Stat 03: Generalized Linear Models (GLMs)

Dr Amar Ahmad

Course Description:

This course provides a comprehensive exploration of Generalized Linear Models (GLMs), a powerful and versatile statistical modeling framework. GLMs extend the concept of linear regression to a wide range of data types, including binary outcomes, count data, and categorical data. Students will learn the theory, application, and implementation of GLMs, as well as model diagnostics and interpretation. Practical exercises and real-world examples will be used to reinforce the concepts learned.

Prerequisites:

- Basic understanding of statistics and linear regression model (Stat 02).
- Familiarity with the statistical software R.

Course Objectives:

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

1. Understand the principles and theory behind Generalized Linear Models.
2. Identify suitable GLMs for various types of response variables.
3. Build, estimate, and interpret GLMs using statistical software.
4. Perform model diagnostics and validate GLMs.
5. Apply GLMs to analyze and interpret real-world data.

Course Outline:

Module 1: Introduction to Generalized Linear Models

- Overview of GLMs and their applications.
- Link functions and the canonical link.
- The exponential family of distributions.

Module 2: Binary and Multinomial Logistic Regression

- Binary logistic regression for binary outcomes.
- Multinomial logistic regression for categorical outcomes.
- Model interpretation and odds ratios.

Module 3: Poisson, quasi-Poisson and Negative Binomial Regression

- Poisson regression for count data.
- Quasi-Poisson regression for count data.
- Negative binomial regression for overdispersed count data.
- Model interpretation and rate ratios.

Module 4: Gamma and Inverse Gaussian Regression
- Gamma regression for continuous, positive outcomes.
- Inverse Gaussian regression for skewed continuous outcomes.
- Model interpretation and mean predictions.

Module 5: Model Diagnostics and Validation
- Residual analysis and deviance statistics.
- Overdispersion and underdispersion.
- Goodness-of-fit tests.

Module 6: Real-World Applications
- Applying GLMs 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
- Harrell, F.E. (2001) Regression Modeling Strategies With Applications to Linear Models, Logistic Regression, and Survival Analysis.