

# DATA130004: COMPUTATIONAL STATISTICS

Fall 2018

---

<b>Instructor:</b>	Nan Zhang	<b>Class:</b>	H6201, Wed 18:30 – 21:05
<b>Email:</b>	<a href="mailto:zhangnan@fudan.edu.cn">zhangnan@fudan.edu.cn</a>	<b>Office:</b>	Zibin N210

---

**Teaching Assistant:** Yunlu Chen, [18210980003@fudan.edu.cn](mailto:18210980003@fudan.edu.cn)

**Course Websites:**

1. <https://zhangnanfudan.github.io/teaching/>
2. Wiki: [shjcx.wang](http://shjcx.wang) (username/password: guest.fudan/shanghai)

**Office Hours:** By appointment.

**Textbook:** Maria Rizzo (2007). *Statistical Computing with R*, CRC

**Main References:** There are various interesting and useful books related to this course. You can consult them occasionally.

- Geof Givens and Jennifer Hoeting (2012). *Computational Statistics*, Wiley
- Norman Matloff (2011). *The Art of R Programming: A Tour of Statistical Software Design*, No Starch Press
- Garrett Grolemund and Hadley Wickham (2017) *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*, 1st Edition, O'Reilly Media

**Objectives:** This course is designed for advanced undergraduates majoring in mathematics, statistics, and computer science. It introduces statistical programming language R and covers modern topics in computational statistics.

**Prerequisites:** Basic knowledge of probability and statistics is assumed. Some coding experience is recommended.

**Tentative Outline:**

- R Basics
- Probability and Statistics Review
- Methods for Generating Random Variables: inverse transformation; acceptance-rejection
- Visualization of Multivariate Data: surface plots, 3d scatter plots, and contour plots
- Monte Carlo Integration and Variance Reduction: example and applications
- Reproducible research, dynamic report generation with *knitr*
- Monte Carlo Methods for Estimation and Hypothesis Tests
- Bootstrap and Jackknife

- Markov Chain Monte Carlo Methods: the Metropolis-Hastings algorithm and the Gibbs sampler
- EM algorithm
- Numerical Methods in R: root-finding in one dimension, numerical integration, linear programming, and etc.

### Grading Policy:

Homework .....	(15%)
Wiki contribution .....	(5%)
Quizzes .....	(20%)
Midterm .....	(30%)
Final .....	(30%)

**Homework:** Problems will be assigned on course website after class meetings and will be due in class on the following week. No late homework will be accepted. Missed homework will receive a grade of zero. Each assignment carries equal weight. You are encouraged to work with other students on the homework problems, however, verbatim copying of homework is absolutely forbidden. Therefore each student must ultimately produce his or her own homework to be handed in and graded.

**Wiki contribution:** Wiki page is designed as a comprehensive resource for this course. Everyone can make contribution. Homework questions and extra exercises will be listed on it and students are assigned to edit solutions or submit R code. Instructor and teaching assistant will help improve and evaluate each student's work.

**Quizzes:** Two in-class quizzes will be arranged accordingly. Questions are conceptual or related to previous homework.

**Midterm:** Midterm exam is also in-class. It is required and there will be no make-up exam. It is expected to cover Chapter 1 to 6.

**Final:** Final exam is scheduled on 18:30-20:30 of January 9, 2019. It will cover all topics along the semester.

Quizzes and exams are closed-book.

### Class Policy:

- Regular attendance is recommended.
- Academic honesty: no plagiarism is tolerated.