第一讲：概率论简介

张鑫

Email: xzhangseu@seu.edu.cn

东南大学　数学学院

参考书目

- 概率论, 张颢, 高等教育出版社
- 概率论 (第二版), 杨振明, 科学出版社
- 概率论基础, 李贤平, 高等教育出版社
- 概率论 (第三版), 苏淳, 冯群强, 科学出版社
- 概率论导论 (翻译版), Joseph K. Blitzstein, Jessica Hwang (张景肖译), 机械工业出版社
- 概率导论 (第二版), Dimitri P. Bertsekas, John N. Tsitsiklis (郑忠国, 童行伟译), 人民邮电出版社
- Probability, A. N. Shiryaev, Springer
- 科学发现纵横谈, 王梓坤, 北京师范大学出版社

参考教材封面

概率论
Probability Theory
张颖 编著

"十二五"普通高等教育本科国家级规划教材
南开大学数学教学丛书
概率论
第二版
杨振明 编
科学出版社

"十二五"普通高等教育本科国家级规划教材
概率论基础
第三版
复旦大学 李贤平
高等教育出版社

中国科学院考研指定参考书
普通高等教育"十一五"国家级规划教材
中国科学技术大学数学教学丛书
概率论
(第三版)
苏淳 冯群强 编著
科学出版社

"十三五"国家重点出版物出版规划项目
名校名家基础学科系列
概率论导论
(翻译版)
Introduction to Probability
[美] 约瑟夫·K. 布利茨斯坦 (Joseph K. Blitzstein) 著
哈佛大学
杰西卡·黄 (Jessica Hwang)
斯坦福大学
张景肖 译
CRC Press
Taylor & Francis Group
机械工业出版社
CHINA MACHINE PRESS

TURING 图灵数学·统计学丛书
美国工程院院士力作，MIT等全球众多名校教材！
Introduction to Probability, 2E
概率导论
[美] Dimitri P. Bertsekas John N. Tsitsiklis 著
郑忠国 童行伟 译
(第2版·修订版)
从直观、自然的角度阐述概率，是理工科学生入门首选。
中国工信出版集团
人民邮电出版社
POSTS & TELECOM PRESS

Graduate Texts in Mathematics
A.N. Shiryaev
Probability
Second Edition
Springer

"十一五"国家重点图书出版规划项目
科学素养大家谈丛书
科学发现纵横谈
——王梓坤教育随笔
王梓坤 著
大连理工大学出版社
DALIAN UNIVERSITY OF TECHNOLOGY PRESS

成绩构成

- 20% - 30%：平时成绩，含平时作业、大作业、讨论报告等。
- 20% - 30%：期中随堂测验
- 50% - 60%：期末考试
- 5 分：课堂及平时表现加分项
- 考勤：不定时随机点名，如被点到未来上课者每次**扣总评成绩 5 分**
- 期末考试过后，凡与考试成绩有关的邮件,QQ 留言，微信留言均不回复，敬请谅解

答疑时间与课程 QQ 群

- 答疑
  - 时间：每周二中午 12:00-13:30
  - 地点：生医楼 206 办公室
- 课程 QQ 群
  - 理科试验班：814238326
  - 强基拔尖班：964600973

作业提交

- 自由组合, 每两人一组并确定组长由组长提交每次作业
- 提交方式: Gitee 线上提交 (Markdown 文件与导出的 pdf 文件)
- 提交时间: 每周日晚 12 点之前
- 如何提交:
  - 注册 Gitee 账号并更新个人信息为" 学号-姓名":
    强基拔尖班: 强基拔尖班邀请链接注册 Gitee 账号.
    理科试验班: 理科试验班邀请链接注册 Gitee 账号.
  - 苹果系统已预装 Git, Window 系统需下载安装 Git:
    下载地址: https://git-scm.com/download/
  - 设置 Gitee SSH 公钥, 以便推拉作业仓库

Gitee SSH 公钥设置

生成 SSH 公钥

> Windows 用户建议使用 **Windows PowerShell** 或者 **Git Bash**，在 **命令提示符** 下无 `cat` 和 `ls` 命令。

1. 通过命令 `ssh-keygen` 生成 SSH Key：

```
ssh-keygen -t ed25519 -C "Gitee SSH Key"
```

- `-t` key 类型
- `-C` 注释

输出，如：

```
Generating public/private ed25519 key pair.
Enter file in which to save the key (/home/git/.ssh/id_ed25519):
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/git/.ssh/id_ed25519
Your public key has been saved in /home/git/.ssh/id_ed25519.pub
The key fingerprint is:
SHA256:ohDd0OK5WG2dx4gST/j35HjvlJlGHvihyY+Msl6IC8I Gitee SSH Key
The key's randomart image is:
+--[ED25519 256]--+
|     .o          |
|    .+oo         |
|   ...O.o +      |
|    .= * = +.    |
|   .o +..S*. +   |
|. ...o o..+* *   |
|.E. o . ..+.O    |
| . . ... o =.    |
|    ..oo. o.o    |
+----[SHA256]-----+
```

- 中间通过三次**回车键**确定

2. 查看生成的 SSH 公钥和私钥：

```
ls ~/.ssh/
```

输出：

```
id_ed25519  id_ed25519.pub
```

- 私钥文件 `id_ed25519`
- 公钥文件 `id_ed25519.pub`

3. 读取公钥文件 `~/.ssh/id_ed25519.pub`：

```
cat ~/.ssh/id_ed25519.pub
```

输出，如：

```
ssh-ed25519 AAAA***5B Gitee SSH Key
```

复制终端输出的公钥。

设置账户 SSH 公钥

用户可以通过主页右上角 「个人设置」->「安全设置」->「SSH 公钥」->「添加公钥」，添加生成的 public key 添加到当前账户中。

> 需要注意：**添加公钥需要验证用户密码**

通过 `ssh -T` 测试，输出 SSH Key 绑定的**用户名**：

```
$ ssh -T git@gitee.com
Hi USERNAME! You've successfully authenticated, but GITEE.COM does not provide shell acce
```

作业提交: Git 命令首次作业提交

1. Clone 作业仓库
   强基拔尖班: git clone git@gitee.com:xzhangseu/a-prob-2024.git
   理科试验班: git clone git@gitee.com:xzhangseu/b-prob-2024.git
2. 创建创建以个人学号-姓名分支:
   git checkout -b 12345-姓名
3. 将个人新建的分支推送至远程:
   git push --set-upstream origin 12345-姓名
4. 书写作业并提交:
   git add 作业文件名
   git commit -m " 第 k 次作业"
   git push

作业提交：Git 命令后续作业提交

1. 更改本地分支到 master 并拉取作业:
   git checkout master
   git pull
2. 拉取作业后更改至分支"12345-姓名" 完成作业并提交:
   git checkout 12345-姓名
3. 书写作业并提交:
   git add 作业文件名
   git commit -m " 第 k 次作业"
   git push

作业格式要求

- 作业文件命名为: 第 k 次作业
- 作业不能仅有公式, 没有语言表达
- 作业内容: 标题要添加" 第 k 次作业 + 每组两人的学号与姓名", 如下图所示

第k次作业

学号1-姓名1　　学号2-姓名2

作业提交

- 完成作业需要学习
  - Git 知识
  - Gitee 使用
  - Markdown 指南
  - Markdown Preview Enhanced 使用指南

Markdown 文档编辑软件及插件

- Visual Studio Code 插件
  - Markdown Preview Enhanced
  - Better Markdown & Latex Shortcuts
  - Auto-Open Markdown Preview
  - Markdown Preview Github Styling
- Obsidian 插件
  - obsidian-latex-suite
  - obsidian-latex
  - math-booster
  - dataview
  - mathlinks
- logseq

一些期望

- 守时：作业按规定时间提交
- 勇于表现，不做沉默者
- 培养批判性思维
- 相互尊重，教学意见及时与老师沟通
- 注重培养书面语言表达能力
- 实事求是、客观的进行教学评价

学习方法

- 课前预习
- 课上学习
- 课后复习
- 每 2-3 周完整复习前 2-3 周的学习内容

一本书三部电影

约翰·克利斯朵夫

肖申克的救赎

闻香识女人

心灵捕手

什么是概率

- 运气、巧合、随机性、不确定性、风险、时运、机会 - 用法模糊
- 概率普遍存在于科学与日常生活中, 但可能非常反直觉
- 单纯依靠直觉, 会出现预测不准确或面临预测过于自信带来严重风险
- 概率论课程旨在搭建一个有原则的量化不确定性和随机性的逻辑框架, 通过学习强化自身的直觉

几个问题

问题 1.1 假设对于概率论课程成绩有两种不同的方式确定, 你选择哪一种?

- 课程成绩由提交作业的次数来决定, 只要至少提交过 6 次作业就给你 60 分;
- 课程成绩由某随机试验的结果来决定, 但我保证: 你的课程成绩小于 60 分的概率为 0.

几个问题

问题 1.2 某人医院体检报告甲胎蛋白 (AFP) 值偏高, 医生建议进行肝细胞活检, 当问及原因时, 医生说: 我们发现在肝癌患者中, 有 99.9% 的患者其甲胎蛋白数值都偏高. 医生给出活检的理由是否充分?

几个问题

问题 1.3 三门问题: 在某电视节目中, 主持人蒙提霍尔会选择一名观众参与一个游戏.

- 三扇门: 一扇门后是一辆汽车, 另外两扇门后各有一只山羊;
- 参与者可从三扇门中任选一扇, 并可以获得该扇门后的奖品;
- 在揭晓参与者所选门后的奖品之前, 主持人会打开其中一个未被选中的一扇门, 并揭晓门后的奖品是山羊.
- 参与者可以坚持原来的选择或换成剩下的那扇门;
- 问题: 参与者是否要改变自己的选择?
- 如果将上面的问题推广至 n 扇门, 其中 k 扇门后是汽车, 那又做何选择? (此处, $n \geq 3, k \leq n-2$)

注 1.1 主持人蒙提霍尔不是在参与者未选择的两扇门中随机选择, 而是总是选取一个后面是羊的门展示.

几个问题

问题 1.4 假设有这样一个赌博游戏: 赌注是 3000 元, 赌博者可以选择掷几个骰子, 只要掷出的骰子有 6 点且没有 1 点, 则赌博者即可获得 10000 元.

- 赌博者应选几个骰子掷胜率更高?
- 如果换做是你, 你是选择玩这个赌博游戏, 还是不玩?

几个问题

问题 1.5 假设某合法赌场内有胜率 1/2 与胜率 1/20 的两类赌局. 现迫于需要, 24 小时之内要在此合法赌场将 100 万元变成 200 万元.

- 一口气将 100 万下注在胜率为 1/2 的赌局上;
- 将 100 万分成 10 份, 每份 10 万下注在胜率为 1/20 的赌局上, 赌 10 次.

问题 1.6 赌徒破产问题: 假设去某赌徒去某赌场玩轮盘赌游戏, 其获胜的概率是 0.48. 该赌徒现有资产 1000 元, 并决定一直玩下去直到资产翻倍或破产, 试问下面哪一种下注方式对赌徒更为有利

- 每次下注 1 元;
- 每次下注 100 元;
- 每次下注 200 元;
- 每次下注 1000 元.

几个问题

问题 1.7 De Mere 分赌注问题

- 甲乙两人各下赌注 m 元;
- 先胜三局者取得全部赌金;
- 甲胜一局而乙胜零局时赌博被迫中止;
- 问赌注如何分?

其他问题

- 买彩票问题;
- 抽签公平性问题;
- 生日相同的概率问题;
- 小概率事件终将发生;
- “一而再再而三”、“智者千虑, 必有一失”等谚语的概率理解;
- “狼来了”故事的概率解释.

概率论的应用领域

- 统计学：概率论是统计学的基础和语言
- 物理学："上帝不掷骰子"，但量子物理学、统计力学均建立在概率基础之上
- 生物学：遗传学、基因序列、群体的生灭、迁移等随机模型
- 计算机科学：随机算法，算法性能、机器学习、人工智能
- 气象学：天气预报
- 赌博：概率论早期研究旨在解决赌博和机会游戏的问题
- 金融学：概率是计量金融的核心
- 政治科学：预测选举结果，民调等
- 保险精算，风险管理理与控制：保费核定及风险量化等
- 医学：随机临床试验
- 化学化工：化学反应动力学反应的时变率及影响因素问题等
- 天文学：湍流理论及星云密度起伏，辐射传递等
- 运筹与控制，排队论
- 生活：生活是不确定的，在不确定中利用概率做决策。
- ……

概率论课程难不难？

- 概率论本质是分析类课程，研究内容与数学分析有很好的类比性
- 数学分析是确定性分析理论，概率论是随机性分析理论
- 难，也不难？
  - 概率论为解决随机问题提供了理论方法，同时也会产生陷阱和悖论
  - 即使是莱布尼茨和牛顿，也无法避免一些概率论上的基本错误
- 如何避免潜在错误
  - 数值模拟：利用 R 语言运行模拟，看看到底谁是对的
  - 完整性检查：使用不同的方法来解决同样的问题，或者检查所得到的答案在简单和极端的情况下是否都是合理的

概率论课程主要内容

- 概率简介 (第 1,2,6,5 章)
  - 事件及其运算, 概率定义及性质
  - 条件概率、事件独立性
  - 全概率公式、贝叶斯公式
- 随机变量 (第 3,4,8,10,9 章)
  - 随机变量定义、性质、结构、研究方法（分布、分布函数）, 分类
  - 多维随机变量定义、研究方法（联合分布、联合分布函数）, 分类
  - 常见的随机变量具体分布类型：离散型、连续型
  - 条件分布函数、随机变量的独立性
  - 分布函数形式下的全概率公式、贝叶斯公式

概率论课程主要内容

- 随机变量的数字特征：期望、方差、协方差 (第 7,11,12 章)
  - 随机变量的积分的定义及性质
  - 期望的定义以及与随机变量积分之间的关系
  - 方差、协方差的定义及性质、随机变量的矩、特征函数、矩母函数
  - 条件数学期望的定义及性质、条件方差
- 随机变量序列的极限 (第 13,14,15 章)
  - 随机变量序列的几种收敛性及相关关系：几乎必然、依概率, 依分布等
  - 弱大数定律与强大数定律
  - 中心极限定理

概率雏形

- 1470 年, 拉丁文的诗书《De Vetula》出版. 上面有首诗记录了 3 个骰子点数和的排列组合结果.

Qvot Punctaturas, et qvot Cadentias habeat qvilibet numerorũ compositorum.

3	18	Punctatura	1	Cadentia	1
4	17	Punctatura	1	Cadentiæ	3
5	16	Punctaturæ	2	Cadentiæ	6
6	15	Punctaturæ	3	Cadentiæ	10
7	14	Punctaturæ	4	Cadentiæ	15
8	13	Punctaturæ	5	Cadentiæ	21
9	12	Punctaturæ	6	Cadentiæ	25
10	11	Punctaturæ	6	Cadentiæ	27

Latin poem De Vetula (~1250?)
by Ovid pseunonim (Richard de Fournival?)
Taken from Kendall (1956)

- 左图为原始印刷, 右图为阿拉伯数字版本. 可以看出, 三个骰子点数相加等于 10 或者 11, 有 6 种组合情况, 但是排列情况有 27 种.

概率雏形

- 140 年后, 即 1610 年左右, 资助 Galileo(伽利略) 的 Tuscany(托斯卡纳) 大公请教了 Galileo 3 个骰子和的问题.
- Tuscany 大公在赌博时发现三个骰子点数和为 10 比点数和为 9 出现的更频繁一些. 他认为有 6 种方式得到 10, 分别为

$$(6,3,1),(6,2,2),(5,4,1),(5,3,2),(4,4,2),(4,3,3),$$

  同样有 6 种方式得到 9, 分别是

$$(6,2,1),(5,3,1),(5,2,2),(4,4,1),(4,3,2),(3,3,3).$$

  因此,9 和 10 出现的频率不应该有差异.
- Galileo 指出了大公的错误, 三个骰子是不同的,$(6,3,1)$ 和 $(6,1,3)$ 是不同的两种情况。可惜伽利略那个时候只考虑了频次, 还没有形成概率的思想。

概率论思想形成

- 1564 年左右, 意大利博学家 Cardano(卡尔达诺) 写了本书《Liber de Ludo Aleae ("Book on Games of Chance")》.
- Cardano 是第一个系统的推算概率的人, 三个骰子和的问题在《Liber de Ludo Aleae 》的第 13 章有提到并且解决.

Consensus sortis in duabus Aleis.

2	12	1	3	11	2	4	10	3. Æqual.
5	9	4	6	8	5	7	8	18. Ad Frit:

Consensus sortis in tribus Aleis tum Frit.

Sortis			Fritilli.	
3	18	1	3	115
4	17	3	4	125
5	16	6	5	126
6	15	10	6	133
7	14	15	7	33
8	13	21	8	36
9	12	25	9	37
10	11	27	10	36
			11	38
			12	26

Circuitus 216.
Æqualitas 108.

- 《Liber de Ludo Aleae 》直到一个世纪后的 1663 年才出版, 伽利略 1642 年去世, 所以也没有机会看到.
- Cardano 在第 14 章中明确定义了“比例”, 若赌局中有利的所有可能的数目为 a, 不利的数目为 b, 则应该根据 a/b 的结果来下注.

French Society in the 1650's

- 赌博是一种流行时尚
- 法律没有禁止
- 当游戏越来越复杂，利益越来越大时，需要数学方法来计算获胜可能性

概率论的诞生

- Cardano 虽说是概率的创派始祖, 但概率论真正变成一门学科的标志性事件则是 1654 年 Pascal(帕斯卡) 和 Fermat(费马) 关于两个赌博问题的通信.
- Pascal(帕斯卡) 就是法国物理学家和数学家帕斯卡, 命名压强单位的物理学家 Pascal.
- Fermat(费马) 就是提出“费马大定理”并且困扰数学家 300 年之久的费马.

两个赌博问题

问题 1.8 De Mere(德梅尔): 事件可能性大小

- 一颗骰子投 4 次至少得到一个六点,
- 两颗骰子投 24 次至少得到一个双六.

问题 1.9 De Mere 分赌注问题

- 甲乙两人各下赌注 m 元;
- 甲乙两人获胜的概率一样, 均为 $\frac{1}{2}$;
- 先胜 t 局者取得全部赌金;
- 甲胜 a 局而乙胜 b 局时赌博被迫中止;
- 问赌注如何分?

Chevalier De Mere

骰子点数问题

- 据经验所知, 一颗骰子投 4 次至少得到一个六点的概率大于 $\frac{1}{2}$;
- 两颗骰子掷一次的结果 6 倍于一颗骰子掷一次的结果, 则两颗骰子投 24 次至少得到一个双六的概率也应大于 $\frac{1}{2}$;
- 但赌场经验却并非如此.
- 考虑成功概率为 p 的随机试验, 连续做 n 次, 我们来确定至少一次成功的概率为 q_0 对应的 n:

$$1-(1-p)^n = q_0 \Rightarrow n = \frac{\ln(1-q_0)}{\ln(1-p)} = \frac{-\ln(1-q_0)}{p+\frac{p^2}{2}+\cdots} \approx \frac{-\ln(1-q_0)}{p};$$

- 两个不同成功概率的随机试验, 连续做 n 次, 要保其至少成功一次的概率相同所需要的 n 近似反比于单次试验成功的概率.

分赌注问题历史

- 分赌注问题最早是由意大利数学家 Paccioli(帕乔利) 在 1494 年提出, 当时 Paccioli 给出的答案是按 $a:b$ 进行分配.
- 50 年后, Tartaglia(塔尔塔利亚) 提出反对意见: 根据这条规则, 如果游戏在一轮之后停止, 那么其中一名玩家就会得到全部赌注.
- Tartaglia 试图修改帕乔利的规则, 以便将上述情况考虑进去, 但最后他怀疑这个问题可能根本没有确定的答案.
- 1537 年, Cardano 也曾经思考过这个问题.
  - Cardano 给出了一个公式 $f(n)=1+\ldots+n$, 甲还剩 $s-a$ 局就可以获胜, 乙还剩 $s-b$ 局可以获胜, 两者的分金比率应该为 $f(s-b):f(s-a)=(s-b)(s-b+1):(s-a)(s-a+1)$.
  - 虽然 Cardano 答案错误, 但 Cardano 已经开始考虑用未来剩余的赌局来决策, 而不是局限于已发生的事件.
- 直到 1654 年, De Mere 向 Pascal 请教, Pascal 和 Fermat 进行书信交流, 并且用不同的解法给出了这个问题的正确答案.
- 后来, 荷兰数学家 Huygens 也参与了 Pascal 和 Fermat 的讨论, 并在 1657 年出版书籍《De ratiociniis in ludo aleae ("On Reasoning in Games of Chance")》.

Blaise Pascal (布莱士·帕斯卡)

- 16 岁时发现著名的帕斯卡六边形定理：内接于一个二次曲线的六边形的三双对边的交点共线;
- 17 岁时写成《圆锥曲线论》，是研究德札尔格射影几何工作心得的论文;
- 设计并制作了一台能自动进位的加减法计算装置，被称为是世界上第一台数字计算器;
- 以积分学的原理解决了摆线问题;
- 与费马相互通信探求赌注分配问题，形成概率论中一个重要的基本概念—数学期望.

Pierre de Fermat (皮耶·德·费玛)

- 独立于笛卡儿发现了解析几何的基本原理;
- 求切线、求极大值和极小值以及定积分方法;
- 费马和帕斯卡在相互通信探求赌注分配问题, 形成概率论中一个重要的基本概念—数学期望;
- 数论中的费马大小定理;
- 光学的最小作用原理.

Christiaan Huygens (克里斯蒂安・惠更斯)

- Christiaan Huygens 在 1657 年写了世界上第一本关于概率论的著作 De ratiociniis in ludo aleae (“On Reasoning in Games of Chance”), 中文译名“论赌博中的计算”.
- 概率论的基本概念、第一次在书中提出期望.

分赌注问题:Fermat 解法

- 显然甲乙距离赢得赌局分别还差 $r = t - a$ 局和 $s := t - b$ 局. 因此, 若赌局不终止, 赌局肯定会在接下来的 $r + s - 1$ 轮内结束.
- 考虑所有 $r + s - 1$ 场赌局的结果, 计算在所有可能结果中, 甲乙最终获胜的概率. 这样一来, 整个问题就简化为一个古典概型的概率计算问题.
- 以上述计算出来的获胜概率作为分配赌注的比例.
- 若 $t = 6, a = 5, b = 3$, 则最多还需要赌 3 场就可以分出胜负. 后续所有可能的结果分别为

$$\{\text{甲甲甲}, \text{甲甲乙}, \text{甲乙甲}, \text{甲乙乙}, \text{乙甲甲}, \text{乙甲乙}, \text{乙乙甲}, \text{乙乙乙}\}.$$

- 所有的 7 种结果中, 乙只有一个结果获胜, 也就是连胜三局, 故赌注分配比例应为 $7:1$.

分赌注问题:Pascal 解法

- 杨辉三角或 Pascal 三角法
  - 后续 $r+s-1$ 场赌局的所有可能结果中:
    使得甲获胜的结果个数为: $n_{甲} = \sum_{k=r}^{r+s-1} C_{r+s-1}^{k}$
    使得乙获胜的结果个数为: $n_{乙} = \sum_{k=s}^{r+s-1} C_{r+s-1}^{k}$
  - 若 $t=6, a=5, b=3$ 即 $r=1, s=3$, 则

$$n_{甲} = C_3^1 + C_3^2 + C_3^3 = 7, \quad n_{乙} = C_3^3 = 1$$

分赌注问题:Pascal 解法

- 递推法
  - 假设甲和乙赌注共 64 个金币, 考虑 $t=6, a=5, b=4$ 的情况:
    如果下一局甲胜, 则甲得到全部 64 枚金币
    如果下一局甲输, 则比分变为 $5:5$, 这时候是平局, 甲乙平分赌注.
  - 不管哪种情况, 甲至少会拿 32 个金币, 剩余的 32 个金币, 可能归甲也可能归乙, 平分. 这样甲应该拿 $32+(64-32)/2=48$ 个金币, 乙则拿 16 个金币.
  - 考虑 $t=6, a=5, b=3$ 的情况:
    - 如果下一局甲胜, 则拿到所有赌注;
    - 如果甲败, 则问题转化为 $t=6, a=5, b=4$ 的情况;
    - 由此可见, 甲至少可拿 48 个硬币, 剩余金币平分;
    - 因此, 甲最终可拿 $48+(64-48)/2=56$ 个金币, 乙拿 8 个金币;
    - 和 Fermat 的 $7:1$ 结果相同.
  - 在递推解法里,Pascal 用到了期望的思想, 尤其是 $48+(64-48)/2$ 这个式子.
    - 下一局甲胜得到 $x=64$ 个金币, 甲输得到 $y=48$ 个金币;
    - 按 Pascal 的式子, 甲拿到的金币为:
      $y+(x-y)/2=(x+y)/2=x*0.5+y*0.5$, 这其实就是条件期望公式.

分赌注问题:Pascal 递推法的一般解法

- 赌博中断时, 甲乙距离赢得赌局分别还差 $r = t - a$ 局和 $s := t - b$ 局;
- 假设中断后赌局继续进行, 甲最终获胜的概率记为 $e(r, s)$, 则有边界条件:

$$e(0, s) = 1, \quad e(r, 0) = 0, \quad e(r, r) = 1/2$$

且有如下递推式

$$e(r, s) = [e(r-1, s) + e(r, s-1)]/2.$$

- 上式用到了全概率公式, 是一个偏差分方程. Pascal 借助算术三角阵, 利用数学归纳法得到

$$e(r, s) = \frac{D_{r+s-1,s-1}}{D_{r+s-1}} (r + s = 2, 3, \cdots)$$

- 这里分子表示算术三角形中第 $r + s - 1$ 行最后 s 项之和, 而分母为第 $r + s - 1$ 行的所有数字之和, 故为 2^{r+s-1}.

Jacob Bernoulli(雅各布·伯努利)

- 概率论与变分法先驱之一;
- 第一次使用“积分”一词;
- 首次给出直角坐标和极坐标下的曲率半径公式;
- 1713 年出版《Ars Conjectandi》(可译为《猜度术》): 提出了概率论中的“伯努利大数定律”-大数定律的最早形式.

Abraham de Moivre(亚伯拉罕·棣莫弗)

- 棣莫弗公式:

$$(\cos x + i\sin x)^n = \cos(nx) + i\sin(nx)$$

- 二项分布的正态逼近: 棣莫弗-拉普拉斯中心极限定理;
- 利用概率生成函数计算均匀分布随机变量和的分布;
- 寿险保险数学: 年金的分析与计算.
- 开创了概率论的现代方法: 1718 年发表了《The Doctrine of Chance》. 在此书中统计独立性的定义首次出现. 该书在 1738 年与 1756 年出了扩展版, 生日问题出现在 1738 年的版本中, 赌徒破产问题出现在 1756 年的版本中.

Pierre-Simon Laplace(皮埃尔-西蒙・拉普拉斯)

- Pierre-Simon Laplace 在他 1812 年的著作《Théorie Analytique des Probabilitiés》(可译为《概率论的解析理论》)中介绍了概率的数学理论及其科学应用;
- Laplace 只考虑了古典概型, 对一般的概率及其应用没有介绍;
- 到 1850 年, 许多数学家发现古典概型对一般场合不合理, 开始尝试重新定义概率.

概率的公理化体系

- 从 17 世纪到 19 世纪长达 300 多年的时间里, 以 Laplace(1749 一 1827), De Moivre(1687-1754), J.Bernouil, D.Bernouli(1700–1782). Pascal、Fermat 等为代表的优秀学者从各自不同的角度对概率论的许多方面进行了深入细致的研究, 取得了大量成果, 可是都无法将概率论纳入到严格的公理化体系中. 而概率论作为数学的重要分支, 其逻辑的严密性和完整性始终受到质疑.
- 数学大师 Hilbert(1862—1943) 在 1900 年提出的 23 个著名问题中, 第 6 个问题是关于“物理定律的公理化”, 其中就涉及概率论的公理化问题.
- 激发了以 Borel(1871 一 1956) 和 Lebesgue(1875—1941) 等为代表的一批学者的研究热情. 他们在测度理论 (Measure Theory) 方面得到的成果为后续研究指明了道路.

Andrey Kolmogorov(安德列・柯尔莫哥洛夫)

- Andrey Kolmogorov 第一个在他 1933 年的著作《Grundbegriffe der Wahrscheinlichkeitsrechnun》(《Foundations of the Theory of Probability》) 中严格的定义了概率;
- 类似于 Euclid 基于公理体系建立几何, 他从基本公理建立了概率理论, 从而使概率论称为一门严谨的数学分支;
- 概率理论的现代研究和测度论非常紧密的结合在一起.

1903.4.25-1987.10.20

Kolmogorov 成就

- 20 世纪最伟大的数学家之一, 在概率论、实分析、拓扑学、泛函分析、微分方程、生物数学、等很多领域都有着开创性的贡献.
- 1933 年, 出版《概率论基础》, 首次将概率论建立在严格的公理基础上, 解决了希尔伯特第 6 问题的概率部分。
- 线性拓扑空间理论的创始人之一, 和美国著名数学家亚历山大（J.W.Alexander）同时独立引入上同调群的概念;
- 40 年代, 在遗传学、弹道学、气象学、金属结晶学等有重要贡献;
- 50 年代, 经典力学、遍历理论、函数论、信息论、算法理论等。
- 柯尔莫哥洛夫的研究几乎遍及数论之外的一切数学领域。1963 年, 在第比利斯召开的概率统计会议上, 美国统计学家沃尔夫维茨 (J.Wolfowitz) 说: “我来苏联的一个特别的目的是确定柯尔莫哥洛夫到底是一个人呢, 还是一个研究机构.”

Kolmogorov 轶事

- 历史学要证实自己的观点需要几个甚至几十个正确证明才行, 数学只需要一个证明就行, 于是转行数学.
- 20 年代的莫斯科大学, 学生被要求在十四个不同的数学分支参加十四门考试; 但考试可用相应领域的独立研究代替. Kolmogorov 从来没有参加考试, 他写了十四个不同方向的有新意的文章。
- Kolmogorov 总是以感激的口气提到斯大林: “首先, 他在战争年代为每一位院士提供了一床毛毯; 第二, 原谅了我在科学院的那次打架。” Kolmogorov 一次在选举会上打了 Luzin（卢津）一个耳光, 他说: “（打架）那是我们常用的方式。” Luzin 在实变函数方面有着很重要的贡献, 但是以打架而论, 远非 Kolmogorov 的对手.

第二讲：事件及其运算关系

张鑫

Email: xzhangseu@seu.edu.cn

东南大学　数学学院

概率：(随机) 事件发生的可能性大小的度量

- 随机现象：一定条件下，并不总是出现相同结果的现象，比如抛一枚硬币，掷一颗骰子等；
- 随机试验：随机现象的实现和对它某个特征的观测；常简称试验
  - 随机试验中要求试验的结果至少 2 个；
  - 每次试验或观测得到其中的一个结果，在试验或观测之前不能预知是哪个结果发生；
  - 一般还要求试验在相同条件下能够重复.
  - 比如：观测把硬币抛 4 次后正面向上的次数；观测某地的温度变化；某电话总机单位时间内转接的电话次数.

样本点与样本空间

- 样本点 (Sample Point): 随机试验的每个可能结果称为一个样本点，常用ω表示.
- 样本空间 (Sample Space): 全体样本点构成的集合，常用Ω表示.
  - 一个试验的样本空间可能是有限的，可数无限的或不可数的;
  - 当样本空间为有限的时，我们可以将样本点可视化为圆点，做一次试验就等同于随机选取一个圆点;
  - 如果所有的圆点都是相同的，那么所有的圆点被抽中的可能性就相同，即后面将要介绍的古典概型.

事件

- 随机事件: 简称事件 (Event), 由部分样本点组成的试验结果
  - 通常指在随机试验中我们所关心的某一部分可能结果
  - 如果按照样本点可视化为圆点，则事件就是一些圆点的集合
- 基本事件: 仅含有一个样本点的随机事件称为基本事件, 它犹如分子中的原子, 在化学反应中不能再分, 所以有“基本”两字.
- 事件发生
  - 事件 A 发生: 事件 A 包含的样本点在随机试验结果中出现即 $\omega \in A$;
  - 通常把事件 A 所包含的样本点构成的集合与事件 A 视为等同, 也记为 A,
- 两个特殊的事件:
  - 必然事件(Ω): 在随机试验中一定会发生的事件;
  - 不可能事件($\emptyset$): 在随机试验中不可能发生的事件;

硬币抛 3 次的随机试验

- 样本点:正正正、正正反、正反正、反正正、正反反、反正反、反反正、反反反，共 8 种可能结果;
- 基本事件: {正正正}、{正正反}、{正反正}、{反正正}、{正反反}、{反正反}、{反反正}、{反反反}，共 8 个基本事件;
- 样本空间 Ω:

$$\Omega := \{正正正, 正正反, 正反正, 反正正, 正反反, 反正反, 反反正, 反反反\}$$

- 随机事件:

$$\begin{aligned} A &:= \{至少出现两次正面\} \\ &= \{正正正, 正正反, 正反正, 反正正\} \\ B &:= \{三次均为正面\} = \{正正正\} \\ \cdots \end{aligned}$$

样本空间示例

- 掷一枚骰子, 观察出现的点数. 则 $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- 考察某一地区的年降雨量, 则 $\Omega = \{x | 0 \leq x < T\}$, 这里 T 表示某个常数, 表示降雨量不会超过 T.
- 将一枚硬币抛两次, 考察正反面出现的情况, 则

$$\Omega = \{\text{正正, 正反, 反正, 反反}\}.$$

- 将一枚硬币抛两次, 考察正面出现的次数, 则 $\Omega = \{0, 1, 2\}$.
- 先后掷两次骰子, 考察两次骰子的点数, 则

$$\Omega = \{(a_1, a_2), a_i \in \{1, 2, \cdots, 6\}, i = 1, 2\}.$$

事件间的并:$A \cup B$

定义 2.1 (事件 A 与事件 B 的并) 记为 $A \cup B$, 其含义为: 事件 A 和事件 B 至少有一个发生, 即由事件 A 与 B 中所有的样本点组成的新事件.用数学符号来描述即为:

$$A \cup B := \{\omega \in \Omega : \omega \in A, \text{ 或} \omega \in B\}$$

例 2.1 考虑掷一枚骰子, 观察出现的点数. 记事件 A 为掷出的点数为奇数, 事件 B 为掷出的点数不超过 3, 则

$$A = \{1, 3, 5\}, \quad B = \{1, 2, 3\}, \quad A \cup B = \{1, 2, 3, 5\}$$

事件间的交:$A \cap B$

定义 2.2(事件 A 与事件 B 的交) 记为 $A \cap B$ 或 AB, 其含义为: 事件 A 和事件 B 同时发生, 即由事件 A 与 B 中公共的样本点组成的新事件.用数学符号来描述即为:

$$A \cap B := \{\omega \in \Omega : \omega \in A \text{ 且} \omega \in B\}$$

例 2.2 考虑掷一枚骰子, 观察出现的点数. 记事件 A 为掷出的点数为奇数, 事件 B 为掷出的点数不超过 3, 则

$$A = \{1, 3, 5\}, \quad B = \{1, 2, 3\}, \quad A \cap B = \{1, 3\}$$

事件的差:$A-B$

定义 2.3(事件 A 与 B 的差) 记为 $A-B$ 或 $A\backslash B$, 其含义为: 事件 A 发生且事件 B 不发生, 即由在 A 中而不在 B 中的样本点组成的新事件. 用数学符号来描述即为:

$$A-B:=\{\omega\in\Omega:\omega\in A\text{ 且}\omega\notin B\}$$

例 2.3 考虑掷一枚骰子, 观察出现的点数. 记事件 A 为掷出的点数为奇数, 事件 B 为掷出的点数不超过 3, 则

$$A=\{1,3,5\},\quad B=\{1,2,3\},\quad A-B=\{5\}$$

对立事件:$\overline{A}$

定义 2.4 (事件 A 的对立事件) 记为 $\overline{A}$ 或 A^c, 其含义为: 事件 A 不发生, 即由在 Ω 中而不在 A 中的样本点组成的新事件.用数学符号来描述即为:

$$\overline{A} := \{\omega \in \Omega : \omega \notin A\} = \Omega - A$$

例 2.4 考虑掷一枚骰子, 观察出现的点数. 记事件 A 为掷出的点数为奇数, 则

$$A = \{1, 3, 5\}, \quad \overline{A} = \{2, 4, 6\}$$

对称差事件:$A\Delta B$

定义 2.5 (事件 A 与 B 的对称差事件) 记为 $A\Delta B$, 其含义为: 事件 A 与 B 中恰有一个事件发生.用数学符号来描述即为:

$$A\Delta B := \{\omega \in \Omega : \omega \in A \cup B \text{ 且} \omega \notin A \cap B\} = (A - B) \cup (B - A)$$

包含关系: $A \subset B$

定义 2.6 (包含关系) 称事件 A 包含于事件 B, 或事件 B 包含事件 A(记为$A \subset B$), 如果事件 A 发生必然导致事件 B 发生, 即属于 A 的样本点必然属于 B.

例 2.5 考虑掷一枚骰子, 观察出现的点数. 记事件 A 为掷出的点数为 2 或 4, 事件 B 为掷出的点数为偶数. 则显然有事件 A 发生必然导致事件 B 发生, 同时易知

$$A = \{2, 4\} \subset \{2, 4, 6\} = B$$

相等关系: $A = B$

定义 2.7 (相等关系) 称事件 A 与事件 B 相等 (记为$A = B$), 如果 $A \subset B$ 且 $B \subset A$, 即事情件 A, B 有相同的样本点.

例 2.6 考虑掷一枚骰子, 观察出现的点数. 记事件 A 为掷出的点数为 $2, 4, 6$, 事件 B 为掷出的点数为偶数. 则显然有

$$A \subset B, B \subset A,$$

同时易有

$$A = \{2, 4, 6\} = \{2, 4, 6\} = B$$

互不相容关系: $A \cap B = \emptyset$

定义 2.8 (互不相容关系) 称事件 A 与事件 B 互不相容, 如果事件 A, B 不能同时发生, 即事情件 A, B 没有相同的样本点.

例 2.7 考虑掷一枚骰子, 观察出现的点数. 记事件 A 为掷出的点数为奇数, 事件 B 为掷出的点数为偶数. 则显然有事件 A, B 互不相容, 并且

$$A \cap B = \{1, 3, 5\} \cap \{2, 4, 6\} = \emptyset$$

对立事件与互不相容事件

- $\overline{\overline{A}} = A, A \cap \overline{A} = \emptyset, A \cup \overline{A} = \Omega$;
- A, B 互为对立事件的充要条件为 $A \cap B = \emptyset, A \cup B = \Omega$;
- 对立事件一定是互不相容事件, 但互不相容事件不一定是对立事件;
- $A - B = A \cap \overline{B} = A\overline{B}$

多个事件的运算

- 有限并事件: $\bigcup_{k=1}^{n} A_k$
- 有限交事件: $\bigcap_{k=1}^{n} A_k$
- 可列并事件: $\bigcup_{k=1}^{\infty} A_k$
- 可列交事件: $\bigcap_{k=1}^{\infty} A_k$

事件运算的性质

(1) 交换律: $A \cup B = B \cup A, AB = BA$;

(2) 结合律: $A \cup (B \cup C) = (A \cup B) \cup C, (AB)C = A(BC)$;

(3) 分配律: $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$,
$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$;

(4) De Morgan 对偶法则:

$$\overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \overline{A}_i$$

$$\overline{\bigcap_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \overline{A}_i$$

事件列的上下极限

定义 2.9 令 $A_n \subset \Omega, n \geq 1$ 为一列集合 (事件), 记

$$\inf_{k\geq n} A_k := \cap_{k=n}^{\infty} A_k, \quad \sup_{k\geq n} A_k := \cup_{k=n}^{\infty} A_k$$

则事件列 A_n 的上下极限定义如下:

$$\liminf_{n\to\infty} A_n := \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k,$$
$$\limsup_{n\to\infty} A_n := \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k.$$

若

$$\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n = A$$

则称 A 为事件列 $\{A_n, n \geq 1\}$ 的极限, 并记作 $\lim_{n\to\infty} A_n = A$.

单调事件列的极限

定义 2.10 事件列 $\{A_n\}$ 称为单调不减的，如果对任意的 n 均有 $A_n \subset A_{n+1}$；事件列 $\{A_n\}$ 称为单调不增的，如果对任意的 n 均有 $A_{n+1} \subset A_n$；

定理 2.1 假设 $\{A_n\}$ 是单调事件列，则

1. 如果 $\{A_n\}$ 是单调不减事件列，则 $\lim_{n\to\infty} A_n = \cup_{n=1}^{\infty} A_n$
2. 如果 $\{A_n\}$ 是单调不增事件列，则 $\lim_{n\to\infty} A_n = \cap_{n=1}^{\infty} A_n$

示性函数 (indicator function)

定义 2.11 设 $A \subset \Omega$, 我们定义事件 (集合)A 的示性函数如下

$$1_A(\omega) = \begin{cases} 1, & \text{如果}\omega \in A; \\ 0, & \text{如果}\omega \notin A. \end{cases}$$

注 2.1

- $A = \{\omega : 1_A(\omega) = 1\}$;
- $1_A \leq 1_B \Leftrightarrow A \subset B$;
- $1_{A\cap B} = 1_A \cdot 1_B$;
- $1_{A\cup B} = 1_A + 1_B - 1_{A\cap B}$;

事件列上下极限的性质

引理 2.1 令 $\{A_n\}$ 为 Ω 的任一子集列，则

$$\begin{aligned}
\limsup_{n\to\infty} A_n &= \{\omega : \sum_{n=1}^{\infty} 1_{A_n}(\omega) = \infty\} \\
&= \{\omega : \omega\text{属于无穷多个}A_n\} \\
\liminf_{n\to\infty} A_n &= \{\omega : \sum_{n=1}^{\infty} 1_{\bar{A}_n}(\omega) < \infty\} \\
&= \{\omega : \text{除去有限多个}A_n\text{外}, \omega\text{属于余下所有的}A_n\}
\end{aligned}$$

证明： 由 $\omega \in \limsup_{n\to\infty} A_n = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k$ 知，对任意的 n，存在 $k_n \geq n$ 使得 $\omega \in A_{k_n}$，从而 $\sum_{n=1}^{\infty} 1_{A_n} \geq \sum_{n=1}^{\infty} 1_{A_{k_n}} = \infty$.

反之，若 $\omega \in \{\omega : \sum_{n=1}^{\infty} 1_{A_n}(\omega) = \infty\}$，则存在 $k_n \to \infty$ 使得 $\omega \in A_{k_n}$，从而对任意的 n，$\omega \in \cup_{k=n}^{\infty} A_k$，即

$$\omega \in \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k = \limsup_{n\to\infty} A_n.$$

事件运算表

符　号	集合论意义	概率论意义
Ω	全集或空间	样本空间、必然事件
$\varnothing$	空集	不可能事件
ω	元素	样本点、基本事件
A	可测子集	随机事件
$\omega \in A$	ω 是 A 的元素	试验结果 ω 属于 A, 事件 A 发生
$A \subset B$	A 包含在 B 中	若 A 发生, 则 B 一定发生; 事件 A 蕴涵事件 B
$A = B$	A 与 B 相等	A 与 B 为同一事件; 它们同时发生或同时不发生
$A \cap B$ 或 AB	交集	表示 A 与 B 同时发生
$AB = \varnothing$	不相交	A 与 B 不相容 (互斥), 它们不可能同时发生
$A \cup B$	并集	表示 A 或 B 发生, A 与 B 中至少有一个发生
A^c	余集	对立事件; A^c 发生表示 A 不发生
$A - B$ 或 AB^c	差集	表示 A 发生, 而 B 不发生
$A \triangle B$	对称差	表示 A 与 B 中恰有一个发生
$\limsup\limits_{n\to\infty} A_n = \bigcap\limits_{k=1}^{\infty} \bigcup\limits_{n=k}^{\infty} A_n$	上极限集合	表示事件序列 $\{A_n\}$ 中有无穷多个事件发生
$\liminf\limits_{n\to\infty} A_n = \bigcup\limits_{k=1}^{\infty} \bigcap\limits_{n=k}^{\infty} A_n$	下极限集合	表示事件序列 $\{A_n\}$ 中至多有有限个事件不发生

例 2.8 设 A, B, C 为某个随机现象的三个事件，试用事件的交，差，并，对立事件表示下列事件

- 事件 A, B 发生而 C 不发生:$AB\overline{C}$;
- 事件 A, B, C 中至少有一个发生:$A \cup B \cup C$;
- 事件 A, B, C 中至少有两个发生: $AB \cup AC \cup BC$;
- 事件 A, B, C 中恰好有两个发生: $AB\overline{C} \cup A\overline{B}C \cup \overline{A}BC$;
- 事件 A, B, C 同时发生: ABC;
- 事件 A, B, C 都不发生: $\overline{A} \cap \overline{B} \cap \overline{C}$;
- 事件 A, B, C 不全发生: $\overline{A} \cup \overline{B} \cup \overline{C}$;

例 2.9 (掷硬币) 掷一枚硬币 10 次, 正面用 H 表示反面用 T 表示.

- 一个可能的结果为 HHHTHHTTHT, 且样本空间就是所有长度为 10 的由 H 和 T 组成的字符串的集合
- 若令 H 为 1，T 为 0 , 则一个结果可用数列 $(s_1, s_2, \cdots, s_{10})$ 表示, 其中 $s_j \in \{0, 1\}$, 样本空间即为这些数列的集合
- 下面我们看几个事件:
  - A_i: 第 i 次掷硬币结果为 H，则

$$A_1 = \{(1, s_2, \cdots, s_{10}) : s_j \in \{1, 0\} \text{ 且} 2 \leqslant j \leqslant 10\}.$$

  - B: 至少有一次掷硬币结果为 H, 则$B = \bigcup_{j=1}^{10} A_j$.
  - C: 所有掷硬币结果都为 H, 则$C = \bigcap_{j=1}^{10} A_j$.
  - D: 至少有两个连续的 H, 则$D = \bigcup_{j=1}^{9} (A_j \cap A_{j+1})$.

第三讲：古典概型与几何概型

张鑫

Email: xzhangseu@seu.edu.cn

东南大学　数学学院

古典概型

定义 2.12 (古典概型:) 有两个条件,

(1) (有限性) 试验结果只有有限个 (记为 n);

(2) (等可能性) 每个基本事件发生的可能性相同.

对于古典概型: $\Omega = \{\omega_1, \cdots, \omega_n\}, \mathcal{F} = \{A : A \subset \Omega\}$, 事件 $A \in \mathcal{F}$ 的概率定义为

$$P(A) = \frac{|A|}{|\Omega|}, \ \forall A \in \mathcal{F}$$

其中, $|B|$ 表示事件 B 所包含的基本事件的个数.

古典概型的两个重要方面

- 确定样本空间与样本点

  例 2.10 掷两个均匀硬币，出现正、反两面的概率是多少？

  - 数学家 d'Alembert(达朗贝尔) 认为掷两个均匀硬币，出现的结果无外乎" 正正"," 反反"," 正反"，因此概率应是 $\frac{1}{3}$;
  - 事实上，该试验样本空间为

$$\Omega = \{(\text{正}, \text{正}), (\text{正}, \text{反}), (\text{反}, \text{正}), (\text{反}, \text{反})\}$$

  - 若硬币不可区分，则样本空间为 $\Omega = \{\{\text{正}, \text{正}\}, \{\text{正}, \text{反}\}, \{\text{反}, \text{反}\}\}$，但样本空间中的样本点并不是等可能出现，所以也不能说出现正、反两面的概率是 $\frac{1}{3}$.

  例 2.11 (Leibnitz 的错误) 掷两个均匀骰子，得到点数和为 11 与 12 的概率相同吗？

- 如何计数：见下面计数原理.

计数原理

加法原理: 考虑试验 A 和 B，若试验 A 有 a 种可能结果，试验 B 有 b 种可能结果，则试验 A 或 B 共有 $a+b$ 种结果

乘法原理 [1]: 考虑一个复合试验，它由两个子试验 A 和 B 构成。假设试验 A 有 a 种可能结果，每一种情况又都对应于试验 B 的 b 种可能结果。那么这个复合试验共有 ab 种可能结果。

[1]通常很容易想当然地认为试验是有先后顺序的，但是乘法原理并没有要求试验 A 必须在试验 B 之前实施。

冰淇淋甜筒购方案计数

例 2.12 (冰淇淋甜筒) 假没某人正在购买一支冰淇淋甜筒，甜筒有蛋卷或华夫饼两种选择，冰淇淋有巧克力 (C)、草莓 (S) 和香草 (V) 三种口味可选。考虑以下问题:

- 该随机试验共有 6 种可能购买方案：无论是先决定甜筒类型还是冰淇淋口味类型，试验都是总共有 $2 \times 3 = 3 \times 2 = 6$ 种可能
- 假设怀特先生某天上下午各买一支冰淇淋甜筒，则该试验共有 $6^2 = 36$ 种可能结果
- 如果只关注这天吃了何种类型的冰淇淋，而不在乎吃冰淇淋的顺序，即 $(cakeC, waffleV)$ 和 $(waffleV, cakeC)$ 之间没有区别，则共有

$$6 * 5/2 + 6 = 21$$

排列组合

- 从 n 个不同的元素中，有放回的取出 r 个元素组成的排列种数为 n^r 种;
- 从 n 个不同的元素中，不放回的取出 r 个元素组成的排列种数为

$$P_n^r = n(n-1)\cdots(n-r+1);$$

- 从 n 个不同的元素中，不放回的取出 r 个元素组成的组合，种数为

$$C_n^r = \frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{n!}{r!(n-r)!};$$

- 从 n 个不同的元素中，有放回的取出 r 个元素组成的组合，种数为 [2]

$$C_{n+r-1}^r.$$

[2] 又称 Bose-Einstein 值，物理学家玻色和爱因斯坦 1920 年做了不可区分粒子的相关研究，成功预测了玻色 - 爱因斯坦凝聚态陌生物质状态的存在

生日问题

例 2.13 若屋子里有 k 个人，假设每个人的生日都等可能为一年 365 天中的一天 (不考虑 2 月 29 日的情况)，而且每个人的生日是相互独立的 (假设这里不存在双胞胎)。试问：屋内有两个及以上的人在同一天过生日的概率是多少？

解： k 个人中没有人生日相同的概率为

$$P(\text{没有人生日相同}) = \frac{365 \cdot 364 \cdots (365-k+1)}{365^k}$$

故至少有两个人生日相同的概率为

$$P(\text{至少有两个人生日相同}) = 1 - \frac{365 \cdot 364 \cdots (365-k+1)}{365^k}$$

可以绘制至少两个人生日相同的概率随 k 变化的函数图。概率首次超过 0.5 是在 $k = 23$, 当 $k = 57$ 时，这个概率超过了 99%.

例 2.14 甲乙丙丁四人进行乒乓球双打练习，两人一对地结为对打的双方，有多少种不同的结对方式?

解: 简单的组合问题: 从四人中选出两人结为一对，剩下的两人结为一对即可。于是可得:

$$\text{有 } C_4^2 = 6 \text{ 种方式. ????}$$

但事实是否如此呢? 注意此时并不要求两队之间的顺序，所以

$$\text{只有 3 种结对方式: } C_4^2/2 = 3.$$

注 2.2

- 在按组合模式分出的“组内”, 元素之间是没有“顺序”的;
- 但是需要指出的是：在“组”之间却存在着“顺序”, 或者叫做“编号”!
- 在按“组合”模式计算时，我们计算的是“取出两个人”的所有不同取法数目.
- 假如我们把取出的两人算一组，而把留下的两人算另一组，那么由于“取出甲乙，留下丙丁”和“取出丙丁，留下甲乙”是两种不同的取出方式，而在这种计算方法中，被算作两种不同的“分组”方式，从而得到 6 种“分组”方式.
- “组合”是一种“有编号的分组模式”，或者说，按照组合模式计算出的分组方式数目中，已经天然地把组的不同编号方式数目计算在内了.

例 2.15 欲将 6 个人分为 3 组，每组 2 人，分别从事 3 项不同工作，求分配方式数.

解: 先取出两人从事第 1 项工作，有 C_6^2 种方式；再取出两人从事第 2 项工作，有 C_4^2 种方式；剩下的两人从事第 3 项工作。所以一共有:

$$C_6^2 \cdot C_4^2 = \frac{6!}{4! \cdot 2!} \frac{4!}{2! \cdot 2!} = \frac{6!}{2! \cdot 2! \cdot 2!} = 90$$

种分配方式.

例 2.16 要把 7 人分为 3 个小组，执行同一种任务，其中一个组 3 人，另两个组各 2 人，求分组方式数.

解: 显然这也是一个“无编号分组”问题。但是却与上面的情况有所不同。因为其中有一个 3 人组，无论是否编号，它都与其余两个组有所区别 (编号无非是为了对分出的组加以区分), 所以在按“有编号分组模式”算出分组方式数之后，只应再除以 2! (即除去两个不加区分的组的排列顺序数), 故得：共有

$$\frac{7!}{3!\cdot 2!\cdot 2!}\cdot\frac{1}{2!}=\frac{7!}{3!\cdot(2!)^3}$$

种分组方式. 为了适应这种分为多个“不同的”组的问题需求，人们总结出如下的“多组组合模式”：

多组组合与排列模式

- 多组组合模式: 有 n 个不同元素，要把它们分为 k 个不同的组，使得各组依次有 $n_1, n_2, \cdots, n_k$ 个元素，其中 $n_1 + n_2 + \cdots + n_k = n$, 则一共有

$$\frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k!}$$

种不同分法.

- 不尽相异元素的排列模式: 有 n 个元素，属于 k 个不同的类，同类元素之间不可辨认，各类元素分别有 $n_1, n_2, \cdots, n_k$ 个，其中 $n_1 + n_2 + \cdots + n_k = n$, 要把它们排成一列，则一共有

$$\frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k!}$$

种不同排法.

例 2.17 一批产品有 N 个，其中废品有 M 个。现从中随机取出 n 个，在以下两种情形下，分别求“其中恰好有 m 个废品”这一事件的概率。(1) 有放回地选取；(2) 不放回地选取.

解： 记 $A=\{$其中恰好有m个废品$\}$, 则

- 有放回情形: $|\Omega|=N^n, |A|=C_n^m M^m (N-M)^{n-m}$, 所以

$$P(A)=\frac{C_n^m M^m (N-M)^{n-m}}{N^n}=C_n^m\left(\frac{M}{N}\right)^m\left(\frac{N-M}{N}\right)^{n-m}$$

- 不放回情形: $|\Omega|=C_N^n, |A|=C_M^m C_{N-M}^{n-m}$, 所以

$$P(A)=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}$$

例 2.18 一个笼子里关着 7 只白猫和 3 只黑猫共 10 只猫，把笼门打开每次只允许钻出 1 只猫。若 10 只猫均钻出笼子，并以 A_k 表示事件“第 k 只出笼的猫是黑猫”，试求 $P(A_k)$，其中，$k = 1, 2, ..., 10$.

方法一：将 10 只猫看成不同的猫

- $|\Omega| = 10!$
- $|A_k| = C_3^1 \cdot 9!$
- $P(A_k) = \dfrac{|A_k|}{|\Omega|} = \dfrac{3 \cdot 9!}{10!} = \dfrac{3}{10}, \quad k = 1, 2, \cdots, 10.$

方法二：仅白猫与黑猫可区别

- 现在只有两种不同的元素，一种有 7 个 (7 只白猫), 另一种有 3 个 (3 只黑猫), 由不尽相异元素的排列模式知

$$|\Omega| = C_{10}^{3} = 120$$

- $|A_k| = C_9^2 = 36$
- $P(A_k) = \frac{36}{120} = \frac{3}{10}, \quad k = 1, 2, \cdots, 10.$

方法三：仅考虑前 k 只出笼猫的排列情况

- $|\Omega| = C_{10}^{k} \cdot k! = \dfrac{10!}{(10-k)!}$
- $|A_k| = C_3^1 \cdot C_9^{k-1} \cdot (k-1)! = \dfrac{3 \cdot 9!}{(10-k)!}$
- $P(A_k) = \dfrac{3 \cdot 9!}{10!} = \dfrac{3}{10}, \quad k = 1, 2, \cdots, 10$

“球-盒”计数问题

- $S(n,k)$: 第二类 Stirling 数, 表示将 n 个元素的集合分解为 k 个非空子集的分法数目;
- 不难验证 $S(n+1,k)=kS(n,k)+S(n,k-1)$, 并可以证明 (枚举空盒的个数，剩下的随便放，盒子是相同的最后要除以 $m!$)

$$S(n,k)=\frac{1}{k!}\sum_{j=0}^{k}(-1)^j C_k^j (k-j)^n$$

- $p(n,k)$: 表示将 n 分解为 k 个正整数之和的分法数目. ($p(n,k)$ 的计算相当复杂, 涉及整数以及集合分拆的问题, 有兴趣的可参阅组合计数方面的文献).

n 个球 r 个盒子	盒可区分		盒不可区分	
	允许空盒	不允许空盒 $n>r$	允许空盒	不允许空盒 $n>r$
球可区分	r^n	$r!S(n,r)$	$\sum_{k=1}^{r}S(n,k)$	$S(n,r)$
球不可区分	C_{n+r-1}^{n}	$C_{(n-r)+r-1}^{n-r}$	$\sum_{k=1}^{r}p(n,r)$	$p(n,r)$

Maxwell-Boltzmann(麦克斯韦-玻尔兹曼) 统计

- 考查热平衡系统以及其中的粒子，在温度充分高且分布的密度足够低，粒子间的量子效应可以忽略.
- 假定能态被标记为 $\{1, 2, \cdots, K\}$，粒子的总数为 N, 那么恰有 N_i 个粒子处于能态 i 上的概率为

$$P_{\mathrm{M-B}} = K^{-N} \frac{N!}{N_1! N_2! \cdots N_K!}, \quad N_1 + \cdots + N_K = N.$$

- Maxwell-Boltzmann 统计最重要的假设是粒子的可区分性.
- 该假设条件下粒子 A 处于能态 1 且粒子 B 处于能态 2，与粒子 A 处于能态 2 且粒子 B 处于能态 1 是不同的两种系统状态.
- 该假设所导出的粒子的能态分布 (即 Boltzmann 分布) 较为符合物理事实.
- 但是其在熵方面却导致了和物理事实严重违背的结果 (所谓的 Gibbs 悖论), 而一旦假设粒子不可区分, 则矛盾会迎刃而解.

Bose-Einstein(博兹-爱因斯坦) 统计

- 当温度下降, 且粒子的空间密度上升时, 粒子间的量子效应变得显著而不可忽略, 粒子在能态间的分布也随之发生变化.
- 根据表现的不同, 可将粒子分为两类: 玻色子 (Boson) 和费米子 (Fermion).
- 无穷多个玻色子能够在同一时间处于 (" 凝聚") 同一能态上, 从而出现统计物理学中著名的 Bose-Einstein 凝聚 (Bose-Einstein condensate) 现象.
- 假定粒子总数为 N 但粒子不可区分, 能态数目为 K, 则玻骰子在各能态上的分布服从 Bose-Einstein 统计,

$$P_{\mathrm{B-E}} = \frac{1}{C_{N+K-1}^{N}} = \binom{N+K-1}{N}^{-1}$$

- Bose-Einstein 统计的粒子能态分布情况与 N 个不可区分的球放入 K 个可区分的盒中的放法形成一一对应.

Fermi-Dirac(费米-狄利克雷) 统计

- 费米子遵守量子力学中的 Pauli 不相容原则，即不同的粒子无法处在相同的能态上 (注意玻色子是不满足这一原则的).
- 因此, 系统中的粒子数目一定比能态数目少, 即 $N < K$.
- 费米子的能态服从 Fermi-Dirac 分布,

$$P_{\mathrm{F-D}} = \frac{1}{C_K^N} = \binom{K}{N}^{-1}$$

- 该分布对应于 N 个不可区分的球放入 K 个可区分的盒中, 且每个盒里不能多于两球的放法数目.
- Fermi-Dirac 统计在帮助人们理解电子的行为方面起到了非常关键的作用.

几何概型的基本思想

1. 可度量性。样本空间 Ω 充满某个区域，其度量 (长度、面积、体积) 大小可用 S_Ω 表示;
2. 等可能性。任意一点落在度量相等的子区域内是等可能的，即任一子区域 A 的概率，只与子区域的度量 S_A 有关，而与子区域的位置无关;
3. 若事件 A 为 Ω 的某个子区域，且其度量大小可用 S_A 表示，则事件 A 的概率为

$$P(A) = \frac{S_A}{S_\Omega}$$

例 2.19 蒲丰投针问题: 平面上画有间隔为 d 的等距平行线，向平面任意投掷一枚长为 l 的针，求针与平行线相交的概率.

解: 注意到

$$\begin{aligned}\Omega &= \{(\varphi, x) : 0 \le x \le d/2, 0 \le \varphi \le \pi\}\\ A &= \{\text{针与平行线相交}\}\\ &= \{(\varphi, x) : x \le \frac{l}{2}\sin\varphi\}\end{aligned}$$

由几何概型知

$$P(A) = \frac{S_A}{S_\Omega} = \frac{\int_0^\pi \frac{l}{2}\sin\varphi d\varphi}{\frac{d}{2}\pi} = \frac{2l}{d\pi}.$$

随机模拟

- l, d 已知，将 π 的值代入即可计算 $P(A)$ 的值;
- 若我们以频率近似概率 $P(A)$, 即

$$\frac{n}{N} \approx P(A) = \frac{2l}{d\pi}$$

则我们可得到 π 的近似值

$$\pi \approx \frac{2lN}{dn}$$

表 1: 历史试验数据

试验者	年份	l/d	投掷次数	数交次数	π 的近似值
Wolf	1850	0.8	5000	2532	3.1596
Fox	1884	0.75	1030	489	3.1595
Lazzerini	1901	0.83	3408	1808	3.1415
Reina	1925	0.54	2520	859	3.1795

会面问题

例 2.20 甲乙约定在下午 6 点到 7 点之间在某处会面，并约定先到者等候另一个人 20 分钟，过时即可离去，求两人能会面的概率.

解: 以 x 和 y 分别表示甲乙两人到达约地点的时间（以分钟为单位），则易知:

$$\begin{aligned}\Omega &:= \{(x,y): 0 \le x \le 60, \quad 0 \le y \le 60\}\\ A &:= \{(x,y): |x-y| \le 20\}\end{aligned}$$

故

$$P(A) = \frac{S_A}{S_\Omega} = \frac{60^2 - 40^2}{60^2} = \frac{5}{9}.$$

例 2.21 在一圆内任取一条弦，问其长度超过该圆内接等边三角形的边长的概率是多少？

解: 易知，圆内接正三角形的边长 $\sqrt{3}R$

1. 注意到弦的长度与它与圆心的距离有关，而与弦的方向无关。显然，当且仅当弦与圆心的距离小于 $R/2$ 时，弦的长超过 $\sqrt{3}R$, 故概率为 $1/2$;
2. 把弦的一端固定，另一端点在圆周上随机移动。若在固定端点作一切线，则与切线交角在 60° 与 120° 之间的弦才能超过 $\sqrt{3}R$. 故所求概率为 $1/3$;
3. 弦长被其中点位置决定，则当且仅当弦的中点落在半径为 $R/2$ 的小圆内时，弦的长度才会超过 $\sqrt{3}R$, 故所求概率为 $1/4$.

第四讲: 黎曼积分、Lebesgue 积分、测度与概率

张鑫

Email: xzhangseu@seu.edu.cn

东南大学　数学学院

黎曼积分

- 函数 $f\colon [a,b] \mapsto \mathbb{R}$;
- 分割 Δ: $a = x_0 < x_1 < \cdots < x_n = b$;
- 黎曼和:

$$I(f,\Delta) := \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1} - x_i),\ \text{其中}\ \xi_i \in [x_i, x_{i+1}];$$

- 令 $|\Delta| := \max_{1\le i\le n} |x_i - x_{i-1}|$, 则若

$$\lim_{|\Delta|\to 0} I(f,\Delta) = \lim_{|\Delta|\to 0} \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1} - x_i) = L := \int_a^b f(x)dx,$$

则称 L 为函数 f 在区间 $[a,b]$ 的积分, 并称函数 f 是黎曼可积的.

黎曼积分示意图

黎曼积分

黎曼积分的缺点

- 一些很简单的函数不是黎曼可积的, 如 Dirichlet 函数:

$$D(x) := \begin{cases} 1, & x\text{是有理数}, \\ 0, & x\text{是无理数}. \end{cases}$$

- 需要一致收敛条件来保证函数积分与极限运算的次序交换, 即若 $f_n(x)$ 一致收敛且黎曼可积, 则

$$\lim_{n\to\infty}\int_a^b f_n(x)dx = \int_a^b \lim_{n\to\infty} f_n(x)dx$$

- 黎曼可积函数空间不完备, 即存在黎曼可积函数列 $f_n(x)$ 逐点收敛到函数 $f(x)$, 但 $f(x)$ 不是黎曼可积.

黎曼积分的等价定义 (I)

- 考虑函数 $f\colon [a,b] \mapsto \mathbb{R}$ 及分割 Δ: $a = x_0 < x_1 < \cdots < x_n = b$, 并令 $|\Delta| := \max_{1\le i\le n} |x_i - x_{i-1}|$ 以及

$$M_i(\Delta) := \sup_{x_{i-1}\le x\le x_i} f(x),\ m_i(\Delta) := \inf_{x_{i-1}\le x\le x_i} f(x),$$

$$M(\Delta) := \sum_{i=0}^{n-1} M_i(x_{i+1}-x_i),\ m(\Delta) := \sum_{i=0}^{n-1} m_i(x_{i+1}-x_i).$$

- 上下积分:

$$\overline{\int_a^b} f(x)dx := \lim_{|\Delta|\to 0} M(\Delta) = \inf_{\Delta} M(\Delta),$$

$$\underline{\int_a^b} f(x)dx := \lim_{|\Delta|\to 0} m(\Delta) = \sup_{\Delta} m(\Delta)$$

- $f(x)$ 黎曼可积当且仅当上下积分相等, 即 $\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx$

黎曼积分的等价定义 (II)

定义 2.13 (分段常值函数或阶梯函数) 函数 $f\colon [a,b] \mapsto \mathbb{R}$ 称为分段常值函数或阶梯函数, 如果区间 $[a,b]$ 存在一个有限区间分割 $I_1, I_2, \cdots, I_n$ 使得 $f(x)$ 在每个区间上 I_i 上恒为常数 c_i, 即

$$f(x) = \sum_{i=1}^{n} c_i 1_{I_i}(x)$$

注 2.3

- 容易验证, $\sum_{i=1}^{n} c_i |I_i|$ 与分段常值函数或阶梯函数 $f(x)$ 表达形式中所选择的区间分割无关.
- 我们记

$$p.c. \int_a^b f(x)dx := \sum_{i=1}^{n} c_i |I_i|$$

并将其称为分段常值函数或阶梯函数 $f(x)$ 在 $[a,b]$ 上的积分.

黎曼积分的等价定义 (II)

定义 2.14 (Darboux integral 达布积分) 设函数 $f\colon [a,b] \mapsto \mathbb{R}$ 为有界函数, 则

- 称
$$\underline{\int_a^b} f(x)dx := \sup_{g\le f,\ g\text{为阶梯函数}} p.c. \int_a^b g(x)dx$$
为函数 $f(x)$ 在 $[a,b]$ 上的达布下积分;
- 称
$$\overline{\int_a^b} f(x)dx := \inf_{h\ge f,\ h\text{为阶梯函数}} p.c. \int_a^b h(x)dx$$
为函数 $f(x)$ 在 $[a,b]$ 上的达布上积分;
- 显然, $\underline{\int_a^b} f(x)dx \le \overline{\int_a^b} f(x)dx$. 如果
$$\underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx,$$
则称函数 $f(x)$ 在 $[a,b]$ 上达布可积.

黎曼积分的等价定义 (II)

定理 2.2 设函数 $f\colon [a,b] \mapsto \mathbb{R}$ 为有界函数, 则函数 $f(x)$ 黎曼可积当且仅当函数 $f(x)$ 达布可积.

黎曼积分的本质

- 黎曼积分思想源于古希腊的“穷竭法”.
  - 求圆的面积: 通过圆外切正 n 边形和内接正 n 边形, 来得到圆面积的上下界, 然后取极限发现上下界都相等, 从而得到圆的面积.
  - 求函数 $f(x)$ 围成的面积: 把 $f(x)$ 与坐标轴围成的区域通过对定义域 $[a,b]$ 进行有限分割并利用分割小区间上函数 $f(x)$ 的最大最小值构造函数外接与函数内接长方形, 然后取分割的极限得到函数 $f(x)$ 围成的面积.
  - 黎曼积分整个过程就是: 有限分割–近似求和–取极限.
- 对函数定义域进行有限区间分割后近似求和取极限, 适用于连续函数, 即函数在小区间内振荡不明显, 但对然而对于振荡很厉害的函数, 哪怕分割很细, 在很细的区间内振荡依然很厉害 (典型的例子是狄利克雷函数).

Lebesgue 积分基本思想

- 黎曼积分是对函数的定义域 $[a,b]$ 进行分割，而 Lebesgue 积分则是对函数 $f(x)$ 的值域进行分割.
- 令 $l<L$ 分别为 $f(x)$ 在区间 $[a,b]$ 上的下确界与上确界，考虑 $[l,L]$ 的分割 Δ_y:

$$l=l_0<l_1<l_2<\cdots<l_n=L$$

Lebesgue 积分基本思想

- 令

$$|\Delta_y| := \max_{1\le k\le n} |l_k - l_{k-1}|,$$
$$E_k := \{x : l_k \le f(x) < l_{k+1}\}, k = 0, 1, \cdots, n-1,$$

- 显然

$$\sum_{k=0}^{n-1} l_k 1_{E_k}(x) \le f(x) < \sum_{k=0}^{n-1} l_{k+1} 1_{E_k}(x), \forall x \in [a, b].$$

因此, 相应的积分应保持上述不等关系, 即

$$\sum_{k=0}^{n-1} l_k \int_a^b 1_{E_k}(x)dx \le \int_a^b f(x)dx < \sum_{k=0}^{n-1} l_{k+1} \int_a^b 1_{E_k}(x)dx.$$

- 类似于黎曼积分的想法, 我们希望当值域分割 $|\Delta_y| \to 0$ 时, 上述左右两端的极限存在且相等.

Lebesgue 积分基本思想

- 若对每一个 $E_k, k=0,\cdots,n-1$, 示性函数 $1_{E_k}(x)$ 黎曼可积, 且 $|\Delta_y|<\epsilon$, 则

$$\sum_{k=0}^{n-1}(l_{k+1}-l_k)\int_a^b 1_{E_k}(x)dx \le \epsilon\sum_{k=0}^{n-1}\int_a^b 1_{E_k}(x)dx = \epsilon(b-a).$$

- 若示性函数 $1_{E_k}(x)$ 不是黎曼可积的, 如何定义:

$$m(E_k) := \int_a^b 1_{E_k}(x)dx,$$

也即集合 E_k 的测度 $m(E_k)$ 如何定义?

- 若我们可以定义 $m(E_k)$, 则我们可以类似于“黎曼和”一样, 我们构造“勒贝格和”, 用面积已知的一些区域逼近曲线下方的区域:

$$\int_a^b f(x)dm := \lim_{|\Delta_y|\to 0}\Big[\sum_{k=0}^{n-1} l_k\cdot m(E_k)\Big].$$

黎曼积与与 Lebesgue 积分的直观解释

- 想像一位零售商, 在一天结束营业时汇总今天的营业收入.
  - 黎曼积分法: 从钱箱中依次取出现金累加计算获得总的营业收入;
  - Lebesgue 积分法: 先将钱箱中的现金按面值进行分类, 用面值乘以该面值现金的数量并累加求和获得总的营业收入.

如何测量集合大小

- 基数或势 (Cardinality): 集合元素的个数;
- 称两个集合的基数相等, 如果两个集合之间存在一一映射;
- $A = \{2, 4, 6, \cdots, \}$ 与 $B = \{1, 2, 3, 4, \cdots\}$ 的基数是相等的;
- 基数或势的优缺点:
  - 在有限集上, 基数还是很有用处的;
  - 但对于无限集情形, 这个度量显得过于粗糙, 例如, 实数集上任意两个区间的基数都是相同的;
- 对于连续的集合, 我们又有自然的 "体积" 的概念:
  一维空间: 长度; 二维空间: 面积; 三维空间: 传统的体积;

可数集与不可数集

- 称集合是可数 (countable) 的, 如果该集合和自然数集 $\mathbf{N}$ 的子集之间存在一一映射, 否则称该集合是不可数 (uncountable) 集.
- 可数集包括有限可数集和无穷可数集.
- 有限个可数集的并仍然是可数集, 可数个可数集的并仍然是可数集.
- 两个可数集的笛卡尔积是可数集.
- 实数区间 $(0,1)$ 是典型的不可数集.
  假设 $(0,1)$ 可数, 其元素可以排列为 $a_1, a_2, a_3, \cdots$. 注意到 $(0,1)$ 中的元素 a_n 可以表示为

$$a_n = 0.a_n^{(1)} a_n^{(2)} a_n^{(3)} \cdots = \sum_{k=1}^{\infty} a_n^{(k)} 10^{-k}$$

  构造 $x = 0.x_1 x_2 x_3 \cdots$, 其中, 对任意的 k, 取 $x_k \neq a_k^{(k)}$. 显然, $x \neq a_k$, $\forall k \in \mathbf{N}$, 矛盾. 故 $(0,1)$ 是不可数集合. 由于

$$f(x) = -\cot(\pi x) : (0,1) \to (-\infty, \infty)$$

  是一一映射, 立刻得到 $\mathbf{R}$ 也是不可数集.

集合的测度

- 对于集合大小, 我们希望定义另外的度量标准: 测度.
  - 希望定义一个测度来测量集合的大小, 能够处理各种情况, 包括有限集、无限集甚至是各种抽象的集合;
  - 希望在某些特定的情况下, 和我们熟知的度量相容: 比如区间的长度, 区域的面积, 空间的体积等度量相容;
  - 希望具有常见度量的一些性质: 可加性.
- 如此一来, 设计一个测度就好像变成了制造一把尺子 (或者量筒: 多维情形).
- 尺子有其适用范围: 用一个普通米尺来测量地球的半径估计会很有困难, 或者用它去测量原子的半径也是不可能的.
- 测度的适用范围在测度论里同样重要: 在定义一个测度的时候, 我们需要同时定义哪些集合是 “可测的”(measurable) – 也就是可以用这个测度来衡量.

区间、矩形、基本集及其测度

- 区间 $I \subset \mathbb{R}$: $[a,b]$, $[a,b)$, (a,b), $\cdots$;
- 矩形 $B \subset \mathbb{R}^d$: $B := I_1 \times \cdots \times I_d$, 其中 I_i 为区间;
- 矩形的测度: $|B| := |I_1| \times |I_2| \times \cdots \times |I_d|$
- 基本集 $E \subset \mathbb{R}^d$: 能够表示为有限个矩形的并的集合;
  - 更进一步, 可以证明基本集一定可以表示为有限个不相交矩形的并, 即
    $$E = \cup_{i=1}^k B_i, \text{其中} B_i, i = 1, \cdots, k, \text{为矩形且互不相交};$$
  - 我们将 $m(E) := |B_1| + |B_2| + \cdots + |B_k|$ 称为基本集 E 的基本测度.

约当测度 (Jordan Measure)

定义 2.15 (约当测度) 令 E 为 $\mathbb{R}^d$ 的有界子集.

- 称
$$m_{*,J}(E) := \sup_{A\subset E, A\text{为基本集}} m(A)$$
为约当内测度;
- 称
$$m^{*,J}(E) := \inf_{A\supset E, A\text{为基本集}} m(A)$$
为约当外测度;
- 称 E 为约当可测的, 如果 $m_{*,J}(E) = m^{*,J}(E)$. 此时称 $m(E) := m_{*,J}(E) = m^{*,J}(E)$ 为集合 E 的约当测度.

约当测度的优缺点

设 E, $F \subset \mathbb{R}^d$ 为约当可测集, 则

- $E \cup F$, $E \cap F$, $E \backslash F$, $E \Delta F$ 均为约当可测集;
- 非负性: $m(E) \geq 0$;
- 有限可加性: 如果 E, F 不相交, 则 $m(E \cup F) = m(E) + m(F)$;
- 平称不变性: $m(x + E) = m(E)$;
- $E_i, i = 1, 2, \cdots$ 约当可测, 但 $\cup_{i=1}^{\infty} E_i$, $\cap_{i=1}^{\infty} E_i$ 不一定约当可测

Riemann 积分与 Jordan 测度之间的联系

定理 2.3 (Riemann 积分与 Jordan 测度之间的联系)

- 如果 E 是区间 $[a,b]$ 上的 Jordan 可测集, 则示性函数

$$1_E(x) := \begin{cases} 1, & x \in E \\ 0, & x \notin E \end{cases}$$

是 Riemann 可积的且 $\int_a^b 1_E(x)\,dx = m(E)$.

- 设 $[a,b]$ 为一区间, $f\colon [a,b] \to \mathbb{R}$ 为一有界函数, 则 f 是 Riemann 可积当且仅当

$$E_+ := \{(x,t) : x \in [a,b], 0 \le f(x) \le t\} \tag{1}$$

$$E_- := \{(x,t) : x \in [a,b], f(x) \le t \le 0\} \tag{2}$$

在 $\mathbb{R}^2$ 上是 Jordan 可测的, 且有

$$\int_a^b f(x)\,dx = m_2(E_+) - m_2(E_-)$$

此处的 m_2 表示二维 Jordan 测度.

Lebesgue 测度

- Lebesgue 外测度 $m^*(E)$:

$$m^*(E) := \inf_{\cup_{i=1}^{\infty} B_i \supset E, B_i, i \geq 1, \text{为矩形}} \sum_{i=1}^{\infty} |B_i|$$

- 显然, $m^*(E) \leq m^{*,J}(E)$, 且存在 $E \subset \mathbb{R}^d$ 使得不等号严格成立;
- Lebesgue 可测集的卡拉西奥多里 (Carathéodory) 准则:

称 $E \subset \mathbb{R}^d$ 为 Lebesgue 可测集, 如果对任意的 $S \subset \mathbb{R}^d$ 均有

$$m^*(S) = m^*(S \cap E) + m^*(S \backslash E),$$

并称 $m(E) := m^*(E)$ 为集合 E 的Lebesgue 测度.

Lebesgue 可测集的代数结构

定义 2.16 (事件或集合的 Borel 运算) 我们将对所给出的一些事件或集合所作的各种 (有限次或可列次) 取余、取交和取并运算以及它们的混合运算都称为 事件或集合的 Borel 运算。

定义 2.17 ($\sigma-$ 代数) 设 $\mathcal{F}$ 为 $\mathbb{R}^d$ 的某些子集构成的集合类, 称 $\mathcal{F}$ 为 $\mathbb{R}^d$ 上的 σ-代数, 如果

(1) $\emptyset \in \mathcal{F}$;

(2) 若 $A \in \mathcal{F}$, 则 $\overline{A} := \mathbb{R}^d - A \in \mathcal{F}$;

(3) 若 $A_i \in \mathcal{F}, i = 1, 2, \cdots,$, 则 $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

命题 2.1 在 $\mathbb{R}^d$ 上的 Lebesgue 可测集具有以下性质:

- 若 E, $F \subset \mathbb{R}^d$ Lebesgue 可测, 则 $E \cup F$, $E \cap F$, $E \backslash F$, $E \Delta F$ 均 Lebesgue 可测;
- 若 $E_i, i \geq 1$,Lebesgue 可测, 则 $\cup_{i=1}^{\infty} E_i$, $\cap_{i=1}^{\infty} E_i$ 也 Lebesgue 可测.

Lebesgue 测度的性质

- 非负性: $m(E) \geq 0$;
- $m(\emptyset) = 0$;
- 可列可加性: 如果 $E_i, i \geq 1$ 互不相交且 Lebesgue 可测, 则

$$m(\cup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} m(E_i);$$

- 平移不变性: $m(x + E) = m(E)$;
- 存在 $F \subset \mathbb{R}^d$ 不 Lebesgue 可测.

不可测集

巴拿赫-塔斯基分球悖论 (Banach-Tarski Paradox)

将一个三维的半径为 1 的实心球用某种巧妙方法分成五等分 (五等分的意思是, 把其中一份旋转平移后可以和另外一份重合), 然后把这五个分块旋转平移后, 可以组合成两个半径为 1 的实心球.

选择公理 (Axiom of Choice): 任意一组 (可能有不可数无限个) 非空集合, 我们都可以从每个集合挑出一个元素.

巴拿赫-塔斯基分球悖论的启示

- 在 $\Omega = \mathbb{R}^3$ 时, 无法构造出一个测度既与常见的体积度量相容且具有平移不变性, 又可以度量 $\mathbb{R}^3$ 的任一子集;
- 类似的, 在 $\Omega = \mathbb{R}$ 时, 同样无法构造出一个测度既与常见的长度度量相容且具有平移不变性, 又可以度量 $\mathbb{R}$ 的任一子集; (不可测集构造详见教材第 33-34 页 2.4 节 "不可测集".)
- 因此, 对于可测空间 $(\Omega, \mathcal{F})$ 中的 $\mathcal{F}$ 通常为 Ω 的某些子集构成的 σ 代数, 而不一定是 Ω 的所有子集构成的集类.

由 Lebesgue 测度到一般抽象测度

- 要度量集合元素的全体: Ω
- 测度的适用范围 (定义域): $\mathcal{F} := \{A \subset \Omega : A\text{可测}\}$, 通常要求 $\mathcal{F}$ 为 σ-代数
- 通常称 $(\Omega, \mathcal{F})$ 为可测空间
- 测度的定义: 称定义在 $\mathcal{F}$ 的集函数 $\mu(\cdot)$ 为测度, 如果
  (1) $\mu(A) \geq 0, \forall A \in \mathcal{F}$;
  (2) $\mu(\emptyset) = 0$;
  (3) μ 具有可列可加性: 若 $A = \cup_{i=1}^{\infty} A_i$, 其中 $A, A_i \in \mathcal{F}$, 且 $i \neq j$ 时, $A_i \cap A_j = \emptyset$, 则

$$\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$$

- 称 $(\Omega, \mathcal{F}, \mu)$ 为测度空间.

概率与测度

- Ω: 概率论中的样本空间对应于测度中要度量的集合所包含元素的全体
- $\mathcal{F}$: 事件域, 即某些随机事件组成的集合, 通常要求其为 σ-代数;
  - 注意到随机事件是随机试验中我们所关心的可能出现的结果, 它由一个或若干个基本事件组成, 因此事件是样本空间的子集, 从而事件域为样本空间的某些子集构成的集合;
  - 概率论中的事件域对应于测度中要度量的集合全体即可测集的概念
- $(\Omega, \mathcal{F})$:可测空间

概率的公理化定义

定义 2.18 (概率的公理化定义) 设 Ω 为样本空间, $\mathcal{F}$ 为 Ω 的某些子集组成的 σ-代数 (事件域), 若定义在 $\mathcal{F}$ 上的集函数 $P(\cdot)$ 满足

(1) 非负性: $P(A) \geq 0$;

(2) 正则性: $P(\Omega) = 1 \Rightarrow P(\emptyset) = 0$(结合可列可加性);

(3) 可列可加性: 若 $A_i, i = 1, 2, \cdots$ 互不相交, 则

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i).$$

则称 P 为 $(\Omega, \mathcal{F})$ 上的概率测度, 简称概率, 称 $(\Omega, \mathcal{F}, P)$ 为概率空间.

关于概率的几个注记

- 从概率的定义可以看出，概率是全集度量为 1 的测度；
- 概率与我们熟知的度量 (长度，面积，体积，质量) 相比，本质上没有太大的区别，只不过概率是用来衡量随机事件发生可能性大小的一种度量，其在全集上的度量为 1；
- 同一随机试验得到的可测空间 $(\Omega, \mathcal{F})$，可定义不同的概率.
  - 对于抛一枚硬币这一随机试验，

$$\Omega = \{\text{正}, \text{反}\}, \text{ 选取} \mathcal{F} = \{\emptyset, \{\text{正}\}, \{\text{反}\}, \Omega\};$$

  - 我们可以构造概率 $P_1(\cdot)$:

$$P_1(\emptyset) = 0, P_1(\{\text{正}\}) = 1/2, P_1(\{\text{反}\}) = 1/2, P_1(\Omega) = 1;$$

  - 也可以构造概率 $P_2(\cdot)$: 其中 $p + q = 1$,

$$P_2(\emptyset) = 0, P_2(\{\text{正}\}) = p, P_2(\{\text{反}\}) = q, P_1(\Omega) = 1.$$

第五讲：概率的公理化体系

张鑫

Email: xzhangseu@seu.edu.cn

东南大学 quad 数学学院

代数 (algebra) 或域 (field)

定义 2.19 (集合的代数 (algebra)) 设 $\mathcal{A}$ 是由 Ω 的某些子集所构成的集合类, 称 $\mathcal{A}$ 为集合 Ω 上的代数, 如果,

1. $\Omega \in \mathcal{A}, \varnothing \in \mathcal{A}$;
2. 若 $A \in \mathcal{A}$, 则 $\overline{A} := \Omega - A \in \mathcal{A}$;
3. 若 $A, B \in \mathcal{A}$, 则 $A \cup B \in \mathcal{A}$.

例 2.22 设 $\Omega = \{a, b, c, d\}$, 则集族 $\mathcal{A} = \{\varnothing, \Omega, \{a\}, \{b, c, d\}\}$ 是代数.

例 2.23 (有限集上的代数) 设 Ω 为有限集, 则集族 2^{Ω} (即 Ω 的所有子集构成的集合) 是代数. 例如, 当 $\Omega = \{a, b, c\}$ 时,

$$\mathcal{A} = \{\varnothing, \Omega, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$$

是代数.

注 2.4

- 代数是" 集合的集合", 即其本身是集合, 其中的元素仍然是集合.
- 代数对于集合的求补、有限交和有限并运算封闭。
- 由于代数仅对集合的有限交并封闭, 若样本空间 Ω 是无限集, 将会导致它将无法包含一些重要的集合 (特别是无限集合).

事件 σ 域定义及其性质

定义 2.20 (事件 σ 域或 σ- 代数) 设 $\mathcal{F}$ 为 Ω 的某些子集构成的集合类，称 $\mathcal{F}$ 为 Ω 上的事件 σ 域或 σ- 代数，如果

1. $\Omega \in \mathcal{F}$;
2. 若 $E \in \mathcal{F}$, 则 $\overline{E} := \Omega - E \in \mathcal{F}$;
3. 若 $E_i \in \mathcal{F}, i = 1, 2, \cdots$, 则 $\cup_{i=1}^{\infty} E_i \in \mathcal{F}$.

定理 2.4 若 $\mathcal{F}$ 为 Ω 上的 σ-代数，则

- $\emptyset \in \mathcal{F}$;
- 若 $E_k \in \mathcal{F}, k = 1, 2, \cdots$, 则 $\cap_{k=1}^{\infty} E_k \in \mathcal{F}$;
- 若 $E_1, \cdots, E_n \in \mathcal{F}$, 则 $\cap_{k=1}^{n} E_k \in \mathcal{F}$ 且 $\cup_{k=1}^{n} E_k \in \mathcal{F}$;
- 若 $E_1, E_2 \in \mathcal{F}$, 则 $E_1 - E_2 \in \mathcal{F}$.

几个例子

例 2.24 (代数而非 σ-代数) 设 Ω 是无限集合, 集族 $\mathcal{A} \subset 2^{\Omega}$ ，满足

$$A \in \mathcal{A} \Longleftrightarrow \#A < \infty \text{ 或} \#\overline{A} < \infty,$$

即 $\mathcal{A}$ 中的集合或者自身是有限集，或者其补集是有限集. 容易验证, $\mathcal{A}$ 是代数, 但不是 σ-代数.

例 2.25 (可数集上 σ-代数) 设样本空间 Ω 为可数集, $\mathcal{F}$ 为其上的 σ-代数, 且 $\forall a \in \Omega$, 有 $\{a\} \in \mathcal{F}$, 则 $\forall A \subset \Omega$, 有 $A \in \mathcal{F}$.

例 2.26 设 $\mathcal{F}$ 为实数集 $\mathbf{R}$ 上的 σ-代数. 若对 $\forall x \in \mathbf{R}$, 都有 $(-\infty, x] \in \mathcal{F}$, 则 $\mathcal{F}$ 包含 $\mathbf{R}$ 上的任意区间.

证明: 仅以 $[a, b)$ 为例. 注意到:

$$(-\infty, a) = \bigcup_{k=1}^{\infty} \left(-\infty, a - \frac{1}{k}\right] \in \mathcal{F}, \quad (x, \infty) = \overline{(-\infty, x]} \in \mathcal{F}, \forall x$$

$$[b, \infty) = \bigcap_{k=1}^{\infty} \left(b - \frac{1}{k}, \infty\right) \in \mathcal{F} \quad [a, b) = \overline{(-\infty, a) \cup [b, \infty)} \in \mathcal{F}$$

σ-代数例子及如何构造

- $\mathcal{F} := \{\emptyset, \Omega\}$;
- $\mathcal{F} := \{A : A \subset \Omega\}$;
- $\mathcal{F} := \{\emptyset, A, \bar{A}, \Omega\}$;
- 如何由 $\{A, B\}$(假设 A, B 为非平凡事件，互不包含且不互为对立事件) 扩充成事件 σ 域?

最小 (生成)σ-代数

定理 2.5 假设 $\mathcal{C}$ 为 Ω 的子集组成的非空集类，则存在唯一 σ-代数 $\mathcal{C}_0$ 使得:

(1) $\mathcal{C} \subset \mathcal{C}_0$;

(2) 若 $\mathcal{G}$ 为任一 σ-代数，且 $\mathcal{C} \subset \mathcal{G}$, 则 $\mathcal{C}_0 \subset \mathcal{G}$.

证明:

- 若 $\mathcal{F}_t, t \in T$ 为一族 σ-代数，则 $\cap_{t\in T}\mathcal{F}_t$ 仍为 σ-代数;
- 令 $\mathcal{C}^{\#} := \{\mathcal{B} : \mathcal{B}为\sigma - 代数, \mathcal{C} \subset \mathcal{B}\}$;
- 令 $\mathcal{C}_0 := \cap_{\mathcal{B}\in\mathcal{C}^{\#}}\mathcal{B}$, 则易证 $\mathcal{C}_0$ 即为定理所求之 σ-代数.

注 2.5

- $\mathcal{C}_0 := \cap_{\mathcal{B}\in\mathcal{C}^{\#}}\mathcal{B}$ 是包含集类 $\mathcal{C}$ 的最小 σ-代数;
- 一般也称 $\mathcal{C}_0$ 为集类 $\mathcal{C}$ 的生成 σ-代数，后面我们常用 $\sigma(\mathcal{C})$ 表示 $\mathcal{C}_0$.

实直线上的 σ- 域

例 2.27 令 $\Omega := \mathbb{R}$, 考虑

$$\mathcal{A}_1 = \{(-\infty, x) | x \in \mathbb{R}\},\ \mathcal{A}_2 = \{(a, b) | -\infty < a < b < \infty\},$$

则 $\sigma(\mathcal{A}_1) = \sigma(\mathcal{A}_2)$

解: 注意到，对任意实数 x, 均有

$$(-\infty, x) = \cup_{n=1}^{\infty}(x - n, x).$$

因此 $(-\infty, x) \in \sigma(\mathcal{A}_2)$, 从而 $\mathcal{A}_1 \subset \sigma(\mathcal{A}_2)$, 进而可得 $\sigma(\mathcal{A}_1) \subset \sigma(\mathcal{A}_2)$.
反之，注意到 $[a, b) = (-\infty, b) - (-\infty, a) \in \sigma(\mathcal{A}_1)$ 以及

$$\{a\} = \cap_{n=1}^{\infty}[a, a + \frac{1}{n}) \in \sigma(\mathcal{A}_1),$$

因此，$(a, b) = [a, b) - \{a\} \in \sigma(\mathcal{A}_1)$, 即 $\mathcal{A}_2 \subset \sigma(\mathcal{A}_1)$, 从而

$$\sigma(\mathcal{A}_2) \subset \sigma(\mathcal{A}_1).$$

Borel 集

例 2.28 设 $\Omega = \mathbb{R}$, 令

$$
\begin{aligned}
\mathcal{C}_1 &= \{(-\infty, x) : -\infty < x < +\infty\} \\
\mathcal{C}_2 &= \{(a, b] : -\infty \le a \le b < +\infty\} \\
\mathcal{C}_3 &= \{[a, b] : -\infty < a \le b < +\infty\} \\
\mathcal{C}_4 &= \{[a, b) : -\infty < a \le b \le +\infty\} \\
\mathcal{C}_5 &= \{(a, b) : -\infty \le a \le b \le +\infty\} \\
\mathcal{C}_6 &= \{A \subset \mathbb{R} : A\text{为开集}\} \\
\mathcal{C}_7 &= \{A \subset \mathbb{R} : A\text{为闭集}\}
\end{aligned}
$$

则以上 7 个集类的生成 σ-代数均相同，称以上集类的生成 σ-代数中的元素为 Borel 集，并记作 $\mathcal{B}(\mathbb{R})$, 即

$$\sigma(\mathcal{C}_1) = \cdots = \sigma(\mathcal{C}_7) := \mathcal{B}(\mathbb{R}).$$

其他集类

定义 2.21 设 $\mathcal{C}$ 为 Ω 的某些子集构成的非空集类,

- 称 $\mathcal{C}$ 为 π 类, 如果 $\mathcal{C}$ 关于有限交运算封闭, 即

$$A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C};$$

- 称 $\mathcal{C}$ 为代数, 如果 $\mathcal{C}$ 关于有限交与取余运算封闭, 即

$$A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}, \bar{A} \in \mathcal{C};$$

- 称 $\mathcal{C}$ 为单调类, 如果 $\mathcal{C}$ 关于单调极限运算封闭, 即

$$A_n \in \mathcal{C}, A_n \uparrow A \text{ 或} A_n \downarrow A \Rightarrow A \in \mathcal{C}$$

- 称 $\mathcal{C}$ 为 λ 类, 如果 $\Omega \in \mathcal{C}$ 并且 $\mathcal{C}$ 对真差运算及上升极限运算封闭, 即
  - $\Omega \in \mathcal{C}$;
  - $A, B \in \mathcal{C}, B \subset A \Rightarrow A - B \in \mathcal{C}$;
  - $A_n \in \mathcal{C}, n \geq 1, A_n \uparrow A \Rightarrow A \in \mathcal{C}$

集类之间的关系

定理 2.6 设 $\mathcal{C}$ 为 Ω 的某些子集构成的非空集类, 则

1. 若 $\mathcal{C}$ 为 $\sigma-$ 代数, 则 $\mathcal{C}$ 一定是:π 类, 代数,λ 类, 单调类;
2. 若 $\mathcal{C}$ 既是代数又是单调类, 则 $\mathcal{C}$ 一定是 $\sigma-$ 代数;
3. 若 $\mathcal{C}$ 既是 π 类又是 λ 类, 则 $\mathcal{C}$ 一定是 $\sigma-$ 代数.

单调类定理

定义 2.22 与生成 $\sigma-$ 代数类似, 我们可以定义包含集类 $\mathcal{C}$ 的最小单调类 $m(\mathcal{C})$ 及最小 λ 类 $\lambda(\mathcal{C})$, 分别称为由集类 $\mathcal{C}$ 生成的单调类与 λ 类.

定理 2.7 (单调类定理) 设 $\mathcal{C}, \mathcal{G}$ 为 Ω 中的集类且 $\mathcal{C} \subset \mathcal{G}$, 则

1. 若 $\mathcal{C}$ 为一代数, 则 $m(\mathcal{C}) = \sigma(\mathcal{C})$;
2. 若 $\mathcal{C}$ 为一 π 类, 则 $\lambda(\mathcal{C}) = \sigma(\mathcal{C})$;
3. 若 $\mathcal{C}$ 为代数而 $\mathcal{G}$ 为单调类, 则 $\sigma(\mathcal{C}) \subset \mathcal{G}$;
4. 若 $\mathcal{C}$ 为 π 类而 $\mathcal{G}$ 为 λ 类, 则 $\sigma(\mathcal{C}) \subset \mathcal{G}$.

如何应用单调类定理

- 问题: 如何证明某 $\sigma-$ 代数 $\mathcal{F}$ 中的所有元素都具有某种性质 (P)?
- 步骤:
  - $\mathcal{G} := \{A \in \mathcal{F} : A\text{具有性质}(P)\}$;
  - 找到某一集类 $\mathcal{C}$ 为代数或 π 类且具有以下性质:

$$\mathcal{C} \subset \mathcal{G}, \quad \sigma(\mathcal{C}) = \mathcal{F};$$

  - 证明 $\mathcal{G}$ 为单调类或 λ 类;
  - 由单调类定理可知:

$$\mathcal{F} = \sigma(\mathcal{C}) \subset \mathcal{G} \subset \mathcal{F}$$

即

$$\mathcal{G} = \mathcal{F}$$

从而 $\mathcal{F}$ 中所有的元素都具有某种性质 (P).

概率的公理化定义

定义 2.23 设 $\mathcal{F}$ 为样本空间 Ω 上的 σ-代数，如果定义在 $\mathcal{F}$ 上的集函数 $P(\cdot)$ 具有:

1. 非负性: $P(A) \geq 0, \forall A \in \mathcal{F}$;
2. 规范性: $P(\Omega) = 1$;
3. 可列可加性：对 $\mathcal{F}$ 中的任何两两互不相容事件列 $\{A_n\}$ 有

$$P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$$

则称 $P(\cdot)$ 为 $\mathcal{F}$ 上的概率测度，而 $P(A)$ 称为事件 A 的概率，$(\Omega, \mathcal{F}, P)$ 称为概率空间.

如何定义概率

- 概率定义在 $\sigma-$ 代数 $\mathcal{F}$ 上, 但是由于 $\mathcal{F}$ 中集合的复杂性, 直接对每个集合赋予概率存在困难;
- 首先在 $\mathcal{F}$ 的子族 $\mathcal{A} \subset \mathcal{F}$ 上定义概率（一般 $\mathcal{A}$ 中的集合较为简单), 然后再将其扩张到 $\mathcal{F}$ 上;
- 如何选择恰当的 $\mathcal{A}$, 使得 $\mathcal{F} = \sigma(\mathcal{A})$?

扩张是否唯一?

例 2.29 (有限集合上的概率) 设样本空间 $\Omega = \{a, b, c, d\}$, σ-代数 $\mathcal{F}$ 由集族 $\mathcal{A}$ 生成, 即

$$\mathcal{F} = \sigma(\mathcal{A}), \quad \mathcal{A} = \{\{a\}, \{a, b\}, \{a, b, c\}, \Omega\}$$

例 2.30 令 $\Omega = \{a, b, c, d\}$, $\mathcal{F} = 2^{\Omega}$, $\mathcal{A} = \{\{a, b\}, \{a, c\}, \{b, d\}, \{c, d\}\}$.

- 容易看出,$\sigma(\mathcal{A}) = 2^{\Omega}$;
- 定义 P_1 和 P_2 为

$$P_1(\{a\}) = P_1(\{d\}) = \frac{1}{6}, \quad P_1(\{b\}) = P_1(\{c\}) = \frac{1}{3}$$
$$P_2(\{a\}) = P_2(\{d\}) = \frac{1}{3}, \quad P_2(\{b\}) = P_2(\{c\}) = \frac{1}{6}$$

- 显然, 它们并不相同, 但它们在 $\mathcal{A}$ 上却完全一致

$$P_1(\{a, b\}) = \frac{1}{2}, P_1(\{a, c\}) = \frac{1}{2}, P_1(\{b, d\}) = \frac{1}{2}, P_1(\{c, d\}) = \frac{1}{2}$$
$$P_2(\{a, b\}) = \frac{1}{2}, P_2(\{a, c\}) = \frac{1}{2}, P_2(\{b, d\}) = \frac{1}{2}, P_2(\{c, d\}) = \frac{1}{2}$$

概率扩张唯一性定理

定理 2.8 (概率扩张的唯一性) 给定样本空间 Ω 和其上的 σ-代数 $\mathcal{F}$. 若

- P_1 和 P_2 是定义在 $\mathcal{F}$ 上的两个概率;
- 集族 $\mathcal{A}$ 关于交运算封闭, 且 $\sigma(\mathcal{A}) = \mathcal{F}$;
- 对任意的 $A \in \mathcal{A}$ 均有 $P_1(A) = P_2(A)$;

则对任意的 $A \in \mathcal{F}$ 一定有 $P_1(A) = P_2(A)$.

证明: 证明思路如下:

- 令 $\mathcal{G} := \{A \in \sigma(\mathcal{A}) : P_1(A) = P_2(A)\}$;
- 注意到 $\mathcal{A}$ 为 π 类且 $\mathcal{A} \subset \mathcal{G}$, 故只需证明 $\mathcal{G}$ 为 λ 类即验证即
  - $\Omega \in \mathcal{G}$;
  - $A, B \in \mathcal{G}, B \subset A \Rightarrow A - B \in \mathcal{G}$;
  - $A_n \in \mathcal{G}, n \geq 1, A_n \uparrow A \Rightarrow A \in \mathcal{G}$
- 利用单调类定理可知 $\mathcal{G} = \sigma(\mathcal{A})$

(概率) 测度扩张定理

定理 2.9 (Carathéodory 测度扩张定理) 设 Ω 为样本空间,$\mathcal{A}$ 为代数. 如果 P_0 为定义在 $\mathcal{A}$ 上的概率, 且满足可数可加性, 那么存在唯一的定义在 $\sigma(\mathcal{A})$ 上的概率 P, 满足

$$P(A) = P_0(A), \quad \forall A \in \mathcal{A},$$

注 2.6 更一般的测度扩张定理详见测度论教材, 如严加安院士的 "测度论讲义" 等.

如何理解概率

- 概率空间定义中的样本空间 Ω 和概率 P 都很好理解，但是 σ-代数 $\mathcal{F}$ 的出现总是显得有些突兀.
- 概率是一个以事件 (集合) 为自变量的函数.
- 给定一个事件 (集合), 概率函数即赋予它一个非负实数, 用以表示该事件 (集合) 的样本点在统计实验的结果中出现的可能性大小.
- 函数的基本概念中，" 定义域" 占据重要位置. 概率函数也不例外.
- 函数的定义域中的数对 "加减乘除" 运算封闭. 类似的, 概率函数应对事件的 "交差并补" 等运算封闭.
- σ-代数 $\mathcal{F}$ 本质上就是概率函数的 "定义域", 对自变量 (事件) 的 "交差并补" 等运算封闭..

概率测度的性质 I

定理 2.10 设 $(\Omega, \mathcal{F}, P)$ 为一概率空间，则其上的概率 P 具有如下性质:

1. $P(\emptyset) = 0$;
2. 有限可加性：若 $A_1, \cdots, A_n$ 为 $\mathcal{F}$ 中两两互不相容事件，则

$$P(\cup_{k=1}^{n} A_k) = \sum_{k=1}^{n} P(A_k)$$

3. 可减性：若 $A, B \in \mathcal{F}$ 且 $A \subset B$, 则 $P(B - A) = P(B) - P(A)$;
4. 单调性：若 $A, B \in \mathcal{F}$ 且 $A \subset B$, 则 $P(A) \leq P(B)$;
5. $P(\bar{A}) = 1 - P(A)$;

概率测度的性质 II

6. 加法公式 (容斥原理): 对任意两个事件 A, B, 有

$$P(A \cup B) = P(A) + P(B) - P(AB). \tag{3}$$

一般的，对任意 n 个事件 $A_1, \cdots, A_n$, 有

$$\begin{aligned} P(\cup_{i=1}^n A_i) &= \sum_{i=1}^n P(A_i) - \sum_{1 \le i < j \le n} P(A_i A_j) + \sum_{1 \le i < j < k \le n} P(A_i A_j A_k) \\ &\quad + \cdots + (-1)^{n-1} P(A_1 A_2 \cdots A_n). \end{aligned} \tag{4}$$

7. 次可加性 (Boole 不等式): 对任意两个事件 A, B, 有

$$P(A \cup B) \le P(A) + P(B).$$

一般的，对任意 n 个事件 $A_1, \cdots, A_n$, 有

$$P(\cup_{i=1}^n A_i) \le \sum_{i=1}^n P(A_i).$$

概率测度的性质 III

8. Bonferroni 不等式:

$$P(\cup_{i=1}^{n} A_i) \geq \sum_{i=1}^{n} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j),$$

$$P(\cup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k), \cdots .$$

9. 下连续性: 若 $A_n \in \mathcal{F}$ 且 $A_n \subset A_{n+1}, n = 1, 2, \cdots$, 则

$$P(\cup_{n=1}^{\infty} A_n) = P(\lim_{n \to \infty} A_n) = \lim_{n \to \infty} P(A_n)$$

10. 上连续性: 若 $A_n \in \mathcal{F}$ 且 $A_n \supset A_{n+1}, n = 1, 2, \cdots$, 则

$$P(\cap_{n=1}^{\infty} A_n) = P(\lim_{n \to \infty} A_n) = \lim_{n \to \infty} P(A_n)$$

11. 概率 P 的可列可加性等价于 P 有限可加且下连续.

加法公式证明 I

注意到

$$A \cup B = A \cup (B - AB).$$

故由概率的有限可加性及减法公式可得:

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

对于一般情形，可由归纳法证明。当 $n = 2$ 时，(4) 式即为 (3) 式。设 (4) 式对 $n-1$ 成立，则先对两个事件 $\cup_{i=1}^{n-1} A_i$ 与 A_n 应用 (3) 得

$$\begin{aligned} P(\cup_{i=1}^{n} A_i) &= P(\cup_{i=1}^{n-1} A_i) + P(A_n) - P((\cup_{i=1}^{n-1} A_i) \cap A_n) \\ &= P(\cup_{i=1}^{n-1} A_i) + P(A_n) - P((\cup_{i=1}^{n-1} (A_i A_n)) \end{aligned} \tag{5}$$

加法公式证明 II

而

$$
\begin{aligned}
P(\cup_{i=1}^{n-1} A_i) &= \sum_{i=1}^{n-1} P(A_i) - \sum_{1\le i<j\le n-1} P(A_iA_j) + \sum_{1\le i<j<k\le n-1} P(A_iA_jA_k) \\
&\quad + \cdots + (-1)^{n-2} P(A_1A_2\cdots A_{n-1}),
\end{aligned}
$$

$$
\begin{aligned}
P((\cup_{i=1}^{n-1}(A_iA_n)) &= \sum_{1\le i\le n-1} P(A_iA_n) - \sum_{1\le i<j\le n-1} P(A_iA_jA_n) \\
&\quad + \cdots + (-1)^{n-2} P(A_1A_2\cdots A_n) \\
&= \sum_{1\le i<j=n} P(A_iA_n) - \sum_{1\le i<j<k=n} P(A_iA_jA_n) \\
&\quad + \cdots + (-1)^{n-2} P(A_1A_2\cdots A_n)
\end{aligned}
$$

故将上面两式代入 (5) 式即得加法公式.

Boole 不等式的证明: 归纳法

- $n=2$ 时显然有

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) \leqslant P(A_1) + P(A_2),$$

- 假定 $n=m$ 时结论成立, 那么 $n=m+1$ 时, 有

$$\begin{aligned} P(\bigcup_{i=1}^{m+1} A_i) &\leq P(\bigcup_{i=1}^{m} A_i) + P(A_{m+1}) \\ &\leqslant \sum_{i=1}^{m} P(A_i) + P(A_{m+1}) \\ &= \sum_{i=1}^{m+1} P(A_i) \end{aligned}$$

Boole 与 Bonferroni 不等式证明: 非归纳法

- 注意到:

$$
\begin{aligned}
\cup_{i=1}^{n} A_i &= A_1 \cup (A_2 - A_1) \cup (A_3 - (A_1 \cup A_2)) \cup \cdots \cup (A_n - \cup_{i=1}^{n-1} A_i) \\
&= A_1 \cup (A_2 \cap \overline{A}_1) \cup (A_3 \cap \overline{A}_1 \cap \overline{A}_2)) \cup \cdots \cup (A_n \cap \cap_{i=1}^{n-1} \overline{A}_i) \\
&= A_1 \cup \cup_{k=2}^{n} (A_k \cap \cap_{i=1}^{k-1} \overline{A}_i)
\end{aligned}
$$

- 从而

概率 P 的下连续性

设 F_n 是 $\mathcal{F}$ 中的一个单调不减的事件序列，即

$$\lim_{n\to\infty} F_n = \cup_{i=1}^{\infty} F_i$$

若令 $F_0 = \emptyset$, 则

$$\cup_{i=1}^{\infty} F_i = \cup_{i=1}^{\infty}(F_i - F_{i-1}).$$

由于 $F_{i-1} \subset F_i$, 显然诸 $F_i - F_{i-1}$ 互不相容，故由可列可加性得

$$\begin{aligned} P(\cup_{i=1}^{\infty} F_i) &= P(\cup_{i=1}^{\infty}(F_i - F_{i-1})) = \sum_{i=1}^{\infty}(P(F_i) - P(F_{i-1})) \\ &= \lim_{n\to\infty}\sum_{i=1}^{n}(P(F_i) - P(F_{i-1})) = \lim_{n\to\infty} P(F_n). \end{aligned}$$

概率 P 的上连续性

设 $\{E_n\}$ 是一列单调不增的事件序列，则 $\{\bar{E}_n\}$ 为单调不减事件序列，由概率的下连续性可得

$$P(\cup_{i=1}^{\infty}\bar{E}_i) = \lim_{n\to\infty} P(\bar{E}_n) = 1 - \lim_{n\to\infty} P(E_n) \tag{6}$$

另一方面，

$$P(\cup_{i=1}^{\infty}\bar{E}_i) = P(\overline{\cap_{i=1}^{\infty}E_i}) = 1 - P(\cap_{i=1}^{\infty}E_i) = 1 - P(\lim_{n\to\infty} E_n) \tag{7}$$

综合 (6) 式与 (7) 式可得

$$\lim_{n\to\infty} P(E_n) = P(\lim_{n\to\infty} E_n)$$

可列可加 $\Leftrightarrow$ 有限可加且下连续

$\Rightarrow$, 显然，下证 $\Leftarrow$。

设 $A_i \in \mathcal{F}, i = 1, 2, \cdots$，是两两互不相容的事件序列，则由有限可加性知，对任意的 n 均有 $P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$, 等式两边同时令 $n \to \infty$ 可得

$$\lim_{n\to\infty} P(\cup_{i=1}^n A_i) = \lim_{n\to\infty} \sum_{i=1}^n P(A_i) = \sum_{i=1}^\infty P(A_i),$$

令 $F_n := \cup_{i=1}^n A_i$, 则易有 F_n 为单调不减事件序列，故有概率的下连续性可得

$$\lim_{n\to\infty} P(\cup_{i=1}^n A_i) = \lim_{n\to\infty} P(F_n) = P(\cup_{n=1}^\infty F_n) = P(\cup_{i=1}^\infty A_i)$$

几个概率空间的例子

例 2.31 (Bernoulli 概率空间) 取 $\mathcal{F}=\{\emptyset, A, \bar{A}, \Omega\}$, 其中 A 为 Ω 的非空真子集。任取两个正数 $p, q(p+q=1)$, 令

$$P(\emptyset)=0, P(A)=p, P(\bar{A})=q, P(\Omega)=1.$$

易证 P 是一个概率测度。从而 $(\Omega, \mathcal{F}, P)$ 是一个概率空间.

例 2.32 (有限概率空间) 样本空间是有限集 $\Omega=\{\omega_1, \cdots, \omega_n\}$. $\mathcal{F}$ 取为 Ω 的一切子集 (共 2^n 个) 组成的集类。设 $\{p_i\}_{1\leq i\leq n}$ 为满足 $\sum_{i=1}^n p_i=1$ 的 n 个非负实数。定义集函数 $P(\cdot)$ 如下

$$P(A)=\sum_{\omega_i\in A} p_i.$$

易证 P 为概率测度，称此 $(\Omega, \mathcal{F}, P)$ 为有限概率空间。特别的若取 $p_i=\dfrac{1}{n}, 1\leq i\leq n$, 则为古典概率空间.

几个概率空间的例子

- 例 2.33 (离散概率空间) 样本空间 $\Omega = \{\omega_1, \omega_2, \cdots\}$ 为可列集，$\mathcal{F}$ 取为 Ω 的所有子集所构成的集类。设 $\{p_i\}_{i\geq 1}$ 为满足 $\sum_{i=1}^{\infty} p_i = 1$ 的一列非负实数。对任意的 $E \subset \Omega$, 定义集函数 $P(\cdot)$ 如下

$$P(E) = \sum_{j:\omega_j \in E} p_j.$$

易证 P 是一个概率测度，于是 $(\Omega, \mathcal{F}, P)$ 是一个概率空间，称为离散概率空间.

- 例 2.34 (一维几何概率空间) 设样本空间 Ω 为实直线上的具有正 Lebesgue 测度的某个 Borel 集，取 $\mathcal{F}$ 为 Ω 的所有 Borel 子集所构成的类。对每个 $E \in \mathcal{F}$, 令

$$P(E) = \frac{m(E)}{m(\Omega)}$$

其中 $m(E)$ 和 $m(\Omega)$ 分别表示集合 E 与 Ω 的 Lebesgue 测度。易证 $P(\cdot)$ 为概率测度.

减法公式妙用

例 2.35 口袋中有编号为 $1, 2, \cdots, n$ 的 n 个球，从中有放回地任取 m 次，求取出 m 个球最大号码为 k 的概率.

解: 令

$$\begin{aligned} A_k &:= \{\text{取出的}m\text{个球的最大号码为}k\}, \\ B_k &:= \{\text{取出的}m\text{个球的最大号码为小于等于}k\} \end{aligned}$$

则易有 $A_k = B_k - B_{k-1}$ 且 $P(B_k) = \frac{k^m}{n^m}$.另一方面，显然有 $B_{k-1} \subset B_k$, 故

$$P(A_k) = P(B_k - B_{k-1}) = P(B_k) - P(B_{k-1}) = \frac{k^m - (k-1)^m}{n^m}$$

对立事件妙用

例 2.36 口袋中有 $n-1$ 个黑球，1 个白球，每次从口袋中随机地摸出一球，并换入一只黑球，求第 k 次摸到黑球的概率.

解: 记 $A := \{第k次取到黑球\}$,则 A 的对立事件为

$$\overline{A} = \{第k次取到白球\}$$

这意味着，第 1 次，$\cdots$，第 $k-1$ 次只能取到黑球，而第 k 次取到白球.

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{(n-1)^{k-1}}{n^k}$$

巧用对立事件解决无空盒问题

例 2.37 将 m 个不同的小球等可能地放入 n 个不同的盒子，$m > n$, 试求无空盒出现的概率.

- 易知 $|\Omega| = n^m$, 记 $E = \{$无空盒出现$\}$, 但 $|E|$ 不易求得;
- 记 $A_k = \{$第k号盒子是空盒$\}$, $|A_k|, |A_{j_1} \cdots A_{j_i}|$ 均易求;
- E 与 A_k 的关系: $\overline{E} = \bigcup\limits_{k=1}^{n} A_k$;
- 加法公式:

$$P(\overline{E}) = P\left(\cup_{k=1}^{n} A_k\right) = \sum_{i=1}^{n} (-1)^{i-1} \sum_{1 \le j_1 < \cdots < j_i \le n} P(A_{j_1} \cdots A_{j_i})$$

巧用对立事件解决无空盒问题

- 计算概率 $P(A_{j_1}\cdots A_{j_i})$:

$$P(A_{j_1}\cdots A_{j_i})=\frac{(n-i)^m}{n^m}=\left(1-\frac{i}{n}\right)^m;$$

- $P(\overline{E})=\sum_{i=1}^{n}(-1)^{i-1}C_n^i\left(1-\frac{i}{n}\right)^m$;
- $P(E)=1-P(\overline{E})=\sum_{i=0}^{n}(-1)^{i}C_n^i\left(1-\frac{i}{n}\right)^m$.

Buffon 投针问题的续一

例 2.38 向画满间隔为 a 的平行直线的桌面上任投一直径为 $l(l<a)$ 的半圆形纸片，求事件 $E=\{$纸片与某直线相交$\}$ 的概率.

解: 设想把半圆形纸片拼成一个圆形，并记

$$F=\{\text{新拼的半圆形与某直线相交}\},$$

则

$$\begin{aligned} E\cup F &= \{\text{直径为}l\text{的圆形与某直线相交}\} \\ E\cap F &= \{\text{长为}l\text{的线段 (即公共直径) 与某直线相交}\} \end{aligned}$$

由前面结论可知 $P(E\cap F)=\dfrac{2l}{\pi a}$. 为计算 $P(E\cup F)$, 可令 d 表示圆心到直线的最近距离，则有 $\Omega=\{d\,|0\le d\le \frac{a}{2}\}, E\cup F=\{d\,|0\le d\le \frac{l}{2}\}$. 因此

$$P(E\cup F)=\frac{l}{a}.$$

从而: $P(E)=\dfrac{P(E)+P(F)}{2}=\dfrac{P(E\cup F)+P(E\cap F)}{2}=\dfrac{\pi l/2+l}{\pi a}.$

Buffon 投针问题的续二

例 2.39 在平面上画有间距为 a 与 b 的水平直线以及垂直直线，向平面随机地抛掷长度为 l 的针，$l < \min\{a, b, a+b-\sqrt{(a+b)^2-\pi ab}\}$，试求事件 E={针与某条直线相交 } 的概率.

- 记 A={针与某条水平直线相交 }, B={针与某条垂直直线相交 };
- $E = A \cup B$, 从而 $P(E) = P(A \cup B) = P(A) + P(B) - P(AB)$;
- $P(A) = \dfrac{2l}{\pi a}, \quad P(B) = \dfrac{2l}{\pi b}$;

Buffon 投针问题的续二

- 计算 $P(AB)$:
  - 设 ρ 和 δ 分别表示针的中点与水平直线的最近距离和与垂直直线的最近距离，设 θ 为针与水平直线的夹角
  - 易知

$$\Omega = \{(\rho, \delta, \theta) | 0 \le \rho \le \frac{a}{2}, 0 \le \delta \le \frac{b}{2}, 0 \le \theta \le \frac{\pi}{2}\},$$

$$AB = \{(\rho, \delta, \theta) | (\rho, \delta, \theta) \in \Omega, \rho \le \frac{l}{2}\sin\theta, \delta \le \frac{l}{2}\cos\theta\}.$$

  - 从而

$$P(AB) = \frac{m(AB)}{m(\Omega)} = \frac{\int_0^{\frac{\pi}{2}} d\theta \int_0^{\frac{l}{2}\cos\theta} d\delta \int_0^{\frac{l}{2}\sin\theta} d\rho}{\frac{a}{2} \cdot \frac{b}{2} \cdot \frac{\pi}{2}} = \frac{l^2}{\pi ab}$$

- 故由加法公式可得

$$P(E) = P(A \cup B) = P(A) + P(B) - P(AB) = \frac{2l(a+b) - l^2}{\pi ab}.$$

配对问题一

例 2.40 设有 n 个孩子，每个人戴着一顶帽子. 如果他们将各自的帽子放在桌子上混杂在一起，然后每人从中随机取一顶帽子戴上，试问至少有一人戴对自己的帽子的概率是多少？

解: 若记

$$A_i := \{第i个人戴对自己的帽子\},$$

则所求事件概率为 $P(\cup_{i=1}^n A_i)$，易知有

$$\begin{aligned}
P(A_i) &= \frac{(n-1)!}{n!} = \frac{1}{n}, \ \sum_{i=1}^n P(A_i) = 1, \\
P(A_iA_j) &= \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}, i \neq j, \\
\sum_{1\le i<j\le n} P(A_iA_j) &= C_n^2 \frac{1}{n(n-1)} = \frac{1}{2!},
\end{aligned}$$

配对问题一（续）

同理可得

$$
\begin{aligned}
P(A_iA_jA_k) &= \frac{(n-3)!}{n!} = \frac{1}{n(n-1)(n-2)}, i \neq j \neq k, \\
\sum_{1\le i<j<k\le n} P(A_iA_jA_k) &= C_n^3 \frac{1}{n(n-1)(n-2)} = \frac{1}{3!}, \\
&\cdots \quad , \\
P(\cap_{i=1}^n A_i) &= \frac{1}{n!}.
\end{aligned}
$$

故由加法公式可得

$$
P(\cup_{i=1}^n A_i) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots + (-1)^{n-1}\frac{1}{n!} \xrightarrow{n\to\infty} 1 - e^{-1}.
$$

$$
e^x = 1 + \sum_{n=1}^{\infty} \frac{x^n}{n!}.
$$

配对问题二

例 2.41 接上面例子，试求下述各事件的概率：
(1) 恰有 k 个人抽到自己礼物；(2) 至少有 m 个人抽中自己礼物。

- 记 E_k ={恰有 k 个孩子戴对自己的帽子 },A_m ={至少有 m 个孩子戴对自己的帽子 };
- $\Omega = n!$, $|E_k| =?$, $|A_m| =?$;
- 若记 D_k ={给定的某 $n-k$ 个孩子均未戴对自己的帽子 },则有 $|E_k| = C_n^k |D_k|$
- D_k 表示给定的某 $n-k$ 个孩子未戴对自己的帽子，不涉及其余 k 个人，故可套用 $n-k$ 情形下的配对问题即

$$\frac{|D_k|}{(n-k)!} = P(D_k) = 1 - P(\overline{D}_k) = \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^{n-k}\frac{1}{(n-k)!}$$

- 因此，

$$|D_k| = (n-k)!P(D_k) = (n-k)!\left\{\frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^{n-k}\frac{1}{(n-k)!}\right\}$$

配对问题二

- 从而

$$
\begin{aligned}
P(E_k) &= \frac{|E_k|}{n!} = \frac{C_n^k |D_k|}{n!} \\
&= C_n^k \frac{(n-k)!}{n!} \left\{ \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^{n-k} \frac{1}{(n-k)!} \right\} \\
&= \frac{1}{k!} \sum_{j=2}^{n-k} (-1)^j \frac{1}{j!} = \frac{1}{k!} \sum_{j=0}^{n-k} (-1)^j \frac{1}{j!}.
\end{aligned}
$$

- 在上式中令 $n \to \infty$, 可得极限概率为 $\dfrac{1}{e \cdot k!}$
- 由于 $A_m = \bigcup\limits_{k=m}^{n} E_k$, 且事件 $E_m, E_{m+1}, \cdots, E_n$ 两两不交，故

$$
P(A_m) = \sum_{k=m}^{n} P(E_k) = \sum_{k=m}^{n} \frac{1}{k!} \sum_{j=0}^{n-k} (-1)^j \frac{1}{j!} = \sum_{k=m}^{n} \sum_{j=0}^{n-k} (-1)^j \frac{1}{k!j!}.
$$

第六讲：条件概率与事件的独立性

张鑫

Email: xzhangseu@seu.edu.cn

东南大学　数学学院

条件思考的重要性

- 条件概率说明了如何以合乎逻辑且相一致的方式将证据纳入人们对世界的理解当中.
- 添加条件是一种非常有用的解决问题的策略.
- 条件既可以作为更新判断依据的方法，也是解决问题的策略.
- 条件是统计学的灵魂.

条件概率的定义

定义 2.24 设 $(\Omega, \mathcal{F}, P)$ 为概率空间. $A \in \mathcal{F}$, $B \in \mathcal{F}$, 若 $P(B) > 0$, 则称

$$P(A|B) = \frac{P(AB)}{P(B)}$$

为"在已知 B 发生的情况下，A 发生的 条件概率", 简称条件概率.

注 2.7 条件概率的两种计算方法:

- 原则性方法: $P(A|B) = \dfrac{P(AB)}{P(B)}$
- 把 B 作为样本空间看待 (经常显得非常方便): $P(A|B) = \dfrac{|AB|}{|B|}$

条件概率的性质

定理 2.11 条件概率是概率，即若设 $P(B) > 0$, 则

- 非负性: $P(A|B) \geq 0, \forall A \in \mathcal{F}$;
- 规范性: $P(\Omega|B) = 1$;
- 可列可加性：若 $A_n \in \mathcal{F}, n = 1, 2, \cdots$，互不相容，则

$$P(\cup_{n=1}^{\infty} A_n|B) = \sum_{n=1}^{\infty} P(A_n|B)$$

证明: 由条件概率的定义易证非负性与规范性。下面说明可列可加性。假设 $A_n, n = 1, \cdots$，互不相容，则 $A_nB, n = 1, \cdots$，也互不相容。从而

$$\begin{aligned} P(\cup_{n=1}^{\infty} A_n|B) &= \frac{P((\cup_{n=1}^{\infty} A_n)B)}{P(B)} = \frac{P(\cup_{n=1}^{\infty}(A_nB))}{P(B)} \\ &= \sum_{n=1}^{\infty} \frac{P(A_nB)}{P(B)} = \sum_{n=1}^{\infty} P(A_n|B). \end{aligned}$$

乘法公式

定理 2.12 (乘法公式)

- 若 $P(A) > 0$, 则

$$P(AB) = P(A)P(B|A) \tag{8}$$

- 若 $P(A_1A_2\cdots A_{n-1}) > 0$, 则

$$P(A_1A_2\cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1}) \tag{9}$$

证明: 由条件概率的定义，移项即得 (8) 式，下证 (9). 因为

$$P(A_1) \geq P(A_2) \geq \cdots \geq P(A_1A_2\cdots A_{n-1}) > 0$$

故 (9) 式中条件概率均有意义。从而由两个事件的乘法公式可得

$$\begin{aligned}
&P(A_1A_2\cdots A_n) = P(A_1A_2\cdots A_{n-1})P(A_n|A_1A_2\cdots A_{n-1}) \\
&= P(A_1\cdots A_{n-2})P(A_{n-1}|A_1\cdots A_{n-2})P(A_n|A_1A_2\cdots A_{n-1}) \\
&= \cdots = P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1})
\end{aligned}$$

条件概率的例子

例 2.42 (两张牌) 洗好一副标准扑克后。从中随机抽取两张牌，无放回地一次抽一张。设 A 事件表示第一张牌为红桃，事件 B 表示第二张牌为红色。求 $P(A|B)$ 和 $P(B|A)$.

解: 由题意易知

$$P(A\cap B)=\frac{13\cdot 25}{52\cdot 51}=\frac{25}{204}$$
$$P(B)=\frac{26\cdot 51}{52\cdot 51}=\frac{1}{2}$$
$$P(A)=\frac{1}{4}$$

从而，

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{25/204}{1/2}=\frac{25}{102}$$
$$P(B|A)=\frac{P(B\cap A)}{P(B)}=\frac{25/204}{1/4}=\frac{25}{51}$$

条件概率的注

- 注意哪些事件放在竖线的哪一边是非常重要的，具体来说就是 $P(A|B) \neq P(B|A)$.
- 无论 $P(A|B)$ 还是 $P(B|A)$ 都是有意义的 (直观上或数学上):
  - 牌抽取的时间顺序并不能决定出现何种条件概率.
  - 在计算条件概率时，我们考虑的是一个事件给另一个事件带来的信息，而不是一个事件是否导致了另一个事件.
- 此外，也可以通过条件概率的直接解释得出 $P(B|A) = 25/51$:
  - 如果第一张抽的牌为红桃，那么剩下的牌就由 25 张红色牌和 26 张黑色牌组成 (所有牌被下一次抽中的可能性是相同的)
  - 所以抽取一张红牌的条件概率是 $25/(25+26) = 25/51$.

条件性思维谬误：控方证人的错误

例 2.43 (控方证人的错误) 1988 年，Sally Clark 由于她的两个孩子在出生不久便死亡，因而被指控谋杀幼童。在审讯期间，控方的一个专家证人证实

- 新生儿因婴儿猝死综合症 (SIDS) 而死亡的概率为 $1/8500$
- 所以两个新生儿由于婴儿猝死综合症死亡的概率为 $(1/8500)^2$, 大约为 7300 万分之一
- 因此，他认为 Clark 清白的概率仅为 7300 万分之一.

解: 这个推理过程至少有两个问题

- 一个家庭内部成员之间死于 SIDS 是否相互独立？
  - 家庭内部成员之间死于 SIDS 相互独立时，“第一个孩子死于 SIDS” 且 “第二个孩子也死于 SIDS” 的概率是相应的两个事件概率相乘
  - 如果遗传因素或其他家庭特有的风险因素导致某些家庭内的所有新生儿面临 SIDS 的风险增加，这种独立性就不再成立了.

条件性思维谬误：控方证人的错误

- 这个所谓的专家将两个不同的条件概率混淆了：$P(\text{清白} \mid \text{证据})$ 和 $P(\text{证据} \mid \text{清白})$ 是不一样的.
  - 专家 1 声称：如果在被告人是清白的情况下，两个孩子死亡的概率很低；那就是说 $P(\text{证据} \mid \text{清白})$ 非常小.
  - 但人们感兴趣的是，给定现在所有的证据 (孩子均死) 条件下，报告人仍清白的概率，即 $P(\text{清白} \mid \text{证据})$.
  - 由贝叶斯准则可知，

$$P(\text{清白} \mid \text{证据}) = \frac{P(\text{证据} \mid \text{清白})P(\text{清白})}{P(\text{证据})},$$

  - 所以为了计算 $P(\text{清白} \mid \text{证据})$, 这里需要考虑 $P(\text{清白})$, 也就是被告清白的先验概率。这个概率是很高的；
  - 虽然 SIDS 造成两个婴儿死亡是很罕见的，但是蓄意杀害两个婴儿的情况也很少见！
  - 基于现有证据的后验概率是对很低的 $P(\text{证据} \mid \text{清白})$ 和很高的 $P(\text{清白})$ 的一个平衡。专家的结果 $(1/8500)^2$ 是有问题的，它只是整个计算式中的一部分.

条件概率的例子：男孩女孩概率问题

例 2.44 (年长的是女孩 vs 至少一个女孩) 某家庭有两个孩子，已知至少有一个是女孩。两个孩子都是女孩的概率是多少？如果条件改为年长的孩子是女孩，那么两个都是女孩的概率又是多少？

解: 假设每个孩子都是女孩和男孩的可能性相同且不相关，那么

$$P(\text{都是女孩} \mid \text{至少有一个是女孩}) \tag{10}$$

$$= \frac{P(\text{都是女孩，至少有一个是女孩})}{P(\text{至少有一个是女孩})} = \frac{1/4}{3/4} = 1/3 \tag{11}$$

$$P(\text{都是女孩} \mid \text{年长的是女孩}) \tag{12}$$

$$= \frac{P(\text{都是女孩，年长的是女孩})}{P(\text{年长的是女孩})} = \frac{1/4}{1/2} = 1/2 \tag{13}$$

男孩女孩概率问题续：随机的一个孩子是女孩

例 2.45 某家庭有两个孩子。随机遇到其中的一个，发现是女孩。给定这个信息后，两个孩子都是女孩的概率是多少？假设随机遇到两个孩子的可能性相同，且与性别无关.

解:

- 直观来看，结果应为 $1/2$.
- 令 G_1, G_2, G_3 分别表示年长、年幼、随机的孩子是女孩这三个事件。由对称性可得: $P(G_1) = P(G_2) = P(G_3) = 1/2$
- 根据朴素概率的定义，或者独立性，可得: $P(G_1 \cap G_2) = 1/4$
- 因此，$P(G_1 \cap G_2 | G_3) = P(G_1 \cap G_2 \cap G_3)/P(G_3) = 1/2$
- 又因为 $G_1 \cap G_2 \cap G_3 = G_1 \cap G_2$, 所以概率为 $1/2$.

注 2.8 假设一个强制性法律规定：如果一个男孩有姐妹则禁止他走出家门。那么这时“随机遇到的孩子是女孩”就等价于“ 至少有一个孩子是女孩”

男孩女孩概率问题续：冬天出生的女孩

例 2.46 某家庭有两个孩子，给定条件至少一个是女孩且在冬天出生，求两个孩子都是女孩的概率。假设四个季节出生的可能性相同且性别和季节是相互独立的.

解: 由条件概率的定义，可得:

$$
\begin{aligned}
&P(\text{两个都是女孩} \mid \text{至少有一个是冬天出生的女孩}) \\
&= \frac{P(\text{两个都是女孩，至少有一个是冬天出生的女孩})}{P(\text{至少有一个是冬天出生的女孩})}
\end{aligned}
$$

由于指定的孩子是在冬天出生的女孩的概率为 1/8, 所以,

$$P(\text{至少有一个是在冬天出生的女孩}) = 1 - (7/8)^2.$$

男孩女孩概率问题续：冬天出生的女孩

利用性别和季节是相互独立的假设，得到:

$$
\begin{aligned}
&P(\text{两个都是女孩，至少有一个是冬天出生的女孩})\\
&= P(\text{两个都是女孩，至少有一个是冬天出生的})\\
&= (1/4)P(\text{至少有一个是冬天出生的女孩})\\
&= (1/4)(1 - P(\text{所有孩子都不是在冬天出生的})
\end{aligned}
$$

合在一起得到，

$$P(\text{两个都是女孩} \mid \text{至少有一个是冬天出生的女孩}) = 7/15$$

结绳问题

例 2.47 n 根绳 $2n$ 个头两两相接，求事件 $A=\{$恰好结成n个圈$\}$ 的概率.

- 设想 $2n$ 个头排成一行，规定将第 $2k-1$ 个头与第 $2k$ 个端头相接;
- 令 B_i 表示第 i 根绳的头与尾恰好相接，则 $A=B_1B_2\cdots B_n$;
- 若以 $n(A)$ 表示事件 A 所包含的样本点个数，则易知

$$n(\Omega)=(2n)!, n(B_1)=2n(2n-2)!\Rightarrow P(B_1)=\frac{n(B_1)}{n(\Omega)}=\frac{1}{2n-1};$$

- $P(B_2|B_1)$ 可看作 $n-1$ 根绳某根绳头尾相接的概率，类比 n 根绳情形可得 $P(B_2|B_1)=\dfrac{1}{2(n-1)-1}=\dfrac{1}{2n-3}$;
- 同理可得

$$P(B_k|B_1B_2\cdots B_{k-1})=\frac{1}{2[n-(k-1)]-1}=\frac{1}{2n-2k+1}, 3\le k\le n;$$

- 利用乘法公式可得

$$\begin{aligned}P(A)&=P(B_1B_2\cdots B_n)=P(B_1)P(B_2|B_1)\cdots P(B_n|B_1\cdots B_{n-1})\\&=\frac{1}{(2n-1)!!}\end{aligned}$$

取球问题

例 2.48 在计算机中输入程序，让它自动完成如下操作:

- 在 $1-\frac{1}{2^n}$ 时刻，往盒中放入标号 $10(n-1)+1 \sim 10n$ 的 10 个球，同时取出标号为 $10(n-1)+1$ 的球，$n \geq 1$;
- 在 $1-\frac{1}{2^n}$ 时刻，往盒中放入标号 $10(n-1)+1 \sim 10n$ 的 10 个球，同时取出标号为 n 的球，$n \geq 1$;
- 在 $1-\frac{1}{2^n}$ 时刻，往盒中放入标号 $10(n-1)+1 \sim 10n$ 的 10 个球，同时随机地从盒中取出一个球，$n \geq 1$.

则在时刻 1, 盒中的球数结果如下

- 盒子中有无穷多个球;
- 盒子变为空的;
- 盒子变为空的概率等于 1?

取球问题

解:

- 记 E={在时刻 1 时盒子变空 };
- $\overline{E}$={在时刻 1 时盒中有球未被取出 };
- 记 A_k={在时刻 1 时 k 号球仍在盒中未被取出 }，则 $\overline{E} = \bigcup\limits_{k=1}^{\infty} A_k$;
- 由概率的次可加性知

$$P(\overline{E}) = P(\bigcup_{k=1}^{\infty} A_k) \leq \sum_{k=1}^{\infty} P(A_k);$$

- 为证 $P(\overline{E}) = 0$，只需证明 $P(A_k) = 0, k \geq 1$;
- 由于证法类似，仅以证明 $P(A_1) = 0$ 为例;

取球问题

- 记 B_n={在 $1-\dfrac{1}{2^n}$ 时刻 1 号球未被取出 }，易知 $A_1=\bigcap\limits_{n=1}^{\infty}B_n$;
- 令 $C_m=\bigcap\limits_{n=1}^{m}B_n$, 则有

$$C_m=\bigcap_{n=1}^{m}B_n\supset\bigcap_{n=1}^{m+1}B_n=C_{m+1},\text{ 且 }\lim_{m\to\infty}C_m=A_1;$$

- 由概率的上连续性知

$$P(A_1)=P\big(\lim_{m\to\infty}C_m\big)=\lim_{m\to\infty}P\big(C_m\big);$$

- 下面求 $P(C_m)$, 即 $P\big(\bigcap_{n=1}^{m}B_n\big)=\prod_{n=1}^{m}\frac{9n}{9n+1}=\prod_{n=1}^{m}\left(1-\frac{1}{9n+1}\right)$;
  - B_1: 在 1/2 时刻，盒中 10 个球，1 号球未被取出，故 $P(B_1)=\frac{9}{10}$;
  - B_2: 在 3/4 时刻，盒中 19 个球，1 号球未被取出，故 $P(B_2|B_1)=\frac{18}{19}$;
  - B_n: $P(B_n|B_1B_2\cdots B_{n-1})=\frac{9n}{9n+1}$.
- $P(A_1)=\prod_{n=1}^{\infty}\left(1-\frac{1}{9n+1}\right)=0$. 同理可证 $P(A_k)=0$, $k=2,3,\cdots$.

罐子模型（波利亚模型）

例 2.49 设罐中有 b 个黑球，r 个红球，每次随机的取出一球，取出后将原球放回，还加进 c 个同色球和 d 个异色球。记 $B_i := \{第i次取出的是黑球\}, R_j := \{第j次取出的是红球\}$. 若连续从罐子中取出三个球，其中有两个红球，一个黑球，则由乘法公式可得

$$
\begin{aligned}
P(B_1R_2R_3) &= P(B_1)P(R_2|B_1)P(R_3|B_1R_2) \\
&= \frac{b}{b+r}\cdot\frac{r+d}{b+r+c+d}\cdot\frac{r+d+c}{b+r+2c+2d} \\
P(R_1B_2R_3) &= P(R_1)P(B_2|R_1)P(R_3|R_1B_2) \\
&= \frac{r}{b+r}\cdot\frac{b+d}{b+r+c+d}\cdot\frac{r+d+c}{b+r+2c+2d} \\
P(R_1R_2B_3) &= P(R_1)P(R_2|R_1)P(B_3|R_1R_2) \\
&= \frac{r}{b+r}\cdot\frac{r+c}{b+r+c+d}\cdot\frac{b+2d}{b+r+2c+2d}
\end{aligned}
$$

显然以上概率与黑球在第几次抽取有关.

罐子模型（波利亚模型）

- 当 $c=-1, d=0$ 时，即为不返回抽样。此时前次抽取结果会影响后次抽取结果，但只要抽取的黑球与红球个数确定，则概率不依赖其抽出球的次序，都是一样的.

$$P(B_1R_2R_3)=P(R_1B_2R_3)=P(R_1R_2B_3)=\frac{br(r-1)}{(b+r)(b+r-1)(b+r-2)}.$$

- 当 $c=0, d=0$ 时，即为返回抽样。此时前次抽取结果不会影响后次抽取结果，故上述三个概率相等且都等于

$$P(B_1R_2R_3)=P(R_1B_2R_3)=P(R_1R_2B_3)=\frac{br^2}{(b+r)^3}.$$

- 当 $c>0, d=0$ 时，称为传染病模型。此时每次取出球后会增加下一次取出同色球的概率，或换言之，每发现一个传染病患者，以后都会增加再传染的概率。与前两种情况一样，三个概率都等于

$$P(B_1R_2R_3)=P(R_1B_2R_3)=P(R_1R_2B_3)=\frac{br(r+c)}{(b+r)(b+r+c)(b+r+2c)}.$$

罐子模型（波利亚模型）

- 从上面的结果可以看出，只要 $d=0$，以上三个概率都相等，即只要抽取的黑球与红球的个数确定，则概率不依赖于抽出黑红球的次序.
- 当 $c=0, d>0$ 时，称为安全模型。此模型可解释为：每当事故发生了 (当红球被取出)，安全工作就抓紧一些，下次再发生事故的概率就会减少，而当事故没有发生时 (黑球被取出)，安全工作就放松一些，下次再发生事故的概率就会增大，此时，上述三个概率分别为

$$
\begin{aligned}
P(B_1R_2R_3) &= \frac{b}{b+r}\cdot\frac{r+d}{b+r+d}\cdot\frac{r+d}{b+r+2d} \\
P(R_1B_2R_3) &= \frac{r}{b+r}\cdot\frac{b+d}{b+r+d}\cdot\frac{r+d}{b+r+2d} \\
P(R_1R_2B_3) &= \frac{r}{b+r}\cdot\frac{r}{b+r+d}\cdot\frac{b+2d}{b+r+2d}
\end{aligned}
$$

罐子模型（波利亚模型）

例 2.50 设罐中有 b 个黑球，r 个红球，每次随机取出一球后将原球放回并加进 c 个同色球，如此反复进行。试证明：在前 $n = n_1 + n_2$ 次取球中，取出了 n_1 个红球和 n_2 个黑球的概率为

$$C_n^{n_1} \frac{a(a+c)(a+2c)\cdots(a+n_1c-c)b(b+c)(b+2c)\cdots(b+n_2c-c)}{(a+b)(a+b+c)(a+b+2c)\cdots(a+b+nc-c)}.$$

两个事件的独立性

- 直观上来说，两个事件的独立性是指：一个事件的发生不影响另一个事件的发生。比如在掷两颗骰子的实验中，第一颗骰子的点数和第二颗骰子的点数是互不影响的.
- 从概率的角度看，$P(A|B)$ 与 $P(A)$ 的差别在于：事件 B 的发生改变了事件 A 发生的概率，也即事件 B 对事件 A 有某种影响。故如果 A 与 B 的发生是相互不影响的，则有

$$P(A|B) = P(A), \quad P(B|A) = P(B).$$

上面两式均等价于

$$P(AB) = P(A)P(B) \tag{14}$$

- 注意到 (14) 式对 $P(B) = 0$ 或 $P(A) = 0$ 仍然成立，为此，我们用 (14) 作为两个事件相互独立的定义.

两个事件独立性的定义

定义 2.25 如果对事件 A 与 B 有

$$P(AB) = P(A)P(B)$$

成立，则称 事件 A 与 B 相互独立, 简称 A 与 B 独立. 否则称 A 与 B 不独立或相依.

注 2.9

- 零概率事件 E 与任何事件相互独立，特别的，不可能事件与任何事件相互独立
- 非零概率互不相容事件，一定不独立
- 非零概率相互独立事件，一定相容

如何确定事件的独立性:

- 实际问题中，两个事件的独立大多根据经验及相互有无影响的直观性来判断.
- 但对于较复杂事件，有无相互影响并不是很直观，则需要验证 (14) 式是否成立来说明独立性.

对立事件的独立性

定理 2.13 若 A 与 B 独立，则 A 与 $\overline{B}$ 独立，$\overline{A}$ 与 B 独立，$\overline{A}$ 与 $\overline{B}$ 独立.

证明: 我们仅证 $P(A\overline{B}) = P(A)P(\overline{B})$, 其余类似可证.

$$\begin{aligned} P(A\overline{B}) &= P(A-B) = P(A) - P(AB) = P(A) - P(A)P(B) \\ &= P(A)(1-P(B)) = P(A)P(\overline{B}) \end{aligned}$$

对于上面的定理直观上来理解也是很容易的：因 A, B 独立，故 A 的发生不影响 B 的发生，从而也不会影响 B 的不发生，$\cdots$

三个事件的独立性

定义 2.26 设 A, B, C 三个事件，如果有

$$\left.\begin{aligned} P(AB) &= P(A)P(B) \\ P(AC) &= P(A)P(C) \\ P(BC) &= P(B)P(C) \end{aligned}\right\} \tag{15}$$

$$P(ABC) = P(A)P(B)P(C) \tag{16}$$

则称 A, B, C 相互独立。如果仅有 (15) 式成立，则称 A, B, C 两两独立.

两两独立与相互独立的关系

- 由定义可知，三个事件相互独立必能推出两两独立.
- 但两两独立未必能推出相互独立，即 (15) 式成立，不一定能推出 (16) 成立
  - 考虑独立投掷两枚均匀硬币的随机试验，设事件 A 代表第一枚硬币正面朝上，事件 B 代表第二枚硬币正面朝上，事件 C 表示两枚硬币结果相同。易知: A B 和 C 是两两独立，但

$$P(A \cap B \cap C) = 1/4 \neq 1/8 = P(A)P(B)P(C).$$

  - 考虑一个均匀的正四面体，第一二三面分别染上红 / 白 / 黑色，第四面同时染上红白黑色。现在以 A, B, C 分别记投一次四面体出现红，白，黑色朝下的事件。则易有

$$\begin{aligned} P(A) = P(B) = P(C) &= 1/2 \\ P(AB) = P(BC) = P(AC) &= 1/4 \\ P(ABC) &= 1/4 \end{aligned}$$

两两独立与相互独立的关系

- 反之，如果 (16) 成立，是否能推出 (15) 成立？
  - 考虑一个均匀的正八面体，第 1, 2, 3, 4 面染上红色，第 1, 2, 3, 5 面染上白色，第 1, 6, 7, 8 面染上黑色。现在以 A, B, C 分别记投一次八面体出现红，白，黑色朝下的事件，则

$$
\begin{aligned}
P(A) = P(B) = P(C) = 4/8 = 1/2 \\
P(ABC) = 1/8 \\
P(AB) = 3/8 \neq 1/4 = P(A)P(B)
\end{aligned}
$$

三个以上事件的独立性

定义 2.27 设 $(\Omega, \mathcal{F}, P)$ 为一概率空间，$A_1, A_2, \cdots, A_n \in \mathcal{F}$, 对任意的 $1 \le k \le n$ 及任意的 $1 < j_1 < j_2 < \cdots < j_k \le n$ 均有:

$$P(A_{j_1}A_{j_2}\cdots A_{j_k}) = P(A_{j_1})P(A_{j_2})\cdots P(A_{j_k}) \tag{17}$$

成立，则称事件 $A_1, \cdots, A_n$ 相互独立.

- (17) 式共有多少个等式?

$$\left.\begin{array}{l} P(A_{j_1}A_{j_2}) = P(A_{j_1})P(A_{j_2}) \\ P(A_{j_1}A_{j_2}A_{j_3}) = P(A_{j_1})P(A_{j_2})P(A_{j_3}) \\ \quad\vdots \\ P(A_1A_2\cdots A_n) = P(A_1)P(A_2)\cdots P(A_n) \end{array}\right\} C_n^2 + \cdots + C_n^n = 2^n - n - 1$$

- 从定义可以看出，n 个相互独立事件中的任取 $m(2 \le m \le n)$ 个事件仍是相互独立的，而且任意一部分与另一部分也是独立的.
- 类似于前面的证明，将相互独立事件中的任一部分换为对立事件，所得诸事件仍是相互独立的.

任意多个事件相互独立

定义 2.28 设 $(\Omega, \mathcal{F}, P)$ 为一概率空间，每个 $t \in T$ 有 $A_t \in \mathcal{F}$. 称 $\{A_t, t \in T\}$ 为独立事件族，如果对 T 的任意有限子集 $\{t_1, t_2, \cdots, t_s\}$, 事件 $A_{t_1}, A_{t_2}, \cdots, A_{t_s}$ 相互独立.

例 2.51 $\mathcal{F}$ 中事件序列 $\{A_n\}$ 为相互独立的充分必要条件是，任意 $n \geq 1$, 事件 $A_1, A_2, \cdots, A_n$ 独立；等价的，任意有限个自然数 $k_1, \cdots, k_s$ 有

$$P(A_{k_1} A_{k_2} \cdots A_{k_s}) = P(A_{k_1}) P(A_{k_2}) \cdots P(A_{k_s})$$

条件独立

定义 2.29 称事件 A 和 B 是关于 E 条件独立的，如果

$$P(A \cap B|E) = P(A|E)P(B|E)$$

- 两个事件可以在给定事件 E 的条件下是条件独立的，但它们不是独立的.
- 两个事件可以是独立但却不是关于 E 条件独立的.
- 两个事件可以关于 E 条件独立但关于 E^c 不存在条件独立.

条件独立不意味着独立

例 2.52 假设有两枚硬币，一枚是均匀的，一枚是不均匀的。从两枚硬币中随机的选一枚硬币并进行抛掷 2 次，若令

$$
\begin{aligned}
F &:= \{\text{选取的硬币是均匀的}\} \\
A_1 &:= \{\text{第一次投掷硬币正面朝上}\} \\
A_2 &:= \{\text{第二次投掷硬币正面朝上}\}
\end{aligned}
$$

则给定 F 为条件，A_1 和 A_2, 是相互独立的，A_1 和 A_2 并不是无条件独立的，因为 A_1 会提供关于 A_2 的信息.

独立不意味着条件独立

例 2.53 假设只有我的朋友 Alice 和 Bob 给我打过电话。每天他俩都会相互独立地决定是否给我打电话。若令

$$A := \{\text{Alice 给我打电话}\}$$
$$B := \{\text{Bob 给我打电话}\}$$
$$R := \{\text{听到电话铃响}\}$$

- 显然，A 和 B 是无条件独立的.
- 但现在我听到一声电话铃响，那 A 和 B 就不再独立了：如果这个电话不是 Alice 打的，那就肯定是 Bob 打的。从而

$$P(B|R) < 1 = P(B|A^cR) = \frac{P(BA^cR)}{P(A^cR)} = \frac{P(BA^c|R)}{P(A^c|R)}.$$

显然: $P(BA^c|R) > P(B|R)P(A^c|R)$

- B 与 A^c 关于 R 不条件独立，A, B 亦是如此.

给定 E 条件独立 *vs* 给定 E^c 条件独立

例 2.54 假设有两种课程：好的课程和坏的课程。在好的课上，如果你努力，就很有可能得到 A. 在坏的课上，教授随机分配给学生分数，而不管他们是否努力。若令

$$G := \{\text{这个课程是好的}\}$$
$$W := \{\text{你学习努力}\}$$
$$A := \{\text{你的得分为}A\}$$

这时，给定 G^c, A 和 W 是条件独立的，但给定 G, A 和 W 却不是独立的！

随机试验的独立性

- 先考虑两个随机试验，假定 $(\Omega_i, \mathcal{F}_i, P_i), i = 1, 2$ 为第 i 个随机试验对应的概率空间。按照之前独立性的理解，两个试验的独立性应当叙述为:

$$\text{对任何的}A_i \in \mathcal{F}_i, i = 1, 2, A_1\text{与}A_2\text{同时}$$
$$\text{发生的概率等于它们各自概率之乘积}$$

- 两个不妥:
  - "A_1 与 A_2 同时发生" 应当是这两个事件的交，但它们分别是两个样本空间 Ω_1, Ω_2 的子集，无法进行运算;
  - 两个概率空间有各自的概率 P_1, P_2, 但涉及两个度验，命题中 "同时发生的概率" 既不能用 P_1 也不能用 P_2 来度量.
- 解决方法：构造可以同时描述两个试验的新概率空间 $(\Omega, \mathcal{F}, P)$.

乘积空间的构造

- 样本乘积空间: $\Omega := \Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \omega_1 \in \Omega_1 且 \omega_2 \in \Omega_2\}$;
- 乘积 σ- 代数 $\mathcal{F}_1 \times \mathcal{F}_2$:
  - 可测矩形集类: $\mathcal{G} := \{A_1 \times A_2 : A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}$;
  - $\mathcal{F}_1 \times \mathcal{F}_2 := \sigma(\mathcal{G})$
- 乘积概率测度:
  - 对于每个可测矩形 $A_1 \times A_2 \in \mathcal{G}$ 定义如下集函数:

$$P(A_1 \times A_2) = P_1(A_1)P_2(A_2), \quad A_i \in \mathcal{F}_i, i = 1, 2. \tag{19}$$

  - 理论上可以证明如上定义在 $\mathcal{G}$ 上的集函数 P 可唯一地扩张为 $\mathcal{F}_1 \times \mathcal{F}_2$ 上的概率测度，称之为 P_1 与 P_2 的乘积 (概率) 测度.
- 在上述乘积测度下

$$\begin{aligned} &P(A_1 \times \Omega_2) = P_1(A_1), \quad P(\Omega_1 \times A_2) = P_2(A_2) \\ &P\big((A_1 \times \Omega_2) \cap (\Omega_1 \times A_2)\big) = P(A_1 \times A_2) \\ &\quad = P_1(A_1)P_2(A_2) = P(A_1 \times \Omega_1)P(\Omega_1 \times A_2) \end{aligned}$$

- $(\Omega_i, \mathcal{F}_i, P_i)$ 的独立性取决于乘积样本空间 $\Omega_1 \times \Omega_2$ 上的概率是否取作由 (19) 所确定的乘积测度

n 个试验相互独立的定义

定义 2.30 设有 n 个随机试验，第 i 个试验的概率空间为 $(\Omega_i, \mathcal{F}_i, P_i)$, $i = 1, \cdots, n$. 代表这 n 个试验的乘积样本空间 $\Omega = \Omega_1 \times \cdots \times \Omega_n$, $\mathcal{F} = \mathcal{F}_1 \times \cdots \times \mathcal{F}_n = \sigma(\mathcal{G})$, 其中 $\mathcal{G}$ 为形如 $B_1 \times \cdots \times B_n (B_i \in \mathcal{F}_i)$ 的可测矩形的全体。如果 $(\Omega, \mathcal{F})$ 上的概率测度 P 是 $P_1, \cdots, P_n$ 的乘积测度，即对任何 $B_1 \times \cdots \times B_n \in \mathcal{G}$ 满足

$$P(B_1 \times \cdots \times B_n) = P_1(B_1) \cdots P(B_n),$$

则称这 n 个度验独立. 如果现设

$$\Omega_i \equiv \Omega_0, \mathcal{F}_i \equiv \mathcal{F}_0, P_i \equiv P_0, i = 1, \cdots, n,$$

即 n 个试验有相同的概率空间，则称它们为 n 个 (重) 独立重复试验. 如果在 n 个独立重复实验中，每次试验的可能结果为两个:A 或 $\overline{A}$, 则称这种试验为 n 重伯努利试验.

彩票问题

例 2.55 某彩票每周开奖一次，每次提供十万分之一的中奖机会，且各周开奖是独立的。若你每周买一张彩票，坚持十年（每年按 52 周计算），试求未中奖的概率.

解: 依假设，每次中奖的概率为 10^{-5}, 于是每次不中奖的概率是 $1-10^{-5}$. 另外十年一共购买 520 次彩票，而每次开奖都是独立的，相当于进行了 520 次独立重复试验. 若记 A_i 为“第 i 次开奖不中奖”，则 $A_1, \cdots, A_{520}$ 相互独立，从而

$$P(A_1A_2\cdots A_{520}) = (1-10^{-5})^{520} = 0.9948$$