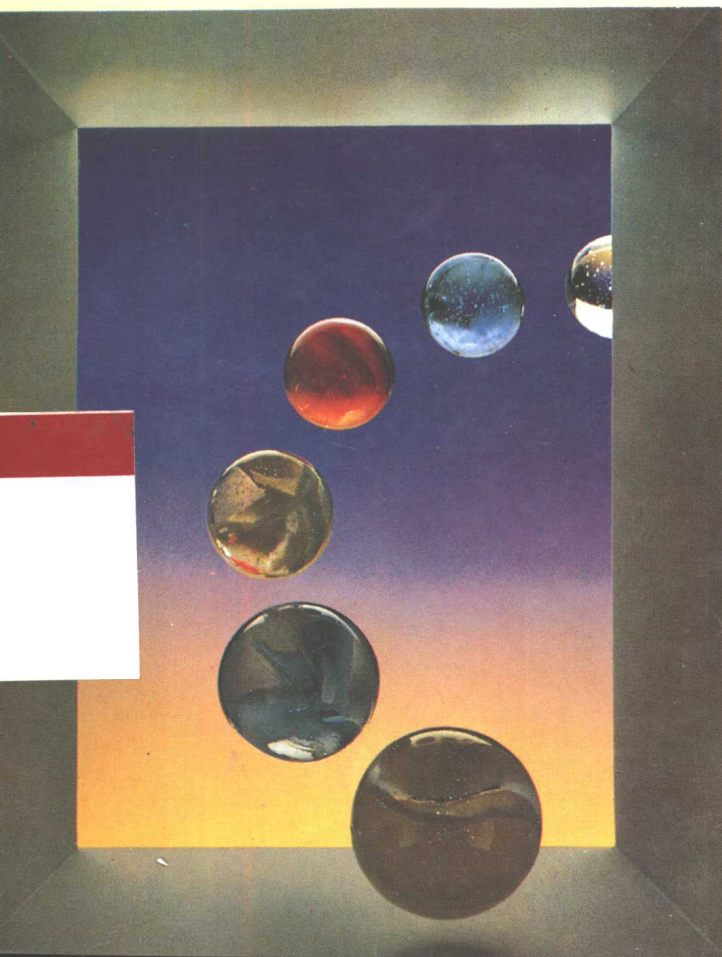


走出混沌

OUT OF CHAOS

走向数学丛书

方兆本 著



走向数学丛书

走出混沌

方兆本 著

湖南教育出版社

走出混沌

Out of Chaos

方兆本 著

Fang Zhaoben

责任编辑：孟实华

湖南教育出版社出版发行

湖南省新华书店经销 湖南省新华印刷二厂印刷

787×1092 毫米 32 开 印张：4.25 字数：100,000

1995 年 8 月第 1 版 1998 年 4 月第 2 次印刷

ISBN 7—5355—2245—9/G · 2240

定价：8.90 元

本书若有印刷、装订错误，可向承印厂调换

《走向数学》丛书编委会

顾问：王 元 丁石孙

主编：杨忠道 冯克勤

编委：李 忠 史树中 王建磐

黎景辉 孟实华

“走向数学”丛书

陳省身題





作者简介

方兆本，男，1945年生，浙江金华人，美国 Pittsburgh 大学统计学博士，教授。1966年毕业于中国科技大学数学系，后任教于该校，其间于1984—1987赴美国 Pittsburgh 大学进修，1989—1990赴加拿大 British Columbia 大学讲学。

主要研究方向是统计数学，可靠性数学及经济数学。著书《非参数统计》、《随机过程》。目前是美國当代统计索引 (CIS) 通讯编委，中国现场统计学会理事，美国 IMS，ASA 会员。

前 言

王 元

从力学、物理学、天文学直到化学、生物学、经济学与工程技术，无不用到数学。一个人从入小学到大学毕业的十六年中，有十三四年有数学课。可见数学之重要与其应用之广泛。

但提起数学，不少人仍觉得头痛，难以入门，甚至望而生畏。我以为要克服这个鸿沟，还是有可能的。近代数学难于接触，原因之一大概是由于其符号、语言与概念陌生，兼之近代数学的高度抽象与概括，难于了解与掌握。我想，如果知道讨论的对象的具体背景，则有可能掌握其实质。显然，一个非数学专业出身的人，要把数学专业的教科书都自修一遍，这在时间与精力上都不易做到。若停留在初等数学水平上，哪怕做了很多难题，似亦不会有助于对近代数学的了解。这就促使我们

设想出一套“走向数学”小丛书，其中每本小册子尽量用深入浅出的语言来讲述数学的某一问题或方面，使工程技术人员，非数学专业的大学生，甚至具有中学数学水平的人，亦能懂得书中全部或部分含义与内容。这对提高我国人民的数学修养与水平，可能会起些作用。显然，要将一门数学深入浅出地讲出来，决非易事。首先要对这门数学有深入的研究与透彻的了解。从整体上说，我国的数学水平还不高，能否较好地完成这一任务还难说。但我了解很多数学家的积极性很高，他们愿意为“走向数学”撰稿。这很值得高兴与欢迎。

承蒙国家自然科学基金委员会、中国数学会数学传播委员会与湖南教育出版社支持，得以出版这套“走向数学”丛书，谨致以感谢。

序

序的另一层含义是 order—有序。混沌是表象的无序、内在的有序。什么是混沌？混沌对不同的人意味着不同的事情。它是科学的新思维、新前沿。

混沌是对初始值的敏感、是拓扑传递性以及周期点稠密。混沌是随机性。混沌是面包师的杰作。混沌中有奇异吸引子。混沌的签名是分形。混沌是周期 3。混沌是正 Lyapunov 指数。混沌是拓扑熵大于 0。混沌是信息之源，是信息的膨胀。混沌是局部的不稳定和整体的稳定。混沌是简单与复杂的统一。混沌仅存在于理念之中，实验中只能得到可以预料的复杂性，非常像混沌而无法确认其是否真是混沌……

统计学、尤其是非线性时间序列是与动力系统理论及混沌科学有深刻内在联系的。概率统计是研究随机现象的科学，而混沌则呈现了随机性，这使概率统计方法可以作为分析混沌数据的有力工具。

统计学与概率论的方法铺设了走出混沌的路和桥。深入地了解混沌，才能走出混沌。走向数学才能走出混沌。

读了这本小册子之后，若读者能对上述说法有些体会那就是作者的最大愿望。

方兆本^①

于中国科技大学

1994.5

① 感谢国家自然科学基金对本书写作的资助。

目 录

前言 (王元)	1
序 (方兆本)	3
第一章 偶然与必然	1
§ 1.1 从掷骰子谈起	1
§ 1.2 正态分布	9
§ 1.3 大数定律与中心极限定理	13
§ 1.4 随机游动与反正弦律	21
第二章 把握随机性	25
§ 2.1 随机过程和时间序列	26
§ 2.2 布朗运动	35
§ 2.3 收敛性与遍历性	42
第三章 混沌	51
§ 3.1 数学家看混沌	53
§ 3.2 倍周期、分叉及 Feigenbaum 常数	66
§ 3.3 奇异吸引子、分形及其维数	71
§ 3.4 熵—不确定性的度量	89
第四章 走出混沌	96
§ 4.1 统计与混沌	97
§ 4.2 分形维数的估计	108
§ 4.3 非线性时间序列模型	113

§ 4.4 非参数回归	124
附记	128
参考书目	129

第一章 偶然与必然

偶然的東西是必然的，必然性本身規定自身為偶然性。

——黑格爾

偶然与必然是人类认识现实世界的一个永恒的话题。有人以为一件事、一个过程、一种关系不是偶然的便是必然的，非此即彼。各门学科所研究的当然是带有普遍性的规律。而能服从普遍规律的事情是必然的。人们还把他所不能理解、无法解释的现象归因于偶然的或者说超自然的力量：鬼使神差！在本书将要展开的关于混沌的话题里读者将会看到偶然与必然有机统一的有趣模式。

§ 1.1 从掷骰子谈起

人们对随机现象的研究已有了 400 多年的历史。在一门叫

概率统计的学科中研究者们正是力图从偶然中悟出必然。谈概率要从扔骰子说起。机会游戏，赌博在西方叫 Gambling。它要靠扔骰子碰运气。据考古学的发现最早的一颗骰子在伊拉克北部出土，是淡黄色陶器制成的立方体。骰子两对面分别是点子 2 和 3、4 和 5 及 6 和 1。时间约在公元前 2000 多年。虽然早在罗马时代就已有法律禁赌，但数千年来却屡禁不止。在数学著作中讨论赌博的最早例子要算 1494 年威尼斯教士 Luca Pacciolo 的著作中给出点数问题的一个解法。天文学家伽利略曾在 1642 年证明了同时掷 3 个骰子在总共 216 种可能出现的情况中能扔出和为 10 点仅有 27 次。也就是说机会为 $27/216=1/8$ 。所谓点数问题是 de Mere 向 Blaise Pascal 提出的。设一个赌徒赌在 8 次扔骰子中掷出一个 6 点。前 3 次他都失败了，赌局因故中断。试问：这时他应分得多少赌注？为此 Pascal 与 Fermat 书信往来研究赌博中的概率问题。什么叫概率，历史上也曾争论不休。一种观点认为它是不确定性的度量；而另一种观点则认为它是人们对某一陈述的信仰程度。以扔骰子出现 6 点为例，概率为 $\frac{1}{6}$ 。依前者这 $\frac{1}{6}$ 是千百次扔骰子所得经验观察的总结，即所谓统计频率；依后者，人们认为骰子 6 面出现的机会均等， $\frac{1}{6}$ 是对其出 6 点的可能性的认识。科学的发展倒不必等待人们对某一定义的普遍认同，科学的理论甚至有时是建立在定义得并不好的名词之上的。比如，人们对力和加速度的定义不甚满意，但卫星仍依照有关的运动理论在太空飞行。概率也是这样，尽管其定义说不清、道不明，但概率模型和统计方法仍在各个领域找到它自身的位置和价值。本书要讨论的混沌遇到同样的麻烦：什么是混沌？至今没有普遍认可的定义或者说其定义至今还是处于混沌之中！

回到最简单的随机游戏，扔一次骰子就相当于作一次随机试验。试验的结果有 1—6 点 6 种可能，它们组成样本空间。样本空间中的点组成事件。比如 $\{2, 4, 6\}$ 三点组成事件“扔出偶数”。而定义在事件所组成的集合上的衡量事件出现可能性大小的一个集合函数就称为该事件的概率。概率作为集合函数其自变量是事件，应变量是 0 和 1 之间的一个实数。它应当满足一定的公理：比如不可能事件的概率为 0，必然发生的事件概率为 1（俗话所说的百分之百的可能）等。本世纪 30 年代，前苏联数学家 Kolmogorov 建立了现代概率论的公理化体系，其数学基础是测度论。测度概念是日常生活中长度、面积、体积概念的自然推广。在本书有关分形维数的讨论中还会碰到测度论的方法与概念。经典概率的计算要依赖于组合计数，通俗而言是要会数数。把总的随机试验的结果数和有利于事件发生的试验结果数数清楚了，它们的比值就是该事件的概率。例如：扔骰子出现偶数的概率为 $3/6=1/2$ 。

科学的惊人成就就在于能对将来将要发生的事件进行预报。天文学家可以预报日蚀的准确时间，概率论之所以越来越受到重视也因为它能在大量试验中预报出某一现象出现的频数。当然，在预报之前我们应该先有一个像行星运动规律那样的理论假定。借助于概率的概念来描述的假设称为是统计假设。提出假设、设计实验、收集数据进行分析整理来检验假设、依据检验确认的假设去预测未来，预见未知就是统计学家的工作全貌。统计学家 Karl Pearson 在 1900 年的一篇文章中引用了 Weldon 扔骰子的实例，他把 12 粒骰子扔了 26306 次，对 5 点或 6 点出现的情况作了记录列在表 1.1 中。表中第 (1) 列为同一次扔的 12 粒中出现 5 或 6 点的粒数，第 (4) 列为观察频数，第 (3) 列则为假定骰子是公平的前提下（即各点出现的概率均同为 $1/6$ ）所

算出的相应的理论频数或称预报，理论值减去观察值的差列在第(2)列。首先，人们不难看到理论值与观察值相差不大，相对误差一般只有 0.01—0.10，且有某种一致性。(3)列理论值的计算要用二项分布。若作一次试验成功机会为 p 则失败的可能为 $1-p=q$ ，独立地(这又是统计学中颇费口舌才能解释清楚的概念)作同样的试验就构成一个 Bernoulli 试验序列。在 n 次 Bernoulli 试验中成功 k 次概率为 $\binom{n}{k} p^k q^{n-k}$ ，成功次数所服从的是二项分布。因 $(p+q)^n$ 的二项式展开第 k 项的值与 k 次成功的概率相同而得名。现在出现 5 点或 6 点的概率为 $2/6=1/3$ ，即 $p=\frac{1}{3}$ 。于是 12 粒骰子有 2 粒为 5 点或 6 点的概率为

$$P\{\text{成功 2 次}\} = \binom{12}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^{10} = \frac{67584}{531441}.$$

(1) 5 或 6 的数目	(2) 超过 I	(3) 预报 I	(4) 观察	(5) 预报 II	(6) 超过 II
0	+18	203	185	187	+2
1	+67	1216	1149	1146	-3
2	+80	3345	3265	3215	-50
3	+101	5576	5475	5465	-10
4	+159	6273	6114	6269	+115
5	-176	5018	5194	5115	-79
6	-140	2927	3067	3043	-24
7	-76	1255	1331	1330	-1
8	-11	392	403	424	+21
9	-18	87	105	96	-9
10	-1	13	14	15	+1
11	-3	1	4	1	-3
12	0	0	0	0	0
总计	0	26, 306	26, 306	26, 306	0

表 1.1

而在 26306 次试验中的近似理论频数应为 $26306 \times P\{12 \text{ 粒中有 } 2 \text{ 粒为 } 5 \text{ 或 } 6 \text{ 点}\} = 3345 \text{ 次}$.

细心的读者还会发现在第(2)列(3)一(4)的误差值中呈现出系统的偏差,上面为+下面为一.一般而言,+、-误差应无规律地掺合呈随机的(又一个不易说清的概念)状态.一串+、-号是否随机的检验将在 § 4.1 讨论.但是现在随机性无处可寻、原因何在?人们便怀疑骰子公平的假设是否真实.依据 12×26306 次掷骰子的实际观察人们可以估出出现 5 点或 6 点的概率 $\hat{p} = 0.3376986$, $\hat{p} > \frac{1}{3}$. 依 $\hat{p}$ 可以算出新的理论频数列第(5)列,再与观察频数(4)列的数据比较,(5)一(4)列入(6)列.这时,可以发现不仅误差减小了,而且+、-相间呈无规律状,随机性更强了.正是这样人们通过建立近似的理论模型来更好地描述客观现实,从偶然中探索必然规律.

读者容易找到各种程度的概率论入门书去了解概率的基本概念,这里引入一些最必须的记号和概念.

定义 样本空间 Ω 的子集合 A (即事件) 上的函数 P 若满足 (i) $0 \leq P(A) \leq 1$;

(ii) $P(\emptyset) = 0$, $P(\Omega) = 1$ ($\emptyset$ 为空集);

(iii) 若 $A_1, A_2, \dots$ 为一串互不相交的子集,则

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$$

称之为**概率**(或**概率测度**).

iii) 中 $\cup$ 表示集合的并,代表两件事至少发生一件, A_i 互不相交则表示事件不同时发生. iii) 又常称作**概率的可加性**.

定义 随机试验结果的全体称为**样本空间** Ω , 样本空间中

的集合称为事件,可以定义概率的事件的全体集合记为 $\mathcal{B}$, $(\Omega, \mathcal{B}, P)$ 称为概率空间 (它是一般测度空间的特例).

$(\Omega, \mathcal{B}, P)$ 便是我们讨论问题的出发点. 为了能比较方便地讨论概率, 我们引入随机变量的概念. 随机变量是指其取值要随机会而变的那种变量, 如果 $\omega \in \Omega$ 为试验结果则取值为实数的变量 $X(\omega)$ 为随机变量. 而 $X(\omega)$ 在某一范围取值则代表了 $\mathcal{B}$ 中的一个事件. 例如, 记

$$X(\omega) = \begin{cases} 1, & \text{扔币出正面;} \\ 0, & \text{扔币出反面,} \end{cases}$$

则 $\{X(\omega) \leq \frac{1}{2}\}$, $\{X(\omega) = 0\}$, $\{0 \leq X(\omega) < 0.9\}$ 都代表了同一事件 {反面向上}. 引入随机变量之后对偶然事件发生的可能性的描述转化为对随机变量的分布的研究. 出现正、反面机会均等就是 $X(\omega)$ 服从两点分布, 其 $P\{X = 1\} = p$ 和 $P\{X = 0\} = 1 - p = q$, 而且 $p = q = \frac{1}{2}$. 研究随机变量既要考查它取值的范围, 又要了解取这些值的可能性的. 这两点的综合就描述为随机变量的分布. 分布本身即代表偶然现象的内在规律. 随机变量的分布常分为离散的与连续的两大类, 后者又有绝对连续与奇异之分. 具体运算中更为方便的是分布函数.

定义 对实数 x , 概率 $P\{X(\omega) \leq x\}$ 称为是随机变量 X 的累积分布函数. 记为 $F(x)$, 若 $F(x)$ 的导数 $F'(x)$ 存在, 则记为 $f(x)$, 并称为是 X 的密度函数.

依微积分基本定理有:

$$\int_{-\infty}^x f(x) dx = F(x),$$