Lakehead University

Biostatistics (Biology 4630/5131, soon to be 3112), Winter term 2016

Instructors:
Lecturer: **Dr. Michael Rennie**
Office: CB4050
Phone: 807-346-8760
e-mail: mrennie@lakeheadu.ca
Office hours: Wednesdays, 11-12 pm
Fridays, 1-2 pm

TA: **Brendan Allan**
Office: CB4030
Phone: TBA
e-mail: ballan@lakeheadu.ca
Office hours: TBA

Orillia Tutorial coordinator: **Dr. Gerardo Reyes**
Office: OA 3009 desk 7
e-mail: greyes@lakeheadu.ca
Phone: TBA
Office hours: by appointment

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: Tuesday and Thursday, 1:00 to 2:30 pm, AT 2021 (Thunder Bay), OA 1025 (Orillia)
TUTORIALS: Mondays 10:30-12:30, ATAC 3003 (Orillia to be organized with Dr. Reyes)

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

#	Date	Topic	Recommended readings
	Jan 4 (Tutorial =T)	SELF-DIRECTED TUTORIAL[1]	
1	5	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.
2	7	Correlation, linear regression, model II regression	Chapter 3 to section 5.3.15; section 5.4, 5.7.
	11 (T1)	*Correlation, regression*	
3	12	Multiple regression	Chapter 6 to section 6.1.5

#	Date	Topic	Recommended readings
4	14	Single factor ANOVA, unplanned contrasts (Assignment 1 posted)	Chapter 8 to section 8.1.5; section 8.3, 8.4
	18 (T2)	*Single factor ANOVA*	
5	19	Type I and II error rates; planned contrasts	Section 8.6, Chapter 3, especially section 3.2
6	21	Random effects (Assignment 1 due)	Section 8.2.1
	25 (T3)	*Estimating variance components*	
7	26	Experimental design	Chapter 7 up to and including section 7.2
8	28	Nested ANOVA (Assignment 2 posted)	Chapter 9 to section 9.1.9
	February 1 (T4)	*Nested ANOVA*	
9	2	Randomized block design	Chapter 10 to section 10.10, 10.14
10	4	Factorial ANOVA (Assignment 2 due)	Section 9.2, up to 9.26; 9.28; 9.2.11; 9.4, 9.5
	8 (T5)	*Blocked design*	
11	9	Mixed effects models (the old way)	
12	11	**MIDTERM**	
	15-19	**READING BREAK** (Assignment 3 posted)	
	22 (T6)	*Factorial ANOVA; working with "real" data*	
13	23	Statistical power	Sections 5.6, 8.9, 9.2.13, 10.10
14	25	Test for heterogeneity of slopes, Analysis of Covariance, comparisons of adjusted means (Assignment 3 due)	Chapter 12, to section 12.4; section 12.5, 12.6, 12.8
	29 (T7)	*Comparing slopes, ANCOVA*	
15	March 1	It's all just linear modelling, man (this is where we blow your mind)	Section 6.1.14
16	3	Multiple testing (Assignment 4 posted)	Section 3.4
	7 (T8)	*Multiple comparisons*	

#	Date	Topic	Recommended readings
17	8	Tests of frequencies	Chapter 14, to section 14.2.2
18	10	Non-parametric tests (Assignment 4 due)	Section 3.3.3, section 5.1.2, Section 8.5.2, 10.5
	14 (T9)	*Non-parametric tests*	
19	15	Randomization tests	Section 3.3.2; readings to be assigned
20	17	General linear models* (Assignment 5 posted)	
	21 (T10)	*Randomization*	
21	22	Mixed effects models* (the new way)	
22	24	Mixed effects models* (continued) (Assignment 5 due)	
	28	**EASTER MONDAY**	**NO TUTORIAL**
23	29	Model selection criteria* (a requiem for the *p*-value)	
24	31	Model selection* (Assignment 6 posted **Graduate students only**)	
	April 4 (T11)	**The R olympics**	
25	5	REVIEW	
	7	Assignment 6 due (**Graduate students only**)	

[1]Computers in the tutorial lab will not yet be ready for this week. Rather, the tutorial this week will be self-directed and will involve students loading R and Rstudio on their personal computers, and going through the introductory R code presented in lecture on Jan 5th. Please make time to complete this task during the first week of classes.
*topics during the final 5 lecture 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 (six for grad students) 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 (five for grad students) 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.

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*, 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 your lakeheadu.ca email address at least twice a week.

Grading (undergraduates):

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

2. Term Test, February 24 [20%]

3. Final Exam [35%]

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

(Calculators are allowed for term test and final exam, but no other materials)

***Grading (graduate students):** *Graduate students will not write exams.* For graduate students, 6 assignments are to be completed, with grades from the best five counting towards 50% of the final grade. 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. 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, and assignments.

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
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.

find something helpful, most often on the first page or two.

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.