

《应用数理统计》教学大纲

一、课程编码：22-025200-B01-17

课内学时： 48 学分： 3

二、适用学科专业：应用统计

三、先修课程：概率论与数理统计

四、教学目标

通过本课程的学习，使研究生：

- 1、了解统计的发展史；
- 2、掌握统计的基本概念，基本方法和基本理论；
- 3、掌握统计机器学习的若干基本概念，基本方法和基本理论；
- 4、能够理论和实际结合，提升处理实际数据分析问题能力。

五、教学方式

课堂讲授、案例讨论与分析。

六、主要内容及学时分配

第1章 大数据与统计学 3 学时

§1.1 什么是数据

§1.2 什么是大数据

案例：布罗德街的水井

案例：啤酒与尿布

案例：谷歌流感预测

§1.3 什么是统计学

第2章 批判性数据思维 3 学时

§2.1 试验设计和抽样方法

案例：物理防晒完胜化学防晒

案例：罗斯福将败给兰登

案例：高考状元的典型特征

§2.2 统计分析

案例：一杯红酒可以代替一小时健身

案例:巧克力可以减肥	
案例:统计学与杀婴母亲	
第 3 章 描述性统计	6 学时
§ 3.1 平均数中位数众数	
案例: 平均数的发展史	
案例: 平均工资的度量	
§ 3.2 统计图	
案例: 丑图百讲	
§ 3.3 可视化	
案例: 北京市空气质量报告	
第 4 章 假设检验	15 学时
§ 4.1 假设检验的基本思想	
案例:Neyman 与 Fisher 关于假设检验的分歧	
§ 4.2 p 值	
案例:美国统计学会关于 p 值的六条准则	
案例:很多心理学论文重复检验失败	
案例: 统计误差	
§ 4.3 Bayes 检验	
案例: Bayes 身世之谜	
§4.4 列联表中的独立性检验	
案例:药效问题中的 Simpson 悖论	
案例:性别歧视问题中的 Simpson 悖论	
第 5 章 Bayes 方法	6 学时
§ 5.1 Bayes 公式	
案例: 失联搜救中的 Bayes 分析	
§5.2 朴素 Bayes	
案例: 垃圾邮件过滤	
§ 5.3 Bayes 网	
第 6 章 回归分析	9 学时
§6.1 线性回归	

案例：北京市空气质量预测

案例：北京市房价预测

§6.2 Logistic 回归

案例：互联网征信中的信用评分模型

案例：基于 UBI 的车险数据分析

§6.3 Lasso 回归

第 7 章 文本挖掘与社交网络分析

6 学时

§7.1 文本挖掘

案例：旧报纸里的英国

§ 7.2 社交网络分析

案例：社交网络中我的标签是否会影响我的社交圈

案例：唐朝诗人之间社交关系分析

七、考核与成绩评定

课程论文结合平时作业和考勤，成绩以百分制衡量。

成绩评定依据：平时作业成绩占 25%，考勤占 15%，课程论文 60%。

八、参考书及学生必读参考资料

王汉生. 数据思维：从数据思维到商业价值. 中国人民大学出版社:2017.

张忠占,徐兴忠. 应用数理统计. 机械工业出版社:2008.

韦来生. 数理统计. 科学出版社:2008.

九、大纲撰写人：董岩