Statistical Methods for Decision Making

Course Description :

Statistical Methods for Decision Making is a branch of applied statistics that emerged in response to machine learning, emphasizing statistical models and assessment of uncertainty. It is important to understand the ideas behind the various techniques, in order to know how and when to use them. One has to understand the simpler methods first, in order to grasp the more sophisticated ones. It is important to accurately assess the performance of a method, to know how well or how badly it is working. Additionally, this is an exciting research area, having important applications in science, industry, and finance. Ultimately, Statistical Methods for Decision Making is a fundamental ingredient in the training of a modern data scientist. This course on statistics will work as a foundation for future Data Science and Machine Learning Courses in this Program.

Learning Outcomes:

After completing this course, you will be able to: • Get a good appreciation of application of statistics in analysis • Get familiar with probability and different probability distributions • Laws of probability, Conditional probability & Bayes theorem • Understand the importance of hypothesis approach and its application • Realize that the hypothesis testing is a must when it comes to problem solving • Perform Test of Hypothesis as well as calculate confidence interval for a population parameter for single sample and two sample cases.

Pedagogy : The course is a mixture of classroom lectures, quizzes, assignments and miniprojects. Jupyter notebook will be the medium of coding in python.

Course Content
Session – 1
Introduction to Statistics • Why Statistics • Statistical Methods • Types of Statistics - Descriptive and Inferential Statistics • Data Sources and Types of Datasets • Attributes of Datasets • Frequency Distribution – Histograms • Cumulative Frequency Distribution • Measures of Central Tendency - Mean, Median, Mode • Measures of Dispersion - Range, IQR, Standard Deviation, coefficient of variation Normal distribution, Chebyshev Rule • Five number summary, boxplots, QQ plots, Quantile plot, scatter plot. Visualization: scatter plot matrix. • Correlation analysis

Session – 2 : Probability and Bayes Theorem • Probablity and Distributions Outline • Probability - Meaning and concepts • Probability - Meaning and concepts • Rules for Computing Probability • Marginal Probability and Example • Bayes' theorem and Example

Session – 3 : Probability Distributions and Sampling • Introduction to Probability distributions • Discrete and continuous distributions • Introduction to the Normal distribution • Sampling distribution and population parameters • Comparing sample and population - The concept of testing hypotheses Sampling error • Defining a confidence interval

Session – 4 Hypothesis Testing • Introduction to hypothesis testing • Defining a null and alternate hypothesis • Types of alternate hypothesis - One tail vs two tail test • Type 1 and type 2 error • Hypothesis testing applications using the z test • Interpreting test results • P-value vs confidence interval approach

Session – 5 Hypothesis Testing Contd.. • Hypothesis testing for small samples - the t distribution • Two sample t-test – unpaired and paired t-test • Test of proportion

Session – 6 ANOVA & ChiSquare Analysing variance between and within groups • Simple one way ANOVA • Chi Square • Intro to the Chi Square test • Tests of independence • Goodness of fit

FAQ:

1. Is this theoretical or with hands-on?
Ans: Theory is important to understand the concepts and different scenario. Hands-on using python is also covered as part of most of the sessions.

2. What are the assessment components for this course?
Ans: Quiz – 40 marks, Graded Assessment – 40 marks, Mandatory assignments – 10 marks, mini project – 20 marks, Attendance – 10 marks (Total 120 which will be scaled down to 100).

3. What is the pattern of End Semester Assessment (ESA)?
Ans: ESA is both handwritten (20 marks) and coded – python based (80) marks.

4. Is there a group assignment?
Ans: There is no group assignment for this course.

5. Are there any textbooks to supplement the material shared?
Ans:
   1. Aczel, Amir, and Jayavel Sounderpandian. *Complete Business Statistics.* McGraw-Hill Education, 2012.
   2. Montgomery, Douglas C., and George C. Runger. *Applied Statistics and Probability for Engineers.* 7th ed. 2018.
   3. Downey, Allen B. *Think Stats: Exploratory Data Analysis in Python.* O'Reilly Media, 2014.
   4. Levine, David M., Kathryn A. Szabat, and David F. Stephan. *Business Statistics: A First Course.* 8th ed. 2021.