21世纪高等院校教材

概率论与数理统计

郭跃华 主编

陆志峰 钱 峰 赵为华 副主编

科学出版社
www.sciencep.com

概率论与数理统计

郭跃华　主编

陆志峰　钱　峰　赵为华　副主编

科学出版社

北　京

内 容 简 介

本书介绍概率论与数理统计的基本概念、基本理论和方法，主要内容包括：随机事件与概率、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、数理统计的基本概念、参数估计、假设检验、统计分析初步、随机过程等. 全书从直观分析入手逐步过渡到严格的数学表述，使学生通过学习，除了掌握概率统计的基本理论外，还能拓宽知识面，具备应用计算机解决问题的能力.

本书可作为普通高等院校的理工(非数学)、经管类各专业本科学生教材，也可供对数学要求较高的高职高专学生使用.

图书在版编目(CIP)数据

概率论与数理统计/郭跃华主编. —北京：科学出版社，2007
ISBN 978-7-03-019573-9

Ⅰ.概… Ⅱ.郭 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(2007) 第121979号

责任编辑：胡华强 李晓鹏 潘继敏／责任校对：陈玉凤
责任印制：张克忠 ／封面设计：王 浩

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

北京市文林印务有限公司 印刷
科学出版社发行 各地新华书店经销
*
2007年8月第 一 版 开本：B5(720×1000)
2007年8月第一次印刷 印张：20
印数：1—5 500 字数：379 000

定价：25.00元

(如有印装质量问题，我社负责调换〈文林〉)

前　言

"概率论与数理统计" 是培养学生利用随机思维模式看待和处理随机现象的一门重要的数学基础课程. 由于随机现象存在的普遍性、研究方法的独特性和教学内容的实用性, 这门课越来越受到人们的重视, 高等学校理、工、管类各专业普遍开设了概率论与数理统计课程. 该门课程研究的目的是从大量数据中研究随机现象蕴含的统计规律. 由于理论方法的应用涉及大量的计算, 其中过于深奥的抽象理论、复杂的公式、陈旧的计算手段限制了它的实用性. 为了适应现代化教学手段和计算机软件应用, 结合理、工、管类各专业的教学实际, 我们编写了《概率论与数理统计》这本教材.

理 (非数学)、工、管类大学生的数学素养主要是通过应用数学来体现, 因此我们在教材中突出数学建模的思想, 以加强应用数学的意识和能力的培养为核心, 结合当前学生状况, 在教学内容中删掉一些过于繁琐的推理和完全可以用计算机代替的计算, 压缩一些论述, 增加一些背景知识的介绍, 尽量增加一些与当今社会生活和现代科技密切相关的实际应用问题与综合应用问题. 本书引入了 Excel 进行统计计算和仿真, 为了不占用课堂时间, 我们以实验形式给出了计算和仿真的具体步骤, 可由学生课后自行利用计算机进行, 以提高学生的实际动手能力, 同时减少了花在繁琐计算上的时间, 使学生有更多的时间用在概念的理解和方法的掌握上.

本教材的主要特色如下:

(1) 增加一些概念背景, 由实际案例引进新概念.

(2) 删除一些复杂的理论推导和繁琐的人工计算, 突出重点, 加强方法的应用.

(3) 计算机工具的广泛使用. 通过计算机仿真实验、函数计算及程序调用, 把 Excel 工具广泛使用于概念的引进和数值计算, 帮助学生形象地理解新概念, 直达核心处理思想, 同时提高学生的动手能力.

(4) 贯穿随机思想, 强调实际应用. 通过应用案例分析方法, 引导学生灵活运用所学知识, 掌握随机处理的基本过程, 提高学生分析问题和解决实际问题的能力.

(5) 融入实验教学内容, 注重理论的实践与应用.

(6) 每一节配备适量与本节内容相关的练习题, 便于学生巩固课堂知识; 每一章配综合习题, 便于学生提高知识水平.

本教材分三部分: 第一部分为概率论 (第 1~5 章), 是后面应用部分的理论基础; 第二部分为数理统计 (第 6~9 章), 重点阐述了参数估计和假设检验, 并介绍了方差分析和回归分析两个常用的统计方法的最基本部分; 第三部分为随机过程 (第 10、11 章), 介绍了一些基本理论和几个常见的随机过程, 并着重讨论了马尔可夫链.

数理统计与随机过程内容相互独立, 可根据专业的需要选用. 本教材主要适用于对理论证明要求不高的大学本科各专业的概率统计课程.

全书各章分别由郭跃华 (第 1、9~11 章)、陆志峰 (第 6~8 章)、钱峰 (第 2、3 章)、赵为华 (第 4、5 章) 执笔. 最后由郭跃华统稿.

全书承蒙南京大学博士生导师、江苏省概率统计学会理事长王金德教授审阅, 并提出了许多宝贵意见, 在此表示衷心感谢. 本书得到科学出版社的大力支持, 在此一并表示衷心感谢.

限于编者水平, 有不当和错误之处, 恳请读者和专家批评指正.

郭跃华

2007 年 05 月

此为试读,需要完整PDF请访问: www.ertongbook.com

目　录

第 1 章　随机事件与概率 ······ 1

1.1　随机事件 ······ 2

1.2　事件的概率与性质 ······ 7

1.3　条件概率 ······ 17

1.4　事件的独立性 ······ 23

应用案例 1 ······ 26

习题 1 ······ 30

第 2 章　随机变量及其分布 ······ 32

2.1　随机变量及其分布函数 ······ 32

2.2　离散型随机变量的分布律 ······ 35

2.3　连续型随机变量的概率密度 ······ 42

2.4　随机变量函数的分布 ······ 49

应用案例 2 ······ 53

习题 2 ······ 55

第 3 章　多维随机变量及其分布 ······ 56

3.1　多维随机变量及其联合分布 ······ 56

3.2　边缘分布 ······ 63

3.3　条件分布 ······ 68

3.4　随机变量的独立性 ······ 72

3.5　二维随机变量的函数的分布 ······ 76

应用案例 3 ······ 83

习题 3 ······ 85

第 4 章　随机变量的数字特征 ······ 87

4.1　随机变量的数学期望 ······ 87

4.2　方差 ······ 95

4.3　协方差、相关系数和矩、协方差矩阵 ······ 101

应用案例 4 ······ 109

习题 4 ······ 112

第 5 章　大数定律与中心极限定理 ········ 114
5.1　大数定律 ········ 114
5.2　中心极限定理 ········ 117
应用案例 5 ········ 122
习题 5 ········ 123
第 6 章　数理统计的基本概念 ········ 124
6.1　基本概念 ········ 125
6.2　抽样分布 ········ 130
应用案例 6 ········ 136
习题 6 ········ 137
第 7 章　参数估计 ········ 139
7.1　点估计 ········ 139
7.2　估计量的评价准则 ········ 145
7.3　区间估计 ········ 148
应用案例 7 ········ 156
习题 7 ········ 157
第 8 章　假设检验 ········ 159
8.1　假设检验的基本概念 ········ 159
8.2　正态总体均值的假设检验 ········ 162
8.3　正态总体方差的假设检验 ········ 172
8.4　非参数假设检验 ········ 178
应用案例 8 ········ 182
习题 8 ········ 185
第 9 章　统计分析初步 ········ 187
9.1　单因素方差分析 ········ 187
9.2　双因素方差分析 ········ 195
9.3　一元线性回归分析 ········ 203
9.4　多元线性回归分析 ········ 219
9.5　非线性回归 ········ 222
应用案例 9 ········ 225
习题 9 ········ 234
第 10 章　随机过程的基本理论 ········ 237
10.1　随机过程的概念 ········ 237
10.2　随机过程的统计描述 ········ 242
10.3　随机过程的基本类型 ········ 245

应用案例 10 ······ 253
习题 10 ······ 257
第 11 章 Markov 链 ······ 258
11.1 Markov 过程的概念 ······ 258
11.2 多步转移概率 ······ 267
11.3 遍历性 ······ 272
应用案例 11 ······ 277
习题 11 ······ 281
参考文献 ······ 282
附表 ······ 283
习题参考答案 ······ 298

第 1 章　随机事件与概率

艾滋病普查

艾滋病 (acquired immune deficiency syndrome, 获得性免疫缺陷综合症, 缩写为 AIDS) 是现在全球较流行的一种致命的接触性传染病. 艾滋病的受害者几乎没有活过五年的. 到 2004 年全球艾滋病病毒感染者人数总计为 3940 万, 死亡人数为 310 万 (联合国艾滋病规划署和世界卫生组织《艾滋病流行最新报告——2004 年 12 月》). 中国卫生部、联合国艾滋病规划署和世界卫生组织 2006 年 1 月 25 日发布的最新数据显示, 中国现在的艾滋病病毒感染者人数为 65 万, 总感染率估计为 0.05%. 为了阻止这种病的传播, 首先要识别艾滋病病毒的携带者, 对这些人采取适当的预防措施以避免传染给其他人. 科学家一直在寻求识别艾滋病病毒携带者的有效方法. 其中一种方法就是血液试验 (ELISA, enzymelinked immunosorbent assay(酶连接免疫吸附测定) 的缩写), 它能检测出身体中艾滋病病毒的某种抗体的存在性. 尽管这种试验是极其精确的, 但它还是会有两种可能的误诊：首先, 它可能会对某些真有艾滋病的人作出没有艾滋病的诊断, 这就是所谓的"假阴性"；其次, 它可能会对某些没有艾滋病的人误诊为患有艾滋病, 这就是"假阳性".

假设血液试验改进到这样的程度, 它能正确识别患有艾滋病的人中的 95%, 因此 5% 的患有艾滋病的人的血液试验结果将是"假阴性". 进一步假设接受血液试验的不带艾滋病病毒的人中 99% 的试验结果为阴性. 这就意味着不带艾滋病病毒的健康人中的 1%, 其血液试验结果将是"假阳性".

美国是艾滋病较为流行的国家之一, 据保守估计大约每 1000 人中就有 1 人受这种病的折磨. 为了能有效地控制和减缓艾滋病的传播, 几年前有人就提议应在申请结婚登记的新婚夫妇中进行艾滋病病毒的血液试验. 该项普查计划一经提出, 立刻就遭到了许多专家学者的反对, 他们认为这是一项既费钱又费力, 同时收效不大的计划. 最终, 此项计划未被通过. 那么, 到底专家的意见对不对? 该普查计划该不该被执行呢? 如果该计划得以实施, 而某人又做了血液试验并且结果是阳性, 他将怎样为之发愁呢? 他真正得了艾滋病的可能性有多大?

在我们的生活中经常会面临许多包含有不确定性的决策问题. 如一笔新投资盈利的可能性有多大? 一项工程按期完成的可能性有多少? 等等.

解决上述这些问题需要运用一些概率论的基本知识. 概率论为解决不确定性问

题提供了最有效的理论和方法. 常言道 “天有不测风云, 人有旦夕祸福”, 风险无处不在, 无时不有. 无论在科研、生产、保险、风险投资、管理决策等许多领域, 还是在日常生活中, 人们都会遇到如何在面临不确定性的情况下做出正确决策的问题. 因此, 概率论有着广泛而重要的应用价值. 本章将介绍概率论的基础知识 —— 随机事件与概率.

1.1　随 机 事 件

1.1.1　基本术语

人类社会和自然界发生的现象多种多样. 其中有一类现象, 其规律是只要具备一定的条件, 某确定的现象一定发生 (或一定不发生). 例如:

如果平面图形是三角形 (条件), 则其内角和一定是 180°(现象);

在一个标准大气压、温度 100°C(条件) 下, 纯水一定沸腾 (现象);

两个同性电荷 (条件) 一定互斥 (现象);

冬天过去 (条件), 春天就会到来 (现象);

在常温 (条件) 下, 铁一定不熔化 (现象);

等等.

定义 1.1.1　在一定条件下, 一定发生 (或一定不发生) 的现象称为**必然现象**.

与必然现象不同, 还存在另一类现象, 其规律是, 在一定条件下, 某种现象可能发生, 也可能不发生. 例如:

抛掷一枚质地均匀的硬币 (条件), 硬币落地后, 带币值的一面朝上 (现象) 可能发生, 也可能不发生;

火炮对坦克射出一发炮弹 (条件), 命中坦克 (现象) 可能发生, 也可能不发生;

同时抛掷两颗质地均匀的骰子 (条件), 点数朝上之和恰好为 6(现象) 可能发生, 也可能不发生;

……

定义 1.1.2　在一定条件下, 可能发生, 也可能不发生的现象称为**随机现象**.

概率论就是研究随机现象中的数量规律的科学.

定义 1.1.3　在一定条件下, 对某随机现象的一次观察或测量称为**随机试验**(简称**试验**), 记为 E. 随机试验具备以下三条性质:

(1) 试验能在相同条件下重复进行;

(2) 试验结果不止一个, 但能明确所有可能的结果;

(3) 试验前不能预知试验会出现哪个结果.

进行一次试验,总要有一个观测目的. 试验中可能观测到多种不同的结果. 以抛一枚硬币为例,如果观测的目的只是看硬币落地后哪一面朝上,由于硬币不变,所以掷硬币的行为可以在相同条件下重复实施. 在掷硬币之前,我们知道出现的结果只能是币值一面或另一面,但具体是哪一面出现不能确切地预料. 可见这种情形是一个随机试验,记为 E_1.

下面再看一些随机试验的例子.

E_2:掷一颗骰子,观察出现的点数;

E_3:一枚硬币重复掷三次,观察其各面出现的情况;

E_4:观察总机每天 9:00~10:00 接到的电话次数;

E_5:从一批电子元件中任取一件测量其使用寿命;

E_6:观察某地区每天的最高温度与最低温度.

定义 1.1.4 设 E 是一试验,其所有可能出现的结果组成的集合称为试验 E 的**样本空间**,记为 Ω(或 S). 样本空间的元素,也就是随机试验的直接结果称为**样本点**.

前面所举的 6 个试验中,若以 Ω_i 表示试验 E_i 的样本空间,i=1,2,$\cdots$,6,则

$\Omega_1=\{H, T\}$,其中 T 表示出现币值一面,H 表示出现另一面;

$\Omega_2=\{1,2,3,4,5,6\}$;

$\Omega_3=\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$;

$\Omega_4=\{0, 1, 2, \cdots\}$,样本点个数为可数无限多个;

$\Omega_5=\{t|t\geqslant 0\}$,$t$ 表示所抽取的元件寿命为 t 小时;

$\Omega_6=\{(x, y)|\quad T_1\leqslant x\leqslant y\leqslant T_2\}$,其中 T_1, T_2 分别是该地区的最低与最高温度.

值得注意的是,对一个随机现象观测的方式和目的有所不同时,相应的样本空间也可能有所不同. 如在 E_3 中考察的是币面出现次数时,样本空间表示为 $\Omega_3=\{0,1,2,3\}$.

定义 1.1.5 随机试验的若干个结果组成的集合称为**随机事件**(简称**事件**),常用大写字母 $A, B, C, \cdots$ 表示. 只含一个试验结果的事件称为**基本事件**.

即事件是样本空间 Ω 的子集,基本事件就是 Ω 的样本点.

在试验 E_2 中,A_i=“掷出 i 点” (i=1, 2, $\cdots$, 6),B=“掷出偶数点”,C=“掷出的点数不大于 4”,则 $A_i=\{i\}(i=1, 2, \cdots, 6)$ 都是基本事件,$B=\{2, 4, 6\}$,$C=\{1, 2, 3, 4\}$.

在试验中,当事件 (集合) 中的一个样本点 (元素) 出现时,称这一事件发生. 如试验 E_2 中,当投掷的结果为 “3 点” 时,事件 A_3, C 均发生.

由于样本空间 Ω 包含了所有的样本点,且是 Ω 自身的一个子集,在每次试验中它总是发生的,所以称 Ω 为**必然事件**. 空集 $\varnothing$ 不包含任何样本点,它也是样本空间的一个子集,且在每次试验中总不发生,所以称 $\varnothing$ 为**不可能事件**.

1.1.2 事件的关系与运算

对于一个随机试验来说,有很多随机事件,有的随机事件可能很复杂. 为了从较

此为试读,需要完整PDF请访问: www.ertongbook.com

简单事件的出现规律中寻求复杂事件出现的规律性, 我们需要研究同一试验的各种事件之间的关系和运算. 由于事件实际上是样本空间中的某一个子集, 因此, 事件之间的关系和运算同集合之间的关系和运算是一致的. 对应着集合的关系和运算, 我们定义事件的关系和运算如下:

设 Ω 是试验 E 的样本空间, A, B, C 及 A_1, A_2, A_3, $\cdots$ 都是事件, 即 Ω 的子集.

1. 事件的包含

如果事件 A 的发生必然导致事件 B 的发生, 则称事件 B**包含**事件 A, 或称事件 A**包含于**事件 B, 记作 $A\subset B$. 显然对任何事件 A, 有 $\varnothing\subset A\subset\Omega$.

2. 事件的相等

$$A=B\Leftrightarrow A\subset B \text{ 且 } B\subset A.$$

3. 事件的并 (和)

"两个事件 A, B 中至少有一个发生" 也是一个事件, 称为事件 A 与 B 的**和**(**并**), 记作 $A\bigcup B$ 或 $A+B$.

"事件 A_1, A_2, $\cdots$, A_n 至少有一个发生" 称为 A_1, A_2, $\cdots$, A_n 的和事件, 记为 $\bigcup\limits_{i=1}^{n}A_i$.

"事件 A_1, A_2, $\cdots$, A_n, $\cdots$ 至少有一个发生" 称为 A_1, A_2, $\cdots$, A_n, $\cdots$ 的和事件, 记为 $\bigcup\limits_{i=1}^{\infty}A_i$.

4. 事件的交 (积)

"两个事件 A 与 B 同时发生" 是一个事件, 称为事件 A 与 B 的**积** (**交**), 记作 $A\bigcap B$ 或 AB.

"事件 A_1, A_2, $\cdots$, A_n 同时发生" 称为 A_1, A_2, $\cdots$, A_n 的积事件, 记为 $\bigcap\limits_{i=1}^{n}A_i$.

"事件 A_1, A_2, $\cdots$, A_n, $\cdots$ 同时发生" 称为 A_1, A_2, $\cdots$, A_n, $\cdots$ 的积事件, 记为 $\bigcap\limits_{i=1}^{\infty}A_i$.

5. 事件的差

"事件 A 发生而事件 B 不发生" 是一个事件, 称为事件 A 与 B 的**差**, 记作 $A-B$.

6. 事件的互斥 (互不相容)

如果事件 A 与事件 B 不能同时发生, 即 $AB=\varnothing$, 称事件 A 与 B**互不相容** (**互斥**). 显然基本事件间是互不相容的.

A_1, A_2, $\cdots$, A_n 两两互斥 $\Leftrightarrow A_iA_j=\varnothing$, $\quad i\neq j$, $i,j=1$, 2, $\cdots$, n.

A_1, A_2, $\cdots$, A_n, $\cdots$ 两两互斥 $\Leftrightarrow A_iA_j=\varnothing$, $\quad i\neq j$, $i,j=1$, 2, $\cdots$.

7. 事件的对立

"事件 A 不发生" 是一个事件, 称为 A 的**对立事件**(**或逆事件**), 记作 $\overline{A}$. 由于 A 也是 $\overline{A}$ 的对立事件, 因此, 称 A 与 $\overline{A}$ 互为对立事件. 由定义知, 两个对立事件一定是互不相容事件; 反之, 两个互不相容事件却不一定是对立事件. 显然有

$$A\overline{A}=\varnothing,\quad A\bigcup\overline{A}=\Omega,\quad \overline{A}=\Omega-A,\quad \overline{\overline{A}}=A.$$

注意 "A 与 B 互相对立" 与 "A 与 B 互斥" 是不同的概念.

事件的关系与运算可用图 1.1 的文氏图直观地来表示, 其中 $A\bigcup B$, AB, $A-B$, $\overline{A}$ 分别为图中的阴影部分.

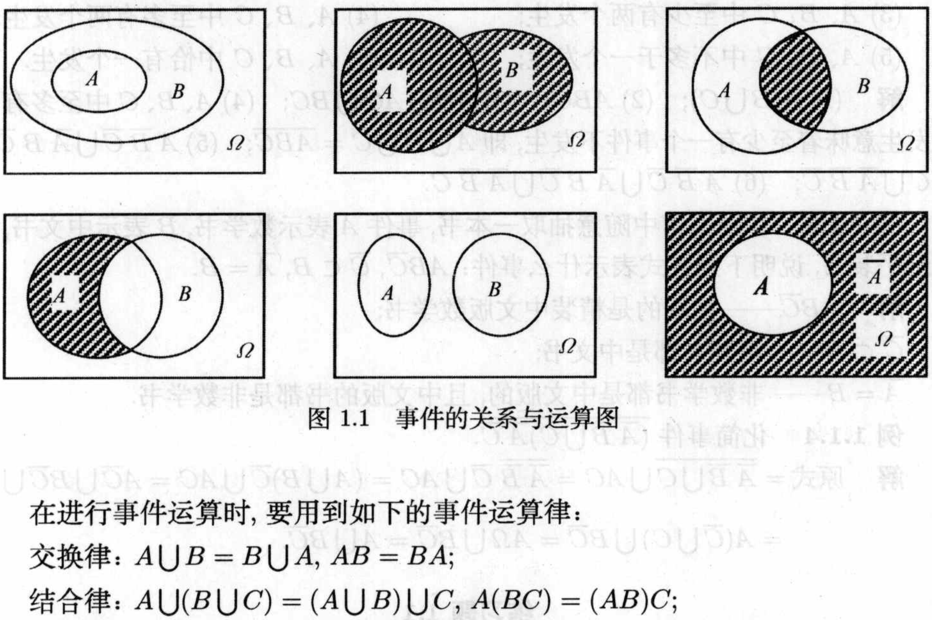

图 1.1 事件的关系与运算图

在进行事件运算时, 要用到如下的事件运算律:

交换律: $A\bigcup B=B\bigcup A$, $AB=BA$;

结合律: $A\bigcup(B\bigcup C)=(A\bigcup B)\bigcup C$, $A(BC)=(AB)C$;

分配律: $A(B\bigcup C)=(AB)\bigcup(AC)$, $A\bigcup(BC)=(A\bigcup B)(A\bigcup C)$;

对偶律 (De Morgan 律): $\overline{A\bigcup B}=\overline{A}\,\overline{B}$, $\overline{AB}=\overline{A}\bigcup\overline{B}$.

对于多个随机事件, 上述运算律也成立. 如 $\overline{A_1\bigcup A_2\bigcup\cdots\bigcup A_n}=\overline{A_1}\ \overline{A_2}\cdots\overline{A_n}$, $\overline{A_1A_2\cdots A_n}=\overline{A_1}\bigcup\overline{A_2}\bigcup\cdots\bigcup\overline{A_n}$ 等.

此外还有

吸收律: $A\bigcup\Omega=\Omega$, $A\Omega=A$, $A\bigcup\varnothing=A$, $A\bigcap\varnothing=\varnothing$;

重余律: $\overline{\overline{A}}=A$;

幂等律: $A\bigcup A=A$, $A\bigcap A=A$;

差化积: $A-B=A\overline{B}=A-AB$.

例 1.1.1　掷两颗骰子, 记事件 A 为 “两颗骰子的点数之和为 5”, 事件 B 为 “两颗骰子中最大点数为 3”, 求 $A, B, A\cup B, AB, A-B, B-A$ 的集合表达式.

解　$A=\{(1,4), (4,1), (2,3), (3,2)\}$;

$B=\{(1,3), (2,3), (3,3), (3,2), (3,1)\}$;

$A\bigcup B=\{(1,4), (4,1), (2,3), (3,2), (1,3), (3,3), (3,1)\}$;

$AB=\{(2,3), (3,2)\}$;

$A-B=\{(1,4), (4,1)\}$;

$B-A=\{(1,3), (3,3), (3,1)\}$.

例 1.1.2　设 A、B、C 表示三个事件, 试用 A、B、C 表示如下事件:

(1) A 发生且 B、C 至少有一个发生;　　(2) A 与 B 发生而 C 不发生;

(3) A、B、C 中至少有两个发生;　　(4) A、B、C 中至多有两个发生;

(5) A、B、C 中不多于一个发生;　　(6) A、B、C 中恰有一个发生.

解　(1) $A(B\bigcup C)$;　(2) $AB\overline{C}$;　(3) $AB\bigcup AC\bigcup BC$;　(4) A、B、C 中至多有两个发生意味着至少有一个事件不发生, 即 $\overline{A}\bigcup\overline{B}\bigcup\overline{C}=\overline{ABC}$;　(5) $A\,\overline{B}\,\overline{C}\bigcup\overline{A}\,B\,\overline{C}\bigcup\overline{A}\,\overline{B}\,C\bigcup\overline{A}\,\overline{B}\,\overline{C}$;　(6) $A\,\overline{B}\,\overline{C}\bigcup\overline{A}\,B\,\overline{C}\bigcup\overline{A}\,\overline{B}\,C$.

例 1.1.3　在图书馆中随意抽取一本书, 事件 A 表示数学书, B 表示中文书, C 表示平装书, 说明下列各式表示什么事件: $AB\overline{C}$, $\overline{C}\subset B$, $\overline{A}=B$.

解　$AB\overline{C}$—— 抽取的是精装中文版数学书;

$\overline{C}\subset B$—— 精装书都是中文书;

$\overline{A}=B$—— 非数学书都是中文版的, 且中文版的书都是非数学书.

例 1.1.4　化简事件 $\overline{(\overline{A}\,\overline{B}\bigcup C)\overline{A}\,\overline{C}}$.

解　原式 $=\overline{\overline{A}\,\overline{B}\bigcup C}\bigcup AC=\overline{\overline{A}}\,\overline{\overline{B}}\,\overline{C}\bigcup AC=(A\bigcup B)\overline{C}\bigcup AC=A\overline{C}\bigcup B\overline{C}\bigcup AC$

$$=A(\overline{C}\bigcup C)\bigcup B\overline{C}=A\Omega\bigcup B\overline{C}=A\bigcup B\overline{C}.$$

练习题 1.1

1. 试写出下列随机试验的样本空间.

(1) 同时抛三颗骰子, 记录三颗骰子点数之和;

(2) 记录一个班级一次考试的平均分数 (以百分制记);

(3) 记录某话务员在一个工作日内接听电话的次数;

(4) 生产产品直到有 10 件正品为止, 记录生产产品的总件数;

(5) 一个口袋中有 5 只外形相同的球, 编号为 1, 2, 3, 4, 5, 从中同时取 3 只球.

2. 将下列事件用事件 A、B、C 表示出来.

(1) A 发生;

(2) 只有 A 发生;

(3) 3 个事件都发生;

(4) 3 个事件中至少有 2 个发生;

(5) 3 个事件中恰好有 2 个发生;

(6) 3 个事件都不发生;

(7) 3 个事件中不多于 2 个发生;

(8) 3 个事件中不多于 1 个发生.

3. 从某班学生中任选一名学生, 设 $A=$\{选出的学生是男生\}, $B=$\{选出的学生是数学爱好者\}, $C=$\{选出的学生是班干部\}, 试问下列运算结果分别表示什么事件.

(1) ABC ; (2)$\overline{A}B\overline{C}$; (3)$\overline{A\bigcup C}$; (4)$A-(B\bigcup C)$.

4. 设 $\Omega=\{1, 2, \cdots, 10\}$, $A=\{2, 3, 4\}$, $B=\{3, 4, 5\}$, $C=\{5, 6, 7\}$, 具体写出下列各式表示的集合.

(1)$\overline{A}\bigcup B$; (2)$\overline{A}B$; (3)$\overline{A}\ \overline{B}$; (4)$\overline{AB\bigcup C}$; (5)$\overline{A(B\bigcup C)}$.

1.2 事件的概率与性质

在 1.1 节介绍了随机试验, 样本空间和随机事件等基本概念. 在一个随机试验中, 对随机事件发生的可能性大小如何作定量分析研究? 以下将要给出的“概率”这一概念正是对随机事件发生可能性大小的一种度量. 为此, 先介绍一点预备知识——频率的概念.

1.2.1 频率

定义 1.2.1 设随机事件 A 在 n 次重复试验中发生了 m 次, 则称比值

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

为事件 A 在 n 次重复试验中发生的**频率**.

由定义, 频率具有如下性质:

(1) 对任何事件 A, $f_n(A)\geqslant 0$(非负性);

(2) $f_n(\Omega)=1$(归一性);

(3) 若事件 A、B 互不相容, 即 $AB=\varnothing$, 则

$$f_n(A\bigcup B)=f_n(A)+f_n(B).$$

一般地, 对任意有限多个两两互不相容的事件 $A_1, A_2, \cdots, A_n$, 即 $A_iA_j=\varnothing(i\neq j, i, j=1, 2, \cdots, n)$, 则有

$$f_n\Big(\bigcup_{i=1}^{n}A_i\Big)=\sum_{i=1}^{n}f_n(A_i)\qquad(\text{有限可加性}).$$

人们经过长期的实践发现, 虽然随机事件在一次试验中可能发生, 也可能不发生, 带有不确定性, 但当重复试验次数 n 充分大时, 随机事件 A 的频率总在一确定的数值附近摆动, 有稳定于确定值的趋势.

频率在一定条件下具有稳定性, 这可以用第 5 章的大数定律加以严格证明, 也可以用试验加以验证. 历史上曾经有一些著名的统计学家进行过抛掷硬币的试验, 其结果如表 1.1 所示.

表 1.1　硬币试验

试验者	投掷次数 n	正面向上次数 m	频率 f_n
蒲丰 (Buffon)	4040	2048	0.5069
皮尔逊 (K.Pearson)	12000	6019	0.5016
皮尔逊 (K.Pearson)	24000	12012	0.5005

模拟试验我们可用 Excel 软件进行仿真.

实验 1.1　抛掷硬币的仿真试验.

具体步骤如下:

第 1 步: 进入 Excel 表格界面, 在 Excel 主菜单中依次选择 "工具"、"数据分析";

第 2 步: 在 "数据分析" 对话框中选择 "随机数发生器", 单击 "确定";

第 3 步: 出现 "随机数发生器" 对话框, 分别输入变量个数 (指试验轮数)、随机数个数 n(试验次数)、分布 (这里选伯努利)、参数 (这里取 0.5)、输出选项, 单击 "确定", 得所产生的 0、1 随机数;

第 4 步: 根据随机数是 1 或 0 确定硬币抛掷后出现正面或反面, 随机数求和可得频数 m, 进一步可计算频率 $\dfrac{m}{n}$.

注　如果点击 "工具" 后未出现 "数据分析" 项, 则需点击 "加载宏", 进行该项功能加载.

与物理化学中的实验类似, 在概率统计中模拟仿真的正确使用常可帮助我们验证结论, 建议读者通过上机实验逐步掌握计算机软件在概率统计中的一些重要功能.

从上面的试验记录可以看到, 在多次重复试验中, 事件 "正面向上" 发生的频率虽不完全相同, 但却在一固定的数值 0.5 附近摆动而呈现出一定的稳定性, 频率随着试验次数 n 的增大越来越有接近于数值 0.5 的趋势. 频率的这种稳定性表明一个随机事件发生的可能性有一定的大小可言: 频率若稳定于较大的数值, 表明相应事件发生的可能性较大; 频率若稳定于较小的数值, 表明相应事件发生的可能性较小. 而频率所接近的这个固定数值就可作为相应事件发生可能性大小的一个客观的、定量的度量.

定义 1.2.2(概率的统计定义) 设随机事件 A 在 n 次重复试验中发生的次数为 m, 当试验次数 n 很大时, 如果频率 $\frac{m}{n}$ 稳定在某一数值 p 的附近摆动, 且随作试验次数 n 的增加, 其摆动的幅度越来越小, 则称数值 p 为随机事件 A 的**概率**, 记作 $P(A)=p$.

概率的统计定义有相当直观的试验背景, 易被人们接受. 但定义本身存在很大的缺陷, 即定义中的"稳定在某一数值 p 的附近摆动"及"摆动的幅度越来越小"等内容模糊不清, 数值 p 无法准确求得, 且在实际问题中, 不可能也不必要对每个事件都要做大量重复试验, 从中得到频率的稳定值. 为了理论研究与实际应用的需要, 由频率的稳定性和频率的性质得到启发, 给出度量随机事件发生可能性大小的概率的公理化定义.

1.2.2 概率的公理化定义

定义 1.2.3 设 Ω 是随机试验 E 的样本空间, 对于 Ω 中的每一个事件 A, 赋于一个满足下列三条公理的实数 $P(A)$:

(1) 非负性: $P(A)\geqslant 0$;

(2) 归一性: $P(\Omega)=1$;

(3) 可列可加性: 若 $A_1, A_2, \cdots, A_n, \cdots$ 两两互不相容, 即 $A_iA_j=\varnothing(i\neq j, i, j=1, 2, \cdots)$, 则有

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

称实数 $P(A)$ 为事件 A 的**概率**.

由概率的公理化定义可推得概率的如下性质:

性质 1 不可能事件 $\varnothing$ 的概率为 0, 即 $P(\varnothing)=0$.

证 因 $\Omega=\Omega\bigcup\varnothing\bigcup\varnothing\bigcup\cdots$, 由定义 1.2.3(3) 可得

$$P(\Omega)=P(\Omega)+P(\varnothing)+P(\varnothing)+\cdots,$$

而 $P(\varnothing)\geqslant 0$, 故由上式知 $P(\varnothing)=0$.

性质 2 概率具有有限可加性, 即若 $A_iA_j=\varnothing(i\neq j, i, j=1, 2, \cdots, n)$, 则

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i).$$

证 $$P\left(\bigcup_{i=1}^{n}A_i\right)=P(A_1\bigcup\cdots\bigcup A_n\bigcup\varnothing\bigcup\cdots),$$

由定义 1.2.3(3) 及 $P(\varnothing)=0$ 得 $P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)$.

此为试读, 需要完整PDF请访问: www.ertongbook.com