南开大学数学教学丛书

概 率 论

杨振明 著

科学出版社

南开大学数学教学丛书

概 率 论

杨振明 著

科 学 出 版 社

1 9 9 9

EA01 1 9

内 容 简 介

本书为南开大学数学系教材.内容包括:事件与概率;随机变量;数字特征与特征函数;极限定理等.本书是作者多年教学工作经验的总结,内容丰富,深入浅出,论述严谨,每一节后都有习题,有助于读者理解书中内容.

本书可供高等学校数学系学生和教师阅读和参考.

图书在版编目(CIP)数据

概率论/杨振明 著. -北京:科学出版社,1999.1

(南开大学数学教学丛书)

ISBN 7-03-006746-0

Ⅰ.概… Ⅱ.杨… Ⅲ.概率论 Ⅳ.O211

中国版本图书馆 CIP 数据核字(97)第 12215 号

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

新蕾印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1999 年 1 月第 一 版 开本:850×1168 1/32

1999 年 1 月第一次印刷 印张:8 5/8

印数:1-2800 字数:231 000

定价:14.50 元

序

海内外华夏炎黄子孙都盼望中国早日成为数学大国，也就是“实现中国数学的平等和独立”[1]。平等和独立是由中国出类拔萃的数学家及其杰出的研究工作来体现的，要有出类拔萃的数学家就要培养一批优秀的研究生、大学生。这批人不在多，而在精，要层次高。也就是要求他们热爱数学、基础扎实、知识面广、能力强。

80 年代中期，国家采纳了陈省身先生的几个建议。建议之一是为培养高质量的数学专业的大学生，需要建立数学专业的试点班。经过胡国定先生等的努力，1986 年在南开大学建立了数学专业的试点班。这些作法取得了成功，并在基础学科的教学中有了推广。1990 年在全国建立“国家理科基础学科研究和教学人才培养基地”。其后南开大学数学专业成为基地之一。从 1986 年到现在的 10 余年中南开数学专业是有成绩的，例如他们四次参加全国大学生数学竞赛获三次团体第一，一次团体第三。在全国和国际大学生数学建模比赛中多次获一等奖。毕业生中的百分之八十继续攻读研究生，其中许多人取得了很好的成绩。

当然，取得这些成绩是与陈省身先生的指导、帮助分不开，是与国内外同行们的支持与帮助分不开的。如杨忠道，王叔平，许以超，虞言林，李克正等或参与教学计划、课程设置、课程内容的制订，或到南开任教等等。有了这些指导、帮助与支持，南开基础数学专业得以广泛吸收国内外先进的数学教学经验，并以此为基础对数学教学进行了许多改革、创新。

这套丛书是南开大学数学专业的部分教材，编著者们长期在南开数学专业任教，不断地把自己的心得体会揉和到基础知识和基本理论的讲述中去，日积月累地形成了这套教材。所以可以说这

[1] 陈省身：《在“二十一世纪中国数学展望”学术讨论会开幕式上的讲话》

些教材不是“编”出来的，而是在长期教学中“教”出来的，“改”出来的，凝聚了我们的一点心血。这些教材的共同点，也是我们教学所遵循的共同点是：首先要加强基础知识、基础理论和基本方法的教学；同时又要适当地开拓知识面，尤其注意反映学科前沿的成就、观点和方法；教学的目的是丰富学生的知识与提高学生的能力，因此配置的习题中多数是为了巩固知识和训练基本方法，也有一些习题是为训练学生解题技巧与钻研数学的能力。

我们要感谢中国科学出版社主动提出将这套教材出版。这对编著者是件大好事。编著者虽然尽了很大努力，但一则由于编著者的水平所限，二则数学的教育和所有学科的教育一样是在不断发展之中，因此这套教材中缺欠和不足肯定存在。我们诚挚希望各位同行不吝指正，从而使编著者更明确了解教材及教学中的短长，进而扬长避短，改进我们的教学。同时通过这套教材也可向同行们介绍南开的教学经验以供他们参考，或许有益于他们的工作。

我们再次感谢帮助过南开的前辈、同行们，同时也希望能继续得到他们和各位同行的帮助。办好南开的数学专业，办好所有学校的数学专业，把中国数学搞上去，使中国成为数学大国是我们的共同愿望！这个愿望一定能实现！

全体编著者

于南开大学

1998 年 6 月

目　　录

第一章　事件与概率

第二章　随机变量

第一章　事件与概率

1.1　基本概念

1.1.1　随机现象

客观世界中存在着两类现象，一类是在一定条件下必然出现的现象，称之为**必然现象**；另一类是在一定条件下可能出现也可能不出现的现象，称作**随机现象**.

在标准大气压下，100° C 的纯水必然沸腾；带异性电荷的小球必然相互吸引；在恒力作用下的物体必然作匀加速运动等等都是必然现象的例子. 以往的各种数学学科的主要内容就是研究必然现象中的数量规律. 例如，在重力作用下，物体的位移随时间变化的函数 $x(t)$，就由二阶微分方程 $x''(t)=g$ 来描述，其中 g 为重力加速度.

随机现象也是广泛存在的. 例如，抛掷一枚硬币，它可能是正面朝上也可能是背面朝上，就是说，“正面朝上”这个结果可能出现也可能不出现；记录一段时间内某电话机的呼唤次数可能较多也可能较少；下一个交易日股市的指数可能上升也可能下跌，而且升跌幅度的大小也不能事先确定. 这些都是随机现象的例子. 许多影响事物发展的偶然因素的存在，是产生随机现象中不确定性的原因. 例如，股市指数的变化取决于金融政策的变化，上市企业的经济状况，股民的经营行为以及其它国家的股市沉浮等诸多不确定因素. 这些因素发展变化的偶然性，决定了股市升跌的随机性.

虽然随机现象中出现什么样的结果不能完全预言，但是可以假定全部可能结果是已知的. 在上述例子中，抛一枚硬币只会有“正面”与“背面”这两个可能结果，电话呼唤次数必定是某个非负整数，而股指的升跌幅度充其量假定它可能是任意的实数. 可见“全部可能结果的集合是已知的”这个假定是合理的，而且它会给我们带来许多方

便.

进行一次试验，如果其所得结果不能完全预言，但其全体可能结果是已知的，则称此试验为**随机试验**. 这里及今后所使用的"试验"这一术语，其含义是广泛的. 它既可以是通常意义下的物理或化学的实验，也可以是对自然或社会现象的观测与记录.

表 1.1 抛硬币试验资料

试验者	抛硬币次数	出现正面次数	出现正面频率
蒲丰	4040	2048	0.5069
德莫根	4092	2048	0.5005
费勒	10000	4979	0.4979
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005
罗曼诺夫斯基	80640	39699	0.4923

虽然一次随机试验的结果不能完全预言，但是，在相同条件下大量重复此试验时，则会呈现出一定的数量规律性. 这一点被历史上许多人的试验结果所证明. 表 1.1 列出了蒲丰等人连续抛掷均匀硬币所得的结果. 从表中的数据可以看到，当抛掷次数很大时，正面出现的频率非常接近 0.5，就是说，出现正面与出现背面的机会差不多各占一半.

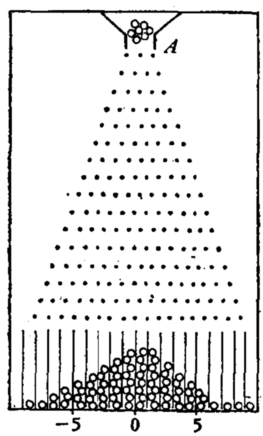

图 1.1 高尔顿板

另一个著名的例子是高尔顿钉板试验. 在一块竖起的木板上钉上一排排间隔相等的铁钉 (见图 1.1). 每一排各个钉子正好对准上面一排两个相邻铁钉的中央. 这样，当小球从两钉之间的间隙下落时，由于下一排铁钉的碰撞，它将以相等的可能性向左或向右落下. 小球再走过两钉间隙时，又遇到下一排铁钉的碰撞. 如此下去，当小球走过最后

一排铁钉的间隙后，便落入下方的被分割成条状的容器中. 让许许多多小球自上而下地重复这一下落过程. 尽管每次指定小球下落的位置无法预言，但许多小球堆积的边缘轮廓线总是两头矮中间高且左右近于对称的钟形曲线. 就是说，落入各条状容器内小球数所占的比例基本保持不变.

上面列举的两个试验的结果表明，在相同条件下大量地重复某一随机试验时，各可能结果出现的频率稳定在某个确定的数值附近. 称这种性质为**频率的稳定性**. 频率稳定性的存在，标志着随机现象也有它的数量规律性. **概率论就是研究随机现象中数量规律的数学学科**.

历史上概率论起源于赌博. 三百多年以前，现代化生产与技术尚处于萌芽状态，对自然界中随机现象规律性的研讨还没有提到议事日程. 但此时，具有悠久历史的游戏与赌博却发达得多. 人们借助于骰子，纸牌以及形形色色的工具进行赌博，遇到了许多无法解释的问题. 由于输赢是无法预言的，而且涉及到金钱的得失，所以了解其中的数量规律就变得越来越迫切了. 17 世纪中叶，赌徒中一些有身分的人开始向他们的数学家朋友请教. 当时在欧洲颇具声望的数学家费尔马、巴斯卡、惠更斯等人都参加了有关的讨论. 于是产生了概率论的基本概念，并给出了计算“等可能型”概率的一套方法.

19 世纪末至 20 世纪初，在现代工业技术蓬勃发展的大潮中，概率论也取得了飞速的发展. 特别是柯尔莫戈洛夫等人建立了概率论的公理化体系，奠定了概率论的严格的数学基础，也沟通了概率论与现代数学中其它分支之间的联系. 近年来，概率论已被广泛应用于自然科学，工程技术，经济理论，经营管理等许多方面. 特别是对金融领域中随机现象的研究与应用有了长足的发展，形成了“金融数学”的重要组成部分. 概率论这个有特色的数学分支呈现出蓬勃发展的局面. 当前，我国高等学校的许多专业都开设程度不同的概率论课程，单从这一点就可看到概率论作为一门基础学科在社会发展中所起的巨大作用.

最后我们指出，在今天的概率论教材中，仍广泛采用“抛硬币”，“掷骰子”等与赌博有关的例子. 除了一些典型例题有很重要的历史

地位外，主要是由于这些例子所涉及的概念简单，规则明了，易于阐明概率论的基本概念与方法.

1.1.2 样本空间

随机试验的每个可能结果称为一个**样本点**, 全体样本点所组成的集合称为**样本空间**. 习惯上分别用小写的 ω 与大写的 Ω 表示样本点与样本空间. 如前所述，我们总假设样本空间 Ω 是确定的.

例 1.1 抛两枚硬币观察其正面与反面出现的情况. 其样本空间由四个样本点组成，即 $\Omega = \{$ (正，正), (正，反), (反，正), (反，反)$\}$. 这里，比如样本点 $\omega =$ (正，反) 表示第一枚硬币抛出正面而第二枚抛得反面.

例 1.2 记录某电话机在一小时内呼唤次数，其样本点是非负整数. 当然，我们可以根据实际情况为电话呼唤次数确定一个上界 N. 但是今后会看到，人们宁愿不确定这个 N, 而假设呼唤次数可以是任何非负整数，即认为样本空间 $\Omega = \{0, 1, 2, \cdots\}$.

例 1.3 接连射击直到命中为止. 为了简洁地写出其样本空间，我们约定以 "0" 表示一次射击未中，而以 "1" 表示命中. 则样本空间 $\Omega = \{1, 01, 001, \cdots\}$, 它是一个可列的无穷集合.

例 1.4 接连不断地射击下去. 沿用上例的记号. 由于这个试验是无休止地反复射击，故其样本点是由 0 与 1 组成的无穷序列. 为确定起见，我们在每个序列之前加上小数点. 于是其样本空间是整数部分为 0 的二进小数全体，即 $\Omega = \{0.a_1a_2a_3\cdots : a_i = 0 \text{ 或 } 1\} = [0, 1]$. 这是一个不可列的无穷集合.

例 1.5 记录某地的最低气温与最高气温. 我们以 x, y 分别表示最低与最高气温，则样本点是数偶 (x, y). 虽然一次试验的结果不能完全预言，但人们总可以确定此时此地气温的下界 a 与上界 b. 于是其样本空间 $\Omega = \{(x, y) : a \leq x \leq y \leq b\}$. 这是坐标平面中一个三角形，包含不可列无穷多个样本点.

不难看出，样本空间 Ω 可以是数集，也可以是任何抽象的集合；可以是有限集，可列集，也可以是不可列的无穷集合；可以是一维的也可以是多维集合. 我们指出，样本空间是研究随机现象的数学模型. 正确地确定不同随机试验的样本点与样本空间是极为重要的. 比如例 1.1 已给出抛两枚硬币观察其正反面出现情况的样本空间，如果我们只关心抛两枚硬币出现正面的个数，则样本空间应改为 $\Omega=\{0,1,2\}$.

1.1.3 事件及其运算

我们时常会关心试验的某一部分可能结果是否出现. 如在例 1.2 中，若以每小时是否达到 5 次电话呼唤来区分这台电话机是否太繁忙，那么"不太繁忙"即不足 5 次的呼唤. 它由样本空间中前 5 个样本点 0, 1, 2, 3, 4 组成. 由于它是由 Ω 中的一部分样本点组成的子集合，故在未来的一次试验中可能发生也可能不发生. 称这种由部分样本点组成的试验结果为**随机事件**，简称作**事件**. 通常用大写的字母 $A, B, \cdots$ 等表示事件. 例如在前述例子中，"不太繁忙"可表为事件 $A=\{$ 至多 4 次呼唤 $\}=\{0,1,2,3,4\}$. 还可以有事件 $B=\{$ 至少 2 次呼唤 $\}=\{2,3,4,\cdots\}$，等等. 于是，如果试验的结果是记录到 1 次呼唤，即样本点 $\omega=1$ 出现，此时事件 A 的要求满足，$\omega\in A$，即事件 A 已发生；事件 B 的要求没满足，$\omega\notin B$，即事件 B 没发生.

我们指出，当样本空间为有穷或至多可列无穷的集合时，可取其任何子集为事件. 而当样本空间为不可列无穷时，比如对例 1.4 中的 $\Omega=[0,1]$，则只能取 Ω 的一部分性质较好的 (称作可测的) 子集作为随机事件. 这一点将在后面的 1.4.2 节作进一步讨论. 由于样本空间 Ω 包含了全部可能结果，因此在每次试验中 Ω 都会发生，故称 Ω 为**必然事件**. 相反，空集 $\varnothing$ 不含任何样本点，每次试验必定不发生，故称 $\varnothing$ 为**不可能事件**. 除此之外，每一随机试验都含有许多随机事件. 由于它们共处于同一试验之中，因而是相互联系着的. 我们有必要弄清它们之间的关系，并引进事件间的运算，以便化复杂事件为简单事件，更好地解决相应的概率问题.

表 1.2 测度论与概率论概念的对照

记　　号	测度论含义	概 率 论 含 义
Ω	空间或全集	样本空间或必然事件
$\varnothing$	空集	不可能事件
ω	元素	样本点
A	可测子集	随机事件
$\omega\in A$	ω 是 A 的元	事件 A 发生
$A\subset B$	A 包含在 B 中	A 发生则 B 必发生
$A=B$	A 与 B 相等	A 与 B 同发生或同不发生
$A\bigcap B$ 或 AB	交集	A 与 B 都发生
$A\bigcap B=\varnothing$	不相交	A 与 B 不相容 (互斥)
$A\bigcup B$	并集	A 与 B 至少一个发生
$\overline{A}$	余集	A 不发生 (对立事件)
$A-B$	差集	A 发生但 B 不发生
$\overline{\lim\limits_{n}}\,A_n=\bigcap\limits_{k=1}^{\infty}\bigcup\limits_{n=k}^{\infty}A_n$	上限集	$\{A_n\}$ 中有无限多个发生
$\underline{\lim\limits_{n}}\,A_n=\bigcup\limits_{k=1}^{\infty}\bigcap\limits_{n=k}^{\infty}A_n$	下限集	$\{A_n\}$ 中至多有限个不发生
$P(A)$	集 A 的测度	事件 A 的概率
$\xi(\omega)$	可测函数	随机变量
$E(\xi)=\int_\Omega \xi dP$	积分	数学期望

注意我们已经开始把概率论的基本概念纳入测度论轨道: 样本空间就是集论中的空间 (即全集), 样本点是空间的元素, 随机事件就是可测子集, 出现的样本点 ω 是集合 A 的元素则意味着事件 A 发生. 循此可沿用测度论的语言完整地叙述概率论的基本概念, 我们将在后面详述. 这里先将其基本概念间的对照简要地列在表 1.2 中, 请读者结合具体例子进一步理解与熟练运用这些记号与概念.

与集合运算类似, 事件的并与交运算也可推广到任意多个事件的情形. 例如 $\bigcap\limits_{n=1}^{\infty}A_n$ 表示事件列中每一个 A_n 均发生, 而 $\bigcup\limits_{n=1}^{\infty}A_n$ 则表示此事件列中至少有一个 A_n 发生.

可以用图示的方法表示事件之间的关系与运算. 设想向正方形 Ω 内任投一点 ω, 则 Ω 与 ω 分别是此试验的样本空间与样本点, 而正方形 Ω 内的子集 A 代表随机事件. 所投的点 ω 落入 A 中, 则事件 A 发生. 于是事件间的几种关系与运算可表示如图 1.2. 这种表示方法称为维恩图.

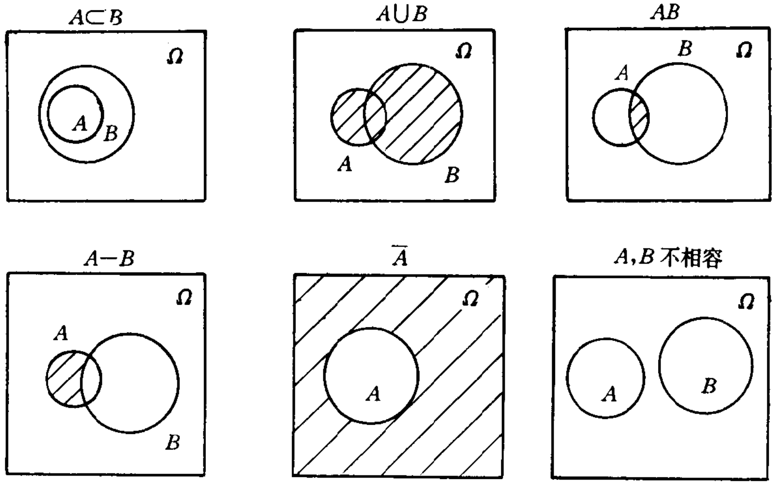

图 1.2　事件运算的维恩图

事件运算有与代数运算类似的交换律、结合律与分配律, 也有作为集合运算的特殊的运算律. 例如有以下对偶律:

$$\overline{A\bigcup B}=\overline{A}\bigcap\overline{B},\qquad \overline{A\bigcap B}=\overline{A}\bigcup\overline{B}. \tag{1.1}$$

我们用概率论语言证明 (1.1) 中第一个等式. 为此需证明"若事件 $\overline{A\bigcup B}$ 发生, 则 $\overline{A}\bigcap\overline{B}$ 发生; 反之, 若 $\overline{A}\bigcap\overline{B}$ 发生则 $\overline{A\bigcup B}$ 也发生". 事实上若事件 $\overline{A\bigcup B}$ 发生, 则事件"A 与 B 至少一个发生"不发生, 则 A 与 B 全不发生, $\overline{A}$ 与 $\overline{B}$ 全发生, 此即事件 $\overline{A}\bigcap\overline{B}$ 发生. 类似可证"若 $\overline{A}\bigcap\overline{B}$ 发生则 $\overline{A\bigcup B}$ 也发生". 至此 (1.1) 中第一个等式得证. 对偶律对于多个事件的交、并运算仍然成立.

1.1.4 频率与概率

本节将通过频率引进事件发生的概率，即介绍概率的“统计定义”. 在一定条件下，将一随机试验重复 n 次，如果其中事件 A 共发生 m 次，则称

$$F(A)=\frac{m}{n}$$

为事件 A 发生的**频率**.

这里试验次数 n 是自然数，A 发生的次数 m 是非负整数. 因而任何事件 A 的频率 $F(A)$ 总是非负的. 其次，由于必然事件 Ω 在每次试验中均发生，故其频率 $F(\Omega)=n/n=1$. 最后，当事件 A 与 B 互不相容，即 A 与 B 不可能同时发生时，事件 $A\bigcup B$ 发生的次数等于 A 发生的次数与 B 发生的次数之和. 于是频率有可加性，即事件 $A\bigcup B$ 发生的频率等于 A,B 各自发生频率之和. 总之，频率 $F(A)$ 有如下基本性质：

1° **非负性**：$F(A)\geq 0$;

2° **规范性**：$F(\Omega)=1$;

3° **可加性**：A 与 B 不相容时 $F(A\bigcup B)=F(A)+F(B)$.

前已指出，事件发生的频率有稳定性，即当重复试验的次数 n 很大时，每个事件 A 发生的频率 $F(A)$ 有一个稳定值. 例如，蒲丰等人的试验结果表明，大量重复地抛一枚均匀硬币时，正面出现的频率稳定在 0.5 附近. 这是因为硬币的质地与形状都是均匀的，正面与反面出现的可能性相等. 于是频率的稳定值 0.5 恰好代表了正面出现可能性的大小. 又如，统计了大量英文文献的字母使用情况，发现字母 E 使用的频率稳定在 0.105 附近，而字母 J 使用的频率则小得多，大约为 0.001. 显然，这种字母使用频率的大小取决于英语本身，由它的特定的构词方法与语法所决定. 频率的稳定值反映了各个字母在英语中出现的可能性的大小.

这种表征在一定条件下事件 A 发生可能性大小的频率稳定值，

就称作事件 A 的**概率**，记为 $P(A)$. 作为频率的稳定值，概率 $P(A)$ 也应具备相应的三条性质：

1° **非负性**：$P(A) \geq 0$;

2° **规范性**：$P(\Omega) = 1$;

3° **可加性**：若 A 与 B 不相容，则 $P(A \bigcup B) = P(A) + P(B)$.

我们指出，在一定条件下进行一随机试验，各确定事件发生的概率是这个试验固有的性质，它是由试验的条件与各事件确定的内含所决定的. 可以通过大量地重复此试验，借助频率的稳定值去认识它，也可以通过其它途径，比如根据对试验机制的具体分析来确定. 严格地说，有些事件的概率是无法通过重复试验来确定的. 例如气象台每晚要发布未来一天中北京市的"降水概率"，这个概率应当是根据采集到的数据经科学的分析与计算得出的. 特定时日的北京市在历史的长河中只有一个，当然无法大量重复地观测其降水情况，从而也无法借助频率的稳定值求出降水概率.

总之，我们借助大量重复试验中频率值呈现的稳定性，来说明表征事件发生可能性大小的概率是客观存在的. 这里所使用的"大量试验中频率的稳定值"，是一种极不规范的说法. 实际上，概率 $P(A)$ 就是当试验次数 $n \to \infty$ 时，频率 $F(A)$ 的一种极限. 只不过其中 A 发生的频数 m 是不能完全预言的数量，从而 $F(A) = m/n$ 也不是通常的数列而已. 其严格的含义将在第四章详加叙述.

1.1.5 习题

1. 设 A, B, C, D 是某一试验中的四个事件. 试用它们的运算表达如下事件： (1) 四个事件中至少一个发生； (2) 恰好发生两个； (3) 至少发生三个； (4) 至多发生一个.

2. 设 A, B, C 为某随机试验中三个事件. 试说明下列关系式的概率含义，并画出相应的维恩图：

(1) $A \bigcap B \bigcap C = A$;　　(2) $A \bigcup B \bigcup C = A$;

(3) $A \bigcap B \subset C$;　　(4) $A \subset \overline{B \bigcap C}$.