Stat 61

Swarthmore College

2022 Fall with Prof Suzy

Calendar (edit 11/9/22)

Week	Topics/Chs to read	M	W	F
1	Review; Ch 1-6	Aug 29 Classes begin	Hw 1 due	Hw 2 due
2	Review; Ch 1-6	Sept 5 No class (holiday)	Hw 3 due	Drop/add ends; Hw 4 due
3	Estimation Part I; Ch 8.1, 8.2, 8.3, 8.4, 8.5.1, 8.6.1	Sept 12 Hw 5 due	Hw 6 due	**Quiz 1**
4	Estimation Part II; Ch 8.5.2, 8.7, 8.8	Sept 19	Hw 7 due	Hw 8 due
5	Estimation Part III; Ch 8.5.3, 8.6.2, 8.6.3, 8.9	Sept 26 Hw 9 due	Hw 10 due	Final exam schedule available;
6	Hypothesis Testing Part I; Ch 9.1, 9.2	Oct 3 **Quiz 2**		Hw 11 due
7		Oct 10 **Fall break**	**Fall break**	**Fall break**
8	Hypothesis Testing Part II; Ch 9.3, 9.4	Oct 17	Hw 12 due	HW 13 due
9	Hypothesis Testing Part III; Ch 9.5, 9.10	Oct 24 HW 14 due	Next semester course schedule available; Hw 15 due	**Quiz 3;** **Traveling**
10	Comparing Two Samples and Experimental Design;	Oct 31 Advising;	Advising;	Advising; Last day to

	Ch 11.1, 11.2, 11.4, 11.5		Hw 16 due	declare CR/NC; Hw 17 due
11	Analysis of Variance Part I; Ch 12.1, 12.2	Nov 7 Advising	Advising; Hw 18 due	Advising; **Quiz 4; Traveling**
12	Analysis of Variance Part II; Ch 12.3, 12.4	Nov 14 Pre-enrollment; Hw 19 due	Pre-enrollment	Hw 20 due
13	Analysis of Categorical Data; Ch 13.1, 13.2, 13.3, 13.4	Nov 21	Hw 21 due	No class (holiday)
14	Linear Least Squares Part I; Ch 14.1, 14.2, 14.3	Nov 28	Hw 22 due	Hw 23 due
15	Linear Least Squares Part II; Ch 14.4, 14.5	Dec 5 Hw 24 due	Last day of class **Quiz 5**	
	Dec 11	**Final and resubmitted quizzes due Sci 158 2:00pm-5:00pm**		

Syllabus

E-mail: sthornt1@swarthmore.edu
Office: SCI 136
Office Hours: M 4:00-5:00pm W 4:00-5:00pm and F 3:00-4:00pm. Please note I will not be available for W or F office hours on Weeks 9 and 11 due to travel commitments.

Class Meeting Times: M/W/F 9:30-10:20am or 11:30−12:20pm
Class Meeting Location: SCI 102

Course Description

Introduction to the mathematical theory of frequentist and Bayesian statistical inference. Topics include parameter estimation, confidence intervals and hypothesis testing, linear regression methods and some basic Bayesian methods. This course will assume you are familiar with STAT 011 topics and have some prior experience with computing. The textbook we will use in class is *Mathematical Statistics and Data Analysis*, 3rd edition, by John A Rice. *You are expected to read along in your textbook each week according to the calendar above*. The class format will begin in a traditional lecture-heavy manner (for Unit 1) but will transition gradually to include more active learning sessions than lectures towards the end of the semester (by Unit 3).

Prerequisite: A grade of C or better in both STAT 051 and MATH 027 (or 028); or permission of the instructor.

Expected knowledge from prior statistical education that will be reinforced in the beginning of the semester

- Demonstrate proficiency with the theory of probability and concept of random variables; familiarity with common distributions and transformations; ability to analytically derive moments of common distributions; understanding of central limit theorem as a theory and in practice.
- Familiarity with sampling considerations and principles of experimental design; ability to plan and execute informative descriptions to accurately summarize data including graphical methods.
- Proficiency with some sort of statistical analysis software.

Learning objectives for Stat 61

1. Understand frequentist and Bayesian methods for parameter estimation and common approaches to evaluate and compare estimators.
2. Work with asymptotic theorems to characterize the behavior of common types of estimators.
3. Understand how to analytically derive a Bayesian posterior distribution and how to interpret a Bayesian credible interval and demonstrate familiarity with different types of priors.
4. Understand the important role of likelihood functions for hypothesis testing in both Bayesian and frequentist frameworks
5. Understand the relationship between frequentist confidence interval estimation and hypothesis testing.
6. Familiarity with common types of optimal testing strategies.
7. Construct and interpret frequentist p-values for hypothesis tests and interpret the error rates.
8. Identify/define model parameters and state statistical inferential questions in terms of these parameters for various common, realistic study settings.
9. Contextualize statistical methods and theory in science and policy at large and develop a habit of mind informed by the stewardly application of such methods.

Grade components

- Daily homework - Worth 20% of overall course grade
  You will be assigned daily homework problems sets consisting of three problems each. These homework sets will be submitted to Gradescope. One of the questions will be graded for correctness, the other two for completion. This will be determined randomly for each problem set. Some homeworks (but not all) will require the use of software. Written or typed homework solutions are both acceptable. Your overall homework grade will be either the mean or median of your individual scores (whichever is highest).

Homework is due (to Gradescope) by the beginning of our next class period however, each student is permitted *a late pass* to use twice during the semester. You are encouraged to work with your classmates on your homework but you are expected to hand in your own unique solution. Assignments with any answers that appear to be copied versions of another's will receive a grade of zero for all matching submissions.

- Semesterly participation - Worth 10% of overall course grade
  More information on this portion of your grade will be given in class.
- 5 Quizzes - Each worth 10% of overall course grade
  These quizzes reinforce the previous lessons and homework assignments as well as prepare you for the material lying ahead. You will have 25 minutes in class to begin working on each quiz and you can take it home with you to complete by the end of the day. Your solutions must be uploaded to Gradescope by midnight that evening. You are not allowed to seek external help from anyone on your quizzes but you may reference your textbook and class notes on your own. All problems will be graded for correctness and I will try to have your grades posted to Gradescope within one week.
- Final test - Worth 20% of overall course grade
  The final test will be cumulative and closed-book but you will be permitted use of a **single double-sided** sheet of personal notes. The final will occur during finals week sometime between Dec 7-11 and no make ups will be given.

Quiz and test policies

During quizzes and tests, you are welcome to use the restroom facilities as you need but I ask that you take turns so only one person is gone from the classroom at a time. I also request all students to put their phones on their desk upside down and leave them there when using the facilities.

Make-up quizzes and tests will **not** be given unless you have notified me at least three weeks in advance of the date or in case of an extreme, urgent emergency.

Finally, although there is built-in flexibility with the due dates of homework assignments, I strongly recommend that you keep up with the homework and reading assignments each week. This will provide you with a framework to better follow along with material in class and will better prepare you for the quizzes and final test.

Office hours etiquette

I encourage you to attend office hours as often as possible! If you are unable to attend my regular office hours then please send me an email so we can arrange another time to meet. Please note that I can only make these accommodations for *academic* scheduling conflicts.

I want to make sure that everyone who shows up gets a chance to ask a question. I prioritize student questions usually by whoever arrives first but sometimes will poll students and prioritize questions that more than one student has in common.

When it's your turn to ask a question, please limit yourself to only one question (even if it is a question with a very short answer). Once everyone present has gotten a chance to ask their question, you will get the opportunity to ask another question.
When others are asking questions, pay attention because you might learn something from the discussion. In fact, when the same question keeps getting repeated in office hours, I turn to students who asked the question before to answer the question when it arises again.

Support outside of class

Peer led Stat Clinics are drop-in study sessions supporting Stat 1, 11, 21, 61. Stat Clinics Session are run by friendly and knowledgeable upperclassmen every night Tuesday, Wednesday, and Thursday night 7-10 pm. Clinics are a wonderful opportunity to study, do homework, meet/work with classmates, and ask questions about statistics and math. Because clinics are drop-in, you are welcome to come and go as you please. Please write your name and course in the sign-in binder and the time you enter and leave the clinic so the Math/Stat Department has a record of your attendance. To make the most of your time at clinic, be sure to first try problems on your own, or bring questions you have from your text or lecture. Having your textbook, lecture notes, and online resources handy is essential because these are helpful resources for both you and the Clinician working with you. There will likely be other students at Clinic with questions for the Clinician, so do not expect to get individual attention the entire time you are at clinic. Be open to working on other problems, thinking about and trying to work through the question you have for the Clinician, working with classmates, or doing other coursework while you wait to speak with the Clinician. For questions about Math & Stat Clinics please visit https://www.swarthmore.edu/math-stat-academic-support/math-and-stat-clinics or contact Laura Dandridge ldandri1@swarthmore.edu, Academic Support Coordinator for the Math/Stat Department.

Stat Clinic 7-10 pm

Tuesday	Jay Leeds	SC 149	Stat 1, 11, 21, 61, Math 66, Math 39
Wednesday	Tarang Saluja	SC 158	Stat 1, 11, 21, 61, Math 69, Math 39
Thursday	Ella Yadav	SC 149	Stat 1, 11, 21, 61

If you are wanting to refine your mathematical skills more generally, the Math/Stat department provides many fun and engaging opportunities to do so! Consider attending Putnam Problem Solving meetings, the Puzzle Solving Collective, GRE Prep Sessions, or help out with the S3P program. Feel free to ask me how to point you in the right direction to get involved with these groups.

And, for more practice with statistical software and data mining, sorting, and visualization techniques, consider attending the Stat Problem Solving sessions which meet about once a month during the regular semester. This year's sessions will be hosted by yours truly.

Course feedback

I will solicit explicit feedback from you twice throughout the semester. These assessments are identical and anonymous. The first time I request your feedback will be about halfway through the semester. Your feedback here helps me identify what is and isn't working for you and make reasonable adjustments accordingly. The last time I request your feedback is at the end of the semester. You are welcome to submit feedback before or after the final exam. This helps me prepare for future sessions of this course.

As you know well, this type of feedback, although informative, is heavily biased. Not only is volunteer response bias inevitable, there are studies that indicate strong emotional and social biases influence course evaluations based on identity factors such as the instructor's race, gender, sexuality, and age and based on external factors that influence respondents' emotions related to the course material. (See, for example [1] and [2], among others.) Because of this, I ask that you keep in mind what you learn about stewardship and ethical statistical practices while filling out these surveys (and other future surveys). For many of you, this is one of the first real-world opportunities you will encounter to apply your deepening statistical knowledge and understanding. (If interested, here is a quick guide for students who wish to avoid bias in their course evaluations.)

If you find that you have more feedback and ideas that you'd like to share with the department at large, please consider joining one of the Math/Stat department student groups. Currently these groups consist of Gender Minorities in Math/Stat (GeMS), Black in Math/Stat (BIMS), and the Math/Stat Student Advisory Council (MSSAC). Participation in these groups is a great way to communicate with the Math/Stat faculty and staff directly as we work together to build a more equitable learning environment for all students. Feel free to ask me about how to get involved!

References

[1] Fan Y, Shepherd LJ, Slavich E, Waters D, Stone M, Abel R, et al. (2019) Gender and cultural bias in student evaluations: Why representation matters. PLoS ONE 14(2). doi:10.1371/journal.pone.0209749.

[2] Mitchell, K., & Martin, J. (2018). Gender Bias in Student Evaluations. PS: Political Science & Politics, 51(3), 648-652. doi:10.1017/S104909651800001X.

Study guide

Unit One - Thoughtful use of estimation techniques

Chapters 1-3

Sections covered: None
Learning objectives: Expected knowledge from prior statistical education that will be reinforced in the beginning of the semester

Notes: This material covers topics that are heavily explored in the prereq Stat 51. We will *not* spend any time in class going over these topics however you *are* responsible for the terms and theorems in these sections. You may want to review the table of contents and bolded terms and main theorems of these chapters as a refresher.

Chapter 4

Sections covered: 4.1 - 4.5
Learning objectives: Expected knowledge from prior statistical education that will be reinforced in the beginning of the semester
Notes: We will spend a little bit of class time reviewing some of the parts of this chapter with a particular emphasis on material that will be useful when we get to Ch 8. The material for this unit will be presented mainly in traditional lecture format but we will start to have regular time in class for group work and problem solving now.

Chapters 5-6

Sections covered: All
Learning objectives: Expected knowledge from prior statistical education that will be reinforced in the beginning of the semester; 2) Work with asymptotic theorems to characterize the behavior of common types of estimators.
Notes: These are very short chapters but they are still important. In class we will emphasize material from Ch 5 on the Law of Large Numbers and Central Limit Theorem but you are expected to be familiar with all of the material in these chapters. We will start to work with a particular analytical tool called a "stakeholder analysis" to delve into these topics.

Chapter 8

Sections covered: All
Learning objectives: 1) Understand frequentist and Bayesian methods for parameter estimation and common approaches to evaluate and compare estimators; 2) Work with asymptotic theorems to characterize the behavior of common types of estimators; 3) Understand how to analytically derive a Bayesian posterior distribution and how to interpret a Bayesian credible interval and demonstrate familiarity with different types of priors.
Notes: We are covering this chapter in a different order than the sections are presented in your textbook so please follow the course calendar. We will continue to use a stakeholder analysis to help reach the learning objectives.

Unit Two - Statistical inference and stewardship

Chapter 9

Sections covered: 9.1 - 9.5, 9.10
Learning objectives: 4) Understand the important role of likelihood functions for hypothesis testing in both Bayesian and frequentist frameworks; 5) Understand the relationship between frequentist confidence interval estimation and hypothesis testing; 6) Familiarity with common types of optimal testing strategies; 7) Construct and interpret frequentist p-values for hypothesis tests and interpret the error rates.

Notes: We are going to spend at least a third of the semester on this chapter alone. Even though we won't cover every single section of the chapter, the material here is deep and is very important for any statistical practitioner to understand. Here you will start to understand what it means to be a steward of statistics and why that is important.

Unit Three - Disciplinary best practices for common study designs and analyses

Chapter 11

Sections covered: 11.1 - 11.2, 11.4 - 11.5
Learning objectives: 8) Identify/define model parameters and state statistical inferential questions in terms of these parameters for various common, realistic study settings; 9) Contextualize statistical methods and theory in science and policy at large and develop a habit of mind informed by the stewardly application of such methods.
Notes: This is the first chapter where we get into more specific (but generally applicable) statistical tests and methods. We will introduce the American Statistical Association's Guidelines for Ethical Practice here as we cover ways to compare two samples.

Chapter 12

Sections covered: All
Learning objectives: 8) Identify/define model parameters and state statistical inferential questions in terms of these parameters for various common, realistic study settings; 9) Contextualize statistical methods and theory in science and policy at large and develop a habit of mind informed by the stewardly application of such methods.
Notes: We will continue to reference the ASA's ethical guidelines with respect to ANOVA methods. At this point in the semester, there will be less material covered in lecture-format and more time in class working through problems together.

Chapter 13

Sections covered: 13.1 - 13.4
Learning objectives: 8) Identify/define model parameters and state statistical inferential questions in terms of these parameters for various common, realistic study settings; 9) Contextualize statistical methods and theory in science and policy at large and develop a habit of mind informed by the stewardly application of such methods.
Notes: We will continue to reference the ASA's ethical guidelines with respect to the analysis of categorical data. Most class time will be collaborative problem solving rather than lectures.

Chapter 14

Sections covered: 14.1 - 14.4; 14.8
Learning objectives: 8) Identify/define model parameters and state statistical inferential questions in terms of these parameters for various common, realistic study settings; 9) Contextualize statistical methods and theory in science and policy at large and develop a habit of mind informed by the stewardly application of such methods.
Notes: We will continue to reference the ASA's ethical guidelines with respect to linear regression. Most class time will be collaborative problem solving rather than lectures.