# Syllabus of Fudan University

**Department: School Data of Science** **Date: Sep 20, 2016**

| | | | |
|---|---|---|---|
| **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, linear regression model. | | |

**Course Requirements:**

Probability Theory, Linear Algebra

**Teaching Methods:**

The course is carried out mostly by conventional lectures, combined with data analytic studies in R. Homework assignments will be given to help the students review the contents and apply their newly acquired knowledge and tools on real-world data problems.

**Instructor's Academic Background:**

Fengnan Gao is an assistant professor jointly appointed by the School of Data Science and Shanghai Center for Mathematical Sciences. He finished hid PhD with Aad van der Vaart from Leiden University in 2016 and joined Fudan afterwards. He has published papers on Electronic Journal of Statistics and Stochastic Processers and their Applications. He has presented his works in Amsterdam, North Carolina, Frejus, Cambridge and Eindhoven.

| Members of Teaching Team | | | | |
|---|---|---|---|---|
| **Name** | **Gender** | **Professional Title** | **Department** | **Responsibility** |
| Fengnan Gao | Male | Assistant Professor | School of Data Science and Shanghai Center for Mathematical Sciences | Main instructor |
| | | | | |
| | | | | |
| | | | | |

**Course Schedule (Please supply the details about each lesson with 48 academic hours in a total of 8 weeks):**

Week 1---- Introduction to the course and R

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

Week 2---- Probability and Random variables

4 Sample space and events, probability, independent events
5 Conditional probability, Bayes' formula
6 Distribution functions and probability functions, mean and variance, moment generating functions

Week 3--- Distributions and Multivariate distributions

7 Discrete random variables, continuous random variables
8 Bivariate distributions, marginal distributions, independent random variables, conditional distributions
9 Multivariate distributions, IID samples, transformations of random variables

Week 4---- Inequalities and Convergence of random variables

10 Probability inequalities
11 Inequalities for expectations
12 Types of convergence

Week 5---- Limit Theorems and Monte Carlo Methods

13 Law of Large Numbers (LLN) and Central Limit Theorem (CLT)
14 Monte Carlo integrals
15 Importance sampling

Week 6---- Introduction to Statistical Inference I

16 Parametric models
17 Nonparametric models
18 Sampling distributions

Week 7---- Mid-term Exam

Week 8---- Introduction to Statistical Inference II & Bootstrap I

22 Fundamental concepts in inference

23 Empirical distributions
24 Simulations

Week 9---- Bootstrap II

25 Bootstrap variance estimation
26 Bootstrap confidence intervals - Approximate normal intervals and Pivotal intervals
27 Bootstrap confidence intervals - Percentile intervals

Week 10---- Point Estimation I

28 Methods of moments estimation
29 Maximum likelihood estimation
30 Properties of MLE

Week 11---- Point Estimation II

31 Asymptotic efficiency
32 Fisher information
33 Parametric bootstrap

Week 12---- Hypothesis testing I

34 EM algorithms
35 Null and alternative hypotheses
36 *P*-values

Week 13---- Hypothesis testing II

37 Two types of errors
38 Wald test, *t*-tests and *t*-intervals
39 Likelihood ratio tests

Week 14---- Hypothesis testing III

40 Asymptotic distribution of likelihood ratio tests
41 Pearson's chi-square test
42 Goodness-of-fit tests

Week 15---- Regression models I

43 Permutation tests
44 Simple linear regression
45 Least squares estimation

Week 16--- Regression models I

46 Prediction
47 Multiple linear regression
48 Model selection

**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.

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


1. Wasserman, L. (2004) *All of Statistics*. Springer (Chapters 1-10)
2. Casella, G.S. and Berger, R.L. (2002). *Statistical Inference* (2nd edition). Duxbury.
3. Knight, K. (2000). *Mathematical Statistics*. Chapman & Hall/CRC.
4. Pawitan, Y. (2001). *In All Likelihood*. Oxford University Press.
5. Zuur, A., Ieno, E. and Meesters, E. (2009). *A Beginners Guide to R*. Springer.


Table column size can be adjusted according to the content. 2016