

**Course Information**  
**Statistical Methods for Data Analysis and Inference**  
**STAT2080/MATH2080/ECON2280**  
**Summer 2019**

- **Lectures**

Monday/Wednesday 6:05 - 8:55 (HH 212)

- **Class Instructor**

Name: Holly Steeves

Office: LSC 7088

Email: holly.steeves@dal.ca

Office hours: Monday/Wednesday 4:00 - 5:30

Office hours will be held in the Math and Stats Learning Center (Chase 119) if available; LSC 7088 otherwise.

- **Marking Scheme**

Assignments 25%

Midterm 25%

Final Exam 50%

If there is a strong improvement from midterm to final, we can discuss alternative marking schemes, but the final decision will be my call.

- **Assignments** There will be 9 assignments, broken up into two due dates. The best 8 of the 9 assignment marks will be used for the assignment grade.

- **Exams**

The midterm will be held in class. It will be held on **Wednesday, July 24, 6:05PM**. You are allowed a 1 one sided legal size (the long one) cheat sheet that is **HAND WRITTEN!** Also ensure you have a calculator (non-programmable)! The final exam is 3 hours long and will be held in class on **Wednesday, August 21, 6:05PM**. You will be allowed 1 two sided legal size cheat sheet that is hand written for the final.

- **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. A natural sequel for this course is STAT 3340.03.

- **Course Prerequisites** STAT 1060.03 or STAT 2060.03 or DISP

The material with which you are expected to be familiar follows. 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 quantiles, interpretation, and post-hoc diagnostics).
- Full understanding of correlation as a measure of dependence, including both parametric (Pearsons) and non-parametric (Spearman's) measures of correlation.
- Full understanding of regression methods for both simple linear regression (assumptions, key quantiles and formulae, implementation, interpretation, and graphical assessment via residuals).
- Basic understanding of multiple regression (assumptions, key quantiles 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. R).
- An online refresher is available at <http://davidmlane.com/hyperstat/>.

- **Textbook**

There is no required text for this course. 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 readings list is found on the course Brightspace site.

- **Course Website**

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. To log in to CAPA, use your net-id (all lower case) for your username and your initial password is your B00# with a capital B.

- **Mathematics and Statistics Student Resource Centre**

The Mathematics and Statistics Student Resource Centre is in Room 119 of the Chase building. There are large tables available for groups to work together, tutors with expertise in Statistics will be there and available to answer questions (on a first come first served basis), please check there for summer hours and availability.

- **Conversion to Letter Grades**

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

- **Course Policies**

- Cell phones and other electronic messaging devices should be turned OFF before class begins.
- If you must use a computer during class, please sit at the back so you do not distract other students.
- If you have to leave early (please do not make this a regular occurrence, it is very disruptive), take a seat near the exit and leave as quietly as possible.
- Be on time for class. Entering late is extremely disruptive to the instructor as well as the other students who arrived on time to learn.
- Refrain from talking when the instructor or a student is speaking to the class. If you have a question, direct it at the instructor not your classmates. That is what I am here for!
- Treat the instructor and your classmates respectfully and with the courtesy with which you would like to be treated.
- Late assignments cannot be submitted and will not be accepted.
- If for any reason you will miss a midterm or final, you will require a doctors note. This will be handled in a way suitable to both parties.

- **Statement of Accommodation**

Students who require academic accommodation for either classroom participation or the writing of tests, quizzes and exams should make their request to

the Office of Student Accessibility & Accommodation (OSAA) prior to or at the outset of each academic term. Please see [www.studentaccessibility.dal.ca](http://www.studentaccessibility.dal.ca) for more information and to obtain Form A - Request for Accommodation.

A note taker may be required to assist a classmate. There is an honourarium of \$75 for the course. If you are interested, please go to the Advising and Access Services Centre, Killam G28 for more information or email [notetaking@dal.ca](mailto:notetaking@dal.ca). Please note that your classroom may contain specialized accessible furniture and equipment. It is important that these items remain in the classroom so that students who require their usage will be able to participate in the class.

- **Statement on Intellectual Honesty**

Please see the section in the undergraduate calendar on Intellectual Honesty: [http://www.dal.ca/content/dam/dalhousie/pdf/university\\_secretariat/POLICY%20on%20Intellectual%20Honesty.pdf](http://www.dal.ca/content/dam/dalhousie/pdf/university_secretariat/POLICY%20on%20Intellectual%20Honesty.pdf).

- **Policy**

Policy If any exam is missed for medical reasons, the students must contact the professor within 24 hours of the exam. If an exam is missed without a valid reason a zero grade may be assigned. Students can work at their own pace on the assignments, following the pace of the lectures. For this reason any unsubmitted assignment gets zero, no exemptions will be granted under any circumstances.

- **Tentative Topic Outline**

Class introduction and material review

The central limit theorem, hypothesis testing, and confidence intervals

One sample t-test and t-interval

Comparison of two means - independent samples

Comparison of two means - permutation tests, Wilcoxon rank-sum test

One way analysis of variance

Bonferroni method for multiple comparisons

Assessing the model assumptions - normal QQ plot, residual plot

Non-parametric one way ANOVA - Kruskal-Wallis test

Two way ANOVA, with and without interaction

Post-hoc comparisons of means

Randomized block design

Comparison of several proportions

Multinomial distribution, goodness of fit test

$\chi^2$  test for contingency tables

Scatterplots, Pearson's correlation

Least squares estimates, ANOVA

Coefficient of determination

Residual plot, remedies, adding variables, transformation

Inferences for slope and intercept, and mean response

Multiple regression basics, types of models, matrix approach

Overall F test, inference for a single  $\beta$

Inferences for the mean at  $x^*$ , prediction intervals for a new  $y$  at  $x^*$   
Partial F test for comparing nested models  
Comparing two regression lines  
One way ANOVA using regression  
Two way ANOVA using regression