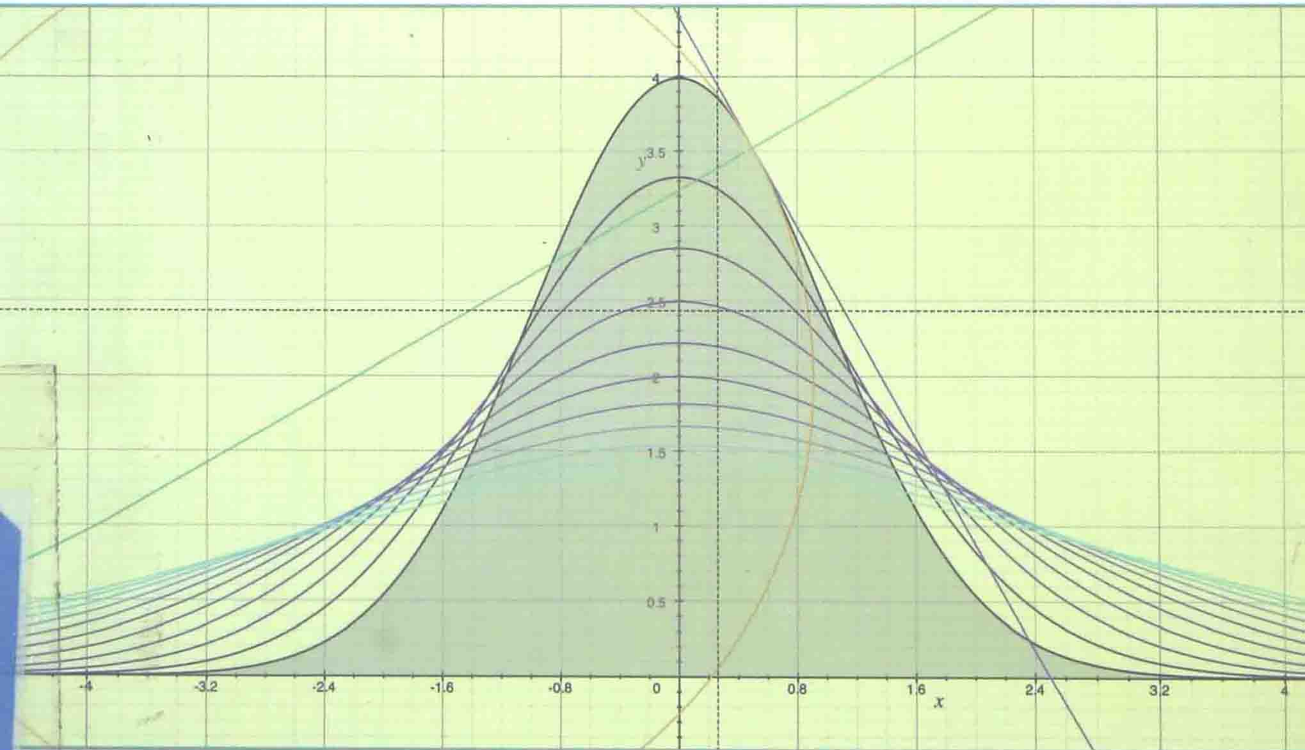
普通高等教育“十一五”国家级规划教材

医药数理统计方法

（第二版）

主 编 祝国强

副主编 刘庆欧



高等教育出版社
HIGHER EDUCATION PRESS

普通高等教育“十一五”国家级规划教材

医药数理统计方法

（第二版）

主　编　祝国强

副主编　刘庆欧

编者（按章节顺序）

杭国明	复旦大学
周书云	南方医科大学
张学良	新疆医科大学
丁　勇	南京医科大学
杨　洁	北京中医药大学
刘庆欧	山西医科大学
郭东星	山西医科大学
罗明奎	第三军医大学
祝国强	第二军医大学
滕海英	第二军医大学
王培承	潍坊医学院

高等教育出版社

内容简介

本书作为普通高等教育“十一五”国家级规划教材，是专门为高等医药类院校本科教育所编写的数学基础课程教材。本版是在教材第一版的基础上，根据第一版在使用过程中的反馈意见修订而成的。

本书系统而简要地介绍了基础概率和统计方法两大部分内容。分为随机事件及其概率、随机变量及其分布、随机变量的数字特征、随机抽样及抽样分布、抽样估计、假设检验、方差分析、正交试验设计与分析、相关与回归分析共九章。本书的特点是内容涵盖广泛，论理深入浅出。与现有的《医药数理统计方法》教材相比有了较大改进与充实，既坚持了数理统计的传统内容，又扩充了一些实用统计方法，有利于数理统计与卫生统计的衔接与沟通。

本书可供高等医药类院校药学、生物技术、中药等各本科专业（含专升本）作教材使用，也可供相关专业的本科及研究生选用，从事医药卫生工作的科技人员也可学习参考。

图书在版编目（CIP）数据

医药数理统计方法/祝国强主编．—2 版．—北京：高等教育出版社，2009.1
ISBN 978-7-04-024858-6

Ⅰ.医…　Ⅱ.祝…　Ⅲ.数理统计—应用—医药学—高等学校—教材　Ⅳ.R311

中国版本图书馆 CIP 数据核字（2008）第 188908 号

策划编辑 宋瑞才　**责任编辑** 蒋　青　**封面设计** 张　楠　**责任绘图** 吴文信
版式设计 王　莹　**责任校对** 金　辉　**责任印制** 陈伟光

出版发行 高等教育出版社
社　　址 北京市西城区德外大街 4 号
邮政编码 100120
总　　机 010-58581000

经　　销 蓝色畅想图书发行有限公司
印　　刷 中青印刷厂

开　　本 850×1168　1/16
印　　张 21.25
字　　数 530 000

购书热线 010-58581118
免费咨询 800-810-0598
网　　址 http://www.hep.edu.cn
http://www.hep.com.cn
网上订购 http://www.landraco.com
http://www.landraco.com.cn
畅想教育 http://www.widedu.com

版　　次 2004 年 8 月第 1 版
2009 年 1 月第 2 版
印　　次 2009 年 1 月第 1 次印刷
定　　价 26.90 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换。

物料号　24858-00

第二版前言

本书是在第一版的基础上，根据几年来的教学改革实践经验全面修订而成的。在修订过程中，保留了原教材的体系和风格，继承并发扬了原教材内容涵盖广泛、理论深入浅出、简明实用、便于教学等优点。同时，新版教材在教学内容和实例上积极开拓创新，使得其更适应当前教学实际的需要。

新版教材增补了部分内容，如最佳普查方案设计、数学期望性质的应用等。新版教材也对部分内容进行了改写，如第六章假设检验重点介绍了临界值法，而把 P 值法及原假设为不等式形式的单侧检验单独列为一节；各章增加了概率统计与医药学结合的实例及习题。同时，新版教材对部分内容也进行了精简，如正态概率纸检验。对第一版中存在的个别问题，这次也一并作了修订。

在新版教材的编写过程中，第二军医大学黄平、刘沛两位老师参与了教材的整理、校对等工作，谨致谢忱。

在本书的修订过程中，许多医科院校的教师提出了不少宝贵的意见和建议，在此表示诚挚的感谢。对新版教材中存在的问题，欢迎广大专家、同行和读者批评指正。

编　者

二〇〇八年九月

第一版序

我们今天生活在一个迅速变化的世界上。如果你关心时事，经常看报、上网或收听新闻广播，就会发现数字的使用十分频繁。只要涉及“何时”、“何地”以及高产、优质、发明创造、企业改制之类的问题，通常都要用数字去回答。为了更好地理解世界上已经、正在或即将发生的事情，寻找挖掘自然与人文的内在规律，学习与应用统计学实在是人生舞台上所必须掌握的知识链中的重要环节。

譬如说，天气的好坏是人们的日常话题，通常是指阴、晴、雨、雾、霜雪，气温，湿度以及紫外线强度，乃至感冒指数等等。天气变化是一种非确定性现象，气象预报必须在研究了大片地区上搜集到的气象资料、卫星云图等之后才能进行分析与预报。就目前的发展水平而言，天气预报不可能做到百分之百的正确。从本质上说，它是一种或然现象。

大家都很关心经济生活，例如商业行情、物价、工资、货币汇率、存贷款利率、股票的涨跌、房价的飙升等等。教育行政部门也必须研究种种数据，以便确立何时何地建造新的学校、需要多少投入等等。总而言之，许多事情都必须在公平、公正、透明的基础上，应用强有力的统计武器来作出正确的决策。

在品尝美味佳肴之前，你得先咬一口。为了判断西瓜的质量，许多人的习惯做法是在西瓜上挖下一小块来试味。其他情况更是举不胜举：电视节目的收视率、散文与小说的评奖、民意测验之类的活动，这些都是某些社团为了收集公众意见而进行抽样调查的例子。机关、工厂、学校、大卖场、规划部门……一般都会通过抽取样本对总体进行各式各样的统计计算、预测结果、评估产品的质量等等。如果抽取随机样本时采用了不适当的方法，往往就会得出错误乃至荒谬的结果，此类例子屡见不鲜。

历史的教训值得记取。1936 年，美国有一家非常畅销的杂志曾经预言：兰登将会当选美国总统，但结果恰恰相反，他的竞选对手富兰克林·罗斯福赢得了美国大选。犯错误的原因是根据大选前夕所作的一次民意测验得出的，而参加这次民意测验的人士是从电话簿上随机抽取而来。然而，1936 年正是美国“大萧条”的一年，只是有钱人家才能装得起电话。归根结底，这次民意测验所依据的“样本”只是有钱人的代表，而不是全体选民的代表，得出的结论之所以大错特错，当然就不奇怪了。

几个世纪以来，概率统计已从一个较脆弱的、不成熟的数学分支日夜长大，发展成为一个内容深刻宽广、影响深远的学科。它从侏儒变成巨人，已经对生物学、医药学、军事运筹学、经济学和心理学、遗传学、地质学、矿物学……的数学化起着中心作用，甚至已经渗透到语言学、体育竞技与文学、艺术的领域。随着时间的推移，概率论与数理统计学已经成为理、工、医、农等各类高等院校中一门必修的学科，它安营扎寨，坐稳了自己的位置。

偶然性的背后往往隐藏着必然的联系。现在，数理统计业已成为实验科学与应用技术研究中必不可少的工具。它在数据变异较大的医药学和生物科学研究中，尤其显得重要。

良好的实验设计是得出科研成果的重要因素和前提。高水平的科研人员不仅应该熟悉自身的专业，还应掌握假设检验、方差分析、多元分析等统计技巧，在实验开始以前以及在进行过

程中都需要认真安排实验设计，才能事半功倍。众所周知，医学科学中的临床实验要比其他科学的实验更为困难得多，无论在环境条件、病例多寡、用药治疗以及效果检查等方面都要受到很大的制约，必须牢固树立人本思想，珍视病人的生命和健康。这就要求当事人了解各种可靠而有效的数理统计的新进展与新方法，例如平稳随机过程、分支过程、随机微分方程、信息熵以及统计决策理论等。一些比较传统的方法如拉丁方、区组设计等也要在大量的应用实践中经受不断的考验、精炼与提高。

本书编者大多来自全国各著名医学院校，拥有深厚的学识和素养以及丰富的教学经验。他们分工合作、集思广益、取精用宏、在广泛汲取以往各种教材优点的基础上编出了适合当前形势、符合客观实际的新教材。本书说理通畅、步骤完整、图表齐全、例题充实，在选择、组织材料方面能做到由浅入深、循序渐进，并顾及前后章节之间的联系，从而易教易学，在新世纪教学改革中迈出了坚实而可喜的一步。

笔者由衷地希望大家会使用它、喜爱它，并在教学实践中对其不断加以修正和提高，使之更臻于完善。

谈祥柏

二〇〇三年十二月三十一日

写于上海大华新村南华苑

第一版前言

概率论与数理统计都是研究随机现象数量规律性的学科。在药学、医学和卫生科技工作中有着广泛的应用。同时，它也是药理学、毒理学、药物动力学等课程的前期基础理论课。

本书以概率为基础、统计推断为中心，重点介绍常用的统计分析方法，包括随机事件及其概率、随机变量及其分布、随机变量的数字特征、随机抽样及抽样分布、抽样估计、假设检验、方差分析、正交试验设计与分析、相关与回归分析共九章内容，书后附有统计软件应用介绍。讲授本书全部内容约需72学时。考虑到我国现有的多种教育层次和医教研卫战线上大量的卫生统计工作者的需求，本书内容涵盖广泛、论理深入浅出，既坚持了传统数理统计的内容，又扩充了一些实用的统计方法，在数理统计与卫生统计的沟通与衔接方面做了有益的尝试。

本书可供高等医药院校药学本科（含专升本）作教材使用，也可供非药学专业本科生及研究生选用，从事医药卫生工作的科技人员也可学习参考。

本书由浙江大学周怀梧教授和第二军医大学谈祥柏教授主审。他们提出了许多中肯而宝贵的意见，在此谨表感谢。第二军医大学高慕勤、徐玲玲两位老师参与了教材的谋划、整理、校对等工作，谨致谢忱。在编辑出版过程中，本书还得到高等教育出版社的热情支持，在此一并致谢。

在编写过程中，本书参考了大量的同类书刊并借鉴了同行们的经验，在此深表谢意。限于编者的水平，书中定有不少缺点、错误，热切希望使用本书的师生和广大读者提出宝贵的批评和建议。

编　者

二〇〇三年十月

目　录

第一章　随机事件及其概率

概率论与数理统计是研究随机现象统计规律的一门数学学科．在医药学领域中，它有着极其广泛的应用，是医药工作者必备的知识．本章首先由随机试验引出研究概率统计中最基本的两个概念——随机事件及其概率．为了能够从简单事件的概率出发，计算复杂事件的概率，本章引进了随机事件的关系与运算，并且讨论了它们的性质．

第一节　随机事件及其运算

一、随机试验

自然界里有各种现象，它们大致可分为两类：一类为确定性现象，另一类为随机现象．什么是确定性现象？什么是随机现象？让我们先做两个简单的试验．

试验Ⅰ：一个盒子中有 10 个完全相同的白球，搅匀后从中任意取一球．

试验Ⅱ：一个盒子中有 10 个相同的球，其中 5 个是白色的，另 5 个是黑色的，搅匀后从中任意取一球．

对于试验Ⅰ，在球没取出之前，我们就能确定取出的球必定是白球．试验Ⅰ所代表的类型——根据试验的条件，在试验之前就能断定它有一个确定的结果，这类试验称为**确定性试验**．确定性试验所对应的现象，即在一定条件下，必然发生或绝不可能发生的现象，称为**确定性现象**．确定性现象非常广泛，例如：

“早晨，太阳必然从东方升起．”（不考虑地球南、北极的情况）

“地球上，在标准大气压下，100℃的水必然沸腾．”

“边长为 a，b 的矩形，其面积必为 ab．”

“两奇数之和为奇数．”

过去我们所学的各门数学课程基本上都是用来处理和研究这类确定性现象的．

对于试验Ⅱ，在球没取出之前，我们不能确定试验的结果（即取出的球的颜色）是白色还是黑色．试验Ⅱ所代表的类型——根据试验的条件，它有多于一种可能的试验结果，在一次试验之前不能确定试验会出现哪一种结果．就一次试验而言，看不出有什么规律，但是，“大数次”地重复这个试验，试验结果又遵循某些规律，称这种规律为“统计规律”．这类试验称为**随机试验**（random trial），随机试验所对应的现象称为**随机现象**．随机现象到处可见，例如：

“某地区的年降雨量．”

“抛一枚硬币，出现正面、反面的情况．”

“某种药物对一种疾病的治疗效果．”

概率论与数理统计就是研究随机现象统计规律的一门数学学科．由于随机现象的普遍性，使得概率统计得到了极其广泛的应用．

上面我们对随机试验作了一般的介绍，现在就随机试验给出一个明确的定义．一个试验如

果满足下列条件：

(1) 试验可在相同的条件下重复进行；

(2) 每次试验的可能结果不止一个，并可事先明确知道试验的所有可能结果；

(3) 进行一次试验之前不能确定哪一个结果会出现.

我们把具有上述性质的试验称为**随机试验**，简称为**试验**.

二、样本空间

在随机试验中，它的每一个可能出现的直接结果，称为**基本事件**，或称**样本点**（sample point），一般用字母 e 表示. 因为随机试验的所有可能结果是明确的，从而所有的基本事件也是明确的. 随机试验的所有基本事件组成的集合称为**样本空间**（sample space），通常用 Ω 表示.

例 1.1 在前述试验Ⅱ中，令

$$e_1=\{\text{取得白球}\},\ e_2=\{\text{取得黑球}\},$$

则样本空间

$$\Omega=\{e_1,e_2\}.$$

例 1.2 掷一枚硬币，观察出现正反面的情况，试验的可能结果有两个：“正面向上”、“反面向上”. 记

$$e_1=\{\text{正面向上}\},\ e_2=\{\text{反面向上}\},$$

则样本空间

$$\Omega=\{e_1,e_2\}.$$

例 1.3 袋中装有 2 只白球（编号为 1 和 2）和 1 只黑球（编号为 3），今从袋中依次不放回地任意摸出两球. 用一有序数组 (i,j) 来表示可能的结果，i 表示第一次摸到 i 号球，j 表示第二次摸到 j 号球，则样本空间

$$\Omega=\{(1,2),(1,3),(2,1),(2,3),(3,1),(3,2)\}.$$

例 1.1，例 1.2 和例 1.3 的样本空间包含的样本点只有有限个.

例 1.4 投掷一枚硬币，直到出现正面为止，观察已经投掷的次数，则该试验的样本空间

$$\Omega=\{1,2,3,\cdots\}.$$

例 1.5 观察某电话交换站在单位时间内收到的呼唤次数，则样本空间

$$\Omega=\{0,1,2,\cdots\}.$$

例 1.4 和例 1.5 的样本空间都是有无限个样本点的样本空间.

例 1.6 测试某种型号节能灯泡的使用寿命. 用 T 表示节能灯泡的使用寿命，此试验的样本空间

$$\Omega=\{T\mid 0\leqslant T<+\infty\}.$$

三、随机事件

在随机试验中，有时关心的是具有某些特征的集合是否发生. 如在例 1.5 中，我们考虑的问题是

$$A=\{\text{某电话交换站在单位时间内收到的呼唤次数是 2 次}\},$$

$B=${某电话交换站在单位时间内收到的呼唤次数不超过 2 次}，

$C=${某电话交换站在单位时间内收到的呼唤次数多于 2 次}.

这些结果是否发生？其中 A 表示一个基本事件，而 B 与 C 是样本空间的某个子集，由多个基本事件所组成．基本事件也称为**简单事件**，相对于简单事件，称 B 和 C 为**复杂事件**．无论是简单事件还是复杂事件，它们在试验中发生与否都带有随机性，所以都叫做**随机事件**(random event)，简称为**事件**．通常，随机事件用大写英文字母 A，B，C 等表示．

在一次试验中，如果出现事件 A 中所包含的某一个样本点 e，则称事件 A 发生，并记作 $e\in A$．由于样本空间 Ω 包含了所有基本事件，而在任何一次试验中，必然会出现 Ω 中的某一基本事件 e，即 $e\in\Omega$，也就是说，在任何一次试验中，Ω 必然会发生，所以我们又称 Ω 为**必然事件**．我们用 $\varnothing$ 表示一个事件，$\varnothing$ 中不包含任何基本事件．所以，在任何一次试验中出现的任何基本事件 e 都不属于 $\varnothing$，$\varnothing$ 也就永远不可能发生，所以称 $\varnothing$ 为**不可能事件**．必然事件和不可能事件的发生与否，已经失去了“不确定性”，本质上它们已经不属于随机事件，但是为了讨论问题方便起见，我们还是把它们作为两个特殊的随机事件来处理．必然事件和不可能事件可理解为随机事件的两个极端情况．

四、事件的关系和运算

在一个样本空间中，可以有多个随机事件．概率论的任务之一，就是研究随机事件的统计规律，通过对较为简单事件规律的研究去掌握更复杂的事件的规律．为此，我们需要研究事件之间的相互关系和运算．

随机事件之间的关系和运算与布尔（Boole）代数中集合之间的关系和运算是完全类似的．20 世纪 30 年代初，冯·米泽斯（Von Mises）开始用集合论的观点来研究随机事件，并使得随后概率论的研究走上了严格化的道路．为了便于比较事件的直观意义与集合论的定义，我们把各符号的两种解释列表如下（见表 1.1）．

表 1.1

符　号	概率论的解释	集合论的解释
Ω	样本空间、必然事件	全集
$\varnothing$	不可能事件	空集
e	样本点、基本事件	元素、点
A	事件 A	Ω 的子集 A
$e\in A$	事件 A 发生	e 是 A 中的元素
$e\notin A$	事件 A 不发生	e 不是 A 中的元素

（一）事件的关系和运算

1. 包含关系

事件 A 发生必然导致事件 B 发生，则称事件 B **包含**事件 A（或称事件 A **包含于**事件 B），记为 $B\supset A$（或 $A\subset B$）.

例如，在例 1.5 中，$A=${某电话交换站在单位时间内收到的呼唤次数是 2 次}，$B=${某电话交换站在单位时间内收到的呼唤次数不超过 2 次}，事件 A 发生必然导致事件 B 发

生， 所以，$B \supset A$.

2. 相等关系

若事件 B 包含事件 A，且事件 A 又包含事件 B，即 $B \supset A$ 且 $A \supset B$，则称事件 A 与事件 B **相等**，记为 $A=B$.

例 1.7 投掷一枚骰子，设事件 $A=$｛骰子出现的点数是奇数｝， $B=$｛骰子出现的点数是 1 或 3 或 5｝， 显然，$B \supset A$ 且 $A \supset B$，所以事件 A 与事件 B 相等.

3. 事件的并（和）

若事件 A 与事件 B 至少有一个发生，这一事件称为事件 A 与事件 B 的**并**（或**和**），记为 $A \cup B$（或 $A+B$).

在例 1.7 中，设事件 $C=$｛骰子出现的点数是 5 或 6｝，则 $B \cup C=$｛骰子出现的点数是 1 或 3 或 5 或 6｝.

类似地，事件 $A_1, A_2, \cdots, A_n$ 中至少有一个发生，这一事件称为 $A_1, A_2, \cdots, A_n$ 这 n 个事件的并（或和），记为 $A_1 \cup A_2 \cup \cdots \cup A_n$（或 $A_1+A_2+\cdots+A_n$），简记为 $\bigcup\limits_{i=1}^{n} A_i$ 或 $\sum\limits_{i=1}^{n} A_i$.

4. 事件的交（积）

若事件 A 与事件 B 同时发生，这一事件称为事件 A 与事件 B 的**交**（或**积**），记为 $A \cap B$（或 AB).

例如，在例 1.7 中，事件 $A=$｛骰子出现的点数是奇数｝，$C=$｛骰子出现的点数是 5 或 6｝，则 $A \cap C=$｛骰子出现的点数是 5｝.

类似地，事件 $A_1, A_2, \cdots, A_n$ 同时发生，这一事件称为 $A_1, A_2, \cdots, A_n$ 这 n 个事件的交（或积），记为 $A_1 \cap A_2 \cap \cdots \cap A_n$（或 $A_1 A_2 \cdots A_n$），简记为 $\bigcap\limits_{i=1}^{n} A_i$（或 $\prod\limits_{i=1}^{n} A_i$ ）.

5. 事件的差

若事件 A 发生而事件 B 不发生，这一事件称为事件 A 与事件 B 的**差**，记为 $A-B$.

例如，在例 1.7 中，事件 $A=$｛骰子出现的点数是奇数｝，$C=$｛骰子出现的点数是 5 或 6｝，则 $A-C=$｛骰子出现的点数是 1 或 3｝，$C-A=$｛骰子出现的点数是 6｝.

6. 互斥关系

若事件 A 与事件 B 不能同时发生，即 $A \cap B=\varnothing$，称该情况为事件 A 与事件 B **互斥**（或**互不相容**).

例如，在例 1.7 中，事件 $A=$｛骰子出现的点数是奇数｝，设事件 $D=$｛骰子出现的点数是 4 或 6｝，此时 $A \cap D=\varnothing$，即事件 A 与事件 D 互斥.

7. 互逆关系

若事件 A 与事件 B 互斥，且在任何一次试验中二者必定有一个发生，即 $A \cap B=\varnothing$ 且 $A \cup B=\Omega$，则称该情况为事件 A 与事件 B **互逆**（或**相互对立**). 称事件 B 为事件 A 的**对立事件**，记为 $\overline{A}$，即 $B=\overline{A}$，由对称性知，$A=\overline{B}$.

例如，在例 1.7 中，事件 $A=$｛骰子出现的点数是奇数｝，设事件 $E=$｛骰子出现的点数是偶数｝，此时 $A \cap E=\varnothing$ 且 $A \cup E=\Omega$，即事件 A 与事件 E 互逆.

若一组事件 $A_1, A_2, \cdots, A_n$ 两两互不相容，且它们的和为必然事件，则称该事件组为**互不相容完备事件组**，简称**完备事件组**（或称 $A_1, A_2, \cdots, A_n$ 为样本空间的一个**剖分**或**分割**).

为了给上述定义一个直观的几何解释，英国逻辑学家文（Venn）为我们提供了一种工具，他使用图示法来表示事件之间的各种关系，我们又称这种图形为**文图**（见图 1.1）.

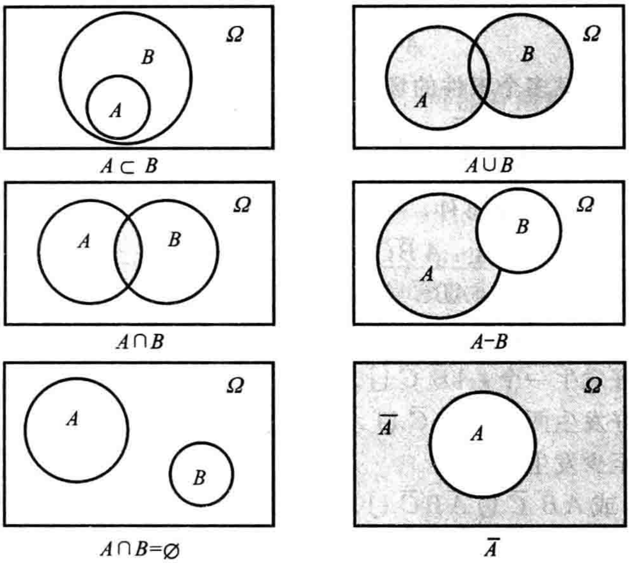


图 1.1　文图

我们把事件的关系与运算的两种解释对照列表如下（见表 1.2）.

表 1.2

符　号	概率论的解释	集合论的解释
$A \subset B$	事件 A 发生必然导致事件 B 发生	A 是 B 的子集
$A=B$	事件 A 与事件 B 相等	集合 A 和集合 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 与事件 B 不能同时发生	A 与 B 没有公共元素
$\overline{A}$	事件 A 的对立事件	A 的补集

（二）事件运算的基本性质

事件运算具有下面的基本性质.

1. 交换律

$$A \cup B = B \cup A,$$
$$A \cap B = B \cap A. \tag{1.1}$$

2. 结合律

$$(A \cup B) \cup C = A \cup (B \cup C),$$
$$(A \cap B) \cap C = A \cap (B \cap C). \tag{1.2}$$

3. 分配律

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

$$(A\cap B)\cup C=(A\cup C)\cap(B\cup C). \tag{1.3}$$

4. 德摩根（De Morgan）原理

$$\overline{A\cup B}=\overline{A}\cap\overline{B},$$
$$\overline{A\cap B}=\overline{A}\cup\overline{B}. \tag{1.4}$$

德摩根原理可以推广到任意多个事件的场合

$$\overline{\bigcup_{i=1}^{n}A_i}=\bigcap_{i=1}^{n}\overline{A_i},\quad \overline{\bigcap_{i=1}^{n}A_i}=\bigcup_{i=1}^{n}\overline{A_i}. \tag{1.5}$$

例 1.8 设 A，B，C 为三个事件，一些事件的表示方法为：

(1) A 发生而 B 与 C 都不发生：$A\overline{B}\overline{C}$ 或 $A-B-C$ 或 $A-(B+C)$；

(2) A 与 B 发生而 C 不发生：$AB\overline{C}$ 或 $AB-C$；

(3) 三个事件都发生：ABC；

(4) 三个事件恰好发生一个：$A\overline{B}\,\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\,\overline{B}C$；

(5) 三个事件恰好发生两个：$AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC$；

(6) 三个事件中至少发生一个：

$A\cup B\cup C$ 或 $A\overline{B}\,\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\,\overline{B}C\cup AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC\cup ABC$；

(7) 三个事件全不发生：

$\overline{A}\,\overline{B}\,\overline{C}$ 或 $\overline{A\cup B\cup C}$ 或 $\Omega-A-B-C$ 或 $\Omega-(A+B+C)$；

(8) 三个事件中至少有一个不发生：

$\overline{A}\cup\overline{B}\cup\overline{C}$ 或 $AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC\cup A\overline{B}\,\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\,\overline{B}C\cup\overline{A}\,\overline{B}\,\overline{C}$.

第二节 随机事件的概率

一、概率的统计定义

对于事件 A，若能用一个数 $P(A)$ 来度量该事件发生的可能性大小，就称这个数 $P(A)$ 为事件 A 发生的概率. 如何来确定 $P(A)$ 这个数呢？我们先来介绍概率的统计定义.

首先给出事件发生的频率的概念.

定义 1.1 若随机事件 A 在 n 次重复独立试验中出现了 m 次，则比值 $\frac{m}{n}$ 称为事件 A 在 n 次试验中出现的**频率**（frequency），记为 $f_n(A)$，即

$$f_n(A)=\frac{m}{n}. \tag{1.6}$$

例 1.9 某地区对一个妇产科医院 6 年中出生的婴儿进行统计，共有新生男婴儿 16 146 人，新生女婴儿 15 248 人，试计算男婴、女婴出生的频率.

解 设 A 表示出生的婴儿为男婴，B 表示出生的婴儿为女婴. 新生男、女婴共有 $16\,146+15\,248=31\,394$ 人，则男婴、女婴出生的频率分别为

$$f_n(A)=\frac{16\,146}{31\,394}\approx 0.514\,3,$$

$$f_n(B)=\frac{15\,248}{31\,394}\approx 0.485\,7.$$

显然，事件的频率具有下列性质：

(1) $0 \leqslant f_n(A) \leqslant 1$；　(1.7)

(2) $f_n(\Omega) = 1$；　(1.8)

(3) $f_n(\varnothing) = 0$.　(1.9)

那么频率与概率之间又有什么关系呢？

以最简单的投硬币试验为例. 投掷一枚均匀硬币，可能有两种结果：出现正面或出现反面. 就一次试验而言，我们根本看不出这些结果的发生有些什么规律，但如果去做大量的试验，就可发现其中的一些规律性. 历史上有些人曾经做过此类试验，结果见表 1.3.

表 1.3

实验者	实验次数	出现正面次数	出现正面的频率
德摩根	2 048	1 039	0.507 3
蒲丰	4 040	2 048	0.506 9
费莱尔	10 000	4 979	0.497 9
皮尔逊	12 000	6 019	0.501 6
皮尔逊	24 000	12 012	0.500 5

从上面的试验中可以看到，当试验次数相当大时，正面出现的频率稳定地在 0.5 左右摆动.

一般地，频率随试验次数的变化而变化，然而，当试验次数足够多时，频率又将稳定地在某个常数附近摆动，此性质称为**频率的稳定性**. 频率的稳定性说明了一个事件发生的可能性大小是事件本身的一种客观属性. 正因为如此，我们便有了下面的随机事件概率的统计定义.

定义 1.2　设在相同条件下，进行大量重复的独立试验，若事件 A 的频率稳定地在某一确定值 p 的附近摆动，则称此数值 p 为事件 A 发生的**概率** (probability)，记作 $P(A)$，即

$$P(A) = p. \tag{1.10}$$

定义 1.2 就是随机事件概率的**统计定义**. 该定义只是对概率的一种描述，它指出了随机事件发生的可能性大小是客观存在的事实，但并不能由此来计算事件发生的概率. 同时定义 1.2 也给出了确定事件概率的近似计算法，即当试验次数充分大时，可用频率作为概率的近似值. 在许多实际问题中，当事件的概率不容易计算时，往往就是用频率近似代替概率，这正是 1946 年由冯·诺伊曼和乌拉姆所建立的蒙特卡罗方法的基本思想.

在例 1.9 中，用频率近似代替概率，那么得到男婴出生的概率为 0.514 3，女婴出生的概率为 0.485 7.

对不同的概率问题，我们有着不同的概率定义，例如，概率的统计定义、概率的几何定义、概率的古典定义等，这些定义都只适用于一类随机现象. 而下面给出的概率的公理化定义，它既概括了这些概率的共性，又避免了各自的局限性，它的问世，很快获得了举世公认，从而使得概率论得到了很大的发展.

下面给出概率的公理化定义.

定义 1.3　设 Ω 是一给定的样本空间，A 为其中的任意一子集，规定一个实数，记作 $P(A)$，若 $P(A)$ 满足下列三条公理：

(1) 非负性：$P(A) \geqslant 0$；　(1.11)

(2) 规范性：$P(\Omega)=1$； (1.12)

(3) 可列可加性：若事件 $A_1,A_2,\cdots,A_n,\cdots$ 两两互斥，即 $A_iA_j=\varnothing(i\neq j)$，则有

$$P(\bigcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}P(A_i). \tag{1.13}$$

则称 $P(A)$ 为事件 A 发生的**概率**.

概率的公理化定义实际上给出了概率的三个基本特性. 无论是概率的统计定义，还是概率的公理化定义，它们都无法精确地计算出某个事件发生的概率. 在实际问题中有许多随机性现象满足一些条件，使得事件的概率能直接得以精确计算，这就是下面要讨论的古典概型问题.

二、概率的古典定义

有这样一类试验，它具有下面两个特性：

(1) 试验中的所有可能结果（即基本事件）只有有限个，而且是两两互斥的；

(2) 每个试验结果出现的可能性相同.

上述两个特性分别称为有限性和等可能性，具有这两个特征的试验称为**古典概型**.

古典概型曾经是概率论发展初期的主要研究对象，它在概率论中有很重要的地位. 一方面，因为它比较简单，许多概念既直观又容易理解，另一方面，它又概括了许多实际问题，有很广泛的应用. 古典概型中随机事件的概率计算，有如下定义.

定义 1.4 设古典概型的所有基本事件为 $e_1,e_2,\cdots,e_n$，事件 A 含有其中的 m 个基本事件，则定义事件 A 的概率为

$$P(A)=\frac{m}{n}, \tag{1.14}$$

其中 n 是基本事件的总数，m 是 A 包含的基本事件数.

定义 1.4 称为概率的**古典定义**. 由（1.14）式规定的概率称为**古典概率**（classical probability），用于古典概型中事件概率的计算.

例 1.10 袋内装有 5 个白球，3 个黑球. 从中任取一球，求取出黑球的概率.

解 设 $A=\{$取出的球是黑球$\}$，记 n_A 为事件 A 包含的基本事件数，n_Ω 为样本空间 Ω 包含的基本事件数，则有 $n_A=3$，$n_\Omega=8$，由（1.14）式得

$$P(A)=\frac{n_A}{n_\Omega}=\frac{3}{8}.$$

例 1.11 设箱中装有 100 件产品，其中有 3 件次品. 为检查产品质量，从中任取 5 件产品，求所取 5 件产品中恰有 1 件次品的概率.

解 设 $A=\{$取出的 5 件产品中恰有 1 件次品$\}$.

在 100 件产品中任取 5 件产品，共有 C_{100}^5 种取法，即 $n_\Omega=C_{100}^5$. A 中包含的样本点个数计算：1 件次品从 3 件次品中取得，共有 C_3^1 种取法；4 件正品从 97 件正品中取得，共有 C_{97}^4 种取法，符合 A 的产品共有 $C_3^1\times C_{97}^4$ 种取法，即 $n_A=C_3^1\times C_{97}^4$. 则有

$$P(A)=\frac{n_A}{n_\Omega}=\frac{C_3^1\times C_{97}^4}{C_{100}^5}\approx 0.138\,1.$$

例 1.12 两封信随机地向标号为Ⅰ、Ⅱ、Ⅲ、Ⅳ的 4 个邮筒投寄，求第二个邮筒恰好被投入 1 封信的概率.

解 设 $A=\{$第二个邮筒恰好被投入 1 封信$\}$.