STA302H1: Methods of Data Analysis I

University of Toronto

Department of Statistical Sciences

Instructor

Guowen Huang
Email: guowen.huang@utoronto.ca
Section: LEC5101
Lectures: Mondays and Wednesdays 6:00PM - 9:00PM, MC 102
Office Hours: Thursday 11:00AM - 1:00PM
Office Address: 149 College Street - Stewart Building in Room 103A (through room 103)

Marking Scheme

All the tests and exam are close book, without any cheat sheet.

- Test 1: 30%, 27th May
- Test 2: 30%, 10th Jun
- Final Exam: 40%, scheduled by Faculty

Course webpage

q.utoronto.ca

Course Outline

Introduction to data analysis with a focus on regression. Initial Examination of data. Correlation. Simple and multiple regression models using least squares. Inference for regression parameters, confidence and prediction intervals. Diagnostics and remedial measures. Interactions and dummy variables. Variable selection. In addition, a brief introduction (review) to R, N(μ, σ), t(n), F(n_1, n_2), $\chi^2(n)$, 95% CI, Central Limit Theorem (CLT), Hypothesis test, t-test, p-value, ANOVA.

Prerequisite

- STA238H1/ STA248H1/ STA255H1/ STA261H1/ ECO227Y1
- CSC108H1/ CSC120H1/ CSC121H1/ CSC148H1
- MAT221H1(70%)/ MAT223H1/ MAT240H1

Test

The tests will be held during class hours

- Location for Test #1 : Exam Centre, Room 100; 255 McCaul Street, Toronto, Ontario;
- Time for Test #1: 6pm-9pm, May 27th;
- Location for Test #2 : Exam Centre, Room 100; 255 McCaul Street, Toronto, Ontario;
- Time for Test #2: 6pm-9pm, Jun 10th;
- Beside pens, bring a calculator, a pencil and your t-card.

Textbook

- Faraway, J.J. (2015), Linear Models with R, 2nd edition, Chapman & Hall/CRC
- Faraway, J.J. (2005), Linear Models with R, 1st edition, Chapman & Hall/CRC.
- Draper N.R., and Smith, H. (1998), Applied Regression Analysis, 3rd edition, Wiley.
- Faraway, J.J. (2016), Extending the Linear Model with R, 2nd edition, Chapman & Hall/CRC.

Lectures

Lectures are mandatory, and are where the majority of the course material will be delivered. They will be a combination of verbal discussion, writing on the chalkboard, and analyzing data in R interactively on the screen.

Here is a schedule of lecture topics:

Lecture #	Date	Topic
1	May 6th	Syllabus, introduction to R, N(μ, σ), t(n), F(n_1, n_2), $\chi^2(n)$, CLT
2	May 8th	95% CI, Hypothesis test, t-test, p-value, ANOVA
3	May 13th	Correlation, Simple Linear regression model using LSE
4	May 15th	Multiple linear regression using MLE
5	May 22th	Diagnostic for linear regression model
6	May 27th	Test #1
7	May 29th	Prediction from linear regression
8	Jun 3rd	Dummy variables and interactions
9	Jun 5th	Model selection and variable selection
10	Jun 10th	Test #2
11	Jun 12th	More on regression
12	Jun 17th	Course (Exame) Review

Missed Tests

If a test is missed for a valid reason, you must submit documentation to the course instructor.

If a test is missed for a valid medical reason, submit a copy of the University of Toronto Verification of Student Illness or Injury form to your instructor within two weeks of the test. Please scan and email me this form. The form is considered received when I email you back saying "Received; feel better!".

Important: The form must indicate that the degree of incapacitation on academic functioning is moderate, serious, or severe in order to be considered a valid medical reason for missing the term test. If the form

indicates that the degree of incapacitation on academic functioning is negligible or mild then this will not be considered a valid medical reason.

If a test is missed for a valid reason then the weight of the test will be added to the final exam.

Computing

You need to use R and RStudio for this course. You can download R from CRAN

You can download RStudio from https://www.rstudio.com/products/rstudio/#Desktop.

Get the open-source version, which is free and runs on Windows/Mac/Linux.

Calculators

You will need a calculator. Any calculator that has logarithmic functions will be sufficient. Calculators on phones or other devices equipped to communicate with the outside world (for example, through the internet or cellular or satellite phone networks) will not be permitted during the term test and the final exam.

TA support

You can ask questions or discuss course material with the instructor or TAs during office hours, at SS623B.

- Wray, Galen
  galen.wray@mail.utoronto.ca;
  Office time: Wed 3:30pm-5:30pm, from the second week;
- Lei, Sun
  liam.sun@mail.utoronto.ca;
  Office time: Fri 10am-12pm, from the second week.

You can send TAs an email to ask any administrative questions.

How to communicate with your instructor

You are welcome to hat to me after lecture or in office hours.

If you need to email me, please consider the following email format:

Email title: STA302 questions;

Email content:

Hi Guowen,

My name is <name>, student number <student number>, and I am a student in your STA302 class. I would like to …..

<more content>

Thank you,

<name>, <student number>

Accessibility needs

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 Services as soon as possible: accessibility.services@utoronto.ca or http://accessibility.utoronto.ca.