

# **STAT 5160 Advanced Statistical Inference II**

## **Section 001      Spring 2023**

**Instructor:** Wei Wang      [wei.wang@pennmedicine.upenn.edu](mailto:wei.wang@pennmedicine.upenn.edu)

**Class hours:** Tuesday and Thursday, 10:15 am - 11:45 am (EST).

**Office hours:** After class, email or by appointment.

### **Course description:**

The course covers both the theories and the applications of frequently used regression models. Topics include multiple linear regression model, two-stage least squares, generalized linear model, quantile regression, generalized estimation equations, multivariate normal distribution, and linear mixed effects models. If time permits, non/semi-parametric techniques will also be studied.

### **References:**

Introduction to Linear Regression Analysis, D. C. Montgomery, E. A. Peck, G. G. Vining.

Generalized Linear Models, P. McCullagh, J. Nelder.

Quantile Regression, R. Koenker.

Analysis of Longitudinal Data, P. J. Diggle, P. Heagerty, K.-Y. Liang, S. L. Zeger.

**Course Prerequisite:** STAT 5150.

### **Required Background:**

Graduate level probability and statistics: conditional expectation/probability/variance, exponential family, maximum likelihood estimation, Fisher information, modes of convergence, central limit theorem, point estimation, hypothesis testing, and confidence intervals.

Calculus: limit, supremum, infimum, continuity, differentiation, integration, Taylor expansion, etc.

Linear algebra: vectors, matrices, matrix multiplication, matrix determinant and inverse, Jacobian matrix and Hessian matrix.

R programming experience: scatter plots, histograms, boxplots, simple simulation, and data summaries.

**Software:** We will use the free statistical computing software R (<http://www.r-project.org/>) frequently in class. You will apply what you learn in class to solve your homework and final exam problems.

**Homework:** There will be biweekly homework assignments. Without a convincing reason, late homework will not be given full credit (25 points off every 24 hours). If you are not sure about your situation, ask the instructor in advance. No last minute notice.

**Exam:** There will be a take home final exam.

**Format:**

You are advised to type your homework or exam solutions. Try your best to copy and paste everything including the relevant R code and output into a single file.

**Grading:** The final grade will be based 60% on homework, and 40% on the final exam. The solution has to be in your own words. No plagiarism or cheating. Identical solutions will be marked zero. Points up to 50% will be deducted if no R code or R output. Independent work is expected for the final exam. No discussion or communication with other people. Otherwise, it will be considered cheating, and the exam paper will be marked 0. Regular classroom attendance and participation is required.