STAT 516 – Statistical Methods II

Spring 2025

Meeting times:	10:05 – 11:20 pm Tuesday and Thursday in LeConte College 107.
Exam dates:	Midterm exams in class on February 20th and April 8th. Final exam at 9:00 am on May 6th.
Instructor:	Dr. Karl Gregory, LeConte College 216 C, gregorkb@stat.sc.edu.
Office hours:	After each class until noon or Tuesday 3:00 – 5:00 pm and Wednesday 1:00 – 2:00 pm.

Link to course website

Course overview

Statistical methods II, as apparent from the title, is a continuation of Statistical Methods I (STAT 515). It builds on the simple linear regression and one-way ANOVA material covered in Statistical Methods I, introducing a series of extensions of these models which allow for multiple covariates and more intricate experimental designs. By the end of the course, students will be familiar with a variety of linear statistical models commonly used by practitioners in agriculture, psychology, biology, manufacturing, medicine, economics, and other disciplines.

Following is an overview of the topics we will cover:

I. Multiple linear regression

We will begin with a review of hypothesis testing for the mean and simple linear regression, which are covered in STAT 515. We will then study multiple linear regression, which aims to predict or explain the variability in a response variable using several predictor variables or covariates (instead of just a single covariate). In the multiple linear regression setting we will consider the effects of multicollinearity (covariates being correlated with one another), the effect of adding a large number of covariates to the linear regression model, and strategies for selecting a set of important covariates from among all the covariates available in a data set.

II. Extensions to one-way ANOVA, factorial designs

Following our discussion of multiple linear regression, we will review the topic of one-way ANOVA, or the comparison of two or more means, which is covered in Statistical Methods I.

Here we will raise the question of how to adjust inferences when making multiple comparisons (for testing multiple hypotheses with the same data) and learn to implement Tukey's HSD and Dunnett's method for making multiple comparisons of means. This will prepare us to study more intricate experimental designs, such as 2-way factorial designs which will be the main focus of this part of the course.

III. Random and mixed effects models

After our study of 2-way factorial designs, we will consider models called random or mixed effects models, beginning with the one-way random effects model. This will prepare us to study more complicated designs including the randomized complete block design and the randomized complete block split-plot design. These experimental designs allow one to accommodate inhomogeneity between groups of subjects or experimental units.

IV. Analysis of covariance

Next we will cover analysis of covariance (ANCOVA), which allows one to account for inhomogeneity of subjects in any of the foregoing experimental designs by incorporating into the model the value of a covariate recorded on each subject, such as age. We will see that this can increase statistical power, that is the ability to identify differences between group means.

V. Logistic regression

We will finish the semester by covering logistic regression, which is a regression model for data with a binary response taking the value 0 or 1, for example indicating whether a subject has an infection or not. In this model the conditional distribution of the response given the covariate values is not assumed to be Normal, but rather a Bernoulli distribution.

Textbook

The course will focus primarily on Chapters 6, 7, 8, 9, 10 and on portions of Chapters 11 and 13 of *Statistical methods, 4th Ed* by Mohr, Wilson, and Freund. I will assign homework problems from this book. I will also post supplemental lecture slides on course website.

Computing

We will use the statistical software R throughout the semester. No previous experience with R is needed. Download the free version of R studio.

Prerequisites

Each student is required to have earned a C or higher in STAT 515, STAT 509, or STAT 512, or some equivalent course.

Grading

Homework will be worth 20% of the final course grade. There will be two midterm exams each worth 25% of the final course grade and a final exam worth 30% of the final course grade. There will be 7 or 8 homework assignments during the semester. The lowest homework score will be dropped when computing the average homework score. Each homework must be typeset using Quarto or R markdown (I will show you how) and uploaded as a single pdf document into Blackboard. If your final exam score is higher than the lower of your two midterm exam scores, your final exam score will replace that midterm exam score in the calculation of your course grade.

The thresholds 90%, 87%, 80%, 77%, 70%, 67%, and 60% will be used to determine the assignment of the letter grades A, B+, B, C+, C, D+, and D, respectively. The grade of F will be assigned to those earning less than 60%.

Find important withdrawal dates here.

Honor code

See the Carolinian Creed in the Carolina Community: Student Handbook & Policy Guide. Violations of the USC Honor Code may result in a 0 for the work in question, and, in accordance with University policy, other punishments up to and including expulsion from the University.

Statement on AI

The exams, which make up 80% of the course grade, will be taken in class without computers or access to phones. Since you must earn 80% of your course grade without the assistance

of AI, I will not very strictly police the use of AI on homework assignments, which make up the other 20% of your grade. However, *letting AI do your homework assignments for you* will likely have the following effects: 1) You will not get the right answers. 2) Even if you get the right answers, you will not learn from the homework assignments the concepts you need to master in order to score well on the exams. 3) Most gravely, you will become less intelligent for failing to flex your brain muscle! If you wish to use AI, I recommend *using AI as a sophisticated search engine*. For example, you could ask AI focused questions like, "How do I make an R plot with tick marks placed exactly where I want them?"

Accommodations

If you require special accommodations, they must be arranged in advance through the `Office of Student Disability Services`.