普通高等教育“十一五”规划教材

21世纪高等院校创新教材

概率论与数理统计及其应用（第二版）

熊德之　张志军　罗　进　主编

科学出版社
www.sciencep.com

普通高等教育"十一五"规划教材

21世纪高等院校创新教材

概率论与数理统计及其应用

(第二版)

熊德之　张志军　罗　进　主编

科学出版社

北京

内 容 简 介

本书在第一版的基础上修订再版．全书共分十章，内容包括随机事件与概率、随机变量及其分布、随机变量的数字特征、大数定律及中心极限定理、概率模型及其应用、数理统计的基本概念、参数估计、假设检验、方差分析与回归分析、SPSS 统计软件介绍与统计模型应用．每章配有习题，书末附有习题答案．

本次修订保持了原教材内容严谨、叙述翔实、突出应用，将数学建模思想、方法和使用软件工具融入一体的优点，修正了某些不妥之处，强化了基础知识和重点内容．本书可作为普通高等学校非数学类专业概率论与数理统计课程的教材，也可供相关教师、考研人员及工程技术人员参考．

图书在版编目（CIP）数据

概率论与数理统计及其应用/熊德之，张志军，罗进主编．—2 版—北京：科学出版社，2009
（普通高等教育“十一五”规划教材．21 世纪高等院校创新教材）
ISBN 978-7-03-023918-1

I. 概… II. ①熊…②张…③罗 III. ①概率论－高等学校－教材②数理统计－高等学校－教材 IV. O21

中国版本图书馆 CIP 数据核字（2009）第 004157 号

责任编辑：江 兰／责任校对：董艳辉
责任印制：彭 超／封面设计：苏 波

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

武汉市新华印刷有限责任公司印刷
科学出版社发行 各地新华书店经销
*
2005 年 12 月第 一 版
2009 年 2 月第 二 版 开本：B5（720×1000）
2009 年 2 月第四次印刷 印张：18 1/4
印数：13 001—19 000 字数：358 000

定价：29.80 元

（如有印装质量问题，我社负责调换）

第二版前言

本书第一版，在概率论与数理统计课程的教学中发挥了积极作用，经过几年的教学实践，教师也积累了新的经验. 为了使本书更好地适应新形势下教材改革精神，体现创新教学理念，打造成精品课程教材，在科学出版社的大力支持下，我们对它作了一次全面的修订.

这次修订，保持了原教材内容严谨、叙述翔实、突出应用，将数学建模思想、方法和实用软件工具融入一体的优点，修正了某些不妥之处，强化了基础知识和重点内容. 其中第一章是重新编写，第四章和第五章作了较大的补充和修改，其他章节也进行了修改与完善. 在修订稿中更加重视教材的实用性，增强趣味性，突出思想方法，不过分追求计算技巧.

这次修订，几位编者的任务作了适当调整，各章具体修订编写的人员如下：第一章，张志军；第二章、第三章，郭光耀；第四章、第六章，罗进；第五章和第十章的第 4 至 5 节，刘任河；第七章、第八章，熊德之；第九章和第十章的第 1 至 3 节，严国义.

书中难免存在的不足之处，诚恳地欢迎广大专家、同行和读者批评指正.

编　者

2008 年 10 月

第一版前言

概率论与数理统计是高等学校的一门重要基础课程，也是应用性极强的一门学科，它在理学、工学、农学、医学、经济学、管理学、军事学和体育科学等学科领域具有广泛的应用. 针对这门课程的特点，我们编写本书的基本思路是：着重概率统计思想和方法的阐述，着重学生应用能力的培养. 本书较详细地叙述了概率论与数理统计中主要概念和方法产生的背景及思想，注重理论联系实际，突出解决问题的思路，详尽介绍各种概率统计模型的掌握与应用；注重将数学建模思想和方法有机地融入到教材中，以期提高学生利用概率统计思想方法建立数学模型和利用统计分析软件解决实际问题的能力.

教材改革是教学改革的重要内容之一. 面向 21 世纪的概率统计教材一方面要在培养和提高学生应用能力上有较大突破，另一方面要在培养学生的数学素质，增强学生学习数学的兴趣上作探索. 一年一度的全国大学生数学建模竞赛已成为高等学校学生参赛人数最多、影响最大的一项课外科技活动，这就要求我们在数学基础课程的教学中介绍数学建模的思想和方法，为学生参赛作必要的准备. 本书也是省级重点教学研究项目《将数学建模思想和方法融入概率统计课程教学中的研究》课题的成果之一. 我们参照教育部高等学校数学与统计学教学指导委员会制订的《非数学类专业数学基础课程教学基本要求》，并按照"十五"国家级规划教材及教育部面向 21 世纪课程教材规划的要求，以及学生参加全国大学生数学建模竞赛的需要，集多年教学之经验，在备课教案和讲义的基础上，编写了这本教材. 为适应不同的教学对象和不同专业类别的教学需要，我们将有些内容打"*"号以便在教学中进行取舍.

本书由熊德之、张志军主编，负责拟定教材的编写大纲，并对全书进行修改、统稿和校审，以及组织在教学中试用. 罗进参加了部分内容的修改和校审工作. 各章的具体编写人员如下：第一章至第三章，郭光耀；第四章，罗进；第五章，刘任河；第六章至第八章，熊德之；第九章、第十章的第 1 至 3 节，严国义；第十章第 4 至 5 节及习题，张志军.

我们深知在众多概率论与数理统计经典教材面前，要写出一本体系结构新颖、特色鲜明且受到欢迎的教材是十分困难的. 在当今改革与探索的年代，我们的教材也算是一家之言. 这本教材的编写仅仅是我们工作的开始，我们将不断探索，为概率统计课程的教学改革尽一份力量.

由于编者水平有限，书中难免有缺点和错误，欢迎广大专家、同行和读者批评指正.

编　者

2005 年 7 月

目　　录

第一章　随机事件与概率

1　随 机 事 件

1.1　随机现象

在自然界和人的实践活动中经常遇到各种各样的现象，这些现象大体可分两类.

一类是确定的，如在一个标准大气压下，纯水加热到 100 ℃时必然沸腾；向上抛一块石头必然下落；同性电荷相斥，异性电荷相吸等，这种在一定条件下有确定结果的现象称为**确定性现象**.

另一类现象是随机的，例如：在相同的条件下，向上抛一枚质地均匀的硬币，其结果可能是正面朝上，也可能是反面朝上，不论如何控制抛掷条件，在每次抛掷之前无法肯定抛掷的结果是什么，这个试验多于一种可能结果，但是在试验之前不能肯定试验会出现哪一个结果. 又如：同一门大炮对同一目标进行多次射击（同一型号的炮弹），各次弹着点可能相同，并且每次射击之前无法肯定弹着点的确切位置. 以上所举的现象都具有随机性，即在一定条件下进行试验或观察会出现不同的结果（也就是说，多于一种可能的试验结果），而且在每次试验之前都无法预言会出现哪一个结果（不能肯定试验会出现哪一个结果）. 这类现象，在一定的条件下，可能出现这样的结果，也可能出现那样的结果，而在试验或观察之前不能预知确切的结果. 但人们经过长期实践并深入研究之后，发现这类现象在大量重复试验或观察下，它的结果却呈现出某种规律性. 例如：多次重复抛一枚硬币大致有一半次数是正面朝上；同一台仪器测量同一物体的重量，所得重量总在真实重量上下波动等. 这种在大量重复试验或观察中所呈现出的固有规律性，就是我们所说的**统计规律性**. 这种在个别试验中其结果呈现出不确定性，在大量重复试验中其结果又具有统计规律性的现象，称之为**随机现象**.

1.2　随机试验

我们遇到过各种试验，这里将把试验作为一个含义广泛的术语，它包括各种各样的科学试验，甚至对某一事物的某一特征的观察也认为是一种试验. 如果这个试验满足：

（1）在相同条件下可以重复进行；

（2）每次试验的可能结果不止一个，并且事先能够确定试验的所有可能结果；

(3) 每次试验总是恰好出现这些可能结果中的一个，但在一次试验之前却不能肯定这次试验出现哪一个结果.

我们称满足以上条件的试验为**随机试验**，也简称为**试验**，本书中所讨论的试验都是指随机试验. 下面举一些随机试验的例子.

E_1：抛一枚硬币，观察正、反面出现的情况.

E_2：掷一颗骰子，观察出现的点数.

E_3：记录一天进入某超市的顾客数.

E_4：测试某种型号电视机的寿命.

E_5：测量某物理量(长度、直径等) 的误差.

进行一次试验总有一个观察的目的，试验中会观察到有多种不同的可能结果. 例如在 E_2 中，如果我们的目的是观察它朝上面的点数，其可能结果是1点、2点、3点、4点、5点、6点；如果我们的目的是观察它朝上面点数的奇偶性，其可能的结果是奇数点、偶数点两个. 至于骰子落在桌面上哪个位置，朝哪个方向滚动等不在目的之列，不算做结果.

1.3 样本空间

对于随机试验，尽管在每次试验之前不能预知试验的结果，但试验的所有可能结果组成的集合是已知的，将随机试验 E 的所有可能结果组成的集合称为 E 的**样本空间**，记为 S. 样本空间的元素，即 E 的每个结果，称为**样本点**，记为 ω.

由于讨论的问题不同，其样本空间差别是很大的. 下面介绍概率论与数理统计中通常讨论的一些情形.

例 1 将两颗骰子掷一次，观察其朝上面点数之和，其样本空间为$\{2,3,4,\cdots,12\}$.

例 2 将两颗骰子掷一次，观察其朝上面点数的奇偶性，其样本空间为{(奇奇)，(奇偶)，(偶偶)}.

例 3 观察一小时中落在地球上某一区域的宇宙射线数，可能的结果一定是非负整数，而且很难指定一个数作为它的上界. 这样，可以把样本空间取为$\{0,1,2,\cdots\}$.

例 4 讨论某地区的气温时，我们可以把样本空间取为$(-\infty,+\infty)$ 或$[a,b]$.

实际上，相同的试验，随着观察的角度不同，其样本空间可以不同，如例 1、例 2；样本点的个数可以是有限个，也可能是无穷多个，如例 2、例 3.

1.4 随机事件

有了样本空间的概念，就可以定义随机事件. 一般称试验 E 的样本空间 S 的子集为 E 的**随机事件**，简称**事件**. 事件既可以看成样本空间的子集，也可以看成由样本点构成的集合. 一般地，用字母 $A,B,C,\cdots$ 表示. 在每次试验中，当且仅当这一集

合中的一个样本点出现时，称这一**事件发生**.

不可能再分的事件称为**基本事件**. 实际上，它是由一个样本点组成的单点集. 由若干个基本事件组成的事件称为**复合事件**. 特别地，样本空间 S 包含所有的样本点，在每次试验中它总是发生的，称为**必然事件**. 空集 $\varnothing$ 不包含任何样本点，它也可以作为样本空间的子集，它在每次试验中都不发生，称为**不可能事件**.

1.5 事件间的关系与事件的运算

概率论的出现，是始于集合论之前的事情. 从本质上说，事件就是集合，事件间的关系与运算就是集合的关系与运算. 下面给出这些关系和运算在概率论中的提法，并根据"事件发生"的含义，给出它们在概率论中的含义.

(1) 若 $A \subset B$，则称事件 B **包含**事件 A，指的是事件 A 发生必然导致事件 B 发生.

(2) 若 $A \subset B$ 且 $B \subset A$，即 $A = B$，则称事件 B 与事件 A **相等**.

(3) 事件 $A \cup B = \{x \mid x \in A$ 或 $x \in B\}$ 称为事件 A 与事件 B 的**和事件**，当且仅当至少 A,B 有一个发生时，事件 $A \cup B$ 发生.

类似地，称 $\bigcup\limits_{k=1}^{n} A_k$ 为 n 个事件 $A_1, A_2, \cdots, A_n$ 的和事件；称 $\bigcup\limits_{k=1}^{\infty} A_k$ 为可列个事件 $A_1, A_2, \cdots$ 的和事件.

(4) 事件 $A \cap B = \{x \mid x \in A$ 且 $x \in B\}$ 称为事件 A 与事件 B 的**积事件**，当且仅当 A,B 同时发生时，事件 $A \cap B$ 发生，也记做 AB.

类似地，称 $\bigcap\limits_{k=1}^{n} A_k$ 为 n 个事件 $A_1, A_2, \cdots, A_n$ 的积事件；称 $\bigcap\limits_{k=1}^{\infty} A_k$ 为可列个事件 $A_1, A_2, \cdots$ 的积事件.

(5) 事件 $A - B = \{x \mid x \in A$ 且 $x \notin B\}$ 称为事件 A 与事件 B 的**差事件**，当且仅当 A 发生，B 不发生时，事件 $A - B$ 发生.

(6) 若 $A \cap B = \varnothing$，则称事件 A 与事件 B 是**互不相容**的，或**互斥**的. 这指的是事件 A 与事件 B 不能同时发生.

(7) 若 $A \cup B = S$ 且 $A \cap B = \varnothing$，则称事件 A 与事件 B 互为**逆事件**，又称事件 A 与事件 B 互为**对立事件**. 这指的是每次试验中，事件 A,B 中必有一个发生，且仅有一个发生. A 的对立事件记为 $\overline{A}$，且有 $\overline{A} = S - A$，$A - B = A \cap \overline{B}$.

在进行事件运算时，要经常运用下述定律：设 A,B,C 为事件，则有

交换律 $A \cup B = B \cup A$，$A \cap B = B \cap A$.

结合律 $A \cup (B \cup C) = (A \cup B) \cup C$，$A \cap (B \cap C) = (A \cap B) \cap C$.

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

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

摩根律 $\overline{A \cup B} = \overline{A} \cap \overline{B}$，$\overline{A \cap B} = \overline{A} \cup \overline{B}$.

一般地，事件运算时的定律可推广到有限，如

$$\overline{A_1 \cup A_2 \cup \cdots \cup A_n} = \overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_n},$$

$$\overline{A_1 \cap A_2 \cap \cdots \cap A_n} = \overline{A_1} \cup \overline{A_2} \cup \cdots \cup \overline{A_n}.$$

关于事件之间的关系、运算与集合之间的关系及运算的类比，见表 1-1.

表 1-1　事件之间的关系、运算与集合之间的关系及运算的类比

符号	概率论	集合论
S	样本空间或必然事件	空间（全集）
$\varnothing$	不可能事件	空集
ω	基本事件（样本点）	元素
A	事件 A	集合 A
$\overline{A}$	A 的对立事件	A 的余集
$A \subset B$	事件 A 发生必导致事件 B 发生	A 是 B 的子集
$A = B$	事件 A 与事件 B 相等	A 与 B 相等
$A \cup B$	事件 A 与事件 B 至少有一个发生	A 与 B 的并集
$A \cap B$	事件 A 与事件 B 同时发生	A 与 B 的交集
$A - B$	事件 A 发生而事件 B 不发生	A 与 B 的差集
$AB = \varnothing$	事件 A 和事件 B 互不相容	A 与 B 不相交

例 5　设 A,B,C 为 S 中的随机事件，试用 A,B,C 表示下列事件：

(1) A 与 B 发生而 C 不发生；

(2) A 发生，B 与 C 不发生；

(3) A,B,C 中恰有一个事件发生；

(4) A,B,C 中恰有两个事件发生；

(5) A,B,C 三个事件都发生；

(6) A,B,C 中至少有一个事件发生；

(7) A,B,C 都不发生；

(8) A,B,C 不都发生；

(9) A,B,C 中不多于一个发生；

(10) A,B,C 中不多于两个发生.

解　(1) $AB\overline{C}$；

(2) $A\overline{B}\,\overline{C}$；

(3) $A\overline{B}\,\overline{C} \cup \overline{A}B\overline{C} \cup \overline{A}\,\overline{B}C$；

(4) $AB\overline{C} \cup A\overline{B}C \cup \overline{A}BC$；

(5) ABC；

(6) $A \cup B \cup C$;

(7) $\overline{A}\ \overline{B}\ \overline{C}$;

(8) $\overline{ABC}$;

(9) $\overline{AB \cup BC \cup CA}$;

(10) $\overline{ABC}$.

例 6 试验 E:袋中有三只球编号分别为 1,2,3,从中任意摸出一球,观察其号码,记 $A=\{$球的号码小于 3$\}$,$B=\{$球的号码为奇数$\}$,$C=\{$球的号码为 3$\}$.试问:

(1) E 的样本空间是什么?

(2) A 与 B,A 与 C,B 与 C 是否互不相容?

(3) A,B,C 对立事件是什么?

(4) A 与 B 的和事件、积事件、差事件各是什么?

解 设 $\omega_i=\{$模到球的号码为 $i\}$ $(i=1,2,3)$.

(1) E 的样本空间为 $S=\{\omega_1,\omega_2,\omega_3\}$;

(2) $A=\{\omega_1,\omega_2\}$,$B=\{\omega_1,\omega_3\}$,$C=\{\omega_3\}$,A 与 B,B 与 C 是相容的,A 与 C 互不相容;

(3) $\overline{A}=\{\omega_3\}$,$\overline{B}=\{\omega_2\}$,$\overline{C}=\{\omega_1,\omega_2\}$;

(4) $A\cup B=S$,$AB=\{\omega_1\}$,$A-B=\{\omega_2\}$.

2 事件的概率

在几何学中线段的长短、平面图形或立体的大小,物理学中物质的多少、质点运动的快慢等,都可以用数值来度量,长度、面积、体积、质量、速度等就是相应的度量.事件在试验中出现的可能性大小,也应该可以用数值度量,这种度量就是“概率”.

概率与长度、面积、体积、质量、速度一样,也是一种度量.具体地说,概率是事件在试验中出现可能性大小的数值度量,我们用 $P(A)$ 表示事件 A 的概率.例如有一批共 100 件产品,其中有 5 件不合格品,则从中随意抽出一件,恰好抽到不合格品的可能性显然是 5%.这时,用 5% 作为事件 $A=\{$抽到不合格品$\}$ 出现的可能性大小的数值度量,即

$$P(A)=P\{\text{抽到不合格品}\}=0.05.$$

在明确概率概念之后,需要解决如何合理地选择或确定这种度量的问题.确定事件在试验中出现的可能性大小的数值度量——事件的概率,大致有以下几种途径:① 直接计算,利用试验结局的某种等可能性或均衡性,直接计算事件的概率,最常用的有古典概型和几何概型;② 用频率估计概率;③ 推算概率,利用概率的性质和公式,由较简单事件的概率推算较复杂事件的概率;④ 利用随机变量的

分布,由随机变量的已知概率分布求与其相联系事件的概率(见第二章).

2.1 古典概型

以掷质地均匀的硬币为例,人们自然想到由于硬币两面是对称的,所以出现正面或反面的可能性都是 0.5. 在概率论研究的初始阶段,主要讨论的随机事件都和上面例子一样具有两条性质:

(1) 试验的可能结果只有有限个;

(2) 试验的每个可能结果在每次试验中发生是等可能的.

对于任意事件 A,对应的概率 $P(A)$ 由下式计算:

$$P(A)=\frac{\text{事件 }A\text{ 包含的基本事件数 }k}{S\text{ 中基本事件总数 }n}=\frac{k}{n}, \tag{1.2.1}$$

并把它称为**古典概型**.

在计算古典概型的概率时,主要是利用排列、组合来求数 k 与 n. 要注意在计算 k 与 n 时是用排列还是组合,与次序是否有关要一致.

例 1 将两颗骰子掷一次,求它们点数之和为 6 点的概率.

解 点数之和为 6 点的事件记为 A,将两颗骰子掷一次,如果考虑其点数之和,其样本空间为

$$S_1=\{2,3,4,\cdots,12\};$$

如果考虑其点数组合,其样本空间为

$$S_2=\{(1,1),(1,2),\cdots,(6,6)\}\text{(共 36 种情形)}.$$

对于 S_1 出现的样本点,从直观上看,各个点数和概率相同显然是错误的. 而 S_2 中出现的各个样本点,从对称性可知其可能性相同. 故所求概率不能直接利用 S_1,而需利用 S_2.

在 S_2 中基本事件总数为 36 种,而点数和为 6 点包含(1,5),(2,4),(3,3),(4,2),(5,1) 5 种情形,故 A 的概率为

$$P(A)=\frac{5}{36}.$$

当样本空间样本点不满足等可能性时,经常考虑转化为另一个满足等可能性的样本空间.

例 2 设袋中有 N 件产品,其中有 M 件次品,从中取产品 n 次,每次随机地取一件. 考虑两种抽取方式:

(1) 先取一件产品后,观察它是否为正品,然后放回袋中,搅匀后再取一件,直到取出 n 件为止,这种抽取方式叫做**有放回抽样**;

(2) 先取一件产品后,观察它是否为正品,然后从剩余的产品中再取一件,直到取出 n 件为止,这种抽取方式叫做**不放回抽样**.

试分别就上面两种抽样方式，求取到 m 件次品的概率.

解 (1) 先考虑有放回情形，此时样本空间 S 的样本点为，第一次抽取时，有 N 种取法；第二次抽取时，仍有 N 种取法……如此下去，一共抽取 n 次，总共有 N^n 种等可能的样本点. 记 A_m 为抽出的产品中有 m 件次品，则 A_m 所含样本点数为，首先在 n 次中选择 m 次，其选择方式有 C_n^m 种，在这 m 次取次品，其方法数为 M^m，然后再在剩下的 $n-m$ 次都取正品，有 $(N-M)^{n-m}$ 种方法，故 A_m 的样本点数为 $C_n^m M^m (N-M)^{n-m}$，则

$$P(A_m)=\frac{C_n^m M^m (N-M)^{n-m}}{N^n}=C_n^m\left(\frac{M}{N}\right)^n\left(1-\frac{M}{N}\right)^{n-m}.$$

(2) 再考虑不放回情形，其中 $m\leqslant M, m\leqslant n$. 先计算样本空间 S 的样本点，从 N 件产品中仍取 n 件，不讲次序，所有样本点的总数为 C_N^n，记 B_m 为抽出的产品中有 m 件次品，则 B_m 所含样本点数为 $C_M^m C_{N-M}^{n-m}$，可得

$$P(B_m)=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n}\quad (m\leqslant \min(n,M)).$$

(1) 中的分布称为**二项分布**，(2) 中的分布称为**超几何分布**.

例 3 一袋中有 a 只红球，b 只白球，从中任意地连续摸出 k 只球，每次摸出的球不放回袋中，试求最后一次摸到红球的概率.

解 将袋中每个球都看成是可以分辨的，抽球与次序有关，则样本空间的基本事件总数为 P_{a+b}^k. 设 $A=\{$第 k 次摸到红球$\}$，则 A 所含基本事件数为先从 a 只红球中任取一只红球，排在最后位置上，有 a 种方法；然后从剩下的 $a+b-1$ 只球中取 $k-1$ 只球任意放在 $k-1$ 个位置，有 P_{a+b-1}^{k-1} 种方法. A 含 $a\times P_{a+b-1}^{k-1}$ 个基本事件，则

$$P(A)=\frac{a\times P_{a+b-1}^{k-1}}{P_{a+b}^k}=\frac{a}{a+b}.$$

从本例可以看出，摸到红球的概率与次序 k 无关，从而可以得到，在抽签中不论先后，每个人抽中的机会都是一样的，所以不必争先恐后.

进一步地，当采用有放回抽球时，每次抽到红球的概率也是 $\frac{a}{a+b}$.

2.2 概率的统计定义

一个随机事件在某次试验中由于受许多无法控制的随机因素影响，不能断言它是否发生. 若在相同条件下，进行了 n 次试验，则将在这 n 次试验中事件 A 发生的次数记为 n_A，称之为事件 A 发生的**频数**，比值 $\frac{n_A}{A}$ 称为事件 A 发生的**频率**，并记为 $f_n(A)$.

由于事件 A 发生的频率是它发生次数与试验次数之比，其大小表示事件 A 发生的频繁程度，频率大就意味着事件 A 在一次试验中发生的可能性大，反之亦然. 因而，

直观的想法是用频率来表示事件 A 在一次试验中发生的可能性的大小.

但实际上,事件 A 在 n 次试验中发生的频率受许多偶然性因素的影响,其表现为 $f_n(A)$ 在 0 与 1 之间随机波动,但这种随机波动当试验次数越小,其幅度越大;试验次数越大,其幅度越小,亦即随着试验次数的增多,频率稳定在某一常数附近,这种“频率稳定性”即通常所说的统计规律性.

历史上通过“掷一枚硬币”的试验来观察“出现正面”这一事件发生的规律,表 1-2 是试验结果记录.

表 1-2　历史上“掷一枚硬币”的试验结果

实验者	投掷次数	出正面次数	频率
蒲丰	4 040	2 048	0.506 9
皮尔逊	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5

又如考察英文中特定字母出现的频率,当观察字母的个数 n 较小时,频率有较大的随机波动,但当 n 增大时,频率呈现出稳定性,表 1-3 是一份英文字母出现频率的统计表.

表 1-3　英文字母出现频率的统计表

字母	频率	字母	频率
E	0.126 8	F	0.025 6
T	0.097 8	M	0.024 4
A	0.078 8	W	0.021 4
O	0.077 6	Y	0.020 2
I	0.070 7	G	0.018 7
N	0.070 6	P	0.018 6
S	0.063 4	B	0.015 6
R	0.059 4	V	0.010 2
H	0.057 3	K	0.006 0
L	0.039 4	X	0.001 6
D	0.038 9	J	0.001 0
U	0.028 0	Q	0.000 9
C	0.026 8	Z	0.000 6

大量实验证实:在多次重复试验中,同一事件发生的频率虽然并不相同,但却

在一个固定的数值附近摆动，呈现出一定的稳定性，而且随着重复试验次数的增加，这种现象愈加显著. 频率所接近的这个固定数值，就可作为相应事件的概率. 随机事件 A 的概率记为 $P(A)$.

定义 1 在大量重复试验中，如果一个事件 A 发生的频率稳定在某一常数 p 附近，这个数 p 就称为 A 的**概率**，记为 $P(A)=p$.

概率的统计定义从直观上给出了概率的定义，但它却有理论和应用上的缺陷. 从理论上说，频率为什么具有稳定性呢？(本问题将在第四章大数定律给出理论上证明.) 在应用上，我们没有理由认为，试验 $n+1$ 次来计算频率总会比试验 n 次更准确、更逼近所求的概率. 因此，我们不知道 n 取多大才行，如果 n 要很大，我们不一定能保证每次试验的条件完全一样，并且从感觉上讲，概率的统计定义也不像一种很严格的数学定义. 因而，要讨论概率的更加严格的数学化定义，下面介绍概率的公理化定义.

2.3 概率的公理化定义

定义 2 设 E 是随机试验，S 是它的样本空间，对于 E 的每一事件 A 赋予一个实数，记为 $P(A)$，称为事件 A 的**概率**(实际上，$P(A)$ 是一个定义域为集合、值域为 $[0,1]$ 的集合函数，简称集函数)，并且集函数 $P(\cdot)$ 满足下列条件：

(1) **非负性**，对于每一个事件 A，有 $P(A)\geqslant 0$；

(2) **规范性**，对于必然事件 S，有 $P(S)=1$；

(3) **可列可加性**，设 $A_1,A_2,\cdots$ 是两两互不相容的事件，即对于 $i\neq j$，$A_iA_j=\varnothing(i,j=1,2\cdots)$，有

$$P(A_1\cup A_2\cup\cdots)=P(A_1)+P(A_2)+\cdots,$$

即

$$P\left(\bigcup_{k=1}^{\infty}A_k\right)=\sum_{k=1}^{\infty}P(A_k). \tag{1.2.2}$$

下面，由概率的公理化定义，得到概率的一些重要性质.

性质 1

$$P(\varnothing)=0. \tag{1.2.3}$$

证 令 $A_n=\varnothing\ (n=1,2,\cdots)$，则 $\bigcup\limits_{k=1}^{\infty}A_n=\varnothing$，且 $A_iA_j=\varnothing\ (i\neq j;i,j=1,2,\cdots)$，由可列可加性知

$$P(\varnothing)=P\left(\bigcup_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}P(A_n)=\sum_{n=1}^{\infty}P(\varnothing),$$

得

$$P(\varnothing)=0.$$

性质 2(有限可加性) 若 $A_1,A_2,\cdots,A_n$ 是两两互不相容的事件，则有

$$P\left(\bigcup_{k=1}^{n} A_k\right)=\sum_{k=1}^{n} P(A_k). \tag{1.2.4}$$

证 令 $A_{n+1}=A_{n+2}=\cdots=\varnothing$，则由可列可加性，得

$$P\left(\bigcup_{k=1}^{n} A_k\right)=P\left(\bigcup_{k=1}^{\infty} A_k\right)=\sum_{k=1}^{\infty} P(A_k)=\sum_{k=1}^{n} P(A_k)+0=\sum_{k=1}^{n} P(A_k).$$

性质 3（单调性） 设 A,B 是两个事件，若 $A\subset B$，则有

$$P(A)\leqslant P(B). \tag{1.2.5}$$

证 因 $B=A\cup(B-A)$，$A\cap(B-A)=\varnothing$，则由性质 2 知

$$P(B)=P(A)+P(B-A),$$

又由非负性知

$$P(B-A)\geqslant 0,$$

故

$$P(B)\geqslant P(A).$$

注 式中等号即使在 A 是 B 的真子集的情况下也不可省.

性质 4 设 A,B 是两个事件，若 $A\subset B$，则有

$$P(B-A)=P(B)-P(A). \tag{1.2.6}$$

证 由性质 3 证明可得.

性质 5 对于任一事件 A，$P(A)\leqslant 1$.

证 因 $A\subset S$，由性质 3 易得.

性质 6（逆事件的概率） 对于任一事件 A，有

$$P(\overline{A})=1-P(A). \tag{1.2.7}$$

证 因 $S=A\cup\overline{A}$，$A\cap\overline{A}=\varnothing$，由性质 2 知

$$P(S)=1=P(A)+P(\overline{A}),$$

于是

$$P(\overline{A})=1-P(A).$$

性质 7（加法公式） 对于任意两事件 A,B，有

$$P(A\cup B)=P(A)+P(B)-P(AB). \tag{1.2.8}$$

证 因 $A\cup B=A\cup(B-AB)$，且 $A\cap(B-AB)=\varnothing$，$AB\subset B$，故由(1.2.4)式及(1.2.6)式，得

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

推论 1 对于任意 n 个事件 $A_1,A_2,\cdots,A_n$，有

$$P\left(\bigcup_{k=1}^{n} A_k\right)=\sum_{k=1}^{n} P(A_k)-\sum_{1\leqslant i<j\leqslant n}^{n} P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}^{n} P(A_iA_jA_k)+\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n). \tag{1.2.9}$$