

ORF 245. Fundamentals of Engineering Statistics

Sebastien Bubeck

Spring 2013. Course information and Syllabus

Lectures: M W F 10:00 - 10:50 am, Bowen Hall 222

Course Summary: A study of fundamentals of statistical methods and their applications, covering the basic concepts of descriptive statistics, probability, discrete and continuous distributions, statistical inference, goodness-of-fit tests and regression analysis.

After this course, you will be able to understand and speak the basic language of statistics and select the right statistical method for your applications. Also, the course will provide the fundamental theoretical background that will allow you to appreciate the strengths and limitations of each method and formulate conclusions accordingly. Finally, and most importantly, you will learn how to use the open source software R (<http://cran.r-project.org/>).

Instructor: SEBASTIEN BUBECK
sbubeck@princeton.edu

Office: Sherrerd Hall 225.

Office hours: Wed. 11:00am-12:00pm, Fri 11:00am-12:00pm

or by appointment.

Assistant instructors:

LU XIA (Head TA)

lxia@princeton.edu

Office: Sherrerd Hall 220.

Office hours: Tue. 1:00-2:00pm, Wed. 9:00-10:00am

WEICHEN WANG

weichenw@princeton.edu

Office: Sherrerd Hall 220.

Office hours: Tue. 9:30-10:30am, Th. 9:30-10:30am

MARTINA MINCHEVA

mincheva@princeton.edu

Office: Sherrerd Hall 222.

Office hours: Mon. 4:30-5:30pm, Th. 2:00-3:00pm

XIAOFENG SHI

xshi@princeton.edu

Office: Sherrerd Hall 222.

Office hours: Mon. 5:30-6:30pm, Wed. 3:00-4:00pm

Precepts:

Mon 3:30-4:20pm EQUAD E225 Martina Mincheva

Mon 7:30-8:20pm EQUAD E225 Xiaofeng Shi

Tue 3:30-4:20pm FRIEND 006 Lu Xia

Tue 7:30-8:20pm EQUAD E225 Weichen Wang

Syllabus:

1. Introduction to data
2. Probability
3. Random Variables
4. Foundations for inference
5. Large sample inference
6. Small sample inference
7. Correlation and Simple Linear Regression
8. Multiple regression and ANOVA

Textbook:

OpenIntro Statistics
Second Edition (2012)
Paperback
ISBN: 1478217200

A printed copy may be purchased from <http://www.amazon.com> for \$9.94

A **FREE** PDF version is available at <http://www.openintro.org>

Reading: The lectures will follow closely the textbook and you are encouraged to bring it with you. In general, it is recommended to read the covered material before and after each lecture. You are responsible for reading the book and understanding the material in the book.

Computation: This course will heavily rely on the use of the statistical software R. It is an open source software available for Mac, Linux and Windows. It can be downloaded for free from <http://cran.r-project.org/>. This software is being used by more and more industries for its flexibility and its constant contributions in the form of packages. The size of current datasets make it mandatory for current students (and future job hunters) to know how to use a statistical software. Many possibilities exist (SAS, SPSS, S-PLUS, ...) but R is likely to become the most widely used software in the future. A calculator will be allowed during the midterm and final exams but no computer.

Precepts: Precepts will cover a mix of course revision, practice exercises and R manipulations. Make sure that R is properly installed for the first precept and bring your laptop to precepts. A precept sheet will be posted on blackboard on Mondays at 12 noon. Presence to the precepts is **Mandatory**.

Homework: A pdf with Multiple Choices Questions will be released on blackboard, every Friday at noon. The answer to these questions should be **submitted directly via blackboard, before 10am the next Friday**. Please ask your preceptor if you need help to enter your responses on blackboard. No late homework will be accepted and a missed homework will receive a grade of zero. All homework will be graded and each assignment will carry the same weight.

Part of the homework will involve using the software R and part of it will involve questions on the statistical concepts covered during the week. On top of the Multiple Choices Questions, each homework might contain one or two exercises that are meant to help you prepare for the midterms/final. These exercises will not be graded, but you are **strongly advised to take them seriously**. A full solution to the homework will be provided on blackboard after the submission deadline.

Class attendance: Attendance of class AND precepts is required. The class covers many conceptual issues and statistical thinking that are not covered in the textbook. They will appear in the midterms and final exams.

Exams and project: Besides homework, three major grades will be assigned.

- In-class midterm exams: Friday March 15 and Friday April 19.
- Take-home project: Dean's date.
- Final exam: date and location to be announced by the registrar.

Missed exams or late projects will receive a grade of zero. Proven emergencies will be dealt with via the Registrar's office. The midterm exams are closed-book and closed-notes. A two-sided, letter-size formula sheet will be allowed. The final exam is open-book and open-notes. Calculators may be used during the exams, but not computers.

The project will be done in groups of three or four. A list of basic questions will be asked in plain English and you will have to answer them using statistical analysis. Each step of the analysis should be backed by a statistical procedure. Here too, creative thinking is encouraged. For example, you can answer relevant questions that have not been asked or look for publicly available data that may give a better perspective to your report.

You will also have the opportunity to write a statistical report on a project of your choice if you are genuinely interested in a particular topic. For this option, you will have to write a proposal and return it to the instructor via email before the last lecture. In this proposal, the following questions should be addressed.

- Why do you want to work in this project?
- What are the main questions and challenges of the project?
- What dataset will you be using? Be precise: nature of the data, availability, size, format, . . .

Include any preliminary result that you may have. The proposals will be reviewed quickly. Please note that this is a one-time submission and all appropriate information should be included in the proposal. There will be only one round of review. If the proposal is not accepted, you will have to complete the regular project.

Schedule and Grading policy:

Homework	15%	Weekly
Midterm	15% (each)	March 15 & April 19
Project	15%	Dean's date
Final	40%	TBD