STA 6246 Linear Models
Spring 2025

Instructor Hani Doss—222 Griffin-Floyd; email: doss@stat.ufl.edu (email is primarily for administrative purposes, not for questions regarding the course material; for such questions, talk to me during office hours). Office Hours: Tu-Th 10:20am–11:20am.

Course Description This course covers the theory of the general linear model, including regression and analysis of variance models, least squares estimates and distribution theory when the errors are normally distributed. A geometric approach is emphasized.

Orientation of the Course The primary purpose of this course is to prepare the students in the Ph.D. program in Statistics for the linear models portion of the Ph.D. qualifying exam in Statistics. As such, I will cover a specific body of material in the theory of linear models that forms the core of what every Ph.D. student in Statistics should know. This is not a course on experimental design, nor is it a course on applied regression. If what you are looking for is a course on these topics, you should drop this course immediately.

Prerequisites Prerequisites are:

- STA 6208 Basic Design and Analysis of Experiments
- STA 6327 Introduction to Theoretical Statistics II
- Familiarity with standard definitions and material on linear algebra (at least at the undergraduate level)
- Some familiarity with R

If you have not taken both of these (or their equivalents), you may not register for this course.

Text R. Christensen, *Plane Answers to Complex Questions: The Theory of Linear Models* (5th edition, 2020), Springer. (The fourth edition is OK also.)

We will use the statistical computing language R (which can be downloaded for free from http://www.r-project.org), and a student who is not familiar with it is strongly advised to become so by the fourth week of the semester.

We will not use the book by Christensen much. However, it is an excellent text, and the approach used in that book is close to the approach we will use. You may not wish to buy it, but you need to be advised that every serious doctoral student in Statistics must own a good book on linear models.

Course Web Page https://users.stat.ufl.edu/~doss/Courses/lm

Grading Your final course grade will be based on the four components below, with the stated weights:

Exam 1:	Thursday February 20, 8:20pm; covers everything up to and including the lectures of Tuesday February 18. Note the evening time slot.	25%
Exam 2:	Thursday March 27, 8:20pm; covers all material after Exam 1 up to and including the lectures of Tuesday March 25. Note the evening time slot.	25%
Final:	Comprehensive. May 1, 10:00am–12:00pm.	35%
HW:	About 8 or 9 homework assignments.	15%

The solutions to the homework assignments must be entirely your own (this applies also to R code).

For Exams 1 and 2, you will have two hours for a 50-minute exam.

Topics

- Overview of general linear models.
- Review of basic linear algebra, including projections, eigenvalues and eigenvectors, and spectral decomposition of symmetric matrices.
- The multivariate normal distribution.
- Least squares estimates: their derivation and basic properties for models of full rank. Relationship between the normal equations and projection operator approaches. Weighted least squares.
- Distribution of least squares estimates under normal errors.
- One-way and two-way analysis of variance.
- Linear regression.
- Implementation in R.
- Introduction to nonparametric regression (local linear smoothing, regression splines, smoothing splines); generalized additive models; implementation of these in R.
- Hypothesis testing. The general linear hypothesis. Discussion of full-rank and less-than-full-rank models. The F test; the likelihood ratio test and connections with the classical tests in large sample theory for parametric models.
- Penalized regression, including ridge regression and lasso.
- Multiple comparisons.
- Bayesian analysis of linear models.

Initial Assignment

1 Read Appendix A of PACQ. You need to have read the first half before the second lecture.

2 Do all the exercises in Appendix A (but for exercises with long lists only do a few parts). This is not to be handed in. You should do this before the third lecture.

3 Read pages 13–22 of the class notes. You should do this before the third lecture. Note: some of this material duplicates, or elaborates on, what is in Appendix A of PACQ.

4 Start to get familiar with R if you aren't already.

Note: Items 1–3 above should be a review.

General Course Policies

- All emails to me *must* have the string "LM" in the subject line (so I can retrieve emails quickly) and must be sent from your official UF mail account.
- Cell phones may not be used; they should be turned off (or set on silent). Laptops must be shut.
- If you have a disability and will request accommodations you should see me as early in the semester as possible.

Policies on Covid Safety

Illness If you are sick, do not come to class. If you are sick and need immediate care, call your primary care provider or the UF Student Health Care Center at 352-392-1161 to be evaluated.

Absences If you test positive, you should not come to class, and as with any excused absence, you will be given a reasonable amount of time to make up missed work.