Statistics for Applications (18.650) - Syllabus

Instructor: Victor-Emmanuel Brunel (Office 2-239b)
E-mail: vebrunel@mit.edu

When? Tuesdays and Thursdays, 11am-12:30pm

Where? Room 2-190

Office hours  Wednesdays, 4pm-6pm, location: 2-239b

Description of the course  Statistics of Applications is a course where we will set the theoretical foundations for statistical methods that are useful in many applications. After briefly reviewing what the students should know from probability theory, we will talk about basic concepts in statistical inference and then, we will have thematic sessions, dedicated to several statistical concepts and methods that are used in practice.

Work required from the students

- Homework: There will be six problem sets for this course. The problem sets will consist of theoretical exercises and applications of the course material on simulated or real data. All programmed algorithms will be have to be coded using the R software (free software, available at https://www.r-project.org/)

- Midterm exam: There will be an in-class midterm exam on Thursday, March 17. It will last 1.5 hours. The study material for the midterm exam will include all topics covered in class until the week before (included).

- Final exam: The final exam will last 2 hours on Tuesday, May 17, from 1:30pm until 3:30pm, in room 2-190.
The final grade will be the average of the total homework grade, the midterm exam grade and the final exam grade, with respective weights 20%, 30% and 50%.

**Homework schedule**

- Homework #1: Thursday Feb 18, due Tuesday Mar 1
- Homework #2: Thursday Mar 3, due Tuesday Mar 15
- Homework #3: Thursday Mar 17, due Tuesday Mar 29
- Homework #4: Thursday Mar 31, due Tuesday Apr 12
- Homework #5: Thursday Apr 14, due Tuesday Apr 26
- Homework #6: Thursday Apr 28, due Thursday May 5

**Prerequisites** The students should have taken a first course in linear algebra, in calculus and in probability.

**Outline of the course** The following outline is not definitive and may change in the course of the semester, depending on the time. Each topic is followed by a suggested reference.

- Review of some probability theory: types of convergence, laws of large numbers, central limit theorem, Slutsky’s theorem, Delta method, Gaussian vectors [5]
- Hypothesis testing for parameters [5]
- Hypothesis testing for model fitting, independence and more [5]
- Linear regression, logistic regression [5]
- Linear regression in high dimension and model selection
- Sampling theory
- Nonparametric estimation [4,5]
- Bayesian estimation [3,5]
- Survey sampling [2]
- Extreme values
- Bootstrap [5]
- Principal components analysis
- Classification [1]
- ...

**Recommended references**  This course has no textbook. However, the students can use the following references (all references other than [5] are more advanced and meant for further reading only):


