# LIN4930/6932 Advanced statistics (Spring 2026)

T 11:45am-1:40pm; Th 12:50pm-1:40pm
Keene-Flint Hall 0117

Instructor: Julian Grove
Office: 4108 Turlington
Email: julian.grove@ufl.edu
Office hour: TBD

**Course objectives**

This course deals with more advanced issues in frequentist approaches to statistical analysis (mixed effects models, contrast coding, among other things), and provides an introduction to Bayesian statistics. Examples come from various fields of linguistics. We will use both R and Stan.

By the end of class, students are expected to:

1. be able to analyze complex quantitative data within frequentist frameworks;
2. be able to approach similar problems from a Bayesian perspective;
3. gain more familiarity with R (and Stan).

**Prerequisite**

LIN4005/LIN6007 Statistics for Linguists, or other equivalent background or coursework. Please contact the instructor to discuss your background if you have not taken LIN4005/LIN6007.

**Course website**

Canvas and TBD.

**Textbook**

There is no required textbook for this course. Readings will be drawn from publicly available resources, chapters from statistics books, and scientific publications.

**Laptop and software**

Students are required to bring their laptops to class, in order to do in-class assignments and group activities. Please have the latest version of R and R studio installed by the first session: R (www.r-project.org); R-studio (www.rstudio.com). Later on, we will use Stan (https://mc-stan.org/).

**Attendance & Participation**

- Coming to class prepared

• Doing the in-class exercises. These are practice exercises illustrating new concepts and techniques. Answers will be provided in class. Doing these exercises is critical, since this allows you to monitor your own understanding and receive immediate feedback.

• Contributing to discussion in class

• Helping others understand the concepts and problems in class

**Assignments**

• Homework assignments are due about every week (about 10 in total), except in weeks when a test is due. Students have one week to make the homework, see schedule.

• Homeworks are made available on and should be submitted through the course website

• Typical homework assignments include: reading a data set into R and conducting analyses or visualizing the data, applying learned concepts to particular research scenarios, interpreting tables and graphs, interpreting results sections of published reports, making exercises for peers in which new terminology, concepts or techniques are illustrated and practiced.

• Note: in some weeks graduate students may get more or other homework exercises than undergraduate students.

**Tests**

• There will be three open book tests over the course of the semester, approximately in weeks 6, 11 and 16. The format of the test (take-home, or in-class) will be decided closer to the dates.

• Test questions and problems are comparable to those practiced in class and in the homework assignments

• Tests will be made available through the course website.

• These tests are non-cumulative, although basic knowledge is assumed of what has been dealt with earlier in the semester (especially concerning R/Stan).

**Participation in LIN/S&P/SLHS experiments (optional, extra credit)**

To encourage awareness of different aspects of experimental research in language-related fields, you can elect to participate in 2 hours of language or communication research during the semester. For a list of studies that qualify, see the SONA website, https://ufl-slhs.sona-systems.com/. For details on how to use SONA, see this document. You need to have participated in 2 hours of experiments before the last session of this course for you to receive credit. This assignment will be worth a maximum of 2% of your course grade. If you choose not to participate or do not qualify for any of the studies, you can receive the same amount of course credit for reading a

short research article (provided by the instructor) and writing a 2 page synopsis. This paper must be turned in no later than the last session of this course.

**Grade breakdown**

| | *Graduates* | *Undergraduates* |
|---|---|---|
| Attendance & Participation | 5% | 5% |
| HWs and in-class assignments | 20% | 35% |
| Critique paper + data set analysis | 15% | N/A |
| Tests | 60% | 60% |
| Extra credit | 2% | 2% |

| | | | | | |
|---|---|---|---|---|---|
| A 94-100 | A- 90-94 | B+ 87-90 | B 84-87 | B- 80-84 | C+ 77-80 |
| C 74-77 | C- 70-74 | D+ 67-70 | D 64-67 | D- 61-64 | E < 61 |

**Academic Integrity**

See https://syllabus.ufl.edu/syllabus-policy/uf-syllabus-policy-links/.

**Attendance Policy, Class Expectations, use of LLMs, Make-Up Policy, Cell-phone use**

• Students are required to submit all assignments and tests before the class period they are due. Please contact the instructor in advance if you need to skip a class, or cannot make a deadline.

• Attendance is mandatory. If you are absent for more than three classes, you will get a warning. If absences persist the instructor can prohibit further attendance and assign a failing grade for excessive
Absences.

• Requirements for class attendance and make-up exams, assignments, and other work in this course are consistent with university policies that can be found at: catalog.ufl.edu/UGRD/academic-regulations/attendance-policies/.

• Students need to disclose any use of AI in their assignments, and need to indicate their prompts and their critical (content) edits of the AI output. AI output without prompts or critical edits will not be accepted. AI needs to be properly cited (https://apastyle.apa.org/blog/how-to-cite-chatgpt).

• Cell-phone use is not allowed during class unless this is part of the course assignments. Laptops and tablets can only be used to take notes and for in-class assignments.

**Campus resources**

See https://syllabus.ufl.edu/syllabus-policy/uf-syllabus-policy-links/.

**Calendar**

| **Week** | **Date** | **Topics** |
|---|---|---|
| Week 1 | 1/10 | Class Introduction<br>Frequentist and Bayesian approaches<br>Refresher of R and terminology |
| Week 2 | 1/17 | Refresher of (logistic) mixed effects models |
| Week 3 | 1/24 | Convergence issues, collinearity, contrast coding |
| Week 4 | 1/31 | Contrast coding |
| Week 5 | 2/7 | Power |
| Week 6 | 2/14 | Poisson regression |
| Week 7 | 2/21 | Capitalizing on random effects |
| Week 8 | 2/27 | Probability, Bayes Rules, Distributions |
| Week 9 | 3/6 | Choosing prior and likelihood |
| Week 10 | **Spring break** | |
| Week 11 | 3/20 | Bayesian linear regression |
| Week 12 | 3/27 | Bayesian mixed-effects models |
| Week 13 | 4/3 | Bayesian mixed-effects models, continued |
| Week 14 | 4/10 | Bayesian prior elicitation |
| Week 15 | 4/17 | Bayesian contrast coding |
| Week 16 | 4/24 | Taking stock |