

I&C

信息与通信工程研究生系列教材

信息与通信工程中的 随机过程 (第二版)

陈 明 编著



 科学出版社
www.sciencep.com

内 容 简 介

本书系统地介绍了从事信息与通信工程领域的科学研究及工程设计所必需的随机数学基础, 内容包括: 随机现象的数学建模和信息与通信工程领域所涉及的随机现象; 概率空间和随机对象; 随机数学分析; 随机信号与线性系统; 信号的统计推断; Markov 链; 排队论; 随机对象的计算机模拟等。本书在注重概念的数学严密性和知识体系的逻辑性基础上, 结合大量信息与通信工程中的问题和范例, 深入浅出地介绍了信息与通信工程领域所必需的随机数学基础。

本书可以作为高等院校信息与通信工程一级学科下各专业的研究生、高年级本科生教材, 也可作为信息与通信工程领域的科研人员及工程技术人员的参考书。

图书在版编目(CIP)数据

信息与通信工程中的随机过程/陈明 编著. —2 版. 北京: 科学出版社, 2005
(信息与通信工程研究生系列教材)

ISBN 7-03-015970-5

I. 信… II. 陈… III. ①信息技术-随机过程-研究生-教材 ②通信工程-随机过程-研究生-教材 IV. ①G202 ②TN911.1

中国版本图书馆 CIP 数据核字 (2005) 第 083069 号

责任编辑: 匡 敏 贾瑞娜/责任校对: 包志虹

责任印制: 钱玉芬/封面设计: 陈 敬

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>

双青印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2001 年 8 月东南大学出版社第一版

2005 年 9 月第 二 版 开本: B5 (720×1000)

2005 年 9 月第一次印刷 印张: 24 1/4

印数: 3 001—6 000 字数: 465 000

定价: 35.00 元

(如有印装质量问题, 我社负责调换<路通>)

前 言

在被称为信息时代的今天,信息与通信技术的发展日新月异。人们所处理的信息种类越来越多,处理手段日益更新,所处理的信息量也越来越大。此外,传输信息的系统——通信系统也不断地更新换代,系统容量越来越大,信息传输速率越来越高,传输的信息种类也越来越丰富。通信技术的发展,正朝着任何人在任何时间和任何地点实现与任何人进行任何种类的信息交换这一“个人通信”的最高目标迈进。

众所周知,信息与通信技术的发展首先依赖于信息与通信理论不断发展。由于信息与通信工程的研究对象涉及大量随机现象,所以用以描述随机现象的概率论、随机过程、数理统计等随机数学理论成了必不可少的理论工具。因此,对于从事这一领域科学研究的人来说,掌握足够的随机数学理论是进行科研的前提条件。

从1996年秋季开始,笔者为东南大学无线电工程系信息与通信工程学科的硕士研究生开设了学位基础课程“随机过程”。这门课的教学目的是使研究生掌握信息与通信领域的科研所必需的随机过程基础理论,为后续课程的学习和将来的科研奠定随机数学的理论基础。

在教学过程中,笔者发现国内关于随机过程方面的教材虽然不下数十种,但在内容的选择上各有所重、深浅不一,缺乏针对信息与通信工程这一学科的专门教材。这些教材有些是针对其他具体学科编写的,有些是针对多学科编写的,因而在内容的选择上,有些在信息与通信工程领域中不常用的知识也占据了很大的篇幅,有些知识为信息与通信工程领域所必需,但却没有提及。

事实上,信息与通信工程领域所需要的随机过程基础理论知识具有鲜明的学科特点。即便在这一学科内部,随着信息与通信技术的发展,所需的科研预备知识的侧重点也发生了一些深刻的变化。例如,信息数字化的发展趋势,使得离散时间随机信号的处理更为常见,关于连续时间随机信号的一些复杂处理已经变得不太常用;网络技术的发展,使得 Markov 链及排队论的知识显得日益重要;计算机仿真成了信息与通信工程领域的重要实验手段之一,这使得随机变量、随机过程的计算机模拟成为必要。

所以,需要针对信息与通信工程学科的知识特点和发展趋势,既照顾知识体系的完整性,又考虑到实际的学科应用背景,对内容进行取舍,并将随机数学的基础理论和信息与通信工程学科的应用实例进行恰当关联,编写一本适用于信息与通信工程学科的《随机过程》研究生教材。这样,可以使得硕士研究生在学业繁重的

就读期间,具有针对性地掌握信息与通信工程领域所必需的随机数学预备知识,为将来后续课程的学习及科研奠定理论基础。

本书正是基于上述背景和出发点编写而成的。在内容的选择上,参考了国内外各种随机过程教材的内容,并结合信息与通信工程的学科特点,根据笔者在东南大学移动通信国家重点实验室从事信息与通信工程领域的科研经验,选择了一些该领域最重要的随机过程基础理论知识作为本书的内容。

具体来说,本书的内容按知识的逻辑结构 and 应用背景可以分为以下几个部分(见图 1):

- 随机数学的方法论 (第 1 章)。
- 概率论和随机过程的基本概念和基础理论 (第 2、3 章)。
- 随机信号处理 (第 4、5 章)。
- 应用于系统分析和网络流量分析的 Markov 链和排队论 (第 6、7 章)。
- 随机对象的计算机仿真 (第 8 章)。

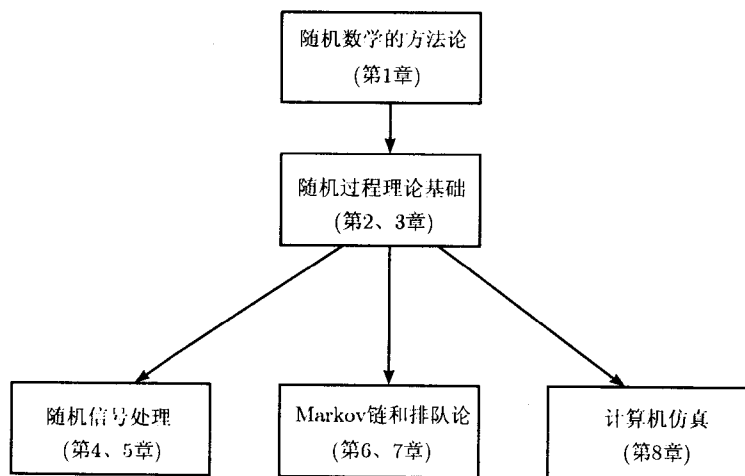


图 1 全书知识的逻辑结构及在信息与通信工程中的应用背景

在本书的编写中,假设读者具有如下预备知识:高等数学中的微积分、常微分方程、线性代数以及信号与系统的初步知识。此外,本书还需要一些泛函分析的基础性概念和广义函数及其导数的概念,这些预备知识已经在附录 A 中给出。

第 1 章介绍了随机现象形成的原因及其数学建模的基本思想,旨在从科学方法论的角度,阐明随机现象的数学建模方法所体现的科学思想。本章还介绍了信息与通信工程中常见的随机对象,旨在阐明随机对象是信息与通信工程领域的常见研究对象。本章内容是全书的纲领,揭示的是随机数学理论中所蕴涵的科学思想,是非常重要的部分,授课教师要花一定时间讲透。

第2、3章是全书的理论基础,这两章内容是以介绍基本的随机数学概念和理论为主线,同时穿插了大量信息与通信工程领域的应用范例。这两章的基本随机数学概念包括概率空间、随机变量、随机向量、随机过程、随机对象的函数、随机序列的收敛、随机过程的微积分、随机过程的正交分解、随机过程的线性变换及随机过程的遍历性等。

第2章详细介绍了概率空间和随机对象的概念。概率空间和随机对象是对随机现象的两种基本建模方法,该章在介绍概率空间的基础上,尝试从一个统一的框架介绍三类最本质的随机对象,即随机变量、随机向量、随机过程。事实上,这三类随机对象只是样本空间维数的不同,它们的样本空间分别是 $\mathbb{R}$ 、 $\mathbb{R}^n$ 、 $\mathbb{R}^\infty$ 。这样一个框架很好地揭示了随机变量、随机向量、随机过程这三类随机对象之间的联系和差别。从这个框架,读者可以很清楚地看到,随机过程实际上是无穷维的随机向量,是随机向量概念的一个推广;此外,有限维的随机变量和随机向量是描述随机过程的必备基础。

值得一提的是,“随机对象”概念的引进以及从“无穷维随机向量”的角度引进随机过程的概念,是本书的两个重要尝试,是笔者在教学过程中摸索出来的介绍随机变量、随机向量、随机过程这三个核心概念的一种直观的、易于理解的方法。这种方法,使得这三个概念牢牢地捆绑在一起,既起到了巩固概率论基础知识的作用,也起到了引进随机过程概念的作用。

该章还介绍了描述这些随机对象的指标,如各种概率函数和数字特征等。该章随机过程部分,还介绍了各种常见的随机过程,如正态随机过程、和过程、Wiener过程、Poisson及其导出过程、更新过程等。通过第2章的学习,读者可以迅速具备概率论的预备知识,并达到对随机过程概念的深刻理解。

第3章在介绍随机对象函数的基础上,介绍了随机变量序列的各种收敛性;在均方收敛的基础上,介绍了均方微积分、二阶矩过程的正交分解、遍历性等概念。这些内容在数学领域内属于“数学分析”的范畴,由于分析的对象是随机对象,所以称之为“随机数学分析”。这一章内容是后续章节内容的理论基础。

第4、5章以信息与通信工程中的“随机信号处理”为应用背景,介绍了随机信号的功率谱密度及带宽、随机信号通过线性系统和信号的统计推断这几个方面的基础理论知识。

第4章介绍了随机信号的功率谱密度和带宽的概念;在此基础上介绍了随机信号通过线性系统的二阶统计特性和概率特性;此外,还介绍带限和带通信号的一些常见性质。

第5章介绍了信号的统计推断,包括信号检测、参数估计和信号的波形估计(滤波)。在检测问题中,介绍了常见的各种检测准则,如Bayes准则、最小化最大风险准则、最大似然准则、最小错误概率准则等;在参数估计中,也介绍了各种估

计准则,如 Bayes 准则、最小化最大风险准则、最大似然准则、线性最小均方误差准则和最小二乘准则等;在滤波问题中,介绍了匹配滤波、白化滤波、Wiener 滤波和 Kalman 滤波。

第 6、7 章以信息与通信工程中的“系统分析和网络流量分析”为应用背景,介绍了 Markov 链及其排队论方面的基础理论知识。

第 6 章在介绍 Markov 过程的基础上,分别介绍了离散时间 Markov 链的状态方程、状态分类、应用举例,以及连续时间 Markov 链的状态停留时间、状态微分方程和应用举例等,嵌入 Markov 链概念的引进,使得可以用离散时间 Markov 链来分析连续时间 Markov 链的极限性质。

第 7 章排队论初步,实际上是第 6 章 Markov 链的理论应用举例,在介绍排队系统基本要素和 Little 定理的基础上,介绍了 M/M 型、M/G/1 和 G/M/1 排队系统。M/M 型排队系统可以用一种特殊的连续时间 Markov 链——生灭过程来建模,M/G/1 和 G/M/1 排队系统可以用嵌入 Markov 链来进行分析。一个排队系统,不仅要对待顾客数、系统顾客数等顾客参量的分布进行分析,还要对等待时间、系统时间、忙期、闲期等时间参量的分布进行分析。

第 8 章以信息与通信工程中的仿真实验为应用背景,介绍了随机过程的计算机模拟方法,也即在计算机仿真实验中,如何生成一个具有指定概率分布或其他统计特性的随机对象。

借助于多媒体的教学手段,本书的内容基本上可以在 60 学时内讲授完毕。建议在使用本书作为教材时,注重基本概念的讲解,让学生对基本概念产生确定的理解,而将大部分非概念性的例子主要交给学生自学,来达到巩固基本概念的目的。“概念是灵魂”,如果学生理解了概念就基本上掌握了这些知识所蕴涵的思想方法,将来在科研中用随机数学的方法对所遇到的问题建立恰当的数学模型就没有困难。否则,对概念一知半解,仅会解一些习题,终究是“影像学习”,根本无法将科研中遇到的实际问题用随机数学的“语言”进行描述,并建立数学模型。

为了激发学生的学习兴趣,本书每章最后给出了一些研究型习题,可以培养学生良好的分析问题、解决问题的能力。

本书的第一版曾于 2001 年 8 月在东南大学出版社出版。此次第二版,虽然知识要点和第一版大致相同,但对各知识要点的内容进行了较大的补充,例题的选用和概念的描述上也做了较大的完善。此外,还对这些知识要点的章节安排做了较大的改动,使得这些知识要点按照一个更加合理清晰的框架呈现在读者面前。

感谢芬兰 NOKIA 公司为本书的出版提供了资助,感谢芬兰 NOKIA 公司的高级研究员 Jorma Lilleberg 和王海峰研究员使得该项资助成为可能,他们也为本书的编写提供了很好的参考资料。

感谢东南大学移动通信国家重点实验室的程时昕教授、尤肖虎教授、赵春明教

授、高西奇教授、曹秀英教授、沈连丰教授、毕光国教授和所有同事、研究生，整个移动通信国家重点实验室不畏劳苦、勇于开拓、互帮互助、团结奋进的精神给了我工作的乐趣和动力。在移动通信国家重点实验室即将成为通信技术国家实验室之际，衷心地祝愿这个团结向上、充满朝气与活力的研究团队在未来的岁月里创造更多的辉煌，为祖国通信事业的发展贡献一份光和热！

感谢东南大学无线电工程系历届学习该课程的硕士研究生，他们对本书提出了许多宝贵的建议；感谢研究生商秋帮助校对了本书的全部草稿；感谢科学出版社高教分社的匡敏女士为本书再版付出的所有努力。

本书虽经多次修改，但由于笔者工作繁忙，时间仓促，书中难免存在不少疏漏和错误之处，恳请读者不吝斧正。任何对本书的指正和建议，可以发送到电子邮箱：chenming@seu.edu.cn，笔者深表感谢！

愿本书能够成为读者掌握信息与通信领域随机数学基础知识的桥梁！

陈 明

2005 年 6 月于南京

本书常用数学记号

$\mathbb{N}$	自然数集
$\mathbb{R}, \mathbb{R}^n$	实数集, n 维实数空间 (n 维欧氏空间)
$\mathbb{Z}, \mathbb{Z}^+, \mathbb{Z}^-$	整数集, 正整数集, 负整数集
$\mathbb{Q}, \mathbb{Q}^+, \mathbb{Q}^-$	有理数集, 正有理数集, 负有理数集
$\mathbb{T}$	连续或离散时间随机过程的时间指标集
j	虚数单位, $j^2 = -1$
π	圆周率
e	自然对数的底
S_X, S_X	随机变量 X 或者随机向量 $\mathbf{X}$ 的样本空间
$A, B, C, X, Y, Z, \dots$	随机变量
$a, b, c, x, y, z, \dots$	一维变量
$\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}, \dots$	随机向量
$\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$	多维向量
$X(t), Y(t), \dots$	连续时间随机过程
$x(t), y(t), \dots$	连续时间函数
$X[n], Y[n], \dots$	离散时间随机过程
$x[n], y[n], \dots$ 或 $x_n, y_n, \dots$	离散时间函数、数列
$P(x), P(\mathbf{x})$	随机变量 X 和随机向量 $\mathbf{X}$ 的概率质量函数
$f_X(x), f_{\mathbf{X}}(\mathbf{x})$	随机变量 X 和随机向量 $\mathbf{X}$ 的概率密度函数
$F_X(x), F_{\mathbf{X}}(\mathbf{x})$	随机变量 X 和随机向量 $\mathbf{X}$ 的概率分布函数
$G_X(z), G_{\mathbf{X}}(z)$	随机变量 X 和随机向量 $\mathbf{X}$ 的概率生成函数
$\Phi_X(\omega), \Phi_{\mathbf{X}}(\omega)$	随机变量 X 和随机向量 $\mathbf{X}$ 的概率特征函数
$P_{X Y}(x), f_{X Y}(x), F_{X Y}(x)$	条件概率质量 (密度、分布) 函数
$E\{X\}$	随机变量 X 的期望
m_X, ψ_X^2	随机变量 X 的均值、均方
$\sigma_X^2, \text{var}\{X\}$	随机变量 X 的方差
R_{XY}, C_{XY}	随机变量 X 和 Y 的相关矩、协方差
$\mathbf{R}_X, \mathbf{C}_X$	随机向量 $\mathbf{X}$ 相关矩阵、协方差矩阵
$\mathbf{R}_{XY}, \mathbf{C}_{XY}$	随机向量 $\mathbf{X}$ 和 $\mathbf{Y}$ 的互相关矩阵、互协方差矩阵
$H(X), H(\mathbf{X})$	随机变量 X 和随机向量 $\mathbf{X}$ 的熵

$m_X(t), \psi_X^2(t), \sigma_X^2(t)$	随机过程 X 的均值函数、均方函数和方差函数
$R_X(t_1, t_2), R_X[n_1, n_2]$	随机过程 X 的自相关函数
$R_X(\tau), R_X[n]$	宽平稳随机过程 X 的自相关函数
$C_X(\tau), C_X[n]$	宽平稳随机过程 X 的自协方差函数
$S_X(f)$	随机过程 X 的功率谱密度
$f(x) * g(x)$	连续函数 $f(x)$ 和 $g(x)$ 的卷积
$f[n] * g[n]$	离散函数 $f[n]$ 和 $g[n]$ 的卷积
Z^*	复数 Z 的共轭
$A^\dagger$	矩阵或者向量 A 的共轭转置
A^T	矩阵或向量 A 的转置
$\bar{A}$	集合 A 的补集
ms lim	均方极限
$\forall$	表示“对于所有的……”

目 录

前言

本书常用数学记号

第 1 章 绪论	1
1.1 自然界的随机现象	1
1.2 随机现象的统计规律	2
1.3 随机现象的数学建模	4
1.3.1 概率空间的建模方法	4
1.3.2 随机对象的建模方法	6
1.4 信息与通信工程中的随机对象	10
1.4.1 信息与通信工程的处理对象	10
1.4.2 信息与通信工程中的典型问题	13
1.4.3 信息与通信工程中的常见随机对象	16
1.5 本章概要和习题	18
第 2 章 概率空间和随机对象	21
2.1 概率空间	21
2.1.1 概率空间的定义	21
2.1.2 条件概率和事件的独立	26
2.2 随机变量	30
2.2.1 概率函数	30
2.2.2 数字特征	41
2.3 随机向量	48
2.3.1 概率函数	48
2.3.2 数字特征	53
2.3.3 条件概率函数和条件数字特征	59
2.4 随机过程	64
2.4.1 随机过程的定义	64
2.4.2 概率函数族	66
2.4.3 矩函数	70
2.4.4 常见随机过程	75

2.5	复随机对象	86
2.5.1	复随机变量	86
2.5.2	复随机向量	88
2.5.3	复随机过程	88
2.6	矩阵随机对象	90
2.6.1	随机矩阵	90
2.6.2	矩阵随机过程	91
2.7	本章概要和习题	92
第3章	随机数学分析	104
3.1	随机对象的函数	104
3.1.1	随机变量的函数	104
3.1.2	随机向量的函数	109
3.1.3	随机过程的函数	112
3.2	随机变量序列的收敛	116
3.2.1	均方收敛	116
3.2.2	其他常见收敛	118
3.3	均方微积分	123
3.3.1	均方连续	123
3.3.2	均方导数	124
3.3.3	均方积分	126
3.4	二阶矩过程的正交分解	128
3.4.1	二阶矩过程的正交分解	128
3.4.2	Fourier 正交分解	130
3.4.3	Karhunen-Loève 正交分解	134
3.5	二阶矩过程的线性变换	138
3.5.1	连续时间二阶矩过程的线性变换	138
3.5.2	离散时间二阶矩过程的线性变换	140
3.6	二阶矩过程的各态遍历性	141
3.7	本章概要和习题	145
第4章	随机信号与线性系统	151
4.1	随机信号的功率谱密度	151
4.1.1	连续时间随机信号的功率谱密度	151
4.1.2	离散时间随机信号的功率谱密度	156
4.2	随机信号的带宽	158

4.3 随机信号通过线性系统	162
4.3.1 连续时间线性系统	162
4.3.2 离散时间线性系统	173
4.4 带限和带通随机信号	180
4.4.1 带限随机信号	180
4.4.2 带通随机信号	184
4.5 本章概要和习题	190
第5章 信号的统计推断	196
5.1 信号的统计推断	196
5.1.1 统计推断	196
5.1.2 信号的统计推断	198
5.2 信号检测	198
5.2.1 信号检测模型	198
5.2.2 常见判决准则	200
5.2.3 应用举例	204
5.3 参数估计	208
5.3.1 参数估计模型	208
5.3.2 常见估计准则	209
5.3.3 应用举例	213
5.4 滤波(波形估计)	216
5.4.1 信号的波形估计	216
5.4.2 匹配滤波	217
5.4.3 白化滤波	219
5.4.4 Wiener 滤波	220
5.4.5 Kalman 滤波	228
5.5 本章概要和习题	234
第6章 Markov 链	241
6.1 Markov 过程	241
6.2 离散时间 Markov 链	244
6.2.1 状态方程	244
6.2.2 状态分类	248
6.2.3 应用举例	262
6.3 连续时间 Markov 链	266
6.3.1 状态停留时间	267

6.3.2	状态微分方程	268
6.3.3	应用举例	271
6.4	嵌入 Markov 链	280
6.5	本章概要和习题	281
第 7 章	排队论初步	287
7.1	排队系统的基本要素	287
7.2	Little 定理	289
7.3	M/M 型排队系统	291
7.3.1	顾客数分析	291
7.3.2	时间分析	299
7.4	M/G/1 排队系统	305
7.4.1	顾客数分析	305
7.4.2	时间分析	310
7.5	G/M/1 排队系统	314
7.5.1	顾客数分析	314
7.5.2	时间分析	317
7.6	本章概要和习题	317
第 8 章	随机对象的计算机模拟	321
8.1	均匀分布随机变量的模拟	321
8.2	具有给定分布随机变量的模拟	323
8.2.1	变换法	323
8.2.2	拒绝法	324
8.3	具有给定协相关矩阵的随机向量的模拟	327
8.4	具有给定功率谱密度的随机过程的模拟	329
8.5	移位寄存器生成的伪随机序列	330
8.5.1	线性移位寄存器	331
8.5.2	m 序列及其性质	333
8.6	Monte Carlo 方法	337
8.7	本章概要和习题	340
参考文献		342
附录 A	泛函分析基本概念	346
A.1	集合和映射	346
A.2	常见泛函空间	348
A.3	广义函数及其导数	351
A.4	习题	355

附录 B 常用数学分式	356
B.1 三角恒等式	356
B.2 不定积分	356
B.3 定积分	357
附录 C 常见随机变量	358
C.1 离散型随机变量	358
C.2 连续型随机变量	359
附录 D Fourier 变换表	362
索引	364

第1章 绪 论

本章在分析随机现象产生原因的基础上,介绍了用概率空间和随机对象的数学模型来描述随机现象统计规律的科学思想。此外,还介绍了信息与通信工程的处理对象、典型问题和常见随机对象,以说明随机现象在信息与通信工程中的普遍存在性,以及概率空间和随机对象的建模方法在信息与通信工程中的重要性。最后还简要介绍了全书内容及其逻辑关联。

1.1 自然界的随机现象

在自然科学的研究中,人们需要对各种自然现象进行观察。在观察过程中,人们发现有一类现象,在发生之前只能知道该现象各种可能的发生结果,但无法准确预知哪一个结果将发生。人们把具有上述特征的现象称为**随机现象**。

随机现象在自然界大量存在,下面是一些例子:

例 1.1 将硬币抛在桌面上,考察硬币的哪一面朝上。在抛之前观察者只知道有两种可能的结果,但却无法准确预测哪一面朝上。

例 1.2 观察明天中午某地点是否下雨。在明天中午未到来之前,观察者只知道“下雨”这个现象可能发生,也可能不发生,但却无法准确预知。

例 1.3 观察某机器的使用寿命。观察者只知道,该机器的使用寿命可以是一个大于零的实数,但具体是哪个实数却无法预知。

例 1.4 观察某天某个时刻某地点的温度。观察者只能知道,温度是某个范围内的一个数,但却无法准确预知这个数的准确值。

为什么会产生随机现象呢?原因是**决定现象发生的因素和机制过于复杂,超过了观察者的认识能力,观察者无法根据自己所掌握的资料准确预测现象的发生。**

以上面所举的抛硬币试验为例,当把一枚均匀的硬币抛向水平桌面时,决定硬币最终是正面朝上还是反面朝上的因素非常多,如抛出时的角度和速度、空气的流动、桌面的平整性等,或许还有其他观察者根本认识不到的因素。此外,这些因素对硬币最终是正面朝上还是反面朝上的影响机制也难以被观察者认识,以致最后的结果——硬币正面朝上还是反面朝上,在现象发生之前,对观察者来说也成了

无法确知的事情。

同样, 决定明天某地点是否下雨的因素有很多, 譬如空气的湿度、温度、流动方向等, 观察者对导致下雨这个现象的所有因素本身就难以一一认知, 并且这些因素怎样导致下雨现象发生的机制也不得而知。所以, 观察者对明天某地点是否下雨不能确定性地预知。

由此可见, 一个现象是否为随机现象, 依赖于观察者本身对现象发生原因的认知程度。当导致某现象发生的因素和作用机制已经被观察者充分认知, 则观察者就可以准确预测该现象的发生; 否则, 当观察者对导致某现象发生的因素和作用机制没有充分认知时, 这个现象对于观察者来说就是随机现象。

1.2 随机现象的统计规律

一个系统可以被抽象为若干个相互作用的因素, 及其这些因素相互作用的机制。分别用符号 $\xi_1, \xi_2, \dots, \xi_N$ 表示某系统的 N 个相互作用的因素, 符号 ϕ 表示这 N 个因素相互作用的机制, 于是该系统的**系统输出**可以用符号抽象地表示为

$$y = \phi(\xi_1, \xi_2, \dots, \xi_N) \quad (1.1)$$

例如, 在抛硬币试验中, y 表示硬币正面朝上或者硬币正面朝下, $\xi_1, \xi_2, \dots, \xi_N$ 表示所有影响硬币朝上或朝下的因素, ϕ 则表示这些因素之间的作用机制。虽然, 观察者难以对所有的因素 $\xi_1, \xi_2, \dots, \xi_N$ 和它们之间的作用机制 ϕ 一一认知, 但是 y 的某个结果一定是若干因素 $\xi_1, \xi_2, \dots, \xi_N$ 按照某个规律 ϕ 作用的结果。

设某个因素 ξ_i 的各种可能状态组成的集合为 $S_i, i = 1, \dots, N$, 当 $\xi_1, \xi_2, \dots, \xi_N$ 在集合 $S_1 \times S_2 \times \dots \times S_N$ 中取某组值时, 将对应系统的一个输出 y , 输出 y 称为该系统的一个**样本点**。定义所有可能的样本点 y 组成的集合, 为该系统的**样本空间**。让系统产生一次输出称为一次**试验**。

当观察者能清晰地认知系统的所有要素 $\xi_1, \xi_2, \dots, \xi_N$ 和作用机制 ϕ , 并且可以准确预知某次试验的输出, 则这个系统被称为**确知系统**; 否则当观察者对组成系统的所有要素 $\xi_1, \xi_2, \dots, \xi_N$ 和作用机制 ϕ 不能完全认知, 在试验之前只知道该系统的样本空间, 而不知道该次试验将输出样本空间中的哪一个样本, 这个系统就被称为**随机系统**。

随机现象实际上是随机系统的输出, 随机现象所有可能发生的结果是该随机系统的样本空间。

由于导致随机系统的系统输出——随机现象——的因素和这些因素间的作用机制已经超过了观察者的认知能力, 产生随机现象的所有因素及其作用机制在观察者面前成了一个“黑箱”。因此, 观察者无法对随机现象进行“从因推果”式的

研究。虽然观察者无法认知“黑箱”随机系统的内部机制，但随机系统的输出样本却是可观察的。通过对随机系统大量试验的观察，人们发现，对于同一个随机系统，不同样本的发生频率具有一定的稳定性。所谓样本的**频率**就是在若干次试验中，某个样本出现的次数占试验总次数的比例。所谓**频率稳定**就是当试验的次数增加时，样本的频率总是在一个常数左右微小波动。

例 1.5 历史上，曾经有人对抛硬币试验进行了观察，如有人抛 4040 次硬币，发现正面出现了 2048 次，正面出现的频率为 0.5069；又有人抛了 12 000 次，结果发现正面出现了 6019 次，频率为 0.5016。人们发现，如果硬币是均匀的，桌面是水平的，则正反面的频率基本上稳定在 1/2 左右。

例 1.6 决定某个汉字在某篇文章的某个位置是否被使用的因素非常复杂，观察者一般无法认知。但每个汉字都有稳定的使用频率。一个汉字的使用频率被定义为该汉字在所统计总字数中所占的比例。这个比例随所统计汉字字数的增加趋于一个稳定的值。

同样，在英文中，各个字母出现的频率也是稳定的。表 1.1 就是一份各字母使用频率的统计表。

表 1.1 英文字母使用频率统计表

字母	频率	字母	频率	字母	频率
空格	0.2	H	0.047	W	0.012
F	0.105	D	0.035	G	0.011
T	0.072	L	0.029	B	0.0105
O	0.0654	C	0.023	V	0.008
A	0.063	F	0.0225	K	0.003
N	0.059	U	0.0225	X	0.002
I	0.055	M	0.021	J	0.001
R	0.054	P	0.0175	Q	0.001
S	0.052	Y	0.012	Z	0.001

例 1.7 观察电话用户的通话时间。假设以分钟为最小单位，并且按照四舍五入的准则只计算整数分钟。决定某个用户某次通话所需时间的因素非常复杂，观察者一般无法预知，但是却可以对大量通话进行观察，结果会发现对于给定的通话时间，其频率具有稳定性。

对于随机系统，其样本点的频率稳定性也被称为样本点的**统计特性**。事实上，样本点的频率稳定性源于随机系统内部不可观察的作用因素及其作用机制的稳定性。所以，样本点的统计特性实际上是**随机系统内在的不可认知的“微观”规律在可观察的“宏观”现象上的反映**(见图 1.1)。对样本点统计特性的认知，也是一种间