Lakehead University

Biostatistics (Biology 3112, 5171), Winter 2020

Instructors:
Lecturer: **Dr. Michael Rennie**
Office: CB 4050
Phone: 807-346-8760
e-mail: mrennie@lakeheadu.ca
Office hour: Monday, 1:00-2:00 pm

TA (Thunder Bay): **Scott Bergson**
Office: CB 4030
Phone: 807-343-8043
e-mail: sbergson@lakeheadu.ca
Office hours: Thursday 10:30-11:30, CB 4030

Text (recommended):
Experimental design and data analysis for biologists.
G. P. Quinn and M .J. Keough, 2002. Cambridge University Press. ISBN: 0 521 00976 6

Class Schedule:

LECTURES: Thunder Bay: Monday and Wednesday, 11:30 am to 1:00 pm, AT 2001

TUTORIALS: Tuesdays 8:30-10:30 am, AT 3003

Lecture Schedule (tentative, will adjust topics as required):

Lecture (L) or Tutorial (T) #	Date	Topic	Recommended readings
L1	Jan 6	I'm a biologist/ecologist/ environmental scientist: what am I doing in a statistics class? *Introduction to R*	Chapter 1; Chapter 2 up to section 2.3 and 2.4.2; Chapter 3 up to section 3.7; Chapter 4, Chapter 19.
T-Intro	7	*Tutorial- getting comfortable with R*[1]	
L2	8	Correlation, linear regression, model II regression	Chapter 3 to section 5.3.15; section 5.4, 5.7.
L3	13	Multiple regression (Assignment 1 posted)	Chapter 6 to section 6.1.5
T1	14	*Correlation, regression*	
L4	15	Single factor ANOVA, unplanned contrasts	Chapter 8 to section 8.1.5; section 8.3, 8.4

Lecture (L) or Tutorial (T) #	Date	Topic	Recommended readings
L5	20	Type I and II error rates; planned contrasts (Assignment 1 due)	Section 8.6, Chapter 3, especially section 3.2; Box 8.4 has a worked example
T2	21	*Single factor ANOVA*	
L6	22	Random effects	Section 8.2.1
L7	27	Experimental design (Assignment 2 posted)	Chapter 7 up to and including section 7.2
T3	28	*Estimating variance components*	
L8	29	Nested ANOVA	Chapter 9 to section 9.1.9
L9	Feb 3	Nested ANOVA, Randomized block design Practice midterm posted (Assignment 2 due)	Chapter 10 to section 10.10, 10.14
T4	4	*Nested ANOVA*	***grad students meet with Dr. Rennie about final projects**
L10	5	Factorial ANOVA; Mixed effects models (the old way)	Section 9.2, up to 9.26; 9.28; 9.2.11; 9.4, 9.5
L11	10	Unbalanced designs in ANOVA; appropriate Sums of Squares Review practice midterm (Assignment 3 posted)	Pages 242-244, section "Unequal sample sizes" ***poll for topics last 2 weeks of class**
T5	11	*Blocked design*	
MIDTERM	12	**MIDTERM**	
	17-21	**READING BREAK**	
L12	24	Statistical power (Assignment 3 due)	Sections 5.6, 8.9, 9.2.13, 10.10 ***grad students submit 1-2 page proposal**
T6	25	*Factorial ANOVA, working with "real" data; midterm questions*	
L13	26	Multiple testing	Section 3.4
L14	Mar 2 (25% of mark needed)	Test for heterogeneity of slopes, Analysis of Covariance, comparisons of adjusted means (Assignment 4 posted)	Chapter 12, to section 12.4; section 12.5, 12.6, 12.8

Lecture (L) or Tutorial (T) #	Date	Topic	Recommended readings
T7	3	*Power analyses, Multiple comparisons*	
L15	4	It's all just general linear modelling, man (this is where we blow your mind); dummy variables	Section 6.1.14
L16	9	Tests of frequencies (Assignment 4 due)	Chapter 14, to section 14.2.2
T8	10	*Comparing slopes, ANCOVA*	
L17	11	Non-parametric tests	Section 3.3.3, section 5.1.2, Section 8.5.2, 10.5
L18	16	Guest lecture, Dr. Cody Dey- data visualization	Section 3.3.2; readings to be assigned
T9	17	*Frequency tests; Traditional non-parametric tests*	
L19	18	Guest lecture, Dr. Cody Dey- data visualization II	
L20	23	General***ized*** linear models* (Assignment 5 posted)	
T10	24	*Data visualisation*	
L21	25	Randomization- permutation tests	
L22	30	Randomization- bootstrapping tests	
T11	31	*Randomization*	
	Apr 1	One of generalized linear models, mixed effects modelling or model selection criteria and **Review** (Assignment 5 due)	Chapter 13 to up to and including section 13.3; assigned reading
	TBD	**Final Exam (Location TBD)**	**Grad students final paper due**

[1]The tutorial this week will be, in part, self-directed; students are strongly encouraged to load R and Rstudio on their personal computers so they can work on assignments, etc. at home (Please make time to complete this task during the first week of classes). Students will go through the introductory R code presented in lecture on Jan 11th, on the machines in AT 3001 and at home using their personal

computers.
*topics during the final 3 lectures and in the last tutorial may vary from this depending on student interests; can be customized if there are specific analyses that the class would like to address.

Assignments: There will be five assignments that are to be completed outside of classes. These will all consist of independent analyses of data sets and a written report for grading. The four assignments in which you do best will be counted in your final grade.

Tutorials: Each week there will be a two-hour tutorial in which you will get practice solving statistical problems using a computer and get comfortable using R. You are not required to submit anything for grading. These are also great opportunities to pick the brain of your TA, instructor, or peers on assignments.

Policy on late assignments or missed work: Failing to submit academic work on time is a serious matter. Students should arrange their schedules so that academic work is a top priority during the school year. *Because only four of the five graded assignments will count towards your final mark (see below), NO medical reasons for failing to submit an assignment on time will be accepted except under the most serious circumstances. A grade of 0 will be assigned to any late or missed assignments.* There is only one term test and only the most urgent medical matter will be accepted as a reason for missing the term test. The only acceptable document for medical emergencies is the 'Lakehead University Medical Certificate" and can be found here, along with instructions and requirements of such exemptions: https://www.lakeheadu.ca/current-students/examination/medical-notes.

Email: In order to receive important course communications, **it is absolutely necessary** that you monitor notices on the course website at least twice a week.

Grading (undergraduates):

1. Best four out of five assignments, 10 points each [40%]

2. Term Test, February 17 [20%]

3. Final Exam [35%]

4. Student engagement (in class, in tutorials, participation on discussion forums, etc) [5%]

(Calculators- NOT phones with calculators, but old-school calculators with no additional functionality- are allowed for term test and final exam, but no other materials)

***Grading (graduate students):** *Graduate students will not write exams.* Assignments will be completed by graduate students, based on the same policy described above. In place of exams, graduate students will meet with the instructor to discuss an appropriate analysis for a dataset of their choosing, and will submit a report at the end of term describing the statistical approach. **A 1-2 page proposal** outlining the dataset and the planned analysis will be submitted around the time of the midterm, **worth 10% of the final grade**. The **final report will be worth the remaining 40%** of the final grade, and will loosely follow a typical scientific report (abstract, introduction, methods, results, discussion), but a heavy emphasis will be placed on the methodological choice of analysis selected in relation to the data

set and experimental design, reporting of results and interpretation of the analysis. Appendices should be included to provide sufficient evidence that assumptions have been tested and have informed the analytical approach presented. Students are encouraged to use their own data for this assignment; if this is not possible, contact the instructor for alternatives. **The remaining 10%** of the course grade for grad students will be for **class participation, and they are expected to have a greater level of engagement.**

Course web page: There will be a course web page through myCourseLink. Stay tuned as I figure out how it works, but it will be a place to find lectures in .pdf format, R code, assignments, and discussion boards.

Discussion board and e-mail policy: Separate forums will be set up for the course in general, R-related questions and possibly additional forums for particular topic areas. Any questions regarding course organization, e.g., assignments, due dates etc. as well as questions regarding course content, e.g., statistical questions, should be posted to the appropriate forum. Students are invited to help answer questions posted to the discussion board as far as possible, particularly with regards to R-help (the best way to learn something yourself is to show someone else how to do it). Entries will be monitored by the course staff and annotated as necessary within two days from posting, and major issues will be addressed in class or during the tutorials. Please keep your questions and answers short and precise and be polite! Using the discussion board gives all students access to the same information. Therefore, the instructor will not answer individual emails about course organization or content, e.g., statistical questions. Students should only send emails to the instructor regarding personal issues that cannot be posted on the discussion board. Emails will usually be answered within two days (three days over weekend).

For help with R:
1. Begin by referring to the documents you have been provided with- the "getting started with R" lecture notes; "An introduction to R", by Venables, Smith and the R Development Core Team, available for download on the course website; other resources on the CRAN contributed documents (http://cran.r-project.org/other-docs.html); all of this stuff is free.
2. Search google with [R] in your search term; e.g. "[R] t-test". Sift through the search results till you find something helpful, most often on the first page or two.
3. Try "?topic" where "topic" is the function or issue you are having with, or help.search(topic) if it's not a function, but something else.
4. Post a question on the R-discussion forum on the course website. Wait for a student to post an answer (may be annotated by course staff within 2 days).
5. If still not answered, ask your TA in the tutorial session.
6. If STILL necessary, ask the instructor after class or during office hours.