through graphical methods

**UNIT – III**

**General Linear Programming Problem (GLP)**

**Introduction to General Linear Programming Problem (GLP):** Definition of GLP, Matrix form representation of GLP. **Variables and Forms in LPP:** Slack variables and their role in LPP, Surplus variables and their significance, Unrestricted variables in LPP, Standard form and canonical form of LPP. **Solutions in LPP:** Solution, Basic solution, Degenerate solution, Basic feasible solution, Optimum basic feasible solution. **Introduction to Simplex Method:** Concept of the Simplex method in solving LPP, Computational procedure of the Simplex algorithm. **Solving LPP using the Simplex Method:** Solving maximization problems (up to three variables), Solving minimization problems (up to three variables). **Problem-Solving and Practice:** Practical examples for converting GLP into standard and canonical forms, Exercises on solving LPP using the Simplex method for maximization and minimization cases

**UNIT – IV**

**Artificial Variable Technique and Advanced Topics in LPP**

**Artificial Variable Techniques:** Introduction to artificial variables in LPP, Big-M method: Concept and computational procedure, Two-phase simplex method: Steps and applications. **Degeneracy in LPP:** Understanding degeneracy in LPP solutions, Methods to resolve degeneracy. **Exceptional Cases in LPP:** Alternative solutions: Identification and interpretation, Unbounded solutions: Concept and graphical representation, Non-existing feasible solutions: Causes and resolution. **Solution of Simultaneous Equations Using Simplex Method:** Application of simplex method for solving simultaneous equations. **Problem-Solving and Practice:** Practical examples for Big-M and Two-phase simplex methods, Exercises to identify and handle degeneracy and exceptional cases in LPP

**UNIT – V**

**Duality in Linear Programming**

**Concept of Duality in Linear Programming:** Introduction to duality and its significance in LPP. **Primal and Dual Problems:** Definitions of primal and dual problems, General rules for converting a primal problem into its dual. **Relationship Between Primal and Dual Solutions:** Statements on the relationship between primal and dual solutions, Practical insights into the interdependence of primal and dual problems. **Solving Primal Problems Using Duality:** Application of duality concepts to solve primal problems. **Dual Simplex Method:** Introduction to the dual simplex method: Computational steps and procedure. **Problem-Solving and Practice:** Examples of converting primal problems into dual form, Exercises on solving primal problems using duality, Practice with the dual simplex method for practical applications.

**Practical Syllabus Credits: 1 (2 hrs/week)**

1. To solve Linear Programming Problem using Graphical Method with

i. Unbounded solution ii. Infeasible solution

1. Solution of LPP with simplex method.
2. Problem solving using Big M - method.
3. Problem solving using Two Phase method.
4. Illustration of following special cases in LPP using Simplex method

iii. Unbounded solution

iv. Alternative or multiple solutions.

1. Problems based on Principle of Duality.
2. Problems based on Dual simplex method.

**III. References**

1. S.D. Sharma: Operations Research, Kedar Nath Ram Nath & Co, Meerut.

2. Kanti Swarup, P.K.Gupta, Manmohn: Operations Research, Sultan Chand and sons, New Delhi.

3. J.K. Sharma: Operations Research and Application, Mc.Millan and Company, New Delhi.

 4. Gass S.I: Linear Programming. Mc Graw Hill.

 5. Hadly G: Linear Programming. Addison-Wesley.

 6. Taha H.M: Operations Research: An Introduction : Mac Millan.

**IV. Suggested Co-curricular Activities:**

1. Training of students by related industrial experts

2. Assignments including technical assignments if any.

3. Seminars, Group Discussions, Quiz, Debates etc on related topics.

4. Preparation of audio and videos on tools of diagrammatic and graphical representations.

5. Collection of material/figures/photos/author photos of related topics.

6. Invited lectures and presentations of stalwarts to those topics.

 7. Visits/field trips of firms, research organizations etc.

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER –V (CBCS, w.e.f 2024-25)**

**COURSE TITLE: OPERATIONS RESEARCH – I**

**Time: 3 Hours Course Code: 24 – STA – 5C12 Max. Marks: 75M**

**SECTION – A**

**Answer ALL the Five Questions. Each Question Carries 2 marks. 5 X 2= 10M**

1. Write the simple definition of OR.
2. Define objective function and constraints.
3. Define Slack and Surplus variables.
4. Define a) objective function b) basic feasible solution.
5. Write the dual to the following LPP

$$MaximizeZ=x\_{1}-x\_{2}+3x\_{3}$$

$$Subjectto$$

$$x\_{1}+x\_{2}+x\_{3}\leq 10$$

$$2x\_{1}-x\_{3}\leq 2$$

$$2x\_{1}-2x\_{2}-3x\_{3}\leq 6$$

$$x\_{1}, x\_{2},x\_{3}\geq 0$$

**SECTION – B**

**Answer any FIVE Questions. Each Question Carries 5 marks. 5X 5 = 25M**

1. Explain the definition and limitations of OR.
2. Explain the procedure of graphical method for solving LPP.
3. Explain the following terms.
4. Slack Variable (b) Surplus Variable.
5. Define a) objective function b) basic feasible solution.
6. Define a) unbounded solution b) convex combination.
7. Prove that Dual of Dual is Primal.
8. Write the advantages of Dual simplex method.
9. Write a short note on post optimal sensitivity analysis.

**SECTION – C**

**Answer any FOUR Questions. Each Question Carries 10 marks. 4X 10 = 40M**

1. Define OR. Write the scope and characteristics of OR.
2. ABC Foods Company is developing a low- calorie high-protein diet supplement called Hi-pro. The specifications for Hi-pro have been established by a panel of medical experts. These specifications along with the calorie, protein and vitamin content of three basic foods are given in the following table.

|  |  |
| --- | --- |
| Nutritional elements | Units of nutritional elements(Per 1000 g serving of basic foods) |
|  | 1 | 2 | 3 |  |
| Calories | 350 | 250 | 200 | 300 |
| Proteins | 250 | 300 | 150 | 200 |
| Vitamin A | 100 | 150 | 75 | 100 |
| Vitamin C | 75 | 125 | 150 | 100 |
| Cost per Serving (Rs.) | 15 | 20 | 12 |  |

` What quantities of foods 1, 2 and 3 should be used? Formulate this problem as an LP model to minimize cost.

1. Solve the following LPP by using graphical method.

$$MaxZ = -8x\_{1}+2x\_{2}$$

 $subjectto$

 $-4x\_{1}+2x\_{2}\leq 1 $

$$5x\_{1}-4x\_{2}\leq 3 andx\_{1}, x\_{2}\geq 0$$

1. Solve the following LPP using Simplex method to

$$MaxZ = 3x\_{1}+2x\_{2}+5x\_{3}$$

$$subjecttox\_{1}+2x\_{2}+x\_{3}\leq 430 $$

$$3x\_{1}+2x\_{2}\leq 260 $$

$$x\_{1}+4x\_{2}\leq 420andx\_{1}, x\_{2},x\_{3}\geq 0$$

1. Solve the following LPP using Two -Phase simplex method

$$MinimizeZ = 3x\_{1}+2x\_{2}+5x\_{3}$$

$$subjectto 4x\_{1}+8x\_{2}+6x\_{3}\geq 64$$

$$3x\_{1}+6x\_{2}+ 12x\_{3}\geq 96$$

$$andx\_{1}, x\_{2},x\_{3}\geq 0$$

1. Solve the following LPP by Dual Simplex method

$$MaximizeZ= – 2x\_{1}-x\_{3}$$

$$subjecttox\_{1}+x\_{2}-x\_{3}\geq 5$$

$$x\_{1}-2x\_{2}+4x\_{3}\geq 8$$

$$andx\_{1},x\_{2},x\_{3}\geq 0$$

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –V (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE 13: STATISTICAL QUALITY CONTROL**

**No. of** Hours/week: 03 **24 – STA – 5C1**3 Credits: 03

**Learning Objectives and Outcomes for Statistical Quality Control (SQC)**

**Learning Objectives (LOs):**

1. Understand the significance of Statistical Quality Control (SQC) in process and product quality management.
2. Learn various SQC techniques, including control charts for variables and attributes, and acceptance sampling.
3. Understand how to construct and interpret control charts for both variables (mean, range, standard deviation) and attributes (p, np, u, C charts).
4. Develop the skills to implement corrective actions for out-of-control processes using SQC methods.
5. Apply acceptance sampling methods for quality control, including the single sampling plan, and compute key quality metrics like AOQ, ATI, and AOQL.

**Learning Outcomes (LOs):**

1. Describe the key components and objectives of Statistical Quality Control and identify the 4 M's.
2. Construct and interpret control charts for variables and attributes to monitor and improve quality.
3. Apply corrective actions for processes that are out of control.
4. Develop and implement acceptance sampling plans, including the computation of AOQ, ATI, and AOQL.
5. Solve practical problems related to sampling plans and interpret the results, including understanding the relationship between sampling inspection and full inspection.

**Course Outcomes (COs):**

| **Course Outcomes (COs)** | **Bloom’s Taxonomy Level** |
| --- | --- |
| **CO1**: Understand the concepts, importance, and applications of Statistical Quality Control (SQC). | **Understanding** |
| **CO2**: Construct and interpret control charts for variables and attributes (p, np, u, C charts). | **Applying** |
| **CO3**: Implement corrective actions for out-of-control processes using SQC methods. | **Applying** |
| **CO4**: Develop and apply acceptance sampling methods to monitor and control product quality. | **Applying** |
| **CO5**: Calculate and interpret key metrics such as AOQ, ATI, AOQL, and apply them in practical quality control situations. | **Analyzing** |

**Program Outcomes (POs):**

1. **PO1**: Understand and apply quality control methods to solve process optimization and product quality issues.
2. **PO2**: Develop problem-solving skills in managing variations and improving processes using statistical techniques.
3. **PO3**: Construct and interpret various control charts for both variables and attributes to monitor and improve quality.
4. **PO4**: Use statistical sampling methods, including acceptance sampling and single sampling plans, for quality assurance in industrial settings.
5. **PO5**: Effectively communicate the results of quality control analyses and actions, both orally and in written reports.
6. **PO6**: Collaborate effectively in teams to identify and solve quality-related problems using statistical techniques.
7. **PO7**: Apply modern quality control techniques in manufacturing and service industries to achieve continuous improvement.
8. **PO8**: Demonstrate the ability to evaluate quality control processes and suggest improvements based on statistical evidence.

**Program-Specific Outcomes (PSOs):**

1. **PSO1**: Apply quality control techniques, including statistical methods and sampling plans, to monitor and control manufacturing processes.
2. **PSO2**: Design and implement control charts and acceptance sampling plans to ensure product and process quality in industrial settings.
3. **PSO3**: Analyze and interpret quality control data to make informed decisions and improve product quality.

**Mapping Table: COs, POs, PSOs with Bloom’s Taxonomy Connectivity**

| **Course Outcomes (COs)** | **Program Outcomes (POs)** | **Program-Specific Outcomes (PSOs)** | **Bloom’s Taxonomy** |
| --- | --- | --- | --- |
| **CO1**: Understand the concepts, importance, and applications of Statistical Quality Control (SQC). | **PO1, PO2**: Understand and apply statistical methods for quality control and process optimization. | **PSO1, PSO2**: Apply SQC methods for process monitoring and control. | **Understanding** |
| **CO2**: Construct and interpret control charts for variables and attributes (p, np, u, C charts). | **PO3**: Construct and interpret control charts for monitoring product and process quality. | **PSO1, PSO2**: Construct and interpret various control charts to monitor product and process quality. | **Applying** |
| **CO3**: Implement corrective actions for out-of-control processes using SQC methods. | **PO4, PO7**: Implement corrective actions to improve processes using statistical quality control methods. | **PSO2, PSO3**: Identify and correct out-of-control processes using SQC methods. | **Applying** |
| **CO4**: Develop and apply acceptance sampling methods to monitor and control product quality. | **PO4**: Develop and implement acceptance sampling methods for quality assurance. | **PSO1, PSO2**: Design and apply sampling inspection plans. | **Applying** |
| **CO5**: Calculate and interpret key metrics such as AOQ, ATI, AOQL, and apply them in practical quality control situations. | **PO5, PO6**: Communicate and analyze quality control results and suggest improvements. | **PSO3**: Analyze and interpret quality control metrics for decision-making. | **Analyzing** |

**Summary of Bloom’s Taxonomy Mapping:**

* **Understanding (CO1)**: The first outcome provides the foundational understanding of Statistical Quality Control, including its significance and objectives.
* **Applying (CO2, CO3, CO4)**: The outcomes encourage students to apply control chart techniques, implement corrective actions, and develop acceptance sampling plans to improve product and process quality.
* **Analyzing (CO5)**: The final outcome emphasizes the analysis and interpretation of key quality control metrics and their practical application in decision-making.

This mapping ensures that the course advances students from understanding the basics of SQC to applying advanced methods and analyzing real-world quality control data to make informed decisions.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –V (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE 13: STATISTICAL QUALITY CONTROL**

**No. of** Hours/week: 03 **24 – STA – 5C1**3 Credits: 03

**II. Syllabus**

**UNIT – I**

**Statistical Quality Control (SQC)**

**Introduction to SQC:** Importance and objectives of Statistical Quality Control, Overview of the 4 M’s of SQC (Machine, Man, Material, and Method). **Causes of Variation:** Types of variation: Assignable causes and chance causes, Understanding and managing variation in processes. **Applications of SQC:** Uses of SQC in process control and product quality control. **Control Charts Technique:** Concept and purpose of control charts, Statistical basis of Shewart control charts, **Problem-Solving and Practice:** Examples of applying control chart techniques, Exercises on identifying and managing variations using SQC methods

**UNIT – II**

**Control Charts for Variables**

**Introduction to Control Charts for Variables:** Purpose and significance of Mean and Range (R) charts, Overview of Mean and Standard Deviation (σ) charts. **Construction of Mean and Range Charts:** Steps for constructing Mean and Range (R) charts, Interpretation of control limits and process stability. **Mean and Standard Deviation Charts:** Construction of Mean Standard deviation (σ) charts when standards are specified, Construction of Mean standard deviation (σ) charts when standards are unspecified. **Corrective Actions for Out-of-Control Processes:** Identifying signals of statistical instability, Steps to implement corrective actions. **Problem-Solving and Practice:** Practical examples for constructing and interpreting Mean, Range, and Standard Deviation charts, Exercises on handling processes that are out of statistical control

**UNIT – III**

**Control Charts for Attributes**

**Introduction to Control Charts for Attributes:** Overview and significance of attribute-based control charts, Applications in monitoring product and process quality. **Fraction Defective Chart (p-Chart):** Concept and purpose of the p-chart, Construction of p-charts when standards are specified and unspecified. **Number of Defectives Chart (np-Chart):** Concept and purpose of the np-chart, Steps for constructing np-charts when standards are specified and unspecified. **Number of Defects Per Unit Chart (u-Chart):** Concept and application of the u-chart, Construction of u-charts for varying sample sizes. **C Chart:** Purpose of the C chart for defect counts, Steps to construct C charts when standards are specified and unspecified. **Corrective Actions for Out-of-Control Processes:** Identifying signals of statistical instability in attribute charts, Implementing corrective actions for process improvement. **Problem-Solving and Practice:** Practical examples of constructing and interpreting p, np, u, and C charts, Exercises on managing processes that are out of statistical control.

**UNIT – IV**

**Acceptance Sampling for Attributes**

**Introduction to Acceptance Sampling for Attributes:** Overview and importance of acceptance sampling in quality control, Concept and objectives of acceptance sampling. **Sampling Inspection Plan:** Definition and structure of a sampling inspection plan, Key components of a sampling plan. **Comparison Between 100% Inspection and Sampling Inspection:** Advantages and disadvantages of 100% inspection, Benefits of using sampling inspection over full inspection. **Procedures of Acceptance Sampling with Rectification:** Steps and procedures involved in acceptance sampling, Role of rectification in acceptance sampling. **Producer's Risk and Consumer's Risk:** Definition and explanation of producer's risk (Type I error), Definition and explanation of consumer's risk (Type II error). **Operating Characteristic (OC) Curve:** Concept of the OC curve in acceptance sampling, Interpretation and application of the OC curve. **Key Quality Metrics:** Acceptable Quality Level (AQL): Definition and significance, Lot Tolerance Fraction Defective (LTFD) and Lot Tolerance Percent Defective (LTPD), Average Outgoing Quality (AOQ) and Average Outgoing Quality Limit (AOQL), **AOQ Curve and ASN:** Construction and interpretation of AOQ curve, Average Sample Number (ASN): Concept and importance. **Average Total Inspection (ATI):** Definition of ATI and its role in sampling plans, Calculation and significance of ATI. **Problem-Solving and Practice:** Practical examples of acceptance sampling plans, Exercises on calculating AOQ, AOQL, ASN, and ATI

**UNIT – V**

**Single Sampling Plan**

**Introduction to Single Sampling Plan:** Overview of the single sampling plan in quality control, Purpose and applications of the single sampling plan. **Computation of Probability of Acceptance:** Definition and concept of probability of acceptance, Calculation using Binomial distribution, Calculation using Poisson approximation, Comparison and application of both methods. **Average Outgoing Quality (AOQ) and Average Total Inspection (ATI):** Calculation of AOQ and ATI in single sampling plans, Significance and interpretation of AOQ and ATI. **Graphical Determination of AOQL:** Concept and importance of Average Outgoing Quality Limit (AOQL), Graphical method for determining AOQL, Interpretation of the AOQL curve, **Determination of a Single Sampling Plan: Lot Quality Approach**: Steps to determine the sampling plan based on lot quality, **Average Quality Approach**: Steps to determine the sampling plan based on average quality, **Problem-Solving and Practice:** Practical exercises on computing probability of acceptance using Binomial and Poisson approximations, Problem-solving related to AOQ, ATI, and AOQL determination, Application of both approaches (lot quality and average quality) in determining sampling plans

**Practical Syllabus**  Practical Credits: 1 2 hrs/week

1. Construction of Mean and R Charts.

2. Construction of Mean and Standard deviation charts.

3. Construction of p Chart for fixed sample size.

4. Construction of p Chart for variable sample size.

5. Construction of np Chart.

6. Construction of C chart.

7. Construction of U chart.

8. Single sampling plan for attributes (OC Curve, Producer’s and Consumer’s risks, AOQ, AOQL, ATI).

9. Determination of single sampling plan by: a) lot quality approach b) average quality approach.

**III. References**

1. Montgomery, D. C. (2008): Statistical Quality Control, 6thEdn., John Wiley, New York.

2. Parimal Mukhopadhyay: Applied Statistics, New Central Book Agency.

3. Goon A.M., Gupta M.K. and Das Gupta B. (1986): Fundamentals of Statistics, Vol. II, World Press, Calcutta.

4. S.C. Gupta and V.K. Kapoor: Fundamentals of Applied Statistics – Chand publications.

5. R.C. Gupta: Statistical Quality Control.

6. Duncan A.J. (1974): Quality Control and Industrial Statistics, fourth edition

7. D.B. Taraporewala Sons and Co. Pvt. Ltd., Mumbai.

**IV. Suggested Co-curricular Activities:**

1. Training of students by related industrial experts

2. Assignments including technical assignments if any.

3. Seminars, Group Discussions, Quiz, Debates etc on related topics.

4. Preparation of audio and videos on tools of diagrammatic and graphical representations.

5. Collection of material/figures/photos/author photoes of related topics.

6. Invited lectures and presentations of stalwarts to those topics.

7. Visits/field trips of firms, research organizations etc.

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER –V (CBCS, w.e.f 2024-25)**

**COURSE TITLE 13: STATISTICAL QUALITY CONTROL**

**Time: 3 Hours Subject Code: 24-STA - 5C13 Max. Marks: 75**

**SECTION –A**

Answer any **FIVE** Questions .Each question Carries **TWO** marks **5 X 2 = 10**

1. Define Natural Tolerance limits?
2. What are Control limits for R –Chart?
3. What are the Control limits for P – Chart for different sample size?
4. Define the term Reliability.
5. Define ASN.

**SECTION –B**

Answer any **FIVE** Questions. Each question carries **FOUR** Marks. **5 X 5 = 2**5

1. Define SQC. What are the uses of SQC?
2. Explain the causes of variation in production process.
3. Discuss control Chart for Number of defectives
4. How can the Shewart control charts be interpreted to draw meaningful conclusions?
5. What do you understand by acceptance sampling plan?
6. Define a consumer and producer’s risk.
7. Explain single sampling plan?
8. Explain sampling inspection?

**SECTION –C**

Answer **ALL** Questions. Each question carries **TEN** marks **4 X 10 = 40**

1. Define Statistical Quality Control and explain the necessity of Statistical Quality Control.
2. Distinguish between process control and product control?
3. Explain in detail the $\overbar{X}$– chart and R – charts. What purposes do they serve?
4. Sixteen boxes of Electric switches each containing 20 switches were randomly selected from a lot of switch boxes and inspected for the number of defects per box were as follows

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Box No** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| **No. Defects** | 12 | 15 | 9 | 14 | 18 | 26 | 8 | 6 | 11 | 12 | 16 | 13 | 19 | 18 | 14 | 21 |

 Construct C – Chart and comment on the nature of control.

1. Define and explain the following terms.
	1. Acceptance Quality Level (AQL)
	2. Lot Tolerance Percent Defective (LTPD)
	3. Average Out going Quality (AOQ)
	4. Average Out going Quality Level (AOQL)
2. Explain about double sampling plan?

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –V (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE 14A: APPLIED STATISTICS**

Theory Credits: 3 **Course Code: 24 – STA – 5C14**A 3 hrs/week

**Learning Objectives:**

* Understand the concept of time series and its components: trend, seasonal, cyclical, and irregular.
* Learn methods for time series analysis, including trend estimation, moving averages, and seasonal components analysis.
* Gain proficiency in constructing index numbers and understanding their importance in economic analysis.
* Understand the significance of vital statistics in demographic studies and its applications in policy-making and planning.
* Learn how to construct and interpret life tables and understand measures of fertility and population growth.

**Learning Outcomes:**

* Students will be able to analyze and forecast time series data using various methods.
* Students will learn to compute seasonal indices and perform deseasonalization for accurate trend analysis.
* Students will develop skills to construct and interpret different types of index numbers and apply them in economic contexts.
* Students will gain an understanding of mortality and fertility rates, applying vital statistics to real-world demographic analysis.
* Students will be proficient in constructing life tables and calculating population growth measures for demographic analysis.

**Course Outcomes (COs):**

1. **CO1:** Apply time series analysis techniques, including trend analysis and moving averages, to real-world data.
2. **CO2:** Compute and interpret seasonal indices using methods such as Ratio to Moving Average and Link Relative Method.
3. **CO3:** Construct index numbers and apply various methods, including Laspeyres, Paasche, and Fisher’s index.
4. **CO4:** Calculate and interpret mortality rates and fertility rates from vital statistics data.
5. **CO5:** Construct life tables and calculate population growth measures like Gross Reproduction Rate (GRR) and Net Reproduction Rate (NRR).

**Program Outcomes (POs):**

1. **PO1:** Understand the basic concepts and methods of time series analysis, vital statistics, and population studies.
2. **PO2:** Develop analytical skills to solve real-world problems in economics and demographics using statistical techniques.
3. **PO3:** Gain the ability to apply statistical methods like index numbers and time series to evaluate economic and health trends.
4. **PO4:** Apply critical thinking in the interpretation and analysis of demographic and economic data.
5. **PO5:** Use appropriate software tools to perform statistical analysis and interpret results effectively.
6. **PO6:** Demonstrate effective communication skills in presenting statistical findings and analyses.
7. **PO7:** Understand the significance of data in decision-making, especially in public health, policy, and economic planning.
8. **PO8:** Apply knowledge of statistical methods to solve problems in areas such as population studies, health, and economics.

**Program Specific Outcomes (PSOs):**

1. **PSO1:** Use time series and index number techniques to analyze economic and demographic data and make informed decisions.
2. **PSO2:** Apply vital statistics in demographic research to understand trends in population health, fertility, and mortality.
3. **PSO3:** Construct and interpret life tables to assess population growth and demographic changes.

**Mapping Summary Table:**

| **Unit No.** | **COs** | **POs** | **PSOs** | **Bloom's Taxonomy Level** |
| --- | --- | --- | --- | --- |
| **Unit 1**: Time Series | CO1 | PO1, PO2, PO3 | PSO1 | Apply, Analyze |
| **Unit 2**: Seasonal Component Analysis | CO2 | PO1, PO3, PO5 | PSO1 | Apply, Analyze |
| **Unit 3**: Index Numbers | CO3 | PO2, PO3, PO6 | PSO2 | Apply, Analyze |
| **Unit 4**: Vital Statistics | CO4 | PO1, PO4, PO7 | PSO2 | Apply, Evaluate |
| **Unit 5**: Life Tables and Population Growth | CO5 | PO2, PO3, PO5 | PSO3 | Apply, Analyze |

**Connectivity with Bloom’s Taxonomy Levels:**

* **Remembering**: Recall key concepts of time series, seasonal components, index numbers, and vital statistics.
* **Understanding**: Explain different methods for time series analysis, constructing index numbers, and interpreting life table data.
* **Applying**: Apply methods such as least squares, moving averages, and ratio to moving average to solve real-world problems.
* **Analyzing**: Analyze complex data to identify trends, seasonal patterns, and make inferences about demographic data.
* **Evaluating**: Evaluate the appropriateness and effectiveness of different statistical methods, including index number construction and life table analysis.
* **Creating**: Construct and interpret time series, index numbers, and life tables to derive insights and forecast future trends.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –V (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE 14A: APPLIED STATISTICS**

Theory Credits: 3 **Course Code: 24 – STA – 5C14**A 3 hrs/week

**II. Syllabus**

**UNIT – I**

**Time Series and Trend Analysis**

**Introduction to Time Series:** Definition and components of time series, Explanation of components: Trend, Seasonal, Cyclical, and Irregular components with illustrations. **Models for Time Series Analysis: Additive Model**: Understanding the additive relationship between components, **Multiplicative Model**: Explanation of the multiplicative relationship between components, **Mixed Model**: Combination of additive and multiplicative models for time series analysis. **Trend Analysis:** Definition and importance of trend in time series data. **Estimation of Trend**: **Free Hand Curve Method**: Graphical method for trend estimation, **Method of Semi Averages**: Dividing the data into two parts and finding averages to estimate trends. **Least Squares Method for Trend Estimation: Linear Trend**: Fitting a straight line to data using the least squares method. **Parabolic Trend**: Fitting a quadratic curve to data using the least squares method. **Moving Averages Method:** Explanation of the moving averages method for smoothing time series data, Use of moving averages to determine the trend component

**UNIT – II**

**Seasonal Component Analysis**

**Introduction to Seasonal Components:** Definition and importance of seasonal variations in time series data, Identifying seasonal patterns and their impact on time series analysis. **Methods for Determining Seasonal Indices- Simple Averages Method**: Calculation of seasonal indices by averaging the data for each season or period. **Ratio to Moving Average Method**: Determining seasonal indices by dividing the observed values by the corresponding moving averages. **Ratio to Trend Method**: Calculation of seasonal indices by dividing the actual data by the estimated trend values. **Link Relative Method**: Use of link relatives to compute seasonal indices for each period.

**UNIT – III**

**Index Numbers:** Definition and purpose of index numbers, Importance of index numbers in comparing relative changes in variables over time or space. **Construction of Index Numbers:** Steps involved in the construction of index numbers, Problems encountered during the construction of index numbers (e.g., selection of base year, choice of weights), Types of index numbers: Price, Quantity, Value Index, etc., **Uses and Limitations of Index Numbers, Types of Index Numbers, Simple Index Numbers**: Calculation and interpretation of simple index numbers, **Weighted Index Numbers**: Simple and Weighted Aggregative Index Numbers, Laspeyres, Paasche , Drobish – Bowley, Marshall Edgeworth and Fisher’s index numbers. **Criterion of a Good Index Number:** Properties of a good index number: Consistency, reliability, and simplicity, Tests of adequacy: Factor reversal, time reversal, etc., **Fisher’s Ideal Index Number:** Derivation and explanation of Fisher’s Ideal Index Number, Advantages and limitations of Fisher’s index. **Special Index Numbers: Cost of Living Index Number**: Purpose, construction, and applications, **Wholesale Price Index Number**: Importance and use in economic analysis.

**UNIT – 4**

**Vital Statistics:** Definition and significance of vital statistics in demographic studies, Importance of vital statistics in public health, policy-making, and planning. **Uses of Vital Statistics:** Role in understanding population dynamics and health trends, Applications in formulating policies for health, education, and welfare. **Sources of Vital Statistics:** Primary sources: Registration systems, censuses, and surveys, Secondary sources: Health organizations, government reports, and demographic studies .**Measures of Mortality Rates: Crude Death Rate (CDR)**: Definition, formula, and interpretation, **Specific Death Rate (SDR)**: Calculation and use of age-specific or cause-specific death rates, **Standardized Death Rate**: Method of standardization using different populations for comparison, Use of age-standardized death rates to compare mortality across populations, **Problems in Mortality Rate Calculation:** Issues such as underreporting, incomplete data, and variations in reporting standards, Problems encountered when comparing mortality rates across different populations or regions

**UNIT – V**

**Life Tables and Population Growth**

**Introduction to Life Table**: Definition and significance in demographic studies, **Columns of a Life Table**: Description of key columns, **Construction of Life Table**: Step-by-step method for constructing a life table from population data, **Uses of Life Table**: Estimation of life expectancy, Understanding mortality patterns across age groups, **Proofs of Life Table Functions**: Mathematical proofs related to life table functions, such as survivor function, death rate, and life expectancy. **Measures of Fertility Rates: Crude Birth Rate (CBR)**: Calculation and interpretation of crude birth rate, **General Fertility Rate (GFR)**: Formula and significance in fertility analysis, **Specific Fertility Rate (SFR)**: Calculation of age-specific fertility rates and their application, **Total Fertility Rate (TFR)**: Definition, calculation, and interpretation of total fertility rate. **Measures of Population Growth: Pearl’s Index**: Calculation and use in measuring population growth, **Gross Reproduction Rate (GRR)**: Definition and calculation, significance in demographic studies, **Net Reproduction Rate (NRR)**: Understanding NRR and its importance in assessing population replacement, **Problems in Population Growth Measures**: Challenges in calculating fertility and reproduction rates due to data quality, underreporting, and changing population structures

**Practical Syllabus** Practical Credits: 1 2 hrs/week

1. Measurement of trend by method of moving averages (odd and even period)

2. Measurement of trend by method of Least squares (linear and parabola)

3. Determination of seasonal indices by method simple averages

4. Determination of seasonal indices by method of Ratio to Moving Averages

5. Determination of seasonal indices by method of Ratio to Trend

6. Determination of seasonal indices by method of Link relatives

7. Computation of simple index numbers.

8. Computation of all weighted index numbers.

9. Computation of reversal tests.

10. Computation of various Mortality rates

11. Computation of various Fertility rates

12. Computation of various Reproduction rates.

13. Construction of Life Table.

**III. References**

1. Fundamentals of Applied Statistics: V. K. Kapoor & S. C. Gupta.

2. Mukopadhyay, P (2011): Applied Statistics, 2nd ed. Revised reprint, Books and Allied Pvt. Ltd.

3. Brockwell, P.J. and Devis, R.A. (2003): Introduction to Time Series Analysis. Springer.

4. Chatfield, C. (2001): Time Series Forecasting., Chapman & Hall.

5. Srinivasan, K. (1998): Demographic Techniques and Applications. Sage Publications

6. Srivastava O.S. (1983): A Text Book of Demography. Vikas Publishing House.

**IV. Suggested Co-curricular Activities:**

1. Training of students by related industrial experts

2. Assignments including technical assignments if any.

3. Seminars, Group Discussions, Quiz, Debates etc on related topics.

4. Preparation of audio and videos on tools of diagrammatic and graphical representations.

5. Collection of material/figures/photos/author photoes of related topics.

6. Invited lectures and presentations of stalwarts to those topics.

7. Visits/field trips of firms, research organizations etc.

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER –V (CBCS, w.e.f 2024-25)**

**COURSE TITLE 14A: APPLIED STATISTICS**

Time: 3Hrs **Course Code: 24 – STA – 5C14A** Max. Marks: 75

**SECTION – A**

Answer any **ALL** of the following questions. Each question carries **2** Marks. **5 X 2 = 10**M

1. What is a mixed model in time series?
2. What is the free-hand curve method for trend estimation?
3. What is an index number?
4. What are vital statistics?
5. What is a Life Table?

**SECCTION - B**

Answer any **FIVE** of the following questions. Each question carries **5** Marks. **5 X 5 = 25**M

1. Define Time Series. Also write its Uses.
2. Explain the additive and multiplicative models of time series.
3. Explain the method of simple averages to determine seasonal variations.
4. What do you understand by the seasonal variations in time series? Give examples.
5. Explain various simple index numbers
6. Explain the Procedure to construct the cost of living index number.
7. Define Vital Statistics and write its Uses.
8. Define central mortality rate and Prove that $q\_{x}=\frac{2m\_{x}}{2+m\_{x}}$

**SECCTION -** C

Answer **FOUR** questions. Each question carries **10** Marks. **4 X 10 = 40 M**

1. Explain the method of fitting parabolic trend in time series analysis.
2. Explain the link relative method of computing the indices of seasonal variation.
3. Explain the Problems involved in the construction of Index numbers
4. Define Index Numbers. Explain various weighted index numbers.
5. Explain various rates of fertility.
6. Explain the construction and Uses of life table.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –V (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE – 14B: INTRODUCTION TO CORE JAVA AND ITS APPLICATIONS IN STATISTICS**

**No. of** Hours/week: 03 **24 – STA – 5C1**4B Credits: 03

**Learning Objectives for the Course**

1. Understand the foundational concepts of Core Java, including its features, environment setup, and programming syntax.
2. Apply Java’s collections framework for efficient data handling and implement statistical computations programmatically.
3. Create and customize statistical visualizations using Java libraries like JFreeChart.
4. Implement advanced statistical methods such as regression analysis and Monte Carlo simulations in Java.
5. Integrate Java applications with databases to manage large-scale datasets effectively.

**Learning Outcomes for the Course**

By the end of the course, students will be able to:

1. Write and execute Java programs using OOP principles for statistical modeling.
2. Perform data handling and basic statistical computations with Java’s collections framework.
3. Develop statistical charts and graphs for data visualization.
4. Apply Java to solve advanced statistical problems, including regression and error analysis.
5. Connect Java applications with databases for handling real-world datasets.

**Course Outcomes (COs)**

1. **CO1**: Explain the features and syntax of Core Java and demonstrate simple programs using control structures and OOP.
2. **CO2**: Utilize Java’s collections framework to perform data handling and basic statistical operations.
3. **CO3**: Develop visualizations for statistical data using Java libraries.
4. **CO4**: Implement regression models, simulations, and error analysis using Java.
5. **CO5**: Apply JDBC to connect Java programs to databases for data storage and retrieval.

**Program Outcomes (POs)**

1. **PO1**: Apply knowledge of programming and statistics to solve real-world problems.
2. **PO2**: Design, implement, and evaluate computational solutions for data-driven challenges.
3. **PO3**: Use modern tools and techniques for effective data analysis and visualization.
4. **PO4**: Demonstrate proficiency in problem-solving through programming.
5. **PO5**: Apply critical thinking and statistical reasoning to develop analytical solutions.
6. **PO6**: Collaborate effectively in multidisciplinary teams to complete projects.
7. **PO7**: Understand professional ethics and responsibilities in data handling and analysis.
8. **PO8**: Communicate technical concepts clearly through documentation and presentations.

**Program-Specific Outcomes (PSOs)**

1. **PSO1**: Develop efficient algorithms and software solutions for statistical computations and analysis.
2. **PSO2**: Apply Java programming techniques to model, simulate, and analyze real-world datasets.
3. **PSO3**: Integrate statistical tools with databases and visualization platforms for comprehensive data analysis.

**Mapping Summary Table**

| **CO** | **PO1** | **PO2** | **PO3** | **PO4** | **PO5** | **PO6** | **PO7** | **PO8** | **PSO1** | **PSO2** | **PSO3** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CO1** | 3 | 3 | 2 | 3 | 2 | 1 | 1 | 2 | 3 | 2 | 1 |
| **CO2** | 3 | 3 | 3 | 2 | 3 | 2 | 1 | 2 | 3 | 3 | 2 |
| **CO3** | 3 | 2 | 3 | 3 | 2 | 2 | 1 | 3 | 2 | 2 | 3 |
| **CO4** | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 3 | 3 | 3 |
| **CO5** | 3 | 3 | 2 | 2 | 3 | 2 | 3 | 3 | 2 | 3 | 3 |

**Legend: 3**: Strongly related, **2**: Moderately related, **1**: Slightly related

**Bloom’s Taxonomy Connectivity**

| **CO** | **Cognitive Level** | **Description** | **Mapped Activity** |
| --- | --- | --- | --- |
| **CO1** | Remember, Understand | Explain Java basics and demonstrate simple programming concepts. | Writing Java programs with basic syntax. |
| **CO2** | Apply, Analyze | Utilize collections for data handling and statistical operations. | Implementing statistical methods in Java. |
| **CO3** | Create, Apply | Develop statistical visualizations using Java libraries. | Designing bar charts, scatter plots, etc. |
| **CO4** | Analyze, Evaluate | Implement regression models and analyze simulation results. | Monte Carlo simulations and regression modeling. |
| **CO5** | Apply, Evaluate | Use JDBC for integrating Java applications with databases. | Real-world case studies and projects. |

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –V (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE – 14B: INTRODUCTION TO CORE JAVA AND ITS APPLICATIONS IN STATISTICS**

**No. of** Hours/week: 03 **24 – STA – 5C14B** Credits: 03

**II. Syllabus**

**UNIT – I**

**Introduction to Core Java**

**Basics of Java Programming:** Overview of Java: Features, JVM, JRE, and JDK. Setting up the development environment (Eclipse, IntelliJ IDEA). Writing and executing a simple Java program. **Core Java Syntax and Fundamentals:** Data types, variables, operators, and control structures (if-else, loops). Arrays and Strings. **Object-Oriented Programming (OOP):** Concepts: Classes, objects, inheritance, polymorphism, abstraction, and encapsulation. Importance of OOP in statistical modeling.

**UNIT – II**

**Data Handling and Statistical Computations**

**Collections Framework in Java:** Arrays, ArrayLists, HashMaps, and LinkedLists. Sorting and filtering datasets. **Basic Statistical Methods in Java:** Descriptive statistics: Mean, median, mode, variance, and standard deviation. Implementation of frequency distribution. **Probability Distributions:** Simulation of normal, binomial, and Poisson distributions using Java.

**UNIT – III**

**Data Visualization with Java**

**Introduction to Java Libraries for Visualization:** Overview of J-FreeChart for statistical plotting. **Creating Visual Representations:** Generating histograms, bar charts, and scatter plots. Customizing charts: Titles, legends, and color schemes. **Integration with Statistical Data:** Representing frequency distributions and regression models graphically.

**UNIT – IV**

**Advanced Statistical Applications in Java**

**Regression Analysis:** Implementing linear regression and calculating R-squared values. Basics of multiple regression modeling. **Statistical Simulations:** Monte Carlo simulations and bootstrapping techniques. Random sampling and hypothesis testing. **Error Analysis:** Handling statistical errors and interpreting results programmatically.

**UNIT – V**

**Database Integration and Real-World Applications**

**Java Database Connectivity (JDBC):** Connecting Java applications to relational databases. Querying, updating, and retrieving statistical datasets. **Big Data Handling:** Introduction to handling large-scale datasets with Java.

**Practical Sessions: Introduction to Core Java and Its Applications in Statistics**

**UNIT – I**

**Introduction to Core Java**

**Practical Sessions:**

1. Setting up Java Development Environment (Eclipse or IntelliJ IDEA).
2. Writing and executing a simple Java program to print statistical data.
3. Implementing control structures (if-else, loops) for basic data handling tasks.
4. Creating a class and object to calculate the sum and average of numbers.
5. Using arrays and strings to process a dataset.

**Reference Books:**

* "Core Java Volume I – Fundamentals" by Cay S. Horstmann and Gary Cornell
* "Java: The Complete Reference" by Herbert Schildt

**Textbooks:**

* "Programming with Java" by E. Balagurusamy
* "Introduction to Java Programming and Data Structures" by Y. Daniel Liang

**UNIT – II**

**Data Handling and Statistical Computations**

**Practical Sessions:**

1. Using arrays and Array Lists to store and retrieve data.
2. Writing a program to calculate descriptive statistics (mean, median, mode, variance, and standard deviation).
3. Implementing frequency distribution using Hash Map.
4. Simulating probability distributions (normal, binomial, and Poisson).
5. Sorting and filtering a dataset using Java collections.

**Reference Books:**

* "Java Programming for Engineers" by Julio Sanchez and Maria P. Canton
* "An Introduction to Statistical Learning" by Gareth James et al. (for statistical concepts)

**Textbooks:**

* "Think Java: How to Think Like a Computer Scientist" by Allen Downey and Chris Mayfield
* "Statistics for Beginners" by Zealure C. Holcomb

**UNIT – I**II

**Data Visualization with Java**

**Practical Sessions:**

1. Setting up JFree Chart library in Java.
2. Creating a bar chart to represent frequency distribution.
3. Generating scatter plots and histograms using statistical data.
4. Customizing charts (adding titles, legends, and color schemes).
5. Visualizing regression models graphically.

**Reference Books:**

* "Java Graphics and GUI Programming" by David M. Geary
* "Statistical Visualization in Java" by Jim X. Chen

**Textbooks:**

* "Java Programming and Object-Oriented Application Development" by Richard Johnson
* "Data Visualization with JavaScript" by Stephen A. Thomas (supplementary for visualization concepts)

**UNIT – IV**

**Advanced Statistical Applications in Java**

**Practical Sessions:**

1. Implementing linear regression in Java and calculating R-squared values.
2. Writing a program to perform multiple regression modeling.
3. Conducting Monte Carlo simulations for statistical problems.
4. Performing random sampling and hypothesis testing using Java.
5. Implementing error analysis and interpreting results programmatically.

**Reference Books:**

* "Java by Comparison: Become a Java Craftsman in 70 Examples" by Simon Harrer et al.
* "Monte Carlo Statistical Methods" by Christian P. Robert and George Casella

**Textbooks:**

* **"**Statistical Methods in Java" by Derek L. Beatty
* "Advanced Java Programming" by Uttam K. Roy

**UNIT – V**

**Database Integration and Real-World Applications**

**Practical Sessions:**

1. Setting up a relational database (MySQL or SQLite) for statistical datasets.
2. Using JDBC to connect a Java application to the database.
3. Querying and updating statistical datasets through Java programs.
4. Creating a program to handle large-scale datasets and represent them graphically.
5. Building a mini-project: Statistical analysis pipeline integrating database and visualization.

**Reference Books:**

* "Java Database Programming Bible" by John O'Donahue
* "Big Data Analytics with Java" by Raju Kumar Mishra

**Textbooks:**

* "Database Management Systems" by Raghu Ramakrishnan and Johannes Gehrke
* "Beginning Java and MySQL" by Jeffrey Hunter

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER –V (CBCS, w.e.f 2024-25)**

**COURSE TITLE: INTRODUCTION TO CORE JAVA AND ITS APPLICATIONS IN STATISTICS**

Time: 3Hrs **Course Code: 24 – STA – 5C14B** Max. Marks: 75

**SECTION – A**

Answer any **ALL** of the following questions. Each question carries **2** Marks. **5 X 2 = 10**M

1. Define JVM, JRE, and JDK.
2. List any four features of Core Java.
3. What is the significance of data types in Java?
4. Explain the term "inheritance" in Object-Oriented Programming.
5. How are Arrays and ArrayLists different in Java?

**SECTION – B**

**Answer any FIVE Questions. Each Question Carries 5 marks. 5X 5 = 25M**

1. Write a Java program to calculate the mean of a set of numbers using arrays.
2. Explain the control structures in Java with examples.
3. Describe the concept of polymorphism and provide a simple Java example.
4. Illustrate the use of Hash-Map for creating a frequency distribution in statistics.
5. Explain the steps to create and customize a histogram using J-FreeChart in Java.
6. Write a program to calculate the standard deviation of a dataset in Java.
7. Discuss the role of Monte Carlo simulation in statistical analysis and how it can be implemented in Java.
8. How is JDBC used to connect Java applications to a database? Provide an example.

**SECTION – C**

**Answer any FOUR Questions. Each Question Carries 10 marks. 4X 10 = 40M**

1. Write a Java program to implement linear regression and calculate the R-squared value for a given dataset.
2. Discuss the importance of Object-Oriented Programming (OOP) concepts such as encapsulation and abstraction in statistical modeling, with examples.
3. Create a Java application that integrates statistical data from a database using JDBC and displays it as a bar chart.
4. Explain probability distributions (normal, binomial, and Poisson) and write a Java program to simulate one of these distributions.
5. Develop a Java application to perform bootstrapping and hypothesis testing for a statistical dataset.
6. Explain the process of error analysis in statistical programming and implement a Java program to calculate the margin of error for a confidence interval.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –V (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE – 15A: OPERATIONS RESEARCH – II**

**No. of** Hours/week: 03 **24 – STA – 5C15A** Credits: 03

**Objective:** To enrich the knowledge of students with advanced techniques of linear programming problem along with real life applications.

**Learning Objectives:**

* Understand the significance of transportation, assignment, sequencing, game theory, and network scheduling problems in operations research.
* Learn how to formulate and solve transportation, assignment, and sequencing problems mathematically.
* Apply optimization techniques like MODI, Hungarian method, and Johnson’s algorithm to solve real-world problems.
* Understand game theory concepts such as pure and mixed strategies and apply them to solve 2x2 rectangular games.
* Analyze project scheduling problems using network scheduling techniques like CPM and PERT.

**Learning Outcomes:**

* Students will be able to formulate transportation and assignment problems and apply optimization techniques to find solutions.
* Students will gain the ability to solve sequencing problems using different algorithms for various machine configurations.
* Students will apply game theory strategies and solve 2x2 games using pure and mixed strategies.
* Students will demonstrate proficiency in calculating project durations using CPM and PERT methods.

**Course Outcomes (COs):**

1. **CO1:** Formulate transportation problems and solve using various methods like North-West Corner, Lowest Cost Entry, and Vogel’s Approximation Method.
2. **CO2:** Apply MODI Method to find the optimal solution for transportation problems and handle unbalanced and maximization transportation problems.
3. **CO3:** Formulate and solve assignment problems using the Hungarian method.
4. **CO4:** Solve sequencing problems for multiple machines using Johnson’s algorithm and other methods.
5. **CO5:** Apply game theory concepts like Max-mini, Mini-max, and saddle points to solve 2x2 games and formulate linear programming problems for larger games.

**Program Outcomes (POs):**

1. **PO1:** Understand and apply the basic concepts of operations research techniques.
2. **PO2:** Formulate and solve optimization problems in various real-life scenarios.
3. **PO3:** Demonstrate proficiency in mathematical modeling and problem-solving techniques.
4. **PO4:** Apply critical thinking and problem-solving strategies in decision-making.
5. **PO5:** Analyze and optimize complex problems using computational tools and techniques.
6. **PO6:** Develop the ability to work on team projects and communicate results effectively.
7. **PO7:** Understand the applications of game theory and optimization techniques in business and economics.
8. **PO8:** Demonstrate skills in project management using network scheduling methods like CPM and PERT.

**Program Specific Outcomes (PSOs):**

1. **PSO1:** Use operations research techniques to optimize processes in supply chain management, logistics, and production planning.
2. **PSO2:** Solve complex real-world problems by applying advanced optimization algorithms in transportation, assignment, sequencing, and game theory.
3. **PSO3:** Implement network scheduling techniques in project management to ensure efficient planning and execution of projects.

**Mapping Summary Table:**

| **Unit No.** | **COs** | **POs** | **PSOs** | **Bloom's Taxonomy Level** |
| --- | --- | --- | --- | --- |
| **Unit 1**: Transportation Problem | CO1, CO2 | PO1, PO2, PO3, PO5 | PSO1 | Apply, Analyze |
| **Unit 2**: Assignment Problem | CO3 | PO2, PO3, PO6 | PSO2 | Apply, Evaluate |
| **Unit 3**: Sequencing Problem | CO4 | PO3, PO5, PO8 | PSO3 | Apply, Analyze |
| **Unit 4**: Game Theory | CO5 | PO2, PO4, PO7 | PSO2 | Analyze, Evaluate |
| **Unit 5**: Network Scheduling | CO4 | PO5, PO8 | PSO3 | Apply, Analyze |

**Connectivity with Bloom’s Taxonomy Levels:**

* **Remembering**: Recall concepts like transportation problem formulation, assignment problem formulation, basic concepts of game theory, and CPM/PERT.
* **Understanding**: Explain optimization methods such as North-West Corner, Vogel’s Approximation, and the Hungarian method.
* **Applying**: Use algorithms like MODI and Johnson’s algorithm to solve real-world problems in transportation, assignment, and sequencing.
* **Analyzing**: Analyze problems to identify optimal solutions and strategies using game theory, and understand project scheduling techniques using CPM and PERT.
* **Evaluating**: Evaluate the effectiveness of different strategies like pure and mixed strategies in game theory and optimal solutions in network scheduling.
* **Creating**: Formulate complex optimization problems and propose solutions using appropriate methods.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –V (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE – 15A: OPERATIONS RESEARCH – II**

**No. of** Hours/week: 03 **24 – STA – 5C15A** Credits: 03

**II. Syllabus**

**UNIT – I**

**Transportation Problem**

**Introduction to Transportation Problem:** Overview and significance of the transportation problem in operations research, Real-world applications of the transportation problem. **Mathematical Formulation of Transportation Problem:** Formulating a transportation problem mathematically, Objective function and constraints in transportation problems. **Initial Basic Feasible Solution (IBFS):** Definition and importance of IBFS, Methods to find IBFS: **North-West Corner Rule**: Step-by-step process, **Lowest Cost Entry Method**: Explanation and application, **Vogel’s Approximation Method**: Concept and computational steps. **Optimal Solution of Transportation Problem:** Introduction to the optimality condition in transportation problems, **MODI Method (U-V Method)**: Detailed explanation and procedure for finding the optimal solution. **Unbalanced Transportation Problem:** Definition and handling of unbalanced transportation problems, Methods for balancing transportation problems before applying solution techniques. **Maximization of Transportation Problem:** Transforming a minimization transportation problem into a maximization problem, Steps for solving maximization transportation problems. **Problem-Solving and Practice:** Practical examples for computing IBFS using different methods, Exercises on applying MODI method for optimal solution, Solving unbalanced and maximization transportation problems

**UNIT – II**

**Assignment Problem**

**Introduction to Assignment Problem:** Overview and significance of the assignment problem in operations research, Real-world applications of assignment problems. **Mathematical Formulation of Assignment Problem:** Formulation of an assignment problem with objectives and constraints, Representation of the assignment problem as a cost minimization or profit maximization problem. **Reduction Theorem (Statement Only):** Statement of the reduction theorem for assignment problems, Purpose and application in simplifying assignment problem solutions. **Hungarian Method for Solving Assignment Problem:** Explanation and procedure of the Hungarian method for solving assignment problems, Steps for solving both balanced and unbalanced assignment problems using the Hungarian method. **Problem-Solving and Practice:** Practical examples on applying the Hungarian method to balanced and unbalanced assignment problems, Exercises on formulating and solving assignment problems with real-life scenarios

**UNIT – III**

**Sequencing Problem**

**Introduction to Sequencing Problem:** Overview of sequencing problems in operations research, Significance and real-world applications of sequencing problems. **Assumptions of Sequencing Problem:** Basic assumptions underlying sequencing problems, Conditions required for sequencing models. **Sequencing of n Jobs on One Machine:** Concept of sequencing for multiple jobs on a single machine, Mathematical formulation and approach for solving the problem. **Johnson's Algorithm for n Jobs on Two Machines:** Explanation of Johnson’s algorithm for sequencing jobs on two machines, Step-by-step procedure for applying Johnson’s algorithm, Solving sequencing problems with n jobs on two machines. **Algorithm for n Jobs on Three Machines:** Overview of the sequencing problem for n jobs on three machines, Step-by-step procedure for solving the three-machine problem, Solving sequencing problems with n jobs on three machines. **Algorithm for n Jobs on m Machines:** Concept and generalization of sequencing problems with n jobs on m machines, Steps involved in solving sequencing problems for multiple machines, Algorithm for sequencing n jobs on m machines. **Problem-Solving and Practice:** Practical examples of sequencing problems on one, two, three, and m machines, Exercises on applying Johnson’s algorithm and other sequencing techniques for various machine configurations

**UNIT – IV**

**Game Theory**

**Introduction to Game Theory:** Overview and significance of game theory in decision-making, Real-world applications of game theory in competitive scenarios. **Two-Person Zero-Sum Games:** Definition and concept of two-person zero-sum games, Explanation of zero-sum principle where the gain of one player equals the loss of the other. **Pure and Mixed Strategies:** Definition and distinction between pure and mixed strategies, Application of pure and mixed strategies in games. **Max-mini and Mini-max Principles:** Explanation of Max-mini principle (Player 1’s strategy), Explanation of Mini-max principle (Player 2’s strategy), Relationship between Max-mini and Mini-max principles. **Saddle Point and Its Existence:** Definition of a saddle point in a game matrix, Conditions for the existence of a saddle point. **Games without Saddle Point and Mixed Strategies:** Concept of games without a saddle point, how mixed strategies are used in such games to find optimal solutions. **Solution of 2 x 2 Rectangular Games:** Steps for solving 2 x 2 games using pure and mixed strategies, Example problems and solutions for 2 x 2 games. **Graphical Method of Solving 2 x n and m x 2 Games:** Explanation of the graphical method for solving 2 x n and m x 2 games, Visual representation of mixed strategy solutions. **Dominance Property:** Concept of dominance in game theory. **Matrix Oddment Method for n x n Games:** Introduction to the matrix oddment method for solving n x n games, Steps for applying the matrix oddment method. **Formulation of Linear Programming Problem for m x n Games:** Concept of formulating a linear programming problem to solve m x n games, Explanation of the linear programming approach for m x n games. **Problem-Solving and Practices.**

**UNIT – V**

**Network Scheduling**

**Introduction to Network Scheduling:** Overview of network scheduling in project management, Significance and applications of network scheduling techniques. **Basic Components of a Network:** Definition and explanation of the basic components of a network: nodes, arcs, events, and activities, Understanding the relationships between events and activities in a network. **Rules of Network Construction:** Principles and rules for constructing a network diagram, Types of network diagrams and their applications (e.g., Activity on Node (AON), Activity on Arrow (AOA)). **Time Calculations in Networks:** Explanation of time-related calculations in project scheduling, Techniques to calculate early start, early finish, late start, late finish, and slack times. **Critical Path Method (CPM):** Introduction to CPM and its importance in project management,Steps involved in constructing a CPM network, Calculation of the critical path and its role in determining project duration. **Program Evaluation and Review Technique (PERT):** Overview of PERT and its use in uncertain project timelines. Difference between PERT and CPM, Steps to develop a PERT network and calculate expected project duration, **Problem-Solving and Practices**

**Practical Syllabus Practical Credits: 1 (2 hrs/week)**

1. IBFS of transportation problem by using North- West corner rule, Matrix minimum method and VAM
2. Optimum solution to balanced and unbalanced transportation problems by MODI method (both maximization and minimization cases)
3. Solution of Assignment problem using Hungarian method (both maximization and minimization cases),
4. Solution of sequencing problem—processing of n jobs through two machines
5. To perform Project scheduling of a given project (Deterministic case-CPM).
6. To perform Project scheduling of a given project (Probabilistic case-PERT).
7. Solution of m x n games by dominance rule.

**III. References**

1. S.D. Sharma: Operations Research, Kedar Nath Ram Nath & Co, Meerut.
2. Kanti Swarup, P.K.Gupta, Manmohan: Operations Research, Sultan Chand and sons, New Delhi.
3. J.K. Sharma: Operations Research and Application, Mc. Millan and Company, New Delhi.
4. Gass: Linear Programming. Mc Graw Hill.
5. Hadly: Linrar Programming. Addison-Wesley.
6. Taha: Operations Research: An Introduction : Mac Millan.
7. Dr. NVS Raju: Operations Research, SMS education.

**IV. Suggested Co-curricular Activities:**

1. Training of students by related industrial experts
2. Assignments including technical assignments if any.
3. Seminars, Group Discussions, Quiz, Debates etc on related topics.
4. Preparation of audio and videos on tools of diagrammatic and graphical representations.
5. Collection of material/figures/photos/author photoes of related topics.
6. Invited lectures and presentations of stalwarts to those topics.
7. Visits/field trips of firms, research organizations etc.

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER –V (CBCS, w.e.f 2024-25)**

**COURSE TITLE: OPERATIONS RESEARCH – II**

**Time: 3 Hours Course Code: 24 – STA – 5C15A Max. Marks: 75**

**SECTION – A**

**Answer ALL the Five Questions. Each Question Carries 2 marks. 5 X 2= 10M**

1. Define Transportation Problem.
2. Define Assignment Problem.
3. What the use of artificial variables.
4. Define Sequencing Problem.
5. What are the Pure Strategy and Mixed Strategy

**SECTION – B**

**Answer any FIVE Questions. Each Question Carries 5 marks. 5X 5 = 25M**

1. Explain the procedure of North - West Corner rule.
2. Explain the algorithm to find IBFS to the TPP by using VAM.
3. Explain the procedure of Assignment problem - Hungarian method.
4. Explain the procedure of sequencing problem for two machines ‘n’ jobs.
5. Define idle time and total elapsed time.
6. Write the procedure for unbalanced assignment problem.
7. Explain the rules of Network construction.
8. Explain the method of project management.

**SECTION – C**

**Answer any FOUR Questions. Each Question Carries 10 marks. 4X 10 = 40M**

1. Solve the following transportation problem using North – West corner rule

|  |
| --- |
| **Sink** |
| **Origin** |  | B | C | D | E | F | **Supply** |
| P | 2 | 11 | 10 | 3 | 7 | 4 |
| Q | 1 | 4 | 7 | 2 | 1 | 8 |
| R | 3 | 9 | 4 | 8 | 12 | 9 |
| **Demand** |  | 3 | 3 | 4 | 5 | 6 |  |

1. Solve the following Assignment problem by using Hungarian method.

|  |  |
| --- | --- |
| **Jobs** | **Machines** |
|  | $$M\_{1}$$ | $$M\_{2}$$ | $$M\_{3}$$ | $$M\_{4}$$ | $$M\_{5}$$ |
| $$J\_{1}$$ | 9 | 22 | 58 | 11 | 19 |
| $$J\_{2}$$ | 43 | 78 | 72 | 50 | 63 |
| $$J\_{3}$$ | 41 | 28 | 91 | 37 | 45 |
| $$J\_{4}$$ | 74 | 42 | 27 | 49 | 39 |
| $$J\_{5}$$ | 36 | 11 | 57 | 22 | 25 |

1. Determine the optimal sequence of jobs that minimizes total based on the following information processing time on machines is given in hours and processing is not allowed.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Jobs** |  | **A** | **B** | **C** | **D** | **E** | **F** |
| **Processing time****in hrs** | **Machine – A** | 8 | 3 | 7 | 2 | 5 | 1 |
| **Machine - B** | 3 | 4 | 5 | 2 | 1 | 6 |
| **Machine - C** | 8 | 7 | 6 | 9 | 10 | 9 |

1. Explain the project elevation review technique for a project management.
2. Explain the procedure of draw a NETWORK diagram and draw the CRITICAL PATH.
3. Explain Two – Person Zero and n sum game.

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –V (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE: BIG DATA ANALYSIS THROUGH APACHE HADOOP**

Theory Credits: 3 **Course Code: 24 – STA – 5C15B** 3 hrs/week

**Learning Objectives**

1. To understand the fundamentals of Big Data and the Hadoop ecosystem.
2. To gain knowledge of Hadoop Distributed File System (HDFS) and its functionality.
3. To learn the MapReduce programming model for processing large datasets.
4. To apply tools like Hive and Pig for Big Data analytics.
5. To explore advanced Hadoop ecosystem tools and implement end-to-end Big Data projects.

**Learning Outcomes**

Upon completing the course, students will:

1. Understand Big Data concepts, challenges, and applications.
2. Demonstrate proficiency in setting up and managing Hadoop clusters.
3. Develop programs using MapReduce for large-scale data processing.
4. Utilize Hive and Pig for data querying and transformation.
5. Integrate Hadoop with other tools for advanced Big Data projects.

**Course Outcomes (COs)**

1. Explain the characteristics of Big Data and the architecture of the Hadoop ecosystem.
2. Demonstrate file management and fault-tolerance capabilities in HDFS.
3. Develop and execute MapReduce programs for data processing tasks.
4. Use HiveQL and Pig Latin for data analysis and transformation.
5. Implement a complete Big Data project using Hadoop tools.

**Program Outcomes (POs)**

1. **PO1:** Apply knowledge of computing and mathematics to Big Data analysis.
2. **PO2:** Analyze problems and identify computational solutions in Big Data contexts.
3. **PO3:** Design and implement solutions for complex data analysis tasks.
4. **PO4:** Use modern tools and techniques for Big Data processing.
5. **PO5:** Understand professional and ethical responsibilities in data analytics.
6. **PO6:** Communicate effectively through reports and visual data presentations.
7. **PO7:** Work collaboratively on multidisciplinary Big Data projects.
8. **PO8:** Engage in lifelong learning to stay updated on emerging Big Data technologies.

**Program Specific Outcomes (PSOs)**

1. **PSO1:** Apply tools and technologies from the Hadoop ecosystem to analyze large datasets.
2. **PSO2:** Develop efficient algorithms for data storage, processing, and analysis.
3. **PSO3:** Deliver solutions for real-world problems using Big Data analytics.

**Mapping Summary Table (CO-PO and CO-PSO)**

| **COs** | **PO1** | **PO2** | **PO3** | **PO4** | **PO5** | **PO6** | **PO7** | **PO8** | **PSO1** | **PSO2** | **PSO3** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CO1** | 3 | 3 | 2 | 2 | 1 | 1 | 1 | 2 | 3 | 2 | 1 |
| **CO2** | 3 | 3 | 3 | 3 | 2 | 1 | 2 | 2 | 3 | 3 | 2 |
| **CO3** | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 3 | 3 | 3 |
| **CO4** | 3 | 2 | 3 | 3 | 2 | 3 | 3 | 2 | 3 | 3 | 3 |
| **CO5** | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |

**Legend:**3 = Strong correlation,2 = Moderate correlation,1 = Low correlation

**Bloom’s Taxonomy Connectivity**

| **Level** | **Key Actions** | **Course Examples** |
| --- | --- | --- |
| **Remembering** | Define, List, Describe | Define the 5Vs of Big Data; List Hadoop ecosystem tools. |
| **Understanding** | Explain, Summarize, Classify | Explain HDFS architecture; Summarize the MapReduce workflow. |
| **Applying** | Implement, Use, Execute | Implement MapReduce programs; Use Hive for querying datasets. |
| **Analyzing** | Differentiate, Compare, Identify | Differentiate Hadoop components; Analyze data processing efficiency. |
| **Evaluating** | Justify, Assess, Validate | Assess the performance of MapReduce; Validate solutions for fault tolerance. |
| **Creating** | Design, Develop, Construct | Design an end-to-end Big Data project; Develop a sentiment analysis application. |

**P.V.K.N GOVT COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (with Mathematics Combination)**

**SEMESTER –V (Under CBCS, SYLLABUS with effect from A.Y 2024-25)**

**COURSE TITLE 15B: BIG DATA ANALYSIS THROUGH APACHE HADOOP**

Theory Credits: 3 **Course Code: 24 – STA – 5C15B** 3 hrs/week

**II. Syllabus**

**UNIT – I**

**Introduction to Big Data and Hadoop**

**Big Data Fundamentals:** Definition, characteristics (5Vs), and challenges of Big Data. Applications and significance of Big Data analytics. **Hadoop Overview:** History and evolution of Hadoop. Core components: HDFS and MapReduce. Hadoop ecosystem: Hive, Pig, HBase, Flume, Sqoop, and YARN. **Hadoop Architecture:** Master-slave architecture. Namenode, Datanode, Secondary Namenode, Job Tracker, and Task Tracker. **Hadoop Installation:** Single-node and multi-node cluster setup (basic overview).

**UNIT – II**

**Hadoop Distributed File System (HDFS)**

**HDFS Fundamentals:** Features of HDFS and its importance in big data storage. File system hierarchy and blocks. Read/write operations in HDFS. **HDFS Architecture:** Role of Namenode and Datanode. Data replication and fault tolerance. **File Management in HDFS:** Command-line interface for HDFS. Common HDFS commands for file operations. **Practical Lab:** Uploading, retrieving, and deleting files in HDFS.

**UNIT – III**

**Data Processing with Map Reduce**

**Map Reduce Framework:** Concept of Map Reduce: Mapper, Reducer, and Combiner. Workflow of MapReduce jobs. **Writing Map Reduce Programs:** Input and output formats (Text Input Format, Sequence File Input Format). Developing a basic word count program. **Advanced Map Reduce Concepts:** Partitioning, shuffling, and sorting. Counters and Joins in Map Reduce. **Practical Lab:** Implementing Map Reduce programs for real-world data processing tasks.

**UNIT – IV**

**Hive and Pig for Big Data Analytics**

**Apache Hive:** Introduction to Hive: Architecture and components. HiveQL: Basics of querying and data definition. Partitioning and Bucketing in Hive. Integrating Hive with HDFS. **Apache Pig:** Introduction to Pig: Architecture and components. Pig Latin scripting for data processing. Load, Transform, and Store operations. Use cases of Pig in Big Data analytics. **Practical Lab:** Querying structured data using Hive. Data transformations with Pig scripts.

**UNIT – V**

**Unit 5: Advanced Hadoop Ecosystem and Big Data Projects**

**YARN (Yet Another Resource Negotiator):** Overview of YARN architecture. Resource allocation and management. **Integrating Hadoop with Other Tools:** Apache HBase for NoSQL database integration. Data ingestion using Apache Flume and Sqoop. Basics of Spark integration with Hadoop. **Big Data Projects:** End-to-end project using HDFS, Map Reduce, and Hive/Pig. Use cases: Log analysis, sentiment analysis, or click stream data processing. **Trends and Future of Big Data Analytics:** Emerging tools and technologies. Ethical considerations and challenges in Big Data.

**Practical Sessions: Big Data Analysis through Apache Hadoop**

**UNIT – I**

***Introduction to Big Data and Hadoop***

**Practical 1: Installing Hadoop on a Single Node and Multi-node Cluster**

* **Objective:** To set up and configure a single-node and multi-node Hadoop cluster.
* **Tools Required:** Hadoop, VirtualBox (or equivalent), Ubuntu Linux, SSH.
* **Steps:**
	1. Install Hadoop on a single-node cluster.
	2. Set up a multi-node Hadoop cluster.
	3. Run a simple Hadoop job to ensure the cluster is set up correctly.

**Practical 2: Exploring the Hadoop Ecosystem Components (Hive, Pig, HBase, etc.)**

* **Objective:** To explore and set up the basic components of the Hadoop ecosystem like Hive, Pig, and HBase.
* **Tools Required:** Hive, Pig, HBase.
* **Steps:**
	1. Install and configure Apache Hive.
	2. Set up Apache Pig and run a simple script.
	3. Create a table in HBase and insert data.

**Reference Books:**

* "Hadoop: The Definitive Guide" by Tom White
* "Hadoop in Action" by Chuck Lam
* "Hadoop Operations" by Eric Sammer

**UNIT – I**I

***Hadoop Distributed File System (HDFS)***

**Practical 1: Managing Files in HDFS**

* **Objective:** To perform basic file operations such as uploading, retrieving, and deleting files in HDFS.
* **Tools Required:** Hadoop, HDFS CLI.
* **Steps:**
	1. Upload files from local to HDFS using hdfsdfs -put.
	2. Retrieve files from HDFS using hdfsdfs -get.
	3. Delete files using hdfsdfs -rm.

**Practical 2: Understanding Fault Tolerance and Data Replication in HDFS**

* **Objective:** To simulate data replication and the fault tolerance mechanism of HDFS.
* **Tools Required:** Hadoop.
* **Steps:**
	1. Run a job that intentionally crashes a DataNode.
	2. Observe HDFS replication and fault tolerance.
	3. Check the status of replication with the command hdfsdfsadmin -report.

**Reference Books:**

* "Hadoop: The Definitive Guide" by Tom White
* "HDFS Architecture and Design" by Shveta Chawla
* "Hadoop in Practice" by Alex Holmes

**UNIT – III**

***Data Processing with MapReduce***

**Practical 1: Writing a Word Count Program Using MapReduce**

* **Objective:** To implement a basic MapReduce word count program.
* **Tools Required:** Hadoop, Eclipse or IntelliJ IDEA.
* **Steps:**
	1. Write Mapper, Reducer, and Driver classes.
	2. Compile and run the program in the Hadoop cluster.
	3. Analyze the output file for word frequency.

**Practical 2: Advanced MapReduce - Using Partitioner and Combiner**

* **Objective:** To optimize MapReduce jobs by implementing partitioners and combiners.
* **Tools Required:** Hadoop.
* **Steps:**
	1. Create a custom Partitioner to ensure data is distributed evenly across reducers.
	2. Implement a Combiner to minimize data transfer between mapper and reducer.

**Practical 3: Handling Large Datasets in MapReduce**

* **Objective:** To process large datasets efficiently using Hadoop’s MapReduce.
* **Tools Required:** Hadoop, Dataset (real-world or synthetic).
* **Steps:**
	1. Write a MapReduce job to process a large dataset.
	2. Analyze the time taken and resource consumption.

**Reference Books:**

* "Hadoop in Practice" by Alex Holmes
* "Hadoop: The Definitive Guide" by Tom White
* "MapReduce Design Patterns" by Donald Miner and Adam Shook

**UNIT - IV**

***Hive and Pig for Big Data Analytics***

**Practical 1: Writing Basic Hive Queries**

* **Objective:** To write basic Hive QL queries for data management and analysis.
* **Tools Required:** Apache Hive, Hadoop.
* **Steps:**
	1. Install and configure Hive.
	2. Create tables in Hive and load data from HDFS.
	3. Run SELECT, WHERE, and GROUP BY queries.

**Practical 2: Data Transformation with Apache Pig**

* **Objective:** To process data using Pig Latin scripts for data transformation.
* **Tools Required:** Apache Pig, Hadoop.
* **Steps:**
	1. Install and configure Pig.
	2. Write a Pig Latin script to filter and transform a dataset.
	3. Store the processed data back in HDFS.

**Practical 3: Partitioning and Bucketing in Hive**

* **Objective:** To use partitioning and bucketing techniques in Hive for optimized querying.
* **Tools Required:** Apache Hive, Hadoop.
* **Steps:**
	1. Create partitioned and bucketed tables in Hive.
	2. Load data into these tables and perform queries.

**Reference Books:**

* "Programming Hive" by Edward Capriolo
* "Pig Programming" by Alan Gates
* "Hadoop for Data Science" by Jason Brownlee

**UNIT - V**

***Advanced Hadoop Ecosystem and Big Data Projects***

**Practical 1: Integrating HBase with Hadoop**

* **Objective:** To store and retrieve real-time data using HBase with Hadoop.
* **Tools Required:** HBase, Hadoop.
* **Steps:**
	1. Install and configure HBase.
	2. Integrate HBase with Hadoop to store and retrieve data.
	3. Perform basic operations like inserting and retrieving data from HBase using Java API.

**Practical 2: Building a Real-Time Data Processing Application with Hadoop and Spark**

* **Objective:** To integrate Spark with Hadoop for real-time data processing.
* **Tools Required:** Apache Spark, Hadoop.
* **Steps:**
	1. Set up Spark and Hadoop integration.
	2. Process real-time data (e.g., social media feeds or logs) using Spark.

**Practical 3: Using Sqoop and Flume for Data Ingestion**

* **Objective:** To ingest data into Hadoop from external sources.
* **Tools Required:** Apache Sqoop, Apache Flume, Hadoop.
* **Steps:**
	1. Use Sqoop to import data from a MySQL database to HDFS.
	2. Use Flume to ingest streaming data into HDFS.

**Practical 4: End-to-End Big Data Project**

* **Objective:** To implement an end-to-end big data project using HDFS, MapReduce, Hive, and Pig.
* **Tools Required:** Hadoop, HDFS, MapReduce, Hive, Pig.
* **Steps:**
	1. Choose a use case (e.g., sentiment analysis, log analysis, or clickstream analysis).
	2. Process the data using HDFS for storage, MapReduce for processing, Hive for querying, and Pig for data transformations.
	3. Present the results.

**Reference Books:**

* "HBase: The Definitive Guide" by Lars George
* "Big Data Analytics with Hadoop" by Vignesh Prajapati
* "Data Science for Business" by Foster Provost and Tom Fawcett

**Additional Reference Textbooks**

* "Hadoop in Action" by Chuck Lam
* "Hadoop: The Definitive Guide" by Tom White
* "Big Data: Principles and Paradigms" by Rajkumar Buyya and S. Thamarai Selvi
* "Big Data Analytics with Hadoop" by Vignesh Prajapati

**P.V.K.N GOVT. COLLEGE (AUTONOMOUS), CHITTOOR**

**DEPARTMENT OF STATISTICS**

**B.Sc., STATISTICS (W.M)**

**SEMESTER –V (CBCS, w.e.f 2024-25)**

**COURSE 15B: BIG DATA ANALYSIS THROUGH APACHE HADOOP**

Time: 3Hrs **Course Code: 24 – STA – 5C15B** Max. Marks: 75

**SECTION – A**

**Answer ALL the Five Questions. Each Question Carries 2 marks. 5 X 2= 10M**

1. What are the five characteristics of Big Data?
2. Explain the role of the Name Node in Hadoop Distributed File System (HDFS).
3. What is the purpose of the Combiner in a Map Reduce job?
4. Describe the key difference between Apache Pig and Apache Hive.
5. What is YARN in Hadoop? Explain its role in resource management.

**SECTION – B**

**Answer any FIVE Questions. Each Question Carries 5 marks. 5X 5 = 25M**

1. Discuss the core components of the Hadoop ecosystem and their roles in Big Data analysis.
2. Explain the architecture of Hadoop. What is the function of the Job Tracker and TaskTracker?
3. Describe how data is stored and managed in HDFS. Explain the significance of data replication.
4. Write a Map Reduce program for word count. Explain each component in the program.
5. Explain the difference between Map Reduce and Hive for Big Data processing.
6. Describe the process of querying data in Hive. How does partitioning improve query performance in Hive?
7. How does Pig simplify data processing tasks in Hadoop? Provide examples of common Pig operations.
8. Discuss the challenges of handling Big Data and how Hadoop addresses these challenges.

**SECTION – C**

**Answer any FOUR Questions. Each Question Carries 10 marks. 4X 10 = 40M**

1. Explain the concept of HDFS and its architecture. How does the data replication mechanism in HDFS ensure fault tolerance?
2. Discuss the steps to install and configure a Hadoop single-node and multi-node cluster. What are the challenges in setting up a multi-node Hadoop cluster?
3. Write and explain a complete Map Reduce program for sorting a list of integers. Describe how partitioning and shuffling work in the process.
4. Describe the architecture of Apache Hive. How does it facilitate data warehousing in Hadoop? Provide an example of creating tables and querying data in Hive.
5. Explain the role of YARN in Hadoop's resource management. How does it improve the efficiency of Big Data processing?
6. Provide a detailed overview of an end-to-end Big Data project using Hadoop. Discuss the integration of HDFS, Map Reduce, Hive, and Pig for real-world data processing.