DALHOUSIE UNIVERSITY

Faculty of Science Course Syllabus
Department of *Mathematics and Statistics*
Stat 3340/Math3340
Regression and Analysis of Variance
Fall, 2024

Dalhousie University is located in Mi'kma'ki, the ancestral and unceded territory of the Mi'kmaq. We are all Treaty people.

We acknowledge the histories, contributions, and legacies of the African Nova Scotian people and communities who have been here for over 400 years.

Instructor: Bruce Smith bruce.smith@dal.ca

Lectures: Tuesday, Thursday 10:05-11:25 , LSC 236

Office hours: Mondays 2-4 PM in Chase 309. Also Thursdays 12:30-1:30 (Sept 12, 19, 26 only), and Wednesdays 10-11 AM beginning October 2.

Learning Centre: There will be a tutor in the learning centre (Chase 119) who is comfortable with Stat3340 material on Monday-Friday, 6:30-7:30 PM.

Course delivery: In person. Lectures will NOT be recorded.

Course Description: A thorough treatment of the theory and practice of regression analysis. Topics include: fitting general linear models using matrices, optimality of least squares estimators (Gauss-Markov theorem), inferences, simple and partial correlation, analysis of residuals, case-deletion diagnostics, polynomial regression, transformations, use of indicator variables for analysis of variance and covariance problems, model selection, and an introduction to nonlinear least squares. This course makes extensive use of computer packages.

Course Prerequisites: Math 1030.03 AND STAT 2060.03 or MATH 2060.03 AND STAT 2080.03 or MATH 2080.03 or ECON 2280.03 or 3338.03, AND STAT 2450.03. (Stat2450 may be taken as a co-requisite.)

Course Objectives/Learning Outcomes

1. Calculate the least squares intercept and slope given (x_i, y_i) data.
2. Evaluate expectation and variance of linear combinations of random variables using matrix operations.
3. Calculate least squares estimates for the multiple regression model using matrix operations.
4. Prove that the least squares estimates are unbiased, and evaluate their covariance matrix.
5. Calculate and interpret entries in the analysis of variance table.
6. Calculate case deletion diagnostics given residuals and leverage values.
7. Determine adequacy of regression models using residual plots, leverage measures and case deletion diagnostics.

DALHOUSIE UNIVERSITY

8. Choose a transformation of the response to produce variance homogeneity, using computer plots and likelihood methods.
7. Select an appropriate multiple regression model given plots and descriptions of the variables.
8. Describe the importance and use of the Gauss-Markov and Cochran's theorem in regression modelling.
9. Construct confidence intervals for model parameters given computer output or formulae for the standard error.
10. Carry out hypothesis tests on single or several model parameters to determine whether some variables can be removed from a model, using computer output.
11. Given replicate observations, test for lack of fit of a model.
12. Use the statistical package R to plot and manipulate date and to fit simple and multiple regression models.
13. Use R to calculate automatic variable selection strategies, given a response and several predictor variables.

Course Materials

- Recommended reading: Introduction to Linear Regression Analysis, 5th Edition, Montgomery, Peck and Vining. A copy will be put on 2 hour reserve in Killam library.
- A linear algebra reference: Matrix Theory and Linear Algebra, Selinger, online at https://www.mathstat.dal.ca/~selinger/linear-algebra/downloads/LinearAlgebra.pdf
- Most assignments will require use of the statistical software package R. You can download the software at http://www.r-project.org/
- Rstudio, with the Markdown addon is required to organize assignments. It is available at https://www.rstudio.com/
- Lecture notes, assignments, etc., are at chase.mathstat.dal.ca/~bsmith/stat3340
- Grades will be posted to Brightspace.

Course Assessment

Component	**Weight (% of final grade)**	**Date**
Midterm exam	20%	Tuesday, October 15 (IN CLASS TIME)
Final exam	40%	(Scheduled by Registrar)
Assignments	*6* assignments totalling 40%,	roughly every two weeks.

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)	

DALHOUSIE UNIVERSITY

Course Policies

Assignments are to be submitted to Brightspace. Each assignment is to be submitted as a single pdf file, with questions answered in the order given.
Late assignments will not be accepted.
It is expected that each student will write up their assignment independently. Students submitting identical assignments will receive a mark of 0 for that assignment.
If you are ill on the day of a test, you must advise me of this fact before the test, and you will need to submit a Student Declaration of Absence form before you can write a make-up test.

Course Content

- Multiple linear regression and its matrix formulation
- Assessing model adequacy and transformations
- Regression diagnostics (influential points and multicolinearity)
- Special types of regression: polynomial and indicator variable regression
- Model building and variable selection
- Special topics (causal inference, spline based models - time permitting)

link to University Policies and Statements

link to Student Resources