AMS 394: Statistical Laboratory

Fall 2023

Course Information:

Instructor: Kai Li
Office: Harriman 202
Class: Tue, Thu 8:30-9:50am, Melville Lbry. W4540
Email: kai.li@stonybrook.edu
Office Hours: Mon, Wed 2:00-3:30pm (in person) or by appointment (online via Zoom)
Course Page: Brightspace

Graduate TA: TBD
Office: Harriman 202
Email: TBD
Office Hours: TBD

Undergraduate TAs:	Zhiqing Chen	Jeffrey Liu
Office:	Harriman 202	Online via Zoom
Email:	zhiqing.chen@stonybrook.edu	jeffrey.liu.1@stonybrook.edu
Office Hours:	Tue, Thu 4:00-5:00pm	Mon 9:00-11:00am

Course Description: Designed for students interested in statistics and their applications. Basic statistical techniques including sampling, design, regression, and analysis of variance are introduced. Includes the use of statistical packages such as SAS and R. Students translate realistic research problems into a statistical context and perform the analysis.

Prerequisites: AMS 310 or AMS 315

Textbooks:

- *Introductory Statistics with R* (ISwR) by Peter Dalgaard, 2nd edition, 2008, Springer
- *Statistical Computing with R* (SCwR) by Maria L. Rizzo, 2nd edition, 2019, Chapman & Hall/CRC
- *Nonparametric Statistical Methods* (NSM) by Myles Hollander, Douglas A. Wolfe and Eric Chicken, 3rd edition, 2013, Wiley

Topics to be Covered: Chapters 1, 2, 3, 4, 5, 6, 7 (excluding Sections 7.3 and 7.4), 8, 11 from ISwR, Chapter 3 (excluding Sections 3.3 and 3.6) from SCwR, and Sections 2.1, 2.2, 2.3, 3.1, 3.7, 4.1, 6.1, 8.1, 8.5, 10.1, 10.2 from NSM.

Homework: Assignments will be given weekly or bi-weekly, posted on Brightspace, to be turned in via Brightspace, as a single PDF file, and must be legibly composed with R Markdown. No late homework will be accepted and will be due at the beginning of class on the due date. If you want to correct your submission (before the deadline), you may upload a revision. Only your last submission will be graded. There will be approximately 7 homework sets, equally weighted. The lowest homework score will be dropped before computing your average.

You may discuss homework problems with other students taking the course and with the instructor and teaching assistants. The work that you turn in should always be your own type-up, and you should show that you personally understand everything that you type. You are not to view other student typeups or use them (or use internet resources) while typing your own solution. Please make certain that your typing is neat and clear, and that you have expressed your reasoning.

Exam: There will be three in-person exams. Midterm 1 will be during class, tentatively scheduled for Thursday, October 5. Midterm 2 will be during class, tentatively scheduled for Thursday, November 9, and will be noncumulative. The final will be held during the university-assigned date/time, on Wednesday, December 12, 11:15-1:45pm, and will be cumulative, but will emphasize more on the recent material; unless otherwise announced, the location will be our regular classroom (Melville Lbry. W4540). All exams are open book and open notes, according to instructions that will be provided with the exams.

Optional Project: There will be an optional individual project for all students, which essentially counts for extra credit (if suitably completed). The project will be assigned tentatively on Tuesday, October 24 at 8:30am. The tentative due date for the project is Friday, December 8 at 11:59pm. Please turn in your project as a single PDF file on Brightspace.

Grading Policy: Your total average score will be calculated based on 15% homework, 25% midterm 1, 25% midterm 2, 35% final, and 5% project. Final grades may possibly be curved based on the overall performance of the class, but such curving may only fairly apply to the entire class, not individually.

Learning Outcomes: Upon successful completion of this course, students should be able to:

- Demonstrate proficiency in handling basic knowledge and operations in R.
- Understand how to read and manipulate data in R.
- Calculate, interpret, and communicate measures of central tendency, dispersion, and graphical representations of data in R.
- Develop a comprehensive understanding of different techniques for generating random variables to conduct simulations and statistical analysis in R.
- Create effective visualizations, such as histograms, box plots, scatterplots, and bar charts, to summarize and communicate statistical findings in R.
- Conduct and interpret t tests and Wilcoxon tests in R.
- Learn how to fit a regression model, interpret its parameters, perform model selection, diagnostic checking, and make predictions in R.
- Calculate, interpret, and conduct hypothesis tests for parametric and nonparametric correlation coefficients in R.
- Perform one-way ANOVA, interpret ANOVA results, and conduct post-hoc tests in R.
- Conduct and interpret proportion tests and chi-square tests in R.
- Develop skills in clearly and effectively communicating statistical results, both orally and in written form, using appropriate statistical terminology and concepts.

Student Accessibility Support Center Statement: If you have a physical, psychological, medical, or learning disability that may impact your course work, please contact the Student Accessibility Support Center, Stony Brook Union Suite 107, (631) 632-6748, or at sasc@stonybrook.edu. They will determine with you what accommodations are necessary and appropriate. All information and documentation is confidential.

Academic Integrity Statement: Each student must pursue his or her academic goals honestly and be personally accountable for all submitted work. Representing another person's work as your own is always wrong. Faculty is required to report any suspected instances of academic dishonesty to the Academic Judiciary. Faculty in the Health Sciences Center (School of Health Professions, Nursing, Social Welfare, Dental Medicine) and School of Medicine are required to follow their school-specific procedures. For more comprehensive information on academic integrity, including categories of academic dishonesty please refer to the academic judiciary website at `http://www.stonybrook.edu/commcms/academic_integrity/index.html`

Critical Incident Management: Stony Brook University expects students to respect the rights, privileges, and property of other people. Faculty are required to report to the Office of Student Conduct and Community Standards any disruptive behavior that interrupts their ability to teach, compromises the safety of the learning environment, or inhibits students' ability to learn. Faculty in the HSC Schools and the School of Medicine are required to follow their school-specific procedures. Further information about most academic matters can be found in the Undergraduate Bulletin, the Undergraduate Class Schedule, and the Faculty-Employee Handbook.

Tentative Course Schedule

Week	Date	Text	Topic	Notes
1	Tue. 8/29	ISwR	Ch1: Basics	
	Thu. 8/31	ISwR	Ch1: Basics	
2	Tue. 9/5	ISwR	Ch2: The R Environment	HW1 due 9/12
	Thu. 9/7	ISwR	Ch2: The R Environment	
3	Tue. 9/12	ISwR	Ch3: Probability and Distributions	HW2 due 9/26
	Thu. 9/14	ISwR	Ch3: Probability and Distributions	
4	Tue. 9/19	SCwR	Ch3: Methods for Generating Random Variables	
	Thu. 9/21	SCwR	Ch3: Methods for Generating Random Variables	
5	Tue. 9/26	ISwR	Ch4: Descriptive Statistics and Graphics	HW3 due 10/3
	Thu. 9/28	ISwR	Ch4: Descriptive Statistics and Graphics	
6	Tue. 10/3		Review	
	Thu. 10/5		Midterm 1	
7	Tue. 10/10		No Class (Fall Break)	
	Thu. 10/12	ISwR	Ch5: One- and Two-sample Tests	HW4 due 10/24
8	Tue. 10/17	ISwR	Ch5: One- and Two-sample Tests	
		NSM	Ch3: The One-Sample Location Problem	
	Thu. 10/19	ISwR	Ch5: One- and Two-sample Tests	
		NSM	Ch3: The One-Sample Location Problem	
			Ch4: The Two-Sample Location Problem	
9	Tue. 10/24	ISwR	Ch6: Regression and Correlation	Project due 12/8
	Thu. 10/26	ISwR	Ch6: Regression and Correlation	HW5 due 11/7
10	Tue. 10/31	ISwR	Ch6: Regression and Correlation	
		NSM	Ch8: The Independence Problem	
	Thu. 11/2	ISwR	Ch11: Multiple Regression	
11	Tue. 11/7		Review	
	Thu. 11/9		Midterm 2	
12	Tue. 11/14	ISwR	Ch7: ANOVA & The Kruskal-Wallis Test	HW6 due 11/28
	Thu. 11/16	ISwR	Ch7: ANOVA & The Kruskal-Wallis Test	
		NSM	Ch6: The One-Way Layout	
13	Tue. 11/21	ISwR	Ch7: ANOVA & The Kruskal-Wallis Test	
	Thu. 11/23		No Class (Thanksgiving Break)	
14	Tue. 11/28	ISwR	Ch8: Tabular Data	HW7 due 12/7
		NSM	Ch2: The Dichotomous Data Problem	
	Thu. 11/30	ISwR	Ch8: Tabular Data	
		NSM	Ch10: Comparing Two Success Probabilities	
15	Tue. 12/5	ISwR	Ch8: Tabular Data	
	Thu. 12/7		Review	
16	Tue. 12/12		Final	11:15-1:45pm