

# Course Information SDS 387: Linear Models Unique number: 57535

## Spring 2015

**Instructor.** Peter Müller. My office is R.L.M 11.174. I will hold office hours from 4-5 on Mo and Wednesday. If you cannot make my office hours and would like to come by, please make an appointment with me. My office telephone is 471-7168.

**Prerequisites.** Graduate standing, knowledge of mathematical statistics at a graduate level and linear algebra at an advanced undergraduate level is required as well as basic coding skills (R, Matlab, or Stata). Some prior knowledge of Bayesian inference is desirable. The course will include a brief review of Bayesian inference (see the textbook below).

**Text.** The textbook for the course is “Plane Answers to Complex Questions: The Theory of Linear Models” by R. Christensen, Springer-Verlag. I will place one copy on reserve in the Math library. The heart of the course will be chapters 2-6, complemented by extensive additional material on Bayesian inference for linear models. As additional text and reference for Bayesian inference we will use “A First Course in Bayesian Statistical Methods” by P. Hoff. (Any other introduction to Bayesian inference is fine – no need to buy this particular book if you have others). P. Hoff’s book is available as e-book at UT libraries. You can get a paperback copy for \$25 from Springer-Verlag.

**Course Grading and Exams.** There will be two midterm exams, each counting 25% of the course grade.

- Exam 1 will be on We, Mar 4 (tentative)
- Exam 2 will be in We, Apr 1 (tentative).
- Exam 3 will be in We, Apr 22 (tentative).

Homework problems will count for the remaining 25% of the course grade.

**Homework.** Homework problems will be assigned throughout the semester. Some problems will involve computational work. I recommend the use of R.

**Group work.** Students are encouraged (but not required) to work on homeworks in groups. Please hand in one assignment with all group member names. No need to register groups besides simply putting the names on the assignment. You can change groups at any time.

**Website.** Please visit the canvas site for this course

<https://utexas.instructure.com/courses/1136021>

Homework assignments will be posted there.

**Students with disabilities.** Please notify me of any modification/adaptation you may require to accommodate a disability-related need. You will be requested to provide documentation to the Dean of Students Office, in order that the most appropriate accommodations can be determined. Specialized services are available on campus through Services for Students with Disabilities.

## Week-by-week schedule

In this class we will discuss the practical application of the projection approach to linear models. The course will begin with a review of essential linear algebra concepts including vector spaces, basis, linear transformations, norms, orthogonal projections, and simple matrix algebra. It continues by presenting the theory of linear models from a projection-based perspective. Still on the projection framework, Bayesian ideas will be introduced. Additional topics include: (i) Analysis of Variance; (ii) Generalized Linear Models; and (iii) Variable Selection

Therefore, the important prerequisites for the class are

- Mathematical statistics at a graduate level and linear algebra at an undergraduate level
- You should know some Bayesian inference. Well, you need not be experts. But you should be able to describe how a posterior distribution summarizes an inference problem, and why posterior simulation is important. The class will include a brief review of Bayesian inference.
- You should know some basic probability. At least you should know how to define a Markov chain. No worries I don't expect that you remember all details about classifying states etc :-)
- You should be familiar with some basic computation, ideally with R (a statistical programming language). Matlab is fine too. The class is not about programming, but you will need to know enough to implement simple algorithms for homework problems etc.

### Text book:

- Plane Answers to Complex Questions: The Theory of Linear Models -Christensen (C)
- A First Course in Bayesian Statistical Methods -Hoff (H)

C will be our main reference.

H is a good review of Bayes, but no need to buy it for the course.

## List of topics to be covered

### Week 1-3: Jan 21, 26/28; Feb 2/4

**Review:** linear algebra, distribution theory & Bayesian inference (some of this material will be reviewed only later when and as needed)

1. **Introduction to Linear Models:** Random Vectors, Matrices
  - a. C: Chapter 1, Appendix A & B
2. **Multivariate Normal Distribution Theory:** C: Chapter 1, Appendix A & B
3. **Conditional Normal Distributions:** C: Chapter 1, Appendix B and C
4. **Chi-Square and non-central Chi-Square:** Quadratic Forms C: Appendix B
5. **Eigenvalues,** Distributions of Quadratic Forms, Orthogonal Projections
  - a. C: Appendix B
6. **More Distributional Results:** C: Appendix C C: Chapter 2

**Week 4: Feb 9/11**

1. **Intro to Bayes:** H: chapters 3,5,6

**Weeks 5-6: Feb 16/18, 23/25**

1. **Models:** Maximum Likelihood Estimates  
C: Chapter 2
2. **Identifiability & Estimation:** C: Chapter 2
3. **Gauss-Markov:** C: Chapter 2
4. **Weighted Regression:** C: Chapter 10

**Week 7: Mar 2/4**

1. **Bayesian Regression:** H: Chapter 9; C: 2
2. Mar 4: Midterm 1
3. **Marginal & Predictive Distributions:** Default Priors

**Week 8: Mar 9/12**

1. **Shrinkage Methods:** C. Chapter 14
2. **Bayesian Shrinkage**

**Week 9: Mar 23/25**

1. **Lasso Shrinkage**
2. **Bayesian Lasso:** Carvalho, Polson Scott, Hans, Park & Casella

**Week 10: Mar 30/Apr 1**

1. **tests:** Likelihood Ratio Tests, F-tests & Analysis of Variance, Bayes Factors versus P-values C:  
Chapter 3 -5; C: Chapter 6-7
2. Apr 1: Midterm 2

**Week 11: Apr 6/8**

1. Bayesian Model Choice/Averaging: Clyde & George, (Statistical Science, 2004) "Model Uncertainty"  
Liang et al. (JASA, 2008) "Mixtures of g-priors"

**Weeks 12-15: Additional topics**

1. **Improper Priors and Model Selection:** Intrinsic Bayes Factors Casella & Moreno (JASA, March 2006)
2. **Outlier Analysis:** Christensen Ch 13
3. **Bayesian Outlier Analysis:** Chaloner & Brant (Biometrika, 1988); Hoeting et al (Comp Stat & Data Analysis, 1996) 5