

BGGN216 Biostatistics Syllabus

SP23 UCSD

Lectures Tue/Thu 12:00 – 1:20PM Pacific Hall 3500
Lab Thu 4:00 – 5:00 PM Pacific Hall 3500

Instructor: Pamela Reinagel
email: preinagel@ucsd.edu
office: Pacific Hall 3100A
office hours: by arrangement

Instructional Assistant: Phoebe Li
email: xil156@ucsd.edu
office hour location: Pacific Hall 3500
office hour: Thu 5:00-6:00 Pacific Hall 3500

Wk	Lectures	Topic
1	April 4,6	Identifying, measuring, controlling, and communicating uncertainty
2	April 11,13	Null Hypothesis Significance Testing e.g., t-test and ANOVA
3	April 18,20	NHST: e.g., non-Gaussian, non-parametric tests, boot-strapping
4	April 25,27	Reproducibility crisis and p-hacking
5	May 2,4	Regression/correlation, common mistakes
6	May 9,11	Effect sizes and confidence intervals
7	May 16,18	Exploratory data analysis and data visualization
8	May 23,25	Introductory Bayesian statistics
9	May 30/Jun1	Reprise: what p values really mean; implications for replications
10	June 6,8	Where to go from here: how, when and why to consult a statistician

The schedule is approximate and subject to adjustment.

Description

This course is primarily intended for first-year students in the Biology and affiliated PhD and MD-PhD programs, and is intended to provide knowledge and skills essential for conducting independent research. **During lectures I will cover** basic conceptual principles and theory underlying the use of statistical methods, including null hypothesis significance testing, regression/correlation, effect sizes, confidence intervals, exploratory data analysis and data visualization, and introductory Bayesian statistics. We will focus on understanding the rationale, proper interpretation, limitations and pitfalls of general classes of methods, rather than focusing on mechanical procedures. We will also discuss reproducibility and reliability of science, including how this should be defined and assessed, obstacles, and best practices. Finally, we will discuss how to know when you need help: where to go, when to go, and what you need to tell a statistician so that they can help you. **The lab section** will be spent workshopping practical applications. Students will be asked to submit or present example experiments arising in their respective fields of research. We will work through these examples as a group over the course of the quarter.

Learning objectives

Students successfully completing this course will be able to

- Explain the purpose of statistical null hypothesis significance tests (NHST)
- Explain what a p value means, and how any test could provide such a measure
- Carry out a few example kinds of NHST correctly and identify when they are applicable
- Explain what statistical power means, what it depends on, and principles of power analysis
- Identify common conceptual or experimental errors that undermine p value validity
- Explain and perform at least one kind of multiple comparison correction

- Correctly compute correlation coefficients and explain the principle behind them
- Explain/use different kinds of effect size measures and confidence intervals
- Graphically present data in ways that are informative and not misleading
- Explain the basic principle behind Bayesian statistics and its pros and cons
- Critically evaluate whether there is a replication crisis in a given field, respond to such claims
- Reason through real-world research problems, evaluate alternative statistical approaches
- Prepare for an effective statistical consultation when help is needed
- Write a valid statistical plan for at least a few basic kinds of experiments

Grades will be based on

50% Participation (lectures and labs) including reading occasional assigned papers

25% Homework – numerical procedures (4 problem sets) and one “statistical plan”

25% An in-class written test on the last day (conceptual explanations)

You may use any method you want to solve numerical homework problems – pen and pencil, calculator, excel, programming in any language, or any statistical package you want, you just need to show your work. We will NOT teach or require MATLAB programming in this class.

You may re-do homework assignments as many times as you want

You will be allowed one chance to improve answers on the final written test

Principles of community

I expect all classroom discussions to be guided by principles of civil discourse, courtesy and respect. Participation is encouraged, so do speak up when you have thoughts or questions, but also take some responsibility for making sure all the other students have a chance to speak up and feel comfortable doing so.