RUTGERS

STATISTICS IN BIOMEDICAL SCIENCES

CTSC 5103S / SGS 16:115:557 (3 Credits)

OVERVIEW

TIMES: Fall 2024 semester on Fridays from 12:30-3:30 PM

LOCATION: Research Tower, Room MED-V12, 675 Hoes Lane West, Piscataway, NJ. This is a relatively small class. In-person class discussions with other students and me are essential for your critical thinking of stat problems.

TEXTBOOK: ***Biostatistics: A Foundation for Analysis in the Health Sciences, 11th Edition*** (2019) by Wayne W. Daniel, Chad L. Cross. ISBN: 978-1-119-28237-2.
https://www.amazon.com/Biostatistics-Foundation-Analysis-Probability-Statistics/dp/1119282373

Suggest you purchase it. The class notes will focus on examples and applications. For detailed explanation, please read the textbook.

Required Software: **R: A Language and Environment for Statistical Computing**
FREE Available from https://www.r-project.org
R Studio, also named **posit**
FREE Available from www.rstudio.com

Instructions how to download and install both **R** and **R Studio**:
https://teacherscollege.screenstepslive.com/a/1108074-install-r-and-r-studio-for-windows

Please install **R** and **R Studio** before your first class and bring your laptop to class for hands-on lab work in every class. A lab time is reserved in each class for your hands-on practice. Sample R codes will be provided for the class materials. No R coding experience is required to register this class.

GRADING

WEIGHTED AVERAGE

	Weight		Grade	Final Score
Attendance & Class Participation	5%		A	90.00 - 100.00
Homework	40%		B+	84.00 – 89.99
Midterm Exam	20%		B	75.00 - 83.99
Class Project	20%		C+	70.00 - 74.99
Presentation	15%		C	65.00 - 69.99
			F	< 65.00

COURSE DESCRIPTION

DESCRIPTION

RUTGERS

This is a graduate level applied statistics course designed for PhD students in the Biomedical Sciences graduate programs in New Brunswick/Piscataway to ensure *Rigor and Reproducibility* training for conducting research.

In this course, we will focus on building your foundation in statistical handling for your biomedical research including rigorous experimental design to answer the question of interest, appropriate analysis methods and accurate interpretation of findings. The course will cover the fundamental statistics methods in biomedical research including estimation, hypothesis testing (parametric and non-parametric), ANOVA, and regression methods (normal, binary, and time-to-event variables).

For more information regarding *Rigor and Reproducibility* training, refer to https://grad.rutgers.edu/academics/academic-enrichment-programs/rigor-reproducibility-training

ATTENDANCE, HOMEWORK, MIDTERM, CLASS PROJECT AND PRESENTATION

ATTENDANCE (5%)

Attendance is of *paramount* importance. The discipline of statistics is essentially a cumulative one, so catching up will be *very* difficult if you fall behind. Attendance and class participation takes 5% of weight. The classroom is facilitated as discussion friendly. Class discussion and lab work are major components of the course rather than just listening to lectures.

Each lecture consists of 3 small sessions (50 minutes each) and two breaks (15 minutes each). The class will promptly start and end on time per schedule. It's a long lecture, so bring some snacks & drink to stay focus. Also bring your laptop. You are expected to run the class notes in R markdown file in class during lecture.

HOMEWORK (40%) AND MIDTERM EXAM (20%)

- Homework problems focus on real problem solving, primarily using R markdown format, which includes your explanation and coding. You are encouraged to discuss with other students when doing your homework. Final project proposal is considered as one homework.
- Midterm exam consists of two parts:
    - Part (1) In-class exam: focus on statistical concepts such as understanding the exploratory data analyses (EDA), model diagnostics, and interpretation, without requirement of extensive computations.
    - Part (2) Take-home exam: focus on real problem solving. It has the same format as homework problems that require practical computations. You are allowed to use textbook, class notes and home works, but not allowed to discuss with other students.

CLASS PROJECT (20%) AND PRESENTATION (15%)

No final exam is planned in this class, but replaced by a class project and presentation.

- **Class Project:**
    - The class project can be from your research topic (preferred) or using the data related to your own expertise.
    - Three students form a project group to collaborate on the class project.

- The project proposal is due at the end of Week 9. To help you drafting the proposal and report, a template is available for you in advance, which includes the essential elements from statistical perspective.
- The project report will essentially follow the proposal.

- **Project Presentation:**
  - Project presentation focuses on clarity including question of interest, choice of appropriate statistical methods, and correct interpretation of results / conclusions. Each presentation is given 15 minutes.

TENTATIVE SCHEDULE

The course notes are primarily in R markdown (.Rmd) format which includes text and executable R codes. Your homework will also be in R markdown format.

Week		Topics	Reading
1	9/6	**Introduction and R Primer** • Installation of R and R Studio • R Markdown, R Package • R Language Basics and R Graphics	Install R and R Studio in advance
2	9/13	**Probability and Distributions** • Normal distribution • Binomial distribution • Bayes' theorem and screening tests	Ch. 3, 4
3	9/20	**Estimation** • Estimation of population mean and population mean difference • Estimation of population proportion and population proportion difference	Ch. 5, 6
4	9/27	**Hypothesis Testing** • Single population mean and two population means • Single population proportion and two population proportions • Sample size calculation	Ch. 7
5	10/4	**Analysis of Variance (I)** • Completely randomized design • Tukey's HSD test	Ch. 8
6	10/11	**Analysis of Variance (II)** • Randomized complete block design • Repeated measures design • Factorial design	Ch. 13
7	10/18	**Nonparametric Methods** • Wilcoxon Signed-rank test, Mann-Whitney test • Nonparametric ANOVA • Permutation test	Ch. 13
8	10/25	**Midterm exam** • Handout Part I: take-home **Chi-square tests and Fisher's exact test**	Ch. 12

		• Chi-square tests (goodness-of-fit, independence, homogeneity) • Fisher's exact test	
9	11/1	**Midterm Exam** • Part I: due prior to class • Part II: in-class part for statistical concepts **Linear Regression (I)** • Simple linear regression	Ch. 9
10	11/8	**Linear Regression (II)** • Multiple linear regression • Model selection	Ch. 10
11	11/15	**Logistic Regression and Project Proposal** • Logistic regression • Finalizing project proposal discussion and Q&A	Ch. 11
12	11/22	**Survival Analysis** • Kaplan-Meier method • Log-rank test • Cox proportional hazards model	Ch. 14
13	11/29	No Class (Thanksgiving Recess)	
14	12/6	**Poisson Distribution and Regression** • Poisson distribution and process • Poisson regression	
15	12/13	**Class Project Presentation** • 15 min each topic	Ch. 4, 11

POLICY

HOMEWORK DUE DATES

Assignments must be completed by the *specified due dates*. Homework assignment will be using R markdown file format and you should also use R markdown and its knitted pdf file to complete your homework including performing analysis and interpreting the results.

MIDTERM EXAM

For midterm in-class part, there will be no early, late, or makeup. The take-home part has 1 week to complete. No late or makeup.

For Midterm exam (in-class part), 2 pages of notes (letter size paper) are allowed. Calculators are allowed.

ACADEMIC INTEGRITY

The University has very strict rules concerning breaches of academic integrity. As such, any student caught cheating *will fail the course* and will likely be recommended for disciplinary action.

Rutgers University takes academic dishonesty very seriously. By enrolling in this course, you assume responsibility for familiarizing yourself with the Academic Integrity Policy and the possible penalties (including suspension and expulsion) for violating the policy. As per the policy, all suspected violations will be reported to the Office of Student Conduct.
Academic dishonesty includes (but is not limited to):

- cheating
- plagiarism
- aiding others in committing a violation or allowing others to use your work
- failure to cite sources correctly
- fabrication
- using another person's ideas or words without attribution
- re-using a previous assignment
- unauthorized collaboration
- sabotaging another student's work

If in doubt, please consult the instructor. Please review the Academic Integrity Policy at: *https://nbacademicintegrity.rutgers.edu*.

GETTING HELP

Many people have questions that do not get answered in class. Therefore, it is *very important* that you make the effort to get help immediately whenever you find yourself struggling. Should my office hours not be convenient, you are *strongly encouraged* to make an appointment with me or the TA at a mutually convenient time. Please do not wait until it is too late. You should understand that poor attendance and/or poor homework effort *will almost certainly lead to a poor grade*. You can also approach Rutgers Learning Centers for additional help. https://rlc.rutgers.edu.