

Department of Biology Course Outline

BIOL5081 Biostatistics Fall, 2016

Course Description

Official Calendar Course Description

This course examines common statistical methods used in biology. Data science and statistical workflows are developed. Descriptive statistics, generalized linear models, regression, nonparametric tests, bootstrapping, randomization tests, tree-based analysis, multivariate statistics, and time-series analysis are considered. The R programming and software environment will be used for data analysis. This course is being offered Fall 2016.

Prerequisites

BIOL 2060, previously numbered BIOL 3090 or an undergraduate course in Statistics. Students who have not taken a statistics course within the last three years are required to audit BIOL 2060 lectures.

Course Instructors and Contact Information

Dr. C. Lortie and Dr. S. Sharma

emails: Lortie: lortie@yorku.ca

Sharma: sharma11@yorku.ca

PH: Lortie: 416.736.2100 ext20558

Sharma: 416.736.2100 ext. 33761

Office hours: Provided as needed & upon request; CJL & SS will also frequently be available after class.

Schedule

Time and location

Thursdays at 11:30am for the first 6 weeks (Sept 8th to Oct 13th, 2016) in TEL 2032*.

Thursdays at 7pm for the latter 6 weeks (Oct 20th to Dec 1st, 2016) in TEL 2118

* Sept 8 class is booked into TEL 2027 (right next door to the other lab).

Module 1. Data science & Fundamental Exploratory Data Analysis (EDA) in r

Christopher Lortie.

<https://cjlortie.github.io/r.biostats/>

Lesson 1. Data science (DS) & Hypothesis Testing (HT). Sept 8th.

want data. have data. will collect data.

assumption: in this course, you are here to work with data.

data literacy IS data science.

Lesson 2. Workflows & Data Wrangling (WDW). Sept 15th.

workflows

reproduce.

openly.

data wrangling
more than half the battle.

Lesson 3. Visualization in r (VR). Sept 22nd.
basic plots.
lattice.
ggplot2.
you need to see data. see the trends. explore them using visuals.

Lesson 4. Exploratory data analysis (EDA) in r. Sept 29th.
fundamental descriptive stats.
GLM. GLMM. Post hoc tests.

Lesson 5. Wrangle, visualize, and analyze (WVA). Oct 6th.
A graduate-level dataset.
Apply your skills from scratch.
A single three-hour block.
As advanced as you want to be.
Fundamental principles demonstrated.
At least two fundamental statistical tests demonstrated.

Lesson 6. Competency test (CT). Oct 13th.
deliberate practice.
tested with feedback.
a virtual repeat of last week but evaluated

Module 2. Advanced environmental data analyses in r.

Sapna Sharma

Lesson 1. October 20th.
Linear-based models and linear-based models lab in R.

October 27th – No class – Reading Week

Lesson 2. November 3rd.
Tree-based models and tree-based models lab in R.
Linear-based model lab due.

Lesson 3. November 10th.
Multivariate analysis I – Indirect Gradient Analysis and Cluster Analysis. Lab in R.
Tree-based model lab due.

Lesson 4. November 17th.
Multivariate analysis II – Direct Gradient Analysis. Lab in R.
Multivariate analysis I lab due.

Lesson 5. November 24th.
Time-series analysis. Lab in R.
Multivariate analysis II lab due.

Lesson 6. December 1st.
Data Analysis workshop.
Time series analysis lab due.

Evaluation

Evaluation Components of Final Grade and related information.

Final course grades may be adjusted to conform to Program or Faculty grades distribution profiles.

Test	25%
Oral presentation/Labs	25%
Written reports	50%

NOTE: An exam or term test worth more than 20% of the final grade may not be given during the final two weeks of classes.

Important Dates

Oct 13th **Test** to demonstrate workflows, data wrangling, and fundamental data wrangling. **25%**

Oct 20th A **written report** to summarize Module A of course due one week after test to more fully analyze project. **25%**

Nov 3rd: Linear models lab due. **5%**

Nov 10th: Tree models lab due. **5%**

Nov 17th: Multivariate analysis lab I due. **5%**

Nov 24th: Multivariate analysis lab II due. **5%**

Dec 1st: Time-series analysis lab due. **5%**.

Dec. 8th: 3 page paper due by email to sharma11@yorku.ca – Sections to include are: Research Question, Methods – Data Acquisition and Data Analysis, and Results & Interpretation. Centred around answering a research question with your own dataset (or one from your lab, or one that is approximately equal to the type of thesis research you are conducting; there are many freely available datasets online, some designed to explicitly fit a specific statistical test). You will compare and contrast 2 statistical approaches to answer your research question. **25%**

Resources

Module 1. Software Carpentry learning materials and open r materials will be provided as needed. All resources from this component of the course listed here: <https://cjlortie.github.io/r.biostats/>

Module 2. There is no required textbook. Suggested readings and resources will be provided. You need to have a Webacadlabs account to use statistical software and resources will be posted on moodle.

Learning Outcomes

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

Module 1. Data science & Fundamental Exploratory Data Analysis in r

1. Complete a data management plan for a graduate-level dataset.
2. Develop a reproducible data and statistical workflow.
3. Design and complete intermediate-level data visualizations appropriate for a graduate-level simple dataset.
4. Complete fundamental exploratory data analysis on any dataset.
5. Appreciate the strengths and limitations of open science, data science, and evidence-based collaboration models.

Module 2. Advanced biostatistics.

1. Conduct basic and advanced statistical analyses in R.
2. Interpretation of data analyses.
3. Analyze data from your thesis or lab group.

Course Content

All details and code for first module of course will be available on GitHub for download.
All details and code for second module of course will be available on Moodle for download.

Experiential Education and E-Learning

GitHub, R Studio, and R will be used to share code, wrangle data, and do statistics.
Students will be provided access to open, datasets from public data repositories.

Other Information

EXPECTATIONS

Attendance is mandatory because the lectures will provide an opportunity for the students not only to listen to summary lectures of the by the professor of statistics but to also engage in the analyses. In the lectures, we will work together to actively handle data and do statistics. All information presented in class is testable.

Course Policies

Alternative dates for assignments/evaluations are not available in this course. If documentation is provided for valid absences on test dates, accommodation will be granted in mutual discussion with the professors.

To promote fairness and student responsibility, all in class exercises are due on the dates specified herein. A 20% penalty will be applied for the first day the exercise is late and 5% every day thereafter. Students who anticipate being unable to submit the exercises on the due date are encouraged to submit early.

Grades on exercises and exams are not negotiable. Every reasonable action is made to ensure advance reminders are provided and instruction. Thus, the course director should only be contacted if there is calculation or clerical error present.

Students are not allowed to record lectures or lab tutorials using their own devices.

University Policies

Academic Honesty and Integrity

York students are required to maintain the highest standards of academic honesty and they are subject to the Senate Policy on Academic Honesty (<http://secretariat-policies.info.yorku.ca/policies/academic-honesty-senate-policy-on/>). The Policy affirms the responsibility of faculty members to foster acceptable standards of academic conduct and of the student to abide by such standards.

There is also an academic integrity website with comprehensive information about academic honesty and how to find resources at York to help improve students' research and writing skills, and cope with University life. Students are expected to review the materials on the Academic Integrity website at - http://www.yorku.ca/spark/academic_integrity/index.html

Access/Disability

York University is committed to principles of respect, inclusion and equality of all persons with disabilities across campus. The University provides services for students with disabilities (including physical, medical, learning and psychiatric disabilities) needing accommodation related to teaching and

evaluation methods/materials. These services are made available to students in all Faculties and programs at York University.

Students in need of these services are asked to register with disability services as early as possible to ensure that appropriate academic accommodation can be provided with advance notice. You are encouraged to schedule a time early in the term to meet with each professor to discuss your accommodation needs. Please note that registering with disabilities services and discussing your needs with your professors is necessary to avoid any impediment to receiving the necessary academic accommodations to meet your needs.

Additional information is available at the following websites:

Counselling & Disability Services - <http://cds.info.yorku.ca/>

Counselling & Disability Services at Glendon –

<http://www.glendon.yorku.ca/counselling/personal-counselling/what-is-counselling/>

York Accessibility Hub - <http://accessibilityhub.info.yorku.ca/>

Ethics Review Process

York students are subject to the York University *Policy for the Ethics Review Process for Research Involving Human Participants*. In particular, students proposing to undertake research involving human participants (e.g., interviewing the director of a company or government agency, having students complete a questionnaire, etc.) are required to submit an *Application for Ethical Approval of Research Involving Human Participants* at least one month before you plan to begin the research. If you are in doubt as to whether this requirement applies to you, contact your Course Director immediately.

Religious Observance Accommodation

York University is committed to respecting the religious beliefs and practices of all members of the community, and making accommodations for observances of special significance to adherents. Should any of the dates specified in this syllabus for an in-class test or examination pose such a conflict for you, contact the Course Director within the first three weeks of class. Similarly, should an assignment to be completed in a lab, practicum placement, workshop, etc., scheduled later in the term pose such a conflict, contact the Course director immediately. Please note that to arrange an alternative date or time for an examination scheduled in the formal examination periods (December and April/May), students must complete an Examination Accommodation Form, which can be obtained from Student Client Services, Student Services Centre or online at http://www.registrar.yorku.ca/pdf/exam_accommodation.pdf

Student Conduct in Academic Situations

Students and instructors are expected to maintain a professional relationship characterized by courtesy and mutual respect. Moreover, it is the responsibility of the instructor to maintain an appropriate academic atmosphere in the classroom and other academic settings, and the responsibility of the student to cooperate in that endeavour. Further, the instructor is the best person to decide, in the first instance, whether such an atmosphere is present in the class. The policy and procedures governing disruptive and/or harassing behaviour by students in academic situations is available at - <http://secretariat-policies.info.yorku.ca/policies/disruptive-and-or-harassing-behaviour-in-academic-situations-senate-policy/>