SIO187 – STATISTICAL METHODS IN MARINE BIOLOGY

Spring 2020
Please note that all times in the written syllabus are in PDT
Lecture: Pre-recorded videos posted to Canvas at midnight Mon/Wed.
Dr. Orenstein will be available for questions on Zoom during the scheduled lecture time: Tue/Thur 9:00am – 10:20

Lab A01: Wed 9:00am – 10:00am, Zoom
Lab A02: Wed 11:00am – 12:00pm, Zoom

Note that the information in this syllabus is subject to change. Any changes will be announced in class and posted to Canvas. Be sure to frequently check the website to keep up-to-date with readings, assignments, and exam schedules.

Important dates

- Mid-term: Tuesday, May 5th, 2020
- Final review: TBD
- Final: Tuesday, June 9th, 2020 (8:00 AM – 11:00 AM)
- Project: Zoom presentations during Week 10 (schedule TBA)

Grading

- Lab assignments = 20%
- Homework = 20%
- Project = 15%
- Midterm = 15%
- Final = 30%

Textbook (required)

- Whitlock & Schluter (2015) *The Analysis of Biological Data, 2nd edition*, Roberts & Company Publishers.

The textbook is available to rent as an e-book for ~$50 from the publisher. Other access options exist, but please be sure to get the second edition. Please contact me if access to the book will be a burden.

Contact information

Instructor
Dr. Eric Orenstein, e1orenst@ucsd.edu
Scripps Institution of Oceanography
Office Hours: Mondays 3:00 – 5:00 pm via Zoom or by appointment

Teaching Assistant
Jonathan Charendoff, jcharend@ucsd.edu
Scripps Institution of Oceanography
Office Hours: Friday 2:00 – 3:00 pm via Zoom

A NOTE ABOUT CORONAVIRUS

As we are all aware, the COVID-19 pandemic has impacted many aspects of our daily lives, including this class. We will conduct this course entirely online at the university's direction. This is a challenging situation for all of us — faculty and staff included — and I urge patience as we figure this out together. I have never taught an online course before and I am sure there will be hiccups. I will try to be as attentive as I can to your needs and requests as we get under way.

If you need general information or assistance from UCSD, I encourage you to look at their new webpage dedicated to the coronavirus response and the VC of Student Affairs' resources. You can find updates about the school's response and how to get help if you need it.

COURSE DESCRIPTION

SIO187 will introduce students to the fundamentals of experimental design and quantitative analysis of data in biological oceanography. Students will learn to test hypotheses with a suite of statistical tests, including: t-tests, ANOVA, regression and correlation analysis, and non-parametric methods. They will also explore the meaning and utility of statistical distributions such as Gaussian, binomial, and Poisson. Hands-on data analysis will be done with the R software package during lab sessions.

LEARNING OUTCOMES

Students should expect to become familiar with some of the tools available to analyze environmental data. They will also be comfortable clearly presenting their results and interpreting the findings of other studies. By the end of the quarter, student will be able to:

1. Understand the foundations of statistics.
2. Recognize the assumptions and limitations of statistical tests.
3. Evaluate and analyze different types of biological data.
4. Interpret and communicate statistical output.
5. Use tools available in R to conduct analyses.

COURSE EXPECTATIONS

This class will challenge you to think critically about how data is collected, analyzed and presented. This is the very core of the scientific method. To excel in this course, you will need to go beyond simply applying equations and doing calculations; rather, you will have to understand these tools and be able to apply them in a deliberate, well-reasoned way. Although the class will introduce you to specific tools and tests, you will need to think more generally about difficulties associated with collecting environmental data.

Developing this level of understanding requires practice. As such, the course will be very focused on practical application and problem solving I will be posting a series of 5-15-minute videos *in leiu* of a complete lecture. Each capsule will focus on a concept or worked problem. This work will be reinforced in lab where you will be working with data using R.

CLASS POLICIES

Academic integrity

Students are expected to adhere to the University's policy on academic integrity. To paraphrase: do not copy another student's work, do not plagiarize work from an outside resource, do not do work for someone else, and do not make use of materials explicitly excluded by the instructor. Basically, be smart, respectful, and collegial. Please feel free to ask me if you have any questions or concerns about a matter related to academic integrity.

You are encouraged to work on homework and labs in groups. Regardless of the amount of collaboration, I will expect you to turn in your own assignments. Do not hesitate to ask for clarification on the expectations for any particular assignment.

Under no circumstances will cheating be tolerated. I will fail any student caught engaging in academic dishonesty. A student caught cheating will be subject to additional penalty from UCSD up to and including suspension.

Lectures

Lectures will be broken up into 5-15-minute video modules. A series of these pre-recorded snippets will be upload to Canvas in a prescribed order, but you may watch them however you like. **I encourage you to watch the videos during our lecture time slot even though they are recorded in advance.** You will be able to ask questions in real time via Zoom during that time period. Please note, that I will not be recording Q&A sessions as these will function more akin to office hours. I intend to make note of common questions and will post a summary to Canvas immediately following the conclusion of lecture.

Labs

Attendance and completion of labs and associated assignments account for 20% of your final grade. Missing more than 4 labs will result in failing the whole course. The labs will be a combination of problem solving, computer assignments, and projects. There will be graded lab assignments that count toward the final grade.

Lab work will be done with the R statistical package, a powerful – and free – computing package. To encourage social distancing, we ask that you install R and RStudio on your personal laptop. **Please get in touch with me if this is a burden in any way.**

The basic R package and instructions can be found at https://cran.r-project.org/. After installing R, install the graphical user interface RStudio from https://www.rstudio.com/products/rstudio/download/ (note that R must first be installed before RStudio). We will go over this in the first lab, but it is not a bad idea to try and get a head start.

Zoom etiquette

Much of the course will take place via Zoom. Please be conscientious when using this tool. When you log-in, you will by default be muted. To ask a question please use the hand-raising tool. When called upon, unmute yourself. Be sure to toggle the mute back on when you are done. This will prevent acoustic chaos in the video chat.

Discussions

We will be making active use of Canvas' discussion feature. Please feel free to post any questions that you might have from lectures, labs, or homework. In all likelihood, one of your colleagues could use clarification on the same point. We will also be using discussions to post frequently asked questions from office hours, etc.

Late work

You are encouraged to complete all assignments on time. The course work is cumulative and catching up might be difficult. Late assignments will be accepted, docked one letter grade per class after the due date. Assignments will be collected via Canvas.

Regrades

If you believe an error has occurred in grading, you may bring it to the attention of the instructor within one week of the assignment being returned. Please do not go to the TA for a regrade. You must submit the regrade request with a written description of the grading error. Regrade requests on an exam completed in pencil will not be accepted.

Missed exams

There are no make-up exams. If you miss an exam, you will be required to provide official documentation of an unavoidable emergency (e.g. serious illness, etc). Without such documentation, you will receive a failing grade for the exam.

Mid-term: A missed mid-term cannot be taken at a later date. With appropriate documentation of an unavoidable emergency, your grade will be scaled to maintain the grade you would receive based upon all other assignments.

Final: A missed final exam cannot be taken at a later date. With appropriate documentation of an unavoidable emergency, you will receive an incomplete for the course. You will then need to make arrangements to complete the class in the following year.

Accessibility

Please contact me in person or via email if you have a disability or condition that might affect your ability to participate in class or require any accommodations. I will work with you and the Office for Students with Disabilities (OSD) to make any necessary accommodations. If you are more comfortable having a staff member at OSD contact me instead, feel free to do so.

SCHEDULE (SUBJECT TO CHANGE)

Meeting	Topic	Reading	Homework
3/31/20	Introduction to stats	Chap. 1	
4/2/20	Plotting data and the elements of a good graph	Chap. 2	**HW1** W&S: 1.14, 1.15, 2.20, 2.24, 2.34 (due 4/9/20)
4/7/20	Describing data: $\bar{Y}, s, Y_{([n+1])/2)}$	Chap. 3	
4/9/20	Samples and Populations (estimation)	Chap. 4	**HW2** W&S: 3.15, 3.17, 4.15, 4.17, 4.23 (due 4/16/20)
4/14/20	Probability	Chap. 5	
4/16/20	Hypothesis testing	Chap. 6	**HW3** W&S: 5.20, 5.21, 5.26, 6.15, 6.21, 6.25 (due 4/23/20)
4/21/20	Proportions	Chap. 7	
4/23/20	Model fitting	Chap. 8	**HW4** W&S: 7.18, 7.21, 8.11, 8.14, 8.17 (due 4/28/20)
4/28/20	Contingency analysis	Chap. 9	
4/30/20	Normal (Gaussian) distribution	Chap. 10	Study, study, study
5/5/20	Mid-term		
5/7/20	One sample inference	Chap. 11	**HW5** W&S: 9.21, 10.15, 10.26, 11.14, 11.18 (due 5/14/20)
5/12/20	Two sample inference	Chap. 12	
5/14/20	Violation of assumptions	Chap. 13	**HW6** 12.18, 12.22, 13.19, 15.23, 15.30 (due 5/21/20)
5/19/20	ANOVA	Chap. 15	
5/21/20	Correlation	Chap. 16	**HW7** 16.15, 16.23, 17.19, 17.23 (due 5/28/20)
5/26/20	Regression	Chap. 17	
5/28/20	Cushion day		
6/2/20	Modern advances and machine learning		
6/4/20	Final review		