浙江省“十一五”重点教材建设项目

大学工科数学核心课程系列教材

概率论与数理统计

主　编　曹飞龙

高等教育出版社
HIGHER EDUCATION PRESS

浙江省“十一五”重点教材建设项目

大学工科数学核心课程系列教材

概率论与数理统计

Gailülun yu Shuli Tongji

主 编 曹飞龙

编 者 银俊成 邹海雷 王 成

高等教育出版社·北京

HIGHER EDUCATION PRESS BEIJING

内容提要

本书是浙江省“十一五”重点教材建设项目成果、“大学工科数学核心课程系列教材”之一，是编者在多年来从事概率论与数理统计教学实践工作的基础上，参照全国工学、经济学硕士研究生入学统一考试数学（一）和数学（三）考试大纲中对概率统计部分的基本要求，认真研究国内外同类教材，取长补短编写而成。

本书系统介绍了概率论与数理统计的基本内容，包括事件与概率、随机变量及其分布、随机变量的数字特征、大数定律与中心极限定理、数理统计的基本概念、参数估计、假设检验、回归分析、常用统计软件简介。本书侧重于基本概念、基本理论和方法，叙述翔实严谨，语言通俗易懂；内容紧扣研究生入学考试大纲，配有适合各层次要求的丰富习题；紧密结合统计软件，介绍了 MATLAB 的概率统计工具箱的功能及应用，并比较了 R、SPSS、SAS、S - Plus、Stata 以及 Eviews 等常见的几款统计软件的特点。

本书可作为高等学校理工科非数学各专业本科生教材或教学参考书，也可供工程技术人员和自学者参考。

图书在版编目(CIP)数据

概率论与数理统计／曹飞龙主编. --北京:高等教育出版社,2012.8

ISBN 978 -7 -04 -035855 -1

Ⅰ.①概… Ⅱ.①曹… Ⅲ.①概率论-高等学校-教材②数理统计-高等学校-教材 Ⅳ.①O21

中国版本图书馆 CIP 数据核字(2012)第 166970 号

策划编辑 杨 波　　责任编辑 杨 波　　封面设计 赵 阳　　版式设计 于 婕
插图绘制 尹文军　　责任校对 刁丽丽　　责任印制 尤 静

出版发行	高等教育出版社	咨询电话	400 - 810 - 0598
社　　址	北京市西城区德外大街 4 号	网　　址	http://www.hep.edu.cn
邮政编码	100120		http://www.hep.com.cn
印　　刷	北京宏信印刷厂	网上订购	http://www.landraco.com
开　　本	787mm×960mm　1/16		http://www.landraco.com.cn
印　　张	16.75	版　　次	2012 年 8 月第 1 版
字　　数	290 千字	印　　次	2012 年 8 月第 1 次印刷
购书热线	010 - 58581118	定　　价	24.80 元

物 料 号　35855 - 00

学数学需掌握理论与方法, 更要培养数学素养

——“大学工科数学核心课程系列教材”序言

我国于 2010 年颁布了《国家中长期教育改革和发展规划纲要 (2010—2020 年)》, 教育部 2011 年开始实施了《高等学校本科教学质量与教学改革工程》,2012 年又出台了“教育部关于全面提高高等教育质量的若干意见”。这为我国实现高等教育强国目标提出了战略部署, 为高校培养具有实践能力、创新精神、国际视野的高素质高级专门人才提出了更高要求。我国高等教育进入大众化教育阶段的新形势、新任务, 让我们深刻地感受到新一轮教育教学改革的必要性和紧迫性, 也为大学工科数学教育教学改革明确了方向。

根据教育部高等学校数学与统计学教学指导委员会制定的“大学数学基础课程教学基本要求”, 结合本科高校的特点, 以培养高素质应用型人才、复合型人才和卓越工程师为目标, 浙江省高等学校数学类专业与数学基础课程教学指导委员会组织编写了大学工科数学核心课程系列教材, 含《高等数学》(上、下册)、《线性代数》、《概率论与数理统计》、《数学物理方程》与《数学软件与大学数学实验》。本系列教材融合了编者的长期教学经验和教学改革成果, 着眼于培养学生学习能力、实践能力和创新能力, 让学生不仅学会数学知识、理论与方法, 还使其数学思想和数学思维得到培养。该系列教材的特点可概括如下:

(1) 问题驱动, 融合背景。在内容的编排上既考虑如何与中学衔接, 又考虑到内在的逻辑关系, 并注意到专业后续课程的需要。教材注重讲解“如何应用基本理论和方法分析与解决实际问题”的思想方法, 适度增加工程技术应用领域中的案例、具有趣味性和活学活用的数学建模内容。融入数学建模方法, 增加 MATLAB 数学软件介绍, 培养学生运用其理论和方法解决实际问题的能力。内容力求简明, 削减一些公式演算内容, 降低对一些定理、公式的证明和一些数学表达形式的要求, 使学生在正确理解的基础上, 能够熟练地应用基本理论与方法。

(2) 剖析思想, 深入浅出。突出思想性, 保持严谨性, 着力揭示数学思想与概念的本质和解决问题的思想方法, 充分体现数学理论的“源”与“流”的关

系。为了使学生的数学思维得到必要的训练,教材的叙述与推理保持应有的严谨性,对于某些重要结论仍保留必要的分析、证明与讨论。讲解力求深入浅出,富有启发性。教材编写尽力从学生熟悉的实例和知识出发,用其熟悉的语言、知识、思想方法或直观的几何形象,凭借联想、类比等思维方法,进行自然的、合乎规律的扩展和深化;尽量以问题为导向,采用“提出问题、讨论问题、解决问题”的方式来展开,以适应学生的思维习惯;语言叙述力求直观清晰、通俗易懂。

(3) 体例新颖,强化训练。以加强学生数学素养的养成、综合应用能力的增强为目标,设计新体例。新体例将充分体现问题引入、数学模型、理论与方法、综合训练、拓展阅读、研究性学习等特点。注重基本运算能力的培养,特别注重数学思想方法的培养,以适量的应用实例为学生提供应用能力训练的素材;为培养学生的综合能力,精心设计了各种不同类型的习题,包括基本理论题、有一定难度的综合题和研究性应用题。书中附录编入 MATLAB 上机演练与实验题目。

(4) 资源丰富,实现共享。系列教材具有丰富的课程数字化资源,教学设计能较好反映教学中重点、难点的巧妙处理。利用现代化网络技术,把这些精彩的数字教学资源 (含讲稿、教案、课件、案例、专题论文和研究性课题) 集成,建成一个数字化的课程教学资源网站,构建一个内容丰富的、师生共享的、“活的”课程教学环境。

该系列教材力求内容简明,体系科学合理,揭示数学思想,融入数学建模,展示思维方法,引入现代数学软件,强化应用能力和创新能力的培养,深入浅出,富于启发性,促进学生学习。本系列教材适合于高等院校工科专业的本科生,也可供理科专业和其他相关专业大学生、研究生学习。

浙江省高等学校数学类专业
与数学基础课程教学指导委员会
主任委员 裘松良
2012 年 5 月 15 日

前言

概率论与数理统计是高等学校理工及管理科学类专业大学生的一门必修的基础课，它在自然科学及社会科学的各个方面有着极其广泛的应用。本书是编者在多年从事概率统计教学实践工作的基础上，参照全国工学、经济学硕士研究生入学统一考试数学 (一) 和数学 (三) 对概率统计部分的基本要求而编写的。本教材的特点是：

(1) 通俗易懂：尽量使用较少的数学知识 (只限于微积分和少量矩阵代数知识)，避免过于数学化的论证，注重思想方法的剖析，保持叙述的严谨性。

(2) 结合考研：本书内容紧扣全国硕士研究生入学考试大纲，采用考试大纲规范的符号和术语。每章配有丰富的习题，分为 A,B 两部分。A 部分是对基本知识和方法的巩固与练习；B 部分是基本内容的提升，其中包含历年的一些考研试题，可供学有余力及立志考研的读者选用。

(3) 紧密结合统计软件：注重数学方法和计算机应用相结合，充分发挥数学软件的作用，提高教学效率，从而有效培养学生应用数学方法解决实际问题的能力。

全书分三部分共 9 章。1 至 4 章为第一部分，讲授概率论的知识，其中包括随机事件、随机变量、数字特征和极限定理等内容，由王成 (wangc628@cjlu.edu.cn) 编写；5 至 8 章为第二部分，讲授数理统计的内容，其中包括基本概念、参数估计、假设检验和线性回归分析等内容，由邹海雷 (zouhailei@163.com) 编写；第 9 章为第三部分，是关于工程领域常用的数学软件 MATLAB 的简单介绍，尤其是对 MATLAB 与 Microsoft Word 紧密结合的适于科技论文写作的 Notebook 以及功能齐全的统计工具箱作了介绍。另外还简单介绍并比较了 R、SPSS、SAS、S-Plus、Stata 以及 Eviews 等常见的几款统计软件，由银俊成 (yjccjlu@gmail.com) 编写。全书由曹飞龙 (feilongcao@gmail.com) 策划，银俊成统稿。全书讲授大概需 64 学时，根据不同学时和不同专业的要求，讲授内容可酌情取舍。

本书可作为高等院校本科非数学专业的理工类概率统计的教材，亦可供有

关技术人员和管理工作者参考。由于编者水平有限，加之时间仓促，书中不当之处在所难免，敬请同行与广大读者不吝赐教。

编者

2012.05.01

目　录

第一章　事件与概率 ········· 1

1.1　样本空间与事件 ········· 1

1.1.1　样本空间与事件 ········· 1

1.1.2　事件的关系及运算 ········· 3

1.2　概率的定义及性质 ········· 5

1.2.1　频率　概率的统计定义及公理化定义 ········· 5

1.2.2　概率的性质 ········· 7

1.3　古典概率模型 ········· 9

1.3.1　古典概型 ········· 9

1.3.2　几何概型简介 ········· 11

1.4　条件概率与全概率公式 ········· 12

1.4.1　条件概率　乘法法则 ········· 12

1.4.2　全概率公式　贝叶斯公式 ········· 14

1.5　独立性 ········· 16

1.5.1　事件的独立性 ········· 16

1.5.2　独立试验序列 ········· 19

习题一 ········· 21

第二章　随机变量及其分布 ········· 24

2.1　随机变量与分布函数 ········· 24

2.1.1　随机变量的概念 ········· 24

2.1.2　随机变量的分布函数 ········· 25

2.2　离散型随机变量 ········· 27

2.2.1　离散型随机变量及其概率分布 ········· 27

2.2.2　常见的离散型随机变量 ········· 28

2.3　连续型随机变量 ········· 31

2.3.1　连续型随机变量及其密度函数 ········· 31

2.3.2 常见的连续型随机变量 …… 33
2.4 随机变量函数的分布 …… 39
2.5 二维随机变量及其分布 …… 41
2.5.1 二维随机变量的联合分布函数 …… 42
2.5.2 二维离散型随机变量 …… 43
2.5.3 二维连续型随机变量 …… 45
2.5.4 独立性 …… 51
2.5.5 二维随机变量函数的分布 …… 53
2.5.6 条件分布简介 …… 58
习题二 …… 60
第三章 随机变量的数字特征 …… 65
3.1 数学期望 …… 65
3.1.1 数学期望的概念 …… 65
3.1.2 随机变量函数的数学期望 …… 69
3.1.3 期望的性质 …… 72
3.2 方差 …… 74
3.2.1 方差的概念 …… 74
3.2.2 方差的性质 …… 77
3.3 协方差与相关系数 …… 78
3.4 矩、协方差矩阵及 n 维正态分布 …… 81
3.4.1 矩与协方差矩阵 …… 81
3.4.2 n 维正态分布 …… 82
习题三 …… 84
第四章 大数定律与中心极限定理 …… 87
4.1 大数定律 …… 87
4.2 中心极限定理 …… 90
习题四 …… 92
第五章 数理统计的基本概念 …… 94
5.1 总体、样本 …… 95
5.2 统计量与抽样分布 …… 96
5.2.1 统计量 …… 96
5.2.2 常用统计分布 …… 97

5.2.3 正态总体的抽样分布 …… 102
习题五 …… 106
第六章　参数估计 …… 108
6.1 点估计 …… 108
6.1.1 矩估计法 …… 108
6.1.2 极大似然估计法 …… 110
6.2 估计量的评选标准 …… 117
6.2.1 无偏性 …… 117
6.2.2 有效性 …… 118
6.2.3 一致性 …… 119
6.3 区间估计 …… 119
6.3.1 单个正态总体均值的区间估计 …… 120
6.3.2 单个正态总体方差的区间估计 …… 123
6.3.3 两个正态总体均值差 $\mu_1-\mu_2$ 的区间估计 …… 125
6.3.4 两个正态总体方差比 σ_1^2/σ_2^2 的区间估计 …… 126
6.3.5 非正态总体参数的区间估计 …… 127
6.3.6 单侧置信区间 …… 129
习题六 …… 131
第七章　假设检验 …… 135
7.1 假设检验的基本概念 …… 135
7.1.1 问题的提出 …… 135
7.1.2 假设检验的基本思想 …… 136
7.1.3 两类错误 …… 136
7.2 单个正态总体参数的假设检验 …… 138
7.2.1 正态总体均值的假设检验 …… 139
7.2.2 正态总体方差的假设检验 …… 142
7.3 两个正态总体参数的假设检验 …… 144
7.3.1 两个正态总体均值的假设检验 …… 145
7.3.2 两个正态总体方差的假设检验 …… 148
7.4 非正态总体参数的假设检验 …… 149
7.5 总体分布的 χ^2 检验法 …… 152
习题七 …… 156

第八章 回归分析 160
8.1 一元线性回归 160
8.1.1 一元线性回归的概念 160
8.1.2 a, b 的最小二乘估计 161
8.1.3 显著性检验 163
8.1.4 预测与控制 167
8.2 多元线性回归 170
习题八 173
第九章 常用统计软件简介 175
9.1 MATLAB 及其应用简介 175
9.1.1 MATLAB 操作入门 175
9.1.2 Notebook 初步 199
9.1.3 常见概率分布的函数 202
9.1.4 参数估计 204
9.1.5 假设检验 206
9.2 R 及其应用简介 211
9.3 SPSS 及其他常用统计软件简介 215
附录 A 常用表 218
附录 B 部分习题解答 244
附录 C 2009 年至 2012 年全国硕士研究生入学统一考试试题 250
参考文献 256

第一章　事件与概率

本章的主要内容有随机事件及其运算、频率、概率的公理化定义及性质、条件概率、全概率公式和贝叶斯 (Bayes) 公式、事件的独立性等. 这些概念是概率理论的基础.

1.1　样本空间与事件

1.1.1　样本空间与事件

在自然界、人类生产生活及科学工作中主要有两类现象.

一类是确定性现象, 这类现象的结果是唯一的, 事先可以预知. 比如, 向上扔一颗石子必然会落地, 同性电荷一定互斥, 太阳总是东升西落等.

另一类是随机现象, 其结果有多种可能并且事先无法准确预测将会发生什么结果. 比如, 掷一颗骰子将会出现几点; 从一批产品中任取 10 件检查, 会有几件是次品; 杭钢集团股票在未来某时刻的价格为多少; 微软网站一天内会被访问多少次等.

概率统计研究的主要目标是随机现象, 对随机现象进行一次观察称为随机试验, 也简称**试验**. 需要注意的是, 虽然我们在试验前无法预测将会发生什么结果, 但是我们可以**明确所有可能的结果的集合**.

试验中每一个可能发生的结果称为一个**样本点**, 常用小写希腊字母 ω 表示. 所有样本点组成的集合称为**样本空间**, 常用大写希腊字母 Ω 表示. 看如下几个例子.

例 1　掷一颗均匀六面体的骰子, 观察出现的点数.

样本空间可写为

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\},$$

其中 ω_i 表示“出现 i 点”, $i = 1, 2, \cdots, 6$. 当然, 也可简单写成

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

例 2 三名同学 (编号 1, 2, 3) 参加百米赛跑, 观察比赛名次 (假设没有名次并列的情况).

设 (i_1, i_2, i_3) 是 (1, 2, 3) 的一个排列, 表示获得第一、第二、第三的同学编号依次为 i_1, i_2 和 i_3. 则样本空间可写为

$$\Omega = \left\{(i_1, i_2, i_3) \middle| (i_1, i_2, i_3)\text{是}(1,2,3)\text{的一个排列}\right\}.$$

例 3 从编号为 1, 2, 3, 4, 5 的五只小球中随机抽取两只, 观察它们的号码.

样本空间可表示为

$$\Omega = \{\omega_{12}, \omega_{13}, \omega_{14}, \omega_{15}, \omega_{23}, \omega_{24}, \omega_{25}, \omega_{34}, \omega_{35}, \omega_{45}\},$$

其中 ω_{ij} 表示取到的两只球的编号为 i 和 j.

例 4 观察某网站一天内被访问的次数.

从理论上讲, 每个自然数都是一个可能的结果, 故样本空间为

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

例 5 观测一部电子产品的使用寿命.

若用 ω_x 表示"测得的使用寿命为 x 小时" $(x \geqslant 0)$, 则样本空间为

$$\Omega = \{\omega_x | x \geqslant 0\}.$$

样本空间 Ω 的任何子集都称为**随机事件**, 简称**事件**. 事件常用大写英文字母 A, B, C 等表示. 若事件 A 中的某个样本点在试验中出现, 我们就说事件 A 发生了, 即**事件 A 发生当且仅当 A 中某个样本点发生**.

比如在前面的例 1 中, 设事件 $A = \{\omega_1, \omega_3, \omega_5\}$, 那么出现"1 点"、"3 点"、"5 点"中的任一个, 事件 A 即发生. 简单地说, 事件 A 即表示"出现奇数点".

再如例 2 中, 设事件 $A = \{(1,2,3), (1,3,2)\}$, 那么若出现样本点 (1, 2, 3) 或 (1, 3, 2), 事件 A 即发生. 显然, 事件 A 即表示"1 号同学获得第一名".

再看例 4, 若要表示"该网站一天的访问次数超过 10 次"这个现象, 可用事件 $A = \{11, 12, \cdots\}$ 来表示; 类似地, $B = \{2, \cdots, 10\}$ 即表示"该网站一天的访问次数至少为 2 次并且不超过 10 次".

注意两个特殊事件 (即 Ω 的两个特殊子集), 一个是空集 $\varnothing$, 因为它不含任何元素, 因此它一定不会发生, 称为**不可能事件**; 另一个是 Ω 本身, 因为它包含了所有的样本点 (即所有可能的结果), 因此它一定会发生, 称为**必然事件**.

由以上讨论, 再次强调如下两点:

(1) 任何事件都是样本空间的子集. 事件可以用集合形式来表示, 也可以用明确的语言来描述;

(2) 事件 A 发生当且仅当 A 中某个样本点发生.

1.1.2 事件的关系及运算

在很多问题中, 经常要用简单事件来表示一些复杂事件, 进而研究其概率. 因此需要讨论事件的各种关系及运算. 我们知道每个事件都是样本空间 Ω 的一个子集, 因此事件的关系和运算完全可以用集合的关系及运算来表示. 下面都是在同一样本空间 Ω 中讨论事件的关系和运算.

一、包含关系

若事件 A 是事件 B 的子集, 则称 A 包含于 B, 或称 B 包含 A, 记为 $A\subset B$, 或 $B\supset A$. 通俗的说, 即是事件 A 发生必然导致事件 B 发生.

譬如在前面例 1 中, 设 $A=\{2\}=$ "出现 2 点", $B=\{2,4,6\}=$ "出现偶数点", 则 $A\subset B$, A 的发生必然导致 B 发生.

若事件 A 与 B 所包含的样本点一样 (即 $A\subset B$ 且 $B\subset A$), 则称 A 与 B 相等, 记为 $A=B$.

二、事件的并

设 A 与 B 是两个事件, 由 A 与 B 中的所有样本点组成的集合称为 A 与 B 的并 (也称 A 与 B 的和), 记为 $A\cup B$(或 $A+B$).

$A\cup B$ 也是一个事件, 易知 $A\cup B$ 发生当且仅当 A 发生或 B 发生.

比如在前面例 1 中, 设 $A=$"出现奇数点" $=\{1,3,5\}$, $B=$"出现的点数不超过 3" $=\{1,2,3\}$, 则 $A\cup B=\{1,2,3,5\}$.

三、事件的交

设 A 与 B 是两个事件, 由 A 与 B 中的公共样本点组成的集合称为 A 与 B 的交 (也称 A 与 B 的乘积), 记为 $A\cap B$ (或 AB).

$A\cap B$ 也是一个事件, 易知 $A\cap B$ 发生当且仅当 A 与 B 同时发生.

比如在前面例 1 中, 设 $A=$ "出现奇数点" $=\{1,3,5\}$, $B=$ "出现的点数不超过 3" $=\{1,2,3\}$, 则 $AB=\{1,3\}$.

事件的并和交可以推广到有限多或可列多的事件, $\bigcup\limits_{i=1}^{n}A_i$(或 $\sum\limits_{i=1}^{n}A_i$) 为 $A_1,A_2,\cdots,A_n$ 的有限并, $\bigcup\limits_{i=1}^{\infty}A_i$(或 $\sum\limits_{i=1}^{\infty}A_i$) 为 $A_1,A_2,\cdots$ 的可列并; $\bigcap\limits_{i=1}^{n}A_i$(或

$\prod_{i=1}^{n} A_i$) 为有限交, $\bigcap_{i=1}^{\infty} A_i$(或 $\prod_{i=1}^{\infty} A_i$) 为可列交.

四、事件互斥

设 A 与 B 是两个事件, 若 $A \cap B = \varnothing$, 即 A 与 B 不会同时发生, 则称 A 与 B 互斥 (也称互不相容).

比如在前面例 2 中, 设

$$A = \text{“1 号同学得第一”} = \{(1,2,3),(1,3,2)\},$$

$$B = \text{“2 号同学得第一”} = \{(2,1,3),(2,3,1)\},$$

显然 $AB = \varnothing$, 即 A 与 B 不会同时发生.

五、事件的差

设 A 与 B 是两个事件, 由在 A 中但不在 B 中的样本点组成的集合称为 A 对事件 B 的差, 记为 $A - B$.

$A - B$ 也是一个事件, 易知 $A - B$ 发生当且仅当 A 发生但是 B 不发生.

比如在前面例 1 中, 设 A= “出现奇数点” $=\{1,3,5\}$, B= “出现的点数不超过 3” $=\{1,2,3\}$, 则 $A - B = \{5\}$.

六、对立事件

事件 A 的对立事件, 写作 $\overline{A}$, $\overline{A}$ 表示由所有在样本空间 Ω 中但不在 A 中的样本点组成的集合 (即 $\overline{A} = \Omega - A$). 易知事件 $\overline{A}$ 发生当且仅当 A 不发生.

显然对立事件是相互的, 即 A 的对立事件为 $\overline{A}$, 而 $\overline{A}$ 的对立事件是 A. $\varnothing$ 与 Ω 互为对立事件.

比如在前面例 1 中, 事件 A= “出现奇数点” $=\{1,3,5\}$ 的对立事件为 $\overline{A} = \{2,4,6\}$, B= “出现的点数不超过 3” $=\{1,2,3\}$ 的对立事件为 $\overline{B} = \{4,5,6\}$.

七、完备事件组

设 $A_1, A_2, \cdots, A_n$ 是一组事件, 若它们两两互斥 (即 $A_iA_j = \varnothing$, $i \neq j$) 且 $\bigcup_{i=1}^{n} A_i = \Omega$, 则称 $A_1, A_2, \cdots, A_n$ 构成一个互不相容的完备事件组, 简称完备事件组.

用集合的维恩 (Venn) 图 (图 1.1), 可以很直观的表示上述集合的关系和运算.

事件的运算 (即集合的运算) 有下面一些性质. 对任意事件 A, B, C, 有

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

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

(3) 分配律 $A(B\cup C)=AB\cup AC,\qquad A\cup(BC)=(A\cup B)(A\cup C)$;

(4) 德摩根 (De Morgan) 定律

$$\overline{A\cup B}=\overline{A}\,\overline{B},$$
$$\overline{AB}=\overline{A}\cup\overline{B}.$$

例 6 某个人加工了三个零件, 设 A_i 表示事件“加工的第 i 个零件是合格品” $(i=1,2,3)$, 则可用 A_1,A_2,A_3 表示下面一些较复杂的事件.

(1) “只有第一个零件是合格品” (记为事件 B) 可表示为

$$B=A_1\overline{A_2}\,\overline{A_3}.$$

(2) “至少有一个是合格品” (记为事件 C) 可表示为

$$C=A_1\cup A_2\cup A_3.$$

(3) “只有一个零件是合格品” (记为事件 D) 可表示为

$$D=A_1\overline{A_2}\,\overline{A_3}\cup\overline{A_1}A_2\overline{A_3}\cup\overline{A_1}\,\overline{A_2}A_3.$$

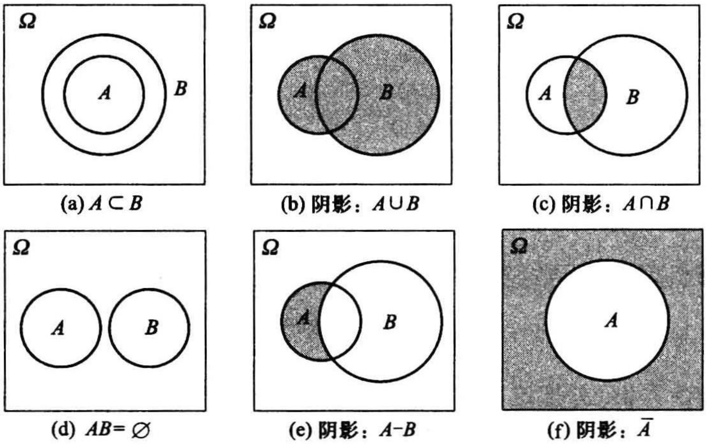

图 1.1 集合的关系与运算

1.2 概率的定义及性质

1.2.1 频率 概率的统计定义及公理化定义

前面我们讨论了事件的概念. 对于随机现象, 人们自然关心事件发生的可能性大小. 我们知道, 除了不可能事件和必然事件以外, 一般的事件在一次试验

中可能发生也可能不发生, 但我们总认为事件发生的可能性大小是客观存在的. 比如掷一枚均匀硬币, 凭直观自然认为出现正面的可能性为 $\frac{1}{2}$. 那么是如何得出这个数的, 最简单直观的想法就是用频率来解释.

定义 1.2.1 在一定条件下将试验进行 n 次, 若事件 A 发生 n_A 次, 则称 A 发生的频率为 $\frac{n_A}{n}$, 记为

$$f_n(A)=\frac{n_A}{n}.$$

像前面掷一枚均匀硬币, 我们可以把硬币抛掷 100 次、 1000 次、 10000 次等, 记录下事件“出现正面”发生的频率. 人们发现在抛掷次数 n 很大时, 事件“出现正面”发生的频率总是很接近 $\frac{1}{2}$, 因此我们说“出现正面”的可能性为 $\frac{1}{2}$. 表 1.1 是历史上一些概率统计学家掷硬币的记录情况.

表 1.1 历史上的掷硬币记录

实验者	抛硬币次数	正面出现次数	正面出现频率
德摩根	2048	1061	0.5181
蒲丰 (Buffon)	4040	2048	0.5069
费勒 (Feller)	10000	4979	0.4979
皮尔逊 (Pearson)	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

对于随机现象, 无数经验告诉我们, 当试验进行的次数 n 增大时, 事件 A 的频率 $\frac{n_A}{n}$ 总逐渐趋向于一个确定的常数①, 我们就用这个常数来反映事件 A 发生的可能性大小, 并把这个常数称为事件 A 发生的概率, 记为 $P(A)$. 概率的这个通俗的定义即为**概率的统计定义**.

概率的统计定义非常简单直观, 有助于人们理解和接受概率的概念. 在很多时候, 人们往往用 n 较大时某事件 A 的频率 $\frac{n_A}{n}$ 作为 A 发生的概率 $P(A)$ 的近似, 这样处理比较方便. 但是概率的统计定义也有缺陷, 不能作为概率论严格的理论基础. 首先, 对于一些无法或很难大量重复的试验, 概率的统计定义就不适用了; 另外, 对于定义中“确定的常数”是多少, 不够明确.

① 关于 $\frac{n_A}{n}$ 趋向于一个确定的常数, 第四章中的大数定律将给出更严格的解释.