Syllabus of Fudan University

Department: School of Data science Date: Oct. 18, 2018

Course Code	DATA130005.01		
Course Title	Statistics: Principles, Methods and R (I)		
Credit	3	Credit Hours	48
Course Nature	☐Specific General Education Courses ☐Core Courses ☐General Education Elective Courses ☐Basic Courses in General Discipline √Professional Compulsory Courses ☐Professional Elective Courses ☐Others		
Course Objectives	The course covers fundamental aspects of probability and statistics methods and principles. Data illustration using statistical package R constitutes an integral part throughout the course, therefore provides the hands-on experience in simulation and data analysis.		
Course Description	The topics covered in this course include: introduction of R, probability, independence, conditional probability, Bayes' formula, random variables and distributions, moment generating functions, probability inequalities, law of large numbers, central limit theorem, point estimation, maximum likelihood estimation, Fisher's information, asymptotic efficiency, Hypothesis testing, Wald's test, t-tests, likelihood ratio tests, permutation tests, confidence intervals.		

Course Requirements:

1. The students are expected to attend all the lectures and pay attention to the classroom discussions.
2. The students are expected to solve the exercise problems independently and hand in their solutions on time; it is especially important to write R code by themselves to gain the first-hand experience in programming.
3. By the end of the course, the students are expected to master the basic concepts taught in the course and solve real-world problems with statistical methodology, they are also expected to be able to read and write R code independently.

Teaching Methods:

The course will be in the principal form of conventional lectures, combined with classroom brainstorming discussions; the contents will be further consolidated with after-class homework so as to review the lectures; practical problems will be discussed and solved

Instructor's Academic Background: (should supply email and the office hour)

Prof. Gao Fengnan, the principal instructor, is the assistant professor at the school of data science of Fudan University. Prof. Gao studied mathematics and statistics in Nanjing University and earned a Bachelor of Science in 2009. He then left China for Master's study in applied mathematics in Kaiserslautern University of Technology and Eindhoven University of Technology, from both of which he obtained Masters of Science. His master's project in Eindhoven was completed in EURANDOM with the guidance of Remco van der Hofstad and Rui Castro. He moved to Leiden at the end of August 2012 to pursue a PhD in mathematics under the supervision of Aad van der Vaart. In the spring of 2016, he accepted an assistant professorship jointly appointed by the School of Data Science of Fudan University and Shanghai Center for Mathematical Sciences. He carries on research in the interdisciplinary frontiers of probability and statistics in the fair city of Shanghai.

Members of Teaching Team

Name	Gender	Professional Title	Department	Responsibility
Gao Fengnan	Male	Middle	School of Data Science	Principal instructor

Course Schedule (Please supply the details about each lesson with 32 academic hours in a total of 16 weeks):

Week 1----

1 What is R? Installing R, help and documentation
2 data import and export, basic data manipulation
3 computing with data, organizing an analysis

Week 2---

4 sample space and events, probability
5 independent events
6 conditional probability, Bayes' formula

Week 3---

7 distribution functions and probability functions 1
8 distribution functions and probability functions 2
9 discrete random variables,

Week 4---

10 continuous random variables
11 bivariate distributions
12 marginal distributions

Week 5---

13 independent random variables
14 conditional distributions
15 IID samples, transformations of random variables

Week 6---

16 probability inequalities 1
17 probability inequalities 2
18 inequalities for expectations

Week 7---

19 Types of convergence
20 law of large numbers (LLN)
21 central limit theorem (CLT)

Week 8---

22 Mid-term exam
23 Mid-term exam
24 Mid-term exam

Week 9---

25 what is statistics? parametric and nonparametric models
26 fundamental concepts in inference
27 empirical distributions

Week 10---

28 method of moments estimation
29 maximum likelihood estimation (MLE)
30 properties of MLE

Week 11---

31 null and alternative hypotheses
32 p-values
33 two-types of errors

Week 12---

34 the Wald test

35 t-tests and t-intervals
36 likelihood ratio tests

Week 13---
37 Pearson's chi-square-test
38 goodness-of-fit tests
39 permutation tests

Week 14---
40 simple linear regression 1
41 simple linear regression 2
42 least squares estimation 1

Week 15---
43 prediction 1
44 prediction 2
45 multiple linear regression 1

Week 16---
46 multiple linear regression 2
47 model selection 1
48 model selection 2

Week 17 & 18 --- Final Exams

The design of class discussion or exercise, practice, experience and so on:

Students may be asked to do small projects so that they can better understand key concepts and statistical methods for solving real problems.

If you need a TA, please indicate the assignment of assistant:

The TA(s) will assist in grading homework and quiz.

Grading & Evaluation (Provide a final grade that reflects the formative evaluation process):

Final grade will depend on the following components with these proportions: homework (20%), midterm (30%), and final exam (50%). Late, poor attendance of the class will be considered for final grade. The students are only allowed to carry a pocket mechanical calculator in the final exam.

Teaching Materials & References (Including Author, Title, Publisher and Publishing time):

1. Wasserman, L. (2004) *All of Statistics*. Springer (Chapters 1-10)
2. Knight, K. (2000). *Mathematical Statistics*. Chapman & Hall/CRC