

STAT 714: LINEAR STATISTICAL MODELS

Fall 2021

| | |
|--|--------------------------------------|
| Instructor: Ray Bai | Time: MWF 1:10 PM – 2:00 PM |
| Email: RBAI@mailbox.sc.edu | Place: Carolina Coliseum 2006 |

Course Page:

<https://blackboard.sc.edu/> (Check regularly for announcements and homework assignments)

Office Hours: Tuesdays and Thursdays 1:00-2:00 pm in Carolina Coliseum 4010C

COVID-19 Policies: All classes and exams will be in person. The class will only be moved online temporarily if the instructor is required to isolate/quarantine. Students are strongly encouraged to wear face coverings and to maintain 3-foot social distancing in the classroom. For university updates on COVID-19, please see <https://sc.edu/safety/coronavirus/>

Course Description: The purpose of this course is to provide an introduction to the theory of linear models. This is a **graduate-level theoretical course** involving rigorous proofs and derivations. **This is *not* an applied linear models course.** If you are looking for an applied data analysis course, please consider enrolling in STAT 704 or another class. Discussion of results from linear algebra will be incorporated into lectures as needed; however, you are expected to have a strong familiarity of this material already.

The course covers the theory of the linear model, including least squares and generalized least squares (GLS) estimation, the Gauss-Markov theorem, distributional theory, hypothesis testing, and shrinkage methods. The tentative schedule is given on the last page of this syllabus.

Learning Outcomes:

1. Develop a rigorous understanding of matrix algebra and apply these insights to the study of linear models.
2. Comprehend the mathematical foundations of linear statistical models and the theory underlying estimation and inference for linear models.
3. Gain familiarity with both “classical” linear regression models and “modern” shrinkage methods for regression models and the bias-variance trade-off for these modeling frameworks.

Prerequisites: A year-long sequence in mathematical statistics and a course in linear algebra.

Main References: We will use a packet of course notes prepared by the instructor. Parts of these lecture notes are *not* complete and will be filled in during lecture. Thus, it is in your best interest to attend lectures.

No textbook is required. However, the following books may be useful as supplementary references:

- Christensen, R. (2002). *Plane Answers to Complex Questions*. Springer.
- Monahan, J. (2008). *A Primer on Linear Models*. CRC Press.

Attendance Policy: Attendance is expected but is not part of the grade.

Computing: This is primarily a theory class so there is minimal computing in this class. Students interested in applied linear models and data analysis should enroll in STAT 704 or another class.

Homework: There will be nine short homework assignments. For each assignment, I will randomly pick two questions to be handed in. In order to master the material, you should spend an ample amount of time on each assignment.

Exams: There will be two midterms and a final. The two midterms will be in class. The final exam will be during the final exam period. *One* double-sided 8.5" × 11" sheet of notes is allowed for the exams. Tablets and calculators are *not* permitted during exams. The dates for Exam 1 and Exam 2 are Monday, September 27 and Wednesday, October 27. The final exam is on Friday, December 10 at 12:30 pm.

Grading: Your grade will be determined according to the following distribution:

- Final Exam: 40%
- Midterm with higher grade: 25%
- $\max\{\text{Homework average, Midterm with lower grade}\}$: 25%
- $\min\{\text{Homework average, Midterm with lower grade}\}$: 10%

The tentative grading scale is as follows: 90-100 for an A, 80-89 for a B+, 70-79 for a B, 60-69 for a C+, 0-59 for a C. This might be adjusted later on.

Honor Code: See the Carolinian Creed in the *Carolina Community: Student Handbook and Policy Guide*. The *minimum* punishment for violations of the USC Honor Code is a grade of zero for the work in question. In accordance with university policy, there may be other punishments, including an automatic F in the class and/or expulsion from the university.

Accommodation: If you need special accommodations for examinations or any other aspects of the course, please contact me before or during the first week of the semester.

Note that reasonable accommodations are available for students with a documented disability. If you have a disability and may need accommodations to fully participate in this class, contact the Office of Student Disability Services by phone (803-777-6142) or e-mail sasds@mailbox.sc.edu. All accommodations must be approved through the Office of Student Disability Services.

Tentative Schedule of Topics:

- **Week 1:** overview of linear models
- **Week 2-5:** review of linear algebra
- **Week 6-7:** normal equations, least squares estimation, identifiability and estimability
- **Week 8-9:** Gauss-Markov Theorem, Gauss-Markov and Aitken models
- **Week 10-11:** multivariate normal distribution, distributional theory
- **Week 12-13:** inference and hypothesis testing for the linear model
- **Week 14-15:** bias-variance trade-off, shrinkage methods for linear regression