

Instructor: Dr. Jason Hoeksema (318 Shoemaker, 662-915-1275, hoeksema@olemiss.edu)

Lecture/Lab: Tuesday & Thursday, 9:30 a.m. – 12:00 (Shoemaker Hall Rm. 225)

Required course materials:

- *Discovering Statistics Using R*, by Andy Field, Jeremy Miles, and Zoe Field. 2012, Sage.

Course grading: Grades will be based on participation in and completion of lab exercises/homework (20%), a midterm exam (25%), a data analysis project (20%), and a final exam (35%).

Grading scale: >90% = A/A-, 80-89% = B+/-, 70-79% = C+/-, 60-69% = D, <60% = F

Learning objectives:

By the end of the course, students will be able to:

- Design efficient sampling schemes and experiments for biological research
- Describe and summarize data using descriptive statistics and graphics
- Conduct inferential statistical tests on a wide variety of data types
- Interpret, explain, and present the results of statistical tests

Course format:

Lectures introduce statistical concepts and approaches to data analysis. Lab exercises and associated homework assignments allow students to practice applying concepts and approaches to analysis of real data on computers, using the open-source statistical software R.

Schedule of lecture topics (relevant chapters from Field et al. in parentheses)

- 27 Aug** Course overview; Intro to Statistics (Ch. 1) and Introduction to R (Ch. 3)
- 29 Aug** Intro to Statistics continued (Ch. 2)
- 3 Sep** Exploring Data with Graphs (Ch. 4)
- 5 Sep** Probability and Hypothesis Testing
- 10 Sep** Probability and Hypothesis Testing (continued: presentations)
- 12 Sep** Correlation (Ch. 6)
- 17 Sep** Linear Regression (Ch. 7, pp. 245-260)
- 19 Sep** Exploring assumptions and handling violations of assumptions
- 24 Sep** Intro to Power Analysis & its application to correlation & regression
- 26 Sep** Comparing two means: t-test, etc. (Ch. 9)
- 1 Oct** t-test (continued, including power & sample size estimation)
- 3 Oct** Introduction to the take-home Midterm Exam (**due Monday, Oct. 7** at 11:59 p.m.)
- 8 Oct** Comparing several means: ANOVA (GLM1) (Ch. 10)
- 10 Oct** Comparing several means: ANOVA (GLM 1) (Ch. 10) continued
- 15 Oct** Analysis of covariance: ANCOVA (GLM 2) (Ch. 11)
- 17 Oct** Analysis of covariance: ANCOVA (GLM 2) (Ch. 11) continued
- 22 Oct** Factorial ANOVA (GLM 3) (Ch. 12)
- 24 Oct** Multiple regression (GLM 4) (Ch. 7, pp. 261-end)
- 29 Oct** GLM overview, review, practice and power analysis

- 31 Oct** Analysis of categorical response variables I: Chi-squared and related (Ch. 18)
- 5 Nov** Analysis of categorical response variables II: Logistic regression (Ch. 8)
- 7 Nov** Mixed models and repeated-measures designs
- 12 Nov** Brief intro to some advanced topics: Meta-analysis, multivariate, and alternatives to null hypothesis testing
- 14 Nov** Work on projects. Come with data and having attempted analysis.
- 19 Nov** Writing workshop: Bring drafts of Written Components of project: Data Analysis Methods, and Results
- 21 Nov** Presentations at 9:30 a.m., written component of project due at 10:00 p.m.
- 26 & 28 November: THANKSGIVING BREAK**
- 3 Dec** Provision of the final exam (take-home, open-book, due at 10:00 p.m. on Monday, December 9th)
- 9 Dec** Take-home final exam due at 10:00pm

Academic misconduct: Students are responsible for abiding by the university's policies on Academic Conduct and Discipline, which are available on the university website.

**Collaboration with your peers is encouraged on lab exercises and homework, but all writing and R code must be your own, with no exceptions. On the take-home midterm exam and final exam, no discussion with other students is allowed, and all answers must be strictly your own work. Violation of this policy or any others will not be tolerated in this course, and possible sanctions for academic misconduct include: failure on the work in question, course grade reduction or failure of the course, disciplinary probation, or suspension or expulsion from the University.

Cell phones, texting, e-mail, and web-browsing in class and meetings: Turn off your phones and other electronic devices before entering class or a meeting with an instructor. Computers are encouraged in class, but may only be used for course work. Texting, e-mailing, internet browsing, and making phone calls are all **strictly prohibited** during class and meetings with the instructor.

Guidelines for data analysis project: You must conduct an analysis of some data related to your graduate research project or your research interests. The main goal is for you to conduct a data analysis similar to what you will need to conduct during your graduate research, or to analyze some data that you already have collected, but have not yet analyzed. The data can be your own, can be borrowed from someone, found on the internet, or made up (e.g., resembling data that would support one of your hypotheses). The analysis can be anything we are learning in class, but it must answer a question related to your research interests. There are two components to the project: a Written Component and a Presentation.

Written component: The written component should have two sections: a *Data Analysis Methods* section, and a *Results* or *Potential Results* section. The instructor will provide examples of how these two sections should be written. The final version should be uploaded as a Word document on Blackboard, with both sections included in the same document. Due date is the same as the day of presentations (by 10pm that night). The Data Analysis Methods should be written in the style of a journal article (past tense) or grant proposal/prospectus (future tense). It should describe the type of analyses conducted, which predictor and response variables were used, whether any transformations were conducted to meet assumptions, what the alpha/significance level was for the tests, and what software (and any extra packages) were used to conduct the analysis (including R and package version numbers, plus any appropriate citations). The Results or Potential Results section should be written in the style of a journal article (Results, past tense) or grant proposal (Potential Results, future tense). This section should include clear conclusion statements, as practiced during the semester, and graphs to illustrate significant results. These same graphs should be used in your presentation. Make sure to provide statistical support in parentheses, including test statistic, degrees of freedom, and P-value). Make sure to describe the result, including the direction of the effect (e.g., positive or negative correlation or slope, or which groups had larger means than others), and refer to a numbered figure to illustrate significant results.

Presentation component: Time limit is 8 minutes, maximum, and will be strictly enforced, so please practice, and shorten if necessary to meet the time limit. Format for presentation (1 slide each unless otherwise noted): Title slide (including your name), brief background (big picture of the scientific problem, bullet points and/or graphics), research question(s), hypotheses (in the style of research hypotheses, not null vs. alternative), data collection methods (how did you get your data?; bullet points and/or graphics), type of analysis (very briefly describe/justify it), Results (1 or 2 slides with figures, no text), Conclusions/Summary (bullet points giving main result(s)), Acknowledgements. Stylistic requirements: **No slides should have large blocks of text on them (brief bullet points at most). **Use large enough fonts to be read from the back of the classroom. **Don't put anything on the bottom 20% of your slides (too difficult to see from back of room). Please upload a Powerpoint presentation to Blackboard, at least 10 minutes before class starts on the day of presentations.