

STATISTICS IN BIOMEDICAL SCIENCES

Fall 2023: CTSC 5103S / SGS 16:115:557 (3 Credits)

Fall Semester Begin-End: 9/5-12/13; Thanksgiving Recess: 11/23-26; Fall Exams Begin: 12/15/2023

OVERVIEW

INSTRUCTOR: Philip He, Ph.D.
EMAIL: PH443@GSBS.RUTGERS.EDU
I do not log into the canvas frequently and may not read the canvas messages. **Please email me directly.**

MEETING TIMES: Wednesdays, 4:00pm - 7:00pm

CLASS WEBSITE: canvas.rutgers.edu

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.

TEACHING ASSISTANT

NAME: Joonwon Lee
EMAIL: jl2815@stat.rutgers.edu

COURSE MATERIALS

TEXTBOOK: **Biostatistics: A Foundation for Analysis in the Health Sciences, 11th Edition** (2019) by Wayne W. Daniel, Chad L. Cross

The text book is essential for the course. **ISBN: 978-1-119-28237-2**
<https://www.amazon.com/Biostatistics-Foundation-Analysis-Probability-Statistics/dp/1119282373>

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**

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), multiple comparison, ANOVA, regression methods (normal, binary, and time-to-event endpoints), microarray methods (if time allows), and clinical statistical methods.

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.

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

Homework and Midterm are not group assignments. However, you can discuss with others for the general approach for your homework problems, without copying others' work.

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

There is no final exam in this class, but replaced by a class project and presentation. The class project can be from your research topic (preferred) or using the data related to your own expertise. At Week 9's class, you are given the opportunity to plan and discuss your class project proposal, and the proposal is due at the end of Week 9. However feel free to discuss your research question of interest and your proposed statistical handling, with other students and me throughout this course. You can also use your manuscript as a class project with a focus on the statistical techniques and interpretation of results. The evaluation of presentation focuses on clarity including question of interest, choice of appropriate statistical methods, and correct interpretation of results / conclusions. Each presentation is given 10 minutes.



TENTATIVE SCHEDULE

There will be 15 modules and 1 module for each week.

Week	Date	Topics	Reading
1	09/06/23	<u>Overview and Descriptive Statistics</u> a. Introduction to R b. Data visualization	Ch. 1, 2
2	09/13/23	<u>Probability and Probability Distributions</u> a. Probability b. Statistical methods in diagnostic medicine and screening test c. Binomial distribution d. Normal distribution	Ch. 3, 4
3	09/20/23	<u>Estimation</u> a. Sampling distribution b. Confidence interval 1) Population means 2) Population proportions	Ch. 5, 6
		<i>Homework 1 Due: Summary Statistics</i>	
4	09/27/23	<u>Hypothesis Testing</u> a. Type I error, Type II error b. Performing hypothesis test 1) on population means 2) on population proportions 3) on location / median 4) goodness-of-fit 5) based on non-parametric methods	Ch. 7, 13
		<u>Analysis of Variance</u>	
5	10/04/23	a. Comparisons between and among means b. Kruskal Wallis test c. Multiple comparisons	Ch. 8
		<u>Analysis of Categorical Data</u>	
6	10/11/23	a. Logistic regression b. Poisson regression a. Chi-square tests (goodness-of-fit, independence, homogeneity, Fisher's exact test) b. Relative risk, Odds ratio, Mantel-Haenszel statistic	Ch. 11, 12
		<i>Homework 2 Due: Hypothesis testing & ANOVA</i>	
		<u>Midterm Exam</u>	
7	10/18/23	<u>Vital Statistics</u> a. Death rates and ratios b. Measures of Fertility, Morbidity	Ch. 15
		<u>Regression</u>	
8	10/25/23	a. Correlation b. Simple linear regression c. Multiple linear regression	Ch. 9, 10



9	11/01/23	<u>Project Proposal</u> a. Class discussion of your project proposal	N/A
10	11/08/23	<u>Survival Analysis</u> a. Kaplan-Meier method b. Log-rank test c. Cox proportional hazard model <i>Homework 3 Due: Categorical Data & Linear Regression</i>	Ch. 14
11	11/15/23	<u>Sample Size Estimation</u> a. Based on Precision of Estimation (proportions) b. Based on Hypothesis Testing <ul style="list-style-type: none"> • Single arm design • Parallel arm design • Various types of endpoints 	Multiple chapters and handouts
12	11/22/23	<u>Statistical Methods in Microarray Data</u> a. Microarray data analysis	Handouts
13	11/29/23	<u>Statistical Practice in Clinical Research and Medical Literature</u> a. Study design in drug development b. Statistical considerations <ol style="list-style-type: none"> 1) Sample size determination 2) Endpoint definitions and data collection 3) Analyses (safety, efficacy) c. Interim decisions planning d. Potential bias, over-interpretation, and misuse <i>Homework 4 Due: Survival Analysis</i>	Handouts
14	12/06/23	Presentation Session 1 (10 min each topic)	
15	12/13/23	Presentation Session 2 (10 min each topic)	



POLICY

HOMEWORK DUE DATES

Assignments must be completed by the *specified due dates*. In each homework assignment, you are expected to use R to perform analysis and interpret the results. There will be no acceptance of late homework assignment submission.

EXAM MAKE-UP

In class exam. There will be no early, late, or makeup midterm exam.

NOTES

For Midterm Exam, 2 pages of notes (letter size paper) are allowed. Calculators are allowed and are most likely needed at exam.



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>.