East China Normal University
GEC International Summer School

MAT 37: Linear Regression Analysis

Term: June 17th to July 19th, 2024
Class Hours: Monday through Friday, 110 minutes each day (2,750 minutes in total)
Instructor: TBD
Home Institution: TBD
Office hours: TBD
Email: TBD

Course Description

Modelling is the key elements in mathematics and statistics, linear regression plays a key role in the concept of modelling, and this course will generalize the simple linear regression to a broader methods in linear regression. Students are expected to gain a thorough knowledge in intermediate linear regression modelling.

Prerequisite: MAT12 Calculus 2, MAT21 Linear algebra, and elementary level statistics/modelling

Course Overview

This course is an intermediate statistic course on topics of applied linear regressions, covering multiple regression, corresponding estimation, diagnosis test, validation, lack of fitness test, variable selections. The focus is the statistical theory behind, with an introductory level of statistical software in SAS and R.

Learning Objective

This course is a study of Ordinary Differential Equations (ODE's), including modeling physical systems.

Topics include:

- Simple Linear Regression;
- Multiple Linear Regression;
- Model Adequacy Checking;
- Transformations and Weighting To Correct Model Inadequacies;
- Diagnostics For Leverage and Influence;
- Indicator Variables.

Required Text

Introduction to Linear Regression Analysis, 6th Edition, Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining.

ISBN: 978-1-119-57875-8

Course Hours

The course has 25 class sessions in total. Each class session is 110 minutes in length, for a total of 2750 minutes of in-class time. The course meets from Monday to Friday. ECNU awards 3 credits for this course. Different universities may count course credits differently. Consult officials at your own home institution.

Attendance

Summer school is very intense and to be successful, students need to attend *every class*. Occasionally, due to illness or other unavoidable circumstance, a student may need to miss a class. ECNU policy requires a medical certificate to be excused. Any absence may impact on the student's grade. Moreover, ECNU policy is that a student who has more than 3 absences will fail the course. Arriving late or leaving early will count as a partial absence.

Grading Policy

ECNU awards grades of A, A-, B+, B, B-, C+, C, D, and F. Most colleges and universities do not award transfer credit for grades of D or F.

In this course, grading will be based on the following:

Assignments*5	5%*5=25%
Midterm Exam	35%

Final Exam   40%

General Expectations

Students are expected to:
- Attend all classes and be responsible for all material covered in class and otherwise assigned. Any unexcused absence may impact a student's grade.
- Arrive to class on-time: Late arrivals are disruptive to your fellow students and to the conduct of the class.
- Complete the day's required reading and assignments before class
- Review the previous day's notes before class; make notes about questions you have about the previous class or the day's reading
- Refrain from texting, phoning or engaging in computer activities unrelated to class during class (不要用手机) It is highly disrespectful to the professor and to the class.
- Participate in class discussions and complete required written work on time.

Course Schedule

The planned schedule sketched out below may be modified to suit the interests or abilities of the enrolled students or to take advantage of special opportunities or events that may arise during the term.

Week 1
- *Day 1*
  - *Review on simple linear regression*
  - *Least-square methods and relationship with MLE*
  - *Covariance and correlation*
  - *Hypothesis testing and prediction*
- *Day 2*
  - *Multiple linear regression*
  - *Matrix and Geometrical view of least square methods*
- *Day 3*
  - *Introduction to statistical software, R, SAS, Python etc.*
  - *Implementation of basic linear models in SAS and R*
- *Day 4*
  - *More on assumptions and asymptotic properties*

- *o Confidence interval of coefficients*
- *Day 5*
  - *o In-class exercises*
  - *o Assignment 1 due*

Week 2

- *Day 1*
  - *o Hypothesis testing: coefficient and mean response*
- *Day 2*
  - *o Hypothesis testing: residuals and assumptions*
- *Day 3*
  - *o Multi-collinearity and Heteroscedasticity*
- *Day 4*
  - *o Standardized Regression coefficients*
- *Day 5*
  - *o Implementation of hypothesis testing in SAS and R*

Week 3

- *Day 1*
  - *o Residual analysis*
- *Day 2*
  - *o Using SAS, and R for residual analysis*
- *Day 3*
  - *o Midterm*
- *Day 4*
  - *o Outliers*
  - *o Lack of fitness of the regression model*
- *Day 5*
  - *o More on diagnostics*
  - *o Variable selection*
  - *o Cross validation*

Week 4

- *Day 1*
  - *o Variance-stabilizing transformations*
  - *o Transformations methods (Box-cox etc)*
- *Day 2*
  - *o Analytical methods for selecting a transformation*

- *Day 3*
  - o *Generalized and weighted least squares*
- *Day 4*
  - o *Regression models with random effects*
- *Day 5*
  - o *In-class exercises*
  - o *Assignment 4 due*

Week 5

- *Day 1*
  - o *Diagnostics For Leverage and Influence*
- *Day 2*
  - o *Indicator Variables and Quantitative Regressor*
  - o *Analysis of Variance*
- *Day 3*
  - o *Final review session*
- *Day 4*
  - o *Final review session*
- *Day 5*
  - o *Final exam*
  - o *Assignment 5 due*

Academic Honesty

Students are expected to maintain high standards of academic honesty. Specifically, unless otherwise directed by the professor, students may not consult other students, books, notes, electronic devices or any other source, on examinations. Failure to abide by this may result in a zero on the examination, or even failure in the course.