

## 复旦大学课程教学大纲

院系：

日期： 2018 年 10 月 18 日

课程代码	DATA130005.01				
课程名称	统计学基础：原理、方法及 R 应用 (I)				
英文名称	Statistics: Principles, Methods and R (I)				
学 分 数	3	周学时	3	授课语言	全英语
课程性质	<input type="checkbox"/> 通识教育专项 <input type="checkbox"/> 核心课程 <input type="checkbox"/> 通识教育选修 <input type="checkbox"/> 大类基础 <input checked="" type="checkbox"/> 专业必修 <input type="checkbox"/> 专业选修 <input type="checkbox"/> 其他				
教学目的	本课程覆盖概率和统计方法及原理的基础知识。统计软件 R 贯穿于课程教学之中，以提高学生在模拟和数据分析方面的动手能力。				
基本内容简介	具体包括：R 语言简介、概率定义、独立性、条件概率、贝叶斯公式、随机变量及分布、矩母函数、概率不等式、大数定律、中心极限定理、点估计、极大似然估计、Fisher 信息、渐近效率、假设检验、Wald 检验、t-检验、似然比检验、置换检验、置信区间等。				
<b>基本要求：（请说明该课程指学生学习、考勤等相关要求）</b> 本课程覆盖概率和统计方法及原理的基础知识。统计软件 R 贯穿于课程教学之中，以提高学生在模拟和数据分析方面的动手能力。一，要求学生做到每节课认真听讲，不迟到，不早退，积极参与课上胡同；二，要求学生按时独立完成作业，尤其是自行完成 R 语言编程作业以锻炼其动手能力。在课程结束之时，要求学生基本掌握课上所提到的基本概念，灵活使用统计方法解决实际问题，能独立阅读和撰写 R 代码。					
<b>授课方式：</b> 本课程以主讲教师讲授为主，兼以课堂学生对若干具体问题分析讨论；通过课后作业对课堂所学内容进行回顾、总结；并用课堂所学知识，解决实际的问题。					

**主讲教师简介：（请注明主讲老师的邮箱联系方式和答疑时间）**

高凤楠，复旦大学大数据学院青年副研究员。本科就读于南京大学统计学专业，2009 年本科毕业后赴德国凯瑟斯劳滕大学及荷兰埃因霍芬理工大学攻读硕士学位，硕士论文在 EURANDOM 研究所 Remco van der Hofstad 及 Rui Castro 指导下于 2011 年 9 月完成。2012 - 2016 年在荷兰莱顿大学师从著名统计学家 Aad van der Vaart 攻读博士。2016 年 9 月之后在由复旦大学大数据学院及上海数学中心合聘并开始任教。主要研究领域包括非参数贝叶斯统计、网络科学中的高维统计推论、复杂网络中的概率方法、社交网络建模及分析。

主讲老师联系方式： 021-65648710 （办公室电话）或 fngao@fudan.edu.cn

**教学团队成员**

姓名	性别	职称	院系	在教学中承担的职责
高凤楠	男	中级	大数据学院	主讲老师

**教学内容安排**（按 36 学时共计 18 周，18 周含考试周，具体到每节课内容）：

**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

**课内外讨论或练习、实践、体验等环节设计：**

学生在作业中会不定期被要求做一些小的项目（projects）以帮助他们理解重要概念，熟悉它们的用处以及通过编程序应用它们。

**如需配备助教，注明助教工作内容：**

希望配备助教帮助批改作业和小测验

**考核和评价方式（提供学生课程最终成绩的分数组成，体现形成性的评价过程）：**

学生的最后成绩将根据平时作业（20%），期中考试（30%），和期末考试（50%）决定。缺席也会影响最终成绩。期末考试考核方式为闭卷。

**教材（包括作者、书名、出版社和出版时间；如使用自编讲义，也请列明）：**

Wasserman, L. (2004) *All of Statistics*. Springer (Chapters 1-10)

**教学参考资料（包括作者、书名、出版社和出版时间）：**

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