Biostatistics (BIO5312) Syllabus, Fall 2017

Course Time and Location
Tuesday 5:30-8pm, BioLife 342

Instructor
Stephanie J. Spielman, PhD
stephanie.spielman@temple.edu
SERC 643

Office Hours
Friday 1-3 pm, Ground floor of SERC *(subject to change)*

Website
Course materials will be hosted on https://sjspielman.org/bio5312_fall2017. We will be using Canvas Instructure (rather than Blackboard) to host grades.

Objective
This course will introduce you to fundamental concepts in biostatistics, data science in biology, and how to apply these concepts using the R statistical programming language. We will additionally cover data visualization and presentation, again using R. In general, this course will emphasize applied biological data analysis over statistical theory. Students are expected to independently make use of (free!) online resources to complete assignments.

Textbook
We will not be using a formal textbook for any assignments in this course. However, I strongly recommend one of the following textbooks as background reading material. Recommended readings from each book, when applicable, will be noted on the course website.
1. **Fundamentals of Biostatistics, B. Rosner**
    a. Recommended for students wishing a more math-oriented reading
2. **Analysis of Biological Data, M. C. Whitlock and D. Schluter**
    a. Recommended for students wishing a more concept-oriented background reading

Requirements

Students are expected to bring a personal laptop to each class. Students must additionally install R (https://www.r-project.org/) and RStudio (https://www.rstudio.com/products/rstudio/download2/), in that order. For RStudio, please download the free open source RStudio Desktop distribution.

Grading and Assignment Policy

Grades will be determined based on weekly homeworks (75% of total) and a final project (25% of total). You are encouraged to work together on homework assignments as needed, but your submitted assignment must be your own work. However, you must complete your final project independently.

Grade boundaries are defined as follows:

A	93-100	C	73-76.5
A-	90-92.5	C-	70-72.5
B+	87-89.5	D+	67-69.5
B	83-86.5	D	63-66.5
B-	80-82.5	D-	60-62.5
C+	77-79.5	F	<60

Note that final grades will generally be rounded up. For example, a final grade of 89.6 will correspond to an A-, but a final grade of 89.3 corresponds to a B+. There will be no exceptions.

Assignments are due by the start of the class (5 pm) and should be turned in either via Blackboard or email. All assignments (unless explicitly stated otherwise) should be written in RMarkdown and compiled to either HTML or PDF for submission.

Late assignments will not be accepted without prior permission from the instructor. If you anticipate that you will be unable to complete an assignment by the due date, please contact the instructor at least 24 hours before the due date.

Email Policy

Emails to the instructor should be restricted to organizational issues, for example requests for an appointment or questions about course organization or logistics. *Technical issues will not be answered via email and must be addressed in person.* Technical issues are defined as questions concerning how to approach problems on assignments, whether a solution is correct, how to use R, or grading issues/questions. Questions about the meaning or appropriate interpretation of a given assignment question are, however, appropriate for email.
Please note that I will not respond to emails between 7 pm - 9 am.

Schedule

Date	Topic
8/29/17	Introduction to Biostatistics ● Descriptive and summary statistics ● Data visualization ● Introduction to base R
9/5/17	Introduction to Data Analysis in R ● The "Tidyverse": Data representation, management, and visualization ● Rmarkdown
9/12/17	Probability and Bayes Theorem
9/19/17	Continuous and Discrete Distributions ● Central Limit Theorem ● Standard Normal distribution and z-scores
9/26/17	Hypothesis Testing I ● Parametric tests for continuous data
10/3/17	Hypothesis Testing II ● Parametric tests for discrete data ● Contingency tables
10/10/17	Hypothesis testing III ● Nonparametric tests
10/17/17	Multiple testing and False Discovery Rate
10/24/17	Correlation and Linear Models I

10/31/17	Linear Models II
11/7/17	Binary classification ● Logistic Regression ● ROC ● Machine learning
11/14/17	Clustering ● K-means clustering ● Principal Components Analysis (PCA) ● Hierarchical clustering
11/21/17	Statistics for Bioinformatics I: Genomic Analysis ● GWAS and RNAseq analysis
11/28/17	Statistics for Bioinformatics II: Evolutionary Analysis ● Markov models of evolution and phylogenetic approaches
12/5/17	Advanced R ● Tidyverse II ● Writing custom packages

Please note that this schedule is subject to change based on student progress. Specifically, the schedule may be amended to prioritize mastery of foundational concepts in Biostatistics over more advanced/special topics.

Disclosures

Civility & Temple's Code of Conduct (CoC)
Violations of the CoC include, but are not limited to: academic dishonesty and impropriety, including plagiarism and academic cheating; interfering or attempting to interfere with or disrupting the conduct of classes or any other normal or regular activities of the University (see: http://policies.temple.edu/getdoc.asp?policy_no=03.70.12)

Disability Disclosure
Any student who has need for accommodation based on the impact of a disability should contact the instructor privately to discuss the specific situation as soon as possible. In addition, Disability Resources should be contacted at 215-204-1280.

Student/Faculty Academic Rights and Responsibilities

The University's policy can be accessed through the following link:

http://policies.temple.edu/getdoc.asp?policy_no=03.70.02

Honor Code

The Academic Honor Code can be accessed through the following link:

http://policies.temple.edu/PDF/398.pdf .