Lakehead University

Biostatistics (Biology 3112, 5171), Winter 2022

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

TAs: **Alex Ross**
Office: CB 4030/ FB 1018
Phone: 807-343-8043 / 807-346-7979
e-mail: aross10@lakeheadu.ca
Office hours: Tuesdays, 11:30-12:30

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: Prior to Jan 23: Monday and Wednesday, 11:30 am to 1:00 pm via ZOOM (see link on course website: other tools > zoom > click on lecture
After Jan 23: Monday and Wednesday, 11:30 am to 1:00 pm, AT3006

TUTORIALS: Prior to Jan 23: Thursdays 12:30 pm to 2:30 pm via ZOOM (see above)
After Jan 23: Thursdays 12:30 pm to 2:30 pm, AT 3001 (via ZOOM for Orillia students)

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

Lecture (L) or Tutorial (T) #	**Date**	**Topic**	**Recommended readings**
L1	Jan 10	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.
L2	12	Correlation, linear regression	Chapter 3 to section 5.3.15; section 5.4, 5.7.
T-Intro	13	*Tutorial- getting comfortable with R*[1]	
L3	17	Linear regression (continued), model II regression (Assignment 1 posted)	Chapter 3 to section 5.3.15; section 5.4, 5.7.

Lecture (L) or Tutorial (T) #	Date	Topic	Recommended readings
L4	19	Multiple regression	Chapter 6 to section 6.1.5
T1	20	*Correlation, regression*	
L5	24	Single factor ANOVA, unplanned contrasts;	Chapter 8 to section 8.1.5; section 8.3, 8.4
L6	26	Type I and II error rates; planned contrasts; Random effects (Assignment 1 due)	Section 8.6, Chapter 3, especially section 3.2; Box 8.4 has a worked example; Section 8.2.1
T2	27	*Single factor ANOVA*	
L7	31	Experimental design (Assignment 2 posted)	Chapter 7 up to and including section 7.2
L8	Feb 2	Nested ANOVA	Chapter 9 to section 9.1.9
T3	3	*Estimating variance components*	
L9	7	Nested ANOVA, Randomized block design **Practice midterm posted**	Chapter 10 to section 10.10, 10.14
L10	9	Factorial ANOVA; Mixed effects models (the old way) (Assignment 3 posted) (Assignment 2 due)	Section 9.2, up to 9.26; 9.28; 9.2.11; 9.4, 9.5
T4	10	*Nested ANOVA*	***grad students meet with Dr. Rennie about final projects**
L11	14	Unbalanced designs in ANOVA; appropriate Sums of Squares Review practice midterm	Pages 242-244, section "Unequal sample sizes"
MIDTERM	**16**	**MIDTERM**	***grad students submit 1-2 page proposal**
T5	17	*Blocked design*	
	21-25	**READING BREAK**	
L12	28	Statistical power (Assignment 3 due)	Sections 5.6, 8.9, 9.2.13, 10.10

Lecture (L) or Tutorial (T) #	**Date**	**Topic**	**Recommended readings**
L13	Mar 2	Multiple testing (Assignment 4 posted)	Section 3.4
T6	3	*Factorial ANOVA, working with "real" data; midterm questions*	
L14	7	Test for heterogeneity of slopes, Analysis of Covariance, comparisons of adjusted means	Chapter 12, to section 12.4; section 12.5, 12.6, 12.8
L15	9	It's all just general linear modelling, man (this is where we blow your mind); dummy variables (Assignment 4 due)	Section 6.1.14
T7	10	*Power analyses, Multiple comparisons*	
	Mar 11	*Final day to withdraw without penalty*	
L16	14	Tests of frequencies- Alex Ross (Assignment 5 posted)	Chapter 14, to section 14.2.2
L17	16	Non-parametric tests- Alex Ross	Section 3.3.3, section 5.1.2, Section 8.5.2, 10.5
T8	17	*Comparing slopes, ANCOVA*	
L18	21	Randomization- permutation tests (Assignment 5 due)	Section 3.3.2; readings as assigned
L19	23	Randomization- bootstrapping tests	
T9	24	*Non-parametric tests and tests of frequencies*	
L20	28	Cody Dey 1	
L21	30	Cody Dey 2 (Assignment 6 posted)	Chapter 13 to up to and including section 13.3; assigned reading
T10	31	*Randomization*	*Frequency tests; Traditional non-parametric tests (posted online)*

Lecture (L) or Tutorial (T) #	Date	Topic	Recommended readings
L22	Apr 4	General***ized*** linear models*	Practice finals posted
L23	6	Mixed effects models* (the new way) **AND/OR** Model selection criteria* (a requiem for the *p*-value) (Assignment 6 due)	
T11	7	*Data visualization?*	
Monday April 11, 2022	**11**	**Final Exam, 1:00 pm Room AT 1010**	**Grad student final paper due Monday April 18, 5:00 pm**

[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 six 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 six 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 six assignments, 15 points each [60%]

2. Term Test, February 22 [15%]

3. Final Exam [20%]

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 25%** 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 5%** of the course grade for grad students will be for **class participation, and they are expected to have a greater level of engagement than undergraduate students.**

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.