DALHOUSIE UNIVERSITY

Faculty of Science Course Syllabus

Department of Mathematics and Statistics

Statistical Methods for Data Analysis and Inference
STAT2080/MATH2080/ECON2280

Winter 2019~~8~~

Instructor(s): Dr. Joanna Mills Flemming Joanna.Flemming@Dal.Ca Chase 103

Dr. Stuart Carson~~Christophe Herbinger~~ st925638~~Christophe.Herbinger~~@Dal.Ca ~~Biology 4056~~Chase 113

Lectures: MWF 10:35am-11:25am LSC C236; 2:35pm-3:25pm LSC ~~Psychology~~ C236~~P5260~~

Laboratories: NA

Tutorials: Thursday~~W~~ 1~~7~~6:35pm-1~~8~~7:25pm, DUNN 117~~MCCAIN ARTS and SS Aud 1~~

Course Description

This is the usual sequel to STAT 1060.03 or STAT 2060.03. This course introduces a number of techniques for data analysis and inference commonly used in the experimental sciences. Topics covered include model building in linear models, multiple regression, analysis of variance, factorial designs, analysis of covariance using the general techniques for linear models and two and three way tables along with logistic regression. A natural sequel for this course is STAT 3340.03.

Course Prerequisites

STAT 1060.03 or STAT 2060.03 or DISP

The materials you are expected to be familiar with are the following: computation and use of various measures of central tendency and variability; preparation and interpretation of graphical displays of data such as boxplots, histograms and scatterplots; normal and t distributions and the use of tables for these distributions; the difference between populations and samples, parameters and estimates; the concept of sampling distributions and why they are important; construction and interpretation of confidence intervals; elements of hypothesis testing; formation of null and alternative hypotheses and computation and interpretation of p-values.

Course Objectives/Learning Outcomes

The main objective of this course is to provide a solid grounding in practical data analysis and common statistical methods encountered in scientific research. To this end the central emphasis of the course is on Analysis of Variance (ANOVA) and Regression.

DALHOUSIE UNIVERSITY

Outcomes:

• Full understanding of the statistical comparison of two means using both parametric and non-parametric methods,

• Full understanding of one-way and two-way analysis of variance (including assumptions, setup, calculations of key quantities, interpretation, and post-hoc diagnostics),

• Full understanding of correlation as a measure of dependence, including both parametric (Pearson's) and non-parametric (Spearman's) measures of correlation,

• Full understanding of simple linear regression (assumptions, key quantities and formulae, implementation, interpretation and graphical assessment via residuals)

• Basic understanding of multiple regression (assumptions, key quantities and formulae, implementation, interpretation and graphical assessment via residuals),

• Experience with the statistical analysis of categorical/count data in one-way and two-way tables (e.g. chi-squared tests and contingency tables),

• Ability to use modern statistical software (MINITAB).

Course Materials

There is a BrightSpace site for the course. This is where assignment information and announcements will be posted. Students are encouraged to use the discussion board for questions about assignments etc. The Brightspace site also contains a link to the course space on the LON-CAPA（Learning Online Network with Computer-Assisted Personalized Approach） server, where class notes, statistical tables and assignments are found.

There is no required text for this course. However, a detailed set of course notes is provided. Readings will be suggested from the books used recently in STAT 1060 (Stats, Data and Models by DeVeaux, Velleman and Bock) and STAT 2060 (Probability and Statistics by J. Devore).

Students will be required to use statistical software as part of this course. Both ~~The use of Statistical Software will be required for this course. We will make use of both the~~ Minitab and R will be demonstrated~~statistical package, as well as R statistical software~~. Minitab is available in c~~C~~ampus computer labs and through Dalhousie ITS free of charge (Windows only, not Mac OS). The state-of-the-art open-source statistical package R is available from www.r-project.org for Mac OS, Windows, and Linux.

~~The Minitab statistical package will be used in the course, and sometimes for demonstration in the lectures. Minitab (or another statistical software) will be required for portions of some assignments.~~

The LON-CAPA e-learning software will be used for assignments and midterms (as well as for disseminating assignment and midterm marks). LON-CAPA can be accessed from the BrightSpace course space, or directly at capa.mathstat.dal.ca . Further details will be provided in class.

DALHOUSIE UNIVERSITY

The Mathematics and Statistics Student Resource Centre is in Room 119 of the Chase building. Please refer to the website (http://www.dal.ca/faculty/science/math-stats/about/learning-centre.html) for the schedule indicating when tutors with expertise in Statistics will be available to answer questions (on a first come first served basis). There are large tables available for groups to work together. Tutors from the Resource Center will also be available in the Learning Commons at the Killam library.

Course Assessment

Component	Weight (% of final grade)	Date
Midterm 1	15%	February 134th
Midterm 2	15%	March 278th
Final exam	45%	(Scheduled by Registrar)
Weekly Assignments	25%	Weekly

Other course requirements-: There is a weekly tutorial ~~tutorial that takes place Thursdays 5:30-6:30PM in Dunn 117. There are no marks associated with this tutorial. The tutorial is~~ delivered by the Teaching Assistant for this~~e~~ course. Its primary purpose is to review assignment materials, and to provide assistance with using statistical software. ~~Tutorials~~These tutorials are not mandatory, but students ~~you~~ are strongly encouraged to attend.

~~None.~~

Conversion of numerical grades to Final Letter Grades follows the Dalhousie Common Grade Scale:

A+ (90-100)	**B+** (77-79)	**C+** (65-69)	**D** (50-54)
A (85-89)	**B** (73-76)	**C** (60-64)	**F** (<50)
A- (80-84)	**B-** (70-72)	**C-** (55-59)	

Course Policies

This course follows the university policy on "missed or late academic requirements due to student absence" for midterm exams:

https://www.dal.ca/dept/university_secretariat/policies/academic/missed-or-late-academic-requirements-due-to-student-absence.html

Students experiencing a short term absence of 3 consecutive days (or less) which result~~ing~~ed in a midterm exam being missed must do the following:

- Contact the instructor(s) by phone or e-mail prior to the scheduled time of the midterm exam,
- Complete a Student Declaration of Absence Form within 3 calendar days of the last day of absence.

Field Code Changed
Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Not Highlight

DALHOUSIE UNIVERSITY

Note that a maximum of one Student Declaration of Absence will be accepted and no make-up midterm will be offered. Instead, the weight of the missed midterm exam will be transferred to the final exam, such that the final exam will then be worth 60%. If a student misses a midterm and does not submit a Student Declaration of Absence, the mark for the missed midterm will be zero.

If there is a legitimate conflict with the time of the midterm exams (e.g. another course or exam scheduled for the same time) students must inform the professors of this at least 3 weeks in advance of the midterm exam and provide details of the conflict.

There will be nine assignments. These will be online assignments delivered using the LON- CAPA software (see http://capa.mathstat.dal.ca). Late assignments are not accepted.

The morning and afternoon sections follow the same schedule. If one section is cancelled on a particular day, the other section will be cancelled as well. Cell phones and other texting devices should be turned OFF before class begins.

Course Content

Listed below in roughly chronological order are the topics to be covered. Note that these may be altered slightly as the term progresses.

- Study design, causal inference and inference to population
- The central limit theorem; hypothesis testing and confidence intervals
- Comparison of two means - paired samples and independent samples
- Comparison of two means - permutation test, Wilcoxon rank-sum test
- One-way analysis of variance
- Bonferroni method for multiple comparisons
- Assessing the model assumptions - residual plot
- Non-parametric one-way ANOVA - Kruskall-Wallis test
- Two-way ANOVA without interaction
- Two-way ANOVA, with interaction, Randomized block design, Post-hoc comparisons of means
- Categorical data, multinomial distribution and goodness of fit test
- Chi-square tests and contingency tables
- Scatterplots, Pearson's correlation, Spearman's rank correlation
- Regression and least squares estimates
- Coefficient of determination, Residual plots, remedies and transformation
- Inference in regression
- Multiple regression basics, hypothesis testing and inference
- Issues in multiple regression
- ANOVA using regression
- Special topics and review