概 率 论

（第二版）

应坚刚　何　萍　编著

博学·数学系列

复旦大学出版社
www.fudanpress.com.cn

内容提要

本书以概率空间和随机变量为主线，力求将概率论的直观思想同严密的数学逻辑结合起来，主要讲述概率论和随机变量的一些基本理论、经典问题，包括一些重要的分布、数学期望、条件概率和独立性、随机变量的各种收敛性以及相互间关系、大数定律、特征函数的方法、中心极限定理等. 本书可作为高等学校理科各专业和其他相关专业的教材，亦可供有关科研人员参考.

第二版前言

讲一门课, 写一本教材, 首先思考的是选取什么内容. 我作为概率论的专业教师, 自工作以来一直教概率论这门课, 送走了一届又一届的学生, 但是每次上课的时候还是不由自主地考虑这门课应该包括哪些内容以及怎么讲这些内容, 所以从第一版出版后, 讲义还是在不断地修改补充删减, 力求让教材更贴合我对大学概率论课程的理解. 虽然每次改动都不大, 但这么多年过去, 积少成多, 感觉到了该出一个新版本的时候了.

学生在概率论这门课程上主要要学些什么? 首先, 我们将讲一些直观经典的例子, 以让读者来感受什么是随机现象, 什么是概率, 也就是说了解这门学科的背景, 这有助于读者理解这门课程为什么要这样来展开. 概率论在17世纪中叶产生到20世纪初建立公理体系而真正成为一个数学分支, 经过了约250年历史, 在这段时间里, 概率论可以说是自由发展着, 每个人都感觉着概率论直观的一面而把它记录下来, 已经积累了丰富的内容. 除了古典概率外, 还知道各种分布, 比如二项分布、Poisson 分布、几何分布、均匀分布、正态分布、指数分布等现在教科书上必讲的经典分布, 也知道它们怎么作为数学模型用在一些具体的问题上, 当然也知道分布的一些基本的数字特征, 比如数学期望和方差、相关系数等, 也知道随机变量独立性应该怎么来描述或者它直观上是什么意思. 令人惊讶的是, 早在1713年, Bernoulli 就知道频率会趋于概率的称之为大数定律的重要思想, 这是大样本统计学的基石; 同样令人惊讶的是在1730 年, De Moivre 进一步精细化了大数定律, 证明了二项分布中心化后的分布趋于正态分布, 发现了概率论中最重要的分布. 以上这些内容是每个学完大学概率论的学生应该了解的基本知识.

在了解了这些基本知识之后, 真正对数学感兴趣的学生必然想要弄清楚这个学科的理论体系, 想要弄清楚每个直观陈述背后的数学含义和逻辑体系, 因此我们将同时介绍概率论的公理体系, 让学生们知道尽管概率论本身研究随机现象, 但作为数学的分支, 概率论有严密的公理和逻辑作为基础, 是一个确定的严格意义上的学科, 没有任何含糊的地方, 它能够比较恰当地解释某些随机现象并对实际的问题有一定的指导意义. 读者应该记住, 数学通常只是把实际问题简单化, 抽象化, 最完美的数学模型也不能完全重现实际问题, 换句话说, 实际世界和数学模型之间还有遥远的距离. 公理化之后, 概率论的研究对象就是概率空间和随机变

量及其分布. 这些研究对象都有很好的实际背景, 但是其数学意义是来自公理体系和逻辑推理的, 直觉可以帮助解决数学问题, 但不能作为解决数学问题的依据. 随后我们将介绍概率论中最重要的一些分布类型, 它们就如同数学分析中重要的函数类型一样, 概率论中最重要的独立性和条件概率也会在数学上给予定义. 接着, 我们会介绍概率研究中的一些数学方法, 比如随机序列的收敛性与分布列的弱收敛性, 然后我们就可以理解概率论中最早也是最重要的一个定理: Bernoulli 大数定律, 它算是概率这个概念的直观意义所在, 也是随机现象中最重要的一个规律. 在这里, 我们还补充了作为大数定律序列中最重要的Kolmogorov 强大数定律及其证明, 其方法漂亮得令人惊叹. 最后, 为了证明概率论另外一个重要定理: 中心极限定理, 我们详细介绍了分析中的最常用, 最重要的方法: 特征函数方法. 这个方法本身没有多少概率气息, 但它在概率论研究中的作用是不可替代的. 最后一章比较抽象, 是为概率论理论研究准备的, 教师可以视课时和学生情况决定是否讲授.

当然以上两部分内容并不是完全分裂的, 在这一版教材中也是交错在一起的. 总体说来, 本次修订相对于上一版还是有较大的改动, 章节的编排和顺序都有所改变, 但是基本的结构内容并没有太多变化. 我们把单调类方法和条件数学期望合在一起把它放在最后一章, 这个内容对于读者进一步学习随机过程和随机分析是绝对重要的. 加星号的补充内容不在大纲或者考试范围之内, 有兴趣的学生可以选择阅读. 加星号的习题也是相对比较难的. 另外参考书目中添加了王梓坤先生不久前再版的经典教材 [13], 它也是我大学学概率论时使用的教材, 印象深刻. 如果对概率论理论感兴趣且想进一步学习测度化概率论, 随机过程和随机分析, 那么可以参考汪嘉冈教授的教材 [14] 和我与合作者编写的研究生教材 [15].

从课时安排看, 本教材的总课时是48个, 第一章初等概率论约6个课时, 第二章随机变量及其分布约3个课时, 第三章条件概率全概率公式约4个课时, 第四章数学期望内容较多, 约需要7个课时, 第五第六两章一维与多维随机变量以及常用分布约10个课时, 第七章三种收敛性约4个课时, 第八章特征函数约4个课时, 第九章中心极限定理约2个课时, 第十章单调类方法独立随机序列存在性与条件期望约8个课时.

尽管讲义着重的是理论, 但更重要的是直觉, 因为理论可以传承, 直觉却只能自己感悟. 本讲义在复旦大学作为数学专业概率论课的教材已经使用多年, 虽经不断的修改补充, 但一定还会有错误或者不妥之处, 望读者见谅.

试读结束：需要全本请在线购买： www.ertongbook.com

应坚刚 12/21/2013
于南京雍园时光青旅

补充 1: 感谢本院2013级学生王博文和吴佩学, 他们仔细阅读了讲义并提醒我改正了讲义中的许多错误.

补充 2: 感谢复旦大学出版社编辑范仁梅女士和陆俊杰先生仔细阅读并改正许多文字表达的错误, 为本书新版的出版提供帮助.

第一版前言

概率论是大学数学各专业的必修课, 作者讲授概率课程多年, 选用过多种不同的概率论教材, 但经常有学生反映为他们讲授的概率论不像是数学, 作为教师, 我们在教授时也有同样的感觉, 因为其中常常有许多概念不能严格地定义, 许多结果不能严格地证明, 原因是严格地定义和证明需要测度论的语言, 而测度论的体系又不可能在概率论课程中详细介绍, 这促使作者尝试编写现在这本教材, 努力用最简单的语言讲清楚每个概念和严格证明每个结果.

本教材分两章, 第一章讲授概率空间和数学期望的公理体系与基本理论, 大多数涉及的例子都是离散的和古典的, 所用的方法是初等的. 第二章讲授连续型随机变量和分布理论, 介绍了收敛的概念及其性质、特征函数及其应用, 还简单介绍了大数定律与中心极限定理. 整个内容除了假设均匀分布的存在性外, 所涉及的每个结果都有严格的证明. 除最后一节介绍特征函数的内容时需要应用一些简单的复变函数知识, 阅读本教材所需的知识限制在数学分析和线性代数的范围内, 但很好的数学素养会对理解有很大的帮助. 我们在编写过程中遵循的宗旨是简明扼要, 期望学生在理解概率的直观与历史背景的同时认识到概率论是严密数学理论的一个分支. 书中还讲述了大量直观的经典例子, 它们是教材的重要组成部分, 期望读者通过这些例子来理解概率论的方法. 如果读者对经典的概率问题感兴趣, 可以进一步阅读 Feller [5], 从直觉和理论两方面来说这都是一部极其经典的概率参考书教材, 较新的 [6]也是不错的参考书. 如果读者对概率的数学理论有兴趣, Billingsley [1] 是一个很好的开始.

概率论是历史悠久且直观背景很强的领域, 但它成为数学的一个分支却还不到百年的历史. 概率论的严格公理体系是建立在测度论上的, 而测度论不能很好体现概率生动而直观的一面, 因此教师在讲授概率论和编写概率论教材时, 常常会陷入注重严格的逻辑体系或者注重直观背景这样两难的选择. 本教材的编写过程也是如此, 虽然我们非常努力地尝试把两者自然地连接起来, 取得我们所理解的某种意义上的平衡, 但是否能达成这一目标依然需要实践的检验. 本教材曾作为复旦大学数学系三年级第一学期部分专业每周4学时概率论课程的讲义试用过, 反响良好. 本教材不是一本通用教材, 希望它的特色适用于那些对概率的数学理论感兴趣的读者, 以帮助他们更好地理解概率论. 我们要感谢复旦大学数学系, 他

们在基金资助和课程安排方面的鼎力支持对于完成教材的编写是至关重要的; 感谢复旦大学出版社的范仁梅女士为本书顺利出版提供的帮助.

应坚刚 复旦大学

何萍 上海财经大学

目　录

第二版前言　i

第一版前言　v

第一章　初等概率论　1

1.1 概率简史 1

1.2 计数 3

1.3 古典概率问题 6

1.4 几何概率问题 21

1.5 随机取个自然数 (*) 23

第二章　概率空间与随机变量　31

2.1 集合 31

2.2 概率空间 32

2.3 随机变量与分布 39

第三章　条件概率与全概率公式　51

3.1 独立性 51

3.2 条件概率 55

3.3 全概率公式与 Bayes 公式 58

第四章　数学期望　68

4.1 期望的定义和性质 68

4.2 期望的计算公式 75

4.3 方差及其不等式 79

4.4 常见分布的期望 80

4.5 大数定律 84

第五章　连续型随机变量　93

5.1 可测性 93

5.2 分布函数的实现 95
5.3 密度函数 98

第六章 随机向量 **111**
6.1 随机向量及联合分布 111
6.2 均匀分布与正态分布 114
6.3 随机向量的函数的分布 118

第七章 随机序列的收敛 **134**
7.1 收敛的不同意义 134
7.2 强大数定律 138
7.3 Kolmogorov 不等式与强大数律 (*) 141
7.4 一致可积性 (*) 144
7.5 依分布收敛 147

第八章 特征函数 **156**
8.1 特征函数 156
8.2 唯一性定理 159
8.3 连续性定理 163

第九章 中心极限定理 **169**
9.1 DeMoivre-Laplace 的估计 (*) 169
9.2 独立同分布场合的中心极限定理 173
9.3 一般中心极限定理 (*) 174

第十章 单调类方法与条件期望 **179**
10.1 单调类方法 179
10.2 独立性 182
10.3 条件期望 186
10.4 鞅与鞅基本定理 (*) 196

参考文献 **202**

第一章　初等概率论

概率论的发展是从古典概率开始的, 理解概率论当然也要从古典概率开始. 古典概率是指我们熟知的那些掷硬币, 掷骰子, 摸球等游戏中产生的概率问题, 在如今的中学教材中也已经出现, 日常生活中更是常见. 它通常具有有限多种可能性且直观上是等概率这样两个特点. 古典概率的问题通常归结于集合计数问题, 我们学习的重点不是怎么解决这些问题, 而是怎么通过这些问题来理解概率论的产生.

§1.1　概率简史

概率论是数学的一个分支, 产生于人们对随机现象的认识与研究. 人类对随机现象的认识和利用应该已经有悠久的历史, 在许多场合都会用到随机现象, 比如赌博, 祭祀等. 特别是赌博, 深植于各种阶层各个民族的文化之中, 正是因为人类无法破解这些随机现象, 赌博作为一种娱乐或者可以让人瞬间贫富转换的战场才始终长盛不衰. 观察概率论的历史, 我们可能还需要感谢一些职业赌徒的执着地想勘破玄机的精神，de Méré 就是记录在案的一个法国贵族赌徒, 他不断地写信给当时最伟大的数学家 Pascal 询问他自己在赌博中迷惑的问题, 比如著名的分奖金问题等, 正是他的这些问题让数学家开始认真思考随机现象中的一些规律性现象, 从而诞生出概率论这个学科.

尽管历史记录认为概率论始于 Pascal 和 Fermat 的通信讨论, 但我相信对随机现象的兴趣与研究一定远早于这个年代, 因为无法想象会没有智者对瞬间暴富的秘法产生兴趣, 由于没有记录, 这些研究湮灭在历史的长河里. 从 Pascal 和 Fermat 那个时间算起, 到 Kolmogorov 的公理体系出现之前, 尽管概率不是严格意义上的数学, 但仍然产生了丰富的结果, 如组成本教材主要内容的大数定律与中心极限定理等, 还有超出本教材范围的随机游动理论, Markov 过程, Lévy 过程, Brown 运动等. 到 20 世纪 30 年代由前苏联伟大的数学家 Kolmogorov 引入了严格的公理体系, 概率论立刻带着丰富的内容进入数学家族, 成为数学的一个重要分支, 直至今天.

本讲义把概率论作为严格的数学理论来介绍, 强调其严密的一面, 但观察概率论的历史, 读者应该认识到概率论的背景和应用是学习概率论所不可缺少的.

试读结束：需要全本请在线购买： www.ertongbook.com

像前言中所说那样: 理论是可以通过教材学的, 直观却是需要自己体验的. 作为数学的一个分支, 如著名概率学者 Feller [5](p.1) 所言: [1] *要注意区分理论的三个方面, 第一是其形式逻辑公理体系, 第二是其直观背景历史演变, 第三是其应用, 对任何一方面的缺失, 都会妨碍你对整个理论的欣赏.*

A. Einstein 有句备受争议的名言: *上帝不掷骰子.* 意思就是说大自然早已确定了所有的规则. 但不论上帝掷不掷骰子, 他至少有时候表现得像在掷骰子, 因为现实世界里确实有许多难以预知结果的现象. 人们称这些现象为随机现象. 概率论是由于人们对随机现象的兴趣发展起来的, 也是研究随机现象的重要工具之一, 但概率论不关心随机现象的成因, 也就是说不研究随机现象是由于自然本身确实不可预知抑或是由于人类的无知. 简单的如掷硬币, 它本质上是一个经典的确定的力学系统, 但现有的测量手段和计算工具远不足以告诉硬币的哪一面会向上, 所以我们一样把它称为随机现象. 概率这个词是人们经常使用的, 用来描述一个随机现象中某个事件出现的可能性大小, 但也许很少有人会仔细地想一想人们在使用这个词的时候的确切意义. 例如, 大多数人都知道掷硬币得到国徽的概率是 $\frac{1}{2}$, 但它的含义究竟是什么呢? 又例如, 若天气预报说明天下雨的概率是 $\frac{1}{2}$, 它的含义又是什么? 是否和硬币的情况一样呢? 其实这两种场合是不同的, 在前一种场合下, 可以说这样的说法是有依据的, 而在后一种场合, 人们只是用概率表达自己的一种信念. 说前一种情况有依据, 是因为我们可以用一种简单的方法进行验证, 例如我们可以任意次地重复地掷硬币, 看其中正面或反面出现的比例, $\frac{1}{2}$ 的意义在于当掷足够多次硬币后, 这个比例大约会是 $\frac{1}{2}$. 而在后一种情况, 概率只是表达说话人语气的一种方式, 不能验证其正确或不正确. 数学中的概率论是以前一种情况为研究对象的. 关于这方面类似问题的解释, 更多地属于哲学范畴, 感兴趣的读者可参考 von Mises 的经典著作 [12].

在 Kolmogorov 把概率作为测度引入了公理体系之后, 概率论成为数学的一个部分, 它很好地表达了绝大多数已知的经典概率问题, 就如同几何学可以表达实际测绘中的问题一样. 严格的理论来自于人们对于直觉的抽象化, 它可以帮助我们触摸到直觉难以触及的深度. 在这里, 我们对概率论的早期历史作一简单说

---

[1] 原文: In each field we must carefully distinguish three aspects of the theory: (a) the formal logical content, (b) the intuitive background, (c) the application. The character, and the charm, of the whole structure cannot be appreciated without considering all three aspects in their proper relation.

明.

(1) 1654年: B. Pascal(1623–1662) 与 P. Fermat(1601-1655) 的 1654 年夏天的一些通信. 在其中, 他们对一些所谓的机会问题进行了探讨, Fermat 建议了古典概率的算法, 而 Pascal 想讨论赌博中的公正问题.

(2) 1657年: Huygens(1629–1695) 出版的书: *On Calculations in Games of Chance.*

(3) 1713年: J. Bernoulli(1654–1705) 出版的书: *The Art of Guessing*, 证明了大数定律.

(4) 1730年: De Moivre(1667–1754) 的工作: The Analytic Method, 证明了中心极限定理.

(5) 1812年: Laplace(1749–1827) 出版的书: *Essai philosophique sur les probabilités.*

(6) 1933年: Kolmogorov 出版的书: *Foundations of Probability Theory*, 概率的公理化推出, 并被大家接受.

还有许多著名学者在概率的早期研究中留下了他们的名字: 如 Poisson(1781–1840), Gauss(1777–1855), Chebyshev (1821–1894), Markov (1856–1922), von Mises (1883–1953), Borel (1871–1956) 等. 然后随着整个数学基础的成熟, 概率也成为一个重要的数学领域.

§1.2 计数

本节简单介绍集合计数的问题, 也就是组合理论, 它是古典概率模型中的主要工具. 最重要的是下面的乘法原理.

定理 1.2.1 设完成一件事情分 r 个顺序的步骤, 第一步有 n_1 种选择, 固定第一步的选择后, 第二步有 n_2 种选择, 固定前两步的选择后, 第三步有 n_3 种选择, $\cdots$, 在固定前 $r-1$ 步的选择后, 第 r 步有 n_r 种选择, 那么完成这件事情共有 $n_1 n_2 \cdots n_r$ 种选择.

从 n 个学生中选 r 个学生排成队列是典型的例子, 这时第一步有 n 种选择, 第一个人选定后, 第二个人有 $n-1$ 种选择, $\cdots$, 由乘法原理共有

$$n(n-1)\cdots(n-r+1)$$

种排列的方法, 记此数为 $(n)_r$, 称为 n 个对象中选 r 个的排列数. 如果 $n=r$, 这等于说 n 个学生有多少种不同的顺序, 共有 $n!$ 种. 这里注意选择数是指固定前面步骤中的选择后的选择数, 单独地看第二步, 它和第一步一样有可能选择到所有的人, 因此也有 n 种选择, 但选好了第一个, 第二步就只有 $n-1$ 个选择了. 这是不可重复的情形, 在其他一些情形下选择也许是可以重复的, 比如选号码, 设 r 个人都从 $1,2,\cdots,n$ 中选个号码, 因为可以重复, 每个人都有 n 种选择, 我们讲选择是独立的, 这时共有 n^r 种不同选择.

将上面 n 个对象取 r 个排列分成为如下两步: 先取出 r 个对象, 然后再将它们排顺序. 第一步的不同选择数称为是 n 个对象中取 r 个的组合数, 记为 $\binom{n}{r}$ 或 C_n^r, 与排列数的不同处是它不计顺序, 第二步是 r 个对象的排列有 $r!$ 这么多选择. 再用乘法原理得

$$(n)_r=\binom{n}{r}\cdot r!,$$

因此

$$\binom{n}{r}=\frac{(n)_r}{r!}=\frac{n!}{r!\cdot(n-r)!}.$$

组合数也被理解为 n 个不同元素的组分为包含 r 个与 $n-r$ 个元素的两个组的分组数. 排列和组合的差别通常也就是考虑顺序和不考虑顺序的差别.

例 1.2.1 掷一个骰子, 有 6 种可能; 两个骰子, 有 36 种可能. 若区别骰子, 两个骰子点数不同有 $6\times5=30$ 种可能, 三个骰子点数互相不同有 $6\times5\times4=120$ 种可能. 若不区别骰子, $2,1,3$ 与 $1,2,3$ 就不区分了, 这时两个骰子点数不同有 $6\times5/2!=15$ 种可能, 三个骰子点数互相不同有 $6\times5\times4/3!=20$ 种可能. 从标准的 52 张扑克牌中取5张牌, 当然通常是不在意牌的顺序的, 共有 $\binom{52}{5}$ 种选择. 其中没有重复的牌点的可能选择共有 $\binom{13}{5}\cdot4^5$ 种, 恰有一个对子的可能选择有 $13\cdot\binom{4}{2}\cdot\binom{12}{3}\cdot4^3$ 种, 恰有两个对子的可能选择有 $\binom{13}{2}\cdot\binom{4}{2}\cdot\binom{4}{2}\cdot11\cdot4$ 种. 恰是三带二(即3张相同数字的加2张相同数字的, 英文是 full house) 的可能选择有 $13\cdot4\cdot12\cdot\binom{4}{2}$ 种. ∎

例 1.2.2 **(分组)** 从 n 个人中取 k 个人的方法数是组合, 它实际上也可以看成为把 n 个人分为两组各有 k 与 $n-k$ 个人, 而且容易验证组合数对应于二项式展开

$$(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k},$$

或者说在左边乘开后的和式中, 其中包含 k 个 x 与 $n-k$ 个 y 的项共有 $\binom{n}{k}$ 个. 考虑更一般的分组: n 个人分为 r 组, 分别各有 $n_1, n_2, \cdots, n_r$ 个人, 这样有多少种不同分法? 第一组的分法有 $\binom{n}{n_1}$ 种, 分好后, 从剩下的 $n-n_1$ 中分第二组, 分法有 $\binom{n-n_1}{n_2}$ 种, 然后再从剩下的 $n-n_1-n_2$ 个人中分出第三组, 依此类推, 由乘法原理, 不同分法共有

$$\binom{n}{n_1}\cdot\binom{n-n_1}{n_2}\cdots\binom{n-n_1-\cdots-n_{r-1}}{n_r}=\frac{n!}{n_1!n_2!\cdots n_r!}$$

种. 我们把右边记为

$$\binom{n}{n_1\ n_2\ \cdots\ n_r}:=\frac{n!}{n_1!n_2!\cdots n_r!},$$

称为多项组合数. 同样有多项式展开公式

$$(x_1+\cdots+x_r)^n=\sum_{n_1+\cdots+n_r=n,n_1,\cdots,n_r\geq 0}\binom{n}{n_1\ n_2\ \cdots\ n_r}x_1^{n_1}\cdots x_r^{n_r},$$

且特别地 $\binom{n}{k}=\binom{n}{k\ n-k}$. ▮

例 1.2.3 满足条件 $x_1+\cdots+x_r=n$ 的正整数值向量 $(x_1,\cdots,x_r)$ 有多少个? 想象有 n 个 1 排成一排, 然后插入 $r-1$ 个隔离板隔成 r 个房间, 第 i 个房间中 1 的个数是 x_i, 这样的分隔恰好给出了上述问题的解. 因此解的个数等于这样的分隔数. 因为 x_i 都是正整数, 故把 $r-1$ 隔板插入 1 之间的 $n-1$ 个空隙就给出了这样一个分隔. 由组合的思想, 这相当于从 $n-1$ 个位置中选择 $r-1$ 个位置, 总数为 $\binom{n-1}{r-1}$. ▮

例 1.2.4 前面看到, 从 $\{1,2,\cdots,n\}$ 中取 k 个不同的数, 如果计顺序, 则有 $(n)_k$ 种不同取法, 也就是说, 集合 $\{(x_1,x_2,\cdots,x_k):\ 1\leq x_i\leq n,$ 且互不相同$\}$ 共有 $(n)_k$ 个元素. 如果不计顺序, 则有 $\binom{n}{k}$ 种不同取法. 也就是说, 满足 $1\leq x_1<$

$x_2 < \cdots < x_k \le n$ 的正整数向量 $(x_1, \cdots, x_k)$ 的个数等于 1 到 n 的整数中取 k 个数的组合数, 即 $\binom{n}{k}$.

然后, 再看取 k 个可以重复的数的取法, 如果计顺序, 那么相当于重复地取 k 次, 每次有 n 种取法, 共有 n^k 种取法. 问题是不计顺序有多少种取法? 那相当于算满足 $1 \le x_1 \le x_2 \le \cdots \le x_k \le n$ 的正整数向量 $(x_1, \cdots, x_k)$ 的个数. 我们用 $A_{n,k}$ 表示满足 $1 \le x_1 < x_2 < \cdots < x_k \le n$ 的正整数向量 $(x_1, \cdots, x_k)$ 组成的集合, 而用 $B_{n,k}$ 表示满足 $1 \le x_1 \le x_2 \le \cdots \le x_k \le n$ 的正整数向量 $(x_1, \cdots, x_k)$ 组成的集合. 作变换

$$\phi((x_1, \cdots, x_k)) := (x_1, x_2 + 1, x_3 + 2, \cdots, x_k + k - 1),$$

它是 $B_{n,k}$ 到 $A_{n+k-1,k}$ 的映射. 显然 ϕ 是单且满的, 因此

$$|B_{n,k}| = |A_{n+k-1,k}| = \binom{n+k-1}{k}.$$

作变换的方法是计数的一个常用方法. 比如我们再来算在 1 到 n 中取 k 个不重叠 2- 组的取法 (不计顺序). 一个 2- 组是指两个连续整数. 如 (1,2), (4,5) 是 1,2,3,4,5 中 2 个 2- 组. 用 $C_{n,k}$ 表示从 1 至 n 个整数中选取的有序的 k 个 2- 组全体. 现在把第 i 个 2- 组 $(n_i, n_i + 1)$ 映射为 $n_i - i + 1$, 这样的映射建立了从 $C_{n,k}$ 到 $A_{n-k,k}$ 的一个一一对应, 因此 k 个不重合 2- 组的取法有 $\binom{n-k}{k}$ 种. ∎

附录 1.2.1 关于和的说明 有限多个数的和是最基本的运算, 但无限多个数的运算就不是那么简单了, 我们在这里作个简单说明. 设 $\{a_i : i \in I\}$ 是实数的集合, 其中 I 是任何指标集, 那么和式

$$\sum_{i \in I} a_i$$

的意义是什么呢? 首先和 $\sum_{i \in I} a_i$ 要有意义的话, 其中最多只能有可数多个数非零, 所以和本质上只能是可列和. 对于可列和 $\sum_{i \in I} a_i$, 这样的写法通常也是没有意义的, 因为和是不是存在及收敛于什么都依赖于求和的顺序, 比如调和级数. 只有当绝对收敛, 即 $\sum_{i \in I} |a_i| < \infty$ 时 (注意因为正项级数求和与顺序无关, 所以这样的写法是有意义的), 求和与顺序无关, 这时 $\sum_{i \in I} a_i$ 这样的写法才有意义.

§1.3 古典概率问题

许多概率论教科书都是直接从公理化开始, 这样的处理从数学上非常简单直

接, 但是作为一本初等概率论教材, 我们希望读者至少能够有机会了解和体会随机现象与概率背后的一些问题以及几个世纪来学者们对这些问题的思考, 因为直接从公理开始不免让初学的读者感觉突兀, 也让具有良好数学背景的读者觉得概率论仅仅是测度论的一个部分. 在这一节中, 让我们通过一些经典例子来看看概率的直观背景, 体会应该怎么样来建立概率论的公理体系. 从数学的角度看这节不是必要的, 特别是对于具有良好数学背景的读者.

随机的特点是有许多可能出现的结果, 而无法给出预测. 我们把这样的事情称为随机试验, 通常用 E 表示. 可能出现的结果全体作为一个集合称为 (该试验的) 样本空间, 可以是有限集, 也可以是无限集, 通常用 Ω 表示, 其中的元素称为样本点或基本事件. 要注意的是这里所说的可能出现的结果依赖于你关心的问题或解决问题的方法. 因此样本空间的选择不是唯一的.

例 1.3.1 掷硬币有两个结果: 反面, 正面. 我们不妨用 T, H 表示, 因此样本空间为 $\Omega = \{\mathrm{T,H}\}$. 要注意我们说掷硬币总是指掷理想化的硬币, 实际上是指有两个可能结果的随机试验, 因为具体的硬币不一定如此, 它可能由于有厚度而站立起来, 不是正面也不是反面. 我们后面讲的各种随机试验都是如此. 掷两枚硬币的样本空间是什么呢? 如果你区别两枚硬币, 那么 $\Omega = \{\mathrm{TT,TH,HT,HH}\}$, 其中第一个字母表示标识的一个硬币的结果, 第二个字母表示另一个硬币的结果. 如果你不能或者不区别硬币, 那么 $\Omega = \{$ 两正, 两反, 一正一反 $\}$. 两者都正确. 问题不在于硬币在物理上是否可区分, 而在于你是否区分它们. 掷两个骰子, 若区别它们, 则样本空间应该有 36 种结果: $\{(i,j) : 1 \leq i, j \leq 6\}$; 若它们无法区别 (或者说不区别它们), 则仅有 21 种结果: $\{\{i,j\} : 1 \leq i \leq j \leq 6\}$; 也许有的游戏只关心偶数与奇数, 那么只有 3 种结果: 都是偶数, 一偶一奇, 都是奇数; 也许有的游戏只关心点数和, 那么只有 11 种结果. ∎

例 1.3.2 掷一个硬币一直到正面出现时停止. 这样的随机试验的样本空间是 $\Omega = \{1, 2, \cdots, n, \cdots, +\infty\}$, 其中 n 表示掷的次数是 n 次, 即前 $n-1$ 次是反面, 第 n 次是正面, $+\infty$ 表示永远都是反面, 不能停止. 这是一个无限多个元素的样本空间. ∎

概率是指一个随机试验中某个所关心的事件发生 (或出现) 的可能性的大小, 因此要理解概率, 首先要理解事件. 仔细地想, 事件是叙述某个条件, 比如说在 “掷骰子出现偶数的概率” 这句话中, 掷骰子 是随机试验, 偶数是事件, 它是指样本空

试读结束: 需要全本请在线购买: www.ertongbook.com