Experimental design and probability

Stanford University, Hopkins Marine Station

Spring 2019

Contents

Syllabus

Course description
Variability is an integral part of biology. Introduction to probability and its use in designing experiments to address biological problems. Focus is on experimental design and the use of linear models in testing hypotheses (e.g., analysis of variance, regression). Students will use R to explore and analyze locally relevant biological datasets. No programming or statistical background is assumed. Prerequisite: consent of instructor.

We will meet for 1.5 hours, twice a week. Classroom activities will be a mix of lecture, discussion, and collaborative problem solving. **Please bring your laptop to all classroom sessions.** Prior to class, students will complete a quiz which will be based on an assigned reading or video. All participants in the course will abide by the Code of Conduct, described below.

Instructor
Robin Elahi
- Office: Dive locker (Boatworks) - Office hours: Drop in, or by appointment via email

Teaching assistant
Nicole Moyen
- Office: Denny Lab (Fisher)
- Office hours: 12-2pm, Tuesday

E-communication
You can email us at:
elahi at stanford
nmoyen at stanford
Please use your Stanford email for all correspondence.

Course website:
https://elahi.github.io/xdp
This course website runs from a repository on GitHub. You can view this repository here:
https://github.com/elahi/xdp
Throughout the quarter, I will update the repository and website to reflect our progress. Please keep checking back to stay up to date.

Times and Rooms:
2:30-4pm; Mondays & Wednesdays; Fisher classroom

Textbook

We will use a variety of resources in this course. The primary one is:

Open intro statistics, 3rd edition - freely available as a pdf (or get a print copy for ~$15).

The others will be online resources for learning the basics of data science using `R`.

Schedule

Week	Topics
1	Intro to course, data collection, and R
2	Numerical and categorical data, data viz
3	Intro to probability
4	Normal distribution, intro to inference
5	Hypothesis testing
6	Inference for numerical data
7	Analysis of variance
8	Regression
9	General linear models, multiple predictors
10	Blocked designs

Details on assignments and readings are on the materials page.

Learning objectives

This course is a practical introduction to data analysis and experimental design for biologists and environmental scientists. Upon completion of the course, students will:

1. Recognize the importance of data collection, identify limitations in data collection methods, and be aware of other sources of statistical bias.
2. Use the programming language `R` to summarize data numerically and visually, and to perform data analysis.
3. Have a conceptual understanding of the unified nature of statistical inference.
4. Apply estimation and testing methods to analyze single variables or the relationship between two variables in order to understand biological phenomena.
5. Model numerical response variables using a single explanatory variable or multiple explanatory variables in order to investigate relationships between variables.
6. Interpret results correctly, effectively, and in context without relying on statistical jargon.
7. Critique data-based claims and evaluate data-based decisions.
8. Design experiments with appropriate considerations of controls, randomization, and blocking.

Evaluation and grading

Here is a breakdown of graded tasks:

- Quizzes (40%)
- Labs (40%)

- Final exam (20%)

Late submissions will not be accepted, but the lowest graded quiz and lab will be dropped.

Tips for success

- Complete the readings before class, to prime you for class lessons (not to mention the quizzes). Then go over the readings again, after class.
- Be an active participant during lectures and labs.
- Ask questions - especially during class or office hours. Ask me, your TA, and your classmates.
- Do not procrastinate - don't let a week go by with unanswered questions as it will just make the following week's material even more difficult to follow.

Notes

We may change the activities schedule or meeting location. You will be notified of any changes.

Use of your personal computer for in-class exercises, such as data analysis and literature searches, is highly recommended. Please let us know if you do not have a computer so we may make classroom arrangements.

Plagiarism, Dishonesty, and Academic Misconduct

It is expected that Stanford's Honor Code will be followed in all matters relating to this course. You are encouraged to meet and exchange ideas with your classmates while studying and working on homework assignments, but you are individually responsible for your own work and for understanding the material. Passing anyone else's scholarly work, which can include: written material, computer code, exam answers, graphics or other images, and even ideas as your own, without proper attribution, is considered academic misconduct.

Plagiarism, cheating, and other misconduct, including bullying, discrimination, and harassment, are serious violations of the University's *Fundamental Standard* and *Honor Code*:

https://communitystandards.stanford.edu/policies-and-guidance

Students with Documented Disabilities

Students who may need an academic accommodation based on the impact of a disability must initiate the request with the Office of Accessible Education (OAE). Professional staff will evaluate the request with required documentation, recommend reasonable accommodations, and prepare an Accommodation Letter for faculty dated in the current quarter in which the request is being made. Students should contact the OAE as soon as possible since timely notice is needed to coordinate accommodations. The OAE is located at 563 Salvatierra Walk (phone: 723-1066, URL: http://studentaffairs.stanford.edu/oae).

Affordability of Course Materials

Stanford University and its instructors are committed to ensuring that all courses are financially accessible to all students. If you are an undergraduate who needs assistance with the cost of course textbooks, supplies, materials and/or fees, you are welcome to approach me directly. If would prefer not to approach me directly, please note that you can ask the Diversity & First-Gen Office for assistance by completing their questionnaire on course textbooks & supplies: http://tinyurl.com/jpqbarn or by contacting Joseph Brown, the Associate Director of the Diversity and First-Gen Office (jlbrown@stanford.edu; Old Union Room 207). Dr. Brown is available to connect you with resources and support while ensuring your privacy.

Roles and Responsibilities

Student: inform the instructor no later than the first week of the quarter of any accommodation(s) you will or may potentially require.

Instructors: maintain strict confidentiality of any student's disability and accommodations; help all students meet the learning objectives of this course.

Before Class

`R` and `RStudio`

Before the first class please read through the computer setup instructions that walk you through how to set up your computer to run R and Rstudio. Even if you have these programs already installed, make sure to check that you are running the latest versions of R and RStudio (which the instructions will tell you how to do).

If you are unfamiliar with R and RStudio:
Intro to R & RStudio
Intro to R markdown

Code of conduct

I will maintain a respectful environment where everyone can engage in an open discussion and bring their strengths and weaknesses to the table without apprehension. Any behavior that detracts from these goals will not be tolerated.

Expected behavior includes (but is not limited to):

- Treating all participants with respect and consideration.
- Communicating openly with respect for others, critiquing ideas rather than individuals.
- Avoiding personal attacks directed toward others.

Unacceptable behavior includes (but is not limited to):

- Behavior that implies or indicates that someone does not belong in the class based on any personal characteristic or identity.
- Any unwanted attention, sexual advances, and comments about appearance.
- Verbal harassment, including comments, epithets, slurs, threats, and negative stereotyping that are offensive, hostile, disrespectful, or unwelcome.
- Non-verbal harassment, including actions or distribution, display, or discussion of any written or graphic material toward an individual or group that ridicules, denigrates, insults, belittles, or shows hostility, aversion, or disrespect.
- Bullying, intimidation, stalking, shaming, and assault.
- Retaliation for reporting harassment.
- Reporting an incident in bad faith.