



引言： 课程简介和概率统计简史

刘 杰

统计与金融系

2020年02月



教材和参考资料

◆ 教材:

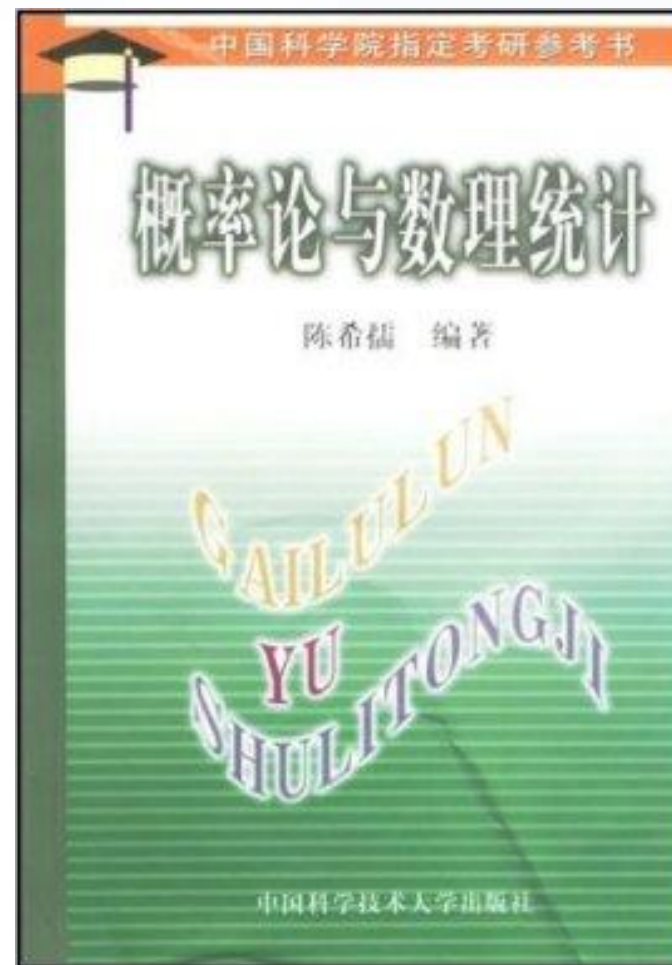
- 《概率论与数理统计》 陈希孺
中国科学技术大学出版社

◆ 习题册

- 《概率论与数理统计习题集》
中国科大统计与金融系

◆ 参考书

- 《概率论》 苏淳 科学出版社
- 《数理统计》 韦来生 科学出版社





课程内容

◆ 概率部分

- 概率定义及其性质(条件概率,独立性)
- 随机变量及其分布(一元,二元随机变量分布, 数字特征)
- 随机变量的收敛性

◆ 统计部分

- 总体与样本
- 参数估计(点估计和区间估计)
- 参数假设检验
- 拟合优度检验(独立性检验)

课程资源

◆ 课程主页: 教务处—BB系统

◆ 论坛: <http://fisher.stat.ustc.edu.cn>

◆ 爱课程网: http://www.icourses.cn/coursestatic/course_3068.html

◆ 成绩评定

30%平时成绩+70%期末考试（左右）

习题课、答疑、随堂提问答题、随机点名等辅助教学和考核方式



教师和助教联系方式

◆ 任课教师

- 姓名：刘杰
- 电话：0551-63607249（办公室：管理科研楼1015室）
- 邮箱：jiel@ustc.edu.cn

◆ 助教

- 钱元辰
电话：15656576201 邮箱：qyc1996@mail.ustc.edu.cn
- 赵家杨
电话：18756080771 邮箱：zhaojy19@mail.ustc.edu.cn



统计学的学科框架

◆ 统计学（0714）是一级学科

➤ 二级学科：

数理统计学，社会经济统计学，生物卫生统计学，
金融统计、风险管理与精算学，应用统计

◆ 学科简介

- 统计是一门艺术，不同于数学，同多学科交叉。
- 统计学是通过搜索、整理、分析、描述数据等手段，以达到推断所测对象的本质，甚至预测对象未来的一门综合性科学。统计学用到了大量的数学及其它学科的专业知识，其应用范围几乎覆盖了社会科学的各个领域。
- 第四轮学科评估，中国科大统计学科A-（全国前5%-10%），刚刚获评国家一流本科。

学校名称	评估结果
北京大学	A+
中国人民大学	A+
南开大学	A
东北师范大学	A
华东师范大学	A
厦门大学	A
北京师范大学	A-
东北财经大学	A-
上海财经大学	A-
浙江工商大学	A-
中国科学技术大学	A-
江西财经大学	A-
清华大学	B+
北京交通大学	B+
北京工业大学	B+
首都师范大学	B+
中央财经大学	B+
对外经济贸易大学	B+
首都经济贸易大学	B+
吉林大学	B+
复旦大学	B+
山东大学	B+



统计学简史

➤ 统计学的英文statistics

最早源于现代拉丁文Statisticum Collegium（国会）、意大利文Statista（国民或政治家）以及德文Statistik，最早是由Gottfried Achenwall于1749年使用，代表对国家的资料进行分析的学问，也就是“研究国家的科学”。

➤ 17世纪，费马与另外一位学者研究赌博，发现概率

100 VS 100, 三局两胜，赌金分配问题

➤ 前苏联Kovgomorov建立概率的公理化体系（类似几何公理化）

➤ 逐步发展完善：数据分析、数据挖掘等

➤ 当今：人工智能和大数据时代

统计实例1：18世纪巴黎男女婴比例

- 公元1814年，法国数学家拉普拉斯在他的新作《概率的哲学探讨》一书中，记载了一下有趣的统计。他根据伦敦，彼得堡，柏林和全法国的统计资料，得出了几乎完全一致的男婴和女婴出生数的比值是22: 21，即在全体出生婴儿中，男婴占51.2%，女婴占48.8%。可奇怪的是，当他统计1745—1784整整四十年间巴黎男婴出生率时，却得到了另一个比是25: 24，男婴占51.02%，与前者相差0.14%。

对于这千分之一点四的微小差异，拉普拉斯感到困惑不解，他深信自然规律，他觉得这千分之一点四的后面，一定有深刻的因素。于是，他深入进行调查研究的，终于发现：当时巴黎人“重女轻男”，有抛弃男婴的陋俗，以至于歪曲了出生率的真相，经过修正，巴黎的男女婴的出生比率依然是22: 21。

- $22 : 21 = 104.8 : 100$ $25 : 24 = 104.17 : 100$



统计实例2：中国历次人口普查统计数据

◆ 人口普查工作包括对人口普查资料的搜集、数据汇总、资料评价、分析研究、编辑出版等全部过程，它的目的是全面掌握全国人口的基本情况，研究制定人口政策和经济社会发展规划提供依据，为社会公众提供人口统计信息服务。

- 1953年第一次人口普查 人口为601938035人 男女比例：106:100
- 1964年第二次人口普查 人口为720370269人 男女比例：104:100
- 1982年第三次人口普查 人口为1031882511人 男女比例：106:100
- 1990年第四次人口普查 人口为1100173810人 男女比例：106.6:100
- 2001年第五次人口普查 人口为129533万人 男女比例：106.74:100
- 2010年第六次人口普查 人口为1,339,724,852人 男女比例：105.2:100
- 2020年第七次人口普查（国务院决定于2020年11月1日零时开展）

目前我国人口刚刚超过14亿。女寿命平均长4-5年。

世界男女比例101.5:100。国际通行安全标准102-107:100。

人口学家认为：新生儿正常的人口毕约为105:100。



统计学实例3：重男轻女与人口膨胀

- ◆ 假设某地重男轻女现象非常严重，所有父母都非要生男孩为止。但政府又怕这样会导致人口膨胀。所以有人提议：

“每对夫妇可以生育，直至有一名男孩，之后必须绝育”。

这个提议引起激烈辩论：

- (a) 通过该提议会不会导致人口膨胀？
 - i) 会导致家庭人口越来越多
 - ii) 有男孩的家庭不再继续增加人口，人口会减少
- (b) 通过该提议会不会导致男多于女？
 - i) 每个家庭都必有男孩，这样会导致男多于女
 - ii) 很多家庭为了生男孩，导致女孩多于男孩

我们可以使用概率论方法来回答这些争论：

(a) 假设每对夫妇都是可以生育的， X 表示每个家庭的子女数， $p=1/2$ 为生男孩的概率，则 X 服从几何分布。于是每个家庭的平均子女数目为 $EX=2$ ，因此不会增加人口。

(b) 假设女孩子的数目为 G ，则 $X=1+G$ (该提议的结果)，所以 $EX=1+EG$ ，所以 $EG=1$ ，所以该提议不会导致男多于女或者女多于男

统计实例4：各种因素对平均寿命的影响

在《统计与真理》一书中，有关于各种因素对平均寿命的影响的统计数据。

以减少的平均寿命进行计算，单位：天。

序号	因素	减少天数	序号	因素	减少天数
1	未婚男性	3500	7	吸烟女性	800
2	未婚女性	1600	8	抽雪茄	330
3	惯用左手	3285	9	饮酒	130
4	30%超重	1300	10	加油烟雾警报	-10
5	20%超重	900	11	长期压抑	4-8年
6	吸烟男性	2250	12	女友肥胖	+57%



统计实例5：心理效应等

- David Phillips 统计25年内美国华裔妇女中秋节前后的死亡率
 - 中秋节前一周比平均低35.1%
 - 中秋节后一周比平均低34.6%

- 生日效应（有时被称为生日劫难，特别是在特指自杀时）是一种统计现象，其中个人的死亡可能性在其生日或接近生日时增加。
 - ✓ 在英格兰和威尔士，瑞士，乌克兰和美国的一般人群以及美国职业棒球大联盟球员等较小人群的研究中已经看到了生日效应。
 - ✓ 在瑞士人中进行类似研究发现，实际生日的死亡率最高（比预期值高17%）。

统计实例6：蒲丰投针 (Buffon)

- 18世纪，法国数学家蒲丰提出的“投针问题”，记载于布丰1777年出版的著作中：“在平面上画有一组间距为 a 的平行线，将一根长度为 l ($l \leq a$) 的针任意掷在这个平面上，求此针与平行线中任一条相交的概率。”

蒲丰本人证明了，这个概率是： $p = 2l/\pi a$

它与圆周率 π 有关，于是人们想到利用投针试验来估计圆周率的值，扔的次数越多，由此能求出越为精确的 π 的值。



试验者	时间	投掷次数	相交次数	圆周率估计值
Wolf	1850年	5000	2532	3.1596
Smith	1855年	3204	1218.5	3.1554
C.De Morgan	1860年	600	382.5	3.137
Fox	1884年	1030	489	3.1595
Lazzerini	1901年	3408	1808	3.1415929
Reina	1925年	2520	859	3.1795



谢谢！