DALHOUSIE UNIVERSITY

Faculty of Science Course Syllabus
Department of Mathematics and Statistics

Statistical Methods for Data Analysis and Inference
STAT2080/MATH2080/ECON2280
Fall 2017

Instructor:	Stuart Carson	st925638@dal.ca	Chase 113
Lectures:	MWF 09:35-10:25	Studley LSC-COMMON AREA C236	
	MWF 13:35-14:25	Studley LSC-COMMON AREA C240	
Laboratories:	NA		
Tutorials:	M 1635-1725	Studley SIR JAMES DUNN 117	

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 material you are expected to be familiar with is the following. The computation and use of various measures of central tendency and variability; the preparation and interpretation of graphical displays of data such as boxplots, histograms and scatterplots; the 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; the construction and interpretation of confidence intervals; the elements of hypothesis testing; the formation of null and alternative hypotheses and the 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 that one encounters in scientific research. Towards this end the central emphasis of the course is on Analysis of Variance (ANOVA) and Regression.

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 regression methods for both simple linear regression (assumptions, key quantities and formulae, implementation, interpretation, and graphical assessment via residuals)

DALHOUSIE UNIVERSITY

• Basic understanding of multiple regression (assumptions, key quantities and formulae, implementation, interpretation, and graphical assessment via residuals),
• Experience in 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 (e.g. MINITAB, R).*

Course Materials

- There is an Brightspace site for the course (https://dal.ca/brightspace). This is where course announcements will be posted. CAPA (http://capa.mathstat.dal.ca) will be used for assignments and the midterms (as well as for disseminating assignment and midterm marks). To log in to CAPA, user your net-id (all lower case) for your username and your initial password is your B00# with a capital 'B'.

- There is no required text for this course. However, a detailed set of course notes will be 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).

- The Minitab statistical package will be required for portions of some assignments, and sometimes used for demonstration in the lectures and tutorials. Students may wish to use other statistical packages, such as R.

- 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 *a schedule and when tutors with expertise in Statistics will be there and 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 Centre will also be available in the Learning Commons at the Killam library.*

Course Assessment

Component	Weight (% of final grade)	Date
Midterm 1	*15%*	*October 16th 18:00-19:30*
Midterm 2	15%	Novermber 20th 18:00-19:30
Final exam	45%	Scheduled by Registrar
Assignments	25%	Weekly or bi-weekly

Other course requirements

NA

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

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

DALHOUSIE UNIVERSITY

If there is a legitimate conflict with the times of the midterm exams (this means another course or an exam scheduled for the same time), students must inform a professor of this at least 3 weeks in advance of the exam with details of the conflict. If an exam is missed for medical reasons, students must contact a professor within 24 hours of the exam and provide a medical excuse within 48 hours. If an exam is missed without a valid reason a zero grade may be assigned.

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-squared 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