

21世纪大学数学创新教材

丛书主编 陈化

概率论与数理统计

(第三版)

刘安平 肖海军 主编
奚 先 田木生



科学出版社

• 21 世纪大学数学创新教材 •
丛书主编 陈 化

概率论与数理统计

(第三版)

刘安平 肖海军 主编
奚 先 田木生



科学出版社
北 京

版权所有,侵权必究

举报电话:010-64030229;010-64034315;13501151303

内 容 简 介

本书内容包括随机事件及其概率、随机变量及其分布、多维随机变量及其分布、随机变量的数字特征、大数定律与中心极限定理、样本与抽样分布、参数估计、假设检验、方差分析、相关与回归分析等。本书知识体系结构完整,例题、习题丰富。其中第1章至第5章为基础部分,可供较少教学学时数使用,第6章与第7章(用*表示)则可供较多学时数使用,第8章至第10章(用**表示)就可供多学时数使用。

本书可作为高等学校本科生(包括理工类与经管类)概率论与数理统计课程的教材或参考用书,也可作为广大概率论与数理统计以及应用统计人员的工具性参考书。

图书在版编目(CIP)数据

概率论与数理统计/刘安平主编. —3版. —北京:科学出版社,2014.1

21世纪大学数学创新教材

ISBN 978-7-03-039431-6

I. ①概… II. ①刘… III. ①概率论—高等学校—教材②数理统计—高等学校—教材 IV. ①021

中国版本图书馆 CIP 数据核字(2013)第 310969 号

责任编辑:王雨舸 / 责任校对:蔡 莹 黄彩霞

责任印制:高 嵘 / 封面设计:蓝 正

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>

武汉市首壹印务有限公司印刷

科学出版社发行 各地新华书店经销

*

2003年8月第 一 版 开本:787×1000 1/16

2009年8月第 二 版 印张:15 1/4

2014年1月第 三 版 字数:331 000

2014年1月第一次印刷 定价:28.00元

(如有印装质量问题,我社负责调换)

“21 世纪大学数学创新教材” 丛书编委会

主 编 陈 化
常务副主编 樊启斌
副 主 编 吴传生 何 穗 刘安平
编 委 (按姓氏笔画为序)
王卫华 王展青 刘安平 严国政 李 星
杨瑞琰 肖海军 吴传生 何 穗 陈 化
罗文强 赵东方 黄樟灿 梅全雄 彭 放
彭斯俊 曾祥金 谢民育 樊启斌

“21 世纪大学数学创新教材” 丛书序

《21 世纪大学数学创新教材》为大学本科数学系列教材,大致划分为公共数学类、专业数学类两大块,创新是其主要特色和要求.经组编委员会审定,列选科学出版社普通高等教育“十一五”规划教材.

一、组编机构

《21 世纪大学数学创新教材》丛书由多所 985 和 211 大学联合组编:

丛 书 主 编 陈 化

常务副主编 樊启斌

副 主 编 吴传生 何 穗 刘安平

丛 书 编 委 (按姓氏笔画为序)

王卫华	王展青	刘安平	严国政	李 星
杨瑞琰	肖海军	吴传生	何 穗	陈 化
罗文强	赵东方	黄樟灿	梅全雄	彭 放
彭斯俊	曾祥金	谢民育	樊启斌	

二、教材特色

创新是本套教材的主要特色和要求,创造双重品牌:

先进.把握教改、课改动态和学科发展前沿,学科、课程的先进理念、知识和方法原则上都要写进教材或体现在教材结构及内容中.

知识与方法创新.重点教材、高层次教材,应体现知识、方法、结构、内容等方面的创新,有所建树,有所创造,有所贡献.

教学实践创新.教材适用,教师好教,学生好学,是教材的基本标准.应紧跟和引领教学实践,在教学方法、教材结构、知识组织、详略把握、内容安排上有独到之处.

继承与创新.创新须与继承相结合,是继承基础上的创新;创新需转变为参编者、授课者的思想和行为,避免文化冲突.

三、指导思想

遵循国家教育部高等学校数学与统计学教学指导委员会关于课程教学的基本要求,力求



教材体系完整,结构严谨,层次分明,深入浅出,循序渐进,阐述精炼,富有启发性,让学生打下坚实的理论基础.除上述一般性要求外,还应具备下列特点:

(1) 恰当融入现代数学的新思想、新观点、新结果,使学生有较新的学术视野.

(2) 体现现代数学创新思维,着力培养学生运用现代数学软件的能力,使教材真正成为基于现代数学软件的、将数学软件融合到具体教学内容中的现代精品教材.

(3) 在内容取舍、材料组织、叙述方式等方面具有较高水准和自身特色.

(4) 数学专业教材要求同步给出重要概念的英文词汇,章末列出中文小结,布置若干道(少量)英文习题,并要求学生用英文解答.章末列出习题和思考题,并列出可进一步深入阅读的文献.书末给出中英文对照名词索引.

(5) 公共数学教材具有概括性和简易性,注重强化学生的实验训练和实际动手能力,加强内容的实用性,注重案例分析,提高学生应用数学知识和数学方法解决实际问题的能力.

四、主编职责

丛书组编委员会和出版社确定全套丛书的编写原则、指导思想和编写规范,在这一框架下,每本教材的主编对本书具有明确的责权利:

1. 拟定指导思想

按照丛书的指导思想和特色要求,拟出编写本书的指导思想和编写说明.

2. 明确创新点

教改、课改动态,学科发展前沿,先进理念、知识和方法,如何引入教材;知识和内容创新闪光点及其编写方法;教学实践创新的具体操作;创新与继承的关系把握及其主客体融合.

3. 把握教材质量

质量是图书的生命,保持和发扬科学出版社“三高”、“三严”的传统特色,创出品牌;适用性是教材的生命力所在,应明确读者对象,篇幅要结合大部分学校对课程学时数的要求.

4. 掌握教材编写环节

(1) 把握教材编写人员水平,原则上要求博士、副教授以上,有多年课程教学经验,熟悉课程和学科领域的发展状况,有教材编写经验,有扎实的文字功底.

(2) 充分注意著作权问题,不侵犯他人著作权.

(3) 讨论、拟定教材提纲,并负责编写组的编写分工、协调与组织.

(4) 拟就内容简介、前言、目录、样章,统稿、定稿,确定交稿时间.

(5) 负责出版事宜,敦促编写组成员使用本教材,并优先选用本系列教材.

《21 世纪大学数学创新教材》组编委员会

2009 年 6 月

前 言

随着科学技术的飞速发展,各种数学方法的应用更加广泛。在许多领域,概率论与数理统计理论已经成为必须掌握的基础知识。

概率论与数理统计课程在培养学生运用概率论与数理统计理论独特的思维方式分析和解决实际问题的能力方面具有重要的作用,并为后续课程以及未来的工作实践提供必要的随机数学基础。该课程是高等学校本科各相关专业(非数学类专业)的一门重要大学数学基础课程,并被列入工学、经济学等硕士研究生入学考试的必考科目。当然,不同的专业对该课程的要求也不尽相同。我们编写本书,就是为了适应教学改革的需要,希望为本课程提供一本合适的教学用书。

本书分为两个部分。概率论部分(第1章至第5章)作为基础知识,为读者提供必要的理论基础。数理统计部分(第6章至第10章)主要讲述了参数估计与假设检验,并介绍了方差分析和回归分析。本书内容包括该课程教学基本要求规定的全部内容,并可满足硕士研究生入学考试中“数学一”考试的基本要求。

本书在第二版出版之后,又经过4年的实践教学,并经过了教学计划的再次修订,第三版就是在这一基础上完成的。这次修订的在章节顺序与原来保持大体一致外,对全书的具体内容、特别是例题与习题作了修改。在选材和叙述上尽量做到联系各专业的实际,注重应用,力求理论描述清晰易懂,并努力做到既便于学生自学,又便于教师教学。

全书共10章,第1、7章由刘安平修订,第2、8章由肖海军修订,第3、9、10章由奚先修订,第4、5、6章由田木生修订。在修订工作中,研究生杨季琛、刘光杰、王苗、隋美玲、李琴、王晓牡、周强、李凯鸿等同学重新修订了习题与例题。全书由刘安平策划,肖海军统稿。

本书适合高等学校本科各相关专业(非数学类专业)作为教材或教学参考书。第1章至第5章作为基础知识教学学时数约为40学时,第1章至第7章教学学时数约为48学时,全部教学内容的教学学时数约为56学时。本书既可供有关教师,也可供工程技术人员等参考。

本书的出版得到了科学出版社、中国地质大学数学与物理学院的帮助和支持。数学教研室的许多教师还提出了大量宝贵的修订意见,在此一并表示衷心的感谢。

由于编者水平有限,书中肯定存在不妥和错误之处,恳请读者批评指正。

编 者

2013年10月于南望山

目 录 | CONTENTS

第 1 章 随机事件及其概率	001
1.1 样本空间与随机事件	001
1.2 事件的频率与概率	006
1.3 古典概型与几何概型	009
1.4 条件概率	014
1.5 全概率公式和贝叶斯公式	017
1.6 事件的独立性	021
习题 1	028
第 2 章 随机变量及其分布	032
2.1 随机变量与分布函数	032
2.2 离散型随机变量及其分布函数	035
2.3 连续型随机变量及其分布函数	040
2.4 随机变量函数的分布	047
习题 2	052
第 3 章 多维随机变量及其分布	056
3.1 二维随机变量的概率分布	056
3.2 边缘分布	061
3.3 条件分布	063
3.4 随机变量的独立性	064



3.5 两个随机变量函数的分布	068
习题 3	075
第 4 章 随机变量的数字特征	078
4.1 数学期望	078
4.2 方差	090
4.3 协方差和相关系数	097
4.4 原点矩与中心矩	102
习题 4	104
第 5 章 大数定律与中心极限定理	108
5.1 大数定律	108
5.2 中心极限定理	112
习题 5	116
* 第 6 章 样本与抽样分布	117
6.1 基本概念	117
6.2 样本数字特征	120
6.3 正态总体的抽样分布	123
习题 6	130
* 第 7 章 参数估计	132
7.1 参数的点估计	133
7.2 估计量的评价标准	138
7.3 参数的区间估计	142
习题 7	155
** 第 8 章 假设检验	159
8.1 假设检验的基本概念	159
8.2 单个正态总体参数的假设检验	161

8.3 两个正态总体参数的假设检验	164
8.4 分布函数的假设检验	167
习题 8	170
** 第 9 章 方差分析	172
9.1 方差分析概述	172
9.2 单因素方差分析	174
9.3 双因素方差分析	178
习题 9	186
** 第 10 章 相关与回归分析	188
10.1 相关分析简介	188
10.2 回归分析的基本概念	192
10.3 一元线性回归模型	193
10.4 一元线性回归模型的显著性检验	199
10.5 一元线性回归模型的预测与控制	201
10.6 可化为一元线性回归的情形	205
习题 10	206
参考答案	208
附表	219
附表 1 几种常用的概率分布	219
附表 2 标准正态分布表	221
附表 3 泊松分布表	222
附表 4 t 分布表	223
附表 5 χ^2 分布表	224
附表 6 F 分布表	226
附表 7 均值的 t 检验的样本容量	
附表 8 均值差的 t 检验的样本容量	

第1章 随机事件及其概率

概率论是研究随机现象统计规律性的科学. 在自然界和人类社会实践中, 广泛地存在着两类不同的现象, 一类是确定性现象, 即一旦某条件实现就必然发生的现象. 例如, “在标准大气压下将水加热到 100°C , 会出现沸腾”; “向上抛一石子必然下落”; “同性电荷相互排斥”等. 另一类现象是非确定性现象, 即在一定条件实现后, 可能产生也可能不产生的现象, 人们称之为随机现象. 例如, 抛掷一枚均匀的硬币, 我们无法预知硬币在某一次抛掷中是正面向上还是反面向上. 又如, 某射击运动员用一支步枪在同一地点进行射击训练, 每次射击的成绩(环数)可能不同, 并且在每次射击之前, 均无法预知其射击后的成绩是多少……但是, 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和观察下, 它的结果却呈现某种规律性. 例如, 多次重复抛掷一枚均匀的硬币得到正面向上大致有一半, 如果我们在同样条件下大量重复抛掷这枚硬币, 那么可以预知的是出现正面与出现反面的次数差不多各占一半. 某射击运动员用一支步枪在同一地点大量重复地进行射击训练, 射击的成绩(环数)按照一定的规律分布等, 这种规律性是在大量的重复试验和观察中呈现出来的, 是随机现象所固有的, 称之为随机现象的统计规律性. 概率论与数理统计就是研究和揭示随机现象统计规律性的一门数学学科.

概率论与数理统计的思想与方法日益渗透到各个领域, 被广泛应用于自然科学、工程技术、经济理论、经营管理、医学、金融保险甚至人文科学等许多方面.

1.1 样本空间与随机事件

1.1.1 随机试验

为了叙述方便, 我们常把对某种现象的一次观察、测量或进行一次科学实验, 统称为一个试验.

下面是一些试验的例子:

E_1 : 抛一枚硬币, 观察正面 H 、反面 T 出现的情况;

E_2 : 将一枚硬币抛掷两次, 观察正面 H 、反面 T 出现的情况;

E_3 : 将一枚硬币抛掷两次, 观察出现正面的次数;

E_4 : 掷一颗骰子, 观察所掷的点数是几;

E_5 : 观察某城市某个月内交通事故发生的次数;

E_6 : 已知某物体的长度在 a 和 b 之间, 测量其长度;



E_7 : 在一批灯泡中任意抽取一只, 测试它的寿命;

E_8 : 向坐标平面区域 $D: x^2 + y^2 \leq 100$ 内随机投掷一点(假设点必落在 D 上), 观察落点 M 的坐标.

上面举出了 8 个试验的例子, 概括起来, 这些试验具有以下的特点:

(1) 可以在相同的条件下重复地进行.

(2) 每次试验的可能结果不止一个, 并且能事先明确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.

在概率论中, 我们将具有上述三个特点的试验称为**随机试验**, 简称为**试验**, 常用字母 E 来表示. 本书中以后所提到的试验都是指随机试验. 进行随机试验的目的是要通过对试验的各种结果的可能性进行分析, 从而找出随机现象的规律.

1.1.2 样本空间

随机试验 E 的所有可能出现的结果所构成的集合称为 E 的**样本空间**, 通常记为 Ω . 样本空间的元素, 即 E 的每一个结果, 称为**样本点**.

下面是前面试验 E_k ($k = 1, 2, 3, \dots, 8$) 的样本空间 Ω_k :

$$\Omega_1: \{H, T\}$$

$$\Omega_2: \{HH, HT, TH, TT\}$$

$$\Omega_3: \{0, 1, 2\}$$

$$\Omega_4: \{1, 2, 3, 4, 5, 6\}$$

$$\Omega_5: \{0, 1, 2, \dots\}$$

$$\Omega_6: \{l \mid a \leq l \leq b\}$$

$$\Omega_7: \{t \mid t \geq 0\}$$

$$\Omega_8: \{(x, y) \mid x^2 + y^2 \leq 100, \text{其中}(x, y) \text{为点 } M \text{ 的坐标}\}$$

注意: 样本空间的元素是由试验的目的所确定的, 例如, 在试验 E_2 和 E_3 中同是将一枚硬币抛掷两次, 由于试验的目的不一样, 其样本空间也不一样.

1.1.3 随机事件

一般地, 我们称随机试验 E 的样本空间 Ω 的子集为 E 的**随机事件**, 简称**事件**, 通常用大写字母 A, B, C 来表示. 在每次试验中, 当且仅当这一子集中的一个样本点出现时, 称这一事件发生.

特别地, 由一个样本点组成的单点集, 称为**基本事件**. 例如, 试验 E_1 中有两个基本事件 $\{H\}$ 和 $\{T\}$; 试验 E_2 有 4 个基本事件 $\{HH\}, \{HT\}, \{TH\}, \{TT\}$.

由 2 个或 2 个以上样本点构成的事件叫复杂事件. 例如, “抛掷一枚骰子出现奇数点”是复

杂事件,由出现1点、3点、5点这3个样本点所构成.

样本空间 Ω 包含所有的样本点,它是 Ω 自身的子集,在每一次试验中它总是发生的,称为**必然事件**.

空集 $\emptyset$ 不包含任何样本点,它也是样本空间 Ω 的子集,它在每一次试验中都不发生,称为**不可能事件**.

1.1.4 事件间的关系与事件的运算

事件是一个集合,因而事件间的关系与事件的运算类似于集合论中集合之间的关系和集合运算.下面给出这些关系和运算在概率中的提法,并根据“事件发生”的含义,给出它们在概率中的含义.

设试验 E 的样本空间为 Ω , $A, B, A_k (k=1, 2, 3, \dots)$ 为 Ω 的子集.

1. 事件的包含和相等

若 $A \subset B$,则称事件 B 包含事件 A ,这指的是事件 A 发生必导致事件 B 发生.

例如,在 E_4 中, A 表示“出现的点数为偶数”, B 表示“出现的点数大于1”. $A = \{2, 4, 6\}$,
 $B = \{2, 3, 4, 5, 6\}$,则 $A \subset B$,又称 A 是 B 的子集.

若 $A \subset B$ 且 $B \subset A$,则称事件 A 与事件 B **相等**,记为 $A = B$.

2. 事件的和

事件 $A \cup B = \{x | x \in A \text{ 或 } x \in B\}$ 称为事件 A 与事件 B 的**和事件**.当且仅当 A, B 中至少有一个发生时,事件 $A \cup B$ 发生. $A \cup B$ 也记为 $A + B$.

类似地,“ n 个事件 $A_1, A_2, \dots, A_n$ 中至少有一个发生”,称为事件 $A_1, A_2, \dots, A_n$ 的**和**,记为 $\bigcup_{k=1}^n A_k$;“可列个事件 $A_1, A_2, \dots, A_n, \dots$ 中至少有一个发生”,称为可列个事件 $A_1, A_2, \dots, A_n, \dots$ 的**和**,记为 $\bigcup_{k=1}^{\infty} A_k$.

例如,在 E_4 中,令 A 表示掷一颗骰子“出现的点数为偶数”, B 表示“出现的点数大于1”, A_k 表示出现的点数为 $k (k=1, 2, 3, \dots)$,那么

$$\begin{aligned} A &= \{2, 4, 6\}, & B &= \{2, 3, 4, 5, 6\}, \\ A \cup B &= \{2, 3, 4, 5, 6\}, & \bigcup_{k=1}^6 A_k &= \{1, 2, 3, 4, 5, 6\} = \Omega \end{aligned}$$

3. 事件的积

事件 $A \cap B = \{x \in A \text{ 且 } x \in B\}$ 称为事件 A 与 B 的**积事件**.当且仅当 A, B 同时发生时,事件 $A \cap B$ 发生, $A \cap B$ 也记为 AB .



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

4. 事件的差

事件 $A - B = \{x \in A \text{ 且 } x \notin B\}$ 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.

5. 互不相容的事件

若 $A \cap B = \emptyset$, 则称事件 A 与 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. 基本事件是两两互不相容的.

6. 对立事件

若 $A \cup B = \Omega$ 且 $A \cap B = \emptyset$, 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**. 这指的是对每次试验而言,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 $\bar{A}$, $\bar{A} = \Omega - A$.

由 $\bar{A}$ 的定义, $\overline{(\bar{A})} = A$, 即 $\bar{A}$ 与 A 互为对立事件. 在一次试验中, $\bar{A}$ 与 A 至少有一个发生,但 $\bar{A}$ 与 A 不会同时发生,即 $A \cup \bar{A} = \Omega$, $A \cap \bar{A} = \emptyset$, 事件 $A - B$ 也可记为 $A \cap \bar{B}$.

在这里,我们可以用文氏图直观地表示事件的关系和事件的运算. 用平面上的一个矩形表示样本空间 Ω , 矩形内的每个点表示一个样本点, 用两个小圆分别表示事件 A 和事件 B , 则事件的关系与运算可以用图 1-1 来表示.

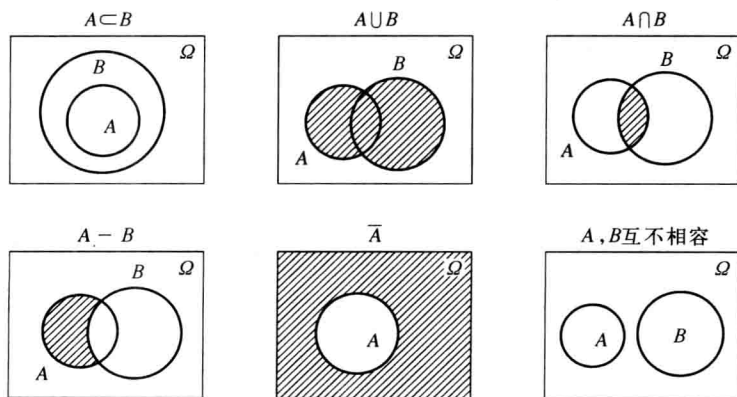


图 1-1

注: $A \cup B$, AB , $A - B$, $\bar{A}$ 分别为图中阴影部分.

7. 事件运算的主要性质

(1) 交换律: $A \cup B = B \cup A, A \cap B = B \cap A$ (2) 结合律: $A \cup (B \cup C) = (A \cup B) \cup C,$

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

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

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

(4) 德摩根(De Morgan)律(对偶律):

$$\overline{A \cup B} = \bar{A} \cap \bar{B}, \quad \overline{A \cap B} = \bar{A} \cup \bar{B}$$

分配律可以推广到有限或可列无穷多个事件的情形:

$$A \cap \left(\bigcup_{k=1}^n A_k \right) = \bigcup_{k=1}^n (A \cap A_k), \quad A \cap \left(\bigcup_{k=1}^{\infty} A_k \right) = \bigcup_{k=1}^{\infty} (A \cap A_k)$$

$$A \cup \left(\bigcap_{k=1}^n A_k \right) = \bigcap_{k=1}^n (A \cup A_k), \quad A \cup \left(\bigcap_{k=1}^{\infty} A_k \right) = \bigcap_{k=1}^{\infty} (A \cup A_k)$$

德摩根律也可以推广到有限或可列无穷多个事件的情形:

$$\overline{\bigcup_{k=1}^n A_k} = \bigcap_{k=1}^n \bar{A}_k, \quad \overline{\bigcup_{k=1}^{\infty} A_k} = \bigcap_{k=1}^{\infty} \bar{A}_k$$

$$\overline{\bigcap_{k=1}^n A_k} = \bigcup_{k=1}^n \bar{A}_k, \quad \overline{\bigcap_{k=1}^{\infty} A_k} = \bigcup_{k=1}^{\infty} \bar{A}_k$$

对偶律是很有用的性质,我们经常要使用到它.

另,易证下列等式的正确性:

$$A \cup A = A, \quad A \cap A = A, \quad A \cup \Omega = \Omega, \quad A \cap \Omega = A,$$

$$A \cup \emptyset = A, \quad A \cap \emptyset = \emptyset, \quad A \cup \bar{A} = \Omega, \quad A \cap \bar{A} = \emptyset,$$

$$A - B = A - AB = A \cap \bar{B}, \quad A \cup B = A + B\bar{A}, \quad A = (A \cap B) \cup (A \cap \bar{B})$$

例 1.1 某灯泡厂取样检查出厂灯泡的寿命,设 A 表示“灯泡寿命大于 1500 小时”, B 表示“灯泡寿命在 1000~2000 小时之间”,试用集合形式写出下列事件:

$$\Omega, A, B, A \cup B, AB, A - B, B - A$$

解 $\Omega = [0, +\infty), A = \{x \mid x > 1500\} = (1500, +\infty),$

$$B = \{x \mid 1000 \leq x \leq 2000\} = [1000, 2000], \quad A \cup B = [1000, +\infty),$$

$$AB = (1500, 2000], \quad A - B = (2000, +\infty), \quad B - A = [1000, 1500]$$

例 1.2 甲、乙、丙三人各向靶子射击一次,设 A_i 表示“第 i 人击中靶子”($i=1, 2, 3$). 试用事件的关系表示下列事件:(1) 仅有乙未击中靶;(2) 甲、乙至少一人击中而丙未击中靶;



(3) 至少两人击中靶; (4) 靶上仅中一弹.

解 (1) 仅有乙未击中靶: $A_1 \overline{A_2} A_3$.

(2) 甲、乙至少一人击中, 而丙未击中靶: $(A_1 \cup A_2) \overline{A_3}$.

(3) 至少两人击中靶: $A_1 A_2 \cup A_2 A_3 \cup A_1 A_3$.

(4) 靶上仅中一弹: $A_1 \overline{A_2} \overline{A_3} \cup \overline{A_1} A_2 \overline{A_3} \cup \overline{A_1} \overline{A_2} A_3$.

例 1.3 一工人生产了 n 个零件, 设 A_i 表示“第 i 个零件是正品” ($i = 1, 2, 3, \dots, n$). 试用文字叙述下列事件:

$$(1) \bigcap_{i=1}^n A_i \quad (2) \overline{\bigcap_{i=1}^n A_i} \quad (3) \bigcup_{i=1}^n \left[\overline{A_i} \cap \left(\bigcap_{\substack{k=1 \\ k \neq i}}^n A_k \right) \right]$$

解 (1) $\bigcap_{i=1}^n A_i$: n 个零件全是正品.

(2) $\overline{\bigcap_{i=1}^n A_i} = \bigcup_{i=1}^n \overline{A_i}$: 至少有一个零件不是正品.

(3) $\bigcup_{i=1}^n \left[\overline{A_i} \cap \left(\bigcap_{\substack{k=1 \\ k \neq i}}^n A_k \right) \right]$: 仅有一个零件不是正品.

1.2 事件的频率与概率

在研究随机试验的各种事件时, 不仅希望知道哪些随机事件可能出现, 还希望知道各种事件出现的可能性的. 希望找到一个合适的数来表征事件在一次试验中发生的可能性的. 为此, 首先引入频率, 它描述了事件发生的频繁程度, 进而引出表征事件在一次试验中发生的可能性的数——概率. 概率是描述事件发生可能性大小的数量指标, 它的定义是逐步形成和完善的.

定义 1.2.1 (频率定义) 在相同的条件下, 进行了 n 次试验, 在这 n 次试验中, 事件 A 发生的次数 n_A 称为事件 A 发生的频数. 比值 $\frac{n_A}{n}$ 称为事件 A 发生的频率, 并记为 $f_n(A)$.

例如, 在某铁矿床中抽取 120 个样品, 铁的品味在 27.5%~32.5% 范围内的样品有 35 个, 我们称数值 35 为“铁品味在 27.5%~32.5% 之间”随机事件的频数, 频数 35 与抽取的样品总数 120 之比称为铁品味落在 27.5%~32.5% 之间随机事件的频率, 即

$$f_n(A) = \frac{35}{120} = 0.292$$

由频率定义, 易知频率具有以下性质:

(1) 非负性: $f_n(A) \geq 0$

(2) 规范性: $f_n(\Omega) = 1$

(3) 有限可加性: 若 $A_1, A_2, \dots, A_k$ 是两两互不相容的事件, 则

$$f_n(A_1 \cup A_2 \cup \dots \cup A_k) = f_n(A_1) + f_n(A_2) + \dots + f_n(A_k)$$

频率的大小反映了事件 A 发生的频繁程度, 频率越大则意味着事件在试验中发生的可能性就越大. 但是, 用频率来表示一个事件发生的可能性的的大小却是行不通的. 这是因为频率具有波动性, 即使在相同条件下重复作多个 n 次试验, 其频率值亦可能不相同. 但是, 人们在实践中发现, 随着试验次数 n 的逐渐增加, 事件 A 发生的频率 $f_n(A)$ 一般会逐渐稳定在一个常数 p 附近, n 越大, 摆动的幅度会越小.

表 1-1 所示是“抛硬币”试验.

表 1-1

实 验 者	投掷次数 n	出现正面次数 k	频率 $f_n(A)$
德摩根	2 048	1 061	0.518 1
蒲 丰	4 040	2 048	0.506 9
K. 皮尔逊	12 000	6 019	0.501 6
K. 皮尔逊	24 000	12 012	0.500 5
维 尼	30 000	14 994	0.499 8

该实验表明, 尽管频率随试验不同而改变, 但是, 当次数 n 越来越大时, 它就愈趋于稳定在 0.5 这个常数附近. 这种在多次重复试验中出现的频率稳定性, 通常称之为随机事件的统计规律性. 当重复的次数 n 增加时, 这种现象就越明显, 频率的这种稳定性, 反映了随机事件本身固有的属性, 也就是说, 事件的概率是客观存在的.

在相同条件下, 重复地进行 n 次试验, 当 n 充分大, 事件 A 发生的频率 $f_n(A) = \frac{n_A}{n}$ 稳定地在某个数值 p 附近摆动, 则称 p 为事件 A 的概率, 记为 $P(A) = p$.

事件的频率稳定性是一个客观存在, 它不断地为人们的实践所证实. 上面我们用频率的稳定值来定义概率. 但是, 在实际中, 我们不可能对每一个事件都做大量的试验, 然后求得事件的频率, 用以表征事件发生可能性的大小. 同时, 为了理论研究的需要, 我们从频率的稳定性和频率的性质得到启发, 给出如下表征事件发生可能性大小的概率的公理化定义:

定义 1.2.2 (概率定义) 设 Ω 为一个随机试验的样本空间, 对 Ω 中任一事件 A , 规定一个实数 $P(A)$. 如果这样的集合函数 $P(\cdot)$ 满足以下 3 个条件, 则称 $P(A)$ 为事件 A 发生的概率.

(1) 非负性: 对于每一个事件 A 有 $P(A) \geq 0$;

(2) 规范性: 对于必然事件 Ω 有 $P(\Omega) = 1$;

(3) 可加性: 设 $A_1, A_2, \dots$ 是可列个两两互不相容事件, 即对于 $i \neq j, A_i A_j = \emptyset$ ($i,$