Stats 02: Introduction to Linear Regression Models

Dr Amar Ahmad

Course Description:

This course provides an in-depth exploration of linear regression models using the R programming language. Linear regression is a fundamental statistical technique for modeling the relationship between a dependent variable and one or more independent variables. Students will learn how to implement linear regression models in R, from data preparation to interpretation of results. Practical exercises and real-world examples will be used to reinforce the concepts learned.

Prerequisites:

- Basic knowledge of statistics and data analysis.
- Familiarity with the statistical software R.

Course Objectives:

By the end of this course, students should be able to:

1. Understand the principles of linear regression and its applications.
2. Perform data preparation and preprocessing for linear regression analysis.
3. Build, evaluate, and interpret simple and multiple linear regression models.
4. Identify and address common issues in linear regression analysis.
5. Use R to visualize and present regression results effectively.

Course Outline:

Module 1: Introduction to Linear Regression model
- Definition and importance of linear regression.
- Types of linear relationships.
- Key terminology: response variable, predictor variable, coefficients.

Module 2: Simple Linear Regression
- Formulating and estimating simple linear regression models.
- Model interpretation and hypothesis testing.
- Visualization of regression results.

Module 3: Multiple Linear Regression
- Extending linear regression to multiple predictors.
- Model building and selection. Interpretation of coefficients.

Module 4: Model Diagnostics
- Residual analysis.
- Detection and handling of outliers.
- Assumption checking: normality and homoscedasticity.

Module 5: Model Interpretation
- Interpreting coefficients in context.
- Practical significance vs. statistical significance.
- Model limitations and caveats.

Module 6: Real-World Applications
- Application to real-world datasets.
- Case studies and examples from various fields.

Assessment Methods:
1. Quizzes and Homework Assignments: Assessing understanding of theoretical concepts.
2. Final Exam: Covering material from the entire course.
3. Group Projects: Applying hypothesis testing to real-world data analysis.
4. Class Participation: Active engagement in discussions and activities.

Grading:
- Quizzes and Homework Assignments: 30%
- Final Exam: 30%
- Group Projects: 30%
- Class Participation: 10%

Textbook:
- Christian Heumann, Michael Schomaker, Shalabh. Introduction to Statistics and Data Analysis with Exercises, Solutions and Applications in R. ISBN 978-3-031-11832-6, ISBN 978-3-031-11833-3 (eBook) https://doi.org/10.1007/978-3-031-11833-3