BIOLOGY 4605/7220 Quantitative Methods in Biology FALL 2021
Version: 13 Oct

Lectures: Mon Wed Fri 12 PM A-1045
Labs: Tue 2-5 or 6-9 PM Online + drop in help

Instructor: David Schneider Office (202 Elizabeth Ave) Tel 864-8393
a84dcs@mun.ca Hours - after class or by appointment (CSF- TBA)

Course Summary. The goal of this course is for you to learn a model based approach to the statistical analysis of research data. Skill and confidence come with practice, so assignments and quizzes will be short and frequent. Lecture material will emphasize principles of good quantitative analysis, illustrated by complete examples. Laboratories will cover the computational aspects of problem solving, with a package of the student's choice.

Goals
1. Principles of good analysis.
2. Skill in application.
3. Capacity for self-instruction.
4. Confer with statistician.
5. Develop critical capacity.
6. Evaluate quantitative presentations.

Evaluation	N	4605	7220
Labs (Peer)	8	30	30
Assignments (P)	7	20	20
Quizzes	7	15	15
Unit 1 Exam	1	5	5
Unit 2 Exam**	1	5	5
Final Exam	1	10	10
Peer Assess	12	15	15
Written Report			30
	37	100	130
**or Worksheet			

Pre-Requisite: 1 course in statistics. 2/3 of students in 2017 considered this unnecessary.

Exams and quizzes are open book, emphasizing ability to use tools, rather than to memorize formulas. Short assignments cover material delivered on the day assigned. Quizzes cover lecture material since the previous quiz. Midterm and final exams will have the same format as quizzes. Examples of quizzes are posted on the course website.

Graduate students (Biology 7220) will be required to prepare a written report on the analysis of a set of data of interest to the student. The topic will be decided during a conference early during the term, then discussed during tutorial sessions. This report will constitute 30% of the final mark.

Contact the instructor if you are unable to complete evaluated work due to acceptable cause.

All course material is on BrightSpace / D2L

Required material: Lecture Notes in Quantitative Biology
Laboratories in Quantitative Biology

Additional material at: www.mun.ca/biology/schneider/b4605

Additional material from web: Review Questions in Quantitative Biology

A calculator is required for quizzes, exams, and Lab 2. The calculator (which can be an app on a portable device) does not require statistical functions but does require y^x and e^x functions.

Policy on labs and assignments. Labs and assignments are due in pdf format on the date stated in the syllabus. Work will be returned to students within a week (usually the next lecture after it is due). Peer assessment will be used so late work will be penalized at 5% off per day (excluding weekends).

Lab 1 is a group project that requires attendance for successful completion.
Labs 2 and 9a are group projects for which attendance is recommended.

Planning for the worst and hoping for the best. The MUN Guide to Campus Recovery
https://www.mun.ca/covid19/return-tocampus/Guide_to_Campus_Recovery.pdf
lists actions we can all take to minimize infection risks. In the event of a rapid rise in COVID19 cases, this course will shift to online for labs and lectures, as delivered in the Fall 2020 term.
For latest information and updates see: https://www.mun.ca/covid19/

About labs. Working together is encouraged in all labs. However, each person is responsible for preparing their *own* written report. Obvious duplicates will be considered misconduct (see below).

About statistical packages. Labs 3, 5, and 6 can be completed in a statistical package or in a spreadsheet using functions and data analysis tools. The GLM labs (5, 6, 7, 8, 9) can be completed in any statistical package with a general linear model routine. Lab 10 can be completed in any package with a logistic regression routine, or with a generalized linear model routine.

Packages that lack a GLM routine and a logistic regression routine are superficially attractive but lack any value in learning principles and best practice in statistical analysis.

Statistical packages consist of line code (you type the command) and a GUI– a graphics user interface (mouse clicks to set up the analysis).

Any of the following packages can be used. The instructor can help you with any of them.

SPSS. Easily learned GUI. Line code for saving procedures. Simple calculations are difficult

Minitab. Easily learned GUI. Line code adequate for GLM and simple calculations.

JMP. GUI for SAS code.

SAS. Relatively easy line code. Gold standard for statistical analysis. Expensive license.

SPlus. GUI for R code. 1-year free license for students is no longer available

RStudio. Open source freeware for editing and executing R code. https://www.rstudio.com/

R. Freeware for statistical computing and graphics. R has overtaken SAS in academic settings because it has no licensing fee. The hidden cost is the learning curve. The course website has code for all the labs.

Q: So which package should I use? *A*: Minitab and JMP are easiest to use. However, MUN no longer pays the rental fee for any package except SPSS.

Please do not print the sometimes verbose files produced by statistical packages. Instead, copy and paste the appropriate sections of output into your lab report or assignment. Note that you will have to use a non-scalable font (`such as courier`) to print or display numerical output (ANOVA tables, *etc.*) without distortion.

Assignments (formative assessment). A1 and A2 will be due the lecture day after the material is presented. Short formative assessements (SA) will be due by the end of the day that the material is presented. Alternates (A3-A5 below) can be substituted for work missed, if desired.

A1. Quantities

In the library or on line, find a journal reporting research results. Open the journal to an article, and list the first defined physical or biological quantity you encounter (if you must move to the next article, then so be it). State the journal name, volume, and page number. For this quantity, provide complete details for each of the 5 components of the quantity: name, symbol, typical value, units, and procedural statement. If a component is not present then state 'not present.'

For each of **3 more** quantities in the journal, complete the following 8-point checklist:

Journal name, volume, and page number _____

name of quantity in words. Present? _____ If so, name is _____

symbol Present in article ? _____

number of values N = _____ or cannot be determined _____

procedural statement Present ? _____

Reproducible by another investigator ? _____

type of measurement scale (nominal, ordinal, interval, ratio) _____

If ratio scale: units = _____

(4 due in all)

Formative Assignments A2-A5 These require data from a text book or published literature. To find published data on a topic of interest to you try Google Scholar. Data often appear in print before 1960, less frequently in subsequent decades. Publications with data are listed on the course website. A3-A5 will not be assigned. They are substitutable for work missed.

A2. Data Equations.. In the published literature find a graph where a regression equation has been displayed. A list of such publications can be found in Brightspace or on the course website: http://www.mun.ca/biology/schneider/b4605/Data/RefswithRegressionEquations.pdf

State the source publication (request bonus point for example not on the course website). Write the equation, write the name of each symbol or parameter value, and give its units. Immediately below the equation (symbolic form) display a data equation for each of 3 different values of the explanatory (X) variable.

A3(alt). Hypothesis testing. Find, in the published literature, two mean values with associated standard deviations and sample size.
1. Report the 6 values with full citation of source of the published data.
2. Compute the t-statistic using the appropriate formula from Ch7.3. State which formula you used and why. Use the generic recipe for decision making with statistics (Ch7.3, Table 7.1) to declare a decision about the two means.

A4(alt). Confidence intervals. For the same data used in A3, compute the confidence limits for each mean. Report all 6 values (means, sd, n), the source of the numbers, and both confidence limits. Use the generic recipe for confidence limits (Ch7.5, Table 7.5a).

To obtain critical *t*-values for confidence intervals, use commands you learned in Lab 3.

A5(alt). Correlation. Find, in the published literature, a table of data that you consider appropriate for correlation. Enter the data into a spreadsheet or statistical package. Compute the mean and variance for each variable. Compute the correlation coefficient. State the source of the data (with full citation), why correlation is appropriate, then display the data (label each column), each mean and variance, and the correlation coefficient. Show calculation of the likelihood ratio from the correlation coefficient. State whether inference to a population is possible and defend your argument for or against calculating a p-value to make the inference.

TABS Teaching assessment by students.
Course evaluation questionnaires are of use to instructors in several important aspects of teaching, especially delivery. Questionnaires at the end of the course are of little use to students, as it is too late for corrective action. No-name questionnaires and polls have been a regular part of these two courses since 1994. They are typically short, a few minutes before the end of a lecture period. Short summaries of the results are usually delivered orally in class.

PABS Peer assessment by students. In the absence of TAs this year, credit will be considered for peer assessment via Peer Scholar in Brightspace. The planned format is synchronous anonymous assessment with discussion. It will be a learning experience for all involved. Peer Assessment will have a weight of 15% if it is used. If not, an alternative discussed and stated early in the course will be used.

* * *

Day	Date	On Web	Topic	Labs	Due	Assesed by
Wed	8-Sep	Ch1	Intro to Course			
Fri	10-Sep	Ch2.1, 2.2	Quantities			
Mon	13-Sep	Ch2.5,2.6	Units, Dimensions		A1 Quantities	A:DCS
Tues	14-Sep	Lab1	Inferential Cards	SN2025	DCS	
Wed	15-Sep	Ch3	Rescaling			
Fri	17-Sep	Ch4	Equations		Lab1, Quiz 1	L:EAG,Q:DCS
Mon	20-Sep	Ch5	Data Eq			
Tues	21-Sep	Lab2	Equations	SN2025 3 PM		
Wed	22-Sep	Ch6.1	Freq Dist I		Quiz2	Q:DCS
Fri	24-Sep	Ch6.2, 6.3	Freq Dist II		Lab2	L:EAG
Mon	27-Sep	Ch7, 7.1,7.3	Inference - 3 Modes			
Tues	28-Sep	Lab3 (+4?)	Freq Dist	C2003	EAG, DCS	
Wed	29-Sep	Ch7.2	Randomization tests		SA/Quiz	DCS
Fri	1-Oct	Unit 1 Exam		A1045	Lab3	DCS
Mon	4-Oct	Ch7.5	Hypoth.Tests & Conf. Limits			
Tues	5-Oct	Lab 5a	Regression	C2003	EAG, DCS	
Wed	6-Oct	Ch8, 9.1	GLM Intro, Regression I		SA/Quiz	DCS
Fri	8-Oct	Ch9.2	Regression II		Lab5a	EAG
Mon	11-Oct	Holiday				
Tues	12-Oct	Holiday				
Wed	13-Oct	Ch9.3, Ch10.2	Regression, t-test		SA	Peer next day
Fri(Tue	15-Oct	Ch10.3, 10.4	1-way ANOVA 2 & 6 PM	C2003	Worksheet	Peer 3 & 7 PM
Mon	18-Oct	Ch11	Rev: 1 Expl Var			
Tues	19-Oct	Lab6a	1-way ANOVA	C2003	DCS	
Wed	20-Oct	Ch12.1	Mult Regr		SA/Quiz	Peer/DCS
Fri	22-Oct	Ch13.1 (13.2)	2-way ANOVA		Lab6a only	EAG
Mon	25-Oct	Ch13.3	Mixed Model - Paired t-test			
Tue	26-Oct	Lab7	2-factor ANOVA	C2003	DCS	
Wed	27-Oct	Ch13.4	Mixed Model - Rand. Block		SA/Quiz	Peer/DCS
Fri	29-Oct	Ch14.1	ANCOVA		Lab7	EAG
Mon	1-Nov	Ch13.6	Nested random effects ANOV/			
Tue	2-Nov	Lab8	ANCOVA	C2003	DCS	
Wed	3-Nov	Ch15 review				
Fri	5-Nov	Unit II Exam or alternative (Wsheet for credit?)		A1045		DCS or Peer
Mon	8-Nov	Ch20.1, 20.4	Correlation & Multivariate Analysis		Lab8	EAG
Tue	9-Nov	Lab9a	GLMM Problem setup	TBA	DCS	
Wed	10-Nov	Ch16, 16.2,3,4	Analysis of Deviance			
Fri	12-Nov	Ch18, 18.1	Logistic regression		SA/Quiz	Peer/DCS
Mon	15-Nov	Ch 18.2	Prospective analysis			
Tues	16-Nov	Lab9b	GLMM execution	C2003	DCS	
Wed	17-Nov	Ch 18.3	Retrospective analysis		SA	Peer/DCS
Fri	19-Nov	Ch 18.6	Logistic ANCOVA		Quiz	DCS
Mon	22-Nov	Ch 17.1	Poisson regression		Lab 9a,b	Peer
Tue	23-Nov	Lab10 optional	Logistic regression		DCS	
Wed	24-Nov	Ch 17.4, 17.5	Contingency tests, Poiss ANCOVA			
Fri	26-Nov	Cn19.1-19.4	Model selection I - EDA and stepwise		Q/SA	DCS/Peer
Mon	29-Nov	Ch19.5	MultiModel Inference (AIC, SIC)			
Tue	30-Nov		Review - GLMM	TBA		
Wed	1-Dec	Worksheet			Q/SA	TBA
Fri	3-Dec	Worksheet	Course review with Q&A		Worksheet	DCS
	15-Dec		Final Exam 9-11 AM	TBA (synchronous)		DCS/EAG

First Day Survey (Anonymous) Biology 4605/7220 Fall 2021

The purpose of this survey is to gather information about student background, including experience and familiarity with online submission of activities (assignments, labs, quizzes, exams) listed on the first page. The information will be used for planning course delivery. The survey assumes that students have Word and a recent version of Adobe Reader with editing tools on their computer. If edit options are not available in your version of Adobe Reader, you can use the signing tool to create a text box.
Save the pdf file and submit it to First Day Survey in Brightspace / D2L.

1. If you were unable to use Adobe, open this survey as a Word Document, add answers in a font of different color. Describe the problem the problem you had with Adobe. Save it as a pdf file and submit it to First Day Survey in BSpace.

2. Which course are you taking? Circle or highlight one. B4605 B7220

   Grad students: list department or program _____

   Honors students: List program _____ and topic _____

3. Using a calculator or spreadsheet can you calculate 5^4 Y / N ? $e^{0.75}$ Y / N

4. Have you ever used the R statistical package ? Y / N
   List other statistical packages you have used. _____ _____ _____ _____

5. Have you ever used one of the functions in a spreadsheet ? Y / N

6. Do you have a laptop or touchscreen device (larger than a phone) that can be used in a classroom setting? Y / N

7. How long has it been since you used logarithms ? _____

8. The variable Y has three values, Y = [9 6 3] What is the mean value of Y ? _____

9. How many online courses have you taken? Circle or highlight one: None A few Many

   Which platforms have you used? WebEx Y/N Zoom Y/N Other _____ _____

   B7220 only: How many courses have you taken with the Brightspace / D2L Learning Management System (LMS) at MUN? None A few Many

   Have you ever taken a course with TopHat?

10. How many university courses in mathematics? _____In statistics? _____

11. How many university courses where you have collected data? _____

12. How many courses where you have analyzed data? _____

13. List any non-classroom experience with quantitative techniques, including any data sets you have collected.

13. Describe what you hope to learn during this course and how it relates to your academic program.