

概率统计（理工类）课程简介

李孟棠，朱彬

中山大学智能工程学院

limt29@mail.sysu.edu.cn, zhub26@mail.sysu.edu.cn

2021年3月2日

概率论和统计学——联系紧密但不同的数学分支

概率论(Probability Theory): 基于公理的演绎式(deductive)理论, 形式上和欧几里得平面几何类似。

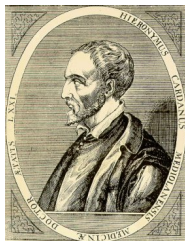
- 给出概率、随机变量及其分布函数、概率密度函数等物体的公理化定义(axiomatic definition);
- 是统计学的理论基础。

统计学(Statistics): 推理式(inferential)理论, 广义上等价于“数据处理”。

- 利用样本(sample)数据中的模式(patterns)对总体(population)做出推理, 同时考虑到数据中的随机性(如误差、噪声等);
- 例如, 研究如何估计某一事件发生的概率。

概率论简史

- 概率的现代数学理论起源于16世纪的Gerolamo Cardano, 17世纪的Pierre de Fermat和Blaise Pascal等人对于博弈游戏(games of chance)的分析。
- 19世纪的Pierre Laplace完善了今天所谓的“古典概型”。
- 早期的概率论主要考虑离散事件, 其方法主要基于排列组合。



- 后来，解析的思想和方法促成了连续变量的引入，并在Andrey Nikolaevich Kolmogorov建立的公理化概率论到达顶点。
- Kolmogorov整合了样本空间(sample space)的概念和测度论(measure theory)，在1933年提出了概率论的公理系统。这成为了无可争议的现代概率论的公理基础。

(更多历史资料请咨询Wikipedia!)



统计学极简史

“Stat-”词源：早期统计学(Statistics)主要服务于国家(states)基于人口和经济数据的政策制定。意大利语“statista”意为政治家(statesman)。

现代数理统计(mathematical statistics)的基础：19世纪末到20世纪初，英国学派。

- **Ronald Fisher爵士，频率学派(frequentist)统计学的奠基人**，发明了“方差”等词汇，在生物遗传、进化等方面做出了统计学的应用。



(更多历史资料请咨询Wikipedia!)

统计学的频率学派和贝叶斯学派：以参数估计为例

频率/Fisherian学派主导了20世纪的统计研究和应用，主要观点：存在一个未知但是**确定性的(deterministic)**的真实参数。

贝叶斯学派(本课程不涉及)主导了20世纪之前的统计研究和应用，主要观点：**没有真实的参数**。反之，应该把参数看成一个“服从一定概率分布的随机变量”，并基于贝叶斯定理（18世纪）做出估计。



两个派别对于看待统计问题有着根本的哲学差异(fundamental philosophical difference)。

为什么要学习概率统计课程？

统计学的别名（“马甲”）：

- **机器学习**（统计学习）
- 统计信号处理
- 模式识别
- 自适应信号处理
- **系统辨识**与控制
- 数据挖掘和信息检索

在这些名字下，统计学在当前仍然是一个极其活跃的科研领域，例如“大数据”的分析。

概率论作为统计学不可分割的理论基础，其方法在统计力学、量子力学等物理领域也有广泛的应用。

本门课程将为这些进阶课程的学习和专业领域的研究打下必要基础。

1 概率论(70%以上的课时)

- 事件与概率的公理化定义
- 随机变量和及其概率分布
- 随机变量的数字特征, 包括期望、方差、相关系数等
- 大数定律和中心极限定理及其应用

2 统计学(不到30%的课时)

- 几种抽样分布
- 参数的矩估计和最大似然估计
- 假设检验

The End