随机信号分析
理论与实践

王仕奎 · 编著

SUIJI XINHAO FENXI
LILUN YU SHIJIAN

东南大学出版社
SOUTHEAST UNIVERSITY PRESS

随机信号分析理论与实践

王仕奎　编著

东南大学出版社
SOUTHEAST UNIVERSITY PRESS
·南京·

内容简介

本书介绍了随机过程的基本理论及在随机信号分析中的应用，包括概率论基础、随机过程及其统计特征、随机信号通过系统分析、窄带随机信号分析和马尔可夫过程及其应用等内容。本书特色在于以创建应用型大学思想为指导，采取理论结合实际的方法，在介绍基本理论的同时，辅以大量的仿真程序，从理论和实践两个方面展示随机信号分析的研究方法和结果，并对仿真的基本要求和技巧进行了较为详细的说明。

本书可以作为电子、信息类本科及研究生教学用书，也可以作为数学专业本科生、相关科研工作者及对MALAB应用感兴趣的人员的参考书。

图书在版编目(CIP)数据

随机信号分析理论与实践 / 王仕奎编著. —南京：东南大学出版社，2016.8

ISBN 978-7-5641-6647-2

Ⅰ.①随… Ⅱ.①王… Ⅲ.①随机信号—信号处理—高等学校—教材 Ⅳ.①TN911.7

中国版本图书馆CIP数据核字(2016)第174221号

随机信号分析理论与实践

出版发行	东南大学出版社
出 版 人	江建中
社　　址	南京市四牌楼2号(210096)
网　　址	http://www.seupress.com
经　　销	全国各地新华书店
印　　刷	南京工大印务有限公司
开　　本	787 mm×1092 mm 1/16
印　　张	18.75
字　　数	445千字
版　　次	2016年8月第1版
印　　次	2016年8月第1次印刷
书　　号	ISBN 978-7-5641-6647-2
定　　价	38.00元

本社图书若有印装质量问题，请直接与营销部联系。电话(传真)：025-83791830

前　言

随机信号分析是随机过程与数字信号处理相结合而产生的一门学科，它研究随机信号的基本理论和实际应用。随机过程是研究随机现象的理论，它和研究确定现象的理论是相对的，是自然科学、工程技术及社会科学各领域中研究随机现象的重要工具。将随机过程运用于随机信号的分析和处理，就产生了随机信号分析理论。随机信号分析以概率论、信号与线性系统分析和数字信号处理等为基础，同时也为数字信号处理及系统分析提供新的方法和理论结果。随机信号分析是通信原理、移动通信和网络技术等理论与技术的重要基础。

随着信息、通信工程理论的发展，随机过程理论的侧重点发生了一些变化：由于计算机和数字信号处理技术的快速发展，计算机仿真和离散数字信号处理技术得到了广泛应用；由于网络和语音识别等技术的快速发展，马尔可夫链与排队论等理论显得日益重要；由于通信原理和移动通信技术的普及，随机信号通过线性和非线性系统的理论及窄带随机信号与系统的理论显得尤其重要。

为了适应上述需求，笔者写作了本书。本书的特色，一是将理论与实践紧密结合起来，大量采用 MATLAB 仿真，其优势是在纯数学的方法无能为力时，起到补充的作用，还可以用来探索新的结果，或者对新的理论进行验证；二是侧重于工程应用，不拘泥于数学上的严格性，尽量避免学究式的纯粹数学推导，对一些重要的概念仅仅通过实例说明其含义，避免枯燥而繁琐的定义，重点阐述随机过程在信号分析中的实际运用。伟大的科学家牛顿有一句名言：一个例子比十个定理更有效(An example is more effective than ten theorems)。本书通过大量的例子，说明基本理论的实际应用，并对仿真技术的要求和基本技巧进行较为详细的说明。通过这些例子，进一步加深读者对基本理论的理解，体会其运用技巧。

为了保持知识的连续性，介绍了概率论和线性系统分析等基础知识。常言说，兴趣是最好的老师，为了提高学习效果，本书尽量举出一些生动有趣的例子，并常常采用一题多解的形式，以使读者体会不同方法的优劣。考虑到该书是为其他相近专业课程打下基础，因此在重点讲述随机过程的基础上，还介绍了随机信号通过线性系统、窄带高斯信号等重要内容。考虑到马尔可夫过程在信息论和信号处理中的广泛采用，对马尔可夫过程及隐马尔可夫过程也进行了较为详细的介绍。

本书如作为电子、信息类等专业的教材，建议基本学时数为48学时，其中理论学时和上机学时各为24学时，以贯彻理论紧密结合实际的主张。

本书的写作得到重庆三峡学院聂祥飞教授、陈立万教授等的大力支持，肖化武副教授、刘毓副教授和丁楠副教授等也提出一些很好的建议，在此一并表示衷心的感谢。

由于作者水平有限，书中疏漏和不妥之处肯定存在，请广大读者不吝批评指正，以便再版时修订，联系方式：wangshikui2002@aliyun.com。

王仕奎

2016年2月于重庆

目　　录

第 1 章　随机分析的数学基础 ……………… 1

1.1　随机现象基本概念 ……………………… 1

1.1.1　随机事件的关系及运算 ………… 2

1.1.2　随机事件的运算律 ……………… 4

1.1.3　概率的定义及其性质 …………… 4

1.1.4　古典概型及其蒙特·卡诺模拟 ………………………………… 5

1.1.5　条件概率与全概率公式 ………… 8

1.1.6　贝努利大数定律及其应用 …… 9

1.1.7　事件的独立性 ………………… 28

1.2　一维随机变量及其概率分布 ……… 30

1.2.1　随机变量及其分布函数……… 30

1.2.2　离散型随机变量及其分布律 … 30

1.2.3　连续型随机变量及其分布律 … 43

1.3　多维随机变量及其分布函数 ……… 56

1.4　随机变量函数的概率分布 ………… 58

1.5　随机变量的数字特征 ……………… 61

1.6　本章小结 …………………………… 66

第 2 章　随机过程基本理论 …………… 68

2.1　随机过程 …………………………… 68

2.1.1　随机过程的定义 …………… 68

2.1.2　随机过程的分类 …………… 72

2.1.3　随机过程的概率分布 ……… 72

2.2　随机过程的数字特征 ……………… 73

2.2.1　数学期望 ……………………… 73

2.2.2　均方值与方差 ………………… 74

2.2.3　自相关函数和协方差函数…… 75

2.2.4　互相关函数和互协方差函数 … 77

2.3　随机过程的平稳性和遍历性 ……… 78

2.3.1　严平稳随机过程 ……………… 78

2.3.2　宽平稳随机过程 ……………… 80

2.3.3　宽平稳过程自相关函数的性质 ………………………………… 85

2.3.4　宽平稳过程的均方遍历性 …… 87

2.4　随机信号的谱分析 ………………… 92

2.4.1　经典傅里叶分析回顾 ……… 92

2.4.2　随机过程的功率谱密度……… 96

2.4.3　随机信号自相关函数的估计 … 99

2.4.4　自相关函数估计的实际应用 ——基音周期分析 …………… 105

2.4.5　平稳随机信号的谱估计 …… 108

2.5　几种典型的随机过程 …………… 117

2.5.1　二阶矩过程 ………………… 118

2.5.2　独立随机过程 ……………… 118

2.5.3　独立增量过程和平稳增量过程 ………………………………… 118

2.5.4　马尔可夫过程 ……………… 119

2.5.5　平稳随机过程 ……………… 119

2.5.6　高斯过程(正态过程) ………… 119

2.5.7　维纳过程 …………………… 125

2.6　本章小结 ………………………… 127

第 3 章　随机信号通过系统分析 ……… 128

3.1　确定信号通过线性系统分析 …… 128

3.2　随机信号通过线性系统分析 …… 131

3.2.1　输出信号的数学期望、方差和自相关函数 ………………………… 132

3.2.2　输出信号与输入信号的互相关函数 ………………………………… 136

3.2.3 输出信号的功率谱密度 …… 137
3.3 白噪声通过线性系统分析 ……… 144
3.3.1 噪声带宽 ……………………… 144
3.3.2 白噪声通过理想线性系统 … 146
3.3.3 白噪声通过线性系统的应用举例 ……………………………… 152
3.4 随机信号通过非线性系统分析…… 155
3.4.1 随机信号通过平方律检波器 ……………………………… 156
3.4.2 随机信号通过半波线性检波器 ……………………………… 162
3.4.3 随机信号通过乘法器 ……… 166
3.4.4 维纳滤波器和卡尔曼滤波器简介 ……………………………… 168
3.5 本章小结 ……………………… 173
第4章 窄带随机信号分析 ……… 175
4.1 窄带随机信号的概念 …………… 175
4.2 希尔伯特变换……………………… 177
4.2.1 希尔伯特变换的定义 ……… 177
4.2.2 希尔伯特变换的性质 ……… 183
4.2.3 希尔伯特变换在通信中的应用 ……………………………… 187
4.2.4 窄带信号的复数表示 ……… 191
4.2.5 希尔伯特-黄变换及其应用 … 193
4.2.6 小波变换及其应用 ………… 199
4.3 窄带随机过程的包络和相位分布 … 205
4.3.1 窄带随机过程的同相和正交分解 ……………………………… 205
4.3.2 同相分量与正交分量的统计特性 ……………………………… 207
4.3.3 窄带高斯过程包络和相位的一维概率密度 …………………… 210
4.3.4 窄带高斯过程包络和相位的二维概率密度 …………………… 212
4.3.5 窄带高斯过程包络平方的概率密度 ……………………………… 214
4.4 余弦信号加窄带高斯过程 ……… 215
4.4.1 余弦信号加窄带高斯过程包络和相位的分布 ………………… 215
4.4.2 余弦信号加窄带高斯过程包络平方的分布 ………………… 218
4.4.3 余弦信号加窄带高斯过程的应用 ……………………………… 218
4.5 χ^2 分布和非中心 χ^2 分布………… 223
4.5.1 χ^2 分布 ……………………… 223
4.5.2 非中心 χ^2 分布 ……………… 224
4.6 本章小结 ……………………… 226
第5章 马尔可夫过程……………… 227
5.1 马尔可夫过程的概念 ………… 227
5.2 齐次马尔可夫链及其平稳分布…… 228
5.2.1 马尔可夫链 ………………… 228
5.2.2 齐次马尔可夫链及其平稳分布 ……………………………… 230
5.2.3 马尔可夫信源及其熵的计算 ……………………………… 239
5.2.4 马尔可夫链的应用举例 …… 243
5.3 连续参数马尔可夫链——泊松过程 ……………………………… 247
5.4 隐马尔可夫模型及其应用 …… 252
5.4.1 HMM 的概念……………… 252
5.4.2 HMM 的三个基本问题及其算法 ……………………………… 254
5.4.3 HMM 的各种不同类型 …… 270
5.4.4 连续参数 HMM …………… 271
5.4.5 HMM 应用举例…………… 272
5.5 本章小结 ……………………… 281
附录A Cauchy-Schwartz 不等式 …… 282
附录B 常用 MATLAB 命令(函数)及其用法 …………………………… 283
附录C 傅里叶变换 ……………… 288
附录D 帕赛瓦定理 ……………… 291
参考文献………………………… 293

第 1 章　随机分析的数学基础

本课程的基础是概率论和数字信号处理等先修课程。本章回顾概率论的基本概念、理论和公式，并不追求数学上的严格性与逻辑性，重点介绍 MATLAB 在概率论中的应用，总结关于概率论仿真的基本规律，并以大量的实例展示 MATLAB 在解决实际问题中的应用。我们认为，尽管纯数学方法有着永恒的价值，它的解答是精确的，可以得到封闭形式的解，但是它有自身的局限性，很多实际问题采用纯数学的方法求解将遇到不可克服的困难。如果将纯数学方法和工程上的仿真技术结合起来，解决问题的范围将大大扩展，因此应该将数学方法和仿真结合起来，而不能片面强调一方面而轻视另一方面。在计算机技术日益发展的今天，利用计算机的强大运算能力解决实际问题显得尤其重要。本章突出了 MATLAB 软件在概率论的计算机仿真中的应用，要求读者熟悉 MATLAB 的工作环境和帮助系统的使用，熟悉常用的控制指令，能创建和操作数组，能对数据进行简单的可视化，逐步掌握常用数学函数的用法。此外，还要熟练掌握各种数字信号处理的基本命令的用法，如快速傅里叶变换(FFT)及其逆变换、基本滤波器设计等。

概率论是研究随机现象的科学，它起源于 17 世纪，最初是为赌博服务的，后来逐步发展为数学的一个分支。在概率论的发展史上，瑞士数学家贝努利由于提出著名的“大数定律”而被称为概率论的奠基人；前苏联数学家柯尔莫哥洛夫为概率论建立了公理化体系结构，奠定了严密的数学基础。随着科学的发展，概率论的应用比以往任何时候都更加广泛，例如从前的天气预报都是“硬”预报，即对未来某一天的天气作出唯一性判断，而现在的天气预报朝着“软”预报即概率预报的方向发展，它指出未来某天出现某种天气的概率为多少，以便人们作出适当的决策。基于概率的软判决在数字通信中运用广泛，例如在 Turbo 码中，基于概率的软判决可以作为边信息输入接收端，以提高对数据正确作出判决的概率。

1.1　随机现象基本概念

概率论的研究对象是随机现象，随机现象是和确定现象相对的，后者是在一定条件下必然发生的现象，如向上投掷一枚硬币，硬币必然要落下；前者是在一定条件下可以出现这种结果，也可以出现那种结果的现象，如向上投掷一枚硬币，硬币落下时可能正面朝上，也可能反面朝上(正面和反面是相对的，假设数字为正面，则花为反面；反之亦然)。随着人类对自然和社会研究的深入，遇到的随机现象会越来越多，必然要引入概率论。例如，物理学中研究自由落体运动，硬币的下落距离和时间有确定的数学关系，如果要进一步深入研究，就会提出这样的问题：硬币落地后是出现正面还是反面?出现正面和反面的可能性各为多少?这就涉及概率论。

随机现象 —— 指在个别试验中呈现出不确定性，而在大量重复试验中又表现出统计规律性的现象。它的反面是确定现象，确定现象在一次试验中肯定发生或肯定不发生。

随机试验 —— 指具有以下几个特点的试验：

(1) 在一定条件下可以重复地进行；

(2) 结果不唯一，并且是明确的；

(3) 试验前不能确定出现哪种结果。

例如，投掷硬币观察出现正面还是反面，抛骰子观察出现点数，这些都是随机试验。随机试验可以大量重复进行，并且表现出一定的统计规律性。我们通过随机试验来研究随机现象。

样本空间和样本点 —— 随机试验所有可能结果组成的集合称为**样本空间**，样本空间一般用 Ω 表示。样本空间的每个元素称为**样本点**。例如，在投掷硬币的随机试验中，样本空间为{正面，反面}，样本点为“正面”和“反面”。

随机事件和基本事件 —— 可能发生也可能不发生的事件叫做**随机事件**，一般用大写字母 A、B 等表示。随机事件是样本空间的一个子集。由样本空间的单个样本点构成的事件称为**基本事件**，基本事件是不能再细分的事件。例如，在 0，1，2，…，9 这十个数字中任取一数，那么“取得的是奇数”和“取得的是偶数”都是随机事件，而“取得的数为 1”，“取得的数为 2”，……，“取得的数为 9”都是基本事件。

必然事件 —— 在一定条件下一定会发生的事件。由于样本空间 Ω 包含了所有的基本事件，它在一次试验中一定发生，所以必然事件也可以用 Ω 表示。Ω 是表示样本空间还是必然事件可由上下文判断。

不可能事件 —— 在一定条件下一定不会发生的事件，一般记为空集 $\varnothing$。

1.1.1 随机事件的关系及运算

随机事件之间具有一定的关系，它们通过一定的运算，可以构成其他事件。

包含关系 —— 若事件 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，记作 $A \subset B$ 或 $B \supset A$。例如，“取得的数为 1”发生，那么“取得的是奇数”一定发生，后者包含前者。

相等关系 —— 若事件 A 和 B 互相包含，则称事件 A 和 B **相等**(等价)，记为 $A = B$。

和事件 —— 表示事件 A 和 B 至少有一个发生，记作 $A \cup B$。和事件的定义可以推广到多个事件的情况。

积事件 —— 表示事件 A 和 B 同时发生，记作 $A \cap B$，有时简记为 AB。积事件的定义同样可以推广到多个事件的情况。

差事件 —— 事件 A 发生而事件 B 不发生，记为 $A - B$，也可以简记为 $A\overline{B}$。

互斥事件 —— 不能同时发生的事件称为**互斥事件**。若 A 和 B 是互斥事件，则它们的积事件是不可能事件，即 $A \cap B = \varnothing$。互斥事件又称为互不相容事件。

对立事件 —— 在一次试验中不能同时发生并且必然有一个发生的两个事件称为**对立事件**。事件 A 的对立事件记为 $\overline{A}$，显然有 $A \cap \overline{A} = \varnothing$ 和 $A \cup \overline{A} = \Omega$。对立事件和互斥事件的关系是：互斥事件不一定为对立事件，但对立事件一定为互斥事件。

事件之间的关系及运算可以通过**维恩图**(Venn Diagram，又称文氏图) 形象表示，如图 1.1 所示。维恩图是维恩于 1880 年在《论命题和推理的图表化和机械化表现》一文中首次采用固定位置的交叉环形式，用封闭曲线(内部区域) 表示集合及其关系的图形，它具有直观的

优点，在表示事件或集合的关系及运算中得到广泛的应用。

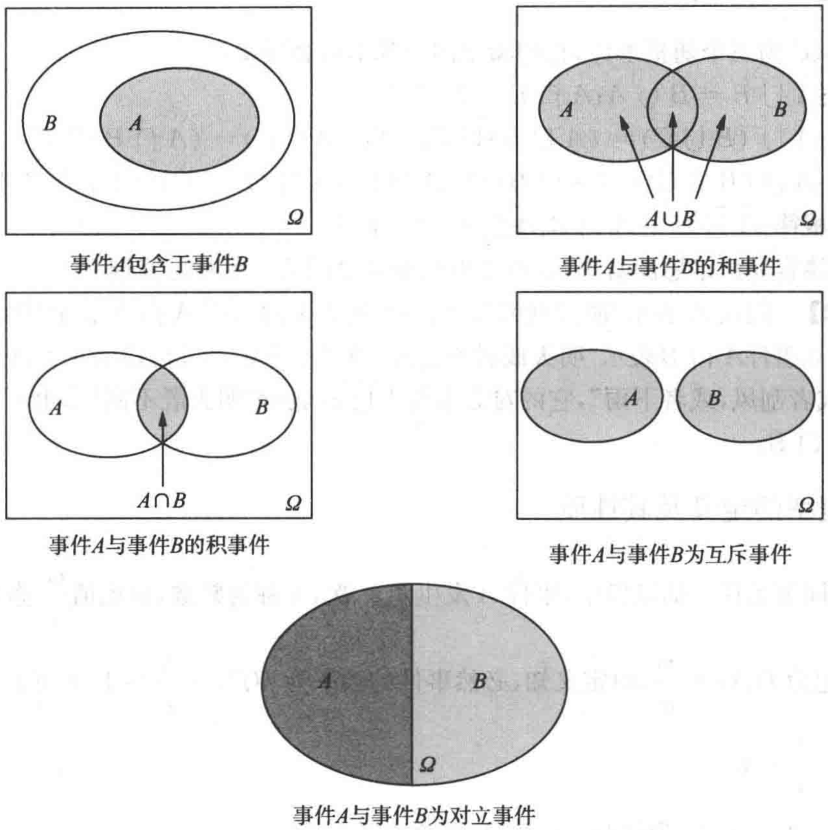

图 1.1　用维恩图表示事件之间的关系及运算

【例 1.1】　如图1.2所示的桥式电路包含5个电阻，每个电阻正常工作分别记为$A_i(i=1,2,3,4,5)$，每个电阻的工作是独立的(事件独立的含义见 1.1.7 节)。将“整个桥式电路正常工作”这件事记为 A，根据物理学知识，A 可以分解为两个互斥事件之和，即 $A=A\overline{A}_3+AA_3$，而 $A\overline{A}_3$ 又可以表示为 $\overline{A}_3\cap(A_1A_2\cup A_4A_5)$，它的含义是当 A_3 断路时，要使得整个电路正常工作，必须保证 1—2 串联支路和 4—5 串联支路中至少有一个是正常的；AA_3 又可以表示为 $A_3\cap(A_1\cup A_4)\cap(A_2\cup A_5)$，它的含义是当 A_3 正常时，要使得整个电路正常工作，必须保证 1—4 并联支路中至少有一个电阻是正常的，并且 2—5 并联支路中至少有一个电阻是正常的。综合上述，A 可以表示为：

$$A=[\overline{A}_3\cap(A_1A_2\cup A_4A_5)]\cup[A_3\cap(A_1\cup A_4)\cap(A_2\cup A_5)]$$

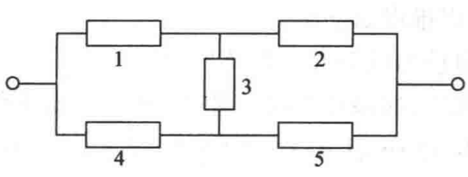

图 1.2　桥式电路

1.1.2 随机事件的运算律

设 A、B、C 为三个随机事件，它们满足四个基本运算规律：

交换律：$A\cup B=B\cup A$；$A\cap B=B\cap A$

结合律：$A\cup(B\cup C)=(A\cup B)\cup C$；$A\cap(B\cap C)=(A\cap B)\cap C$

分配率：$A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$；$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$

德·摩根律：$\overline{A\cup B}=\overline{A}\cap\overline{B}$；$\overline{A\cap B}=\overline{A}\cup\overline{B}$

以上规律容易由维恩图看出，也可以根据事件之间的关系及运算推出。

【例 1.2】 假设 A 表示“明天刮风”，B 表示“明天下雨”，则 $A\cap B$ 表示“明天刮风且下雨”，它的对立事件 $\overline{A\cap B}$ 表示“明天或者不刮风，或者不下雨”，又可以表示为 $\overline{A}\cup\overline{B}$。$A\cup B$ 表示“明天或者刮风，或者下雨”，它的对立事件 $\overline{A\cup B}$ 表示“明天既不刮风，也不下雨”，又可以表示为 $\overline{A}\cap\overline{B}$。

1.1.3 概率的定义及其性质

在 n 次同等条件下的试验中，事件 A 发生了 m 次，m 称为**频数**，而比值 $\frac{m}{n}$ 称为事件 A 发生的**频率**，记为 $f(A)=\frac{m}{n}$。由定义知，必然事件的频率为 $f(\Omega)=\frac{n}{n}=1$，不可能事件的频率为 $f(\varnothing)=\frac{0}{n}=0$。

如果 $A_1,A_2,\cdots,A_n$ 为两两互斥事件，那么有

$$f(A_1\cup A_2\cup\cdots\cup A_n)=f(A_1)+f(A_2)+\cdots+f(A_n)$$

当试验次数 n 增大时，事件 A 发生的频率 $\frac{m}{n}$ 表现出一定的稳定性，从频率的稳定性出发可以引入概率的概念。

定义 1.1.1 设随机事件为 A，对 A 赋予一个实数 $P(A)$，满足三个性质：

(1) 非负性，即 $P(A)\geqslant 0$；

(2) 规范性，即 $P(\Omega)=1$；

(3) 可列可加性，即对任意不同的 i 和 j，随机事件 A_i 和 A_j 两两互斥，则有

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

则称实数 $P(A)$ 为事件 A 的**概率**。

从概率的定义出发，可以推出如下性质：

性质 1.1.1 不可能事件的概率为 0，即 $P(\varnothing)=0$。

性质 1.1.2 有限可加性：如果对任意不同的 i 和 j，随机事件 A_i 和 A_j 两两互斥，则有

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

性质 1.1.3 若事件 B 包含事件 A，则有：

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

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

性质 1.1.4　对于任意事件 A，有 $0 \leqslant P(A) \leqslant 1$。

性质 1.1.5　互逆事件概率之间的关系：$P(\overline{A}) = 1 - P(A)$。

性质 1.1.6　多除少补原理：设 $A_1, A_2, \cdots, A_n$ 为任意 n 个事件，则有如下概率关系式

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \leqslant i < j \leqslant n} P(A_i A_j) + \sum_{1 \leqslant i < j < k \leqslant n} P(A_i A_j A_k) + \cdots + (-1)^{n-1} P(A_1 A_2 \cdots A_n) \tag{1.1}$$

1.1.4　古典概型及其蒙特·卡诺模拟

古典概型又称等可能概型，它具有如下特点：

(1) 样本空间只包含有限个样本点；

(2) 每个样本点发生的概率相等。

古典概型是概率论发展史上早期研究的对象。

设古典概型中样本空间为 Ω，它包含的基本事件总数为 n，而随机事件 A 的基本事件数为 k，则事件 A 发生的概率为

$$P(A) = \frac{k}{n}$$

实际问题中，样本空间的大小一般较易计算，但是某一事件所包含的基本事件数却很难计算。这时要计算事件的概率，可以利用计算机的强大运算能力进行大量的试验，用事件发生的频率代替概率，其基本思想即蒙特·卡诺方法，它的理论根据即后面讲到的贝努利大数定理。理论上，随机事件发生的概率为 p，那么一次试验中事件发生的次数(即期望值)为 p，n 次独立试验中事件发生的次数为 $m = np$，于是事件的概率即为 $p = m/n$，m 为试验中某事件发生的总次数，n 为试验的总次数。实践表明，随着试验次数的增加，事件发生的频率 m/n 确实趋于一个稳定的值，即事件的概率。下面用一个例子进行说明。

【例 1.3】　有高矮各不相同的 100 名同学随机地排成一个 10×10 的方阵。每行取最高的一个同学，一共 10 个高个子，记为集合 T；每列取最矮的一个同学，一共 10 个矮个子，记为集合 S。问题：(1) T 中最矮的同学(记为 $T[S]$)和 S 中最高的同学(记为 $S[T]$)相比谁更高？(2) 如果 $T[S]$ 和 $S[T]$ 一样高(即为同一个人)，求其概率为多少。

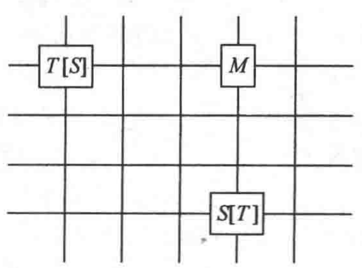

图 1.3　$T[S]$ 和 $S[T]$ 不同行也不同列的情况

解： 对第一个问题的回答可以分四种情况讨论：

(1) 当 $T[S]$ 和 $S[T]$ 既不在同一行，也不在同一列时，如图 1.3 所示。设 $T[S]$ 所在的行与 $S[T]$ 所在的列交于 M。因为 $T[S]$ 是一行中最高的，所以 $T[S]$ 比 M 高；又因为 $S[T]$ 是一列中最矮的，所以 M 比 $S[T]$ 高。在这种情况下，$T[S]$ 一定比 $S[T]$ 高，即高个子中最矮的比矮个子中最高的要高。

(2) 当 $T[S]$ 与 $S[T]$ 在同一行时，显然 $T[S]$ 比 $S[T]$ 高，因为 $T[S]$ 是一行中最高的。

(3) 当 $T[S]$ 与 $S[T]$ 在同一列时，显然 $T[S]$ 也比 $S[T]$ 高，因为 $S[T]$ 是一列中最矮的。

(4) 当 $T[S]$ 和 $S[T]$ 为同一人时，两者一样高。

综合上述四种情况，只要 $T[S]$ 和 $S[T]$ 不是同一个人，那么 $T[S]$ 一定比 $S[T]$ 高。但是存在一种情况，即 $T[S]$ 和 $S[T]$ 是同一个人，这一点可以很简单地加以验证。直观上，$T[S]$ 和 $S[T]$ 为同一个人的概率很小。

下面回答第二个问题，即 $T[S]$ 和 $S[T]$ 一样高的概率是多少。

记{$T[S]$ 和 $S[T]$ 为同一人}为事件 A，显然 A 为随机事件，其概率不容易得出封闭形式的解（精确解）。100 个同学的所有排列组合数为 100!（不考虑通过方阵的转置得到相同的排列），$100! \approx 9.33 \times 10^{157}$，这是一个天文数字。$T[S]$ 和 $S[T]$ 为同一个人的排列情况很难罗列，所以事件 A 包含的基本事件总数难以计算。显然，基本事件即每一种排列都是等可能的，总数为 100!，而事件 A 包含的基本事件总数不容易求得，可以通过大量试验的方法，用频率代替概率。只要试验次数足够大，那么用频率代替概率就会足够精确。由于 100 个同学高矮各不相同，可以按照从矮到高的顺序排序，分别赋予整数 1,2,…,100。进行 10 000 000 次试验，统计 $T[S]$ 和 $S[T]$ 为同一个人的频率。试验的 MATLAB 代码如下：

```
clc, clear all;
tic
jj = 0;
N = 10000000;                          % 试验次数为 10000000
for ii = 1:N
      A = randperm(100);               % 100 个同学随机排成一行
      B = reshape(A, 10, 10);          % 将 100 个同学排成 10 × 10 的方阵
      C1 = max(B');                    % 每一行找出一个最高的同学,组成一个向量
      D1 = min(C1);                    % 每一列找出一个最矮的同学,组成一个向量
      C2 = min(B);                     % 找出高个子集合中最矮的同学
      D2 = max(C2);                    % 找出矮个子集合中最高的同学
      if (D1 - D2) <= eps
            jj = jj + 1;               % 如果 C2 与 D2 为同一个人,计数器加 1
      end
end
rate = jj/N                            % 计算频率
time = toc
```

在 Intel Celeron CPU G1820(2.70 GHz) 上运行结果为：

```
rate =
  1.0990e - 004
time =
  169.1524
```

因为试验次数非常大(一千万次),根据贝努利大数定律,用频率代替概率可以得到足够的精度,大致认为 $T[S]$ 和 $S[T]$ 为同一个人的概率为 0.011%,而前者比后者高的概率约为 99.989%,即在任一次试验中,几乎可以肯定地说,高个子中的矮个子比矮个子中的高个子要高。同时可以看出,10 000 000 次重复试验所花的机器时间约为 169 秒,即接近 3 分钟,充分体现了计算机运行速度快的优点。

编程要求及说明:

(1) 掌握 tic 和 toc 的用法。tic 和 toc 成对使用,统计某一段程序的运行时间,格式如下:

```
tic
operations
toc
```

tic 为 operations 运行开始计时,toc 根据 operations 结束时间,计算出 operations 的运行时间并赋给 toc。

(2) 掌握数值数组的创建和操作,能够用数组模拟并解决实际问题。本例用 randperm(100) 模拟 100 名学生的一个随机排列,然后用 reshape(A, 10, 10) 将 100 名学生排成 10×10 的方阵,并用 max 和 min 命令找出所有行(列)的最大或最小值,组成一个向量,再将 max 中的最小值与 min 中的最大值进行比较,如果相等,计数器加 1。

(3) 掌握 max 和 min 命令的用法。当 A 为 $n \times m$ 矩阵时,max(A) 得到每一列的最大值并组成一个行向量,如运行以下代码:

```
A = randperm(16);
B = reshape(A, 4, 4)
C = max(B)
```

运行结果为:

```
B =
    8    11     6     7
   15    14     3     4
    5    10    12     2
    9    13    16     1
C =
   15    14    16     7
```

要得到矩阵每一行的最大值并组成一个向量,需要对矩阵求转置。

(4) 掌握 eps 的用法。eps 表示大于 1 的最大双精度数值与 1 之间的差值,即 eps $= 2^{-52}$。由

于计算机采用浮点法表示整数，当两个整数差值的绝对值不大于 eps 时，才能认为它们是相等的。注意，eps 不是计算机能表示的最小正数，事实上，计算机能表示的最小正数比 eps 要小得多。

如果未作说明，本书所有程序都是在 Intel Celeron CPU G1820 个人计算机上运行的。

1.1.5 条件概率与全概率公式

定义 1.1.2 考虑当事件 A 发生时事件 B 发生的**条件概率**，记为 $P(B \mid A)$，它定义为

$$P(B \mid A) = \frac{P(AB)}{P(A)} \tag{1.2}$$

其中 $P(A) > 0$。

由式(1.2) 容易得出，当 $P(A) > 0$ 时

$$P(AB) = P(B \mid A) \cdot P(A) \tag{1.3}$$

式(1.3) 称为**乘法公式**，它很容易推广到多个事件积事件的情况。

下面介绍划分的概念。设样本空间为 Ω，$B_1, B_2, \cdots, B_n$ 为一组事件并且两两互斥，它们的和事件为样本空间 Ω，则称 $B_1, B_2, \cdots, B_n$ 为样本空间 Ω 的一个划分。

对于样本空间 Ω 的任何一个事件 A，它的概率可以由划分 $B_1, B_2, \cdots, B_n$ 的条件概率得出，即

$$P(A) = P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \cdots + P(A \mid B_n)P(B_n) \tag{1.4}$$

式(1.4) 称为**全概率公式**，在概率论中具有很重要的地位。

【例 1.4】 某种商品的商标为“MAXAM”，其中有两个字母脱落，有人捡起随意放回。求放回后仍为“MAXAM” 的概率。

解法一：用全概率公式。“MAXAM” 由三个不同的字母组成：M、A 和 X，其中 M 和 A 各两个，X 只有一个。当脱落的两个字母都是 M 或 A 时，随意放回后将以概率 1 还原；当脱落的两个字母为不同的字母时，还原的概率为 1/2。记 $A = \{$脱落的两个字母不同$\}$，$B = \{$放回后仍为“MAXAM”$\}$，根据全概率公式有

$$\begin{aligned} P\{B\} &= P\{A\} \cdot P\{B \mid A\} + P\{\overline{A}\} \cdot P\{B \mid \overline{A}\} \\ &= \left(1 - \frac{C_2^2 + C_2^2}{C_5^2}\right) \cdot \frac{1}{2} + \frac{C_2^2 + C_2^2}{C_5^2} \cdot 1 \\ &= \frac{3}{5} \end{aligned}$$

解法二：用 MATLAB 进行试验。建立字符串数组“MAXAM”，随机地选择任意两个字母，对选择的两个字母进行随机重排，如果原排列不发生改变，计数器加 1，否则计数器不变，完成一次试验。重复进行大量(这里为 10 000 000 次) 试验，统计原排列保持不变的频率。试验的 MATLAB 代码如下：

```
clc; clear;
tic
n = 0;                                        % 计数器初始化为 0
N = 10000000;                                 % 试验次数为 10000000
for ii = 1:N
     b = 'MAXAM';                             % 将商标 MAXAM 组成一个字符串数组
     a = randperm(5);
     b = [b(a(1)), b(a(2))];                  % 模拟脱落的两个字母
     c = randperm(2);
     if (b(1) = = b(c(1)))&(b(2) = = b(c(2)))  % 模拟随机放回两个字母并判断是否正确放回
        n = n + 1;                            % 如果正确放回计数器加 1
     end
end
rate = n/N                                    % 计算正确放回的频率
time = toc
```

运行结果为：

```
rate =
  0.5999
time =
  113.9250
```

试验结果 0.5999 和理论值 0.6 非常接近，运行时间接近 2 分钟。比较以上两种解法，可以看出试验方法的好处，即只要满足两个条件：① 试验次数充分大；② 正确地模拟随机事件，那么试验得到的概率值可以和理论上概率的"真实值" 非常接近。

编程要求及说明：

(1) 掌握字符串数组的创建。将字符串的字符放在单引号内，就创建了一个字符串数组，如 b = 'MAXAM' 就创建了字符串数组 MAXAM。对其中每一个字符的标识通过用数组名加上表示字符的位置的整数即可，如 b(4) 标识第四个字符 'A'。

(2) 正确模拟脱落的两个字母。这里用一个整数重排命令 randperm(5)，然后任选两个字符，如第一个字符和第二个字符，就可以模拟任意脱落的两个字母。再用一个整数重排命令 randperm(2)，就可以模拟按照相同的顺序及相反的顺序放回字母。

(3) 掌握逻辑与(&) 的用法。

采用试验法求随机事件概率的理论依据是大数定律，它在概率论和数理统计的理论和应用中都十分重要。如前所述，概率论是研究随机现象的统计规律性的科学，但是统计规律性只有在相同条件下的大量试验中才能体现出来。

1.1.6　贝努利大数定律及其应用

在大量的重复试验中，随机事件发生的频率具有稳定性，在某种意义下逼近一个常数，即概率。大数定律是描述频率稳定性规律的定律，正是有了大数定律，概率这一概念才有了

客观意义。从某种意义上来说，概率论的历史是从大数定律出现后才真正开始的。大数定律有多个，下面介绍频率稳定于概率的贝努利大数定律。

定理 1.1.1 贝努利大数定律

设 $\xi_n(n=1,2,\cdots)$ 为相互独立同分布随机序列，且 $P\{\xi_n=1\}=p, P\{\xi_n=0\}=q$，其中 $q=1-p,\ 0<p<1$，则 $\{\xi_n\}$ 服从大数定律，即若令 $\bar{\xi}_n=\dfrac{1}{n}\sum\limits_{i=1}^{n}\xi_i$，则有

$$\lim_{n\to\infty}\bar{\xi}_n=p \quad (P)$$

即随机序列 $\bar{\xi}_n$ 依概率收敛于常数 p。

由于贝努利大数定律的证明要用到切比雪夫不等式，下面先叙述切比雪夫不等式并证明。

定理 1.1.2 切比雪夫不等式

若随机变量 ξ 的方差 $D(\xi)$ 存在，则对任意的 $\varepsilon>0$，有以下不等式成立：

$$P\{|\xi-E(\xi)|\geqslant\varepsilon\}\leqslant\frac{D(\xi)}{\varepsilon^2}$$

证明： 设 ξ 的分布函数为 $F(x)$，则

$$\begin{aligned}
P\{|\xi-E(\xi)|\geqslant\varepsilon\}&=\int_{|x-E(\xi)|\geqslant\varepsilon}\mathrm{d}F(x)\\
&\leqslant\int_{|x-E(\xi)|\geqslant\varepsilon}\frac{[x-E(\xi)]^2}{\varepsilon^2}\mathrm{d}F(x)\\
&\leqslant\frac{1}{\varepsilon^2}\int_{-\infty}^{\infty}[x-E(\xi)]^2\mathrm{d}F(x)\\
&=\frac{D(\xi)}{\varepsilon^2}
\end{aligned}$$

贝努利大数定律的证明：

因为 $D(\bar{\xi}_n)=\dfrac{1}{n^2}\sum\limits_{i=1}^{n}D(\xi_i)=\dfrac{1}{n^2}npq=\dfrac{1}{n}pq$，由切比雪夫不等式得，对任意 ε，有

$$0\leqslant P\{|\bar{\xi}_n-p|\geqslant\varepsilon\}\leqslant\frac{pq}{n\varepsilon^2}$$

因此得

$$\lim_{n\to\infty}P\{|\bar{\xi}_n-p|\geqslant\varepsilon\}=0$$

贝努利大数定律以严格的数学形式表达了频率的稳定性，即当独立试验次数很大时，事件发生的频率与概率发生较大偏差的可能性很小。用于实际中，即只要试验次数足够大，那么就可以用事件发生的频率代替概率。试验法在概率的理论值不易得出或根本不能给出封闭解的条件下尤其有益，如例 1.3 所示，因为试验的“近似值”给我们提供了可能性大小的概念，它给理论值的推导提供了一条线索或启示。

下面给出另一个计算概率的例子，这里概率值虽然从理论上可以求得精确值，由于计算