

|                                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------|-----------------------|---------------------------------|
| <b>Semester: 4</b>                                                                                                                                                                                               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| <b>Programme: Data Science</b>                                                                                                                                                                                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| <b>Course: Inferential Statistics</b>                                                                                                                                                                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| <b>Paper code: C2DS250421T</b>                                                                                                                                                                                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               | <b>Credits: 4</b>     |                                 |
| <b>Hours/week: 4</b>                                                                                                                                                                                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| <b>Category: Core/MDC/SEC/VAC: Core</b>                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| <b>Theory / Practical / Composite: Theory</b>                                                                                                                                                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| <b>No of Module: 1</b>                                                                                                                                                                                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
|                                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| <b>Course Outcome:</b>                                                                                                                                                                                           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| 1. <b>Remember</b> fundamental definitions and properties of sampling distributions (Chi-square, t, F), estimation criteria (sufficiency, unbiasedness, consistency), and the terminology of hypothesis testing. |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| 2. <b>Understand</b> the conceptual paradigms of statistical inference, the implications of the Central Limit Theorem in data science, and the structural logic underlying the Gauss-Markov model.               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| 3. <b>Apply</b> point and interval estimation methods, such as Maximum Likelihood and the Method of Moments, to derive parameters and confidence intervals for real-life datasets.                               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| 4. <b>Analyze</b> experimental designs and business data by performing one-way and two-way ANOVA and investigating the estimability of parametric functions within linear models.                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| 5. <b>Evaluate</b> the validity of statistical models and the strength of experimental results using p-values, power functions, and non-parametric goodness-of-fit tests like the Kolmogorov-Smirnov test.       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| 6. <b>Create</b> robust inferential frameworks by synthesizing Monte-Carlo simulations and integrated sampling distribution theories to model and predict complex real-world data fluctuations.                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| <b>Prerequisites: Basic knowledge about any prior course</b>                                                                                                                                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| <b>SYLLABUS</b>                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                               |                       |                                 |
| <b>UNIT</b>                                                                                                                                                                                                      | <b>CONTENT</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | <b>HOURS or<br/>NUMBER<br/>OF<br/>CLASSES</b> | <b>CO<br/>Mapping</b> | <b>COGNI<br/>TIVE<br/>LEVEL</b> |
| 1.                                                                                                                                                                                                               | <b>Sampling Distributions and Simulation</b><br>Concept of i.i.d random variables, parameters & statistics. Sampling fluctuations and sampling distributions of statistics. Central and non-central Chi-square, t, and F distributions (with applications in data analysis). Sampling distribution of sample mean and variance & their independence under normality assumption. Sampling distribution related to simple linear regression. Central Limit Theorem and its implications in data science. Concept | 15                                            | CO1,<br>CO2,<br>CO6   | K1, K2,<br>K6                   |

|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |    |          |        |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|----------|--------|
|    | of Monte-Carlo Simulation and Applications                                                                                                                                                                                                                                                                                                                                                                                                                              |    |          |        |
| 2. | <b>Fundamentals of Statistical Inference</b><br>Paradigms: Frequentist vs Bayesian.<br>Parametric vs Non-Parametric.<br>Estimation vs Hypothesis testing (concepts and illustrations).                                                                                                                                                                                                                                                                                  | 2  | CO2      | K2     |
| 3. | <b>Estimation</b><br>Point & interval estimation. Criteria for a good estimator: sufficiency, unbiasedness, minimum variance, efficiency, consistency (concept and application). Accuracy vs precision in practice. Estimation methods: method of moments and maximum likelihood. Confidence interval & coverage. Confidence intervals of mean and variance for univariate and bivariate normal population. Approximate confidence interval for binomial proportion(s). | 10 | CO1, CO3 | K1, K3 |
| 4. | <b>Hypothesis Testing</b><br>Simple & Composite hypothesis. Null and Alternative hypothesis. Type I and II errors, level of significance, size of a test. Power and power function. p-value and its interpretation. Tests of mean and variance of univariate—and bivariate normal distribution. Tests of proportion. Likelihood ratio test. Chi-square tests for categorical data in contingency tables.                                                                | 10 | CO1, CO5 | K1, K5 |
| 5. | <b>Linear Models and ANOVA</b><br>Gauss-Markov model, parametric functions, estimability, concept and application of Best Linear Unbiased Estimator. One-way and two-way ANOVA (fixed effects models). Applications in experimental design, business & biological studies.                                                                                                                                                                                              | 9  | CO2, CO4 | K2, K4 |
| 6. | <b>Non-Parametric Methods</b><br>Motivation for non-parametric inference in data science. Runs test for randomness. Empirical distribution function. Tests for location and scale: Sign test & Mood's test. Goodness-of-                                                                                                                                                                                                                                                | 6  | CO5      | K5     |

|                                                                                                                                                                                                                                    |                                                                                         |  |                                |  |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------|--|--------------------------------|--|
|                                                                                                                                                                                                                                    | fit tests: Kolmogorov-Smirnov test, Q-Q plot, Shapiro-Wilk test, Anderson-Darling test. |  |                                |  |
| <b>Text Books</b>                                                                                                                                                                                                                  |                                                                                         |  |                                |  |
| 1. Rohatgi V.K. and Saleh, A. K. Md , E. (2009): <i>An Introduction to Probability and Statistics</i> , 2nd edition (Reprint), John Wiley and Sons.                                                                                |                                                                                         |  |                                |  |
| 2. Johnson, R.A. and Bhattacharya, G.K. (2001): <i>Introduction to the theory of Statistics</i> , 3rd edition (Reprint). Tata McGraw-Hill Pub. Co. Ltd.                                                                            |                                                                                         |  |                                |  |
| 3. Casella, G. & Berger, R.L. (2021): <i>Statistical Inference</i> . Cengage Learning.                                                                                                                                             |                                                                                         |  |                                |  |
| 4. Faraway, J. J. (2014). Linear models with r. crc press. <i>Boca Raton, Florida</i> .                                                                                                                                            |                                                                                         |  |                                |  |
| <b>Suggested readings</b>                                                                                                                                                                                                          |                                                                                         |  |                                |  |
| 1. Ismay, C., & Kim, A. Y. (2019). <i>Statistical inference via data science: a modern dive into R and the Tidyverse</i> . Chapman and Hall/CRC.                                                                                   |                                                                                         |  |                                |  |
| 2. Moulin, P., & Veeravalli, V. V. (2019). <i>Statistical inference for engineers and data scientists</i> . Cambridge University Press.                                                                                            |                                                                                         |  |                                |  |
| 3. Caffo, B. (2016). Statistical inference for data science. <i>British Columbia, UK: Leanpub</i> .                                                                                                                                |                                                                                         |  |                                |  |
| <b>Online Resources</b>                                                                                                                                                                                                            |                                                                                         |  |                                |  |
| 1. <a href="https://www.coursera.org/learn/inferential-statistics">https://www.coursera.org/learn/inferential-statistics</a>                                                                                                       |                                                                                         |  |                                |  |
| 2. <a href="https://ocw.mit.edu/courses/18-650-statistics-for-applications-fall-2016/video_galleries/lecture-videos/">https://ocw.mit.edu/courses/18-650-statistics-for-applications-fall-2016/video_galleries/lecture-videos/</a> |                                                                                         |  |                                |  |
|                                                                                                                                                                                                                                    |                                                                                         |  |                                |  |
|                                                                                                                                                                                                                                    |                                                                                         |  |                                |  |
| <b>Evaluation</b>                                                                                                                                                                                                                  | Theory CIA: 30<br>Semester Exam: 70<br>Total: 100                                       |  |                                |  |
| <b>Paper Structure for Theory Semester Exam Module :</b>                                                                                                                                                                           | Short questions (5 marks each)                                                          |  | Long questions (15 marks each) |  |
|                                                                                                                                                                                                                                    | 5 out of 7                                                                              |  | 3 out of 5                     |  |

### Course outcomes (COs) and Cognitive Level Mapping

| COs | CO Description                                                                                                                                                                                                | Cognitive levels |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|
| CO1 | <b>Remember</b> fundamental definitions and properties of sampling distributions (Chi-square, t, F), estimation criteria (sufficiency, unbiasedness, consistency), and the terminology of hypothesis testing. | K1               |
| CO2 | <b>Understand</b> the conceptual paradigms of statistical inference, the implications of the Central Limit Theorem in data science, and the structural logic underlying the Gauss-Markov model.               | K2               |
| CO3 | <b>Apply</b> point and interval estimation methods, such as Maximum Likelihood and the Method of Moments, to derive parameters and confidence intervals for real-life datasets.                               | K3               |
| CO4 | <b>Analyze</b> experimental designs and business data by performing one-way and two-way ANOVA and investigating the estimability of parametric functions within linear models.                                | K4               |

|            |                                                                                                                                                                                                         |    |
|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>CO5</b> | <b>Evaluate</b> the validity of statistical models and the strength of experimental results using p-values, power functions, and non-parametric goodness-of-fit tests like the Kolmogorov-Smirnov test. | K5 |
| <b>CO6</b> | <b>Create</b> robust inferential frameworks by synthesizing Monte-Carlo simulations and integrated sampling distribution theories to model and predict complex real-world data fluctuations.            | K6 |