

University of Toronto Department of Mechanical and Industrial Engineering

MIE1413: Statistical Models in Empirical Research (START: TBD)
Winter 2021

Prof. Birsen Donmez Office: RS305A Office hours: Tues 9:30-11:00 or by appointment
donmez@mie.utoronto.ca

Recommended Text (can be found at Engineering Library Reserves):

(K) Applied Linear Statistical Models, 5th Edition, Kutner, Nachtsheim, Neter, & Li

(F) Applied Regression Analysis and Generalized Linear Models, 2nd Edition, John Fox Appendices and datasets:
<http://socserv.socsci.mcmaster.ca/jfox/Books/Applied-Regression-2E/index.html>

(R) An R Companion to Applied Regression, 2nd Edition, John Fox, Sanford Weisberg

(M) Generalized Linear Models with Applications in Engineering and the Sciences, 2nd Edition, Myers, Montgomery, Vining, & Robinson

Course Objectives:

This course covers statistical models frequently used in empirical research. For various observational and experimental data, students will be proficient in generating relevant hypotheses to answer research questions, selecting and building appropriate statistical models, and effectively communicating these results through interpretation and presentation of results. By the end of the course, students will be able to:

1. Formulate and apply hypothesis testing;
2. Conduct both parametric and nonparametric statistical analyses including linear regression, ANOVA, mixed linear models, and generalized linear models (cluster analysis and time series models will be covered at an introductory level);
3. Address assumptions and limitations of statistical models;
4. Assess whether or not the results statistically support the original research question;
5. Assess the validity of the use of various statistical tests in the literature (journal articles and conference papers)
6. Communicate interpretation of statistical results, both written and orally.
7. Become familiar with one or more statistical software packages.

Homework Assignments: If there is a homework assignment, it will be collected at the beginning of class. A student (randomly selected) will present the homework solutions to the rest of the class. This presentation will count towards the participation grade.

- Unless otherwise noted, assignments are individual efforts.
- Case Studies are papers that you are to analyze from a statistics viewpoint. The analysis is not a formal research report, bullets are fine, but should address whether or not the general study design/data collection method was appropriate, if the appropriate statistical tests were used (and if not, what should have been used), if the results were presented in a clear and concise manner, and what the overall contribution of the statistical analysis was to the main theme/research questions of the paper?

Software: In class we will cover R and SAS. R is a free software package, available at <http://www.r-project.org/>. SAS is free to students: http://www.sas.com/en_us/software/university-edition.html. For the homework or the project assignments, students are encouraged to use either R or SAS. Please talk to the instructor if you want to use another package.

Final Project: For the final project, students will either conduct an experiment or analyze observational data that they acquired. Students are encouraged to relate this assignment to their current research. Students will form hypothesis, analyze the results, and communicate the conclusions through both a written report in research paper

format as well as in a presentation at the conclusion of the semester. Students will submit a proposal for the project mid-semester. Students can work in pairs on the project (no more than 2 to a project; a group of 3 requires prior instructor approval). If 2 people turn in a single assignment, it is expected that the analyses will be more detailed and include more tests/variables than that for a single student.

Below is a list of lessons learned from past regarding the project:

- Clearly state your research question(s)
- Specify independent variables (or predictor variables for observational studies) and dependent variables
- Be mindful of overall type I error: do not fish for significant results.
- If you have multiple independent/predictor variables, make sure you address multicollinearity
- Ensure you submit an REB application if this is part of a funded research program and you are involving human subjects
- Address how you checked assumptions
- If you ran an experiment,
 - o be mindful of statistical power: often students select too many independent variables for their sample size
 - o making a mistake in the design and conduct of your experiment is expected! If you discover an error once it is too late, just simply address what happened in your report and explain how you would correct this in the future. You will likely get a better grade for a mea culpa and lessons learned!
 - o depict your test matrix

Grade Basis:

Final Project: Written	30%
Final Project: Oral	10%
Final Exam	25%
Homework:	15%
Participation:	5%
Case Studies	15%
Total	100%

Academic Honesty Policy: UofT’s academic integrity policy states that “Honesty and fairness are considered fundamental values shared by students, staff and faculty at the University of Toronto. The University’s policies and procedures that deal with cases of cheating and plagiarism are designed to protect the integrity of the institution. As a result, the University treats cases of cheating and plagiarism very seriously. Any student accused of committing an academic offence will find that the accusation is dealt with formally and that the penalties can be severe if it is determined that they did, in fact, cheat.” (<http://www.utoronto.ca/academicintegrity/>)

Specifically in this class, collaboration is allowed for the projects but case summaries and homework are individual efforts. All referenced work should be appropriately cited. If there are any questions, feel free to contact the course instructor.