

应用统计学

Instructor: 郝壮

haozhuang@buaa.edu.cn
School of Economics and Management
Beihang University

February 27, 2022

自我介绍

北京航空航天大学经济管理学院

数量经济与商务统计系 助理教授

教育经历

经济学博士, 华盛顿州立大学, 2018

经济学硕士, 北京师范大学, 2014

工学学士, 同济大学, 2011

研究兴趣: 应用微观经济学, 健康经济学, 教育经济学, 劳动经济学

主讲课程: 概率论(本), 应用统计学(本), 高级计量经济学(博), 高级计量经济学II(博). 办公室: A1005

课程基本信息

课程的性质： 核心专业类必修课

主讲老师： 郝壮

学分/学时： 3/48 (24+24)

课表： 1-16周 (8周数理统计+8周应用统计)

先修课程： 数学分析, 高等代数, 概率论

开课对象： 经管学院二年级本科生

教学目的和要求:

通过本课程的学习掌握统计推断的基本理论和方法, 初步具备基本的统计思维, 以及应用统计推断方法分析和解决实际问题的能力;

针对实际的社会和经济管理问题, 采用正确的统计分析方法建立合理的定量分析模型, 探索内在的数量规律性, 进而做出正确的决策.

授课方式: 课堂讲授为主, 辅以上机操作

统计软件: STATA/R

- 教材：** 1. 茆诗松, 程依明, 濮晓龙. 概率论与数理统计教程(第3版). 高等教育出版社, 2019年.
2. 徐哲, 石晓军, 杨继平, 王惠文. 应用统计学:经济与管理中的数据分析. 清华大学出版, 2011年第一版

- 课件 (登录北航课程中心: <http://course.buaa.edu.cn/>)
- Lawrence C. Hamilton, (2012). Statistics with STATA (Version 12). Cengage Learning.
- 盛骤, 谢式千, 潘承毅. 概率论与数理统计. 浙江大学. 高等教育出版社, 2008年6月第4版
- 吴喜之, 统计学-从数据到结论, 中国统计出版社, 2004
- Nathan Yau 著, 鲜活的数据—数据可视化指南, 人民邮电出版社, 2012

教学方法, 课内外教学环节及基本要求

讲授教学法为主, 课堂讨论和学生自学为辅.

本课程教学环节包括讲授、课堂讨论、课下作业等.

考核方式及成绩评定

考核方式： 平时成绩(40%)和期末考试(笔试闭卷)成绩(60%)。

平时成绩： 占40%(出勤、随堂测验、日常作业、上机实验、课程报告、中期考评)。

出勤： 随机点名. 缺勤3次不能参加期末考试.

作业禁止抄袭, 以理解课上知识和锻炼解题能力为目的, 对总评成绩影响有限.

作业提交： 课代表交给助教

作业包括： 个人作业+小组作业(期末案例).

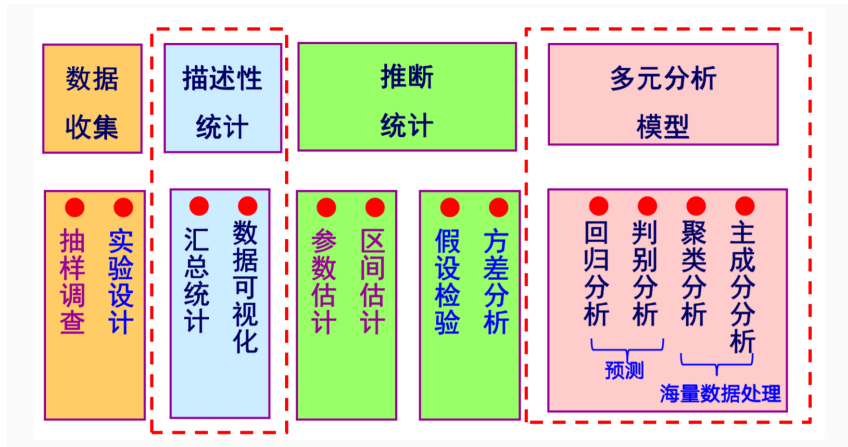
课程助教(作业与答疑): 田曼 (21硕)
405045518@qq.com

助教答疑时间: (提前邮件预约)

选课代表: 收发作业 (学习委员)

统计学

一门关于数据资料的收集、描述、分析、解释的学科(数据
⇒ 结论).



教学大纲(根据接受程度适当调整教学进度)

数理统计

- §1 第五章 统计量及其分布(2周)
- §2 第六章 参数估计(3周)
- §3 第七章 假设检验(2/3周)
- §4 第八章 方差分析(1周)

应用统计

- §1 描述性统计(1周)
- §2 相关分析与一元线性回归分析 (2周)
- §3 多元线性回归分析 (1/2周)
- §4 聚类方法(1周)
- §5 主成分分析 (1周)
- §6 分类方法: 判别分析, logistic (1周)

几点建议

1. 时间关系, 课件上的例子不能全部讲完, 教材中的证明无法全部证完需要大家课后自行完成.
2. 鼓励课上提问, 鼓励课上动笔算.
3. 本课选取的教材是数学专业的本科教材, 证明和理论概念很多, 一些概念在经济管理中不常用, 不求全部掌握.

应用统计学案例(期末小组作业)提交要求

1、选题：

- 国内外值得关注的经济、社会热点问题
- 在日常生活中发现的具有“公共性”的问题

2、注意：在能力、资源、时间方面的可操作性

- 比较熟悉问题的基本概念，具有一定的先验知识
- 数据有可把握的来源
- 可以找到较丰富的参考文献
- 时间计划合理

3、研究要求：综合运用多种统计分析方法，对相关问题的数据进行描述和分析，并形成相应的研究报告。

鼓励采用一些课堂上没有讲到的内容，如机器学习方法，计量经济学方法等。

二、研究报告的格式

- 标题、作者、班级
- 摘要(中英文)、关键字
- 引言(研究意义、文献综述、内容框架)
- 正文(可以分成若干节：介绍数据来源，描述性统计、数据分析等)
- 研究结论或建议
- 研究体会(可选)
- 参考文献与数据来源

提交报告时需要标明小组分工！

小组成员要求： 3-5人

要求： 分工明确， 共同训练; 不 “打酱油”

时间安排：

1. 第3次课上交分组名单和拟选题(学习过程中不断完善、随时沟通)
2. 第16周交研究报告, 并进行典型研究的交流(可选, 需报名).

范例：中国入境旅游发展状况分析

1. 引言：研究目的、意义、文献综述
2. 整体情况分析：旅游业对整个GDP贡献、旅游业对服务业的贡献、入境旅游和国内旅游份额比较
3. 中国入境旅游发展历史：历年接待人数、入境旅游外汇收入
4. 需求分析：游客特征、主要客源市场、旅客行为分析、重游率、游因率、驻留率、各省差别
5. 供给分析：硬件分析、软件分析
6. 供需矛盾：旅游业的区域结构失衡、差距分析
7. 建议
8. 数据来源、参考书目

部分数据来源渠道

- Google
- 各种统计年鉴(北京图书馆, 年鉴室)
- 《中国经济统计快报》
- 中国统计数据(中国网的宏观统计数据)
- 国家统计局官网
- 国家旅游局
- 数据堂(行业数据、网络数据, 需要积分才能下载)
- UCI: 机器学习领域的公开数据集(免费)

部分数据来源渠道(国际经济与金融数据)

- 世界银行(涵盖了全球各种类型的统计数据库, 例如世界发展指标, 国际债务统计, 全球经济监测, 教育统计, 社会性别统计, 国际比较项目, 等等)
- 国际货币基金组织IMF(国际收支统计BOP, 国际金融统计IFS, 贸易方向统计DOT, 政府财政统计GFS)
- IMF数据图形可视化服务网址
- 新浪财经(提供全球宏观经济数据, 是世界银行的世界发展指标的中文版)

http://sem.buaa.edu.cn/News/MenuTemplates/IntroMenu5_Resource.aspx?MenuId=549

- 色诺芬数据库(经济、金融)
- 高频数据 (股票)
- 国泰安数据库 (金融)
- 中宏数据库(宏观经济、产业经济)
- BvD数据库(2010年7月1日-2011年6月30 国际商品贸易分析库Chelem)

常用的统计和数据处理软件

- Stata
- R
- SPSS
- Eviews
- Gauss
- Excel
- SAS
- (Python)

R和Stata的比较

R的优缺点

- 优点1: 免费开源
- 优点2: 较大使用社群, 可个人书写软件包, 可扩展性强
- 缺点1: 大量软件包为个人书写, 有时较难保证其正确
- 缺点2: 绝大多数命令为代码完成, 学习曲线较陡峭

Stata的优缺点

- 优点1: 商用软件, 标准化命令包较多, 软件支持好
- 优点2: 即可通过书写代码也可通过UI进行鼠标点击进行操作, 学习曲线较平滑
- 优点3: 较大使用社群, 可个人书写软件包, 可扩展性强
- 缺点: 付费软件, SE版本相对便宜, MP版本较昂贵

R和Stata的比较

- 均为非常常用的功能强大的统计软件, 均可实现本课和未来所有课程的所有功能.
- 两种软件没有对错高低之分, 在工作中使用哪种软件主要是习惯性问题
- 本课的课堂展示教学软件为Stata; 同时为了统一教学内容, 提供R代码供学有余力和感兴趣的同学参考

R软件的装载与学习

安装包括两部分： R (核心软件) 与 RStudio (集成开发环境, IDE)

- R: <https://cran.r-project.org/>
- RStudio:
<https://rstudio.com/products/rstudio/download/>

R语言学习参考资料

- 《R语言入门与数据分析》 <https://www.bilibili.com/video/BV19x411X7C6?p=1>

拓展： Stata软件的学习参考资料

- 《Stata 16 统计分析与应用》 <https://www.bilibili.com/video/BV1Fv411t7aB?from=search&seid=3528457206971009838>

作业0

1. 请同学们下载课程中心所有材料并仔细阅读
2. 确认已准备好上课所需教材和教学软件
3. 安装统计软件.