

STA302-Methods of Data Analysis I-Winter 2019

Class time and location: Tuesday 10-12am Thursday 10-11am in HS 610

Instructor: Fode Tounkara

Office: SS 6011

E-mail: f.tounkara@utoronto.ca

Office hours: Thursday 2-4 pm or by appointment

Course Description

The course is intended to gain a solid understanding of the theory and application of linear regression analysis and practical skills for developing linear regression models for inference, prediction and interpreting the results. You will learn about the following: initial examination of data, correlation, simple and multiple regression using least squares, key concepts in inference for regression parameters for normally distributed errors including confidence and prediction intervals, model diagnostics and remedial measures when the model assumptions are violated, interaction and dummy variable, ANOVA and model selection.

Important Note: I strongly encourage you to attend lectures and take notes. You are also strongly encouraged to take advantage of the office hours to discuss any questions that you have.

Course Objective

At the end of this course, student will:

- Know how to perform a linear regression and a correlation analysis
- Know practical skills for making inference and prediction under linear regression models
- be able to use R software to interpret

Prerequisites

Students should have a second year statistics course such as STA 248H1/ STA 255H1/ STA 261H1/ ECO227Y1. **Students of any major are discouraged from taking only STA 302.** Students are expected to also have the mathematics pre-and co-requisites required by students in these second-year statistics courses. You still need to know basic matrix operations. A short but good review of matrix is available on the website at https://onlinecourses.science.psu.edu/statprogram/matrix_review. Also, for a good review of statistics background you should have for this course, I found very helpful the document created by Pr. Alison Gibbs

<http://www.utstat.utoronto.ca/alisong/Teaching/1112/Sta302/backgroundstats.pdf>

Resources

Course webpage

- Lectures notes, assignments are available through the learning portal at <https://q.utoronto.ca>.

Two recommended textbooks

1. *A Modern Approach to Regression with R*, by Simon J. Sheather
2. *Applied Linear Regression Models*, 4th edition by Kutner, Nachtsheim, and Neter

Software

In this course, **Statistical programming language and environment R** will be used to analyse data. **R** is freely available for download at

- **R** : <https://www.r-project.org/>

Rstudio is a good integrated development environment to **R**. It is also freely available at:

- **Rstudio** : <https://www.rstudio.org/>

Good online reference

- <https://cran.r-project.org/doc/contrib/Torfs+Brauer-Short-R-Intro.pdf>

For assignment, you will use **Rmarkdown** to write your solution.

- **Good online reference for Rmarkdown** : <https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf>

Evaluation

Grading Distribution

Grades will be calculated using two schemes. The final course grade will be the largest of these two grades.

Assessment	Scheme 1	Scheme 2
Group of Assignments	20%	15%
Midterm Test	30%	35%
Final Exam	50%	50%

Assignments

- Two assignments
 - A1 : Thursday, February 7th, due 23:59
 - A2 : Thursday, March 21th, due 23:59
- Must be written as PDF files using **R Markdown**
- Must be submitted via **Quercus**

- late assignment submission may be penalized without documentation of a valid reason

Midterm exam

- In class, Tuesday February 26th (provisional)

Final exam

- Scheduling by the faculty

Course grades may vary dependent on class performance.

Note: I do not negotiate grades unless a mathematical error has been made on my part.

Missed Tests:

If a test is missed for a valid reason, you must submit the University of Toronto Student Medical Certificate, completed by your doctor, to your instructor within one week of the test. Print on it your name, student number, and date. If documentation is not received in time, your test mark will be zero. If a test is missed for a valid reason, its weight will be shifted to the final exam.

Marking concerns:

Any requests to have marked work re-evaluated must be made in writing within two weeks of the date the work was returned to the class. The request must contain a justification for consideration.

How to communicate with your instructor

- Questions about course material such as:
 - How do I do question 3.7 in the textbook?
 - What is standard deviation?
 - When is the midterm?

should be posted on the discussion forums on Piazza. Questions can be posted anonymously (so that the author is anonymous to other students but not to the instructors), if desired.

- For private communication, such as:
 - I missed the test because I was ill.

e-mail your instructor, and include your full name and student number..

Note: I will only respond to e-mails you send me if they come from your e-mail account *@utoronto.ca*.

University of Toronto academic integrity

You are responsible for knowing the content of the University of Toronto's Code of Behaviour on Academic integrity at http://www.governingcouncil.utoronto.ca/Governing_Council/policies.htm. If you have any question about what is or is not permitted in this course, please do not hesitate to contact your instructor.

Students with Disabilities

I am committed to teach every student in this course. The University of Toronto is committed to accessibility. If you require accommodations for a disability, or have any accessibility concerns about the course, the classroom, or course materials, please contact Accessibility Service as soon as possible at <https://www.utoronto.ca/accessibility>. Students who may need course adaptations because of disability are welcome to make an appointment to see me during office hours.

Your responsibility

The course is designed to actively engage you in the course material. We hope you'll find the subject of statistics interesting, challenging, and fun, and an excellent opportunity to truly learn the material. In order for these sessions to be effective, preparing by learning about the week's concepts through the notes is essential.