BIOMEDSCI 552: Statistical Thinking

Course Lead: Eric Lofgren
Email: Eric.Lofgren@wsu.edu
Office: AC 311

Credit Hours: 1
Semester: Fall 2024, Aug. 19th to Sept. 26th
Classroom: Wegner G1
Meeting Times: Tuesday & Thursday, 3:00 – 4:15 PM
Learning Management System: Canvas

Course Description

This course will introduce students to the core concepts and methods for *statistical thinking*, the skillset necessary to reason about data, how it is obtained, and how to use statistics to interrogate that data and build an understanding of the world around us. The intent of the course is to ensure that students approach the data analysis problems that will arise in their academic careers in a rigorous and thoughtful manner, rather than relying on rote algorithms to obtain a statistical result. Course material will focus on applied examples drawn from both laboratory-based biomedical science as well as population health science (i.e. epidemiology and disease ecology). Additionally, it will familiarize students with good coding and documentation habits early in their academic careers. This course is expressly *not* intended as a substitute for one or more classes in specific statistical methods.

Course Materials

There is no required textbook for this course. For specific lectures, a small set of relevant papers from the literature, online tutorials, etc. will be utilized. These will be provided ahead of class via Canvas.

Assessments

There will be three means of assessment for this course problem sets, class participation and discussion, and a take-home group final.

- Problem Sets: Three problem sets will be provided during a Thursday class, due before class the following Thursday posted to Canvas. These problem sets will require you to practice and demonstrate understanding of the concepts introduced in preceding classes, usually in the form of a set of applied problems. These problem sets may be done in small groups.

- **Class Participation:** Data analysis is rarely done in a vacuum, and team science requires interaction with your peers. You will be required to engage actively when discussing readings or problems in class, asking questions during lecture, etc.

- **Final Project:** You will select a topic of interest to you, based on your own work or as part of the material we have covered over the course of the semester. The purpose of this project is to identify a problem – including prior research in the area, and outline how a statistical approach may aid you in answering it. Your work will be presented as a 5-page final paper.

Assessment	**Percentage**
Problem Sets	40% (13.33% each)
Participation	30%
Final	30%

Grading: Letter grades will be given according to the following ranges:

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 (60–66.99%), F (less than 60%).

Expectations for Student Effort

For each hour of lecture equivalent, students should expect to have a minimum of two hours of work outside class.

Participation and attendance

Consistent attendance in this course is essential. Beyond the participation aspect of assessing your performance, the range of topics covered – and the pace at which they will be covered – leaves very little time for catch up. Discussions in class will often extend the material and discussion beyond the readings or lecture slides. Additionally, consistent attendance and active participation will allow you to meet and interact with your colleagues who, while possessing different academic backgrounds and pursuing research in several different areas, are all part of the same graduate program.

Policy on Late Assignments

Assignments that are submitted after the beginning of class on Thursday will receive one-half of the total graded points (i.e. a maximum of 5% of your total grade). A late final will lose 10% off the final grade for each day late it is, beginning the day after it is due at 12:01 AM.

Learning Outcomes

Students who successfully complete this course will be able to:

- Engage with faculty, colleagues and collaborators from a broad range of disciplines working in the immunology and infectious diseases field
- Utilize modern software engineering tools used in the field of data analysis
- Build familiarity the core mathematical, statistical and computational concepts underlying statistical methods
- Place their own subject matter problems in a statistical context

Collaboration Policy

Collaboration is encouraged on all problem sets, and necessarily a component of your class participation evaluation. While you are welcome to work in groups for the problem sets, the work you submit should be original to you – that is, it should be written in your own words, reflecting your own understanding of the solutions your group arrived at. If you do elect to work in a group, please list the names of the other students you worked with for your assignment. Your final project must be completed independently.

The use of solutions from sources outside the class, such as internet help sites, will be considered a violation of the academic integrity policy.

Weekly Course Outline

August 20: Computational Tools in Data

This week will introduce the R programming language, including an introduction to several key libraries as well as how to install them, as well as the use of the version control system Git and the online repository management system GitHub.

August 22: Thinking About and Interacting with Data

This week will explore the conceptual foundations of data itself – what is data, and how can it be categorized? What is the difference between prospectively and retrospectively collected data? Between active and passively collected data? What is "Big Data" and how has it impacted both data science and the fields that use data science techniques?

August 27th: No Class

August 29th: No Class. Problem Set 1 Due.

Sept 3rd: Probability, Part 1

Probabilistic questions underly most, if not all, questions in data analysis. This week will cover the basics of probability theory, including independence, conditional probability, and Bayes' Theorem.

September 5th: Probability, Part 2

This week will cover discrete and continuous random variables, pdfs and cdfs, as well as the common distributions encountered in the biomedical sciences (e.g., Normal distributions, Poisson distributions, and exponential distributions).

September 10th: Sampling

This week will cover the concept of sampling – what is it, how is it performed, what is a sampling distribution, and how the concept of sampling impacts how we approach designing and analyzing experimental and observational data.

September 12th: Estimation. Problem Set 2 Due

This week covers the concept of estimation – what is an estimator, bias and variance, as well as how uncertainty is captured and quantified.

September 17: Basic Statistics and Exploratory Data Analysis

Exploratory data analysis (EDA) is a collection of methods to explore, assess and visualize the properties of your data to help inform your analytical approach. This course will cover the basic methods of EDA, including data visualization and basic summary statistics.

September 19th: Hypothesis Testing

Hypothesis testing is at the core of many (but by no means all) study designs in the biomedical sciences. This week covers what it means to test a hypothesis, Type I and Type II error, the notion of a *p-value*, and several applied examples.

September 24th: Parametric vs. Nonparametric Statistics

Much of what we have covered so far in this course considers *parametric* statistics, where the goal of our analysis is to estimate a parameter from the population. In this week, we cover what a parameter is, what parametric statistics are, and the circumstances under which we may wish to use non-parametric statistics, as well as what information we can glean from that.

September 26th: Bayesian Reasoning and Causal Inference. Problem Set 3 Due

Many models in data science seek to go beyond describing or predicting the relationship between two variables to instead make causal arguments about them. This week will cover the basics of causal inference in data science, as well as provide an overview of Bayesian methods – a collection of techniques that supplement the estimates from a data set with prior knowledge.

COVID-19 Policy

Students are expected to abide by all current COVID-19 related university policies and public health directives. These directives may be adjusted to respond to the evolving COVID-19

pandemic. Directives may include, but are not limited to, compliance with WSU's COVID-19 vaccination policy, wearing a cloth face covering, physically distancing, and sanitizing common use spaces. All current COVID-19 related university policies and public health directives are located at https://wsu.edu/covid-19/. Students who do not comply with these directives may be required to leave the classroom; in egregious or repetitive cases, student non-compliance may be referred to the Center for Community Standards for action under the Standards of Conduct for Students.

Assigning Incompletes

University policy (Acad. Reg. #90) states that Incompletes may only be awarded if: "the student is unable to complete their work on time due to circumstances beyond their control".

WSU Reasonable Accommodation Statement

Students with Disabilities: Reasonable accommodations are available for students with a documented disability. If you have a disability and need accommodations to fully participate in this class, please either visit or call the Access Center to schedule an appointment with an Access Advisor. All accommodations MUST be approved through the Access Center. For more information contact a Disability Specialist on your home campus. Pullman or WSU Online: 509-335-3417, Washington Building 217; http://accesscenter.wsu.edu, Access.Center@wsu.edu.
All students requesting reasonable accommodation must meet with the instructor prior to or during the first week of the course to review all proposed accommodations in relation to course content and requirements. Exceptions to this timeframe will be granted only upon a showing of good cause.

WSU Academic Integrity Statement

Academic integrity is the cornerstone of higher education. As such, all members of the university community share responsibility for maintaining and promoting the principles of integrity in all activities, including academic integrity and honest scholarship. Academic integrity will be strongly enforced in this course. Students who violate WSU's Academic Integrity Policy (identified in Washington Administrative Code (WAC) 504-26-010(3) and -404) will fail the course, will not have the option to withdraw from the course pending an appeal, and will be reported to the Office of Student Conduct.
Cheating includes, but is not limited to, plagiarism and unauthorized collaboration as defined in the Standards of Conduct for Students, WAC 504-26-010(3). You need to read and understand all of the definitions of cheating: http://app.leg.wa.gov/WAC/default.aspx?cite=504-26-010. If you have any questions about what is and is not allowed in this course, you should ask course instructors before proceeding. If you wish to appeal a faculty member's decision relating to academic integrity, please use the form available at https://conduct.wsu.edu/.

Accommodation for Religious Observances or Activities

Washington State University reasonably accommodates absences allowing for students to take holidays for reasons of faith or conscience or organized activities conducted under the auspices of a religious denomination, church, or religious organization. Reasonable accommodation

requires the student to coordinate with the instructor on scheduling examinations or other activities necessary for course completion. Students requesting accommodation must provide written notification within the first two weeks of the beginning of the course and include specific dates for absences. Approved accommodations for absences will not adversely impact student grades. Absence from classes or examinations for religious reasons does not relieve students from responsibility for any part of the course work required during the period of absence. Students who feel they have been treated unfairly in terms of this accommodation may refer to Academic Regulation 104 – Academic Complaint Procedures.

Classroom Safety Statement

Classroom and campus safety are of paramount importance at Washington State University, and are the shared responsibility of the entire campus population. WSU urges students to follow the "Alert, Assess, Act," protocol for all types of emergencies and the "Run, Hide, Fight" response for an active shooter incident. Remain ALERT (through direct observation or emergency notification), ASSESS your specific situation, and ACT in the most appropriate way to assure your own safety (and the safety of others if you are able).
Please sign up for emergency alerts on your account at MyWSU. For more information on this subject, campus safety, and related topics, please view the FBI's Run, Hide, Fight video and visit the WSU safety portal.

Lauren's Promise

I will listen and believe you if someone is threatening you.

Lauren McCluskey, a 21-year-old honors student athlete, was murdered on Oct. 22, 2018, by a man she briefly dated on the University of Utah campus. We must all take actions to ensure that this never happens again.

If you are in immediate danger, call 911.

If you are experiencing sexual assault, domestic violence, and stalking, please report it to me and I will connect you to resources or call the National Alternatives to Violence at 877-334-2887 (24-hour crisis hotline).

Any form of sexual harassment or violence will not be excused or tolerated at Washington State University. WSU has instituted procedures to respond to violations of these laws and standards, programs aimed at the prevention of such conduct, and intervention on behalf of the victims.

- National Resources for Domestic Violence, Sexual Assault, and Stalking:
- National Domestic Violence Hotline: 1-800-799-SAFE
  and https://www.thehotline.org/help/
- RAINN (Rape, Abuse & Incest National Network): 1-800-656-HOPE (4673)
  and https://www.rainn.org/
- Victim Connect Resource Center: 1-855-484-2846 and https://victimconnect.org/

- Love is Respect – National Dating Abuse Hotline: 1 (866) 331- 9474
    - Text: 22522 and https://www.loveisrespect.org/
- Anti-Violence Project Hotline: 212-714-1141 and https://avp.org/get-help/