QBS 120: Foundations of Biostatistics I
Statistical Theory for the Quantitative Biomedical Sciences

Fall 2020

Instructor: Rob Frost, PhD, Assistant Professor of Biomedical Data Science
Office: Rubin 704, Email: rob.frost@dartmouth.edu
Office hours: Zoom by appointment

TAs: Courtney Schiebout
Email: Courtney.T.Schiebout.GR@dartmouth.edu
Office hours: see below

Mirjana Stevanovic
Email: Mirjana.Stevanovic.GR@dartmouth.edu
Office hours: see below

Overview

This is a time-intensive graduate-level course in mathematical statistics designed to teach the fundamental knowledge of statistical theory required to read and, with further study, contribute to the statistical methodology literature. An in depth overview of statistical estimation and hypothesis testing will be provided, including the method of least squares, maximum likelihood methods, asymptotic methods, and correction for multiple comparisons. The basic elements of statistical design and sample size calculations will be introduced. Resampling strategies will be discussed in the context of the bootstrap, as well as simulation as a tool for statistical research. The emphasis will be on theory used in modern applications in biomedical sciences, including genomics, epidemiology, and clinical and health services research. The statistical program language R will be leveraged for computational examples, problem sets and exams.

Time and place

Lectures will be on Tuesday/Thursday 8:30-10am (all times EST) in the Kellogg auditorium. Lectures will also be streamed live via Zoom and recorded by the classroom video equipment. Zoom info:

- https://dartmouth.zoom.us/j/92669128421?pwd=T0drcjZPWGlkdXIwaEZONFRuSHhxQT09
- Meeting ID: 926 6912 8421, Passcode: 198413

Office hours:

- Courtney
  - Thursdays, 11:30AM-12:30PM
  - https://dartmouth.zoom.us/j/92520751088?pwd=cEtwYnU4Ylcwc2dYTDNPZHEwc2RoUT09
  - Passcode: 936799
- Mina
  - Mondays, 2-3PM
  - https://dartmouth.zoom.us/j/95812871304?pwd=Z1FvTi8rWC8wZ3hXKzZ1SXU3ZkpBZz09
  - Passcode: 646076
- Courtney and/or Mina
  - Tuesdays, 10AM-11AM
  - https://dartmouth.zoom.us/j/92454891135?pwd=QURmTFRGWjB5Njc1Um5RZ0VCSlZmQT09
  - Passcode: 010178

Prerequisites

QBS 120 is a fast-paced, calculus-based graduate mathematical statistics course with a strong theoretical component. It is assumed that students are comfortable with multivariate calculus, mathematical proofs, linear algebra and R programming. A strong internal motivation to learn the material and complete challenging assignments is essential to success in this class. Students should expect to spend 10-15+ hours per week outside of class. **Students are strongly encouraged to review the content and level of theory in the class textbook (Rice, see below) prior to registering for the class.** Versions of prior problem sets (and solutions) are also available on request to help students assess the class workload, theoretical component and assumed mathematical and computational background.

Summer Review

For students lacking R experience or whose math skills are a bit rusty, I recommend they review the following topics over the summer (note: some of this will be covered during the QBS boot camp):

- **R statistical programming language.** Students without prior R experience should work through a basic R tutorial/overview (e.g., "An Introduction to R" in the R help documentation or a 3rd party book or online tutorial). Students should also learn a framework such as R Markdown or Sweave that combines R logic with structured text. Since handwritten work is not allowed for problem sets or exams, skill in one of these frameworks will be a huge time saver. Learning to navigate an R IDE such as R Studio is also recommended.
- **Math (calculus, linear algebra, "pre-calculus" topics).** The "Mathematics Survival Kit" (https://www.maplesoft.com/products/MSK/MSK_TableOfContents.pdf) is a good resource and students should get free access during the QBS boot camp. Many other free online resources exist including comprehensive sites (e.g., Khan Academy, https://www.khanacademy.org/math/calculus-1) and more concise reviews (e.g., http://pages.stat.wisc.edu/~ifischer/calculus.pdf, https://www.cs.cmu.edu/~zkolter/course/linalg/index.html).

Requirements and Grading

QBS 120 is graded on the HP/P/LP/NC scale. Grades will be determined using the following approximate weighting: problem sets (50%), take home final (50%). Problem sets must be completed electronically (i.e., no hand written work) and one or two random problems will be graded. Students are expected to submit corrected assignments after official solutions have been posted (these are graded for completion).

Textbook

Most lectures will be based on *Rice, John A. (2007) Mathematical Statistics and Data Analysis (3rd edition). Duxbury Advanced Series.* Specific lectures may include readings from other sources. Students are expected to complete the assigned reading prior to the relevant lecture. A more applied discussion of these topics can be found in *Rosner, Bernard. (2010) Fundamentals of Biostatistics (7th edition)*. A more theoretical discussion of these topics can be found in *Casella and Berger (2002). Statistical Inference (2nd edition)*.

Problem Sessions

There will be weekly Zoom-based, TA-lead problem sessions on Fridays from noon-1pm:

- https://dartmouth.zoom.us/j/99458033670?pwd=YjVnbk5vSE1Mbmh4K0dIcDFnbGd0QT09
- Passcode: 947004

Schedule

Date	Topic	Readings	Out	Due
Week 1, 9/15	Class overview, review of probability	Rice, Chapter 1		
Week 1, 9/17	Random variables	Rice, Chapter 2	PS1	

Week 2, 9/22	Joint distributions	Rice, Chapter 3		PS1
Week 2, 9/24	Expected values, variance, standard deviation	Rice, Chapter 4.1-4.2	PS2	
Week 3, 9/29	Covariance, correlation, conditional expectation, moments	Rice, Chapter 4.3-4.6		PS2, PS1 corrections
Week 3, 10/1	Limit theorems	Rice, Chapter 5	PS3	
Week 4, 10/6	Distributions derived from normal	Rice, Chapter 6		PS3, PS2 corrections
Week 4, 10/8	Finite population sampling	Rice, Chapter 7.1-7.3	PS4	
Week 5, 10/13	Parameter estimation	Rice, Chapter 8.1-8.5		PS4, PS3 corrections
Week 5, 10/15	Parameter estimation, continued	Rice, Chapter 8.5.1-8.8 (not 8.6)	PS5	
Week 6, 10/20	Hypothesis testing	Rice, Chapter 9.1-9.3		PS5, PS4 corrections
Week 6, 10/22	Hypothesis testing, continued	Rice, Chapter 9.4-9.9	PS6	
Week 7, 10/27	Summarizing data	Rice, Chapter 10		PS6, PS5 corrections
Week 7, 10/29	Two sample tests	Rice, Chapter 11.1-11.3	PS7	
Week 8, 11/3	Sample size and power	Rice 11.2.2, https://stats.idre.ucla.edu/other/mult-pkg/seminars/intro-power/		PS7, PS6 corrections
Week 8, 11/5	ANOVA	Rice, Chapter 12	PS8	
Week 9, 11/10	Multiple testing	Rice 11.4.8, 12.2.2, https://doi.org/10.1002/sim.6082		PS8, PS7 corrections
Week 9, 11/12	Categorical data analysis	Rice, Chapter 13	PS9	
Week 10, 11/17	*No class*		**Final**	PS9, PS8 corrections
Week 11, 11/23				**Final**

Dartmouth Title IX Notice

At Dartmouth, we value integrity, responsibility, and respect for the rights and interests of others, all central to our Principles of Community. We are dedicated to establishing and maintaining a safe and inclusive campus where all have equal access to the educational and employment opportunities Dartmouth offers. We strive to promote an environment of sexual respect, safety, and well-being. In its policies and standards, Dartmouth demonstrates unequivocally that sexual assault, gender-based harassment, domestic violence, dating violence, and stalking are not tolerated in our community.

The Sexual Respect Website (sexual-respect.dartmouth.edu) at Dartmouth provides a wealth of information on your right with regard to sexual respect and resources that are available to all in our

community. Please note that, as a faculty member, I am obligated to share disclosures regarding conduct under Title IX with Dartmouth's Title IX Coordinator.

Should you have any questions, please feel free to contact Dartmouth's Title IX Coordinator (Kristi.Clemens@Dartmouth.edu) or the Deputy Title IX Coordinator for Geisel (Leslie.Henderson@Dartmouth.edu) or for Guarini (Gary.Hutchins@Dartmouth.edu).