BMEG 802 – Advanced Biomedical Experimental Design and Analysis

Introduction

Joshua G. A. Cashaback, PhD

WELCOME TO UD!

BMEG802: No Prerequisites

- See statistics and probability primer

Tuesday & Thursday, 3:30-3:45pm

- Lecture Days
- “Lab” Days

Instructor

Joshua Cashaback, Ph.D

- Ph.D. (Biomechanics, McMaster)
- Postdoctoral Fellow (Computational Neuroscience, Western)
- Postdoctoral Fellow (Neuromechanics, Calgary)
- Assistant Professor (Biomedical Engineering, Delaware)

Contact Information

- Office Hours: By Appointment (Zoom)
- Before and After Lectures (time permitting)
- During Labs
- Utilize TA office hours, labs, before and after class for questions
- email: joshcash@udel.edu (subject title: BMEG802)

Teaching Assistant

Adam Roth

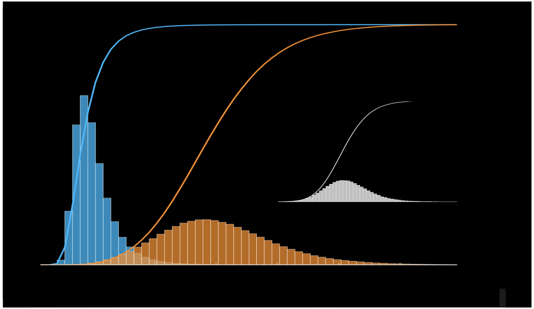
- Physics Undergrad
- PhD Candidate (Computational Neuroscience, Biomedical Engineering, Delaware)

Contact Information

- Office Hours: Wednesday (Zoom)
- Office: STAR Complex, 2nd floor
- Utilize labs, before and after class for questions
- email: aroth@udel.edu (subject title: BMEG802)

Course

Understanding statistical analyses is an essential skill for scientists in academia and industry. Here we will discuss hypothesis testing, simulate datasets via sampling, and perform parametric and nonparametric tests. In additional to traditional tests (mean comparison, regression, ANOVA) we will also introduce some advanced techniques (maximum likelihood, Bayes, bootstrapping, and MCMC).



Course Objectives

- Understand the role and philosophy of hypothesis testing
- rationale and limitations of statistical tests
- Identify the appropriate statistical tests to use.
- Perform basic to advanced statistical models.
- Build a stronger understanding of probability by using sampling and resampling techniques
- Develop an ability to write and communicate statistical results.

Modules

Module 1: Statistics and Probability Primer: We will review the building blocks of probability & statistics.

Module 2, Fundamentals: Next, we will discuss several basic tests such as regression, hypothesis testing, effect size, and power analysis.

Module 3, Omnibus Tests: This section will cover ANOVA (between, within, and mixed), ANCOVA, and nonparametric equivalents.

Module 4, Advanced Techniques: Finally, we will introduce some more advanced techniques, including maximum likelihood, Bayes' Theorem, bootstrapping, and MCMC.

Grading

- 70% Biweekly Assignments (7 x 10% each)
- 15% Midterm (Take home)
- 15% Final (Take home)

Grading Scale

Letter Grade	Percent Grade
A	93-100%
A-	90-92.99%
B+	87-89.99%
B	83-86.99%
B-	80-82.99%
C+	77-79.99%
C	73-76.99%
C-	70-72.99%
D+	67-69.99%
D	63-66.99%
D-	60-62.99%

Lectures

- I'll highlight main concepts
- My role is to introduce you to the basics of each topic and get you started with examples
- It is your responsibility to dig deeper when you need and / or want to
- Graduate studies – learning to be indepedent and thinking on your own!

Stats Courses Often Get a Rap

- Courses often focus on rote memorization of recipes of statistical tests
- Often left with little understanding of how and why (even after getting a high grade)
- Limited statistical repertoire
- Inability and / or desire to confront new statistical challenges in own research
- Science is about exploration! You should be able to, and want to, adapt what you know to learn new things

My High-Level Goals for You

- Understand Material (Visual Examples, Simulations, etc.)
- Learner Centered Environment (In class problem-sets during Labs)
- Actively Engage and Interact (Both sides, ‘connect-the-dots’)
- Improve Communication Skills
- Set you up for SUCCESS!

My Statistics Goals for You

- understand the logic and rationale behind statistical approaches
- enable to reason your way out of any statistical jam
- the goal here is NOT to rely on performing low-level arithmetic or memorizing equations (I can never remember these).
- remember the underlying concepts is far more important
- build a statistical repertoire
- use the appropriate test
- learn a high-level program to execute stats.

On your end. . .

- Please, be respectful of others (e.g., don't be disruptive during lectures, understand people have different perspectives and approaches).
- Submit assignments on time (10% deductions per day)
- Use office hours, labs, and class for questions. Don't wait until last minute to ask questions.
- Academic Integrity (see syllabus)
- Disputing marks (legitimate reasons only)
- First offering of course, so please give me feedback if something is unclear
- FUN LEARNING!

Your Goals in Grad School

- learning to explore, visualize, and modeling your data should not be a chore — this is the fun part of science!
- You should be excited about:
    i. your data
    ii. the idea of obtaining more data
    iii. looking at data in many different ways
    iv. thinking about a model for your data
    v. thinking about what your data means
    vi. exploring and analysing your data in new ways
    vii. thinking about how to make or change experiments to ask new questions.

Textbook

- none required
- Designing Experiments and Analysing Data: A Model Comparison Perspective (3rd Edition) by Scott E. Maxwell, Harold D. Delaney and Ken Kelley. Routledge (2017). ISBN: 978-1138892286 (recommended)
- Biostatistics: A Foundation for Analysis in the Health Sciences Wayne W. Daniel, Chad L. Cross, 11th Edition (used in years past; 10th addition free online)

Programming

- Use any language you please (R, Python, MatLab, SPSS, JMP, etc.)
- All coding examples will be done in R
- IN THIS ORDER, download:
  - https://www.r-project.org
  - https://rstudio.com/products/rstudio/download/
- Bring a laptop to lecture / lab, or pair up with someone that does

Assignments

- Rstudio, Latex or Word
- https://pages.uoregon.edu/koch/texshop/obtaining.html
- Well written and correctly interpret statistics.
- Spend time making nice graphs. You'll be grateful you did when you start publishing your research articles.
- https://ggplot2.tidyverse.org

QUESTIONS ABOUT THE CLASS?

SOME HISTORY ON THE PHILOSOPHY OF STATISTICS

MD Chapter 1 & 2

- Idea of pure science
- Philosophical stances on science
- Historical Review
- Gets you thinking abou the logic of science and experimentation

Assumptions

- Lawfulness of Nature
    a. regularities exist, can be discovered, and are understandable
    b. Nature is uniform
- Causality
    a. events have causes; if we reconstruct the causes, the event should occur again
    b. can we ever prove causality?
- Reductionism
    a. Can we ever prove anything? What is proof?

Assumptions

- Finite Causation
  a. causes are finite in number and discoverable
  b. generality of some sort is possible
  c. we don't have to replicate an infinite # of elements to replicate an effect
- Bias toward simplicity (parsimony)
  a. seek simplicity and distrust it
  b. start with simplest model: try to refute it; when it fails, add complexity (slowly)
- Progress occurs by falsifying theories

Logical Fallacy

- Fallacy of inductive reasoning (affirming the consequent)
  a. Predict: if theory T, then data will follow pattern P
  b. Observe: data indeed follows pattern P
  c. Conclude (Falsely): therefore theory T is tru
- example
  a. A cough is one symptom of COVID
  b. I have a sore throat
  c. Therefore, I have COVID

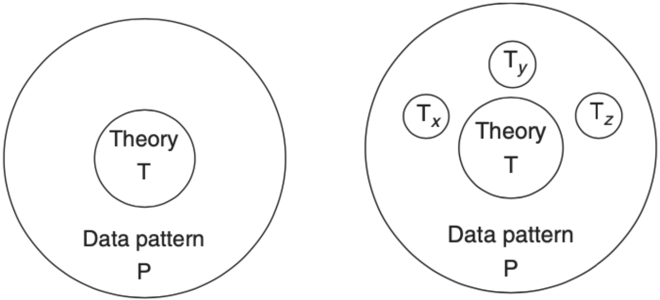
Of course other things besides COVID can cause a cough. Common cold. Or yelling. Or cancer.

Falsification is better

- Falsification:
  a. Predict: If theory T is true, then data will follow pattern P
  b. Observe: data do not follow pattern P
  c. Conclude: theory T cannot be true
- We cannot prove a theory to be true
- We can only prove a theory to be false

Falsification is better

- Falsification:
    a. Predict: If theory T is true, then data will follow pattern P
    b. Observe: data do not follow pattern P
    c. Conclude: theory T cannot be true
- We cannot prove a theory to be true
- We can only prove a theory to be false



Karl Popper

- Theories must have contrete predictions
- Constructs (measures) must be valid
- Empirical methodology must be valid

Basis of Interpreting Data

The Fisher tradition

- statistics is not mathematics
- statistics is not arithmetic or calculation
- statistics is a logical framework for:
  a. making decisions about theories
  b. based on data
  c. defending your arguements
- Fisher (1890-1962) was a central figures in modern approaches to statistics
- The F-test is named after him

Fundamental idea

- THE critical ingredient in an inferential statistical test (frequentist approach):
  a. determining the Probability, assuming the null hypothesis is true, of obtaining the observed data

Calculating Probabilities

- Calculation of probability is typically based on probability distributions
  a. discrete (e.g., binomial)
  b. continuous (e.g., z, t, F)
- We can also compute this probability without having to assume a theoretical distribution
  a. Resampling techniques (e.g., bootstrapping)

Basis of Interpreting Data

- design experiments so that inferences drawn are fully justified and logically compelled by the data
- theoretical explanation is different from the statistical conclusion
- Fisher's key insight
  a. randomization
  b. assures no uncontrolled factor will bias results of statistical tests

QUESTIONS???

Before next class

- Download R, Rstudio
- Perform R Resources
  i. RStart.pdf
  ii. RPLOT.pdf & States03.csv (read in csv files from folder too)
  iii. RDistributions.pdf
- Spend the time working through all these documents in detail before next class!
- Google is your friend
- Try. And then Try Again!
- Ask a classmate, before / after class, lab days, office hours.

Next class

- LAB DAY
    - Statistics and Probability primer (summary of some undergraduate level stats)