醉汉的脚步

随机性如何主宰我们的生活

The Drunkard's Walk: How Randomness Rules Our Lives

■【英】列纳德·蒙洛迪诺 / 著 ■ 郭斯羽 / 译 ■

数学圈丛书 MATHEMATICAL CIRCLES



湖南科学技术出版社
Hunan Science & Technology Press

数学圈丛书

醉汉的脚步

随机性如何主宰我们的生活

The Drunkard's Walk: How Randomness Rules Our Lives

■【英】列纳德·蒙洛迪诺 / 著 ■ 郭斯羽 / 译 ■



湖 南 科 学 技 术 出 版 社

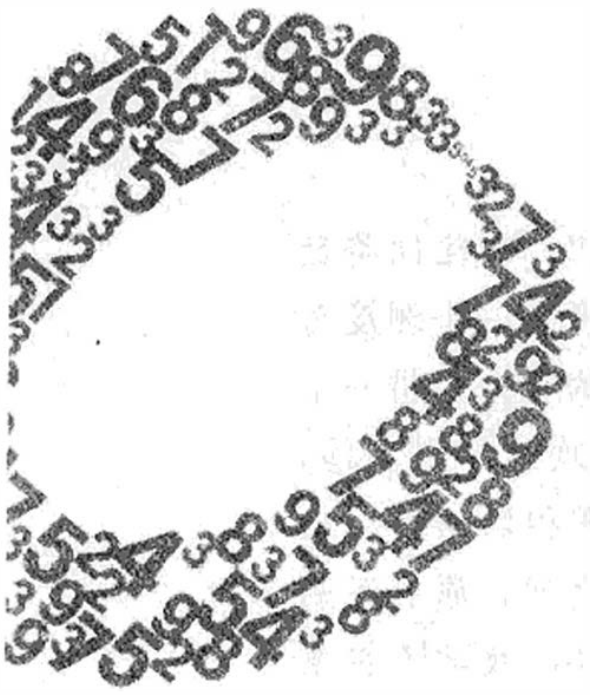
The Drunkand's Walk: How Randomness Rules Our Lives © 2008 by Leonard Mlodinow
Simplified Chinese translation copyright © 2010 by Hunan Science & Technology Press Published by arrangement with Writers House, LLC
ALL RIGHTS RESERVED
湖南科学技术出版社通过博达著作权代理公司获得本书中文简体版中国大陆地区出版发行权。
著作权合同登记号：18-2008-093

图书在版编目（CIP）数据

醉汉的脚步 /（英）列纳德·蒙洛迪诺著 ；郭斯羽译. -- 长沙 ：湖南科学技术出版社，2010.5
（数学圈丛书）
书名原文：The Drunkard's Walk : How Randomness Rules Our Lives
ISBN 978-7-5357-6225-2
Ⅰ. ①醉… Ⅱ. ①蒙… ②郭… Ⅲ. ①数学—普及读物 Ⅳ. ①O1-49
中国版本图书馆 CIP 数据核字(2010)第 091545 号

数学圈丛书
醉汉的脚步 随机性如何主宰我们的生活
著　　者：【英】列纳德·蒙洛迪诺
译　　者：郭斯羽
责任编辑：吴　炜　孙桂均
出版发行：湖南科学技术出版社
社　　址：长沙市湘雅路 276 号
　　http://www.hnstp.com
邮购联系：本社直销科　0731-84375808
印　　刷：湖南凌华印务有限责任公司
　　（印装质量问题请直接与本厂联系）
厂　　址：长沙市长沙县黄花镇黄花印刷工业园
邮　　编：410013
出版日期：2010 年 6 月第 1 版第 1 次
开　　本：710mm×1000mm　1/16
印　　张：16.5
插　　页：2
字　　数：202000
书　　号：ISBN 978-7-5357-6225-2
定　　价：35.00 元

（版权所有 · 翻印必究）



总 序

欢迎你来数学圈，那是我们熟悉而陌生的园地。

我们熟悉它，因为几乎每个人都走过多年的数学路，从123走到6月6（或7月7），从课堂走进考场。然后，我们把它留给最后一张考卷，解放的头脑，不再为它留一点儿空间。我们也陌生，模糊的记忆里，是残缺的公式和零乱的图形，是课堂的催眠曲，是考场的蒙汗药……去吧，那些被课本和考卷异化和扭曲了的数学；忘记那一朵朵恶之花，我们会迎来新的百花园。

“数学圈丛书”请大家走进数学圈，也走近数学圈子里的人。这是一套新视角下的数学读物，它不专门传达任何具体的数学知识和解题技巧，而以“非数学的形式来普及数学”，着重宣扬数学和数学家的思想和精神。它的目的不是教人学数学，而是改变人们对数学和数学家的看法，把数学融入大众文化，回到人们的生活。读这些书不需要智力竞赛的紧张，而是要一点儿文艺欣赏的平和。你可以怀着360样心情来享受数学，经历它的趣味和生命，感悟符号背后的情感和人生。



2 醉汉的脚步

没有人怀疑数学是文化的一部分，但偌大的“文化”，却往往将数学排除在外。当然，从人数来看，数学家在文化人中顶多占一个测度为零的空间。但是，数学的每一点进步都影响着整个文明的根基。借一个历史学家的话说，“有谁知道，在微积分和路易十四的政治原则之间，在古典的城邦和欧几里得几何之间，在油画的空间透视和以铁路、电话、远距离武器制胜空间之间，在对位音乐和信用经济之间，原有深刻的一致关系呢?”（斯宾格勒《西方的没落·导言》）所以，数学不在象牙塔，就在身边。上帝用混乱的语言摧毁了石头的巴比塔，而人类用同一种语言建造了精神的巴比塔，那就是数学。它是艺术，也是生活；是态度，也是信仰；是最复杂的简单，也是最单纯的完美。

数学是生活。当然，我们的意思不是说生活离不开算术，技术离不开微积分；而是说数学本身也能成为大众的生活态度和生活方式。很多人感觉数学枯燥无味，是因为他把数学从生活中赶走了。当你发现一个小公式也像一首小诗那么多情的时候，还忍心把它忘记吗？大家能享受“诗意的生活”，从这点说，数学是一样的。

数学的生活很简单。如今流行着很多深藏“大道理”的小故事，那些道理多半取决于讲道理的人的态度和立场。它们是多变的，因为多变而被随意扭曲，因为扭曲而成为多样选择的理由。在所谓“后现代”的今天，似乎一切东西都成为多元的，人们像浮萍一样漂荡在多样选择的迷雾里，起码的追求也失落在“和谐”的“中庸”里。数学能告诉我们，多元的背后存在统一，极端才是和谐的源泉和基础。从某种意义说，数学的精神就是追求极端，它永远选择最简的、最美的，当然也是最好的。数学决没有圆滑的道理，也不为模糊的借口留下一点儿空间。

数学生活也浪漫。很多人怕数学抽象，却喜欢抽象的绘画和怪诞的文学。可见抽象不是数学的罪过。艺术家的想象力令人羡慕，而数学家的想象力更多。希尔伯特说过，如果哪个数学家一旦改行做了小说家（真的有），我们不要惊奇——因为那人缺乏足够的想象力做数学家，却足够做一个小说家。懂一点儿数学的伏尔泰也感觉，阿基米德头脑的想

象力比荷马的多。我们认为艺术家最有想象力，那是因为我们自己太缺乏想象力。

数学是明澈的思维。生活里的许多巧合——那些常被有心或无心地异化为玄妙或骗术法宝的巧合，也许只是自然而简单的数学结果。以数学的眼光来看生活，不会有那么多的模糊。有数学精神的人多了，骗子（特别是那些穿戴科学衣冠的骗子）的空间就小了。无限的虚幻能在数学找到最踏实的归宿，它们“如龙涎香和麝香，如安息香和乳香，对精神和感官的激动都一一颂扬。”（波德莱尔《恶之花·感应》）

数学是奇异的旅行。数学在某个属于它们自身的永恒而朦胧的地方，在那片朦胧的土地上，我们已经看到了三角形的三个内角和等于180度，三条中线总是交于一点而且以1∶2的比例分划每一条中线；在那片朦胧的土地上，还存在着无数更令人惊奇的几何图形和数字的奇妙，等着我们去和它们相遇。

数学是纯美的艺术。数学家像画家和诗人，都创造“模式”，不过是用思想来创造，用符号来表达。数学的思想，就像画家的色彩和诗人的文字，以和谐的方式组织起来。数学的世界里没有丑陋的位置。在数学家的眼里，自己笔下的公式和符号就像希腊神话里的那位塞浦路斯国王，从自己的雕像看到了爱人的生命。在数学里，在那比石头还严密的逻辑里，真的藏着数学家们的美的追求，藏着他们的性情和生命。

数学是精神的自由。惟独在数学中，人们可以通过完全自由的思想达到自我的满足。不论王摩诘的“雪地芭蕉”还是皮格马利翁（Pygmalion）的加拉提亚（Galatea），都能在数学中找到。数学没有任何外在的约束，约束数学的还是数学。

数学是永不停歇的人生。学数学的感觉就像在爬山，为了寻找新的山峰不停地去攀爬。当我们对寻找新的山峰不再感兴趣，生命也就结束了。

不论你是不是知道一点儿（或很多）数学，都可以走进数学圈，孔夫子说过：“知之者不如好之者，好之者不如乐之者。”只要“君子乐

之”，就走进了一种高远的境界。王国维先生讲人生境界，是从“望极天涯”到“蓦然回首”，换一种眼光看，就是从无穷回到眼前，从无限回归有限。而真正圆满了这个过程的，就是数学。来数学圈走走，我们也许能唤回正在失去的灵魂，找回一个圆满的人生。

1939 年 12 月，怀特海在哈佛大学演讲“数学与善”时说：“因为有无限的主题和内容，数学甚至现代数学，也还是处在婴儿时期的学问。如果文明继续发展，那么在今后两千年，人类思想的新特点就是数学理解占统治地位。”这个想法也许浪漫，但他期许的年代似乎太过久远——他自己曾估计，一个新的思想模式渗透进一个文化的核心，需要 1000 年——我们的希望是，这个过程会快一点儿，更快一点儿。

最后，我们借从数学家成为最有想象力的作家的卡洛尔笔下的爱丽丝和那只著名的“柴郡猫”的一段充满数学趣味的对话，来总结我们的数学圈旅行：

“你能告诉我，我从这儿该走哪条路吗?”

“那多半儿要看你想去哪儿。”猫说。

“我不在乎去哪儿——”爱丽丝说。

“那么你走哪条路都没关系，”猫说。

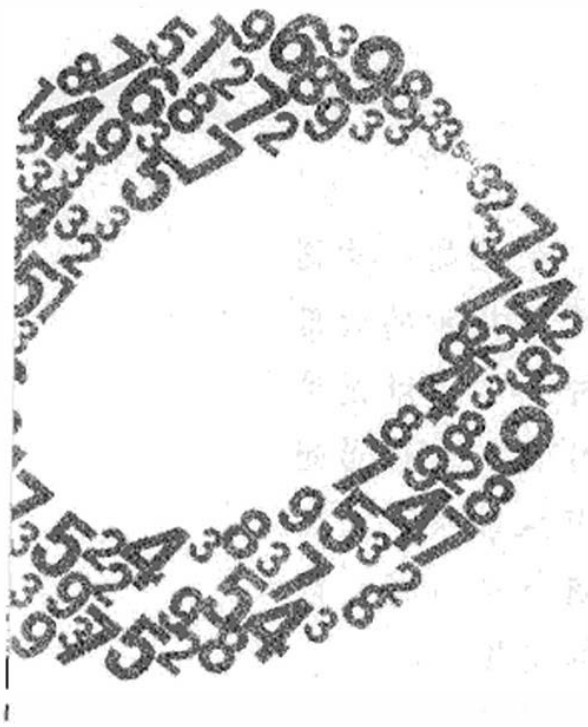
“　　只要能到个地方就行，”爱丽丝解释。

“噢，当然，你总能到个地方的，”猫说，“只要你走得够远。”

我们的数学圈没有起点，也没有终点，不论怎么走，只要走得够远，你总能到某个地方的。

李泳

2006 年 8 月



前　言

早几年前，有人靠着一张尾数为48的彩票，赢得了西班牙全国乐透大奖。对此“成就”大感自豪的中奖者，披露了带来财富的那个理论。“我连着7个晚上都梦到了数字7，”他说，“而7乘7等于48。”[1]对乘法表掌握得更好些的人也许会窃笑不已，但对于这个世界，每个人都会创造出他自己的看法，再以之过滤和处理所感知的一切，从日常生活里冲刷着我们的数据之海中抽取出意义。我们也经常犯下一些错误，虽然不如前面那人那么明显，但却同样重大。

人类直觉不适于处理涉及不确定性的情势，这一事实早至20世纪30年代就已为人所知。当时，研究者发现人们既不能构造出一个能通过随机性检验的数列，也不能可靠地分辨某个给定数列是否是随机产生的。一个新的学术领域在过去几十年里逐渐浮出水面，它研究的是当信息不全或不完美时，人们如何进行判断与决策。研究证明，一旦偶然性牵涉其中，人们的思维处理通常就会产生严重的问题。这些研究工作综合了许多学科，如数学与传统科

学，乃至于认知心理学、行为经济学以及现代神经科学。最近的一次诺贝尔经济学奖使该类研究成为了正统，但尽管如此，研究所揭示的大部分经验教训却仍未脱出学术圈子而流入普通人的头脑之中。本书就是试图改观这种情况的一个尝试，它讲述的是主宰偶然性的原理、这些思想的发展及其在政治、商业、医学、经济、体育、休闲和其他涉及人之因素的领域中发挥作用的方式。书中同样谈到了我们进行选择的方式，以及在面对随机性或不确定性时导致错误判断和糟糕决定的过程。

信息的缺乏常常导致互相矛盾的解释。为什么确认全球变暖会如此费力？为什么有的药品声称安全而之后又从市场被召回？为什么有人会不出所料地不赞同我那个巧克力奶昔乃是维护心脏健康的饮食之必需成分的观点？其原因就在于此。对数据的误读会很不幸地带来或大或小的负面后果。例如我们将要看到的那样，医患双方都常常错误解释了有关药品有效性和重要医学检查结果之意义的统计数据；家长、老师和学生们误解了如SAT这类考试的显著性，品酒师在给葡萄酒评分时也犯下同样的错误；而投资人则从基金的历史表现中得出了并不成立的结论。

体育领域中已发展出一种文化，在这种文化下，基于我们对相关性的直觉，球队的成败常常大部分都归结于教练的个人能力。结果当球队失利时，教练就常常被炒了鱿鱼。但对主要体育项目中解聘教练这种做法的数学分析表明，它们平均而言对球队表现并无影响[2]。类似现象也存在于企业界。这里的人们认为首席执行官（CEO）们拥有可以成就或搞垮整个公司的超人能力。但在柯达（Kodak）、朗讯（Lucent）、施乐（Xerox）以及其他公司，这种能力被一次次证明不过是幻觉而已。例如在1990年代，当加里·温特（Gary Wendt）在杰克·韦尔奇（Jack Welch）手下掌管GE投资公司（GE Capital Services）时，他被认为是这个国家中最聪明的生意人之一。后来，温特赌上了名誉，试图拯救深陷麻烦的金融公司Conseco以赢得4500万美元奖金。温特来当头儿，Conseco的麻烦也就到了头，对于这种说法，投资者们显然颇为认可：一年之内，公司股票涨到了原来的3倍。但两年后，温特突然辞了职，

Conseco破了产，而公司股票也成了“仙股”①[3]。温特接手的是否根本就是个不可能之任务？是不是他大权在握时却睡着了？还是他的冠冕根本就来自颇有问题的假设，例如管理者拥有几近绝对的能力来影响公司，或者某人以往的单单一次成功就足以成为其未来表现的可靠指标？在任何特定的场合中，人们都不能在未曾仔细检查当前状况的细节时，就对这些问题的答案充满信心。在书中的若干例子里，我就将进行这样的检查。但更重要地，我将叙述用来发现偶然性之蛛丝马迹所需的工具。

顶着人类的直觉破浪前行，是一件困难的事。如所要看到的那样，人类头脑之构造，是要给每一个事件找出确定理由，因此它难以接受无关或随机因素所造成的影响。因此，我们首先要认识到，成败有时并非来自于过人的能力或无能，而是来自于如经济学家艾智仁（Armen Alchian）所说的“幸运的环境”[4]。随机过程就本性而言非常基本，在日常生活中也无所不在，但大多数人却不了解它，或者很少想到它。

书名《醉汉的脚步》（*The Drunkard's Walk*）来自于一个描述随机运动的数学术语。当分子飞越空间并不断撞击其他分子或被其他分子撞击时，它的脚步就是如此。分子的路径可以用来比拟我们的生活、我们从大学到工作、从单身到建立家庭、打高尔夫球时从进第1洞到进第18洞之间的过程。惊人之处在于，用于理解醉汉脚步的工具，同样也能用于理解日常生活中的事情。本书的目的，就是说明偶然性在我们周围世界中所扮演的角色，以及如何在世事之中发现它正发挥着作用。希望在这趟随机世界的旅行之后，读者您将开始以一种全新的视角来看待生活，并对日常世界有了更深刻的理解。

---

① 原文直译为公司股票的交易价格仅仅只有几个便士或美分而已，这里便用了“仙股”这个贴切但不正式的名字。



目 录

第一章　透过随机性的目镜凝视 …… (1)

偶然性所扮演的隐蔽角色——为什么人会被老鼠击败

第二章　真理与半真理的法则 …… (21)

概率的基本原理以及对它的滥用——为什么动听的故事常常比不可靠的解释更不可能为真

第三章　寻找穿越可能性空间之路 …… (41)

考虑随机性状况的方法框架——由瘟疫肆虐时的一个意大利赌徒到《让我们做个交易》

第四章　追寻通往成功的路径 …… (61)

如何数清事件可能发生方式的数目，以及这个问题之所以很重要的原因——期望的数学含义

第五章　针锋相对的大数定律与小数定律 …… (82)
几率在我们的观察结果中得到反映的程度
芝诺悖论，极限的概念，以及在轮盘赌上制胜
赌场

第六章　假阳性与好错误 …… (105)
如何根据以往事件或新知识调整期望——从医
学检查到辛普森案中的条件概率错误，以及检
察官谬误

第七章　测量与误差定律 …… (125)
测量中的意义与无意义——钟形曲线和葡萄酒
评分、政治民意调查、分数以及行星的位置

第八章　混沌中的秩序 …… (147)
大数是如何抹去随机性之无序的——或说为何
2 亿驾驶员能形成一个具有习惯性的生物

第九章　模式的错觉与错觉的模式 …… (170)
我们为何常常被偶然性事件中的规律性所愚
弄——连续 100 万个 0 或华尔街巨头的成功
可能是随机的产物么

第十章　醉汉的脚步 …… (193)
为什么偶然性是一个比因果性更为基本的概
念　　布鲁斯·威利斯、比尔·盖茨以及生
活的事故常态理论



致谢 …… (221)

注释 …… (223)

第一章

透过随机性的目镜凝视

记得在十来岁时，我曾看着黄色的安息日烛火，在为它供应燃料的白色石蜡圆柱上随意舞动。那时的我还太年轻，觉不到烛光的浪漫；但火焰形成的摇曳影像，却仍使我感到烛光的不可思议。它飘忽不定，变幻无方，时大时小，而这些看来都没有明显的原因或计划。我相信在那火焰后面，一定有着它的节奏和成因，有着某种科学家们能用数学方程来预测和解释的模式。“生活并非如此，”父亲却这么说道，“有时，无法预见的事情是会发生的。”他说起了在布痕瓦尔德（Buchenwald）那个他被关押、忍饥挨饿的纳粹集中营里，他曾从面包铺里偷了一条面包。铺子的面包师傅让盖世太保(Gestapo)① 把所有可能犯下这一罪行的嫌犯集中起来，排成一排。“谁偷了面包?”面包师傅问道。如果无人作答，他就会叫卫兵将嫌犯们一个接一个地枪毙，直到他们全部被杀死或是有人出来认罪为止。为了放其他人一条生路，我父亲站了出来。父亲不想给自己装点上一层英雄主义色彩，他说这只是因为他已经预计到，不管怎样他都会被枪毙。但面包师傅没有叫人打死我父亲，反而给了他一个待遇优厚的助手职位。“这就是个碰运气的事，”父亲说，“这跟你怎么做毫无关系。但如果事情不是这样发生的话，你永远都不会被生出来了。”我震惊了：我竟然应该为了我的存在而感谢希特勒！德国人杀害了我父亲的妻子和两个年幼的孩子，彻底抹去了他的前半生。

---

① 纳粹德国的秘密警察机构。

因此，若不是因为战争，我父亲就永不会移民到纽约，永不会遇到同为难民的我母亲，当然也永不会生下我和我的两个兄弟。

父亲很少谈及这场战争。当时我还不怎么觉得，只是多年后才开始慢慢认识到，当父亲向我谈起他痛苦的过去时，并非只希望我知道这些经历，而更多地是想给我上一堂关于生命的大课。战争是个极端的情况，但机遇在生活中的角色却不是由种种极端状况来体现的。如同那烛火一样，生命的主线总是被各种随机事件诱往新的方向，这些随机事件——连同我们的反应——就决定了我们的命运。如此所造成的结果，则是人生既难以预测，也难以解释。在注视罗夏墨迹（Rorschach blot）① 时，你可能会看到麦当娜（Madonna）而我却看到一只鸭嘴兽，而同样地，我们在商务、法律、医药、体育和媒体中所遇到的数据，或孩子的三年级成绩单，都能以多种方式来解读。不过解释偶然性在事件中所扮演的角色，可不像解释一个罗夏墨迹：有的解释方法是对的，有的则是错的。

我们在不确定局面下进行评估和选择时，常常会依赖于直觉。假设我们碰到了一只微笑的剑齿虎，我们就必须确定这微笑是因为它又肥又快活呢，还是因为饿得半死的它看到了我们这顿美餐。这时，直觉的处理无疑更具进化优势。但现代世界有着完全不同的物种力量平衡，同样的直觉在今天则伴随着缺点。用习惯性的思维方式去对待今天的老虎，可能会导致非最佳甚至是不当的决策。对于研究大脑如何处理不确定性的人而言，这个结论一点儿也不奇怪：许多研究表明，人类大脑中对不确定局面进行评估的部分，和处理情感这一人类特性——常常也被认为是非理性的主要来源——的部分，两者之间存在着紧密联系。例如，通过功能性磁共振成像

① 罗夏墨迹实验是一种心理学实验，最初由赫尔曼·罗夏（Hermann Rorschach）于1921年提出。被试者被要求观察10个特别设计的墨迹，并根据他们觉得墨迹像什么东西来分析其性格。

(Functional magnetic resonance imaging，fMRI）得到的图像表明，风险和回报是由多巴胺能系统的某些部分来估定的，而多巴胺能系统正是对于动机和情感过程非常重要的大脑奖励回路①[1]。这些图像还显示，陷在不确定性中进行决策时，同样与人类情绪状态、特别是恐惧感有关的扁桃体，处在了激活状态。[2]

人们对涉及偶然性的形势进行分析的机制，是进化因素、大脑构造、个人经验、知识及情感共同形成的复杂产物。事实上，人类对于不确定性的反应是如此复杂，以至于某些时候大脑的不同部分会得出不同结论，并显然通过相互斗争来最终确定哪个结论能够做主。举个例子，如果每 4 次中有 3 次，你吃完虾后脸肿得有正常大小的 5 倍大，那么“逻辑的”大脑左半球会试图找出一个模式来描述其中的规律，而“直觉的”大脑右半球则简单地告诉你“离虾远点”。至少这就是研究者们通过不那么痛苦的实验所发现的结论。这个实验游戏称为概率猜测，其中代替虾和组胺②的，是一系列展示给被试者的两种颜色——比方说绿与红——的卡片或灯光。通过实验安排，这些色彩以不同的概率出现，但除此之外别无其他模式。比如在类似红-红-绿-红-绿-红-红-绿-绿-红-红-红这样的颜色序列中，红色出现的次数是绿色的两倍。被试者的任务，是通过一段时间的观察，来预测下一次将出现的色彩是红还是绿。

玩这个游戏可以有两种基本策略。策略一是总以出现最为频繁的那个颜色作为猜测结果。老鼠和其他非人类动物就喜欢走这条路。该策略总能保证得到一个较为确定的预测成功率，但同时也得接受如下事实，即我们无法做得更好一些。比方说，如果绿色出现的比

---

① 多巴胺（dopamine）是一种神经递质，其主要功能之一，就是会产生愉悦感，并鼓励个体重复造成这些愉悦感的行为。

② 组胺（histamine）是广泛存在于人体的一种自体活性物质，与靶细胞的组胺受体结合后，可产生若干病理生理学反应，如造成毛细血管扩张而出现红斑、造成毛细血管通透性增加而形成肿块、刺激神经末梢引起痛痒的感觉等。

例为 75%，那就总是猜下一个出现的仍是绿色，此时的预测准确率将为 75%。策略二则是将你所预测的绿/红色比例与所观察到的绿/红色比例“匹配”起来。如果绿色和红色的确按某种模式出现，而我们又找到了这一模式，该策略就能每次都做出正确预测。但如果颜色的出现完全随机，那最好还是坚持第一种策略。在绿色以 75% 的比例随机出现的情况中，采用策略二，在 10 次中大概只能猜中 6 次。

人类通常都试图猜出红/绿出现的模式，这就让我们输给了老鼠。但还有一些人，他们遭受了称为开颅术（split brain）的外科手术术后大脑损伤，其大脑左右半球之间的通讯被切断了。如果让这些病人进行上面的概率实验，并让他们只使用左眼来观察彩色灯光或卡片，同时只用左手来给出其预测的话，这就相当于只对他们的右半大脑进行实验；而如果使用的是右眼和右手，那么实验就是针对左半大脑进行的了。研究者们如此实验后发现，同一个病人，其大脑右半球总是选择猜测更频繁出现的颜色，而左半球则总是试图找出一个模式。[3]

在面临不确定性时能做出明智的评估和选择，是一种罕有的能力。但与任何技能一样，经验可以使之得到改善。在后面的章节中，我将细究偶然性在周围世界中所扮演的角色，还有那些经多个世纪发展起来的、能帮助我们理解这一角色的思想，以及常常使我们误入歧途的各种因素。英国哲学家和数学家罗素（Bertrand Russell）曾写道：

“我们都从‘朴素的现实’——也就是‘事物就是它们看上去的那样’这一教条——出发，认为草是绿的，石头是硬的，雪是冷的。但物理学使我们确信，草的绿、石头的硬和雪的冷其实并非是我们通过自身经验所知道的草的绿、石头的硬和雪的冷，而是某种非常不同的东西。”[4]

接下来，通过随机性的目镜，我们将看到在生活中也同样有许

多事情，它们并非如表面所见，而是别的非常不同的东西。

2002 年，诺贝尔委员会将诺贝尔经济学奖授予了一位名叫卡尼曼（Daniel Kahneman）的科学家。如今的经济学家什么问题都研究——他们解释教师的工资为何如此之低，足球队何以能值这么多钱，而猪的生理功能又是如何对养猪场规模形成了限制（一头猪的排泄物是一个人的 3～5 倍，因此，数千头猪的一个农场，常常比它旁边的城市产生更多的废物）[5]。尽管经济学家发起了这么多了不起的研究，2002 年诺贝尔奖的引人注目之处，在于卡尼曼并非是一名经济学家。他是一名心理学家。数十年中，他与已故的特沃斯基（Amos Tversky）一道，研究和阐明了对随机性之误解的各种不同类型，而正是这些误解，造成了本书中将要谈及的许多常见的谬误。

要理解随机性在生活中扮演的角色，最大的挑战在于，尽管随机性的基本原则产生于日常的逻辑，但其导致的许多后果是违反直觉的。卡尼曼和特沃斯基的研究本身就是被一个随机事件所激发的。20 世纪 60 年代中期，当时还是希伯来大学（Hebrew University）心理学副教授的卡尼曼，接受了一桩相当不刺激的杂差：为一群以色列空军飞行教官讲授有关行为矫正的传统观点及其在飞行训练心理学中的应用。卡尼曼强调指出，对正确行为进行奖励能起到矫正作用，而惩罚错误行为却没有同样效果。这时一名学员打住了他，并说出了自己的看法。这个看法使卡尼曼恍然而悟，并在之后几十年中引导了他的研究[6]。

“我常常为飞行员所做出的漂亮机动而热烈地表扬他们，但下一次他们总是飞得更糟糕些，”这名飞行教官说道，“我也曾为飞行员糟糕的机动冲他们尖叫，而大体上，他们下一次总是要飞得好些。所以请不要告诉我奖励有效而惩罚无效，这与我的经验不符。”其他的教官也都同意这一观点。对卡尼曼而言，飞行教官的经历听来不假，但他也坚信动物实验所给出的结果，即奖励比惩罚更为有效。

这个明显的悖论使他陷入了沉思，然后一个想法出现了：飞行教官的尖叫的确发生在飞行员改进其飞行之前，但不同于这个表象，尖叫并不是导致改进的原因。

道理何在呢？答案就在于一种被称为回归均值的现象中，即在任何一系列随机事件中，一个不寻常的事件后面，总是更可能跟着一个相对更平凡的事件，而这完全都出自偶然。飞行教官的例子中，回归是这样来做的：每个飞行学员都具备一定的战斗机驾驶能力。飞行技术水平的提高牵涉许多因素，还需要大量的练习。因此，尽管他们的技术通过飞行训练在慢慢提高，但仅仅比较连着的两次飞行，是难以察觉这个提高的。所以任何特别好或特别糟的飞行表现，基本都是个运气问题。如果飞行员完成了一次优异的着陆，远好于其正常水平，那么第二天他就有很高的机会表现得更接近于正常水平——换而言之，比这次糟糕一些。如果教官因为这次的优秀表现表扬了他，那这个表扬看来就没起到好作用。而如果这名飞行员的着陆非常糟糕——飞机冲出跑道尽头并栽进了基地自助餐厅盛玉米杂烩的大桶里——那么他同样有很高的机会在第二天飞得更接近正常水平——也就是飞得更好。假设教官有着为其糟糕表现而尖叫“你这头笨猩猩!”的习惯，那么表面上看来，这个批评就收到了好效果。如此就出现了一种表面上的模式：学员飞得好，表扬就起不到好效果；而如果学员的表现糟得让教官高分贝地将他与低等灵长类相提并论，那之后其表现就有所改进。卡尼曼课上的那些飞行教官们，就从这样的经验中得出结论，认为尖声训斥是一件有力的教育工具。但实际上，尖叫根本没有改变什么。

这个直觉的错误引起了卡尼曼的思索。这种误解是不是普遍现象？我们是不是也像这些飞行教官一样，相信粗暴的批评能改进孩子的行为或雇员的表现？面对着不确定性时，我们是否还做出了其他误判？卡尼曼知道，人类出于必要而使用了某些策略，以降低进行判断的复杂性，这个过程中，对概率的直觉就扮演了重要角色。

你是否也曾在享用了从街头小贩购买的看似甜美可口的酸橘汁腌鱼炸玉米粉圆饼后觉得不舒服？你并不会有意识地回忆起所有那些你曾为之掏腰包的类似的食品摊，并数一数在饕餮之后的那个夜晚猛灌了多少次碱式水杨酸铋（Pepto-Bismol）①，然后再得出一个数值估计量；你只是让直觉来完成这个判断。但20世纪50年代与60年代早期的研究表明，这种情况下，人们对随机性的直觉常常会导致失败的结果。卡尼曼思忖道，对于不确定性的误解到底有多广泛呢？对于人类决策而言，这又意味着什么？几年后，卡尼曼邀请了一位副教授同事特沃斯基，在其研究小组的学术例会上作了一次来宾讲座。然后在午餐时，卡尼曼向特沃斯基提到了他所酝酿的想法。之后的30年中，特沃斯基和卡尼曼发现，即使是久经世故的被试者，当涉及随机性过程——不论是军事或体育比赛中的形势、生意上的窘境或是医疗方面的疑问时，人们的信念和直觉都常常有负所托。

设想你所写的一部关于爱情、战争以及全球变暖的惊险小说，在4个出版商那儿都被退了稿。直觉以及胸口堵得慌的感觉大概在说，这些出版业专家的拒绝，只能意味着手稿一无是处。但这直觉是否正确？你的小说真的卖不出去吗？根据经验，人人都知道即使连着几次扔硬币都是正面朝上，也并不意味着硬币的两面都是正面。出版物的成功是否会如此难以预料，以至于哪怕我们的小说注定要登上畅销书榜，却仍然有许多出版商不能发现它的闪光点，并寄出那些写满了“谢谢”和“不，谢谢”的信？在1950年代，有本书被多个出版商退稿，并得到了诸如“非常乏味”、“一部典型的家庭内部斗嘴皮、微不足道的烦恼以及青春期情感的沉闷记录”以及“即使本书在5年前趁着它的主题（第二次世界大战）还没有过时的时候面世，我也看不出它有任何机会”之类的评论。但安妮·弗兰克（Anne Frank）的这本《安妮日记》（*The Diary of a Young Girl*）

① 一种可用于治疗急、慢性腹泻的药物。

至今已卖出了3000万册，并因此成为有史以来最畅销的书籍之一。退稿信同样寄给了希尔维亚·普拉斯（Sylvia Plath）①，因为她“肯定没有足够的真正的天分来引起我们的注意”；将乔治·奥威尔（George Orwell）② 的《动物庄园》（*Animal Farm*）退稿的原因则是“动物故事在美国没法卖出去”；而由于“又是波兰和富有的犹太人”，以撒·巴什维斯·辛格（Isaac Bashevis Singer）③ 也被退了稿。而走红之前的托尼·席勒曼（Tony Hillerman）④ 被代理商所抛弃，并被建议“把所有那些印第安人的玩意儿都去掉”。[7]

这些并非孤立的误判。实际上，许多日后取得巨大成功的书，都不得不从退稿的经历中熬出头来——不是一次退稿，而是反复多次的退稿。举例来说吧，相比于约翰·格里森姆（John Grisham）⑤、西奥多·盖泽尔（Theodore Geisel，瑟斯博士）⑥ 和罗琳（J. K. Rowling）⑦ 的作品，现在大概没几本书会具有同样明显而全方位的吸引力。但在成名之前，那些手稿都遭到了一次次地退稿。格里森姆的《杀戮时刻》（*A Time to Kill*）被26个出版商退了稿，而他的第二部作品——《公司》（*The Firm*）——似似是在一个私抄本流传于好莱坞并得到60万美元的电影版权费后，才引起了出版商的兴趣。瑟斯博士的第一部儿童读物《想起我在桑树街见过它》

---

① 希尔维亚·普拉斯（1932～1963）是美国女诗人、小说家与儿童作品作家。

② 乔治·奥威尔（1903～1950）原名埃里克·亚瑟·布莱尔（Eric Authur Blair），英国作家，代表作为《动物庄园》和《1984》。

③ 以撒·巴什维斯·辛格（1902～1991）是出生于波兰的美国作家，曾获诺贝尔文学奖。

④ 托尼·席勒曼（1925～2008）是美国侦探小说家，因撰写纳瓦霍部落（美国最大的印第安部落）警察系列神秘小说而著名。

⑤ 约翰·格里森姆（1955～）是美国畅销小说作家，其多部小说被改编为电影。

⑥ 西奥多·盖泽尔（1904～1991）是美国作家与卡通画家，笔名瑟斯博士（Dr. Seuss），著有多部儿童读物。

⑦ J·K·罗琳（1965～）是英国作家及系列魔幻小说《哈利·波特》的作者。



(*And to Think That I Saw It on Mulberry Street*) 被 27 个出版商退稿。而罗琳的《哈利·波特》(*Harry Potter*) 第一部的手稿经历了 9 次退稿[8]。相对于这些最终成功的人，则是这个行当中谁都再清楚不过的硬币的另一面：那许许多多潜力无穷却永未成功的作者，那些在最初 20 次退稿后就封笔的格里森姆们，或在最初 5 次退稿后就放弃的罗琳们。这其中的一位——约翰·肯尼迪·图尔 (John Kennedy Toole)① ——在多次退稿后，不再指望作品能得以出版，并选择了自杀。他的母亲留住了手稿。11 年后，《笨伯联盟》(*A Confederacy of Dunces*) 出版了。它赢得了普利策小说奖，并售出了近 200 万本。

一部伟大小说的创作——或说一件珠宝或一块洒上了巧克力碎块的曲奇饼，和这部小说即刷成书后在数以千计的零售店前高高堆起的书堆——或说成套的珠宝或成袋的曲奇饼，这两者间有着一道随机性和不确定性的鸿沟。之所以各领域的成功人士都几乎无例外地属于特定的一类人——那些从不放弃的人，其原因就在于此。

许多发生在我们身上的事情——职场上、投资上以及在生活中或大或小决策方面的成功——都是随机因素的结果，这些因素的影响一点儿也不比我们的能力、勤奋与为把握机会而做好准备的影响来得小。因此，我们所认识的现实，并不是人或环境的直接反映，而是被不可预见或不断变化的外部力量所随机化后的模糊映像。这并非说能力无关紧要——能力正是增加成功机会的因素之一——但行动与结果之间的联系，可能并非如我们所乐于相信地那么直接。因此，理解过去不容易，预测将来同样不容易。在这两种情况下，如果能超越肤浅的解释去观察问题，我们将受益匪浅。

我们总是习惯性地低估随机性的影响。股票经纪人推荐投资拉

---

① 约翰·肯尼迪·图尔 (1937～1969) 是美国小说家。

美共同基金，因为在 5 年时间里，它的收益“把国内基金赢得屁滚尿流”；医生则将三酰甘油水平的上升归咎于我们的新习惯——在尽责地喂给孩子们一顿芒果和脱脂酸奶的早餐后，享用一块搭配牛奶的 Hostess Ding Dong 蛋糕。人们可能会也可能不会接受股票经纪人或医生的建议，但却很少有人质疑这些建议是否有足够的数据支持。在政治、经济和商业世界中，甚至当整个职业生涯和数百万美元都被孤注一掷时，偶然事件也常常被明显地错误解释为成功或失败。

好莱坞就是一个好例子。好莱坞游戏中的赏（与罚）是否确有所值？或者在票房的成（与败）中，运气扮演的角色比人们所认为的还来得重要得多？我们都理解天赋并不保证成功，但认为成功必然来自天赋则颇有诱惑性。不过，无人能预知一部影片是否能够走红的这个观点，在好莱坞一直是个让人不快的疑虑，至少自小说家与剧本作家威廉·高曼（William Goldman）在 1983 年的经典之作《银屏交易中的冒险》（*Adventures in the Screen Trade*）中阐明以来就是如此。高曼在书中引用前电影监制大卫·皮克尔（David Picker）的话说：“如果我对所有被我毙掉的片子说‘行’，而对接受的片子说‘不行’，最终结果可能也差不多。”[9]

这并不是说一部家庭作坊制作的神经兮兮的恐怖电影会与——比如说——耗费了 8000 万美元的《驱魔人前传》（*Exorcist: The Beginning*）同样容易成为票房明星，虽然这种事情实际上在若干年前就发生过：《女巫布莱尔》（*The Blair Witch*）只花了制片人仅仅 6 万美元，却带来了 1.4 亿美元的国内票房收入，是《驱魔人前传》的 3 倍有余。不过高曼所谈论的并不是这个。他所谈及的只限于专业制作的好莱坞影片，并且是好得足以由著名发行商来发行的影片。高曼并不否认一部电影的票房表现受到多种因素的影响，他真正说的，是这些因素如此复杂，而一部电影从被开绿灯开始制作到周末首映的路途，又是如此容易受到不可预见和不可控制因素的攻击，因此，根据各种因素来对一部尚未杀青的电影之潜力进行预测，其

实并不比靠扔硬币来猜测要好多少。

不可预测性的例子在好莱坞并不难找。影迷们应该都还记得电影公司对于百万巨制《伊斯达》（*Ishtar*）[沃伦·比蒂（Warren Beatty）＋达斯汀·霍夫曼（Dustin Hoffman）＋5500 万美元预算＝1400 万美元票房收入]和《幻影英雄》（*Last Action Hero*）[阿诺·施瓦辛格（Arnold Schwarzenegger）＋8500 万美元＝5000 万美元]所抱有的期待。另一方面，我们大概还能回想起环球影城（Universal Studios）的主管们对年轻导演乔治·卢卡斯（George Lucas）的《美国风情画》（*American Graffiti*）所持的悲观看法。该片的拍摄只花了不到 100 万美元，而尽管遭受了种种怀疑，它却带来了 1.15 亿美元的收入。不过这个成绩仍不能阻止主管们对卢卡斯的下一个点子——他自己称之为“编自《威尔斯航行日志》的歼星者路克历险记”的故事——怀有更为悲观的疑虑。按他们的话说，这部片子根本就无法制作。最终 20 世纪福克斯（20th Century Fox）拍摄了影片，但公司对该片的信心也仅仅到此为止：公司付给卢卡斯区区 20 万美元作为编写剧本和导演的酬劳，而作为交换，卢卡斯得到了影片的续拍和销售权。预算为 1300 万美元的《星球大战》（*Star Wars*）最终带来了 4.61 亿美元的收入，而卢卡斯则赢得了一个帝国。

为一部电影开绿灯放行的决定，一般比影片制作完工要早个几年，而在这些制作和营销的年头中，影片又会受到许多不知哪里冒出来的不可预测因素的影响，至于观众那难以揣摩的口味就更不必说了。考虑到这些事实，高曼的理论就一点儿也不牵强了（该理论同样也得到了经济领域中更近期研究的支持[10]）。尽管如此，电影公司行政主管们获得的评价，却不是根据其实际管理能力来做出的，虽然这些能力对于美国钢铁公司（United States Steel Corporation）的头头或派拉蒙（Paramount Pictures）的头头而言，都同样至关重要。实际上，对电影公司的行政主管来说，他们得到的评价取决于

是否能够选中票房热门。如果高曼是正确的，那这种能力实际只是个错觉，而且无论这些主管们如何大吹特吹，他们谁都配不上那2500万美元的合同。

要确定在最终结果中到底有多少应归功于能力，又有多少应归功于运气，这可不是个不动脑子就能获得答案的问题。随机事件的出现常常像一盒麦片中的葡萄干那样，一群群、一溜溜、一簇簇地。尽管幸运女神会公平地赐予潜能，但在实际结果上可就算不得公平了。这就意味着，如果让10个好莱坞主管每人扔10次硬币，尽管他们都有同样的机会成为赢家或输家，但最终总是会有确定的赢家和输家。在这个扔硬币的例子中，至少有一个人能扔出8次或更多的正面或反面的可能性为2/3。

假设卢卡斯拍了一部新的《星球大战》电影，然后决定在一个测试市场上做个疯狂的实验。他用两个不同的名字——《星战前传A》和《星战前传B》——来发行该片完全相同的两个拷贝，每部都有各自的市场营销计划和发行安排，并且除了在预演公告和广告中，一个说的是《星战前传A》而另一个说的是《星战前传B》之外，这些计划和安排的细节都完全相同。让这“两”部影片进行竞争，哪一部会更受欢迎呢？比方说，我们观察最初的2万名电影爱好者，并记录下他们所选择的影片（我们忽略掉那些铁杆影迷，他们会两部都看，并坚持认为两部影片中有一些微妙但颇有意义的差别）。由于电影本身以及营销策略都完全相同，因此，可以数学化地对这个实验建立如下模型：将这些观众排成一行，然后依次让他们扔硬币，扔出正面就看《前传A》，否则就看《前传B》。由于这枚硬币正面或背面朝上的机会相等，你可能会认为，在这场试验性的票房战争中，每部电影都应该有大概一半的时间在比赛中领先。随机性的数学说的却是另外一回事：最可能出现的领先权交替的次数是0次，即在这场争夺2万名顾客的比赛中，两部影片中的某一部会从头到尾都保持领先，并且这个可能性是领先权不停交替的可能性的88

倍[11]。我们得到的教训并不是说所有影片都毫无差别，而是说哪怕所有的影片都分毫不差，也总会有某些影片比别的影片有更好的票房。

不管是在好莱坞还是在其他什么地方，公司会议室里的讨论中都不会提及上面这些问题。因此，随机性的典型模式——表面上的“热”或“冷”的曲线，或是一簇一簇的数据——都照例被错误地加以解释。更糟糕的是，这些模式成为了下一步行动的依据，好像它们真的体现了某个新趋势似的。

雪莉·兰辛（Sherry Lansing）就是现代好莱坞历史上任命和开除都最为高调的例子之一。兰辛掌管派拉蒙多年，并取得了巨大的成功[12]。在她任职期间，派拉蒙靠《阿甘正传》（*Forrest Gump*）、《勇敢的心》（*Brave Heart*）以及《泰坦尼克》（*Titanic*）夺得了最佳影片奖，并诞生了收入最高的两年。然后，兰辛的名声就一落千丈，而在派拉蒙经历了如 *Variety* 杂志所说的“很长时间的不佳票房”[13]后，她终于被抛弃了。

利用数学语言，可以对兰辛的命运给出一短一长的两个解释。先来看看短的解释。请看下面这串百分数：11.4、10.6、11.3、7.4、7.1、6.7。发现了什么吗？是的，兰辛的老板萨姆纳·雷德斯通（Sumner Redstone）也发现了。对他而言，趋势已经足够明显，因为这6个数字给出了派拉蒙的动作影片组在兰辛就职最后6年中的市场份额。这个趋势使得《商业周刊》（*Business Week*）怀疑兰辛“大概就是不再拥有好莱坞的‘热手’了”[14]。不久，兰辛宣布她将离开，而几个月后，一个名叫布拉德·格雷（Brad Grey）的有天分的经理被提升到了台面上。

一名无疑是颇有天赋的人，怎么可能在带领公司走过7个了不起的年头后，事实上是在一夜之间就突然走了麦城呢？有大量的理论可以解释兰辛早期的成功。当派拉蒙干得不错时，兰辛受到了赞扬，因为她造就了好莱坞营运最佳的公司之一，而且还有着将老套

故事转化为 1 亿美元票房的点石成金术。而当她的运气变糟后，修正主义者们就占了上风。兰辛那重拍老片或拍摄续集的倾向，现在成了一种缺点。最毒辣的评论，大概是将其失败归因于她那“半吊子品味”的说法。如今，因为赞成拍摄《时间线》（*Timeline*）和《古墓丽影 2》（*Tomb Raider：The Cradle of Life*）这样的票房毒药，她饱受指责。兰辛反对冒险、观念过时、与潮流脱节之类的说法突然就成了正统。但是，相信迈克尔·克莱顿（Michael Crichton）的畅销书能为一部大有希望的影片提供素材，这也应当受到批评么？而当第一部《古墓丽影》带来 1.31 亿票房的时候，所有那些关于劳拉[①]的批评又在哪里？

即使有关兰辛缺点的那些理论确有道理，可她的下坡发生得多么突然呀！是因为她一夜之间突然变得反对冒险和脱离潮流吗？不，是因为派拉蒙的市场份额突然下降了。头一年兰辛还高处云端，下一年她就成了午夜剧场喜剧演员们用来制造连珠妙语的笑料。如果她跟好莱坞的其他人一样，在肮脏的离婚案中变得沮丧，或者因为挪用财物被起诉，或者加入了某个邪教组织，则这种改变多少还能理解。但情况并非如此，而她显然也没有遭受什么大脑皮质损伤。实际上，兰辛的批评家们能提供的说明其失败原因的证据，就只有这些失败本身。

事后来看，兰辛遭解雇很显然是由于业界对于随机性的误解，而不是由于她那些有瑕疵的决策：当兰辛离开派拉蒙时，公司次年的影片已在出品过程中，因此，如果我们想大致知道，在一个兰辛

① 《古墓丽影》系列游戏与电影的女主人公。

得以留任的平行宇宙①中她做得如何，只需看看她离职次年的数据就可以了。得益于《世界大战》（*War of the Worlds*）和《最长的一码》（*The Longest Yard*）等影片，派拉蒙收获了10年中最好的一个夏天，市场份额也回弹至了将近10%。这不仅仅是讽刺——这再次体现了随机性那回归均值的一面。一篇*Variety*头条对此写道："分手的礼物：老领导的影片为派拉蒙注入反弹动力"[15]，但人们仍然会不自禁地想到，如果Viacom（派拉蒙的母公司）能多一点点耐心，这个头条可能就会变成："标志性的一年将派拉蒙和兰辛的职业生涯拉回正轨"。

兰辛开始时好运当头，结束时则厄运缠身，不过，她的运气还有可能更糟：她可能一开始就霉运罩顶。这就发生在了哥伦比亚影业（Columbia Pictures）一位名叫马克·坎顿（Mark Canton）的主管身上。刚被聘用时，他被说成是对票房有着透彻了解和充满激情的人，但在最初几年的票房成绩令人失望后，他被解雇了。一位不愿透露姓名的同事评价他"没有能力分辨成功和失败的影片"，而另一位则称他"过分忙碌于拉拉队表演"。这个丧失名誉的人离去时，在制片流水线上留下了《黑衣人》（*Men in Black*）（5.89亿美元的全球票房）、《空军一号》（*Air Force One*）（3.15亿美元）、《第五元素》（*The Fifth Element*）（2.64亿美元）、《甜心先生》（*Jerry*

---

① 平行宇宙是为了解决量子物理中的"测量难题"而提出的一个假设。量子测量难题可以通过"薛定谔猫"来加以说明：设想有一只猫与杀猫剂一道置于一个密闭的空间中，同时有一个量子过程（例如一个电子到达靶后落到靶的左边还是右边）以五五开的概率使杀猫剂被释放出来，那么根据量子理论，在没有进行测量前，猫就处于一种半死半活的状态，但一旦进行了观察，那么猫要么死了，要么活着。为了解决这个看来十分荒谬的问题，平行宇宙作为一种假设被提了出来，它认为所有可能的量子世界都同样实在且平行存在，一旦进行了测量，则宇宙就产生分叉一分为二，其中一个宇宙中的猫死了，而另一个宇宙中的猫还活着，两个世界都同样实在，但每个世界的居住者都智能察觉到他们所处的那个宇宙。（见《上帝与新物理学》，保罗·戴维斯著，徐培译，湖南科学技术出版社，1992年。）

Maguire）（2.74 亿美元）和《狂蟒之灾》（*Anaconda*）（1.37 亿美元）这样的影片。正如 *Variety* 所说的，坎顿的遗产“红得发紫”[16]。

瞧，这就是好莱坞，一个迈克尔·奥维兹（Michael Ovitz）作为迪斯尼的总裁在此工作了 15 个月后、带着 1.4 亿美元的解聘红包离开的小城，一个哥伦比亚影业的头头戴维·伯格尔曼（David Begelman）因伪造罪与挪用公司财产而被解雇、却又在几年后被聘为米高梅（MGM）CEO 的地方。我们将在后面的章节中看到，这些毒害了好莱坞的误判，同样也毒害着人们对生活各方面的理解。

我自己对于随机性那隐形影响的领悟，产生于大学时代。当时我选修了一门概率论的课程，并开始将其原理应用于体育比赛上。这事儿做起来不难，因为如同在电影业中一样，大多数体育运动的成就都很容易量化，数据也都是现成的。我所发现的是，正如在体育比赛中所学到的持之以恒、勤学苦练和团队合作，也能应用于生活中的各种努力一样，在体育比赛中学到的关于随机性的教训同样如此。因此我接下来要讲一讲棒球手罗杰·马立斯（Roger Maris）和米奇·曼托（Mickey Mantle）的故事。这个故事对所有人来说都包含着深刻教训，哪怕他们连棒球和乒乓球都分不清。

那是在 1961 年，我才刚刚到读书的年纪。但我仍记得《生活》（*Life*）杂志封面上的马立斯，以及他那更受欢迎的纽约扬基队队友曼托的面孔。这两名棒球选手在当年投入了一场历史性的竞争，以追平或打破 1927 年贝比·鲁斯（Babe Ruth）创下的 60 个本垒打的心爱纪录。那是一个唯心主义大行其道的时代，我的老师会讲一些诸如“我们需要更多像贝比·鲁斯那样的英雄”或“我们从来没有过不老实的总统”之类的话。由于鲁斯的传奇已被神圣化，所以任何可能对他形成挑战的人，最好都能配得上鲁斯之挑战者的这个名头。勇敢而毫不懈怠、承受着膝伤坚持战斗的击球手曼托，是当年

的球迷以及记者们毋庸置疑的最爱。曼托长相讨人喜欢，脾气性格也好，给人的印象就是那种人人都希望由他来创造纪录的彻头彻尾的美国男孩。另一方面，马立斯却是个粗暴而神秘兮兮的家伙，一个从未在一年中打出过多于 39 个本垒打的失败者，而 39 个本垒打离 60 个本垒打根本还差了十万八千里。他被认为是个讨嫌的人，那种从不接受采访也不喜欢孩子的人。在这场创纪录大战中，球迷和记者都坚决站在曼托一边——我喜欢马立斯。

最终，尽管曼托受伤的膝盖尽了全力，他却只打出了 54 记本垒打。马立斯依靠 61 个本垒打打破了鲁斯的纪录。在鲁斯的整个职业生涯中，曾有 4 次在一个赛季中击出了 50 记或更多的本垒打，并 12 次摘得联赛本垒打之王的桂冠。而在创造了纪录之后，马立斯再也没有击出 50 甚至是 40 个本垒打，也再没有在联赛中占过鳌头。这样的总体表现无疑给愤愤不平的情绪火上浇油。后来的岁月中，马立斯不断遭到球迷、体育评论员以及有时来自其他球员的无情批评。他们的判决是：成为冠军的压力把他彻底压垮了。一位著名的棒球老前辈说道："马立斯没有权力去打破鲁斯的纪录。"[17]这句话也许是真的，但原因却并非如这位老前辈所想。

多年以后，受到那门大学数学课的影响，我开始学着从另一个角度来考虑马立斯的成就。为了分析这场"鲁斯-曼托竞赛"，我重新阅读了《生活》的那篇旧文，并在其中发现了一个有关概率论以及如何用它来预测竞赛结果的简短讨论[18]。我决定建立一个我自己的本垒打数学模型如下。任何一次击球（也就是一次通往成功的机会），其结果当然主要依赖于球员的能力，但同样也依赖于许多其他因素的相互作用：球员的健康状况、风、太阳或球场灯光、投球的质量、比赛形势、是否正确猜出了投手投球的方式、挥棒时手眼是否协调、在酒吧遇到的那位浅黑皮肤的姑娘是否让他睡得太晚或早餐吃的辣奶酪热狗和蒜香薯条是否把肚子弄坏了，如此等等。如果没有这些不可预测的因素，球员要么每次击打都打出本垒打，要么

每次都打不出。实际上，对于每次击球，我们只能说这名球员有一个确定的概率击出本垒打，也有一个确定的概率不能击出本垒打。球员每年获得的数百次击球机会，通常都将随机因素平均掉了，最终得到的是某个典型的本垒打产量，它随着球员技术的提高而增加，并不可避免地因为那个在他英俊面庞上刻出皱纹的过程①的影响而下降。但有时这些随机因素并没有被完全平均掉。这种情况发生得有多频繁呢？它所造成的偏差又有多大呢？

根据球员每年的统计数据，我们可以估计出他每次走到本垒[19]前击球时，能打出本垒打的概率。在 1960 年，也就是马立斯创纪录之年的前一年，他每 14.7 次击球中能击出 1 记本垒打（与他表现最佳 4 年的总平均值差不多）。让我们称这个水平为“正常马立斯”。我们可以用下面的方法，来为正常马立斯的本垒打技术建模：设想有一枚硬币，平均抛出 14.7 次（而不是 2 次）时，就有 1 次是正面朝上。当马立斯每次上垒击球时，扔一下这枚硬币，如果扔出正面，就给他记上一次本垒打。以这种方式来进行马立斯的——比方说—— 1961 赛季的比赛的话，只要在那个赛季中马立斯每次获得击球机会时，都扔一下这枚硬币，扔硬币的结果就能给出一大堆五花八门的“1961 年赛季”。在这些“赛季”中，马立斯的技术水平，与“正常马立斯”的本垒打数量是相称的，而这些模拟赛季的结果，就能说明当技术没有突然急剧提高时，“正常马立斯”在 1961 年所能指望达到的成绩范围　　即他的“正常”本垒打能力加上纯运气因素所能达到的成绩范围。

要真真正正地完成这项实验，我得需要一枚相当奇怪的硬币、一个非常强壮的手腕以及一段时间的脱离大学岗位的工作。不过，随机性的数学理论使我能利用方程与计算机来完成这个分析。在许多个虚拟的 1961 赛季中，正常马立斯的本垒打数量并不令人吃惊地

---

① 指岁月的流逝。

落在了一个对他而言是正常的范围之内，在某些模拟赛季中他打中得多一点，在某些中打中得少一点；而在一个赛季中打中了比正常数量多得多或少得多的本垒打的情况则很罕见。那么，以马立斯的天赋却获得鲁斯式的结果，这样的情况有多常见呢？

我本以为马立斯追上鲁斯纪录的可能性，大概跟杰克·惠特克（Jack Whittaker）那次中彩票的机会差不多：几年前他在一个便利店买早餐饼干时，多扔了 1 美元买了张全国“强力球”彩票，最终却赢得了 3.14 亿美元①。对于一个天分更差些的球员而言，追上鲁斯纪录的机会跟这差不多，但正常马立斯尽管没有鲁斯的天赋，其本垒打水平仍比平均水平高出许多，因此，仅仅靠偶然性而创造纪录的机会，对正常马立斯而言绝非可以忽略不计：差不多每 32 个赛季中，他就能在 1 个赛季追平或打破鲁斯的纪录。这概率听起来似乎不是很高，估计也没有人会对马立斯这个特定的人或 1961 年这特定的一年下注，来打赌此年此人将破纪录。但这个概率却能给出惊人的结论。为什么这样说？现在让我们来问一个更有趣的问题：假设所有的球员都具有正常马立斯的天分，再来考察从鲁斯完成纪录到“类固醇时代”（在这个时代中，由于药物的使用，本垒打与以前相比变得非常寻常）开始时，这之间所有 72 年的时间。在这个时期内，仅仅由于运气，某个球员在某个时候追平或打破鲁斯的纪录，其可能性有多少呢？相信马立斯仅仅只是凑巧成为那个异常赛季的幸运得益者的观点，是否真的合理呢？

历史表明，在这段时期内，每 3 年中，大概就有 1 名球员兼具与 1961 年的正常马立斯差不多的天赋和运气。总的来说，仅仅依靠运气，这些球员中的某一个能够追平或打破鲁斯的纪录的可能性，

---

① 惠特克是在 2002 年 12 月 25 日中奖的，这也是当时美国彩票史上单张彩票的最高中奖金额。但财富似乎并没有带来更多的好运。之后惠特克不断被卷入法律纠纷、盗窃与家庭悲剧之中。

要比50%多一点。换句话说，在70年的时间里，一名实力更接近于40记本垒打的球员产生出一个60记或更多本垒打的随机峰值，这个可能性是比较靠得住的——这个现象有点像在使用糟糕的线路进行通话时，在背景杂音中所不时听到的大声噼啪声。而同样靠得住的是，不管那个破纪录的“幸运儿”是谁，人们仍将把他神化或贬低——当然还有看不到头的分析。

我们永远不能确切地知道，相比于其职业棒球生涯中的其他年头，1961年的马立斯是一个水平好得多的球员呢，还是个仅仅受到幸运女神垂青的人。但如已故的古尔德（Stephen Jay Gould）① 或诺贝尔桂冠获得者珀塞尔（E. M. Purcell）② 这些卓越的科学家对棒球和其他体育项目进行的细致分析表明，我所描述的扔硬币模型与球员和球队的真实表现吻合得很好，包括他们的巅峰与低谷[20]。

当看到体育——或其他领域——中的辉煌成就时，我们应该记住，非凡的事情可以在没有非凡的原因时发生。随机事件常常看似非随机的事件，而在解释世事时，必须注意不要把这两者给弄混了。尽管花费了许多个世纪，科学家们最终还是学会了透过表面的秩序，去看到隐藏在自然界和日常生活之后的随机性。本章对于随机性运作方式投以了匆匆数瞥；在下面的章节中，我将在随机性的历史发展这一情境里，考察随机性的中心思想，并说明这些思想与我们目标——提供一个观察日常环境的新视角，并更好地理解随机性这一自然界基本特征与我们自身体验间的联系——之间的关联。

---

① 斯蒂芬·杰·古尔德（1941～2002）是美国的史前生物学家、进化论生物学家和科学史学家，同时还是一名著名的科普作家。

② 爱德华·米尔斯·珀塞尔（1912～1997）是美国物理学家，因发现磁共振而获得1952年诺贝尔物理学奖。

第二章

真理与半真理的法则

仰望晴朗无月的夜空，人的肉眼能看到数以千计的闪烁光源。随意散布的星星中，有着许多模式：这里有一头狮子，那里是一个水斗。发现模式的能力既能是一种力量，也能是一种弱点。牛顿（Isaac Newton）思索物体下落的模式，创造了万有引力定律；另一些人则发现，当他们穿着某双脏袜子时，运动成绩会明显提高，所以从此以后他们就不再穿干净的袜子。在自然界所有这些模式中，如何去分辨出真正有意义的那一些呢？本质而言，这个区分是一个实践问题，因此你也许不会太过吃惊地发现，与几何这门由一帮榆木脑袋的哲学家们所创造的整套公理、证明和定理中诞生的学科不同，随机性的理论却萌发自那些专注于咒语和赌博的头脑之中。在这幅场景中，我们更容易想到的是手中的骰子或魔法药水，而不是书本或卷轴。

根本而言，随机性理论就是常识的成文化。但它同时也是一个充满微妙之处的领域，一个伟大的专家们犯下了著名的错误、而专业赌徒们却得到了声名狼藉的正确结论的领域。要理解随机性并克服我们的错误概念，需要经验和大量细致的思考。因此，我们将以概率论的若干基本法则以及揭示、理解和应用这些法则时所面临的挑战，来作为本书旅途的起点。对概率法则的直觉观念所进行的经典探索之一，就是由卡尼曼与特沃斯基这一对在阐明人们错误概念方面进行了许多工作的专家所进行的实验[1]。请尽管来加入这个实验——并了解一些你自己的概率直觉。

设想一位名叫琳达的女性，31 岁，单身，坦率直言，非常聪明。大学时主修哲学，其间非常关心歧视与社会公正问题，并参加了反核游行。在对 88 名被试者进行了这样一番描述后，特沃斯基与卡尼曼让他们根据自己认为的可能性，以 1 分（最可能）到 8 分（最不可能）来给下列陈述评分。按最可能到最不可能的顺序排序后，结果如下：

陈　述	平均可能性评分
琳达在女权运动中表现活跃	2.1
琳达是一名治疗精神病的社会工作者	3.1
琳达在一个书店中工作，而且还参加了瑜伽班	3.3
琳达是一名银行出纳，并且在女权运动中表现活跃	4.1
琳达是一名小学教师	5.2
琳达是女选民联盟的一名成员	5.4
琳达是一名银行出纳	6.2
琳达是一名保险推销员	6.4

乍看之下，结果中似乎没什么不寻常的地方：对琳达的描述，实际上就是根据一名典型的女权主义者而非一名银行出纳或保险推销员来设计的。但现在让我们将注意力集中在其中的 3 种可能性及其得分上，并按最可能到最不可能的顺序排列如下。这也是 85％的实验结果中给出的顺序：

陈　述	平均可能性评分
琳达在女权运动中表现活跃	2.1
琳达是一名银行出纳，并且在女权运动中表现活跃	4.1
琳达是一名银行出纳	6.2

如果这些结果看上去也没啥好奇怪的话，那就说明卡尼曼和特沃斯基成功地愚弄了你，因为琳达是一名银行出纳并且在女权运动中表现活跃的可能性，比琳达是一名银行出纳的可能性还要大，这就违背了有关概率的第一条——也是最基础的一条——定律：两个事件同时发生的概率，永远不会比各事件单独发生的概率大。为什么呢？这只是简单的算术：事件 A 发生的概率＝事件 A 和 B 都发生的概率＋事件 A 发生而事件 B 不发生的概率。

卡尼曼和特沃斯基并未对此结果感到惊讶，因为被试者面对的是一长串的可能性，当这些可能性被打散后，上面 3 个陈述之间的联系就很容易被人遗忘。因此，卡尼曼和特沃斯基向另一组被试者描述了琳达的情况，但只使用下面 3 个陈述进行打分：

琳达在女权运动中表现活跃

琳达是一名银行出纳，并且在女权运动中表现活跃

琳达是一名银行出纳

令他们惊讶的是，87％的被试者仍然认为琳达是一名银行出纳并且在女权运动中表现活跃的可能性，比琳达是一名银行出纳的可能性要大。因此，两位研究者更进了一步：他们明确地要求 36 名还算老练的研究生，在了解了概率第一定律之后，再来考虑他们的回答。即便有这样的提示，仍有两名被试者执着于那个不合逻辑的答案。

在这个顽固的错误中，卡尼曼和特沃斯基注意到了一件有趣的事情：如果问题与被试者所了解的琳达的情况没什么关系，同样的错误就不会发生。例如，假设卡尼曼和特沃斯基问的是下列陈述中哪一个最可能：

琳达拥有 IHOP（International House of Pancakes）① 的特许经营权

琳达做了变性手术，现在名叫拉里

琳达做了变性手术，现在名叫拉里，并且拥有 IHOP 的特许经营权

这时几乎没人会认为最后一个陈述比其余两个更可能。

卡尼曼和特沃斯基由此得出结论认为，由于开始所给的琳达的特征描述使“琳达在女权运动中表现活跃”这一细节听起来很正确，因此当这一细节与银行出纳的猜想加在一起时，整个陈述的可信度就增加了。但在琳达的嬉皮岁月和她生活于我们星球上的第四个 10 年之间，可能会发生许多事情。她可能会皈依某个原教旨主义的宗教派别，可能与某个剃着平头的男人结了婚并在左边屁股上纹了一个纳粹十字，或者因为忙于生活中的其他方面而不再积极于政治活动。这些情况——以及许多其他情况　　中的每一个都可能使她不再是女权运动的活跃分子。因此，加入的细节实际上减少了整个陈述发生的可能性，哪怕这个细节看来似乎使得整个描述更显精确。

如果有关某事的细节能契合到人们对此事认知的心理图景之中，那么描述中的细节越多，它看来就越真实，而人们也就会认为它越可能发生——尽管任何不确定性细节的加入，都只会使总的复合描述变得更不可能。概率的逻辑与人们对不确定事件的评价之间的这个矛盾，令卡尼曼和特沃斯基大感兴趣，因为在现实中，这可能会导致不公平或错误的评判。一名被告在发现尸体后离开了犯罪现场，或一名被告在发现尸体后因为他担心因这宗令人毛骨悚然的谋杀案被起诉离开了犯罪现场，两者中哪一个更可能呢？总统将增加对教

---

① IHOP（国际松饼屋）是美国的一家连锁餐馆，主营早餐食品。松饼有点类似于我们国内的煎饼，口感柔软。

育的联邦资助力度，或者总统将通过取消其他资助所节余的资金来增加对教育的资助力度，前者是不是比后者更可能？你的公司明年将增加销售，以及你的公司由于经历了一个丰收年而将在明年增加销售，前者是不是比后者更可能？所有这些情况中，尽管后者的发生概率总比前者小，但它们听起来都显得更为可能。或者如卡尼曼和特沃斯基所说："一个好的故事常常比一个不那么令人满意的（解释）更不可能。"

卡尼曼和特沃斯基还发现，即使是受过高度训练的医生，也会犯同样的错误[2]。他俩给一群实习医生叙述了一个严重的医学问题：肺部血栓（在肺部的血液凝块）。如果人得了这个小毛病，就可能出现一系列症状中的一个或多个。某些症状——如半身不遂——并不常见，而另一些——如呼吸急促——则很可能发生。下列情况中，哪一个更可能发生呢？是血栓患者只表现出半身不遂，还是患者同时表现出半身不遂并且呼吸急促？卡尼曼和特沃斯基发现，91%的医生相信，血栓仅仅导致一种罕见症状的可能性，要小于导致一种罕见症状加一种常见症状复合的可能性。（这些医生为自己辩解道，患者并不会走进诊室，然后说出"我的肺里有个血栓，猜猜我有啥症状"之类的话）。

多年后，卡尼曼的一名学生与另一名研究者一道，发现律师在进行判断时，也沦为了相同偏见的受害者[3]。不论是刑事案或民事案，当事人通常都依靠律师来了解进入审判程序后可能发生的情况。无罪或调解或进行数额不等的赔偿，各自的可能性有多少？尽管律师不会用具体的概率值来表达意见，他们却会根据对各种结果相对可能性的个人预计来提供建议。研究者在此同样发现，律师认为那些细节更丰富之联合表述的概率更高。在葆拉·琼斯（Paula Jones）对时任总统的克林顿的民事诉讼①中，200 名执业律师被请求预测案

① 1994 年，琼斯指控克林顿对其进行了性骚扰。

子不会走完全部司法程序的可能性。对某些被试者，诉讼提前结束的可能性被分解为各种具体的原因，例如调解、撤回指控或被法官驳回，而对其他被试者，则只要求他们简单地给出审理是否完成的预测。在对这两个组进行比较时，研究者发现，相比于另一组，那些被提供了案子提前结束之具体原因的被试者，更倾向于预测审理将早早结束。

评判环境中各种现象间有意义的联系，这种能力可能是如此重要，以至于在某些情况下，即使追寻的是海市蜃楼般的幻影，却也仍然值得。假设一个饥饿的穴居人看到了远处岩石上一点模糊的绿色，那他可以置之不理而错过一条肥美可口的蜥蜴，也可以跑去突袭却发现那实际上只是几片稻草叶，但前种做法显然将导致更大的代价。因此，根据这个理论，我们可能已进化得通过偶尔犯下后一种错误，来避免犯下后果更严重的前一种错误。

在数学发展的故事中，古希腊人是以现代数学处理方式发明人的面貌出现的，这种方式就是：公理、证明、定理、更多的证明、更多的定理，等等，等等。但到了 20 世纪 30 年代，捷克裔美国数学家哥德尔（Kurt Gödel）① ——爱因斯坦（Albert Einstein）的朋友——证明这条路子有些缺陷：他证明，大多数的数学必定是自相矛盾的，否则就一定包含了某些无法证明的真理。但数学仍然毫无争辩地按照古希腊的——欧几里得（Euclid）的——方式前进着。古希腊人这些几何天才创造了一个小小的公理集，即那些无需证明就

① 库尔特·哥德尔（1906～1978）是一位逻辑学家、数学家和哲学家。他证明了紧接着给出的哥德尔不完备性定理。关于该定理的科普书籍，有兴趣的读者可以参见克莱因（Morris Kline）的《数学：确定性的丧失》（*Mathematics: The Loss of Certainty*，李宏魁译，长沙：湖南科学技术出版社，1997 年 6 月第 1 版）。哥德尔出生于当时奥匈帝国的 Brünn，即今天捷克共和国的布尔诺（Brno）。因此，也有资料称哥德尔为奥地利-美国籍数学家。

被接受的陈述，并依此证明了许多美丽的定理，详细给出了线、面、三角形与其他几何形状的性质。他们由这些知识认识到　　举例而言——地球是一个球体，甚至还计算出了它的半径。人们一定会想，为什么这样一个文明，它产生了如欧几里得《几何原本》（*Elements*）第一卷之命题 29“一条直线与两条平行直线相交，则所成的内错角相等，同位角相等，且同旁内角的和等于二直角”① 这样的定理，却未能创造一个理论，来证明在扔两个骰子时，把你的 Corvette 跑车押在扔出两个 6 点上是很不明智的做法。

事实上，古希腊人没有 Corvette 跑车，也没有骰子。不过他们确实有赌瘾，还有足够多被宰杀的牲畜，因此，他们玩的是由牲口的踵骨制成的距骨。一个距骨有六面，但落下后只能停在其中的四个面上。现代的学者注意到，由于距骨本身的构造，它停在这四个面上的机会并不均等：停在其中两个面的机会大概各有 10%，而停在另两个面的机会则各有 40%。一个常见的投距骨游戏要使用 4 个距骨，而最好的结果是扔出的四个面都不相同（这被称为掷出了一个维纳斯）。这种情况很罕见，但并非最罕见。扔出维纳斯的可能性约为 384/10000，不过缺少随机性理论的古希腊人并不知道这一点。

在寻求神谕时，古希腊人也使用距骨。提问者能从神谕中得到据称是神所做的回答。历史学家希罗多得（Herodotus）以及如荷马（Homer）②、埃斯库罗斯（Aeschylus）③ 和索福克勒斯（Sophocles）④ 这些作家的记述表明，许多杰出古希腊人的许多重要抉择，都是根据神谕而做出的。但尽管扔距骨在赌博和宗教中的地位如此

---

① 引自兰纪正，朱恩宽译《几何原本》，第 2 版，西安：陕西科学技术出版社，2003，第 26 页。

② 荷马（约前 9 世纪～前 8 世纪）是古希腊盲诗人，整理编写了《荷马史诗》（《伊利亚特》和《奥德赛》）。

③ 埃斯库罗斯（约前 525～前 456）是古希腊悲剧作家，被誉为“悲剧之父”。

④ 索福克勒斯（前 496～前 406）是古希腊悲剧作家。

重要，古希腊人却从未试图去理解其规律。

为什么古希腊人没能发展出概率理论呢？一个答案在于许多古希腊人相信，未来是按神的意志而发展的。如果扔距骨的结果意味着“跟那个在学校木板房后的摔跤比赛中把你牢牢扣住的又矮又壮的斯巴达姑娘结婚”，古希腊人不会觉得这是随机过程中的一个走运（或不走运）的结果，而只会视之为神的意愿。据此观点，通过理解随机性来寻找答案根本就是南辕北辙。因此，对随机性的数学预测似乎就没有出现的可能了。另一个答案大概就恰恰存在于那个使古希腊人成为如此伟大数学家的哲学观念中：他们执着于通过逻辑和公理证明而得的绝对真理，却对不确定的断言大皱眉头。例如，在柏拉图（Plato）《对话录》的《斐多篇》（*Phaedo*）中，西米阿斯（Simmias）对苏格拉底（Socrates）说，“可能性的论据都是骗人的东西”，并指出“除非在使用过程中非常谨慎，否则它们将很容易具有欺骗性——不管是在几何中或是在其他事情上”[4]。这正预见到了卡尼曼和特沃斯基的工作。在《泰阿泰德篇》（*Theaetetus*）中，苏格拉底说，任何“在几何中使用概率和似然性进行论证”的数学家“都不配被称为第一流”[5]。不过，即使古希腊人认为概率学家的确配得上第一流的称号，在没有保留大规模记录的当时，要创造出一个自洽的概率理论，也仍然十分困难，因为，在根据以往来估计事情发生频率——然后就是概率——时，人的记忆力实在是差得可怜。

所有第五个字母为 n 的六字母英文单词，和所有以 ing 结尾的六字母英文单词，哪一个更多？大多数人选择了后者[6]。为什么呢？因为相比于一般性的第五个字母为 n 的六字母单词，人们更容易记起以 ing 结尾的单词。但我们用不着去查《牛津英语字典》（*Oxford English Dictionary*）——甚至不用知道如何数数——就可以证明这个猜测不对：第五个字母为 n 的六字母单词必然包含了所有以 ing 结尾的六字母单词。心理学家称这类错误为易取性偏误，因为在重构过去时，我们对那些最为生动从而也最容易回忆的事物，赋予了无

根据的重要性。

易取性偏误的可恶之处，在于它扭曲了我们对过去以及周围事物的认知，从而偷偷地扭曲了我们对世界的看法。举例来说，人们总倾向于过高估计无家可归者中患有精神疾病的比例，因为当碰到一个行为并不奇怪的无家可归者时，他们并不会注意到此人，也不会跟朋友们谈及这个凑巧遇见的不起眼的人；但如果碰到的是一个边走路边重重跺脚、还唱着《圣者进行曲》并朝着并不存在的同伴挥舞手臂的流浪者，那人们就的确更容易记住这件事情[7]。在杂货店的 5 个收银台中，你排到最费时的那一队，可能性有多大？除非你被黑魔法所诅咒，否则这个可能性约为 1/5。既然如此，为什么当你后来回想时，会觉得是什么超自然的把戏让你选中了这条队伍呢？因为当一切都天下太平时，你只会把注意力集中到其他更重要的事情上，但如果排在你前面的那位购物车里只放了一件商品的女士，一定要搞清楚为什么她买的鸡肉要 1.5 元，而在肉类柜台中的标价却是 1.49 元时，这一定令你印象深刻。

易取性偏误对判断和决策可能造成的影响，在一个模拟审判得到了彻底展示[8]。在这项研究中，某名司机被指控在开车撞上一辆垃圾车时是在醉酒驾驶，而陪审团所得到的证据里面，支持和推翻这个指控的证据分量相等。关键在于一组陪审员得到的无罪证据十分“苍白”“垃圾车车主在质询中承认，他的车是灰色的，因此在夜晚难以被看见。”；而另一组陪审员得到的是一个更“生动”的版本：“垃圾车车主在质询中承认，他的车是灰色的，因此在夜晚难以被看见。车主强调说，他的车之所以是灰色的，是因为‘灰色不显脏。不然你想要我怎么样？把它漆成粉红色？’”有罪证据也同样有两个版本，只不过给第一组的是“生动版”而第二组是“苍白版”。当这些陪审员被要求给出一个有罪/无罪的分值时，说明生动的那一方最终总是胜出。如果还延期 48 小时再裁决的话，这种生动所造成的影响就更加明显（大概是因为此时所要回忆的是更为久远的事情）。

通过扭曲对过去的看法，易取性偏误使任何从既往中获得有意义结果的企图变得复杂。对我们是如此，对古希腊人也是如此。不过在通往古希腊人那可能的随机性理论的道路上，还另有一个重大阻碍，一个非常现实的阻碍：尽管基本概率只需要算术知识，古希腊人却不懂算术，至少不懂那种形式便于使用的算术。举例来说，在公元前 5 世纪，在处于希腊文明顶峰时期的雅典，人们使用字母代码来记数[9]。希腊字母表 24 个字母中的前 9 个代表了我们称为 1 到 9 的 9 个数字；接下去的 9 个字母则代表 10，20，30，…；而最后的 6 个字母加上 3 个额外的符号则代表了前 9 个整百数（100，200，…，900）。如果你现在都还觉得做算术有麻烦的话，试试从 ΩΨΠ 中减去 ΔΓΘ 吧！更糟的是，根本没人在乎这些个位数、十位数和百位数的书写次序：有时百位数被写在最前面，有时又在最后面，有时就根本没什么次序。最后一点：古希腊人没有零。

到亚历山大（Alexander）于公元前 331 年侵入巴比伦帝国时，古希腊人才有了零的概念，但即使在当时，尽管亚历山大已经开始用零来表示一个数的缺失，它也不是作为一个数字被使用的。现代数学中，数字 0 有两个关键性质：在加法中，它是唯一不改变被加数值的数；而在乘法中，它是唯一与任何数相乘却不改变自身值的数。这个概念直到公元 9 世纪才由印度数学家摩诃毗罗（Mahāvīra）所引入。

甚至在一个可用的记数系统被发展起来后，仍需要许多个世纪，人们才开始将加减乘除作为算术的基本运算——并慢慢认识到一个便利的符号系统能使这些运算容易得多。因此，直到 16 世纪，西方世界才真正为概率论的发展做好了准备。不过，在理解随机性方面走出第一步的，却是古希腊人的征服者——罗马人，尽管他们笨拙的计算体系仍然是一个障碍。

罗马人通常对数学不屑一顾，至少对古希腊人的数学是如此。

按罗马政治家西塞罗（Cicero）（前 106～前 43）的话说："希腊人赋予几何学家以最高荣誉；相应的，希腊人所创造的事物中，没有比数学更为辉煌的东西。但作为数学这门艺术的极限，我们确立了它在测量与计数中的实用性。"[10]的确，如果设想古希腊人的教科书注重于抽象三角形之全等性证明的话，典型的罗马教材则注重于如何在敌人占据着对岸时确定河流宽度之类的事情[11]。由于侧重不同，那么对于古希腊人中诞生了如阿基米德（Archimedes）、丢番图（Diophantus）、欧几里得、欧多克索斯（Eudoxus）、毕达哥拉斯（Pythagoras）和泰勒斯（Thales）这样的数学巨星，而罗马人却没能产生哪怕一位数学家的事实，也就不必觉得奇怪了[12]。在罗马文化中，享乐与战争——而非真理和美——占据着中心地位。也正是因为注重实用，罗马人看到了理解概率的价值。因此，一方面看不出抽象几何能有什么用处，另一方面，西塞罗却写道"偶然性正是生活之向导"[13]。

西塞罗或许是古典时期最伟大的概率大师。他用概率来反驳关于赌博的胜利是因于神之干涉这一普遍接受的解释。他写道："那些经常赌博的人迟早都会扔出一次维纳斯；而不时地，他们也的确能连着扔出两次甚至三次维纳斯。我们难道会打算因此而低能地确信，这种事情的发生是因为维纳斯的亲身介入，而不是因为纯粹的好运气？"[14]西塞罗相信，即使某事件的发生只是纯运气的结果，它仍然可以预计和预测。他甚至使用统计论述来奚落对人们星相学的迷信。尽管星相学在罗马是非法的，但它仍然生存了下来，而且处境还挺不错。这惹恼了西塞罗。他注意到在公元前 216 年的坎尼（Cannae），率领着大约五万迦太基及其盟国部队的汉尼拔（Hannibal），将一支人数多得多的罗马军队击溃，并杀死了 80000 名士兵中的 60000 人。"难道所有在坎尼倒下的罗马人，他们的占星都一样？"西塞罗这样问道，"但他们最终的命运都相同。"[15]西塞罗如果能知道在数千年后，《自然》杂志上发表的一个对星相学预测之正确性所进

行的科学研究，也得出了同样结论的话，他一定会大感激励[16]。另一方面，《纽约邮报》（*New York Post*）却在今天建议说，由于我是射手座的，所以我必须客观地看待批评，并相应作出似乎必要的不论什么改变。

西塞罗最终在随机性领域留下的主要遗产，就是他所使用的probabilis这一术语——我们现在所使用的“概率”一词的词源。不过，查士丁尼大帝（Justinian）在公元6世纪所编撰的罗马法典《汇编》（*the Digest*），才是概率一词作为该领域日常用语而出现的第一份文本[17]。要正确了解罗马人将数学思想应用到法律理论中的做法，就必须了解有关历史的来龙去脉：黑暗时代的罗马法是依照日尔曼部落的做法而制订的。这可没什么好的。举个作证方面的例子吧，设想现在有一名丈夫被指控和为他妻子做外袍的织娘有了一腿，而他对此完全否认，那么这个证词的真实性，并非来自于它们是否能经受住一场尖锐控辩中的种种严厉盘问，而是看他能不能一直咬定这个说法，哪怕当烧红的铁块烙在身上时也绝不改口。（如果我们把这个习俗再带回到现在的法律中，那通过庭外调解来解决的离婚案就会多得多了。）如果一名被告说马车夫根本就没有试图使车停下来，而另一名专家证人却说，蹄印表明车夫已经踩了刹车，那么日耳曼人的条令为这种情况提供了一个简单的办法：“让两边各挑出一个人来，用矛和盾干一仗，输者即为伪证，要被砍去右手。”[18]

作为这种基于武力之审判的替代或至少是一种补充，罗马人试图以数学的精确，来提供一个解救这种陈旧武断体系之不足的良方。在这个历史背景下来看，罗马人有关公正的观念中，就使用了先进的智识概念。由于认识到证据和证词常常会互相矛盾，而解决这些矛盾的最好方法，莫过于将无法避免的不确定性因素加以量化，罗马人据此创造出了“半证据”的概念，并应用于那些没有完全可靠的理由来相信或不相信证言证据的案件中。在某些案子里，与证据有关的罗马律令，还包含了甚至更为精细的证明力度。如在教令中

就规定，“主教不应被判有罪，除非有 72 名证人证明他有罪……枢机司铎不应被判有罪，除非有 44 名证人；罗马城的枢机助祭不应被判有罪，除非有 36 名证人；副助祭、侍僧、驱魔人、诵经人或看门人不应被判有罪，除非有 7 名证人。”[19]要想在这样的条文下被认定有罪，嫌犯不仅要犯下罪行，还得靠出售“罪行观摩券”来凑到足够多的证人才行。尽管如此，能认识到证词真实性存在差异的可能性以及这些真实性之组合规则的必要性，本身就是一个很好的开端。于是，恰恰在看来不大可能的罗马审判庭上，诞生了一套基于概率的系统化规则。

不幸的是，如果还蹒跚于Ⅷ和XIV这样的数字中，对数量的灵活使用就仍然十分困难。尽管罗马法有了一定的法律理性和一致性，却最终仍然缺乏数学有效性。例如，在罗马法中，两个“半证据”可以构成一个完整的证据。对于一个还不惯于通过量化方式来思考的头脑而言，这听起来挺合理，但在熟知分数运算的今天，问题就出来了：如果两个半证据可以构成一个完全确定的证据，那么三个半证据可以弄出个什么东西？根据概率组合的正确方式，不仅两个半证据只能给出一个仍不确定的证据，任意有限多个部分可靠的证据加在一起，也还是不可能给出一个完全确定的证据，因为概率的组合不是用加法，而是乘法。

这就是我们的下一条定律——概率组合规则：如果两个可能事件 A 与 B 相互独立，则 A 与 B 同时发生的概率，等于各自单独发生的概率之积。假设一名已婚者平均每年有约 1/50 的可能性会离婚，而一名警察每年有约 5000 分之一的机会殉职，那么一名已婚警察在同一年中离婚并殉职的可能性有多大？根据上面的规则，如果这两个事件相互独立，则概率约为 $1/50 \times 1/5000 = 1/250000$。不过这两个事件当然并非独立，而是互有关联的：如果你死了（真倒霉！），那你就没办法再离婚了。因此，运气差到上面这种地步的可能性，实际要比 250000 分之一少一点点。

为什么是乘而不是加呢？设想你做了一套收藏卡，上面印着你迄今为止通过因特网约会服务所碰到的 100 个家伙，这些家伙的个人网页上挂着的相片一个个看来都像汤姆·克鲁斯（Tom Cruise），而实际长得却跟丹尼·迪维图（Danny DeVito）有得一比[①]；再假设在每张卡片的背面，都记下了与此人有关的一些数据，诸如是否诚实和是否有吸引力等；最后，再假设这些预期的柏拉图恋人中，就诚实及吸引力两者中的每一项指标而言，都有 1/10 的可能性得到一个“是”。那么，在你收藏的 100 个人中，能在诚实和有吸引力两项上都合格的人有多少呢？让我们拿诚实来做第一个考察的对象吧（拿吸引力来考察也是一样的）。既然有 1/10 的卡片上的诚实评价为“是”，那 100 张卡片中就有 10 张这种卡片；在这 10 张卡片中，有吸引力的是多少呢？又是 1/10，也就是硕果仅存的 1 张卡片。第一个 1/10 将可能的选择减少到了 1/10，第二个 1/10 又再进一步地降到了 1/10 的 1/10，也就是 1/100。这就是要使用乘法的原因。因此，如果你的要求不仅仅是诚实和有吸引力，那你就得一直这样乘下去，所以……得，祝你好运吧！

在继续我们的话题之前，有一个重要的细节值得再注意一下：*如果两个可能事件 A 与 B 相互独立*。假设客机上只剩一个座位，却还有两名乘客可能会上来；再假设根据以往的经验得知，预定了座位的旅客中有 2/3 会实际来乘坐航班。根据乘法规则，登机口的服务人员可以算出，大概有 2/3×2/3 即约 44％的可能性，她会要面对一位满脸不快的顾客；另一方面，因为两名乘客都没来而使得飞机带着一个空座位起飞的可能性则是 1/3×1/3，或说只有约 11％。但这是在两名乘客来或不来是相互独立的假设下得到的结果。如果他们是一道的旅伴，那上面的分析就错了：两人同时乘机的可能性跟一个人来乘机的可能性一样，都是 2/3。因此，一定要记住，利用乘

① 只需去谷歌搜索一下两人的相片，就知道作者为何有此说法了。

法将简单概率组成复合概率的做法，只有在这些事件的发生不以任何方式相互依赖时，才是正确的。

我们方才所使用的规则，也可用于有关半证据的罗马法中：两个独立的半证据同时都错的可能性是 1/4，因此两个半证据构成的不是一个完整证据，而是完整证据的 3/4。在该用乘法的地方，罗马人用了加法。

在某些情况下，概率必须相加，这就是下一条定律。这时，我们希望知道两个事件中至少有一个发生的概率，而不是两个同时发生的概率。该定律是：如果一个事件有多个不同且互不重叠的可能结果 *A*、*B*、*C* 等等，那么结果 *A* 或者结果 *B* 发生的可能性等于结果 *A* 和结果 *B* 各自单独发生的概率之和，而所有可能结果（*A*、*B*、*C* 等）各自单独发生的概率之和等于 1（即 100%）。如果你想知道两个独立事件 *A* 与 *B* 同时发生的可能性，就用乘法；如果想知道两个互斥事件 *A* 或 *B* 发生的可能性，就用加法。回到客机的问题上来：登机口的服务员何时应将概率相加而非相乘呢？如果她想知道的是两名乘客都来或者都不来的可能性的话，就应该用加法将两个概率相加，根据我们上面的计算可知，这个概率是 55%。

这三条定律虽然简单，却构成了概率论的大部分基础。如果使用得当，我们将能更好地洞察自然及日常世界的运作。在平时的决策中，我们一直在使用这些定律，但跟罗马人一样，我们并非总是使用得当。

看看过去，摇摇脑袋，然后写下《腐化堕落的罗马人》（*The Rotten Romans*）（Scholastic，1994）这类标题的书，并不是什么难事。但为了避免毫无缘由的自满自得，我们来看看迄今所讨论的这些基本定律在司法体系中的若干应用方式，并以此来结束本章。这些讨论所揭示的结果，应该足以让任何沉醉于文化优越感的人冷静下来。

好消息是我们不再有所谓的半证据了。但我们确实有着类似于999000/1000000证据这样的东西。举例来说，我们常常看到DNA分析专家在罪案庭审中作证，称在罪案现场所获取的DNA样本与某嫌犯的DNA样本相匹配。这些匹配有多确定呢？最初引入DNA证据时，不少专家证明说DNA检验中不可能存在假阳性。如今的DNA专家一般认为，随机选一个人，其DNA与罪案中的样本匹配的可能性，不到100万分之一或10亿分之一。在这样的可能性下，人们很难责怪法官这样想道：让这家伙死在牢里吧。但还有另一个通常不会告诉陪审团的统计数据，即整个实验操作过程中可能出错的统计数据。例如，在采集和使用样本时，可能会无意将样本混合或调换，或是曲解了结果，或是结果报告出了错。这些错误都很少发生，但却远非随机匹配那样罕见。例如费城犯罪实验室就承认曾在一起强奸案中把被告和受害人的参考样本给对调了，而名为Cellmark Diagnostics的检验公司也承认犯过一个类似的错误[20]。不幸的是，在法庭出示的有关DNA的统计数据是如此有力，以至于尽管有11名证人提供的不在场证据，俄克拉何马州的一个法庭仍然判处了一位名叫蒂莫西·达拉谟的男子以超过3100年的徒刑。后来才发现，在初始分析时，实验室未能将强奸犯的DNA和达拉谟的DNA从被检液体中完全分离，而当两者的DNA进行比较时，就得到了阳性结果。这个错误在后来的重检中被发现，达拉谟被释放了，但已经在监狱中呆了差不多4年。[21]

对人为因素造成的错误率，不同人的估计值互不相同，但许多专家认为它大致是1%左右。不过，由于从未衡量过许多实验室的错误率，因此法庭通常不会接受此类统计数据证言。就算法庭接受了这些有关假阳性的证言，陪审员又会如何来评价它们的效力呢？多数陪审员会假设，在给定了这两类错误概率——10亿分之一的随机匹配概率和1%的因实验室错误造成的匹配概率——一时，总错误率应是介于两者之间的一个什么值，比如说500万分之一，而对大多数

的陪审员而言，根据这个概率来提出质疑仍非足够合理。运用概率论的法则，却能得到一个非常不同的结果。

求解的思路是这样的：既然两种错误都非常罕见，那我们就可以忽略随机匹配和实验室错误同时发生的可能性，而只来看看发生了其中一个或另一个错误的可能性。由加法法则给出的可能性为实验室错误的概率（1%）＋随机匹配的概率（10 亿分之一）。由于后者要比前者小 1000 万倍，因此，较大的那个 1%的概率，能很好地近似这两个概率之和。因此，在给定了这两种可能的错误源后，我们应当忽视自负的专家们给出的那些有关随机匹配错误率的证言，而将注意力集中在可能性高得多的实验室人为错误上，而这恰恰就是法庭通常不允许律师提出的数据！因此，那些被再三重复的 DNA 检验绝对可靠的宣言实际太过夸张了。

这并不是一个孤立的问题。相比于许多世纪前在罗马发生的情况，在现代法律体系中使用数学所碰到的问题一点也不小。最能展现概率在法律中的运用和误用的著名案例之一，就是 1968 年在加州最高法院审理的“公诉人对科林斯”（People v. Collins）案件[22]。以下就是法庭在裁决时所掌握的有关本案的一些事实：

1964 年 6 月 18 日，大约上午 11：30，刚买完东西的胡安妮塔·布鲁克斯夫人，正沿着洛杉矶城圣佩德罗区的一条小巷回家。她当时拖着一个柳条筐小车，里面放着所购的杂货，钱包就放在购物袋顶上。她用一条藤条拖着小车。当她弯腰去拣一个空硬纸盒时，突然被人推倒在地。她既没有看见这个人，也没有听见其逼近。她在摔倒时感到了晕眩和一阵疼痛，但仍然努力看到了一名年轻女子正从现场跑开。根据布鲁克斯夫人的陈述，这名年轻女子看来大约重 145 磅，穿着“深色的什么衣服”，有着“介于暗金黄色和亮金黄色之间”的发色，但比被告珍妮特·科林斯上庭时的发色要深。事情发生后，布鲁克斯夫人马上发现她装有 35～40 美元现金的钱包不

见了。

与发生抢劫差不多同一时间，居住在小巷当头街上的约翰·巴斯正在屋前给草坪浇水。他的注意力被小巷中传来的“许多喊声和尖叫声”所吸引。当他朝那个方向看过去时，见到一名女子从巷子里跑出来，并跳上了他的街对面的一辆黄色汽车。他无法给出这辆汽车的牌子和型号。汽车马上发动，并开了一个大圈来绕过另一辆停放着的车辆。在这条窄路上绕圈时，它一度离巴斯仅 6 英尺，而后者看到开车的是一名黑人男子，留着短的络腮胡子……其他证人则分别描述该车为黄色、黄色带灰白车顶和黄色带蛋壳白色车顶。该车大小据述为中到大型。

事情发生几天后，一名洛杉矶警官发现了一辆黄色带灰白车顶的林肯车停在被告家前，并与被告进行了交谈，说明他正在调查一起抢劫案。他注意到疑犯很符合那一对犯下罪行的男女的描述，只是现在这名男子没有大胡子，不过他自己承认他有时会留大胡子。当天晚些时候，洛杉矶警察逮捕了这两名疑犯——马尔科姆·里卡多·科林斯和他的妻子珍妮特。

指控夫妻俩的证据并不充分，因此，该案在很大程度上依赖于受害人与证人巴斯对疑犯的认定。对于检方而言很不幸的是，受害人与证人都不是证人席上的模范标兵。受害人无法认定珍妮特就是罪犯，也根本没见过那名开车的司机；巴斯没见过抢劫者，而且在警察局给出的那一列待指认者中，也不能确认马尔科姆就是那名司机。因此，这件案子看来无法再进行下去了。

然后，本案的关键证人登场了。根据加州最高法院判决书中的描述，该证人不过是“一所州立大学中的数学讲师”。据其证词，被告是“一名金发梳马尾辫的高加索女子……（以及）一名留有短的络腮胡子的黑人男子”这一事实，就足以给法庭上的夫妻俩定罪了。为了说明这一观点，检方给出了下面的表格，我从最高法院判决中

逐字引用如下：

特　　征	独立发生的概率
部分为黄色的汽车	1/10
短胡子男性	1/4
络腮胡子黑人男性	1/10
梳马尾辫的女孩	1/10
金发女孩	1/3
汽车中人种不同的夫妻俩	1/1000

检方传唤的这名数学讲师，将乘法规则应用在这些数据上，而由乘积得到的结论，是出现一对符合所有这些不同特征的夫妻的概率为 1200 万分之一。证人据此推理认为，法庭上这对夫妻无辜的概率就是 1200 万分之一。检方随即指出，这些独立的概率都是估计值，并请陪审员们用自己的估计值来做做这道算术题。检察官称，他本人相信这些值都是相当保守的估计，如果用他自己的估计值，那得到的概率更接近于 10 亿分之一。陪审团接受了这一结果，并认定夫妻俩罪名成立。

这个计算错在何处呢？首先，正如我们已经看到的，要通过各组成概率相乘来得到复合概率，这些组成概率所对应的范畴应相互独立，而在本案中，所列特征却明显不是相互独立的。例如，表中称看到一名“络腮胡子黑人男性”的可能性是 1/10，而看到一名“短胡子男性”的机会是 1/4，但大多数络腮胡子男性同样也留短胡子，因此，如果你看到了一名“络腮胡子黑人男性”，那么此人同时也留着短胡子的可能性就不再是 1/4，而远比这个值要大了。如果把“络腮胡子黑人男性”这个特征去掉的话，上述问题就能得到纠正，此时得到的概率乘积约是 100 万分之一。

法庭的那个分析中还有一个错误：真正与案件有关的概率并不是上面所给的那个——也就是随机选出一对夫妻，他们符合对疑犯之描述的概率。真正相关的概率应该是一对符合所有上述特征的夫妻，他们有罪的可能性到底是多少。前面一个概率可能是 100 万分之一，但对后者而言，由于与案发地相邻的区域中有数百万人口，因此，假定在这一地区内有 2～3 对夫妻符合上述描述，应该是一个合理的估计。在这种情况下，仅根据前面的证据（这些证据差不多也就是检方所掌握的所有材料了），要判断符合这些描述的夫妻有罪，正确的可能性不过是 1/2 或 1/3。这样的概率根本就不足以对嫌犯提出合理的质疑。由于这些原因，最高法院最终推翻了对科林斯的定罪。

在现代法庭中使用概率和统计仍然是一个争议性话题。在科林斯案中，加州最高法院嘲弄了其所谓的“数学之审判”，但却给更多“数学技巧的适当应用”敞开了大门。其后的若干年中，法庭很少考虑数学论据。但即使律师与法官们没有明确地使用概率或数学定理，他们也确实经常使用这类推理；陪审员在判断证据的有力程度时的做法也一样。不仅如此，由于有必要对 DNA 证据的效力进行评价，统计论据正变得越来越重要。不幸的是，这不断增长的重要性，却并未促使律师、法官或陪审团对概率的理解也随之增长。如同在南加州大学执教概率与法律的托马斯·里昂（Thomas Lyon）所解释的那样：“没几个学生会在法律课上使用概率，也没有几个律师觉得概率在其中应该占有一席之地。”[23] 如同在其他领域一样，在法律中，对随机性的理解同样可以揭示隐藏的真理，但只对那些掌握了这个真理发掘工具的人而言是如此。在下一章中，我们将来说说系统化研究这些工具之第一人的故事。

第三章

寻找穿越可能性空间之路

1576 年之前的几年中，人们常能看到一名着装古怪的老人，迈着怪异而不规则的步伐，在罗马的大街上游荡，还不时发出几声既不针对特定听众、也根本没人听的喊声。作为著名星相师、宫廷贵族的医师、帕维亚大学医学院主席，这个人曾一度被整个欧洲所称颂。他创造了许多不朽的发明，其中包括组合锁以及如今用于汽车上的万能接头的前身。他发表了 131 篇著作，涉及哲学、医学、数学和科学等领域的广泛课题。但在 1576 年，他不过是一个有过去却没将来的人，一个生活在阴暗和无望的贫困之中的人。这一年夏末，他坐在桌前，写下了最后的话语——为他那 16 年前才 26 岁就被处死的长子、也是他最钟爱的儿子所写的一首颂诗。老人死于 9 月 20 日，离他 75 岁的生日只差了几天。他 3 个儿子中的两个都先他而离世，而去世时他那唯一健在的儿子，正当着受雇于宗教裁判所的职业拷问者——一份赏赐他提供不利于他父亲之证据的优差。

吉罗拉莫·卡尔达诺（Gerolamo Cardano）死前烧掉了 170 篇未发表的文章草稿[1]。清理遗物的人发现了其中未被烧毁的 111 份，其中之一是一篇由 32 个短小章节构成的论文，写就于数十年前，而且从外观来看还常常被改动。论文名为《论赌博》（*The Book on Games of Chance*）。这是历史上第一本关于随机性理论的著作。到卡尔达诺生活的时代为止，人们已经在赌博或处理其他不确定性方面实践了数千年：我能在渴死之前穿越这片沙漠吗？当大地像现在这样震动的时候，仍然呆在山崖下面是不是很危险？那个喜欢在岩

石上画水牛的穴居人女孩对我露齿一笑，是不是表示她喜欢我？但直到卡尔达诺之前，还没有人对赌博或其他不确定过程进行过详尽分析。卡尔达诺对于随机性之作用机制的洞察，形成了被称为样本空间定律的原理。样本空间定律是一种新思想和新的方法论，并成为了后世对不确定性进行数学描述的基础。该方法十分简单：我们使支票簿收支保持平衡的思路的随机性定律形式对应物，就是样本空间定律。这个简单的方法却能使我们系统化地解决诸多问题，而如果没有这个新方法，那些问题几乎会令人绝望地困惑。为了说明这条定律的运用及其威力，考虑一个描述十分简单的问题。解决这个问题并不需要高等数学知识，却很可能比随机理论历史上的任何其他问题都难倒了更多的人。

作为期刊专栏，《大观》（*Parade*）杂志的《来问玛里琳》（*Ask Marilyn*）被公认为了不起的成功。刊登于350期报纸上，并有着足以令人自豪的近3600万份的发行量，这个始于1986年的问答式专栏今天仍然充满活力。专栏中的问题本身可能与答案一样富有启发性，而这个专栏也可以看作是一个（不科学的）调查美国人脑子里在想些什么的盖洛普民意测验（Gallup Poll）。例如下面这些问题：

当股票市场在交易日停盘时，为什么每个人都边微笑边鼓掌地围站着，而不管今天的股票是涨还是跌？

一位朋友怀了对异卵双胞胎。双胞胎中至少有一个是女孩的机会是多大？

当你在路上驶过一只死黄鼠狼时，为什么过了10秒钟你才闻到那股臭味？假设你并没有碾过它。

显然，美国人都很现实。这里需要注意的，是每个问题都含有一定的科学或数学成分，这也是许多在该专栏得到答复的问题的一

个共同特征。

可能有人　特别是对于数学和科学还知道那么一点的人——会问："谁是这个当家的玛里琳?"呃……玛里琳就是玛里琳·莎凡(Marilyn vos Savant)，因为多年占据着吉尼斯世界纪录名人榜中有记录的世界最高 IQ（228）宝座而名声远扬，也因为嫁给了罗伯特·贾维克（Robert Jarvik）——贾维克人工心脏的发明者——而广为人知。但不管著名人物们的成就如何，他们被人们牢记的原因，有时却是一些他们希望从未发生过的事情（"我没有和那个女人发生过性关系"①)。玛里琳的情况也大体如此。最使她出名的，就是她对于下面这个 1990 年 9 月某个周日出现在专栏中的问题所作的回答(我稍微改动了问题陈述中的一些用词)：

假设有一个游戏，游戏者可以在 3 张门中做出选择：一张门后是一辆汽车，而另两张门后则是山羊。当游戏者选定了某一张门后，游戏的庄家（他知道各个门后面是什么东西）将打开另外两张门之一，门后是一只山羊。然后他问游戏者："你想更换你的选择，去选另外那张没打开的门么?"请问，如果游戏者的确改变了选择，这个改变是否对他有利?[2]

这个问题的提出，受到了电视游戏节目《让我们做个交易》(*Let's Make a Deal*）的启发。该节目播放于 1963～1976 年，在 1980～1991 年还曾改头换面地重新出现过几次。节目最吸引人的，就是英俊和蔼的主持人蒙提·霍尔（Monty Hall)，以及他那身材惹火的陪衬式助手、1957 年度的加州 Azusa 小姐卡萝尔·梅里尔(Carol Merrill)。

对于节目创始人而言，在播放了近 27 个年头的 4500 期后，节

---

① 美国总统克林顿在莱温斯基丑闻中所撒的那个非常著名的谎。

目的主要遗产却是上面这个数学概率问题，这不能不说是一件令人吃惊的事情。玛里琳的读者们对于这个问题的热烈甚至是激烈的回应，使她和《让我们做个交易》这个节目都长留青史。毕竟，这个问题看起来挺傻的：还有两扇门可以选择——打开对的那扇，你就赢了；打开另一扇，你就输了——因此，看来不言而喻的事实是，不管你改变或不改变选择，赢的机会总是 50 对 50。还有比这更简单的么？问题在于，玛里琳在专栏中说，改变选择的赢面更大。

尽管公众对于数学问题向来表现出可以预料得到的死气沉沉，但这一次，玛里琳读者们的反应，就如同她是在鼓吹放弃加利福尼亚并将之交还给墨西哥一样激烈。她对于一个再显而易见不过之事实的否定，带来了一场邮件的大雪崩——据她估计，可能有 10000 封信之多[3]。如果问问美国人植物是否产生了空气中的氧气，或光是否比声音传播得更快，或你是否无法通过煮沸带放射性的牛奶来使之变得更安全时，回答为“否”的比例都达到了两位数（分别为 13%、24%和 35%）[4]。但在玛里琳这事儿上，美国人团结了起来：92%的人认为玛里琳错了。

许多读者看来好像非常失望：这样一个他们在各类问题上都如此信任有加的人，怎么会在这样一个简单问题上犯了糊涂呢？她的错误是不是美国人那令人哀叹的无知的一个象征？将近 1000 名博士——其中有许多数学教授，而且似乎对这个错误特别生气——给玛里琳写了信[5]。“你错得离谱，”一位乔治·梅森大学（George Mason University）的数学家写道：

“让我来解释一下吧：如果一扇失败的门被打开了，那这个信息将使剩余两个选择各自的获胜概率变为 1/2，因为两者之中的任何一个都没有理由比另一个更可能获胜。作为一名职业数学家，我对于普通民众在数学技能上的匮乏而忧心。请您承认自己的错误，以此为改变这种状况而做出贡献，并请在今后更加细心一些。”

从狄克森州立大学（Dickinson State University）寄来的是这么一封信："在至少 3 位数学家的纠正之后，您却依然没有认识到自己的错误，这使我震惊。"而来自乔治敦大学（Georgetown）的则是："到底要多少个愤怒的数学家才能使您改变您的想法?"美国陆军研究所里的某位如此评论道："如果所有这些博士都错了的话，这个国家就有大麻烦了。"各种反馈的数量是如此之多，且持续时间是如此之长，以至于在这个问题上投入了颇多一部分的栏目空间之后，玛里琳决定不再继续讨论这个问题。

那位写信的陆军博士说得大概没错：如果所有这些博士都错了，那将是出大麻烦的信号。但玛里琳确确实实是对的。20 世纪最杰出的数学家之一的保罗·爱尔特希（Paul Erdös)① 在得知此事后说："这不可能。"甚至在看了正确答案的正式数学证明之后，他仍不相信并且恼火了。直到一名同事利用计算机仿真将这个游戏重复了数百次，并让爱尔特希亲眼看到，改变选择相比于不改变选择，其胜率之比为 2∶1，在这之后，爱尔特希才勉强承认他错了[6]。

如此显然的一件事怎么可能错呢？用哈佛大学一位概率统计专业教授的话来说，个中原因在于"我们大脑的构造天生就不能很好地处理概率问题"[7]。伟大的美国物理学家理查德·费恩曼（Richard Feynman）曾告诉我说，如果只是将别人的推导看了又看，就永远别想理解任何物理学问题。他说，真正去理解一个理论的唯一途径，就是自己去完成它的推导（或者可能最后证明它是错的!）。对于我们这些不是费恩曼的人而言，重新证明他人的工作，是个让自己丢掉在大学里的饭碗并在家得宝（Home Depot)② 的收银员岗位

① 保罗·爱尔特希（1913～1996）是匈牙利数学家，也是当代最伟大的数学家之一，在组合论、数论、概率论、集合论、数学分析等许多数学领域中都做出了贡献。

② 美国的一家家居建材连锁超市。

上实践数学技能的好方法①。但蒙提·霍尔问题不需要任何专业数学知识就能解答。你不需要微积分、几何、代数甚至安非他明——据称爱尔特希就很喜欢使用这种兴奋剂[8]（传说爱尔特希在戒用了一个月之后说道："以前，当我看着一张白纸时，脑子里满是各种各样的点子；而现在，我所能看到的就只有一张白纸了。"）你所需要的，不过是对于概率的运作方式以及样本空间定律的理解而已，而样本空间定律正是卡尔达诺在 16 世纪首先著诸于文、用于分析随机性的理论框架。

吉罗拉莫·卡尔达诺并不是试图打破 16 世纪欧洲的智识氛围的离经叛道者。对他而言，犬吠意味着所爱的人的死亡，而屋顶上几只乌鸦发出的嘎嘎叫声，则意味着一场严重的疾病正在到来。他跟其他人一样相信命运、运气，以及根据行星和恒星的排列预测未来。尽管如此，如果他也玩扑克的话，他可不会把运气赌在中间缺一张牌的顺子上。赌博是卡尔达诺的第二天性，他对赌博的感觉存在于骨子里而非脑袋里。因此，他对于各种可能发生的随机结果间数学关系的理解，就凌驾于"这种理解根本就注定无用"之类的宿命论信念之上了。卡尔达诺的工作也走在了当时粗糙的数学之前，因为在 16 世纪早期，代数乃至算术都还处在石器时代，甚至连等号都有待发明。

根据卡尔达诺的自传以及同时代人的一些文章，他配得上在历史上大书一笔。有些文献内容相互矛盾，但有一件事可以肯定：生于 1501 年的卡尔达诺，并不是一个值得投资的孩子。尽管——或者说因为——他的母亲基娅拉已经有了 3 个儿子，她却不喜欢小孩子。又矮又壮、脾气火爆且男女关系混乱的基娅拉怀上吉罗拉莫后，就

---

① 作者在此似乎是小小地调侃了一下大学的教学科研人员所面临的处境：如果只是重复他人的工作而没有自己原创性的成绩，那么在大学里是很难立足的。

配置了一剂16世纪的事后避孕药：由苦艾、烧过的大麦粒和撑柳根酿成的酒。她喝下了这剂药酒，希望能将胎儿打掉。但药酒虽然使她恶心无比，未出世的吉罗拉莫却对之凛然不惧，而且显然十分满意于药物在母亲血液中留下的代谢产物。基娅拉其他的尝试也都以差不多的方式宣告失败。

基娅拉与吉罗拉莫的父亲费吉奥·卡尔达诺并没有结婚，但他们经常表现得如同一对正式夫妻——他们那高声的对骂可是相当地出名。在吉罗拉莫出生前一个月，基娅拉离开了米兰的家，搬到南边20英里的帕维亚与她妹妹同住。在3天痛苦吃力的分娩之后，吉罗拉莫终于冒出了头。看到这孩子的第一眼肯定让基娅拉觉得，她总算还是可以摆脱这个臭小子了。他看上去十分脆弱，更糟糕的是，他不哭也不闹，就静静地躺在那里。产婆预言他将在出生后的一小时内死掉。但如果基娅拉梦想着“美好解脱”的话，她又再次失望了：婴儿的乳母用温酒给他泡了个澡，然后，吉罗拉莫复活了。不过他良好的健康状况只维持了几个月，就与乳母和3个同母异父的兄弟在大瘟疫中病倒了。黑死病——这场瘟疫的另一个称呼——实际上是3种不同的疾病：淋巴结鼠疫、肺鼠疫和败血性鼠疫。卡尔达诺患上的是三者中最常见的淋巴结鼠疫，其典型症状之一就是令人痛苦难当、鸡蛋大小的腹股沟淋巴结肿块，也是该病名称的由来。一旦出现了这样的肿块，患者的期望寿命就只有一周左右了。

黑死病最初于1347年跟随着从东方返回的一支热那亚舰队，由西西里东北城市墨西拿的一个港口进入欧洲[9]。舰队很快被隔离，所有船员都死在了船上——但老鼠活了下来，并一路小跑地上了岸，身上带着病菌以及传播病菌的跳蚤。接下来的瘟疫大爆发在两个月内夺去了该城一半人口的性命，并最终导致了全欧洲25%～50%的人口死亡。后继的传染病不断到来，使得欧洲人口在数个世纪中一直无法增长。1501年可算不上什么好年头，因为当年在意大利也有瘟疫流行。吉罗拉莫的乳母和兄弟们都死了，这个走运的孩子却幸

存了下来，而且没留下什么后遗症，只除了外貌上的损伤：他的鼻子上、额头上、脸颊以及下巴上都留下了瘤子。他注定要活过将近75个年头，一条满是冲突的人生之路，以及饱受打骂的幼年。

吉罗拉莫的父亲大致可以算个操作员。就职业而言，他是名几何学家，而且曾是莱昂纳多·达·芬奇（Leonardo da Vinci）的密友。几何学家从来就不是个赚钱的职业，费吉奥也的确常为租金而发愁，因此他做起了事务咨询的买卖，为豪门望族们提供法律与医学方面的意见。费吉奥宣称他是米兰的一位戈弗雷多·卡斯蒂格里奥尼（Goffredo Castiglioni）——更有名的称呼是教皇塞莱斯廷四世（Pope Celestine IV）——之兄弟的后人。得益于此，咨询买卖最终兴旺了起来。吉罗拉莫五岁时，他父亲将他带入了生意之中　姑且这么说吧。实际上，他的父亲在他背上绑了个筐，装上沉重的法律与医学书籍，然后拽着小男孩满城跑去会见主顾。吉罗拉莫后来写道："有时，当我们走在街上时，父亲会命令我站住，然后打开一本书，拿我的头当桌子，阅读一些长长的段落，并在我因为承受不住那巨大重量而摇晃时，用他的鞋尖刺我，让我不要动来动去。"[10]

1516年，吉罗拉莫下了决心，认为他最好的机会在于医学领域，并宣布要离开米兰的家，去帕维亚学习。费吉奥却希望他能学习法律，以便赚取100克朗的年薪。一场大吵之后，费吉奥缓和了下来，但问题却没有解决：如果没有这笔年薪，吉罗拉莫怎样才能维持在帕维亚的生活和学习呢？于是吉罗拉莫开始为人占星，并教小学生几何、炼金术和天文学，并将所挣的钱积攒起来。其间的某一天，他发觉自己在赌博方面颇有天分，而这种天分带来金钱的速度，比其他方法要快得多。

在卡尔达诺的时代，对任何一个有兴趣赌博的人来说，每座城市都是拉斯维加斯。赌纸牌，赌骰子，赌西洋双陆棋，甚至赌国际象棋；到处都可以下注。卡尔达诺将这些赌博分为两类：一类需要某种形式的策略或技巧，而另一类则完全靠运气。在国际象棋这一

类赌博中，卡尔达诺需要冒上被某个16世纪的鲍比·费舍尔（Bobby Fischer）① 击败的危险，但如果是对那一双小小的立方体下注，他的机会就跟其他人一样大了。不过，在这后一类赌博中，他的确比别人更有优势，因为相比于所有的对手，他对各种局面下获胜的机会，有着更好的理解。于是，作为进入博弈世界的第一步，他玩起了那些纯靠运气的赌戏。没过多长时间，他就为自己存下了1000克朗的教育基金——比父亲希望他得到的10年年薪还要多。1520年，他在帕维亚注册成为一名学生，不久后，便开始了赌博理论的写作。

正如古希腊人和罗马人拥有古希腊语这个优势一样，在卡尔达诺生活的年代，他也拥有了理解事物的一个优势，因为在将算术作为有力工具而使用的方向上，印度人已迈出了最初的一大步。正是在这个时期，以10为底的位置相关计数法发展了起来，并于公元700年前后成为了标准做法[11]。印度人还在分数算术方面取得了巨大进展，而分数运算对于概率分析至关重要，因为事情发生的可能性总是比1小。印度人的这些知识为阿拉伯人所采用，并最终传入欧洲。最初的运算符号缩写——p表示加法，m表示减法——于15世纪被使用，而＋和－也在同一时期由德国人引入，不过当时只是用来表示箱子超重或欠重。如果注意到当时还没有等号，我们大概就能在一定程度上认识到卡尔达诺所面临的挑战了。等号于1557年由工作于牛津和剑桥的罗伯特·雷科德（Robert Recorde）发明。受到几何学的启发，他认为没有什么东西能比两条平行直线更为相似，因此，这样的两条平行直线段应当被用来表示相等性。而归功于某位英国国教牧师的乘法符号×，则要到17世纪才会登上历史舞台。

---

① 鲍比·费舍尔（1943～2008）是美国国际象棋大师与国际象棋世界冠军，并常被认为是有史以来最伟大的棋手之一。

卡尔达诺的《论赌博》涵盖了纸牌、骰子、西洋双陆棋和距骨等游戏。该书并非完美，它反映了卡尔达诺的性格，疯狂的想法，狂野的脾气，从事任何事情时所怀有的激情——以及他生命和时代中的动荡。在某些问题上，卡尔达诺还犯了错。但《论赌博》仍然是在概率论战场上登陆建立的滩头阵地，是人类认识不确定性之本性的探索过程中取得的第一个成功。而卡尔达诺用来解决偶然性问题的方法，则让人同时惊讶于其威力和简单性。

并非书中的所有章节都是处理技术性问题的。如第 26 章的标题是《教得好也能玩得好吗?》（他最后下的结论是："知道一件事，和实际去做这件事，似乎完全是两码事。"）第 29 章名为《论赌徒的性格》（"有些人话太多，让自己和别人都丧失了正常的理智。"）这看起来都更像是《亲爱的艾比》（*Dear Abby*）① 而非《来问玛里琳》。但在同一本书中，卡尔达诺在第 14 章《论点数的组合》（论可能性）里，叙述了"一条普适的规则"——即样本空间定律。

"样本空间"一词所表达的，是某个随机过程的可能结果，可被视为一个空间中的点。简单情况下，这个空间可能仅包含少数几个点，但在更复杂的情况中，它可能如我们生活于其中的空间那样，是一个连续统。但卡尔达诺并没有称其为"空间"：数的集合可形成一个空间的描述，还要再过一个世纪才会出现，那时，天才的笛卡儿（René Descartes）② 发明了坐标，并统一了代数与几何。

用现代语言来说，卡尔达诺的规则看起来差不多是这个样子的：假设一个随机过程有若干等可能性的结果，这些结果中有些好（即赢的结果），有些不好（输的结果）。那么，获得一个好结果的概率，就等于所有那些好结果所占的总比例。这个所有可能结果所构成的

---

① 一个知名的建议专栏。

② 勒奈·笛卡儿（1596～1650）是法国哲学家、数学家和物理学家，解析几何的创始人。

集合称为样本空间。换而言之，如果扔出一个骰子6面中的任一面，其可能性均等，那么这6个可能出现的面就形成了样本空间，而如果对其中两个面下注，获胜的概率就是2/6。

对于所有结果发生的可能性都相等这个假设，我们再多说一句。很显然，该假设并非总是成立。如果我们对奥普拉·温弗瑞（Oprah Winfrey）① 成年时的体重进行观测，那么可能观察到的体重所构成的样本空间（曾经）是145磅到237磅，而在这整段时间内，并非所有的体重区间都以等可能性出现[12]。不同的可能结果具有不同的概率，这使得问题变得复杂，但通过仔细计算来为每个可能结果赋予一个适当的概率，我们可以解决这个问题。不过让我们先来看几个例子，正如卡尔达诺所分析的那些例子一样，其中的结果都具有等可能性。

卡尔达诺规则的力量与某些微妙之处密不可分，其中之一就是“结果”一词的含义。迟至18世纪，多部概率论著作的作者、著名法国数学家达朗贝尔（Jean Le Rond d’ Alembert）② 在分析扔两枚硬币的问题时，还错误地使用了这个概念[13]。扔出两枚硬币，则正面朝上的硬币之可能数目为0、1或2。然后，达朗贝尔推论说，既然有这样3种结果，那么每种结果出现的可能性就一定是1/3。他错了。

卡尔达诺的工作，其最大不足之一在于没有系统分析事件序列不同的可能发生方式，例如连续扔多次硬币所能得到的结果，等等。正如下一章中所说，要到他过世后的下一个世纪，才有人来做这个工作。不过，扔两次硬币的问题已经足够简单，我们可以很容易地将卡尔达诺的方法用于其上。关键是要认识到，所谓可能的结果，

---

① 奥普拉·温弗瑞是美国电视访谈节目“奥普拉脱口秀”的主持人，被称为“脱口秀女王”，同时也是最有影响力的黑人女性之一。

② 让·勒隆·达朗贝尔（1717～1783）是法国数学家、物理学家和天文学家，在数学分析领域做出了基础性的工作，也为牛顿力学体系的建立做出了卓越贡献。

应该是描述硬币两次掉落情况的数据，而不是由这些数据计算得到的正面朝上的次数。换句话来说，可能结果并不是0、1或2次正面朝上，而应该是可能出现的投掷结果序列（正，正）、（正，反）、（反，正）和（反，反）。构成样本空间的，便是以上这4种可能结果。

按卡尔达诺所说，我们接下来应该排列出这些结果，并记录每个结果中正面朝上的硬币数目。在这4种可能结果中，仅仅只有（正，正）这一种能得到两枚正面朝上的硬币。类似地，只有（反，反）这一种让我们空手而还。但对于一枚硬币为正面的情况，则有两种结果能满足要求：（正，反）和（反，正）。在此，卡尔达诺的方法就说明了达朗贝尔的错误之处：出现0个或2个硬币正面朝上的机会各为25%，但出现1个硬币正面朝上的机会则为50%。如果卡尔达诺以3赔1的赔率下注于出现1次正面朝上，那么在一半的赌局中，他会输掉赌注，但在另一半的赌局中，却会赚到三倍于赌本的钱。对于一个正在为读大学而积攒学费的16世纪的小男孩来说，这可是个了不得的好机会　　即使在今天，只要找得到愿意出这个赔率的人，那也仍然是个了不得的好机会。

经常出现于初等概率课程中的一个相关问题，就是所谓的“两女儿问题”。它与我从《来问玛里琳》专栏所引用的问题之一颇为类似。假设一名母亲怀上了一对双胞胎，现在她想知道两个都是女儿或一个儿子一个女儿等情况的概率。这个问题的样本空间，由两个孩子按出生先后顺序排列所得的性别列表构成：（女孩，女孩）、（女孩，男孩）、（男孩，女孩）以及（男孩，男孩）。这个样本空间与前面那个扔硬币问题的样本空间其实是一回事，只不过空间中点的命名方式有所不同：“正”变成了“女孩”，而“反”变成了“男孩”。数学家们给这种某个问题不过是另一问题换了个说法的情况起了一个怪名字：同构性。两个问题同构常常就意味着我们可以省掉很多麻烦。在目前的例子中，同构性意味着两个孩子都是女孩的可能性，

等于两枚硬币都是正面朝上的可能性。由此，甚至不需要进行任何分析，我们就能知道问题的答案是一样的：25%。现在，我们就能回答玛里琳专栏里的那个问题了：两个孩子中至少有一个女孩的可能性，等于两个都是女孩的可能性，再加上两个孩子中仅有一个是女孩的可能性，也就是25%加50%，等于75%。

在两女儿问题中，常被进一步问及的问题是：如果两个孩子中有一个是女孩，那么两个孩子都是女孩的可能性是多少？大家可能会推理如下：既然已知一个孩子是女孩，那只要考虑剩下那个孩子的性别即可，而这个孩子是女孩的可能性是50%，所以两个都是女孩的可能性，也就是50%。

这个回答并不正确。为什么呢？尽管问题的描述中说有一个孩子是女孩，但却并没有说是哪一个。这就使情况发生了变化。如果你对此感到糊涂，也请不用担心，这是很正常的，因为这个问题恰恰能展现卡尔达诺方法的威力——它使得推理过程清晰而明了。

两个孩子中有一个是女孩的新信息，意味着我们可以把两个都是男孩的可能，从需要考虑的情况中去掉。因此，在运用卡尔达诺的方法时，我们将（男孩，男孩）这个结果从样本空间中剔除，而新的样本空间中就只剩下了3种结果：（女孩，男孩）、（男孩，女孩）和（女孩，女孩）。其中只有（女孩，女孩）——即两个孩子都是女儿——这一个结果是我们所希望的，因此，两个都是女孩的可能性是1/3，即33%。下面我们来看看，为什么不具体指明哪一个孩子是女孩这一点，会成为一个十分重要的因素。如果问题问的是当第一个孩子是女孩时，两个都是女孩的可能性是多少，那么我们就应该从样本空间中把（男孩，男孩）和（男孩，女孩）这两个可能都去掉，而所求的概率就是1/2，即50%。

我们应当信任玛里琳，因为她不仅试图提高公众对于基本概率知识的理解，而且还勇于继续发表相关的问题，哪怕是在经历了令人沮丧的蒙提·霍尔问题之后也仍然如此。作为讨论的结束，我们

再来看看她专栏中的另一个刊登于 1996 年 3 月的问题：

我爸爸从广播里听到了下面这个故事。杜克大学（Duke University）有两名学生，他们在整整一个学期的化学测验中都拿了 A。但在期终考的头天晚上，他们跑到另一个州去参加派对，返回学校时考试已经结束了。他们向教授解释说车胎没气了，并希望能进行补考。教授同意了，出了一份卷子，然后将两人分在两个教室里考试。第一道题（在试卷的一面上）有 5 分，而翻到另一面，则是有 95 分的第二题："是哪个轮胎没气了?"请问两名学生给出相同答案的可能性是多少？我爸爸和我都觉得是 1/16，对吗?[14]

不，这个答案不对：如果两名学生在撒谎，那么他们碰巧选中相同答案的可能性，其正确值为 1/4（如果你需要一点点帮助来找到原因的话，请看看本书最后的注解）[15]。好了，现在我们已经熟悉了将问题分解为一系列的可能情况的做法，接着我们就可以运用样本空间定律来解决蒙提·霍尔问题了。

如前所述，要理解蒙提·霍尔问题，并不需要什么数学方面的训练，但确实需要一点细致的逻辑思维。因此，如果你一边在看《辛普森一家》（*Simpsons*）的重播，一边在读本书的这一段时，最好把两件事里的某一件推后一些。好消息是，要理解我们的问题，只需几页纸就够了。

在蒙提·霍尔问题中，你面对的是 3 张门：一张后面是个值钱的东西，比如说一辆崭新锃亮的红色玛莎拉蒂（Maserati），而在另两张门后，则是极为无趣的什么玩意儿，比如说塞尔维亚文的莎翁全集。假设你选择了 1 号门。这时，样本空间由以下的 3 个可能结果构成：

玛莎拉蒂在 1 号门后。
玛莎拉蒂在 2 号门后。
玛莎拉蒂在 3 号门后。

每种可能的发生概率都是 1/3。我们假定大多数人都对玛莎拉蒂更感兴趣，因此第一种结果是获胜的结果，而你猜对的可能性为 1/3。

根据问题所说，接下来发生的，就是庄家——他知道每扇门后面是什么东西——打开了那两张未被选中的门之一，其后放着的是那两套莎翁全集中的一套。在打开这扇门时，庄家实际上已经根据他所掌握的信息，来避免打开停着玛莎拉蒂的那扇门。因此，这一行动并不是完全随机的过程。现在考虑两种情况。

第一种情况：你一开始就选对了。我们称这种情况为“幸运一猜”。这时，庄家将随机地打开 2 号门或 3 号门，而如果你在庄家开门后改变选择，那你就成为了托拉克（Torlak）方言版《特洛埃勒斯与克雷雪达》（*Troilus and Cressida*）的主人，而不是爽爽地飙一次车。在“幸运一猜”中，不更换选择的结局更好——但发生“幸运一猜”这个情况的概率，却只有 1/3。

另一种情况则是你开始时选错了，我们称之为“错误一猜”。你开始时选错的可能性是 2/3，因此，出现“错误一猜”的机会，要比出现“幸运一猜”的机会大上一倍。“错误一猜”的情况与“幸运一猜”相比有何不同呢？在“错误一猜”中，那辆玛莎拉蒂就藏在你没有选的那两扇门中某一扇的后面，而在另外一扇的后面则是塞尔维亚语的莎翁全集。与“幸运一猜”不同，在“错误一猜”的情况中，庄家并非随机地打开一扇你没有选中的门。既然他不想让玛莎拉蒂暴露出来，所以他有选择性地打开了正好没有玛莎拉蒂的那张门。换句话说，在“错误一猜”中，庄家介入了迄今为止还是随机的游戏过程，而这一介入，使得过程不再随机：庄家利用他的知识，

使结果偏离了真正随机的结果，他确保你如果改变选择就一定能得到那辆迷人的红色轿车，从而打破了游戏的随机性。由于这个介入，如果你发现自己落在了“错误一猜”之中，那么改变先前的选择就能获胜，否则就输了。

总而言之：如果开始选中了“幸运一猜”（可能性为 1/3），那么不改变选择就能获胜；如果开始选中的是“错误一猜”（可能性为 2/3），那么由于庄家的行为，你可以通过改变选择而获胜。因此，最后的决定在于下面这个猜测：你到底是落在了哪种情况之中？如果你觉得是第六感或命运指引你作出了最初的选择，那或许就不该更改选择。但除非你能用脑电波把一把银匙子弯成脆饼干，那么落在“错误一猜”与落在“幸运一猜”这两者的可能性之比为 2 比 1，因此你最好还是改变主意并更换选择。电视节目的统计数据证实了这一论断：身处该问题所述局面的人之中，相比于不改变选择的人，那些改变选择的人获胜的次数差不多是前者的两倍。

理解蒙提·霍尔问题是很困难的，因为不仔细思考的话，庄家所扮演的角色就会被忽视，正如同你忽视母亲所做的一切那样。但庄家实际上将这个游戏修整得对游戏者更为有利。为了使庄家扮演的角色更明显，让我们假设游戏中有的不是 3 张门，而是 100 张门；你仍然选择其中一扇，但现在选对的可能性就只有 1/100，而玛莎拉蒂藏在剩下的某扇门后面的可能性则是 99/100。跟前面一样，庄家将剩下的 99 张门中的 98 张都打开，并确保这些门后都不是玛莎拉蒂。现在，你就该感谢庄家的介入了，因为只剩下 1 张门来代表那未被选中的 99 张门，而玛莎拉蒂藏在门后的可能性则是 99/100！

如果蒙提·霍尔问题出现在卡尔达诺的时代，他会是玛里琳还是爱尔特希？样本空间定律对这个问题的处理很漂亮，但我们却无法确知卡尔达诺到底会是哪一个，因为该问题最早的（以另外一个名字的）表述，出现于 1959 年马丁·加德纳（Martin Gardner）发表在《科学美国人》（*Scientific American*）杂志的一篇文章中[16]。

加德纳称之为“一个把人搞糊涂的极好的小问题”，并称“相比于数学领域中其他任何一个分支，概率理论最容易使专家犯下大错”。当然，对数学家而言，犯下这个错误是件很没面子的事情，但对赌徒而言，却根本就是饭碗攸关的问题。因此，由卡尔达诺这位赌徒来搞出第一个系统的概率理论，倒是再恰当不过了。

卡尔达诺十来岁的某一天，他的一位朋友突然死去了。几个月后，卡尔达诺注意到这位朋友的名字不再被任何人所提及。这事儿使他悲伤而且印象深刻。生命不过是过眼云烟，而又有谁能改变这个事实呢？在他看来，要改变这一点，就必须在自己身后留下些什么东西——要么是一名继承人，要么是不朽的成就，或者两样都有。在自传中，卡尔达诺描述了他那希望在世上留下印记的“不可动摇的野心”[17]。

获得医学学位后，卡尔达诺回到米兰找工作。他在大学时写了一篇《论医生之互不相同的观点》，在文中，他把当时的医学干脆地称为一堆骗人的东西。米兰医师学院在他找工作时还了这份情：他们拒绝承认他的医学学历。这意味着他无法在米兰行医。因此，卡尔达诺用他剩下的靠教书和赌博攒下的积蓄，在米兰东边的小镇皮奥维·迪·萨科（Piove di Sacco）买了座小屋。当时镇上疾病肆虐，而且又没有医生，因此，他满心在这里做笔好买卖。不过这个市场调研却有着致命的缺陷：镇上之所以没有医生，是因为人们更愿意找巫师或牧师来治病。在数年刻苦的工作学习后，卡尔达诺发现他的收入几乎为零，而空闲时间倒是颇多。后来发生的事情证明，这段时间对他而言是一个幸运的修整期，因为他利用这一机会开始立言著书，其中之一就是《论赌博》。

在萨科呆了 5 年后，卡尔达诺于 1532 年搬回了米兰，希望在那里将著作出版，并再次申请加入医师学院。他又一次在两条战线上遭到了彻底拒绝。“在那些日子里，”他写道，“我感到如此疲惫，以

至于我会去拜访算卦的和玩弄巫术的家伙们，希望能为那一个接一个的麻烦找到解决之道。”[18]一个巫师建议他将月光遮挡在外，另一个则建议他在起床时打 3 个喷嚏并敲打木头。卡尔达诺遵行了所有这些处方，但没有一个能改变他的霉运。因此他只好戴上头巾，半夜里鬼鬼祟祟地从一幢房子跑到另一幢房子，秘密医治那些或者负担不起执业医生的诊疗费或者在他们的治疗下不见好转的病人。他在自传中还写道，为了弄一点额外的金钱来贴补这样辛苦赚来的收入，他“被迫再次拿起了骰子，以求养活我的妻子。在骰子上，我的知识战胜了运气，而我们也才得以买到食物，继续生活，尽管我们租住的地方仍是家徒四壁。”[19]至于《论赌博》，他在今后的岁月中不断进行审校改进，却再也不曾试图将它出版了。这也许是因为他认识到，把每个人都教得跟自己一样善于赌博，可不是什么好主意。

卡尔达诺最终达成了他的人生目标，他有了继承人，也有了名声——当然还有大笔的财富供其攫取。他将大学时的论文《论医生之互不相同的观点》改头换面，把原来那个带股子学究气的标题，换成了铿锵有力的《论普通用药中的不良做法》，并成书出版。该书畅销一时，而财富从此也开始慢慢增长。后来，他那些秘密病人之一的某知名奥斯丁会小修道院院长，病情突然（而且从各种可能性来看，都应该是碰巧）好转了。他将痊愈归功于卡尔达诺的照料，卡尔达诺作为一名医生的名声也因此扶摇而上，一飞冲天，以至于医师学院不得不授予他学院成员的地位，并任命他为院长。与此同时，他出版了更多的书籍，销量都还不错，特别是为普通大众所写的《算术练习》（*The Practice of Arithmetic*）一书。数年后，他出版了一本技术性更强的《*Ars magna*》——《伟大的艺术》。这是一部代数论著，其中给出了第一个对负数的清晰描述，以及对若干代数方程的著名分析。到他 50 出头即 1550 年代中期时，卡尔达诺到达了人生的巅峰，成为了帕维亚大学医学院的主席，也成为了一名

有钱人。

不过他的好运没能持续多久。很大程度上，使卡尔达诺走下坡路的，正是他遗产的另一半——他的孩子们。女儿基娅拉（以他母亲的名字命名）在16岁那年勾引了大儿子吉奥瓦尼并因此怀孕。尽管流产成功，但她却丧失了生育能力。这于她倒没啥不好，因为她在乱交方面实在称得上大胆，甚至在婚后也是如此。最终她感染了梅毒。吉奥瓦尼后来也成了一名医生，但很快就作为一名卑鄙的罪犯而更加出名了。他因此不得不近似被勒索地与一个采金者家庭联姻，因为他们掌握了他用毒药谋害一名低级市政官员的证据。与此同时，卡尔达诺那从小就热衷于虐待动物的小儿子阿尔多，将这热情变成了一份差使——宗教裁判所中自由职业式的拷问者。与吉奥瓦尼一样，他还是名兼职骗子。

结婚数年后的一天，吉奥瓦尼把一瓶神秘的混合物交给了他的一名佣人，要他把这瓶东西混入为妻子所做的蛋糕中。他妻子在享用了这份甜点后就倒下了。尽管吉罗拉莫花了一大笔钱聘请律师，尽管他想方设法去暗中操纵，尽管他为儿子提供了有利的证言，但当局把种种证据联系起来后，年轻的吉奥瓦尼仍然在不久后于狱中被处决。资金与名誉都耗尽之后，卡尔达诺在老对头手下就不堪一击了。米兰元老院将他从获准演讲的名单中抹去，指责他犯下了鸡奸与乱伦的罪行，并将他逐出了这个省份。当卡尔达诺在1563年末离开米兰时——他在自传中写道——他被“再次贬为了一个穷光蛋，财富消失了，收入没有了，租金被扣了，书也被没收了[20]。”他的思想这时也开始消失了，对他而言，生活成为了不连贯的片段。一个名叫尼可罗·塔塔利亚（Niccolò Tartaglia）① 的自学成才的数学家给了他最后一击，因为他对卡尔达诺在《*Ars magna*》一书中公开了他求解某些方程的秘方而怒不可遏。他哄着阿尔多提供了不利于

① 塔塔利亚（1499～1557）是意大利数学家，一元三次方程解法的发现者。

其父的证据，而作为交换，阿尔多得到了博洛尼亚城公共拷问者与行刑人的官方委任。卡尔达诺被短暂地关押了一段时间，然后静悄悄地在罗马度过了人生的最后几年。《论赌博》最终于 1663 年出版，此时距年轻的卡尔达诺第一次将这些文字书于纸上，已经过去了 100 多年，而且其他人也已重新发现并超越了他的分析方法。

第四章

追寻通往成功的路径

在卡尔达诺的时代，如果有赌徒了解了他关于机遇性的数学成就，就能跟不那么聪明的对手赌博并赚到可观的利润。换了在今天，卡尔达诺同样也能写出诸如《与容易上当的笨蛋玩扔骰子之傻瓜指南》这样的书，并靠它名利双收。但在他自己的时代，卡尔达诺的工作却没有激起多少涟漪，《论赌博》一书在他死后很久也未能出版。为什么卡尔达诺工作的影响如此之小呢？正如已经提到过的，对他的前人而言，通往概率理论的障碍之一，在于缺乏一个良好的代数符号体系；到卡尔达诺那时候，虽然符号系统有了改进，但也只能算处在婴儿时期。另外还有一个障碍需要去除：卡尔达诺的工作进行于一个人们认为神秘咒语比数学运算更为有用的时代中。如果人们没打算找寻自然的秩序，也不曾想过创造对事件的数值描述，那阐述随机性对这些事件之影响的理论，自然也就注定不为人赏识了。后来的历史表明，如果卡尔达诺再多活个几十年，他能达到的成就及这些成就所能得到的接受程度，就会完全是另一码事，因为在他死后的几十年中，欧洲的思想与信念发生了历史性的转变，这个转变常常被尊称为科学革命。

对于刚摆脱中世纪桎梏的欧洲而言，科学革命是对当时流行思维方式的一次反叛。在中世纪这个刚刚成为历史的时代中，人们关于世界运转方式的观念，还未曾经受过任何系统性的追根究底。某个小城的商人把被绞死者的衣服偷走，相信这有助于啤酒的销售；而另一个小城教区的居民则相信，裸体围着教堂中的祭坛绕圈并唱

着亵渎的祈祷，可以治愈疾病[1]。有人甚至相信在“错误”的厕所中如厕会带来坏运气。事实上，该名证券交易员是在 2003 年向 CNN 记者坦白这个秘密的[2]。是的，有些人至今仍然迷信，但到了今天，只要你有兴趣，至少我们已经拥有了用来证明或证伪类似行为之功效的智力工具。不过，对于卡尔达诺同时代的人来说，如果他们扔骰子赢了，他们也不会对这些经验加以数学分析，而是跑去做一次感恩祷告，或者从此再也不洗那双幸运袜子。卡尔达诺自己也相信，连续的失败之所以发生，是由于“幸运之神有些不快”了，而要改变此种状况，方法之一就是把骰子狠狠地扔一下。如果幸运之 7 彩票的结果可玩弄于股掌之间，干吗还要对数学卑躬屈膝呢？

常被视为科学革命转折点的那个时刻，出现在 1583 年，距卡尔达诺之死才不过 7 年。比萨大学（University of Pisa）一名年轻的学生正坐在大教堂中。传说他并没有聆听教堂中进行的仪式，而是怔怔地看着一个东西，一件他觉得比那些仪式更有启发性的事情：一座巨大吊灯的摆动。伽利略（Galileo Galilei）① 用自己的脉搏计时，发现吊灯摆过一个长弧形所用的时间，似乎跟它摆过一个较短弧形的时间是一样的。这个观察似乎暗示着：摆完成一次摆动所需的时间，与摆动的幅度无关。伽利略的观察精确而实用，它虽然很简单，但却代表了一条描述物理现象的新道路：科学必须专注于经验与实验，即大自然的运作方式，而不应集中于直觉所认定或是对我们更具吸引力的什么东西。最重要的是，这个描述必须通过数学来完成。

应用其科学技巧，伽利略写了一篇有关赌博的短文《骰子游戏之思考》。这项工作是在其资助人托斯坎尼（Tuscany）大公的要求下完成的。困扰大公的是这样一个问题：为什么扔 3 次骰子，得到

---

① 伽利略（1564～1642）是意大利数学家、物理学家、天文学家、科学家和哲学家，近代实验科学的先驱。他推翻了亚里士多德“物体下落速度与重量成正比”的错误学说，同时还是天文望远镜的发明者。

总点数为 10 点的情况比得到 9 点的情况出现得更频繁？这个频率上的差别只有区区 8%左右，而且，扔出 10 点或 9 点的情况都并非很常见，因此，大公肯定在这个骰子游戏上花费了足够多的时间才能发现这个微小差别，而这似乎意味着大公更需要的应该是 12 步赌瘾戒除计划而非伽利略。不管原因如何，伽利略对此问题并没多大兴趣，也不免嘟嘟囔囔牢骚满腹。但正如任何一个不想丢掉饭碗的顾问那样，他咽下了嘟囔，干了他该干的活。

单独扔 1 次骰子，得到任何一个特定点数的机会是 1/6。但如果扔 2 次骰子，各个可能的总点数出现的机会就不再相等了。例如，扔出 2 点的机会是 1/36，而扔出 3 点的机会则要高 1 倍。造成此种差异的原因在于，只有一种情况能扔 2 次骰子而得到 2 点——2 个都是 1 点；而要扔出 3 点，则有 2 种情况能满足要求——先扔出 1 点再扔出 2 点，或者先 2 点后 1 点。这个思路将我们引向了理解随机过程的下一大步，也就是本章的主题：分析事件不同的出现方式之数量的系统性方法的建立。

要理解大公的困惑，关键在于我们应该把自己想象成一名犹太法典（Talmud）学者，然后再来看这个问题：与其试图去解释为什么 10 点比 9 点出现得更频繁，不如问问为什么 10 点不应该比 9 点出现得更频繁？有一个让人很容易接受的理由，来说明 10 点和 9 点应当出现得同样频繁：我们都能以 6 种方式，来扔 3 次骰子而得到 10 点和 9 点。扔出 9 点的方式是（621）、（531）、（522）、（441）和（333）；对于 10 点则是（631）、（622）、（541）、（532）、（442）和（433）。根据卡尔达诺的样本空间定律，得到一个好结果的概率，等于所有好结果在所有可能结果中所占的比例。既然 9 点和 10 点都能通过同样多的方式来构成，为什么一个会比另一个更可能呢？

原因在于——如我所言——原始形式的样本空间定律，只适用于具有等可能性的结果，而上述的点数组合，其发生的可能性却并

不相等。比如（631）这个结果——即扔出 1 个 6 点、1 个 3 点和 1 个 1 点——发生的可能性，6 倍于扔出（333）这个结果的可能性，因为扔出 3 个 3 点的方式只有 1 种，而扔出 1 个 6 点、1 个 3 点和 1 个 1 点的方式却有 6 种：先 6 后 3 后 1，或先 1 后 3 后 6，如此等等。现在，我们换种方式来表示扔骰子的结果：我们用逗号隔开的三元组来表示投掷结果，并以此来保留扔骰子的次序。简单来说，（631）这个结果由以下的可能组合所形成：（1，3，6）、（1，6，3）、（3，1，6）、（3，6，1）、（6，1，3）和（6，3，1），而（333）这个结果则只能是（3，3，3）。对扔骰子的结果如此分解后，新形式的结果就具有了相等的可能性，而样本空间定律也就可以使用了。这时，扔出 10 点的方式有 27 种，而扔出 9 点的只有 25 种，伽利略由此得出结论，即扔 3 次骰子得到 10 点的可能性，是得到 9 点的可能性的 27/25，即约 1.08 倍。

这个问题的解决过程中，伽利略隐含地使用了我们的下一条重要原则：一个事件发生的可能性，依赖于这个事件所有可能发生方式的数量。这并不惊人，真正让人吃惊的，是它能造成的影响之大，以及计算这个量的困难程度之高。假设全班 25 名 6 年级学生们做了一次有 10 道判断题的小测验，让我们来算算某个学生可能达到的测验结果有多少种：她可能 10 题全对；她可能只错 1 题——出现这种情况的方式有 10 种，因为总共有 10 道题供她挑出一个来犯错；她可能只错 2 题——这种情况对应了 45 种方式，因为这 10 道题能组成 45 对不同的题目构成的题对；等等。作为这些组合方式的结果，假设一群学生完全都是靠碰运气来猜答案的话，那么平均而言，对应着每一个得满分的学生，将会有 10 个得 90 分的学生以及 45 个得 80 分的学生。得到 50 分上下的可能性当然就更高了。对于有 25 名学生的班级，如果也都是靠猜的话，那么至少有一名学生能得到 B（80 分）或更高分的概率，差不多是 75%。因此，一名资深教师在多年从教中所碰到的那些考前不好好复习、测验时多少在赌一把的

学生中，很可能会有一些人的最后得分为 A 或 B。

几年前，加拿大彩票管理部门决定，将一些以往积累下来的无人认领奖金返还给彩民，而在这个过程中，他们就结结实实地领教了认真数数的重要性[3]。他们购买了 500 辆小汽车作为奖品，然后编写了一个计算机程序，从 240 万个彩票订户号中随机抽出 500 个，以决定获得奖品的赢家。官方公布了这 500 个未排序的获胜号码，并承诺每个号码都能赢得一辆汽车。使他们大为尴尬的是，有一个人（正确地）宣称他赢得了两辆汽车。官员们哑然失色——计算机怎么可能从 200 多万个数字中，将同一个数字随机挑中两次呢？是不是程序有问题？

这些彩票官员碰到的计数问题，等价于另一个被称为“生日问题”的情形：如果希望一群人中有某两人在同一天生日（假定每天出生的可能性都相等）的可能性大于 50%，这个群体至少应该包括多少人？大多数人都觉得答案应该是一年中天数的一半，也就是大约 183 人。不过这是另一个问题的正确答案：如果要一群人中有某个人与你同一天生日的可能性大于 50%，这个群体至少应有多少人？假如没有限制到底哪两个人同一天生日，则形成这一对生日相同的人的许多组合方式，就戏剧性地改变了答案。实际上，这个数小得令人吃惊：仅仅只需要 23 个人就够了。如果像加拿大彩票的情况那样，从 240 万个数字中挑数字，我们当然需要远多于 500 个数字，来保证出现重复数字的可能性大于 50%，但仅仅是 500 个数字中发生重复的可能性，也绝非可以忽略不计：实际上，这个概率大约是 5%。不大，但仍然需要解决，比如这样：计算机每抽取一个数字后，就与所有已抽出的数字进行比对，以确定是否发生重复抽取。根据记载，加拿大彩票部门要求这个幸运儿放弃他所得汽车中的一辆，但遭到了拒绝。

另一个令许多人瞪大眼睛的彩票神秘事件，发生在 1995 年 6 月 21 日的德国[4]。这个怪异的事情产生于一个称为乐透 6/49 的彩票

抽奖中，中奖的 6 个号码从数字 1 到 49 中随机抽得。这天的中奖号码是 15—25—27—30—42—48，而在 1986 年 12 月 20 日，完全相同的中奖号码也被抽到过。在本彩票 3016 次抽奖的历史上，出现重复的中奖序列，这可是头一遭。发生这种事情的可能性有多大呢？可并不像你想的那么糟。算一算就会发现，在这么多个抽奖的年头中，在某个时刻发生中奖号码重复的机会大概是 28%。

既然在随机过程中，某个结果的可能发生方式的数量，是决定该结果发生概率的关键因素之一，那么关键问题就在于，如何才能计算出某事发生的可能方式有多少种？伽利略似乎没有注意到这个问题的重要性，也并没有在他的骰子问题上进一步深入，而且，他文章的第一段就写道，仅仅是因为“被命令”如此，他才写下了这篇关于骰子的文章[5]。1633 年，作为伽利略提出通向科学的新方法的奖赏，宗教裁判所将他判了刑。但科学与神学从此便永远地分道扬镳了：现在，分析“如何（去做到）”的科学家们卸下了神学家们所问的“为什么（会如此）”的这个包袱。不久后，从小就接受伽利略之科学哲学教育的新一代学者中的一员，就将把可能事件计数的分析，带到全新的高度和理解水平，而如果没有这些理解，今日科学中的绝大部分都将无法开展。

随着科学革命欣欣向荣，随机性研究的前线从意大利移到了法国。在那里，抛弃了亚里士多德（Aristotle）而跟随伽利略的新型科学家们，将研究推进到了比卡尔达诺和伽利略更远也更深的境地。这一次，新成就的重要性终于为人所知，并形成了一股席卷欧洲的浪潮。新想法和新思路再次诞生于博弈这个大背景下，但这些科学家中的第一人，却更大程度上是由数学家变为赌徒，而不是像卡尔达诺那样从赌徒变成数学家。他的名字是布莱士·帕斯卡（Blaise

Pascal)①。

帕斯卡 1623 年 7 月出生于巴黎南边 250 英里多一点的克莱蒙费朗（Clermont-Ferrand）。他的父亲发觉了儿子的聪明才智，正好又搬到了巴黎，因此，他推介 13 岁的帕斯卡加入了一个新成立的交流团体。团体成员们都称之为梅赛纳学院——以创建该团体的黑袍修士的名字（Marin Mersenne）命名。梅赛纳的小团体包括著名哲学家与数学家笛卡儿，以及业余数学天才费马（Pierre de Fermat)②。这些睿智的思想家与自大狂的奇特混合——再加上梅赛纳从中搅和——必定对十来岁的布莱士产生了巨大的影响。他与费马和笛卡儿有了私人往来，并迅速认识到科学新方法的深层基础。“让所有亚里士多德的信徒们……”他写道，“认识到实验才是物理学中必须遵从的真正主人。”[6]

但一个有着神圣信仰的乏味的书呆子，怎么会跟城里的赌博圈扯上关系呢？帕斯卡一直断断续续地经受着胃痛，他吞咽困难，而要让吞下的东西不再冒出来也不容易，此外，他还遭受着使人体质变差的虚弱、严重的头痛、一阵阵地冒汗以及腿部部分瘫痪的折磨。他坚忍地遵从医生的建议，包括放血、服泻药以及驴奶和其他“令人作呕的”药剂，而这些药剂差不多又都被吐了出来——按他姐姐吉尔贝特的说法，完全是一种“分毫不差的虐待”[7]。帕斯卡这时已经离开了巴黎，但在 1647 年夏天，已经 24 岁并开始绝望的他，又和妹妹杰奎琳再次归来，希望能找到更好的治疗。新医生们提出了一个在当时而言最为先进的建议：帕斯卡“应该放下一切持续的精神负担，并尽可能多地找乐子来散心。”[8]因此，帕斯卡教会了自己

---

① 帕斯卡（1623～1662）是法国数学家、物理学家、哲学家和散文家，主要贡献有射影几何中的帕斯卡定理、帕斯卡三角形以及物理学中的帕斯卡定律。为纪念他的贡献，帕斯卡的名字又被用作压强的单位。

② 费马（1601～1665）是法国数学家，被誉为“业余数学家之王”，其贡献涉及解析几何、微积分、概率论、数论和光学。尤以数论中的费马大定理而广为人知。

好好休息和放松，并开始在其他追求享乐的年轻人的陪伴下打发时间。1651 年，布莱士的父亲去世，帕斯卡也突然间成了一个 20 来岁的遗产继承人。他很好地使用了这笔财产，至少从他医生建议的角度而言是这样。传记作家们称 1651～1654 年这几个年头为帕斯卡的“入世时期”，而吉尔贝特则称之为“他生命中最被糟蹋的时光”。[9]尽管在自身思想提升方面进行了一些努力，他的科学研究却毫无起色。不过根据记述来看，他此时的健康状况倒是达到了最佳状况。

在历史上，有关随机性的研究，常常得益于一些本身就是随机的事件。帕斯卡的成就便是这样一个例子：正是由于他放弃了研究，才最终使他走向了对机遇的研究。所有这一切，都源于一个派对伙伴把他介绍给了 45 岁的势利之徒安托万·贡博（Antoine Gombaud）。拥有德·梅勒骑士头衔的贵族贡博一直自视为调情大师，而他那罗曼史纠葛的长长目录，也表明他的确无愧于这一称号。但德·梅勒还是一名喜欢下重注的专家级赌徒，而且他赢得如此频繁，以至于有人怀疑他作弊。一次，他深深困惑于一个小小的赌博问题，便就找到了帕斯卡来寻求帮助。德·梅勒由此引发的研究，使帕斯卡的科研枯竭期终告结束，也使德·梅勒确立了在思想史上的地位，并解决了那个伽利略的大公扔骰子问题中遗留的未解之处。

那是 1654 年。德·梅勒提出的问题称为“点数问题”：假设你和对手来玩一个游戏，其中你们俩得分的机会均等，而总分首先达到某个给定值的一方为赢家。当某方领先时，游戏被中断，以便赌徒们对最终胜者进行下注。那么，怎样才能最公平地定出赔率呢？德·梅勒注意到，在给定了游戏中断时两名游戏者各自获得的分数时，这个确定赔率的方法，应能反映每名游戏者最终获胜的概率。但如何计算这两个概率呢？

帕斯卡认识到，不管答案是什么，其计算方法在当时都属未知。而不管这个方法是什么，它都可能在任何类型的竞争场合中具有重要的应用意义。但正如在理论研究中常常发生的那样，帕斯卡发现

自己对想要采取的方法也并不确定，甚至还感到困扰。最终，他决定寻找一位合作者，或至少是一位他可以与之讨论想法的数学家。梅赛纳这位了不起的交流者此时已亡故数年，但帕斯卡与梅赛纳学院的网络仍保持着联系。因此，从 1654 年起，就肇始了数学史上最伟大的书信往来之一——帕斯卡与费马之间的系列通信。

1654 年时，费马在图卢兹的 Tournelle（刑事法庭）身居高位。开庭时，你也许能看见一个身着华丽长袍的费马，判处罪人以火刑。但闭庭时，他就会将其分析技巧转向更为温柔一些的数学探索。费马也许只是个业余爱好者，但他通常都被认为是所有时代中最伟大的业余数学家。

费马并不是靠什么特别的野心或成就而获得他的职位的，他的方法相当老套：当他的上级一个个死于瘟疫时，他就一步步稳健地爬了上去。事实上，当帕斯卡的信寄到费马手中时，他本人也刚从一次瘟疫中慢慢康复。费马的朋友伯纳尔·梅东（Bernard Medon）甚至还宣告他已病故。不过他却并没有死，而尴尬却可以想见颇为高兴的梅东，则收回了他的公告。不过，费马无疑曾一只脚踏进了鬼门关。虽然费马比帕斯卡年长 22 岁，但他实际将比这位新的通信伙伴还要晚几年去世。

正如将要看到的，在生活中任何两者间存在竞争的领域里，都会出现这个点数问题。在帕斯卡和费马的通信过程中，他们各自发展出了一套方法，并解决了几个不同版本的点数问题。但后来的事实证明，帕斯卡的方法更简单——甚至可以称之为完美，而且又足够普适，能应用于许多我们日常经历中遇到的问题。由于点数问题最初出现于赌赛中，因此，我将以一个体育比赛的例子来加以说明。在 1996 年的世界职业棒球大赛中，亚特兰大勇士队在头 2 场比赛里都击败了纽约扬基队。而最先获得 4 场比赛胜利的球队将被加冕为冠军。勇士队赢得头 2 场比赛的事实，并不必然意味着他们实力更强，但不管怎样，这可以被看作一个征兆，一个勇士队球员们的确

更强一些的迹象。不过，考虑到我们希望达到的目的，我们将维持以下假设，即在一场比赛中，两队具有相等的获胜可能性，而头 2 场比赛不过正好是勇士队赢了而已。

在这个假设下，扬基队获胜的恰当赔率是多少呢？也就是说，扬基大翻盘的可能性是多大？为了计算这个结果，让我们来数一数所有扬基队获胜的可能方式，然后再跟他们落败的可能方式进行比较。系列赛的两场比赛已经打完了，那么还剩下 5 场可能的比赛。既然每场比赛有两种可能结果——扬基获胜（Y）或勇士获胜（B），那么就有 2^5 即 32 种可能的系列赛结果。比如说，扬基队可能赢 3 场输 2 场：YYYBB；或者他们可能交替获胜：YBYBY。（在后一种情况中，由于勇士队在第 6 场比完后就已经获得了 4 场胜利，那最后一场比赛就不再会进行了。不过我们等下再来谈这个问题。）因此，扬基队翻盘获得冠军的可能性，就等于这些比赛结果序列中，他们至少赢得 4 场比赛的个数，再除以这些可能序列的总数 32；而勇士队获得冠军的可能性，则等于他们至少赢得 2 场比赛的序列个数除以 32。

这种计算方式看起来有点怪，因为如所提及的，序列中包括了（如 YBYBY 这样的）两队在勇士队赢得所需的 4 场胜利后仍然继续系列赛的情况，可一旦勇士队赢了 4 场，那么这两支队伍肯定就不会再比第 7 场了。但数学可不会搭理人的这些怪规矩，不论球员们比赛或不比赛，都不会影响到这种结果的存在性。比方说，假设你玩扔硬币，如果扔出正面就算赢。那么，扔两次硬币的可能结果有 2^2 即 4 个：正反、正正、反正和反反。如果扔出的是前两个结果，那么就不用再扔第二次硬币，因为第一扔时你就已经赢了，但你获胜的概率仍然是 3/4，因为 4 个可能结果中有 3 个包含了一次头朝上。

因此，为了计算扬基队和勇士队各自获胜的可能性，只需简单地数一数，系列赛剩下的 5 场比赛，能产生哪些可能的比赛结果序列。首先，扬基队如果在 5 场中胜 4 场，那他们就能获胜，而这个

结果，可按以下 5 种方式之一来发生：BYYYY、YBYYY、YYBYY、YYYBY 或 YYYYB。此外，扬基队还可赢下所有这 5 场比赛而凯旋班师，此时只有一种情况：YYYYY。再来看看勇士队：如果扬基队只赢了 3 场，那他们就是冠军，而这种结果对应了 10 种可能方式（BBYYY、BYBYY、……），或者只赢了 2 场（也同样有 10 种可能的发生方式），或者只赢了 1 场（5 种可能发生方式），或者连一场也赢不了（只有 1 种可能发生方式）。将这些结果加起来，我们发现，扬基队获得冠军的可能性是 6/32，或者说约 19%，而勇士队则为 26/32，即约 81%。根据帕斯卡和费马的说法，如果系列赛在第 2 场后被中断，那么所下赌注就应以 19：81 的比例加以瓜分，而这也是此时为扬基队获胜所应设置的赔率。最终，扬基队确实来了个大翻盘，连胜了接下来的 4 场比赛并加冕为冠军。

同样的推理方法也可用于系列赛开始之时——即还没有进行任何一场比赛之前。如果两队赢得一场比赛的可能性相等，那么可以发现，他们——理所当然地——具有相同的赢得整个系列赛的机会。不过，类似的推理还同样可以用于单场比赛获胜机会不相等的情况中，这时，只需将原来简单地累计所有可能方式的个数，稍加修改为对每个可能方式，根据其相对概率乘上一个权值即可。在系列赛开始时进行这个分析就能发现，在一个 7 场比赛构成的系列赛中，较弱的队伍也有相当大的夺冠机会。例如，假设一支好队伍能保证在 55%的比赛中击败对手，那么即便如此，弱队也仍能在 10 次 7 场系列赛中获胜 4 次；如果平均而言，较强队在与较弱的对手交手时，每 3 次中可以获胜 2 次，那么在这样的 7 场系列赛中，弱队仍然可以赢得 5 次系列赛中的 1 次。不过说实话，体育联赛中的确没有什么办法来改变这种情况。在 2/3 胜率这个机会不均等的例子中，需要通过至少有 23 场比赛的系列赛来决定冠军归属，才能达到所谓的统计显著性，即弱队在这样的比赛中，只有 5%或更少的机会赢得冠军（见第五章）。而对于 55：45 这样的强弱比例，最短的具有统计

显著性的“世界职业棒球大赛”需要比 269 场比赛。好个冗长的赛事！因此，体育锦标赛非常有趣而刺激，但“世界冠军”的桂冠，却并不是说明某队确实是最强者的可靠标志。

如我所言，同样的方法并不仅限于游戏、赌博与体育比赛。如果两个公司展开狭路相逢的竞争，或者一个公司中的两名职员之间也一较高下，那么，尽管每个季度或每个年度都会出现赢家和输家，但是，要通过简单的胜负记录，来可靠地确定哪个公司或职员更好一些，可得花上几十年乃至几个世纪才行。比方说，如果职员 A 确实更好一些，而且长期来看，他在 100 次业绩竞赛中有 60 次胜过了职员 B，那么在 5 战 3 胜制的比较中，较差的 B 仍有将近 1/3 的机会获胜。利用短期结果来评估能力，是相当危险的。

以上的计数问题都很简单，不费什么力气就可以干好。但是，当涉及的数量更大时，计数过程就会变得十分困难。考虑下面这个问题：你正在筹备一场婚礼中 100 位宾客的接待工作，而每张桌子能安排 10 个人。你不能让罗德表哥和你的朋友艾米坐在一起，因为 8 年前他们发生过感情纠葛，并且艾米把罗德给甩了；另一方面，艾米和利蒂希娅都想挨着你那身材健美的鲍比表弟坐；露兹姨妈则最好坐到没有隔墙之耳的那一桌去，否则那些争风吃醋的调情，恐怕会在接下来的 5 年中，为假日聚餐时的流言蜚语提供源源不断的原料。你仔细考虑着这种种情况。先仅仅考虑一下第一张桌子吧。从 100 个人中挑出 10 个来，有多少可能的挑选方式？这个问题也等价于问，从 100 支基金中挑出 10 个来投资，或者从一块硅晶体的 100 个位置中挑出 10 个来放锗原子，有多少种挑选方式？类似问题一次又一次地出现于随机性理论中，而不仅仅只是在前述的点数问题中才存在。当数量很大时，一五一十地列出所有可能性来计算可能方式的个数，这种做法冗长乏味或根本就不可行。帕斯卡的真正成就就在于此：他提出了一种普适而系统的计数方法，使我们可以通过公式或读图表的方式，来得到这类问题的答案。该方法根基于一种

奇特的三角形数字排列。

作为帕斯卡工作核心的这个计算方法，其实在 1050 年前后，就已经被中国数学家贾宪①所发现，并在 1303 年由另一位中国数学家朱世杰②发表；卡尔达诺在 1570 年也同样对之进行了讨论。但帕斯卡将它整合到了更大的概率理论的整体中，因此发现这个方法的功劳大部分都归给了他[10]。不过这些已有的成果并没有让帕斯卡感到不安。“别说我做的事情里什么新玩意都没有，”帕斯卡在自传中争辩道，“这些数字的排列就是新的。打网球时，我们击打的都是同一个球，但总有某个人击出的落点更好。”[11]如下这个帕斯卡所发明的图形因此被称为帕斯卡三角形③。在图中，我将帕斯卡三角形截止在第 10 行上，但这个三角形可以向下一直无限延伸。实际上，要扩展这个三角形的做法很简单：除了处在顶端的 1 之外，三角形中的每一个数，都是它左上方与右上方数字之和（如果这两个位置上没有数字，就以 0 代替）。

将若干个物体从有着同样多或更多物体的集合中取出，如果想知道有多少种取法的话，帕斯卡三角形会是一个非常有用的工具。在婚礼宾客问题中，是这么来干的：要计算从 100 位宾客中挑出 10 位来安排他们的座位，有多少种不同挑选方式的话，我们先沿着图中左侧的数字一直往下数，直到标有 100 的那一行为止。这里的三角形并没有画得那么深，不过我们姑且假装它画到了好了。第 100 行的第一个数字给出的，是从 100 名宾客中挑出 0 位的挑选方式个

---

① 贾宪（约 11 世纪）是我国北宋数学家，主要贡献是创造了“贾宪三角”（后因杨辉抄录又称“杨辉三角”，也即文中所说的帕斯卡三角形）和增乘开方法。

② 朱世杰（1300 前后）是我国元朝数学家和科学家，主要著作有《算学启蒙》和《四元玉鉴》。

③ 帕斯卡三角形在我国数学文献中称为杨辉三角形，系由我国南宋数学家杨辉（13 世纪中叶）在《详解九章算法》中给出。——编者注

数，当然，这样的方式只有 1 种：谁都不选。这个答案对于任意多数目的宾客而言同样正确，也正因为如此，每行的第一个数字都是 1。第 100 行的第二个数字说的，则是从 100 名宾客中挑出 1 位的挑选方式个数，而这样的方式有 100 种：你可以仅仅挑宾客 1 号，或仅仅挑宾客 2 号，如此等等。这个推理同样适用于每一行，因此每行的第二个数字都简单地等于该行的序号。该行中的第三个数字，则表示从 100 名宾客中挑出不同的两位的挑选方式个数，其余依此类推。因此，我们要找的答案——选出 10 名不同的宾客的方式数量——就是第 100 行的第十一个数字。不过，即使我们真把帕斯卡三角形画得深到包含了第 100 行，这个数字也大得无法在一页中全部写出。实际上，如果有来宾抱怨座位安排的话，你可以告诉他，要每一种可能情况都考虑到的话，可得要多长的时间啊：假设在每种情况上花 1 秒钟来考虑，那么得用 10 万亿年才能考虑所有的可能情况。当然了，这位不高兴的客人肯定会认为你是在装腔作势。

行											
0						1					
1					1		1				
2				1		2		1			
3			1		3		3		1		
4		1		4		6		4		1	
5	1		5		10		10		5		1
6		1		6		15		20		15	
7			1		7		21		35		
8				1		8		28		56	
9					1		9		36		
10						1		10		45	

帕斯卡三角形

为了能用上这里画出的帕斯卡三角形，现在假设宾客名单中就

只有 10 个人。这时，与问题有关的，就是所给帕斯卡三角形中最底下标着 10 的那一行。行中的数字就表示 10 位宾客中的 0 位、1 位、2 位……宾客的不同坐席安排方式的数量。你也许看出来了，这些数字在前面那个 6 年纪测验的例子中也出现过——在 10 道判断题测验中，学生做错若干道题目的可能方式之数量，就等于从 10 名宾客中挑选若干名来安排座位的可能方式之数量。这就是帕斯卡三角形如此有力的原因之一：同样的数学可以用于许多不同的场合。至于扬基对勇士的世界职棒大赛的例子，我们原来只是乏味地数出了剩余 5 场比赛的所有可能结果，而现在，我们可以直接从三角形中得到扬基队获得 0 场、1 场、2 场、3 场、4 场或 5 场比赛的可能方式数：

1 5 10 10 5 1

从这行数字，一眼就可以看出，扬基队获得 2 场胜利的可能性（10 种方式）是获得 1 场胜利的可能性（5 种方式）的两倍。

一旦学会了这种方法，那突然间我们就会发现，帕斯卡三角形的应用无所不在。我的一位朋友曾为一个刚起步的计算机游戏公司工作。她常常谈到，尽管市场部主任承认小规模的小组座谈会只适合提供“仅仅是定性的结论”，她却仍然不时地报告座谈会成员达成了 4 比 2 或 5 比 1 的“压倒性的”一致意见，好像这样的数据真有什么意义似的。假设你自己来组织一个小组座谈会，会上，6 名成员将对开发中的新产品加以检查并给出意见。如果实际有半数人喜欢该产品，那么这个小组座谈会能多精确地反映出这一真实的偏好性呢？与这个问题相关的，是帕斯卡三角形中标着 6 的那一行，它表明了所有可能的包括 0 人、1 人、2 人、3 人、4 人、5 人或 6 人“小小组”的组成方式数量，这些“小小组”成员对产品表示了喜欢（或不喜欢）：

1 6 15 20 15 6 1

从这些数字可以看出，有 20 种情况中，小组成员会分为一半对一半的两群人，一群喜欢该产品，而另一群不喜欢。这种情况精确地反映了人们对新产品的真实态度。但除此之外，同样还有 1 + 6 + 15 + 15 + 6 + 1 = 44 种其他方式，它们给出的或喜欢或不喜欢的一致性意见，实际却不具代表性。因此，如果不慎重的话，被座谈会结论误导的几率就是 44/64，或说约 2/3。这个例子并非说共识完全由随机因素造成，而只是说我们不能假定这个共识具有显著性。

事后证明，帕斯卡和费马的分析是通往一致的随机性理论的第一大步。他们那著名通信的最后一封写于 1654 年 10 月 27 日。几周后，帕斯卡精神恍惚地坐了两个钟头。有些人称这种恍惚为一种神秘体验，而其他人则悲叹他终于失去了正常的心智。不管如何去描述，帕斯卡在事后就完全变了一个人，而这个改变，将引导他在随机性概念上做出另一个奠基性的贡献。

在 1662 年帕斯卡去世的几天后，仆人发现他的一件夹克上，很奇怪地鼓出一块。仆人扯开了夹克内衬，在里面找到了几张叠起来的羊皮纸和普通纸张。显然，帕斯卡在生命的最后 8 年中，每天都把它们带在身边。在这些纸上，用他的笔迹潦草地写着一系列不连贯的词语和词组，所标的日期为 1654 年 11 月 23 日。这些文字就是对那恍惚的两小时的说明，其中他描述了上帝是如何降临到他面前，并在那两小时内将他从堕落的生活方式中解救出来。

这次启示之后，帕斯卡与大多数朋友断绝了关系，并称他们为“可怕的联系”[12]。他卖掉了马车、马、家具、藏书——除了《圣经》之外的所有东西。他将钱财施舍给穷人，给自己留下的只有那么一点点，以至于不得不靠乞讨或借债才能果腹。他戴着一条朝内有尖刺的铁腰带，以便总是处于不舒适之中，而一旦他发现自己有

高兴起来的危险，就将尖刺刺入自己的身体。他结束了数学与科学方面的研究。对于儿时对几何的迷恋，他写道："我差不多都记不起世界上还有几何这么回事了。我现在认识到几何是如此地毫无用处……我很可能连想也不会再想到它了。"[13]

但帕斯卡却仍然多产。在经历了那恍惚之后的岁月里，他记下了他对上帝、宗教和生命的思索。这些思索后来出版为一本名为《沉思录》（*Pensées*）的书，时至今日仍不时重印。尽管帕斯卡对数学大加抨击，但他对人世生活之轻浮无益所做的观察，却是一个数学论述，并在其中将他的数学概率武器初试锋芒地直接应用于一个神学问题，从而做出了另一个贡献，其重要性不亚于早期在点数问题上所获的成就。

《沉思录》中的数学概括于手稿中的两页纸上。这两张纸的正反面都写满了朝向各异的字迹，到处都是删涂与更正。在纸上，帕斯卡详细给出了一个关于人对上帝尽责之好处与坏处的分析，就如同用数学求解所下赌注是否明智一般。他那了不起的创造，就是用来计算这些好处与坏处的总收支的方法，即今日被称为数学期望的概念。

帕斯卡的论述如下：假设我们承认不知道上帝到底是否存在，从而对这两种可能，我们都分别给予50%的几率。如果现在要决定是否应该虔诚生活，那应该如何给这两个几率赋以权重呢？帕斯卡认为，如果行为虔诚而上帝又确实存在，那么我们所获得的价值——永恒的快乐——将是无穷大。另一方面，如果上帝不存在，那么我们所损失的东西——或者说负回报——却很小：不过是这些虔诚被白费了而已。然后帕斯卡提出，为了将这些得失进行加权，应将每种可能结果的概率与其对应的回报相乘，并把结果相加，从而得到一种平均或说期望的盈利。换而言之，与我们的虔诚所对应的期望回报，就是无穷大的一半（上帝存在时你的收获），再减去某个小数字的一半（上帝不存在时你的损失）。对于无穷大，帕斯卡已

经了解得足够多，因此他知道无穷大的一半仍然是无穷大，故而虔诚的期望回报就是正的无穷大。结论是：任何一个有理性的人，都应该遵从上帝的律法。这个论证如今被称为帕斯卡的赌注。

期望是一个十分重要的概念，不仅在赌博中如此，在所有决策中皆然。事实上，帕斯卡赌注常被认为是博弈论——关于博弈中最优策略的定量研究——这门数学分支的成立标志。我必须承认，这个思考问题的方式让人上瘾，以至于有时我把它用得过了点头。“那个停车计时器这次要扣我多少钱?”我问儿子。标记上说是 25 美分。没错，不过我每 20 次左右就有 1 次会停车超时，而这时就能看到车上放着张单子，上面写着 40 美元。因此我解释道，这个 25 美分的咪表费用只是个残酷的诱惑，而实际开销是 2.25 美元（这额外的 2 美元，便是我那 20 次中就有 1 次的超时几率乘以 40 美元的罚款所产生的费用）。“那么我们这条车道的情况又如何?”我问儿子，“它是收费公路吗?”嗯，我们已经在现在的房子里住了 5 年了，粗略算来，我沿着这条车道回家，也有大概 2400 次了，其中有 3 次被伸出的栅栏柱撞断了后视镜，每次都花了我 400 美元。因此我儿子告诉我，这就相当于在路上放一个收费箱，每次我回来时，就得扔 50 美分进去。他完全理解了期望这个概念（他还建议我，在没有喝早咖啡前，不要送他们去学校）。

通过数学期望的透镜来看这个世界，常常会得出令人惊讶的结果。比如最近有一个通过邮件进行的抽奖，大奖是 500 万美元[14]。只需寄出一封信，就能参与抽奖并有机会赢得大奖。参与次数没有限制，但一封信只能代表一次参与。这次抽奖的赞助者显然希望能有约 2 亿人次来参加，因为本次活动那难懂的条款中指出，获胜机会是 2 亿分之 1。那么，参与这样的“免费抽奖”值不值得呢？将获胜的机会乘以它的回报，我们可以求得每次参与大概值 1/40 美元，或者说 2.5 美分——远低于所付的邮费。事实上，这场竞赛的大赢家是邮局，而如果关于参与次数的估计是准确的话，邮局将由此赚

取近 8000 万美元的邮资。

下面是另一个疯狂游戏。假设加州向它的公民们提出如下提议：对于所有那些支付了一美元或两美元的参与者们，其中绝大多数人什么也得不到，某一个人将得到一笔财富，而还有一个人会以惨烈的方式被处死。会不会有人参与这个游戏呢？有，而且还十分踊跃。这个游戏名叫加州彩票。尽管该州没有用我所说的方式来大做宣传，但事实上这个彩票就是如此行事的。在每次抽奖中，每个获得大奖的幸运儿身后，都有数百万的其他竞争者驾车往返于彩票发行点以购买彩票，其中就有一些人在路上死于交通事故。根据全国公路交通安全局的统计，并且对每个人的驾驶路程、他/她购买的彩票张数以及有多少人卷入了常见的交通事故这些方面做出一些假设后，我们能发现，这个死亡率的合理估计值差不多就是 1 人/次抽奖。

各州政府都倾向于忽视有关彩票之潜在不良后果的争论。很大程度上，这是由于他们对数学期望了解充分，从而可以通过一定的安排，使得彩民们所购的每张彩票之期望回报——总奖金额除以售出彩票的张数——要小于彩票的价格。通常而言，这笔可观的差额就进入了州金库的保险箱中。但 1992 年，澳大利亚墨尔本的一些投资人发现，弗吉尼亚的彩票违反了上述原则[15]。该博彩从 1 到 44 中挑出 6 个数字来决定中奖者。如果能找到一个足够大的帕斯卡三角形，就可以发现，从 44 个数字中挑出 6 个来，不同的方式有 7059052 种。彩票的奖金池中有 2700 万美元，而如果把二等、三等和四等奖都算进来，总奖金额就上升为 27918561 美元。这些聪明的投资人进一步推得，如果他们对可能的 7059052 种组合的每一个都下一注，那彩票总价值应等于奖金额，即每张彩票应值 2790 万美元除以 7059052，或者说约 3.95 美元。而弗吉尼亚州的大脑定出的彩票价格是多少呢？照例是 1 美元。

这些澳大利亚投资人很快就在本国和新西兰、欧洲及美国找到了 2500 个平均愿出资 3000 美元的小投资者。如果计划成功，这笔

3000 美元的投资将带来 10800 美元的回报。计划还有一些风险。其一，由于他们不是唯一的彩票购买者，因此，可能另有一名甚至不止一名彩民也会买中号码，而这就意味着他们将不得不与人分享奖金。在已举办的 170 次抽奖中，有 120 次是无人获奖，一人获奖的不过是 40 次，而两人获奖的则仅有 10 次。如果这些数据反映了各种获奖情况的真实概率，那就表明投资者们有 120/170 的机会独享奖金，有 40/170 的机会得到一半奖金，而有 10/170 的机会只获得 1/3 的奖金。通过帕斯卡的数学期望原理来重新计算，可知期望奖金额为（120/170×2790 万美元）+（40/170×1395 万美元）+（10/170×697.5 万美元）=2340 万美元，对应于每张彩票 3.31 美元。相对于 1 美元的彩票价格，即使除去其他开支，也仍然是个巨大的回报。

还有另一个危险：要赶在最后时限前完成所有彩票的购买，这可是一场后勤噩梦。这将花费相当一部分投资本金，而且没有明显的奖金回报。

投资团成员们进行了周密准备。他们按规则要求手工填写了 140 万张彩票，每张彩票可抽奖 5 次。他们将一批批购买者安排到 125 个彩票零售点，并获得了许多杂货铺的合作，这些杂货铺从所售出的每张彩票中获得一定的利润。方案于最后时限前 72 小时启动。杂货铺的售货员们轮班工作，以便卖出尽可能多的彩票。有一个铺子在最后 48 小时中卖出了 75000 张彩票；一个连锁店则收到了足够购买 240 万张彩票的支票，将打印彩票的工作分配到各个加盟店，还雇用了快递来收集这些彩票。即使如此，投资团最终还是没能在规定时间内完成计划：他们仅仅只买了 7059052 张彩票中的 500 万张。

获奖彩票公布后的几天中都无人领奖。投资团中奖了，不过这几天工夫都在寻找中奖彩票。然后，当州彩票官员发现投资团的所作所为后，他们阻止了奖金的支付。接着就是一个月的司法争吵，直到这些官员终于明白，他们并没有合理原因来取消投资团的中奖。

最终，他们支付了这笔奖金。

帕斯卡为随机性研究贡献了计数方法以及数学期望概念这两个思想。尽管他与数学断绝了往来，但如果身体还撑得住，谁知道他还会发现些什么呢？但他最终没能撑住。1662 年 7 月，帕斯卡身患重病。医生们开出了常用的处方：给他放血，并使用烈性泻药、灌肠剂和催吐剂。有那么一阵子，他的状况有所好转，但之后，病症再次复发，更伴随了严重的头痛、晕眩与痉挛。帕斯卡发誓说，如果他能活下来，就将生命奉献于帮助穷人。他还要求把自己移送到一所收治不治患者的医院去，以便在死时有人做伴。几天后，在 1662 年 8 月，他真的去世了，时年 39 岁。尸体解剖发现他的死因是大脑出血，还有那些折磨他一生的病痛对肝脏、胃部和肠道所造成的损伤。

第五章

针锋相对的大数定律与小数定律

卡尔达诺、伽利略和帕斯卡都假定问题中的概率为已知。如伽利略就假设骰子6个面朝上的可能性彼此相等。但这种“知识”有多可靠呢？大公的骰子很可能是按各面平等的原则来设计的，但这并不意味着预期的平等当真就得到了实现。伽利略也许观察了若干次投掷的结果，并记下各面出现的频度，来对这个假设进行了一定的检验。但是，如果把这个检验重复上几次，那他很可能会发现，频率分布每次都会稍有出入。对于他需要解释的那个8%的微小差异，分布中的任何小小出入都可能对结论造成显著影响。随机性的早期研究要能真正应用于真实世界，就不得不面对下面的问题：隐藏的概率与观察结果之间，其联系究竟如何？从实用角度而言，说骰子扔出2点的可能性是1/6，这到底是什么意思？如果这句话并不意味着在任何一系列投掷中，每6次投掷就严格地扔出1次2点，那么对于扔出2点的几率为1/6的信念，到底扎根于何处？如果医生说一种药的有效率为70%，但还有1%的情况会出现严重的副作用，这是什么意思？或者一次民意调查发现某位候选人的选民支持率为36%，这又是什么意思？这些都是十分深刻的问题，并与随机性的概念——一个时至今日仍然为数学家们所乐于争论的概念——本身有关。

在最近的一个温暖的春日里，我就与一名自希伯来大学前来访问的统计学家进行了一次这样的讨论。这位统计学家——莫希（Moshe）　　当时在加州理工的午餐桌上与我对面而坐。在将一勺

勺脱脂酸奶送进嘴的间隙中，莫希表达了一个意见，即真正的随机数是不存在的。“根本就没有这种事情。”他说，“噢，他们出版了许多数表，写了许多计算机程序，但他们只不过是在自己骗自己而已。从来没有一个人找到过一个比扔骰子更好的产生随机性的方法，而扔骰子同样也做不到真正的随机。”

莫希朝我挥动着他的白色塑料勺子。他已经有些激动了。在他对随机性的感受以及他的宗教信仰之间，我觉察到了一种联系。莫希是一名犹太东正教徒，而我知道，许多信仰宗教的人，会有些难以想象上帝竟然能允许随机性的存在。“假设你想得到 N 个由 1 到 6 的数字所构成的一串随机数，”他告诉我，“那么就把骰子扔上 N 次，并记录下所得的点数。这是一个随机数串吗?”

不　　他宣称道　　因为没人能做出一个完美的骰子。总会有某些点数出现得比其他点数更频繁。也许要扔上 1000 次或 10 亿次，但你总归能注意到这样的偏好性。他说，任何人造出来的东西，都注定要受此缺陷的影响，因为人类无法达到完美。这也许是对的。不过大自然却可以达到完美，而真正的随机事件确实在原子层面上发生着。实际上，这正是量子理论的基础。因此，我们就把剩下的午餐时间都花在了对量子光学的讨论上。

今天，通过投掷大自然完美的量子骰子，最尖端的量子发生器能产生出真正的随机数。在过去，产生随机性所必需的完美性，的确是一个难以把握的目标。在这方面最具创造性的方法之一，来自于 1920 年前后纽约市的 Harlem 犯罪辛迪加[1]。由于每天都需要由 5 位数字构成的随机数来进行非法彩票活动，这些骗子们将鼻子伸进了权力机关，并利用美国财政收支的最后 5 位数字来产生所需的随机数。（在撰写本书时，美国政府已负债 8995800515946.50 美元，或者说平均每人负债 29679.02 美元。因此，这些骗子在今天可以通过人均负债来获得所需的 5 个数字!）他们那所谓的“财富”彩票不但违反了刑法，还违反了科学规律，因为根据一条被称为本福德定

律（Benford's law）的规则，按这种累积方式所产生的数字并不随机，而是更偏好于那些较小的数字。

本福德定律并不是由一个叫本福德的家伙发现的。它的发现者是美国天文学家西蒙·纽康姆（Simon Newcomb）。1881 年前后，纽康姆注意到一个现象：在对数表书中，用来处理 1 开头的数字的那几页，相比于那些处理 2 开头一直到 9 开头数字的书页，要来得更脏些，磨损得也更厉害，特别是 9 开头的数字所在的那些书页，看上去干净簇新。纽康姆假定在长期使用过程中，磨损度与使用次数成正比，从而得出结论说，那些与他使用同一本对数书的科学家们，他们所处理的数据满足如上的数字使用频数分布。福兰克·本福德（Frank Benford）于 1938 年在纽约州斯克内克塔迪（Schenectady）的通用电气研究实验室（General Electric Research Laboratory）中，也细细地查看了对数表，并发现了同样的事情，之后，这条规律就获得了它今天所冠的名字。不过两人都没能证明这条定律，这个证明直到 1995 年才由佐治亚理工学院（Georgia Institute of Technology）的数学家特德·希尔（Ted Hill）完成。

根据本福德定律，1 到 9 这 9 个数字的出现频率并不相等。相反地，数据最高位为 1 的情况约占 30%，为 2 的机会约占 18%，如此一直到数字 9，它出现在最高位的机会大概是 5%。一个类似但较少提及的定律是关于最低位数字的。许多类型的数据都服从本福德定律，特别是财务数字。实际上，这条定律简直就是为在大量财务数据中寻找欺诈行为而量身定做的。

本福德定律的著名应用之一，涉及一名叫做凯文·劳伦斯（Kevin Lawrence）的年轻企业家。他集资 9100 万美元以建立一个连锁高科技健康俱乐部[2]。大把钞票到手后，劳伦斯迅速行动起来。他雇用了一群行政人员，并很快花掉了投资人的钱，跟他募集到这些资金一样快。要不是因为一个细节，这本来倒也没什么，不过他和他那一大帮子人将大部分钱用在了购买私人物品而不是公务上。

几栋私人住宅，20 艘私人游艇，47 辆汽车（包括 5 辆悍马、4 辆法拉利、2 辆道奇 Viper、2 辆 DeTomaso Pantera 以及 1 辆兰博基尼 Diablo），2 块劳力士表，1 个 21 克拉的钻石手镯，1 把 20 万美元的日本武士刀，以及 1 台商用棉花糖机，这些东西很难被解释为业务上的必要开支。于是，劳伦斯和同伙们试图通过一个复杂的银行账户与皮包公司网络，来制造业务欣欣向荣的假象，从而掩饰他们贪污公款的痕迹。不幸的是，一位名叫达雷尔·多雷尔（Darrell Dorrell）的疑心很重的法务会计师，将他们的支票与转账单编成了 70000 个数据的一张表，然后将其中数字的分布情况与本福德定律进行了对比。这些数据没能通过定律的检验[3]。这当然只是调查的开始，但之后，传奇就不出意料地真相大白了。在 2003 年感恩节的前一天，穿着浅蓝囚服、被律师簇拥着的劳伦斯被判处 20 年徒刑，不得假释，从而为故事画上了句号。IRS 也对本福德定律进行了研究，以辨认出缴纳税款中的骗局。有研究者甚至将它用到了克林顿 13 年中的退税款上。它们通过了定律的检验[4]。

可想而知，Harlem 辛迪加和它的彩票顾客都没有注意到中奖号码中的这种规律性，但如果像纽康姆、本福德或希尔这样的人也来玩这个彩票，原则上他们就可以利用本福德定律，来下一些赢面更大的注，从而在他们的学者薪水之外，赚上一笔颇为不错的外快。

1947 年，兰德公司（Rand Corporation）的科学家们迫切需要一个大大的随机数表。当然，他们的目的高尚得多：通过一种被恰如其分地称为蒙特卡洛法的方法，利用这些随机数去帮助求取某些数学方程的近似解。为了产生这些数字，他们使用了电子噪声——一种电子轮盘赌。那么，电子噪声是随机的吗？这个问题就如随机性的定义本身一样微妙。

1896 年，美国哲学家查尔斯·山德斯·皮尔士（Charles Sanders Peirce）写道，随机采样是“根据某个规则或方法进行的采样，当这个规则或方法被重复无限次，那么就长期而言，它从某个物集

抽取出其中任何一个的次数，跟从其他任何一个具有相同元素个数的集合中抽取任何一个的次数，两者同样频繁。”[5] 这被称为随机性的频数解释。另一个主要解释叫做主观解释。频数解释根据采样结果进行判断，而主观解释则根据采样的产生方式来进行判断。根据主观解释，如果我们既不知道也无法预测产生某个或某些数字的过程将会给出什么样的结果，那这个或这些数字就被认为是随机的。

这两种解释之间的差别，比看上去的要来得大。打个比方，在一个完美的世界里，根据第一种定义，扔骰子是随机的，但根据第二种定义则不是，因为尽管扔出各个点数的机会都相等，但我们可以（在一个完美的世界中）根据所知的精确物理条件，在每次投出骰子前就确定将扔出的点数。但在我们不完美的真实世界中，扔骰子在第二种定义下是随机的，而在第一种定义下却非如此，因为正如莫希所指出的，由于骰子的不完美性，它的各个点数不会以相同的频率出现，但尽管如此，由于我们的能力有限，因此没有先验知识来知道某个点数会比另一个点数出现得更频繁。

为了搞清楚他们的随机数表是否真的随机，兰德公司的科学家们对它进行了许多不同的检验。更细致的检查显示，正如莫希那个存在着原则性不完美的骰子一样，这个随机数生成系统似乎是有偏的[6]。科学家们对系统进行了若干精巧的修正，但却从未能完全排除其中的规律性。如莫希所言，很讽刺地，彻底的混沌本身也是一种完美。不过，可以证明这些兰德随机数已具有足够的随机性，从而可以发挥作用了。兰德公司在 1955 年，以一个易记的名字出版了这些数表：《100 万个随机数》（*A Million Random Digits*）。

兰德科学家们在研究中所碰到问题，在差不多一个世纪前，就由英国人约瑟夫·贾格（Joseph Jagger）以某种抽象的方式所发现了，这个问题叫轮盘赌问题[7]。贾格是约克郡一个棉花厂的工程师和机修师，因此，他对于机械的能力及缺点有一种直觉式的感觉。1873 年的一天，他将这种直觉与创造性思维从棉花转向了金钱。他

所考虑的问题是，蒙特卡洛的轮盘赌到底有多完美呢？

轮盘赌——至少根据传说，是帕斯卡在瞎想着永动机时发明出来的——基本而言是一个很大的碗，其中被隔成了许多形如馅饼上切下的窄窄一片的扇形部分（称为槽）。当轮子转动时，一颗石弹子先在碗沿上弹跳，并最终落入以数字 1 到 36 再加一个 0（美国的轮盘上还有一个 00）所标记着的间隔中的某一个。轮盘赌的存在是一个极好的证据，来证明所谓可靠的通灵者根本就不存在，因为在蒙特卡洛轮盘赌中，如果对某个号码的间隔下注 1 美元，而弹子又恰好落到了这个间隔里，那么赌场就会付你 35 美元（再加上你下注的那 1 美元）。如果灵媒的确存在，我们就应该在赌场这种地方，看到他们一个个喧哗不已、手舞足蹈地推着一车又一车的钞票；但实际上，我们所看到的灵媒大都出现在网上，自称为“什么都知道而且什么都看得到的泽尔塔”，提供 24 小时的免费在线爱情建议，并与大概 120 万（据谷歌）的其他网络灵媒激烈竞争。对我而言，未来有如在浓雾中一般模糊不清，甚至过去的事情也日益如此。但我的确知道一件事：要去赌欧式轮盘赌的话，我输的机会将是 36/37，而赢的机会只有 1/37。这就意味着每下 1 美元注，赌场就会赢得（36/37×1 美元）一（1/37×35 美元），即 1/37 美元，大约 2.7 美分。我可以觉得这 2.7 美分是为了欣赏那小小石球在闪亮的大轮子上蹦来蹦去所付的费用，也可以觉得这是为了争取一次被闪电击中（当然是以一种好的方式）的机会所付的价钱，至于到底是哪种看法，就得根据当时的精神状态而定了。至少轮盘赌的机器就应该以这种方式来行事。

但果真如此吗？贾格认为，只有当轮盘处于完美的平衡时，机器才能满足这样的要求。而他与机器一起混得太久了，因此他持有与莫希同样的观点。他很愿意去赌一赌这些轮盘并不完美。因此，他带着全部积蓄来到了蒙特卡洛，雇了 6 名助手，每人负责赌场里 6 台轮盘赌中的 1 台。每天，他的助手都要观察所负责的轮子，记下

赌场开张的 12 个小时内每次轮盘赌中获胜的数字。晚上回到旅馆后，贾格就对这些数字加以分析。6 天下来，他在其中 5 台上没有发现任何不均匀的数字分布，但在第 6 台上，有 9 个数字很明显地比其他数字出现得更加频繁。因此，在第 7 天，他就开赴赌场，并在这 9 个出现得更频繁的数字——7、8、9、17、18、19、22、28 和 29　　上投以重注。

当晚赌场关门时，贾格赢了 7 万美元。这个胜利并非无人察觉。其他赌徒围到他的桌旁，扔下钞票来一同分享好运。赌场所有的巡查都紧盯着贾格，试图找出他好运的原因，或者更棒地，在他作弊的时候抓个现行。大赌 4 天之后，贾格已经堆起了 30 万美元，而赌场经理只能绝望地希求能摆脱这个神秘的家伙，或至少阻止其计划的继续进行。也许你会以为这个目的是靠某个布鲁克林的壮汉来达成的①，不过赌场职员的做法可要聪明得多。

第 5 天，贾格开始输钱了。这个失败如同前面的成功一样，并不是马上就能察觉得到的。在赌场搞小动作的之前或之后，贾格都一直有输有赢，只不过现在他输多赢少，而之前则输少赢多。要照赌场在每一注上小小的正常盈利，贾格可得加倍勤恳地在赌桌上工作好一阵，才能把钱输光，但在饕餮了 4 天赌场的金钱之后，他可不想因为一点风声鹤唳就收手不干了。到这次转运最终使他停手为止，贾格已经输掉了一半的家产。我们不难想象他的心情——更不用说那些跟随者的心情——是多么失落和苦涩。方案怎么会突然就行不通了呢？

最终，贾格狡猾地观察到，在他赢钱的那许多个钟头里，他曾留意到的轮子上的一条细微划痕，现在却不见了。赌场会不会如此好心地给划痕补上一道漆，以便他能漂漂亮亮地把赌场搞破产呢？贾格猜他们可不会有这份好心。因此，他仔细地检查了其他几台机

---

① 作者此处意指找人用暴力来强迫贾格止步。

器。其中一台是有划痕的。赌场经理没有猜错，贾格那大获成功的时光，肯定以某种方式，与他所赌的那台机器有关系，因此他们连夜调换了轮子。贾格换到了这台有划痕的机器，又开始赢钱了。不久，他赚到的钱就超过了之前所达到的高度，已经有将近 50 万美元了。

之后发生的事情对贾格而言很不幸。赌场经理最终集中全力来对付他的计划，并找到了一个新方法来阻止他。他们决定在每晚关门后，就把那些下注槽全部转离原先的位置，因此机器的不平衡性每一天都会偏爱于不同的数字　　那些贾格不知道的数字。贾格又开始输钱了，并最终离开了赌场。他就此终结了赌博生涯，带着 32.5 万美元离开了蒙特卡洛。以现值来算，这笔钱大约有 500 万美元。回到家乡后，他辞去了厂里的工作，并投资于房地产。

贾格的方法看似稳妥可靠，但实际却并非如此。即使对于具有完美平衡性的轮子而言，0、1、2、3 等点数也不会以绝对相等的频率出现，那些出现频率领先的数字不会礼貌地停下来，等待掉队者们赶上来。实际上，肯定会有某些数字出现得比平均水平更频繁，而另一些则达不到平均水平。所以，即使有了 6 天的观察结果，贾格仍有出错的可能。他所观察到的特定数字所表现出的更高出现频率，有可能也是随机的结果，而并不表示它们的出现概率确实更高。也就是说，贾格同样要面对本章开始时的问题：对于一系列未知概率，我们能期望对系统的观察结果与这些概率的吻合程度到底有多高呢？正如帕斯卡的工作是在（科学）革命的新氛围中得以完成，这个问题同样是在一次革命　　这次是数学革命　　期间才得到答案，这个革命就是微积分的发明。

1680 年，一颗巨大的彗星掠过了太阳系中我们近旁的空间。它离地球如此之近，使得它所反射的那微不足道的一部分阳光，都足以让它成为夜空中引人注目的主角。当这颗彗星处在地球轨道中我

们称为 11 月的那一段时，人们第一次发现了它。之后的几个月中，它都是人们热烈而详细调查的对象，其运行路线也得到了细节极其详尽的记录。1687 年，牛顿将用这些数据作为他那个重力的平方反比定律的例子。一个晴朗的夜晚，在瑞士那块被称为巴塞尔（Basel）的小小土地上，另一个注定成就伟大的人物，也同样注意着这颗彗星。这是名年轻的神学家。当他凝视着彗星那明亮而弥漫的光芒时，他意识到他希望为之奉献一生的，不是教会，而是数学[8]。伴随着这个意识，不仅萌发了雅可布·伯努利（Jokob Bernoulli）① 个人职业生涯的改变，也诞生了数学史上最伟大的家族树：从雅可布出生到 1800 年间的一个半世纪中，伯努利家族产生了许多后人，他们中约一半都颇具天分，包括 8 位知名的数学家，而其中 3 位［雅可布，他的弟弟约翰（Johann），以及约翰的儿子丹尼尔（Daniel）］更是被归入了历史上最伟大的数学家这一群体中。

当时，彗星被神学家和普通大众认为是神之愤怒的标志，而且上帝肯定是被气得发疯，才会放出像 1680 年的彗星这样一个信号——它占据了大半个肉眼可见的天空。一个牧师称它为“那来自全能而神圣之上帝、书写并呈现于无力而不洁的凡人之子眼前的天之警告”。他还写道，这彗星预示着国家或城镇的“精神或俗务中值得注意的变化”[9]。雅可布·伯努利却另有看法。1681 年，他出版了一本小册子，名为《一个新发现的方法，它如何能将彗星或说扫帚星的路线简化为若干基本定律，且如何能预测它的出现》。

在这颗彗星的问题上，伯努利比牛顿可早了 6 年。至少，如果他的理论是正确的话，那就的确是早于牛顿了。他的理论不正确，但却仍然是一个公开的宣言，宣称彗星所遵循的，是自然规律而非上帝的奇思怪想。这是一件很勇敢的事情，特别是在一年前——几

---

① 雅可布·伯努利（1654～1705）是瑞士数学家，主要贡献包括超越曲线及微分方程求解，以及概率论中的大数定理等。



乎是伽利略被宣判有罪之后 50 年　巴塞尔大学的数学教授彼得·梅杰伦（Peter Megerlin）因接受了哥白尼的科学体系而遭到了神学家们的全面攻击，并被禁止在这所大学中授课。一道无法跨越的鸿沟横亘于巴塞尔的数学家-科学家与神学家之间，而伯努利坚定地站在科学家的一方。

伯努利的天分很快就使他为数学家团体所拥抱。当梅杰伦于1686 年晚些时候去世后，伯努利继承了其数学教授的职位。伯努利当时正在研究机遇性博弈问题。对他影响最大者之一，是荷兰数学家与科学家克里斯蒂安·惠更斯（Christiaan Huygens）①。惠更斯改进了望远镜，第一个理解了土星光环，创造了第一台摆钟（基于伽利略的思路），并进一步发展了光的波动学说，除此之外，他还受到帕斯卡与费马观点的启发，写了一本关于概率的数学初级读本。

对伯努利而言，惠更斯的书确实带来了灵感。不过，他还是在惠更斯的理论中看到了严重的局限性。对于机遇性博弈而言，它也许够用了，但是对于更具主观性的生活中的方方面面又如何呢？怎样才能给法庭证言的可信度赋予一个确定的概率呢？或者在英格兰的查理一世（Charles Ⅰ of England）和苏格兰人的女王玛丽（Mary, Queen of Scots）两者中，分辨出哪个的高尔夫打得更好（这两人都是颇为热衷的高尔夫球手）？伯努利相信，理性决策要成为可能，就必须有一个可靠的数学方法来确定概率。他的观点反映了那个时代的文化，在这种文化中，通过与概率期望相一致的方式来管理事务，被认为是某人具有理性的标志。但在伯努利看来，限制着旧的随机理论的，不仅仅是主观性。他还认识到，这个理论的设计不能处理那些我们一无所知的局面，虽然此时各个结果的概率原则上可以被定义，但实际却是未知的。这就是我与莫希谈论的问题，也是贾格

---

① 惠更斯（1629～1695）是荷兰物理学家、天文学家和数学家，对于力学和光学有杰出贡献，是近代自然科学的重要开拓者。

需要解决的问题：一个不完美的骰子扔出 6 点的机会是多大？你感染上瘟疫的可能性有多大？你的胸甲能挡住对手长剑一刺的可能性又是多大？在主观性和不确定的局面中，指望惠更斯书中所预设的那种先验概率能现成得到，伯努利认为这简直是一种“精神错乱”[10]。

正如贾格后来所做的那样，伯努利认识到问题的答案就在于：相比于依赖那些塞到我们手中的概率，我们更应该通过观测来找出这些概率。作为一名数学家，他努力使这个想法精确化。假定我们观察了一定次数的轮盘赌，我们能多么接近地去确定隐含的概率，而对于所得概率的正确性，又能抱以多高的信任度呢？下一章将回到这些问题上来，但它们却不完全是伯努利所能解决的问题——他回答的是另一个与之紧密相关的问题：真实获得的结果能多好地反映那些隐含的概率？伯努利认为，当实验次数增加时，显然有理由期望所观察到的各种结果出现的频率，能越来越精确地反映出各自的真实概率。他肯定不是第一个相信这一点的人，但却是第一个对该问题进行规范处理，将想法转化为证明，并对其进行量化的人。他问的问题是：需要多少次实验才足够呢？我们对所得结果又能在多大程度上确定？同时，在解决这些问题时，他也认识到微积分这个新学科的重要性，并成为最初认识到这一点的那些人中的一员。

事后证明，伯努利在巴塞尔被提名为教授的那一年，是数学史上具有里程碑意义的一年：就是在这一年，戈特弗里德·莱布尼茨(Gottfried Leibniz)① 发表了其革命性的论文。作为 1684 年所发表的关于微分计算论文的补充，他在新论文中阐述了积分运算的原理。1687 年，牛顿在常被简称为《原理》的《自然哲学之数学原理》

① 莱布尼茨（1646～1716）是德国科学家、数学家、物理学家、历史学家和哲学家，其贡献与影响涉及范围极广，包括微积分、高等数学、计算机科学、物理学与哲学。

(*Philosophiae Naturalis Principia Mathematica*) 中，对相同的问题给出他自己的版本。而伯努利解决有关随机性问题的关键点，就在于这些进展之中。

到莱布尼茨和牛顿发表他们的论文时为止，他们都已经在这个问题上花费了多年的时间。但是，他们那几乎相同的论文发表时间，带来了谁最早提出新思路的争论。伟大的数学家卡尔·皮尔森(Karl Pearson)①（我们将在第八章中碰到他）认为，数学家“流传于后世的（声望），很大程度上并不在于他们做了些什么，而在于他们同时代的人认定他们做了些什么”[11]。牛顿和莱布尼茨大概会赞同这个观点。两人实在难分先后，而失败者的痛苦也尽人皆知。当时的争论并没有得出一个清楚的结论。德国人和瑞士人通过莱布尼茨的著作学习微积分，而英国人与许多法国人则使用牛顿的书。从现代立场来看，这两种微积分之间的差别十分微小，但长期以来，牛顿的贡献则常常得到强调，因为他似乎的确是更早一点想到这个新主意，而且在《原理》中建立现代物理学时，他运用了这个发明。正是现代物理学的建立，使《原理》成为了或许是有史以来所有科学书籍中的最伟大者。不过莱布尼茨却发展出了一套更好的表示方法，并为我们所更常使用。

这两位的著作都不好懂。除了是最伟大的科学书籍之外，牛顿的《原理》还被称为“有史以来最难以理解的书籍之一”[12]。按雅可布·伯努利的一位传记作家的话来说，莱布尼茨的著作“根本没人看得懂”，因为它不仅仅是不清晰，而且还充满了印刷错误。雅可布的弟弟约翰则称它为“谜题而非解答”[13]。实际上，两部著作都是如此难以理解，以至于有学者猜测，他们两人可能都是故意将文章弄得晦涩难懂，从而使业余爱好者们知难而退。但对雅可布·伯努利而言，这谜一般的风格倒不失为一种优点，因为它的确能将天

① 卡尔·皮尔森（1857～1936）是英国数学家，统计学的创始人。

分的良莠分个清楚，而他的智力则属于前者一类。因此，一旦他参透了莱布尼茨的思路，就拥有了一件在全世界也只有屈指可数的几个人所拥有的武器。有了它，伯努利就能轻而易举地解决足以令他人望而却步的难题。

微积分以及伯努利的工作中，最中心的概念就是序列、级数和极限。对数学家或对其他人而言，序列这个词的意思几乎相同：按顺序排列的若干元素（如点或数）。级数则不过是一个数的序列的加和。不那么严密地来说，如果某个序列的组成元素看起来似乎趋向于某处——比如一个特定的端点或数值——则此处就被称为该序列的极限。

尽管微积分对序列的理解达到了新高度，但与许多其他概念一样，古希腊人对之早已熟知。实际上，公元前 5 世纪的古希腊哲学家芝诺（Zeno），就用一个令人惊讶的序列，造出了一个至今仍然让哲学专业的大学生们——特别是在几杯啤酒下肚之后——争论不休的悖论。芝诺悖论大致是这样的：设想有一个学生想朝 1 米外的门口走去（1 米只不过是为了叙述方便，但下面的内容对于 1 英里或其他任何距离而言都是一样的）。在这名学生到达门口之前，她首先必须到达这段距离的中点；而为了到达这个中点，她又必须到达这一半距离的中点，也就是 1/4 距离点；如此等等，直至无穷。换句话说，为了达到目的地，她必须走过如下这个序列中的距离：1/2 米、1/4 米、1/8 米、1/16 米，等等。芝诺因此说，由于这个序列无穷无尽，所以她必须越过无穷多的有限长距离，而这将耗费无穷多的时间。芝诺由此得出结论：她哪儿也去不了。

许多个世纪中，从亚里士多德到康德（Immanuel Kant）① 的哲

① 康德（1724～1804）是德国哲学家和天文学家，德国古典哲学的创始人和古典美学的奠基人，代表著作有《纯粹理性批判》和《实践理性批判》等。

学家们对这个困境反复争辩。犬儒派哲学[1]的创始人提奥奇尼斯(Diogenes the Cynic)用了一个经验主义的方法来解决这个问题：他只是简单地走了几步，然后指出，物体实际上确实在运动。对于我们这些并非哲学专业学生的人而言，这个答案听起来挺不错。但这可打动不了芝诺。芝诺当然清楚他的逻辑证明与他所感觉到的运动之间的矛盾，只不过跟提奥奇尼斯不同，芝诺更相信逻辑。而且他的努力也并非徒劳无功。即使是提奥奇尼斯自己大概也不得不承认，他的回答使我们面对着一个令人困惑（而且最终被证明是十分深刻）的问题：如果我们的感觉是正确的，那么芝诺的逻辑错在何处呢？

让我们看看芝诺悖论中的这个距离序列：1/2 米，1/4 米，1/8 米，1/16 米，…（这些距离将越来越小）。这个序列有无穷多项，因此，我们不能通过直接相加来得到整个序列的总和。但我们注意到，尽管距离项的个数有无穷多，这些项的值却一个比一个小。在这个序列中，无穷无尽的项数和各项对应的距离无休无止地缩短，这两者会不会平衡在一个有限的总和上呢？序列、级数和极限等概念所解释的正是这一类问题。要看看是怎么做的，现在我们不再试图去计算，在芝诺那无穷多间隔之后，这名学生到底走了多远，而是一次只挑出一个间隔来考察。下面就是这个学生在最初几个间隔后所走过的距离：

第一个间隔后：1/2 米
第二个间隔后：1/2 米 + 1/4 米 = 3/4 米
第三个间隔后：1/2 米 + 1/4 米 + 1/8 米 = 7/8 米
第四个间隔后：1/2 米 + 1/4 米 + 1/8 米 + 1/16 米 = 15/16 米

1/2 米，3/4 米，7/8 米，15/16 米，…，这些数字中存在着一

① 古希腊的一个哲学流派，主张清心寡欲，鄙弃俗世的富贵荣华，倡导回归自然。

种模式：它们的分母是 2 的幂，而分子总比分母小 1。根据这个模式，我们可以猜出，在经过了头 10 个间隔后，这名学生将走过 1023/1024 米的距离；而在头 20 个间隔后走过的距离是 1024575/10248576 米；等等。从这个模式可以看出，间隔越多，所走过的距离也越长，在这一点上，芝诺是正确的；但他说总距离将趋向无穷，这可就错了。上面这些数字看似越来越接近于 1，或者用数学家的话来说，1 米就是这个距离序列的极限。于是一切都说得通了：尽管芝诺把这段旅途剪成了无穷多的间隔，但她最终仍然只需要越过 1 米的距离。

芝诺悖论与这段旅程需要的时间而非长度有关。如果这名学生被迫一步一步地走完芝诺剪出来的这些间隔，那的确会碰到时间问题（更不用说她还得克服困难去走出比毫米还小的碎步）！但如果她可以按常速行动，而不必停顿在芝诺想象中的那些检查站前的话——为什么要停下来呢？——那么穿越每一段芝诺间隔的时间，就与其距离成正比，既然总距离有限，总时间也有限，那么对我们所有人而言，很走运地，运动总算还是可能的。

虽然现代的极限概念直到芝诺甚至伯努利死后很久　　直至 19 世纪[14]——才出现，但正是这个概念，蕴涵了微积分的核心思想。雅可布·伯努利正是运用了这一核心思想，来处理概率与观察结果之间关系的问题。特别地，伯努利研究了任意大数量之重复实验的极限情况。将一枚（平衡的）硬币扔上 10 次，也许会有 7 次是正面，但如果扔上恒河沙数那么多次，则最可能得到的结果将非常接近于 50%。1940 年代，有一位名为约翰·克里奇（John Kerrich）的南非数学家，他决定以实际的实验来验证这一结论。他把一枚硬币扔了又扔，扔了肯定得有恒河沙数次——实际上，是 10000 次——然后记录下每次扔出的结果[15]。你大概会想，这位克里奇总该有些什么别的更要紧的事做吧？不过当时他是一名战囚，当德国人于 1940 年春入侵丹麦时，他很不走运地正在哥本哈根访问。根据

克里奇得到的数据，在头 100 次中，只有 44%的次数扔出了正面；但到了 10000 次的时候，得到的结果就很接近于对半开了：50.67%。那么，如何对这个现象加以量化呢？这个答案就是伯努利的成就。

用科学史学家和哲学家伊恩·哈金（Ian Hacking）的话来说，伯努利的工作“出现在公众的面前，闪亮地预示着我们如今所知的有关这个成就的一切：深奥的数学，不受限制的实际应用，生动的二元性，以及对哲学化的恒久欢迎。概率到此时才完全浮现”。用伯努利自己更谦虚的话来说，他的研究被证明具有“创新性，以及……高度的实用性”。他还写道，这个研究过程充满了“令人殚精竭虑的困难”[16]。他在这上面工作了整整 20 年。

雅克布·伯努利将他 20 年努力所达到的巅峰称为“黄金定理”，而它的那些仅存在若干细微技术性差别的不同现代版本则有许多名字：伯努利定理，大数定律，以及弱大数定律。正如已经说过的，使用大数定律这个词，是由于伯努利定理考虑的，是大量观察的结果如何反映了隐含的概率。不过我们将沿用伯努利的用词，称之为黄金定理，因为我们将以这条定理的最原始形式来进行讨论[17]。

尽管伯努利感兴趣的是现实世界中的应用，但他最乐于使用的某些例子，却与一件大多数家居中都没有的物品有关：一个装满了彩色鹅卵石的瓮。在一个例子中，他假设这个瓮里有 3000 颗白鹅卵石和 2000 颗黑鹅卵石，也就是 60：40 的白黑比例。我们蒙上眼睛，从这个瓮里“归还式地”取出一系列鹅卵石，“归还式”的意思是指，每次抽取下一枚鹅卵石之前，要把前次取出的石子放回瓮里，从而保证 3：2 的白黑比不会发生改变。这样一来，抽到白色石子的先验概率就是 5 分之 3，或者说 60%，而在本例中，伯努利最关心的问题就是：经过一系列抽取得到的白色石子数量，能以怎样的严格程度来符合这个 60%的比例？发生这个符合的概率又是多少？

这个瓮的问题是个好例子，因为用来描述从瓮里抽取鹅卵石的数学，同样也能用来描述任何一系列具有两种可能结果的试验，只要这些结果随机出现，且各次试验相互独立。这种试验如今被称为伯努利试验，一系列的伯努利实验则称为一个伯努利过程。当一个随机试验有两种可能结果时，其中之一常被标记为“成功”，而另一个则是“失败”。这样的标记并不严格符合它们的字面意义，有时候，它们根本与这两个词语的日常含义无关　　比如说，如果你看这本书时迫不及待地要读下去，那它就是个成功；而如果你只是在壁炉里的木柴都烧完之后，再用它来给你和心上人延续温暖，那它就是个失败。扔一枚硬币，投票给候选人 A 或 B，生男或生女，买或不买某件商品，病愈或未愈，甚至死亡或生存，这些都是伯努利试验的例子。如果问题可以通过“是”或“不是”来回答，那具有多个结果的行为也常常能用伯努利试验来描述，比如“骰子是不是扔出了 4 点?”或“北极还有没有冰?”，等等。因此，伯努利写的虽然是石子和瓮，但他所有的例子都能同样应用于许多其他类似的场合。

理解了这一点后，让我们再回到瓮上面来，其中装的鹅卵石有 60％是白色的。如果你从瓮里（归还式地）取出 100 颗鹅卵石，你也许会发现其中恰好有 60％是白色的。不过你同样有可能只抽到了 50 或 59 颗白色石子。那么，你取出的石子中，正好有 58％到 62％是白色鹅卵石的机会是多少？59％到 61％的是白色的机会又是多少？如果你不是取 100 颗，而是取 1000 颗或 100 万颗鹅卵石，那么对所得结果的信任又能增加多少？我们当然不能百分之百地确信所得结果，但是否能通过抽取足够多的鹅卵石，来获得 99.9999％的确定度，以保证取出的石子中，白色石子的比例在 59.9％到 60.1％之间？伯努利的黄金定理所解决的，就是诸如此类的问题。

要使用黄金定理，必须做出两个选择。首先，必须给定一个可容忍的误差范围。那一大串试验所得的结果，应该多大程度上接近

于真实的 60%的比例？我们必须给定一个接近范围，比如 60%加减 1%或 2%或 0.00001%。其次，必须给定一个可接受的不确定度。我们永远无法 100%地肯定试验会给出期望的结果，但可以确定，在 100 次试验中，有 99 次的结果会令人满意，或者在 1000 次中有 999 次是如此。

黄金定理指出，我们总能取出足够多的鹅卵石，以使我们能几乎确定所得的白色鹅卵石比例很接近于 60%，而不论这个几乎确定和接近的定义是何等地严苛；而且，在给定了几乎确定和接近的具体定义后，定理还给出了用来计算这个“足够”的试验次数的数学公式。

定律的第一部分是观念上的胜利，也是定理中唯一能幸存到各个现代版本中的那部分。对于第二部分，也就是伯努利的公式，重要的是理解下面这一点：黄金定理虽然给出了一个足以满足所要求的置信度与精度的试验次数，但却并不意味着无法用更少的试验来达到同样的目标。这对定理的第一部分并无影响，因为第一部分只是说这个特定的试验次数总是有限的。伯努利希望他的公式所给出的答案具有实际可行性，但不幸的是，在大多数实际应用中很难做到这一点。例如下面就是伯努利自己解出的一个算例，尽管我改变了行文的顺序：假设有 60%的巴塞尔选民都支持该市市长，那么，需要对多少选民进行民意测验，才能以 99.9%的概率保证，调查结果所得的市长支持率在 58%到 62%之间，即调查结果的精度为正负 2 个百分点？（为了与伯努利问题保持一致，假设这些被调查者都是随机选取的，而且是归还式选取的。换句话说，可能会对同一名被调查者进行不止 1 次的测验。）答案是 25550，而在伯努利的时代，这差不多也就是巴塞尔的总人口数。伯努利并没有无视这个不实用的结果，他同样也知道，熟练的赌徒们能通过远少于数千次的试验，直觉地猜出某个新赌赛的成功概率。

伯努利的数值估计远非最优的原因之一，在于他的证明使用了

许多近似。另一个原因在于他选择了 99.9%作为置信度标准——即他要求所得结果错误（与真实概率偏差超过 2 个百分点）的机会小于 1/1000。这个标准过于苛刻了。伯努利称该标准为符合道德的确定性，即在他看来，一个有理性的人，在进行有理性的决策时，应该使用这个确定度。今天，我们已抛弃了所谓符合道德的确定性这一提法，而更偏爱于另一个将在最后一章中遇到的所谓统计显著性标准，该标准意味着在 20 次中，只有不到 1 次会得到错误答案。由此也许可以看出，从那时到现在，时代发生了多大的变化！

利用当今的数学方法，统计学家们已证明，对上述的民意调查，只需抽查区区 370 名被调查者，就能得到具有统计显著性且精度为正负 5%的结果。如果调查 1000 人，那么结果落在真实答案（巴赛尔市长那 60%的支持度）的 2 个百分点的误差范围之内的可能性为 90%。尽管有种种局限，伯努利的黄金定理仍然是一个里程碑，因为它至少从原理上证明：足够大的样本能几乎肯定地反映出被采样群体的真实组成。

在真实生活中，我们常常不是通过数千次的试验来观察某人或某物的表现。因此，如果说伯努利的错误在于确定性标准过于严格，那么在实际生活中，我们则常常犯下相反的错误：我们会认为一个样本或一系列试验的结果代表了真实情况，而实际上，这些样本或试验却容量过小，以至于无法给出可靠的结论。举例来说，如果在伯努利那个年代，只对刚好 5 位巴塞尔居民进行调查的话，根据类似于在第四章中的计算，得到正确结果——即目标人群中的 60%或说 5 位中的 3 位支持市长——的可能性，仅仅只有 1/3。

仅仅 1/3？难道在对某个选民样本进行调查时，市长支持者在人群中的真实比例，不是最可能得到的结果吗？事实上，1/3 的确是最可能的结果：调查发现 0 个、1 个、2 个、4 个或 5 个支持者的机会都比发现 3 个支持者的机会小。但是，发现 3 个支持者的可能性并

不是很大：因为还存在着很多不能代表真实情况的结果，而它们发生的可能性之和，就比调查结果正确反映人群比例的可能性大 1 倍。所以，通过对 5 名选民进行的调查，每 3 次中就会有 2 次得到“错误”的支持率；事实上，每 10 次中大概就有 1 次，你将得到所有 5 名选民都喜欢或不喜欢市长的结果。如果将注意力放在容量为 5 的样本集上的话，市长真实的受欢迎度很可能被严重高估或低估。

一个小样本能真实反映隐含的概率，这个误解或说错误的直觉之存在是如此广泛，以至于卡尼曼和特沃斯基给它起了个名字：小数定律[18]。小数定律并不是一条真正的定律，而只是一个讽刺性的名字，用来描述对实际并不够多的数量使用大数定律的这种被误导的做法。

如果人们只把这个（不正确的）小数定律用在那些坛坛罐罐的事情上，倒也没什么大不了的影响。但如之前所说的，生活中的许多事情都是伯努利过程，因此直觉常常使我们对所观察到的事物做出错误解释。这也是为什么——像在第一章中谈过的——当人们观察了兰辛或坎通们那屈指可数的几个成功或不那么成功的年头时，就认为他们以往的表现能准确预示其将来的表现。

现在，我们把上面的思路用到第四章中简略提到过的一个例子中：两个公司之间或公司内两个职员之间的竞争。现在，考虑一下财富 500 强公司的 CEO 们。让我们假设，由于他们各自的知识与能力，每名 CEO 每年都有一定的成功概率（不管他/她的公司如何定义这个成功）。同时，为了简化问题，再假设这些 CEO 成功年度的出现概率，与白石子或市长支持者的比例一样，为 60%（真实值高一点或低一点，都不会影响论证的本质）。这是否就意味着，我们应当指望在某个给定的 5 年时间段内，这名 CEO 应该精确地有 3 个好年头？

并非如此。先前的分析表明，即使这些 CEO 们都有一个先验的、过得去的 60% 的成功率，在某个给定的 5 年期间，特定某名

CEO的表现能准确反映这一成功率的可能性，也仅仅只有1/3！套到财富500强的例子里，这就意味着在过去5年中，约有333名CEO表现出来的业绩并不能准确反映他们的真实能力。更进一步地，我们可以指望有大概1/10的CEO们纯粹因为运气也会连赢或连输5年，这又告诉我们什么呢？评价一个人更可靠的方法，应该是分析他们的能力，而不仅仅是看看记分牌。或如伯努利所言："我们不应该以成败论英雄。"[19]

我们需要一些特殊的品质才能摆脱小数定律的影响。任何人都可以躺在沙发里，指着成绩表的最下一行来大发评论；但要评定一个人的真实知识与能力，则需要信心、思考、良好的判断以及——呃——胆量。你不能在会议中就这么跳出来对着同事们大吼："不要炒了她，她只不过正好碰到了一个极其糟糕的伯努利序列的结果。"当然，如果你跳出来说那个志得意满、刚成为代理史上卖出最多丰田凯美瑞的家伙"不过是碰到了随机起伏而已"，这同样也不会为你赢得朋友。这种事情很少发生。执行官们的成功岁月总被归因于他们的聪明才智，并通过深刻地事后之明回过头来加以解释；而当他们不成功时，我们又常常假定，这些失败准确反映了他们的天分与能力在坛子中所占的真实比例。

另一个与大数定律有关的错误观念，就是仅根据事情最近是否发生过，而认为它应该更可能或更不可能发生。根据事件的近期发生情况，而觉得某个具有确定发生概率的事件，其发生的可能性增加或减少了，这叫做赌徒谬误。比方说，即使克里奇在最初100次中只扔出了44次正面，这枚硬币也不会变得对反面更偏心一些，以便其出现次数能赶上来！但赌徒谬误就是诸如"她的运气用光了"或"该轮到他了"之类想法的根本所在。这根本不会发生。不论真假，连续的好结果不会让你倒霉；而连续的坏结果，也很不幸地不代表好运正在为你而备好。不过，赌徒谬误仍有意无意地影响着的人，可比你想象的要多得多。人们总是期待好运气出现在坏运气之

后，或者担心坏运气就跟着好运气而到来。

记得几年前在一艘游轮上，我看到一个充满热望的矮胖子，一边冒汗，一边疯狂地把硬币一个接一个地塞进老虎机里，它能吞多快，他就塞多快。他的同伴看到我正注视着他们，便简单地解释到："他的好时辰快到了。"尽管我很想对他说：不，他的好时辰还没到，但终于还是走开了。刚走几步，一阵突如其来地闪光、不停的铃响、这一对朋友发出的绝非些许的叫声以及听来肯定有个好几分钟的硬币流泻出老虎机的声音，让我停下了脚步。我知道现代的老虎机都已计算机化，其输赢都由一个随机数发生器所驱动，而且根据法律或规章，这个随机数发生器必须产生名副其实的真正的随机数，来保证手柄的每一次掰动都完全独立于之前的掰动。但人们却仍然……唉，我们只能说，赌徒谬误是一个太强大的错觉。

伯努利给出黄金定理的手稿中断得十分突兀，尽管在开头部分，他承诺将给出定理在包括公民事务及经济问题等若干不同问题上的应用。统计学史学家斯蒂芬·施蒂格勒（Stephen Stigler）写道，似乎是"伯努利在看到 25550 这个数字后，就真真切切地放弃了"[20]。事实上，当 1705 年 8 月时年 50 岁的伯努利死于"慢性发热"时，他正准备将手稿出版。出版商们请求约翰·伯努利来完成手稿，但他以过于繁忙为由而拒绝了。这听起来可能很奇怪，不过伯努利家本来就是个奇怪的家族。如果要评选史上最不开心的数学家的话，选约翰·伯努利大概不会差得太多。在许多历史文献中，他被描述为嫉妒、虚荣、敏感、固执、脾气糟糕、自卖自夸、不诚实的登峰造极的骗子。在数学方面他颇有成就，但同样为人所熟知的，则是当他的儿子丹尼尔在一场他本人也参加的竞争中获得最终奖赏后，他就抄袭了丹尼尔的一本水力著作，篡改了出版日期以使他的书看来更早出版，然后借口丹尼尔试图剽窃他哥哥以及莱布尼茨的想法，而把丹尼尔踢出了法兰西科学院。

当约翰被请求去完成他过世兄长的手稿时，他刚从荷兰的格罗宁根大学（University of Groningen）来到巴塞尔不久，获得了一个希腊语教授而非数学教授的职位。雅可布本人当时也觉得这次职业变动颇为可疑，特别是他觉得约翰并不懂希腊语。在写给莱布尼茨的信中，雅可布怀疑约翰来到巴塞尔是为了篡夺数学教授的位子。的确，雅可布刚去世，约翰就获得了这个职位。

在两人成人后的大部分岁月中，约翰和雅可布相处得并不融洽。在数学著作和书信中，他们颇有规律地相互羞辱着对方，一位数学家对此写道："这些愤怒而激烈的语言，通常是为偷马贼所保留的。"[21]因此，编辑雅可布遗稿这个任务，就落到了食物链的更下一层——雅可布的侄子尼古拉（Nikolaus），雅可布的另一位同样名叫尼古拉的兄弟的儿子——身上。小尼古拉当时只有 18 岁，不过他也曾是雅可布的学生之一。不幸的是，他感到不能胜任这项任务，也可能部分是因为他清楚地知道，莱布尼茨并不赞同他伯父关于理论应用方面的想法。因此，手稿又沉睡了 8 年。最终，手稿以《猜度术》（*Ars conjectandi*）为名于 1713 年出版。和帕斯卡的《思想录》一样，该书至今仍被印刷出版。

雅可布·伯努利证明，通过数学分析，我们可以了解自然系统中隐含的概率，是如何在这些系统所产生的数据中得到体现的。至于伯努利没有回答的那个问题——如何从产生的数据中，推得隐含的事件概率——它的答案还要过几十年才会到来。

第六章

假阳性与好错误

1970 年代，哈佛大学某心理学教授的班上，来了一个外表奇特的中年学生。最初几堂课之后，他来到教授面前，解释他为什么来上这门课[1]。在我自己的从教经历中，虽然也碰到过一些有礼貌的学生跟我解释选我课的原因，但却从没有学生觉得必须要做出这个解释。所以我以为我可以很高兴地猜想他的理由大概是："因为我对这门课很着迷，而且你是一位好老师。"并抛开这个问题不理。不过他却另有原因：他需要帮助，因为在他身上发生了一些很奇怪的事——他妻子不等他开口，就会把他所想的事说出来，可现在却要跟他离婚；一个同事在喝酒时跟他提到了解聘的事，而两天后，他就丢了工作。他还声称，他长久以来经历了许多倒霉事以及他认为的令人着恼的巧合。

这些事情开始时让他困惑；然后，与大多数人一样，一个能使这些事情与他所坚信的世界之行事方式相吻合的思想模型逐渐成形。这个最终理论却与常人的想法大相径庭：他是一个精心策划的秘密科学试验中的被试对象。他相信这个试验正由一大帮阴谋家实施着，而领导者是著名心理学家斯金纳（B. F. Skinner）。他还相信，一旦试验完成，他就将名声大振，并可能被选举到一个高级公职上。他来上课的原因就是，他希望知道如何利用所积累的证据来检验这个假设。

课程结束后几个月，这名学生再次求助于教授。他说试验仍在进行中，他对前雇主提起了诉讼，而前雇主则找了一名精神病医生，

来证明他患上了妄想症。

医生所指称的妄想错觉之一，就是该学生所创造和宣扬的一名虚幻的18世纪的神甫。医生还特别嘲笑了学生所宣称的内容，即该神甫是名业余数学家，在空闲时间创建了一个十分怪异的概率理论。学生声称，神甫名叫托马斯·贝叶斯（Thomas Bayes）①。他还说，贝叶斯的理论所描述的，是当某事件发生时，如何估计另一个特定事件发生的可能性。特定的某个学生成为某个庞大秘密的试验心理学家们阴谋中的被试对象，这个可能性有多大？无可否认，这个可能性不大。但如果某人的妻子能在对方开口前就说出对方的心思，并且同事们能在酒桌旁的不经意交谈中预言他职业生涯的命运时，情况又如何呢？这名学生称，贝叶斯理论能说明如何根据新证据来调整概率的初始估计值。他还据此在法庭上出示了与其假设有关的一堆莫名其妙的公式和计算，得出的结论则是，这些附加证据意味着他的阴谋论假设，其成立的可能性为100万分之999999。而另一边的精神病医生则宣称，这位牧师数学家及其理论，都是学生精神分裂的想象所臆造的产物。

学生希望教授能帮他驳倒这个观点。教授同意了。他有足够理由去反驳这个观点，因为1701年出生于伦敦的贝叶斯，的确是滕布里奇韦尔斯（Tunbridge Wells）教区的一名神甫。他去世于1761年，安葬在伦敦一个名叫本希尔菲尔兹（Bunhill Fields）的公园中，与他那同为神甫的父亲约书亚（Joshua）葬在一起。而且，贝叶斯也的确创造了“条件概率”理论，展示了概率论如何由原先仅能处理独立事件，而扩展到处理那些存在相互关联的事件。例如，一个随机挑选的人患有精神疾病的概率，以及一个随机挑选的人相信他妻子能“读”到他心中所想的概率，两者都十分小；但是，在某人相信妻子会读心术的前提下，此人患有精神疾病的概率则要高得多；

---

① 贝叶斯（1702～1763）是英国数学家，主要贡献为概率论中的贝叶斯理论。

而且，在某人患有精神疾病的前提下，他相信妻子有读心术的概率，也同样高得多。所有这些概率是如何联系起来的？这就是条件概率所要研究的对象。

于是，教授提供了一份证言，解释了贝叶斯的存在及其理论，尽管他也并未支持前学生所声称的、可以证明其神志正常的可疑计算。这个故事中令人感到悲哀的，不仅仅是那名中年精神分裂者，也包括和他唱对台戏的医学与法律团队。经受精神分裂的痛苦是件很不幸的事情；但即算药物可以帮助对抗疾病本身，它们却不能对抗无知。我们即将看到，在许多医学诊断以及法律判决的严重错误中占据中心地位的，正是对贝叶斯思想的无知。而这个无知，却很少在医生或律师的职业训练中被提及。

我们同样在日常生活中进行着贝叶斯推理。有这么一部电影，说的是一名律师，有着很棒的工作，迷人的妻子，幸福的家庭。他爱妻子和女儿，却仍然觉得生活中似乎缺少些什么。一天晚上，当他乘火车回家时，无意瞥见一个美丽的女人，带着沉思的表情，从一个舞蹈班的窗口朝外凝望着。第二天晚上，他的目光追寻着这个女人；第三天晚上也是如此。他的火车每多经过一次舞蹈班，他在爱情魔咒中就更陷进一步。终于，在一个夜晚，他冲动地跳下车，报名参加了舞蹈班，希望能遇到这个女人。但当远处的凝视变成面对面的相对后，她那令人难忘的吸引力逐渐消失了。他确实恋爱了：不是和她，而是和舞蹈。

他将这不理智的情感当作秘密保守着，不让家人和同事们知道，并为他那越来越多不在家中度过的夜晚寻找种种借口。终于，妻子发现他并不像所宣称的那样经常加班到很晚。妻子想，他有了外遇而用谎言来掩饰下班后的行踪的可能性，要比他不是因为外遇而撒谎的可能性高得多，所以，她的结论是丈夫有了外遇。这个结论当然是错的，但她犯的错误却不仅仅是这个结论，而是整个推理过程本身：她搞混了当她丈夫有外遇时表现得鬼鬼祟祟的概率，以及当

她丈夫表现得鬼鬼祟祟时，他是在搞外遇的概率。

这是个常见的错误。设想老板这次回复你电子邮件的时间比平时要长。许多人会将此视为幸运星坠地的表现，因为如果幸运星坠落了，那么老板回邮件的速度比以前慢的可能性就高了。但也许老板仅仅是因为比平时更忙些，或是她母亲正好生病，回邮件才慢了一些。因此当她回邮件比较慢时，你的幸运星掉下来的可能性，比当你的幸运星正在往下掉时，她回邮速度变慢的可能性，前者要小得多。许多阴谋论都源于对这个逻辑关系的错误理解，也就是说，当一系列事件是巨大阴谋的产物时，这些事件发生的可能性，与当一系列事件已经发生时，这些事件证明存在着巨大阴谋的可能性，两者被混为一谈了。

贝叶斯理论所讨论的全部内容，就是当其他事件已经发生，或说给定其他事件发生的前提下，对于某事件发生的可能性所造成的影响。为了看看这个影响究竟为何，我们转到与第三章中的两个女儿问题相关的另一个问题上来。设想一个远房表亲有两个孩子。回想一下，在两个女儿问题中，我们知道这两个孩子中有一个或两个是女孩，而我们要搞清的，则是到底有几个女孩，一个还是两个？如果一个家庭中有两个孩子，那么如果其中至少有一个是女孩的话，两个都是女孩的可能性为多大？第三章并没有用如上的语言来讨论这个问题，但“如果”两字，将问题变为了一个条件概率问题。如果没有这个“如果”，两个孩子都是女孩的可能性是 1/4，对应的 4 种可能的出生顺序为（男孩，男孩）、（男孩，女孩）、（女孩，男孩）和（女孩，女孩）。但知道至少有一个是女孩的这个额外信息时，两个都是女孩的可能性就变为了 1/3，这是因为如果至少一个孩子是女孩的话，那么两个孩子的性别就只有 3 种可能情况——（男孩，女孩）、（女孩，男孩）和（女孩，女孩），其中 1 种正好对应了两个孩子都是女孩的结果。这也许就是理解贝叶斯思想的最简单方法——也就是个算账的问题嘛。首先，把样本空间——也就是所有可能情况的

清单——写下来，如果这些情况的可能性不等，就将其各自的概率也一同记下（在分析任何容易犯糊涂的概率问题时，这的确是个好办法）。接着，把被条件（在现在的问题中，就是“至少有一个女孩”这个条件）所否定掉的那些可能性划掉，剩下的就是条件满足时的可能情况，以及它们的相对概率。

这个方法看来似乎理所当然。自信爆棚的你也许会觉得，不用亲爱的贝叶斯神父来帮忙，自己也能想通这一点，于是发誓在下次进浴缸泡澡时，一定要抓上本别的什么书而不是现在的这一本。因此，在进一步讨论前，我们来看看两个女儿问题的一个略显不同的变体，其答案大概会让人更多些震惊。[2]

这个变体如下：某家庭有两个孩子，如果两个孩子之一是名叫佛罗里达的女孩，那么两个孩子都是女孩的概率有多大？是的，我说的是一个名叫佛罗里达的女孩。这个名字看起来好像是随便取的，但实际却并非如此。虽然这是以古巴移民、橙子以及那些为享受棕榈树和有组织的 bingo 赌戏而卖掉大房子的北方老人而出名的一个州的名字，但它同样也是一个真实的人名。实际上，佛罗里达是 20 世纪最初 30 年左右的时间中，1000 个最常见的美国女性名字之一。我很仔细地挑出这个名字，因为现在的谜题，部分就在于佛罗里达这个名字的什么东西——如果真有某个东西的话——影响了我们感兴趣的概率？不过我已经把自己的后话都讲了。在继续之前，请思索一下这个问题：在这个“名叫佛罗里达的女孩”问题中，两个都是女孩的可能性，是否仍为 1/3（就如两个女儿问题的答案一样）？

我很快就会证明，这个答案是“否”。有一个女孩名叫佛罗里达的事实，将我们所寻找的概率变成了 1/2：如果你想不通，也不必忧虑。理解随机性乃至所有数学问题的关键，并不在于能仅靠直觉就立刻知道问题的答案，而在于掌握能求得答案的工具。

质疑贝叶斯之存在的人们，在一件事情上是对的：贝叶斯从未

发表过哪怕一篇论文。对于他的生平，我们几乎一无所知，但他多半是出于自己的快乐才进行研究的，因此也不觉有多大必要去与他人交流研究成果。从这个角度以及其他角度来说，他和雅可布·伯努利正好相反：伯努利反对神学研究，而贝叶斯却由衷信奉；伯努利追求名声，而贝叶斯对之毫无兴趣；最后，伯努利定理关心的是，假如你打算用一枚平衡的硬币进行多次投掷，那么可以期望得到的正面朝上的次数是多少，而贝叶斯研究的却是伯努利最初的目标，即当你得到了正面朝上的次数时，你能在多大程度上信任硬币的平衡性。

使贝叶斯留名至今的理论，在 1763 年 12 月 23 日而最初为人所知。当时，另一位牧师兼数学家理查德·普莱斯（Rechard Price)① 在皇家学会（the Royal Society）——英国的国家科学院——宣读了一篇论文。这篇由贝叶斯撰写的论文名为《通往机遇学说的一个问题之解决的短文》，并于 1764 年发表在皇家学院的《哲学汇刊》（*Philosophical Transactions*）上。贝叶斯在其遗嘱中将该文连同 100 英镑一起留给了普莱斯。贝叶斯在遗嘱中用“我估摸着是纽因顿格林的一名传教士”来称呼普莱斯，而在写完遗嘱的 4 个月后，贝叶斯就离开了人世[3]。

尽管贝叶斯只是漫不经心地提到了普莱斯，后者却并不是一个阴暗角落中的传教士。他是宗教自由的著名倡导者，本杰明·富兰克林（Benjamin Franklin)② 的朋友，一个深受亚当·斯密（Adam Smith)③ 信赖而对其《国富论》（*The Wealth of Nations*）草稿中某些部分提出批评的人，以及一位有名的数学家。此外，他还被公

---

① 理查德·普莱斯（1723～1791）是英国的道德哲学家与政治哲学家，同时还在统计与金融方面做出过贡献。

② 富兰克林（1706～1790）是美国科学家、发明家、政治家、外交家、哲学家和文学家，美国独立战争的领袖，参与起草了《独立宣言》和美国宪法。

③ 亚当·斯密（1723～1790）是古典经济学的主要创立者。

认为保险精算学的创建者。1765 年，一个名为“公平社会”的保险公司派了 3 个人，向他寻求帮助。他因此建立了这个学科领域。在这次见面的 6 年后，他将研究成果发表于《评继承支付》（*Observations on Reversionary Payments*）一书中。尽管直至 19 世纪该书都还是精算师们的圣经，但由于某些糟糕的数据和估计方法，他似乎低估了人们的期望寿命。由此造成的大为膨胀的人寿保险费让他那些公平社会公司的伙伴们发了财；另一方面，英国政府根据普莱斯的表格制订了养老金支付标准，而当这些领取养老金的人们没有按预计的速度自然退出历史舞台时，倒霉的政府就大大地亏了一笔。

如我所述，贝叶斯是在试图回答那个激发了伯努利的相同问题时，发展出条件概率的：我们如何才能根据观察来推出其后隐含的概率？如果一种药物在临床试验中仅仅治愈了 60 名病人中的 45 名，那么对于药物在下一个病人身上有效的可能性，这个数据所能告诉我们些什么？如果药物是对 100 万名病人中的 60 万名有效，那么其有效率落在 60%附近的可能性显然很高。但当结果是来自规模更小的试验时，从中能得出什么结论呢？贝叶斯同时还问了另一个问题：如果在试验之前有理由相信该药物的有效率只有 50%，那么今后评估有效率时，试验所得的新数据应占有多大的权重呢？我们大多数的生活经验都类似于此：观察一个相对较小的结果样本集，由此获得若干信息，然后对产生出这些结果的内在性质加以判断。但这一系列推理应当如何进行呢？

贝叶斯是通过一个类比来考虑的[4]。设想有人提供了一张方桌和两个球。我们将第一个球滚落在桌子上，并且使它以等可能性停止在桌上的任意一点。我们现在要做的，是在看不到桌子和球的情况下，确定这个球停止在由左至右的坐标轴上的哪一点，而使用的工具就是那第二个球。我们以相同的方式反复将第二个球在桌上滚动，同时，一名合作者将观察第二个球的停止位置，是在第一个球的左边还是右边，并把这个结果记录下来。试验结束后，合作者将

会把第二个球落在两个大范围内的次数分别告诉我们。第一个球即对应我们希望获得的未知量，第二个球代表的则是我们能够得到的证据。如果第二个球都停在了第一个球的右边，我们就可以很有把握地认为，第一个球停在了离桌子最左边很近的地方；如果第二个球并非如此一贯地停在这个范围内，那么我们对于上述结论的信心也相应降低，或者去猜想第一个球实际落在了比我们假设的位置更靠右的地方。贝叶斯所表明的，就是如何根据第二个球的数据，来确定第一个球落在左-右坐标轴上任意给定位置处的准确概率，以及在给定附加数据的情况下，如何去修正概率的初始估计。按贝叶斯的术语，初始估计称为先验概率，而修正后的新猜测称为后验概率。

贝叶斯之所以设计出这个游戏，是因为它可以对应生活中的许多决策。在药物试验的例子中，第一个球的位置代表了药物的真实有效率，而第二个球所提供的，则相当于在病人身上得到的试验数据。第一个球的位置同样还能代表一部影片的受欢迎与否、产品质量、驾驶技术、勤奋与否、固执与否、天分高低、能力高低或任何能决定事情成败的东西。而第二个球所提供的东西，就代表了我们的观察或收集得到的数据。贝叶斯理论所说明的，就是如何估计未知概率，然后根据新的数据来对其加以修正。

如今，贝叶斯分析已广泛应用于科学和产业的各个领域。例如，用来确定汽车保险费率的模型中，就包含了一个描述在单位驾驶时间内，个人遭遇 0 次、1 次或更多次交通事故的概率。针对我们的目的，可以设想一个简化的模型，其中每个人都归入两种类型中的一种：平均每年都至少遭遇 1 次事故的高危司机类型，和平均每年遭遇不到 1 次事故的低危司机类型。假设你申请保险时的驾驶记录表明，过去 20 年中你一次事故也没有发生，或是在这 20 年中，你碰到了 37 次事故，那么保险公司将很清楚应该把你归入哪一种类型。但如果你是一名新手，那应该把你放在低危类型（遵守限速规定且志愿成为模范驾驶员的好孩子）呢，还是高危类型（一边痛饮那只

剩半瓶的 2 美元一瓶的布恩农场牌苹果酒、一边在梅因街飙车的坏孩子)？由于公司并没有你的个人数据　或说对“第一个球的位置”完全一无所知　它也许会以相等的先验概率将你归入任何一类之中，或者根据新司机的一般性数据，来估计你属于高危类型的初始可能性为（比方说）1/3。后一种情况下，公司可能会将你作为一个杂合体来建模——即 1/3 的高危加 2/3 的低危——并按 1/3 的高危保险费加 2/3 的低危保险费来收钱。经过一年的观察——也就是贝叶斯那第二个球被滚动一次后——公司就会使用新的数据来重新评价模型，调整之前所设的 1/3 和 2/3 的比例，再重新计算出应该收取的保费。如果这一年中你没有发生事故，那么低危的比例以及对应的低费用的比例就会相应增加；而如果发生了两次事故的话，那么这些比例就会相应降低。精确的调整幅度由贝叶斯定理给出。按同样的方式，保险公司今后就能周期性地调整评估值，以便反映出你到底是无事故类型的驾驶员，还是那个一连两次都因为在单行道上开错方向、左手拿着手机、右手还拿着甜面包圈而造成事故的家伙。这也就是保险公司给“好司机”打折的原因：无交通事故实际上提升了一名司机属于低危类型的后验概率。

显然，贝叶斯定理的许多细节都十分复杂。但正如我在分析两个女儿问题时所指出的，关键在于利用新的信息去“裁剪”样本空间，并相应调整各个概率。在两个女儿问题中，初始样本空间为（男孩，男孩）、（男孩，女孩）、（女孩，男孩）和（女孩，女孩）；但如果我们得知至少有一个是女孩，样本空间就缩减为（男孩，女孩）、（女孩，男孩）和（女孩，女孩），从而使两个都为女孩的可能性变为了 1/3。让我们使用这个简单策略，来看看当有一个孩子是名为佛罗里达的女孩时，会发生些什么。

在名叫佛罗里达的女孩问题中，新信息不仅与孩子的性别有关，而且还与女孩的名字有关。既然初始样本空间是一张所有可能情况的清单，那在现在的例子中，它就是一张包括了性别与姓名的清单。

我们以女孩 F 来表示“名叫佛罗里达的女孩”，以女孩 NF 来表示“名字不是佛罗里达的女孩”。于是样本空间写下来就是这么个样子：(男孩，男孩)、(男孩，女孩 F)、(男孩，女孩 NF)、(女孩 F，男孩)、(女孩 NF，男孩)、(女孩 NF，女孩 F)、(女孩 F，女孩 NF)、(女孩 NF，女孩 NF) 和 (女孩 F，女孩 F)。

现在我们来裁剪样本空间。既然已知两个孩子之一是名为佛罗里达的女孩，故样本空间可以缩小为 (男孩，女孩 F)、(女孩 F，男孩)、(女孩 NF，女孩 F)、(女孩 F，女孩 NF) 和 (女孩 F，女孩 F)。至此，现在的问题就与原先的两个女儿问题有所不同了。由于一个女孩叫或不叫佛罗里达的概率并不相等，因此，这个样本空间中的元素并非是等概率事件。

在 1935 年，即社会福利局 (Social Security Administration) 提供姓名统计数据的最后一年，大约每 30000 名女孩中，就有 1 人叫做佛罗里达[5]。由于这个名字现在已慢慢消失，为了叙述方便，让我们假定如今一个女孩名为佛罗里达的概率是 100 万分之 1。这也就意味着，如果我们知道某个女孩的名字不是佛罗里达，那并不稀奇；但如果我们知道某个女孩的名字确实是佛罗里达，那简直就等于买彩票中了大奖。即使不考虑父母通常都不会给孩子们起重名的这个事实，两个女孩都叫佛罗里达的可能性也仍然微乎其微，因此我们完全可以忽略掉两个女孩都叫佛罗里达的可能性，而剩下的样本空间中，就只剩下 (男孩，女孩 F)、(女孩 F，男孩)、(女孩 NF，女孩 F) 和 (女孩 F，女孩 NF)，而它们十分近似于等概率事件。

既然在 4 个元素中有 2 个——或者说有一半——对应了这一家子有两个女儿的情况，因此问题的答案就不再如两个女儿问题那样是 1/3，而变成了 1/2。这个增加的信息——关于女孩姓名的知识——造成了答案的不同。

要是仍然有些人想不通，那可以用另一种方法来理解：设想 7500 万个有两个孩子且至少有一个是女孩的家庭，聚集在一个非常

巨大的房间中。正如两个女儿问题告诉我们的那样，在房间里，大概有 2500 万个家庭中，两个孩子都是女孩，而其他的 5000 万个家庭则是一男一女（这其中又有 2500 万个家庭，其中女孩是姐姐，而剩下的相等数量的家庭中，女孩是妹妹）。接下来，就是我们的裁剪了：我们只留下那些有一个名叫佛罗里达的女儿的家庭。既然起名为佛罗里达的概率是 100 万分之 1，那么那 5000 万个有一个女孩的家庭中的大概 50 个留了下来；对于那 2500 万个两个女孩的家庭，也有大约 50 个家庭会留下来，其中 25 个的长女名叫佛罗里达，而另外 25 个则是次女起了这个名字。这就好比把女孩们当作彩票，而名叫佛罗里达的女孩则是中奖的彩票。尽管有一个女孩的家庭比有两个女孩的家庭多了一倍，但这些两个女孩的家庭，每个都拥有两张彩票，因此在最终的中奖者中，一个女孩的家庭跟两个女孩的家庭，两者数量应大致相等。

至此，对于名叫佛罗里达的女孩这个问题，我通过可能会让人不快的细节进行了描述，这种细节有时就让我上了邻居派对的黑名单。我之所以这么做，并非想让你们也碰到同样的遭遇，而是因为这个例子中的内在关系十分简单，而同样的推理可以使许多生活中的真实局面变得更加清晰明了。下面，我们就来说几个这样的实际情况。

我与贝叶斯神父最值得记住的遭遇，发生在 1989 年某个周五的下午。当时，医生打电话告诉我，我将有 999/1000 的可能性在 10 年内死去。他补充道：“我真的十分抱歉。”好像他对某些病人说抱歉时并非真心诚意似的。在回答了几个病情发展的问题后，他挂断了电话，很可能是去向另一位病人报告他或她周五下午的号外。很难描述甚至是回忆我到底是如何度过那个周末的，不过简单来说，那个周末我可没有去迪斯尼乐园。可为什么有了这个死亡判决，我却还好好地在这里，并写书来谈论它呢？

这次历险始于妻子和我申请人寿保险之时。申请中需要进行血检。一两周后，我们的申请被拒绝了。这个向来节约的保险公司，通过两封内容几乎相同的短信告诉了我们这个消息，而给我妻子的信相较于给我的信，不过就多了一个词。给我的信中说道，公司由于“你的血检结果”而否决了我的保险申请，而给我妻子的信中则说是由于“你丈夫的血检结果”而拒绝了她的申请。这多出来的“丈夫”一词，显然是好心的保险公司所乐意提供的、关于我俩投保被拒之原因的线索，因此，我带着不祥的预感去看了医生，进行了一次 HIV 检查。检查结果呈阳性。尽管开始时，我因为过于震惊而没有询问这个几率是否真的可靠，但后来我了解到，医生是如下得到那个 1/1000 的健康概率的：在 1000 个没有艾滋病毒的血液样本中，有 1 个的 HIV 检查结果会呈阳性。这听起来似乎跟他说的是一回事，但事实并非如此。医生将“如果我没有感染 HIV 而检查结果呈阳性”的概率和“如果我的检查结果呈阳性而我并没有感染 HIV”的概率搞混了。

为了明白医生的错误，我们来运用一下贝叶斯方法。首先是定义样本空间。我们可以把所有曾接受过 HIV 检查的人都包括进来，但如果多用一点与我有关的附加信息来定义样本空间，那结果会更准确一些：我们只考虑所有曾接受过 HIV 检查的、异性恋的、不滥用静脉注射吸毒的美国白人男性（我们将会看到，这个附加信息将给出怎样不同的结果）。

现在已经知道了哪些人应该被归到样本空间中。接下来，我们将这些人加以分类。与本问题有关的类型不再是男孩或女孩，而是检查呈阳性且确实为 HIV 感染者的人（真阳性）、检查呈阳性但没有感染 HIV 的人（假阳性）、检查呈阴性且没有感染 HIV 的人（真阴性）以及检查呈阴性但实际感染了 HIV 的人（假阴性）。

最后，让我们问一问，每种类型的人数有多少呢？考虑一个 10000 人的初始人群。根据疾病控制与防治中心的统计数据可以估

计，在 1989 年，大约每 10000 名接受检查的异性恋不滥用毒品静脉注射的美国白人男性中，就有 1 名感染了 HIV[6]。假设假阴性率十分接近 0，这就意味着，大约每 10000 人中有 1 人会因为真实感染而被检测到阳性。此外，由于假阳性率为——如医生所说的——1000 分之 1，因此，大概另有 10 人虽然没有感染 HIV，但还是被查出为阳性。而样本空间的 10000 人中剩下的 9989 人，检查结果将为阴性。

现在裁剪样本空间，使之仅包含那些检查结果为阳性的人。这样得到的，就是 10 个假阳性和 1 个真阳性，换句话说，在 11 个被查出为阳性的人中，仅仅只有 1 个真地感染了 HIV。医生告诉我检查结果出错——即我实际上十分健康——的可能性是 1/1000，而他真正应该说的是："不用担心，你没有感染的机会不小于 10/11。"就我的情况而言，这个筛检明显是被我血液中的某种标记物所糊弄了，虽然筛查所要发现的病毒其实并不存在。

在评价任何诊断检查的结果时，知道假阳性率是十分重要的。举例来说，某个检查能查出 99%的恶性肿瘤。这个诊断率听起来令人印象深刻，但我却可以很简单地设计出一个能查出所有 100%肿瘤的检查方法，而这不过是将每个被检查者都归为患有肿瘤的一类。我的方法和一个真正有用的方法之间的差别，关键就在于我的检查假阳性率很高。但上面的遭遇表明，已知假阳性率，并不足以确定检查的有用性，还应该将假阳性率与疾病的真实流行情况加以比较才行。如果是十分罕见的疾病，那么即使是低的假阳性率，也不意味着阳性结果就代表了确实患病；而如果是常见疾病，阳性结果就更可能具有意义。现在假设我是同性恋，并被查出阳性，再来看看疾病的真实流行程度对阳性结果诊断意义的影响。假定在 1989 年的男同性恋人群中，那些接受检查的人被感染的可能性为 1%左右。这就意味着在 10000 个检查结果中，我们得到的不是之前的 1 个、而是 100 个真阳性，连同 10 个假阳性。在这种情况下，这个阳性结果

将以 11 分之 10 的比率，来表明我的确被感染了。确定是否来自于高危人群之所以十分有助于评价检查结果，其原因就在于此。

贝叶斯理论表明，B 发生时 A 也将发生的概率，一般不同于 A 发生时 B 也将发生的概率[7]。医生这个职业中的一个常见错误，就是未能清楚认识到这一点。比如一些在德国和美国进行的研究中，已知有 7%的乳腺 X 射线检查在实际没有肿瘤时仍给出阳性结果，在此前提下，研究者们请医师估计一下，一名无症状的、40 至 50 岁且乳腺 X 射线检查结果为阳性的妇女，她真正患乳腺癌的可能性有多大[8]。此外，这些医生们还被告知，实际的乳腺癌发病率约为 0.8%，而真实的假阳性率约为 10%。把所有这些数据凑拢来，利用贝叶斯方法就能得出：真正因癌症而导致的乳腺 X 线检查阳性结果的比例，仅仅只有 9%左右。但在德国医生组中，有 1/3 的医生认为这个概率为 90%，而所有估计值的中值为 70%。在美国医生组中，100 名医生中就有 95 名估计这个概率应该在 75%上下。

类似问题还出现在运动员兴奋剂检测中。我们再次看到，被反复提及但实际却并不与问题直接相关的数据，就是假阳性率。它扭曲了我们对运动员真正犯禁之可能性的看法。例如世界级的跑步运动员及 1983 年 1500 米和 3000 米世界冠军获得者史兰尼（Mary Decker Slaney），当她于 1996 年在亚特兰大参加美国奥运会选拔赛以求复出赛场时，却被指控违反了禁止使用等同于睾丸激素类药物的禁令。经慎重商讨，IAAF［2001 年起国际田联（International Association of Athletics Federations）的官方名称］确认史兰尼“违反了药物禁令”，从而事实上结束了她的职业生涯。根据史兰尼一案中的若干证言，她接受的尿检，其假阳性率可高达 1%。这个数字大概能让许多人心安，因为她的确违规的可能性似乎为 99%。但我们已经看到，这并不正确。假设有 1000 名运动员接受了检查，而每 10 名运动员中就有 1 名的确服用了禁药。检查能以 50%的概率发现真

正服禁药的运动员。那么，每一千名受检的运动员中，有 100 名的确违规，而其中 50 名将会被尿检查出。同时，在那无辜的 900 名运动员中，也有差不多 9 个人的结果呈阳性。因此，尿检阳性并不意味着史兰尼的确有罪的可能性为 99%，而是 50/59 = 84.7%。换个方式来说，对于史兰尼确实违规的确信度，应该跟她扔骰子没扔出 1 点的可能性相近，而这样的确信度，显然为合理质疑留下了足够的空间。更重要的是，这个例子表明，进行大规模的检查（每年有 90000 名运动员接受尿检）并按前述过程来作出判决，会使许多无辜者蒙受冤屈[9]。

在法律圈中，有时将与上述相反的错误称为“检控官谬误”，因为检察官们常常使用这种靠不住的论断，来诱导陪审团根据单薄的证据对疑犯定罪。考虑一下发生在英国的莎莉·克拉克案（Sally Clark）[10]。克拉克的第一个孩子于出生后 11 周死亡。据称其死亡原因是婴儿猝死综合征（Sudden Infant Death Syndrome，SIDS）——当婴儿意外死亡而尸检又无法查出致死原因时所做出的诊断。克拉克再次怀孕，而这次她的孩子在第 8 周时死亡，死因再次被认为是 SIDS。事情发生后，克拉克被捕并被指控将两个孩子窒息致死。在法庭上，检方传唤了资历颇深的儿科医生罗伊·梅道（Roy Meadow）爵士来证明，由于 SIDS 的罕见性，两名婴儿都死于 SIDS 的可能性只有 7300 万分之 1。除此之外，检方没有拿出任何不利于被告的实质证据。那么，这个几率是否足以对克拉克定罪呢？陪审团觉得够了。因此，1999 年 11 月，克拉克太太入狱服刑。

梅道爵士估计的是一名婴儿死于 SIDS 的机会：1/8543。既然现在有两个孩子，那么他们都死于 SIDS 的几率，就是把上面这个数字跟自己再乘一次，从而得到了 7300 万分之 1。不过这个计算隐含的假设，是两起死亡事件相互独立——也就是说，即使在哥哥或姐姐死于 SIDS 的情况下，也没有任何环境或遗传因素会增加第二个孩子的死亡风险。但实际上，在审判结束数周后出版的《英国医学杂志》

（*British Medical Journal*）之编者按中，给出两名兄弟姐妹都死于SIDS几率的估计值为275万分之1[11]。这个机会当然还是很小。

要理解克拉克之所以是蒙冤人狱的关键，仍在于考虑相反的错误：我们真正要找的，并不是两个孩子都死于SIDS的概率，而是两个孩子的死亡，都是由SIDS所造成的概率。克拉克人狱两年后，皇家统计学会（Royal Statistical Society）通过一份新闻稿加入了这个话题，声明称陪审团的决定是基于“一个严重的、被称为检控官谬误的逻辑错误。陪审团应当衡量的，是对婴儿死因的两种相互对立的解释：SIDS或谋杀。两起死亡都是由于SIDS或谋杀所造成的可能性都十分小，但显然其中之一就发生在本案之中。真正重要的，是造成死亡的相对似然性……而不仅仅是［死于SIDS的情况］到底有多么不可能……”[12]一名数学家后来对一个家庭因为SIDS或谋杀而失去两个孩子的相对似然性作出了估计。根据可得到的数据，他的结论是，两名婴儿是SIDS受害者的可能性，9倍于谋杀致死的可能性。[13]

克拉克家因此提起上诉，并且在上诉庭中聘请了统计学家作为专家证人。上诉以败诉告终，但他们仍然继续寻求死因的医学解释，并于此期间发现了一个被隐藏的事实：检方的病理学家隐瞒了第二名婴儿在死前遭遇病菌感染的事实，而这个感染可能会导致其死亡。由于这个发现，法官撤销了有罪判决。在坐了差不多3年半的牢房后，克拉克被释放出狱。

著名律师、哈佛大学法学院教授阿兰·德什维兹（Alan Dershowitz）也曾同样成功地运用了检控官谬误——以在审理辛普森（O. J. Simpson）谋杀前妻妮可·布朗·辛普森（Nicole Brown Simpson）及其男性同伴的案件中，为辛普森进行辩护。辛普森这名前橄榄球明星的案子，是1994～1995年的媒体大热点之一。警方掌握了足够的不利证据：他们在其住所发现了一只染血的手套，而且与在谋杀现场发现的那只似乎正好是一对；在两只手套上、辛普森

的白色福特 Bronco 中、卧室中的一双袜子上以及车道和房屋中，都发现了与妮可血型相符的血迹；更进一步地，罪案现场提取到的血样 DNA 分析结果，与辛普森的 DNA 相吻合。辩方除了以种族主义——辛普森是一名非裔美国人——指责洛杉矶警察局，并对警方的可靠性以及所提供证据的权威性提出指责之外，大概就帮不上更多的忙了。

检方决定在案子开场时，将焦点集中在辛普森对妮可的暴力倾向上。检察官将开庭后最初 10 天的时间花在了展示被告虐待妻子的证据上，并声称仅仅根据这些证据，就已经有很好的理由怀疑被告谋杀了被害人。如他们所言："打耳光就是杀人的前奏。"[14] 而辩方律师就利用检方的这一策略，来作为指责检方不诚实的出发点。他们争辩说，检方花了两周的时间，不过是在试图误导陪审团。这些辛普森之前狠揍妮可的证据根本说明不了什么。下面就是德什维兹的推理过程：在美国，有 400 万妇女每年都遭到丈夫或男友的殴打，而在 1992 年，根据 FBI 犯罪报告汇编，共有 1432 名——或者说 2500 名中有 1 名——妇女真正被丈夫或男友杀害[15]。因此辩方称，在那些抽伴侣耳光或殴打伴侣的男人中，没有几个真正发展到了谋杀的地步。这对不对呢？对。确信无疑？是。与本案有关？无关。与本案有关的数字，不是一个殴打妻子的男人会进而杀害她的概率（1/2500），而是一名遭殴打并被谋杀的妻子是被她的施虐者所杀害的概率。根据 1993 年的美国及其领地犯罪报告汇编，德什维兹（或检方）应当提供的概率是下面这个：1993 年，在所有遭谋杀的被虐美国妻子中，差不多有 90%是被施虐者所杀害。这个统计数字却没有出现在法庭上。

当判决时刻即将到来时，长途电话的通话量掉了一半，纽约股票交易所的交易量则下降了 40%。据估计，大概有 1 亿人打开电视或收音机聆听最终判决：无罪。德什维兹大概对于他误导陪审团的行为问心无愧，因为按他的话说："法庭宣誓——'陈述事实、事实

之全部以及仅仅事实’——只适用于证人。辩护律师、检察官和法官都不用进行这个宣誓……事实上，可以很公平地说，美国司法系统是建立在不说出全部事实的基础之上。”[16]

尽管条件概率是认识随机性的一场革命，贝叶斯本人却一点也不革命。他的工作虽然于 1764 年发表在地位尊崇的《哲学汇刊》上，但却在无人理睬中慢慢憔悴。让科学家们注意到贝叶斯的观点，并最终向世界说明如何根据观察结果来推算出隐藏的真实概率，这一重任就落在了另一个人——法国科学家与数学家拉普拉斯（Pierre-Simon de Laplace）① ——的肩上。

你大概还能记得，在我们实际进行一系列扔硬币试验之前，伯努利的黄金定理能告诉我们观察到某个特定试验结果的可能性——如果硬币是公平的话。你可能也还记得，这个定理并没有告诉我们，在进行了特定某一系列投掷之后，投掷结果能表明硬币是公平的可能性有多大。按同样的思路，如果我们知道一名 85 岁老人能活到 90 岁的可能性为 50 对 50，那么黄金定理就能指出，1000 名 85 岁老人组成的群体，其中一半在接下来 5 年中过世的几率有多大；但如果某个群体中有一半的人在 85 岁到 90 岁的年龄段去世，定理却不能告诉我们，这个群体中个体生存率为 50 对 50 的可能性有多高。再换个例子。假设福特知道，它生产的每 100 辆汽车中，就有 1 辆存在传动故障。那黄金定理能告诉福特，在一批 1000 辆汽车中，有 10 辆或更多会存在传动问题；但如果福特在一个 1000 辆汽车的样本中，发现了 10 辆存在传动问题的汽车，定理却不能得出，这批汽车中存在传动问题车辆的平均数为 100 辆中有 1 辆的似然性有多大。这些情况中，后一个问题常常更加有用：除了与赌博有关的情况之

① 拉普拉斯（1749～1827）是法国数学家和天文学家，天体力学的主要奠基人，分析概率论的创始人，应用数学的先驱。

外，我们通常没有关于所需概率的理论知识，因而必须在一系列观察之后对其进行估计。科学家们也发现他们处境相同：他们一般不是在给定了某个物理量的值之后，去得到对该物理量测量的结果是这个或那个的可能性，而是在给定了观测结果后，去试图发现某个物理量的真实值。

我着重强调了这两类问题的区别，因为它十分重要。这个区别定义了概率与统计之间最根本的不同：前者关心的是根据确定概率来进行预测，而后者考虑的是根据观测数据来推出那些概率。

拉普拉斯阐明的就是后一类问题。他并不知道贝叶斯的理论，因此不得不亲自再重做一遍。在他框架中的问题如下：给定一系列观测值时，我们对被观测量的真实值所能做出的最佳估计是什么？这个最佳估计落在真值“附近”——不管这个附近的定义如何苛刻——的机会是多少？

拉普拉斯的分析始于 1774 年的一篇论文，分析过程本身却持续了 40 多年。尽管拉普拉斯是一个才智超群而且有时也相当慷慨的人，但他也不时地借用他人成果却不表示任何的承认或谢意，同时他还是个不知疲倦的自吹自擂者。最重要的是，拉普拉斯是棵墙头草，根据风向来确定自己的朝向。这个性格使他得以连续从事其开创性的研究，而几乎未受周围动荡时局的打扰。法国大革命之前，拉普拉斯获得了一个油水颇丰的王家炮兵考官职位，在职期间，他幸运地主持了一个名叫拿破仑·波拿巴（Napoleon Bonaparte）的有前途的 16 岁考生的考试。当革命于 1789 年到来时，他曾在短期内被怀疑为反革命分子，但与许多人不同，他毫发无伤地脱身而出，公开声明了他那“对于王权无法磨灭的痛恨”，并最终从共和国获得了新的荣誉。到 1804 年，当他的老熟人拿破仑将自己加冕为帝时，他又立刻告别了共和主义，而于 1806 年获得了伯爵爵位。波旁复辟后，拉普拉斯在他的专著《概率的分析理论》（*Théorie analytique des probabilités*）1814 年版中对拿破仑大加抨击地写道：“对于一个

企图统治世界的帝国，精通几率计算的人们能够以非常高的概率来预言它的破灭。”[17] 而早一点的 1812 年版则被献给了“拿破仑大帝”。

拉普拉斯的政治灵活性对数学而言是件幸事，因为他的分析最终比贝叶斯的理论具有了更丰富的内容和更完整的理论。有了拉普拉斯所打下的基础，我们将在下一章离开概率论的领域，而进入统计学的地盘。这两者的连接点，就是所有数学和科学学科中最重要的曲线之一——钟形曲线，又被称为正态分布。这条曲线，以及随之而来的新的测量理论，将是下一章的主题。

第七章

测量与误差定律

不久前的一天，儿子阿列克赛回到家中，通报了他最近一次英语作文的分数——93 分。在这种情况下，我通常会祝贺他得了个 A；可这次他只得了一个分数较低的 A，而且我知道他可以得到更好的分数，所以我通常还会多句嘴说，这个分数表明，如果他能再努力一点点，那下次就能得到更高的分数。不过这次的情况可不一般，而我认为 93 分是一个令人大吃一惊的对作文真实质量的低估。现在你可能觉得这些话似乎更多地在说我自己，而不是阿列克赛。果真如此，那你可说到点子上了：实际上，这段前言的确是完全说我自己的，因为是我写的那篇作文。

是是是，我的做法很可耻。作为辩护，我应当指出，通常我不会帮阿列克赛写作文，就像我不会在他的功夫课上替他挨踹在脸上的那一脚一样。不过，那天阿列克赛照例在交作业的头天深夜来到我跟前，希望我点评一下他的作品。我说我回头再看看。当我在计算机上看他的文章时，刚开始只是作了一两处不值一提的小改动，然后，我这个无情的改写者就渐渐发现自己深陷其中：这里重新排一下文字，那里干脆重新写过。到完成时，不但儿子已经睡着，而且文章也成了我自己的作品。第二天早上，我睡眼惺忪地承认没有把原始版本另存一份加以保留，因此我告诉他，直接把我的那份交上去就成。

他把成绩单递给了我，上面写着几句鼓励的话语。“还不坏，”他告诉我，“93 分确实更像个 A一而不是个 A，但那天也真是很晚

了，而且我相信，如果那天你更清醒些的话，肯定会做得更好。”我可一点儿也高兴不起来。首先，一个 15 岁孩子对你说这些本该是你对他说的话，可丝毫也不叫人感到舒服，而且我能觉察到他的话一点儿也不真诚；但除此之外，我的作文——一个至少我母亲认为是专业作家的人的作品——怎么可能在高中英语课上拿不到高分？显然，在这个问题上我并非孤军奋战。事后，我听说了另一个有着类似经历的作家的事情，不过他的女儿拿的可是个 B。很明显，这位拥有英语博士学位的作家，他的文笔好得至少足以应付《滚石》（*Rolling Stone*）、《时尚先生》（*Esquire*）和《纽约时报》（*New York Times*），但却对付不了《英语 101》。阿列克赛还试着用另一个故事来安慰我：他的两名朋友有次曾将完全相同的作文一起交了上去。他觉得他们很蠢，而且肯定会被逮住，但操劳过度的老师不仅没有注意到这两篇文章完全雷同，而且还给其中一个打了 90 分（A），却给另一个打了 79 分（C）。[听起来很怪，除非你跟我一样，熬个通宵来给高高一摞卷子打分，而只有旁边重播的《星际旅行》（*Star Trek*）来打破这单调。]

数字似乎总是带有权威性。人们差不多都会——至少下意识地——认为：如果教师按 100 分制来打分，那即使是 1 分 2 分的微小差别，也一定真地意味着些什么。但如果连 10 个出版商都相信最早那部《哈利·波特》的手稿不值得出版的话，可怜的芬尼根太太（并非前面那位老师的真名）怎么可能如此精细地区分出两篇作文的好坏，来给一个打 92 分而另一个打 93 分呢？即使能够接受作文质量可以在一定程度上加以定义的观点，也仍应认识到分数并不是作文质量的描述，而更多地是对作文质量的测量。随机性对我们最重要的影响方式之一，就是影响测量。在这个作文的例子中，测量装置是教师，而教师给出的评价如同任何测量值一样，容易受随机变化和误差的影响。

投票也同样是一种测量。此时所测量的，不是每位候选人在投

票日当天得到多少人的支持，而是有多少人真正在乎这次选举并不嫌麻烦地跑去投票。某些合法选民也许会发现，他们的名字并没有出现在已登记的选民名单上；另外一些人则可能误将选票投给了他并不支持的人。当然计票也有误差。有些选票被不适当地接受或拒收；还有些选票则凭空消失了。在大多数选举中，这些因素累加的总和并不足以影响选举结果。但在候选人高下难分时，它就能产生实质影响。我们常常通过一次或多次重新计票来处理这种情况，好像这第二次或第三次计票，会比第一次计票所受的随机性影响更少似的。

例如，在 2004 年华盛顿州州长竞选中，尽管最初的计票表明共和党候选人靠着总数约 300 万张选票中的 261 张而获胜，但民主党候选人最终被宣布胜出[1]。由于第一次的得票数如此接近，该州法律要求进行重新计票。这次还是共和党赢了，只是领先差距缩小到了 42 票。两次计票结果这 219 票的差别，实际上已经是新的领先票数的好几倍，至于是不是有人觉得这是个坏兆头，我们就无从得知了。但头两次计票的后果，则是第三次纯“手工”计票。这个 42 票的胜利，对应着每 70000 张选票赢 1 票的优势，因此，手工计票的作用，可以被比拟为让 42 个人从 1 数到 70000，并希望平均数错的次数为每人 1 次。所以选举结果并不令人吃惊地再次改变，民主党人反赢了 10 票。当后来将新发现的 700 张“丢失选票”包括进来后，这个优势变成了 129 票。

上面的计票和投票过程都非尽善尽美。比方说，如果因为邮局犯错，每 100 个原本打算投票的选民中，就有 1 个未能收到通知投票地点的邮件而没有投票，那么在华盛顿州的这场选举中，仅这一项错误就能产生 300 名本想投票却因为政府的过失而没有投票的选民。与所有的测量一样，选举并非精确无误，重新计票也是如此。因此，当选举结果极其接近时，也许我们应该接受现实，或干脆扔硬币决定胜负，而不是把票重新数了又数。

测量的不精确性是18世纪中叶的一个主要问题。那时，天体物理和数学研究者的首要任务就是使牛顿定律与所观测到的月球与行星运行相吻合。要从若干彼此不同的观测值中产生单独一个数，方法之一就是对这些观测值求平均，或者说取它们的均值。现在看来，似乎是年轻的牛顿在进行光学研究中，最早将此方法用于上述目的[2]。但如同在许多其他事情上一样，牛顿的做法在当时显得十分另类。在牛顿那个时代乃至其后的一个世纪里，大多数科学家都不使用均值，而是从观测结果中挑出一个“黄金数”——那个他们的第六感认定的最可靠结果。他们之所以这么做，是因为他们不是把测量值的变化视为测量过程不可避免的副产品，而是视其为实验失败的结果——有时甚至还会带来道德上的麻烦。实际上，他们很少公开同一个量的多个测量结果，否则就觉得这等于承认对结果进行了修补，从而带来信任危机。但到了18世纪中期，情况开始发生变化。今天，计算天体的完整运行轨迹——一系列的近圆形椭圆——是一件十分简单的事情，早熟的高中生们甚至可以一边戴着耳机听音乐，一边就把它给解决了。但是要更精细地描述行星的运动，则不仅需要考虑太阳引力，还需要考虑其他行星的引力，以及行星和月球的形状与完美球体之间的偏差。直至今日，这仍是一个巨大的难题。人们不得不将复杂的近似数学公式与不完美的观测结果加以弥合，以完成前述的任务。

在18世纪晚期出现的对测量之数学理论的需求，还另有一个原因：1780年代的法国兴起了一种新的严格的实验物理学方式[3]。在这之前，物理学由两种相互分离的实践方式所组成。一方面，数学科学家们研究牛顿关于运动和重力之理论的精确推论；另一方面，一群有时被称为实验哲学家的人们，则对电、磁、光和热进行经验主义式的研究。这些实验哲学家们——常常是业余爱好者——相比于面向数学的研究者而言，不那么注重严密的科学方法论。因此，一场对实验物理进行改革并使之数学化的运动发动了起来。其中，

拉普拉斯再次扮演了主角。

正是通过被尊为现代化学之父的同胞安托万-洛朗·拉瓦锡(Antoine-Laurent Lavoisier)① 的工作，拉普拉斯才对物理学发生了兴趣[4]。拉普拉斯和拉瓦锡共同工作了多年，但在度过动乱岁月方面，拉瓦锡却未能表明他跟拉普拉斯一样是个老手。为了赚钱支撑那许多的科学实验，他成为国家保护的收税官这个特权私人组织中的一员。这种职位大概从来也不会激发起你的公民同胞们邀你到家中享用热乎可口的姜饼配咖啡的热情；而法国大革命到来时，它更成为后果特别严重的铁证。1794 年，拉瓦锡与组织的其他成员一同被捕，并很快被判处死刑。一直以来都是一名专注的科学家的拉瓦锡，请求能再给他一点时间，来完成若干研究以留给子孙后代。对这个请求，主持法官给出了他那著名的答复："共和国不需要科学家。"现代化学之父很快被砍了头，尸体扔进了一个合葬墓。据称他还指示助手去数一数，看他那被砍下的头颅还能说出几个字来。

拉普拉斯、拉瓦锡以及其他一些人——特别是进行电磁实验的法国物理学家查理-奥古斯丁·德·库仑（Charles-Augustin de Coulomb)② ——的成果，改变了实验物理学的面貌，此外在 1790 年代，还对一种新的有理单位制之发展作出了贡献。这个新单位系统——公制系统——的目的是要替代那些各自为政、阻碍科学发展并经常导致贸易纠纷的多个单位系统。公制系统由法王路易十六（Louis XVI）所指派的一个小组发展起来，在路易被推翻后为革命政府所采用。讽刺的是，拉瓦锡就曾是这个小组的成员之一。

天文学和实验物理学方面的需求，意味着在 18 世纪晚期到 19 世纪早期，很大一部分数学家的任务，就是理解和量化随机误差。他们的努力带来了一个新领域：数学统计学。它提供了一整套工具

① 拉瓦锡（1743～1794）是法国化学家，近代化学的奠基人之一。
② 库仑（1736～1806）是法国物理学家和工程师。

来解释观测值和实验数据。统计学家有时认为，现代科学就是围绕着测量理论的建立而成长起来的。统计学同样为解决现实问题——诸如药物有效性或政治家受欢迎程度等——提供了工具，因此，正确理解统计推理，不仅在科学工作中十分有用，在日常生活中也是如此。

尽管测量总伴随着不确定性，但在说明测量结果时，测量中的不确定性却很少被谈到。生活中有不少自相矛盾的地方，这就是其中之一。如果一个挑剔的交警告诉法官说，她的雷达枪测出你在限速 35 英里/小时的路段开到了 39 英里的时速，你通常无法逃脱被开超速罚单的命运，即使雷达枪的读数实际上常常会发生数英里之多的改变[5]。尽管对许多学生（及其家长）而言，只要能让 SAT 数学考试成绩从 598 分提高到 625 分，要他们跳楼都没问题，但没有几位教育者指出，有研究表明，如果想多拿个 30 分，那只要简单地多考几次就行，而且达到目的的可能性还不算小[6]。有时，毫无意义的差别甚至能成为新闻。近年的某个 8 月份，劳工部就业统计数据显示失业率为 4.7%，而在 7 月份，这个比例是 4.8%。这个改变马上就上了头条，比如像《纽约时报》上的这个：《上月的职位与薪水中等增长》[7]。但用 *Barron's* 杂志经济版编辑基恩·爱泼斯坦（Gene Epstein）的话来说："仅仅数字上的改变，并不一定意味着事情本身就真的发生了变化。比如，在任何时候，失业率都可能发生 0.1 个百分点的变化……这个变化如此之小，以至于我们无法确定是否真有改变发生。"[8]换句话说，如果劳工部在 8 月份统计了失业率，并在一小时后再重新统计一次，那仅仅由于随机误差的作用，就很可能使两次结果之间相差至少 0.1 个百分点。如果是这样的话，《纽约时报》的头条是不是该换成《下午两点的职位与薪水中等增长》？

当测量对象是一个如阿列克赛的英语课作文那样的主观量时，测量中的不确定性造成的问题就更大。比如，宾夕法尼亚州克莱瑞

恩大学（Clarion University of Pennsylvania）的一群研究者们，收集了120个学期的试卷加以非常仔细的审查。可以确定，这种仔细程度是我们孩子的作业永远享受不到的：每张试卷都由8名教员独立打分。按A到F七等级制打分的最终结果，有时会相差2个甚至更多等级。平均而言，这个差别约为1个等级[9]。由于学生的前途常常依赖于类似的评价，因此，这个不精确是十分不幸的。但要知道，不论从方法上或是理念上而言，任何一所大学中都有着从卡尔·马克思（Karl Marx）式到格劳乔·马克思（Groucho Marx）①式的各式各样的教授，如此一来，分数上的差别也就可以理解了。但是，如果在一定程度上控制这些影响打分的因素，即固定打分的教师，并指示他们按某个固定的打分依据来进行，又会如何呢？艾奥瓦州立大学（Iowa State University）的一名研究者将大约100名学生的作文，交给了一群主修修辞与专业交流的博士生，并事先按评分标准对他们进行了高强度训练[10]。每篇作文由两名独立评分者打分，等级为1到4等。当比较他们的分数时，两名打分者意见一致的情况仅仅发生在约一半的试卷上。在其大学入学作文考试分数上进行的一项分析中，得克萨斯州立大学（University of Texas）也发现了类似的结果[11]。即使是可敬的大学委员会（College Board），它所期望的也不过是，当由两位打分者进行评判时，“在按6分制打分的SAT作文考试中，92%的作文所得的两个分数之差在±1分之内。”[12]

另一个被赋予超出其所能确保之可信度的主观性测量，就是葡萄酒的评分。回溯到1970年代，当时的葡萄酒业还处于沉睡之中。行业在增长，但主要来自于廉价的低等佐餐酒的销售。在1978年发生了一件事，而葡萄酒业的迅猛发展常被归功于此：小帕克（Rob-

① 格劳乔·马克思（1890～1977）是美国著名的喜剧演员。

ert M. Parker Jr.）这位律师出身的自封的酒评家，决定在评论之外，再用 0 到 100 的数字给葡萄酒打分。之后，大多数其他酒类出版物都沿用了这套体系。如今，美国的葡萄酒年销售额超过了 200 亿美元，而那些数以百万计的葡萄酒迷们，绝不会在把钱放到柜台上之前，不先看看酒的评分是否支持他们的选择。因此，比如那次《葡萄酒鉴赏家》（*Wine Spectator*）杂志给 2004 年份 Valentín Bianchi 的阿根廷赤霞珠打了个 90 分而非 89 分时，这刚刚多出来的一分，就转变成了 Valentín Bianchi 销售额上的巨变[13]。实际上，如果看看你当地的葡萄酒店，那些因吸引力不够而常常沦为促销品和低价货的酒，得分一般都在 80 多 90 分差点的这一档。但如果对 90 分的 2004 年份 Valentín Bianchi 阿根廷赤霞珠，在比方说一小时之后再重新给它打一次分，那新得分为 89 分的可能性是多少呢？

威廉·詹姆斯（William James）在其 1890 年的《心理学原理》（*The Principle of Psychology*）中认为，品酒专家们能将其能力发挥到如此的程度，来判断出某杯 Madeira 白葡萄酒是盛自酒瓶的上部还是下部[14]。在这么多年来参加的品酒会中，我注意到，如果我左边那个大胡子的家伙嘟哝了一声“大鼻子”（这酒闻起来挺香），其他人也肯定会发出同意的共鸣。但如果以独立无讨论的方式来各自打分，那么常能发现，那个大胡子家伙写的是“大鼻子”；另一个剃了头的家伙则潦草地写道“没鼻子”；而这个烫发的金发女人写的则是“有趣的鼻子，似乎有欧芹及刚晒过的皮革的味道痕迹”。

从理论的观点来看，有许多理由去质疑葡萄酒评分的显著性。首先，味道的感知依赖于味觉和嗅觉刺激之间复杂的相互作用。严格来说，味觉来自舌头上的 5 种感知细胞：咸、甜、酸、苦和鲜味感知细胞。最后一种细胞对某些氨基酸成分（例如普遍存在于酱油中的某些成分）发生反应。但如果这些就是味觉的全部，那任何东西——你最爱的牛排加烤土豆加苹果派大餐或美味的意大利肉酱面——就都可以通过餐桌盐、蔗糖、醋、奎宁及味精模仿出来。走

运的是，使我们食指大动的并非就是这些味道，而是还掺和了嗅觉。饮用两杯浓度完全相同的糖水时，如果在其中一杯里加入一些（无糖的）草莓香精，那么你会觉得这杯水更甜。嗅觉能够解释这一点[15]。我们所感知的葡萄酒味道，是 600 到 800 种可挥发性有机组分的大杂烩，同时在舌头上和鼻子中所产生的效果[16]。有研究证明，即使是受过品味训练的专业人员，也很少能分辨出一个混合物中多于 3 种或 4 种的组分[17]，因此，分辨产生葡萄酒味的大杂烩就真是一个问题了。

我们的预期同样能影响味道的感知。1963 年，3 名研究者秘密地在白葡萄酒中加入了一点点红色食用色素，使之变成了玫瑰红色。然后，他们请一群专家给染了色的酒的甜度打分，并与未加色素的酒进行比较。由于对酒的甜度所抱有的期望，专家们都认为假的玫瑰红葡萄酒比白葡萄酒更甜。另一组研究者则交给一群酒学专业学生两杯葡萄酒样品。这两杯样品都是同样的白葡萄酒，但其中之一加入了无味的葡萄花青素染料，使之看上去像是红葡萄酒。同样由于对不同品种酒的味道预期，这些学生觉得红色酒和白葡萄酒在味道上存在差别[18]。2008 年的一项研究中，志愿者们被要求给 5 瓶酒打分。他们给标价为 90 美元的酒的分数，要高于另一瓶标价为 10 美元的酒，但实际上，狡猾的研究者们在两个瓶子里灌的是相同的酒。而且，实验中还对被试者的大脑进行了磁共振成像。结果表明，当被试者们品尝他们相信是更贵的酒时，被认为是对愉快体验进行编码的那个大脑区域，确实更为活跃[19]。但在评判这些品酒行家之前，应该考虑到下面这一点：某研究者先问明 30 个喝可乐的人是否更喜欢可口可乐或百事可乐，然后再让他们品尝放在一起的两个牌子的可乐，以检验他们的偏好。30 人中有 21 人都声称，品尝确认了他们的选择。实际上，不厚道的研究者将可口可乐倒进了百事可乐的瓶子，而将百事可乐倒进了可口可乐的瓶子[20]。在进行评价或测量时，我们的大脑并非仅依赖于直接的感知输入，而是额外还结合

了其他的信息源——比如期望。

品酒还常常被期望偏误的反面——缺乏相关背景因素——所欺骗。把一大块山葵放在你的鼻子下面，你大概不会把它跟一瓣大蒜搞混；当然，你大概也不会将大蒜的味道误认为是——打个比方——你运动鞋里的味道。但努力去分辨一杯清澈液体的气味，则根本就是徒劳。在缺乏其他相关因素的情况下，搞混气味的可能性相当大。至少当两名研究者让专家们判断一系列 16 种随机抽取的气味时，情况就是这样：大概每 4 种气味中，这些专家们就会认错 1 种[21]。

这么多令人产生怀疑的原因，促使科学家们设计了多种方式，来直接测量品酒专家们对味道的分辨力。一种方法是酒味三角。它并非是真正的三角形，而是个比喻：每位专家被提供了 3 杯酒，其中两杯完全相同。他们的任务是找出那个不同的样品。1990 年的一项研究中，专家能正确识别出不同样品的几率仅为 2/3，即大概每 3 次中有 1 次，这些品酒宗师们就会在品酒挑战中，不能区别出比方说有着“野草莓、甜黑莓和覆盆子的馥郁醇香”的黑比诺，和有着“干李子、黄樱桃和丝滑黑醋栗的独特气味”的黑比诺[22]。还是这个实验中，一群专家被要求在包括酒精含量、是否含丹宁酸、甜度和果味度等 12 个方面给多种酒来打分。专家们的意见在 9 个方面上明显不一致。最后，在根据其他专家的描述来指出相应的酒的测试中，被试者的正确率仅仅只有 70%。

酒评家们都很清楚所有这些困难。“从许多层面而言……（这个给酒评分的体系）是没有意义的。”《葡萄酒与烈酒杂志》（*Wine and Spirit Magazine*）的编辑这样说道[23]。而按《葡萄酒狂热者》（*Wine Enthusiast*）某前编辑的话说：“你对此了解越深，就越能了解到这个东西是多么地被误导却又误导人。”[24]但评分体系却仍然发展蓬勃。为什么呢？酒评家们发现，当他们试图用星级制或简单的好、坏或者糟透了这样的口语描述来概括酒的质量时，他们的意见

并不令人信服；但如果用的是数字，那么购买者对于评价的态度就简直是崇拜了。数字分数尽管十分可疑，却能让购买者相信，他们从不同种类、不同酿酒商和不同年份葡萄酒的稻草堆中挑中的，就是那一枚金针（或银针，这就要看他们的预算而定了）。

要使利用数字来下定论的质量测量方式确实能应用于一种酒——或一篇作文　　之上，测量理论就必须解决两个关键问题：如何根据一系列不同的测量值来决定这个数？给定有限个测量值时，怎样才能估计出此数即为正确答案的概率？我们现在就来看看这两个问题，因为无论数据来源为客观或主观，问题的答案都是测量理论所要达到的目标。

理解测量的关键，在于理解由于随机误差所造成的数据变化的本性。假设将若干种酒提供给 15 个酒评家，或在不同时间重复提供给某 1 个酒评家，或将两者结合进行，那通过得分的平均数或说均值，就可以干净利落地汇总酒评家对各种酒的意见。但重要的不仅仅是均值：如果所有 15 份酒评都认同某种酒的得分是 90 分，这传达了一种信息；而如果 15 个分数是 80、81、82、87、89、89、90、90、90、91、91、94、97、99 和 100，这传达的又是另一种信息。两组数据均值相同，但数据与均值之间的差别却不同。由于数据点的分布方式如此重要，数学家们就创造了一个数值量度，来描述数据中的波动。这个数值被称为样本标准差。数学家有时还利用这个波动的平方来进行量度，它称为样本方差。

样本标准差描述了一簇数据与其均值之间的接近程度，或者用更实用的话来说，描述了数据中不确定性的大小。该值较小时，数据都落在均值附近。如所有酒评分都是 90 分的数据，其样本标准差为 0，这告诉我们，所有数据都与均值相同。但当样本标准差较大时，数据就没有簇集在均值附近。那个取值在 80 到 100 的葡萄酒分数集，其样本标准差为 6。根据一个经验规则可知，大多数评分与均

值的差都落在 6 分之内。此时我们对这种酒所能真正给出的评价，是它的分数很可能落在 84 分到 96 分之间。

在判断测量结果的真实意义方面，18、19 世纪的科学家们面临着与可疑的酒评同样的问题。如果由一群研究人员来进行一系列观测，那他们各自得到的结果几乎总会相互有所不同。一个天文学家可能碰到了不利的大气条件；另一个的望远镜可能被微风给吹动了些；第三个则可能是刚刚与威廉·詹姆斯品完 Madeira 酒之后回来。1838 年，数学家与天文学家贝塞尔（F. W. Bessel）① 归总出了发生在望远镜观测过程中的 11 类随机误差。即使是同一名天文学家进行重复测量，不可靠的视力或温度对测量机构的影响之类的变数，也会导致测量结果的改变。因此，天文学家们必须要知道，在给定一系列有差异的测量结果时，如何才能确定天体的真实位置。但尽管酒评家和科学家面对着相同的问题，却不能仅仅因此而认为各自的解决之道也相同。能否找到具有普适性的随机误差之特征？还是说随机误差的特征依赖于具体的问题情境？

最早认为不同类型的测量数据也拥有共同特征的人之一，就是雅可布·伯努利的侄子丹尼尔。1777 年，他将天文观测中的随机误差，与弓箭手所射箭支的飞行轨迹偏差做了类比。他推导说，在两种情况下，目标——被测量的真实值或靶心——都应该落在中心附近的某处，而观测结果则应聚集在它周围，而且，离目标较近的结果应该比远离目标的结果更多。他给出的描述这个分布的定律并不正确，但重要之处在于，他洞悉到弓箭手误差的分布，也能反映天文观测误差的分布。

误差分布服从某个普适定律——有时称为误差定律——是作为测量理论之基础的中心法则。它神奇地隐含着下面的推论，即当满足某些非常常见的条件时，通过同一个数学分析，就能根据测量值

① 贝塞尔（1784～1846）是德国天文学家和数学家，天体测量学的奠基人。

确定任意类型的真值。运用这样一条普适法则，根据天文学家的观测数据集来确定某天体真实位置的问题，就等价于在只知道箭支落点处小洞的位置时确定靶心位置，或根据一系列分数来确定葡萄酒“品质”的问题。这也是数学统计得以成为贯通一体的学科而非仅仅是一袋子小诀窍的原因：不管所重复的测量是为了确定圣诞节凌晨 4 点时木星的位置，还是为了确定某条生产线出品的提子面包的重量，误差的分布总是一样的。

这并不是说随机误差是影响检测结果的唯一一种误差。如果一群酒评家中的一半只喜欢红葡萄酒，另一半只喜欢白葡萄酒，而除此之外，他们都完全赞同彼此的看法（而且观点也完全一致），那么一种特定的葡萄酒所获分数的分布，就不会遵循误差定律，而是形成两个尖峰，一个尖峰对应于红葡萄酒爱好者所给的分数，而另一个则对应于白葡萄酒爱好者所给的分数。但即使在一些该定律的适用性不那么明显的场合——从橄榄球比赛中获胜方领先的分数分布[25]到 IQ 分数，误差定律通常都确实适用。许多年以前，我曾握有数千张某消费软件的登记卡。一位朋友为 8 岁和 9 岁的孩子设计了这个软件，但软件销售不如预料中来得好。到底是什么人在购买该软件呢？制作了一些表格后，我发现软件的最大购买人群是 7 岁的孩子，一个不受欢迎但并非不可预期的目标人群上的错配。但真正让我震惊的，则是我为了看看随着购买者年龄逐渐偏离 7 岁这一均值时，相应的购买者数量减少的情况，而画出的条形图。这图形非常眼熟——就是误差定律中的那条曲线。

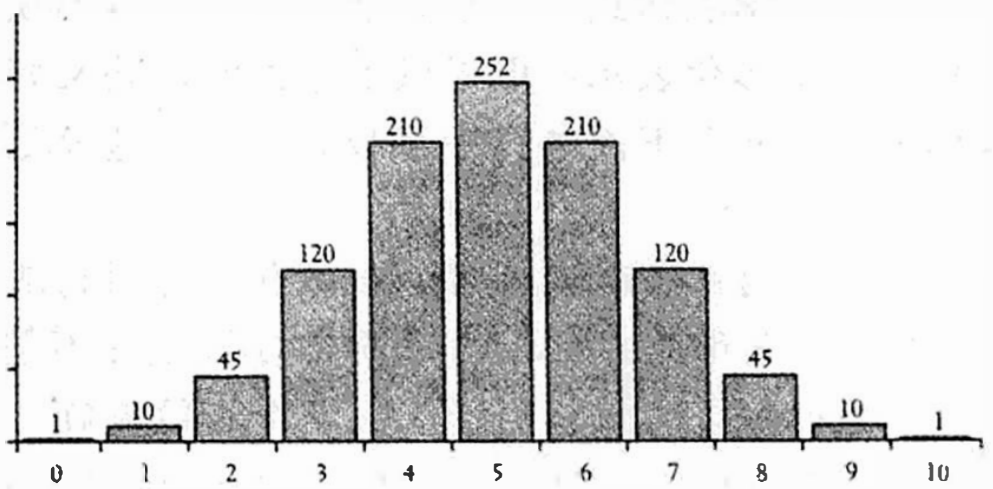
能够怀疑到箭手和天文学家、化学家和市场营销人员遇到的是同一条误差定律，这是一回事；但发现定律的具体形式，却是另一回事。在分析天文数据的需求驱动下，18 世纪晚期，类似丹尼尔·伯努利和拉普拉斯这样的科学家们，设计了一系列有瑕疵的候选假设。最终结果表明，描述误差定律的正确数学函数——钟形曲线——其实由始至终就在他们的眼皮子底下。好几十年前，在一个

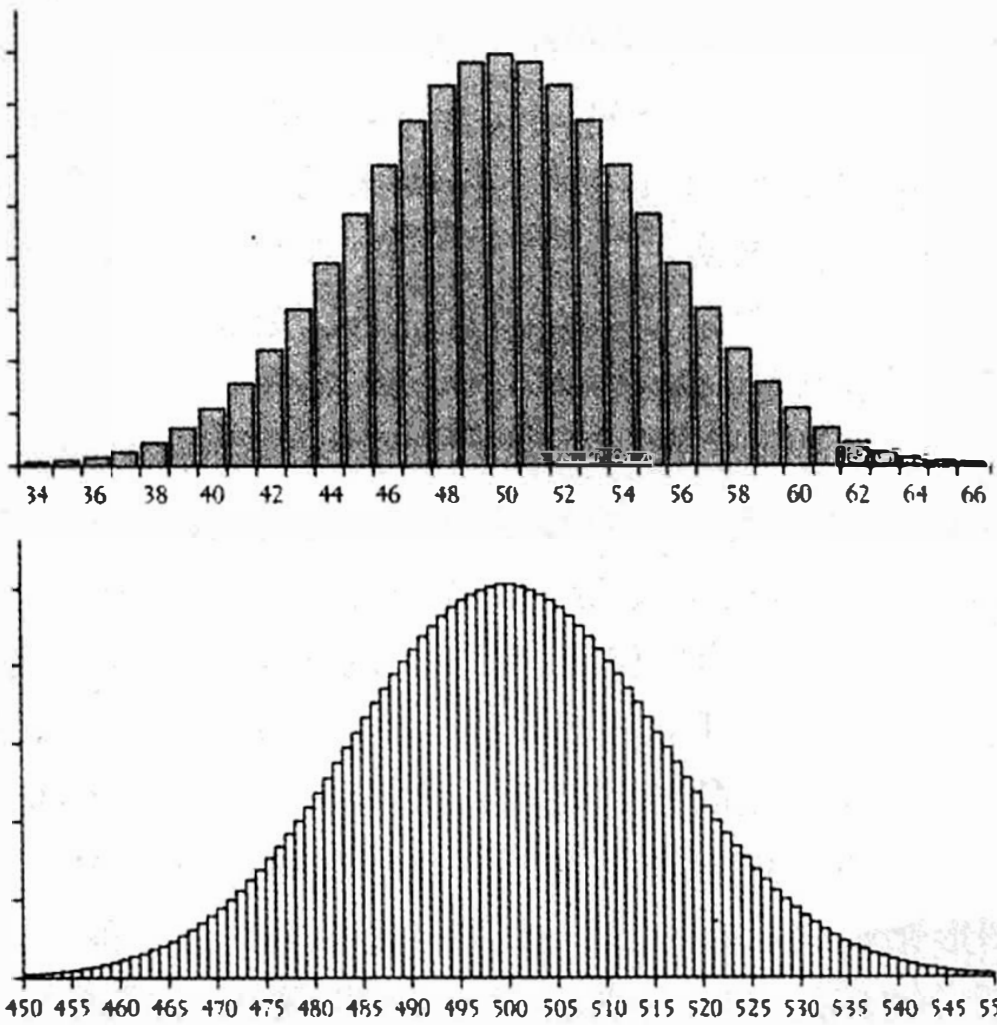
不同的场合中，这个函数就已经在伦敦被发现了。

在揭示钟形曲线重要性的3个人中，其发现者分到的功劳却是最少。亚伯拉罕·棣莫弗（Abraham De Moivre）的突破产生于1733年，当时他正是65岁左右的年纪，而直到5年后他的《机率论》（*Doctrine of Chances*）第二版出版，人们才知道了这个突破。棣莫弗要找的，是那个被我们在第10行就截断的帕斯卡三角形，当它继续向下深深延伸几百行或几千行时，这些延伸区域中数字的近似值。这个追寻最终通向了钟形曲线。雅可布·伯努利在证明他自己版本的大数定律时，也不得不为这些行中的数字的某些性质而大伤脑筋。这些数字可以非常大，例如在帕斯卡三角形第200行中的一个系数，有着59个数位！在伯努利的时代——实际上在计算机出现之前——这样的数字显然非常难以计算。而我提到过，伯努利在证明定理时进行了种种近似，并最终降低了结果的实用性，这就是原因所在。而利用所发现的曲线，棣莫弗可以得到远好得多的近似系数值，从而极大地改进了伯努利的估计。

如果像在登记卡上所做的一样，我们也将帕斯卡三角形某行的数字，以高度在条形图中表示出来的话，那么棣莫弗推导得到的近似就是一个很显然的结果。比如，在帕斯卡三角形第三行中的3个数字是1、2、1。在条形图中，第一个条形为1个单位高；第二个的高度是第一个的一倍；而第三个条形的高度又再次为1个单位。现在来看看第五行中的五个数字：1、4、6、4、1。这时的图中将有五根条形，且同样起始于矮条，在中间上升到顶点，然后又对称地降下来。那些非常下面的行，其系数会形成有非常多条形的条形图，但行为方式却是一样的。帕斯卡三角形的第10行、第100行和第1000行所对应的条形图在139页中给出。

如果用一条曲线将这些条形的顶部连起来，就出现了一个特征形状——它们趋近于一口钟的形状。如果再将曲线弄光滑些，就可





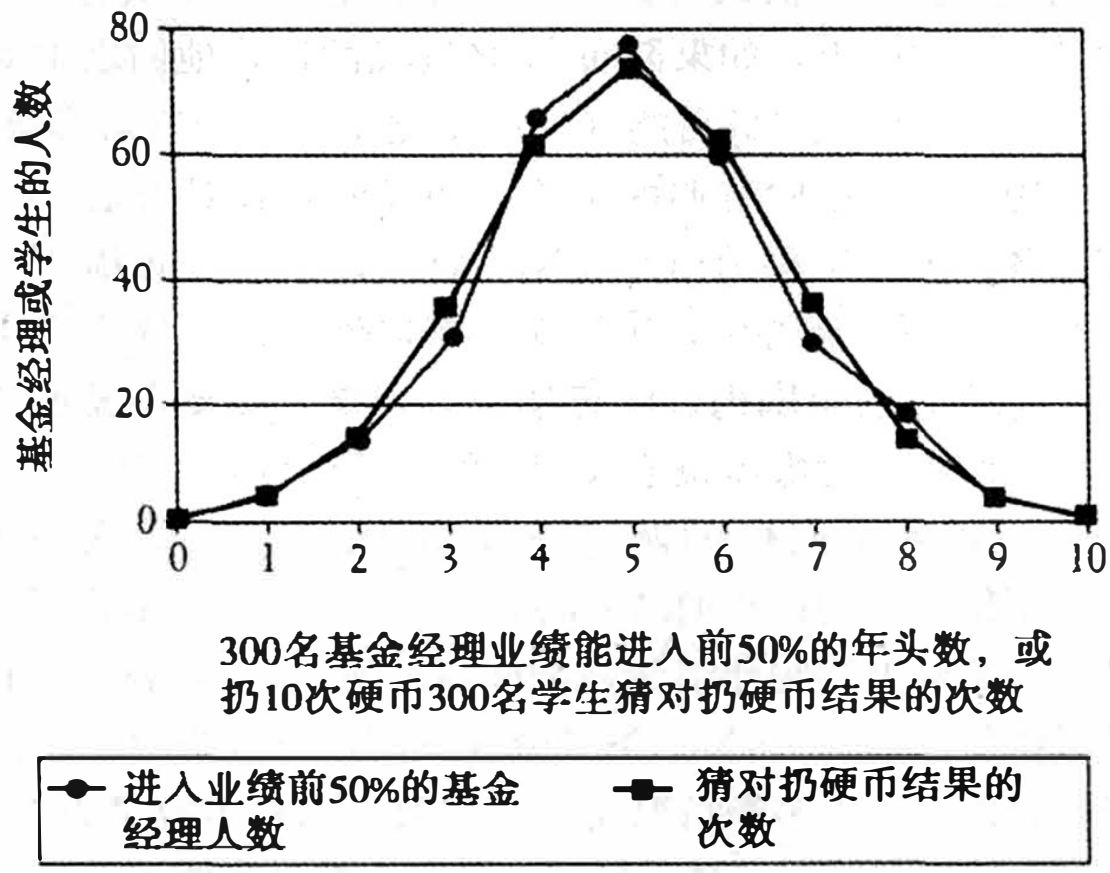
以上图中的条形高度代表帕斯卡三角形（原书第 72 页）中第 10、第 100 和第 1000 行中各系数的相对幅值。横轴上的数字代表对应系数的序号。按惯例，这些序号由 0 而不是 1 开始。（中图和下图中的横轴进行了截断，省略了那些其条形高度可以忽略的系数。）

以写出它所对应的数学表达式。这条光滑的钟形曲线不仅仅是帕斯卡三角形里数字的图形化表示，更提供了一种既精确又好用的估计帕斯卡三角形中那些非常下面的行中系数的方法。这就是棣莫弗的发现。

今天，钟形曲线通常被称为正态分布，有时也叫高斯分布（我们将在后面看到这个名称的由来）。正态分布实际上并不是一条固定的曲线，而是一族曲线，其位置与形状由两个参数来确定。第一个参数确定了曲线峰值出现的位置，在此图中，峰的位置分别为 5、50 和 500；第二个参数则确定了曲线的散布程度。这个散布程度的现代名称——标准差——直到 1894 年才出现，同时也正是早先所说的样本标准差这个概念的理论对应物。粗略来说，标准差等于曲线上值约为最大值的 60%之两点间宽度的一半。如今，正态分布的重要性已远远超出了近似帕斯卡三角形中数字的用途。实际上，它是我们所发现的最普遍的数据分布方式。

当用于描述数据分布时，钟形曲线表明，进行多次观测后，大多数结果将落在均值附近。曲线的峰就表明了这一点。不仅如此，当曲线对称地朝两边逐渐降低时，它给出的，是得到特定观测结果——大于均值与小于均值的结果——的次数是如何逐渐减小的。在服从正态分布的数据中，大约 68%（差不多 2/3）的观测值将落在距均值为 1 倍标准差的范围之内，大约 95%落在 2 倍标准差的范围内，而 99.7%的观测值都落在了 3 倍标准差的范围内。

为了图形化地表示这一点，让我们看看 141 页上的图。在图中，以正方形标出的，是 300 名学生对 10 次扔硬币结果进行猜测的数据[26]。水平轴上的数字，代表这 10 次中猜对的次数，从 0 次到 10 次。垂直方向上画出的，是正好猜中水平轴中相应次数的学生个数。曲线的形状像一口钟，中心处在猜中 5 次的位置，该点处的曲线高度对应了 75 名学生。在左边，差不多是 3 次到 4 次中间的地方，曲线高度大致为最高值的 2/3，即约 51 名学生；而在右边，这个高度



猜对扔硬币的结果与成功选中股票之间的比较

所对应的点落在了 6 次与 7 次的中间。对于猜测扔硬币的结果这样的随机过程而言，产生这个数量级标准差的钟形曲线，是一个典型的结果。

在同一张图上，我们还画出了另一组以圆形标出的数据。这组数据给出的，是 300 名基金经理的表现。此时，水平轴表示的不是 10 次扔硬币中猜对的次数，而是 10 年中某个经理的表现高于这个群体平均水平的年份数量。注意两条曲线之间的相似性！在第九章中，我们将回到这个问题上来。

要感性认识正态分布与随机误差的关系，一个好办法是考虑民意调查　　或说抽样　　过程。你也许还能回想起第五章中那个巴塞尔市长支持度的调查。在城里，有确定的一部分选民支持市长，也有确定的另一部分不支持他。简单起见，假设这两种人各占 50%。

如我们已经看到的，被调查人群有可能不能正好反映这个 50 对 50 的对立情况。实际上，如果询问 N 名选民的话，他们之中支持市长的人数为任意给定数量的可能性，就正比于帕斯卡三角形第 N 行中的数字。因此，根据棣莫弗的工作，如果调查者对大量选民进行调查，那么各种调查结果的出现概率，就能以正态分布描述。换而言之，大约在 95%的调查中，所得支持率将落在 50%这个真实支持率附近 2 个标准差的范围内。民意调查方用*误差幅度*来表示这一不确定性。如果调查者对媒体说某次调查的误差幅度是正负 5%，他们的意思是指，如果将同样的调查重复多次，那么 20 次中有 19 次(95%)，所得结果会在正确结果的 5%误差范围内（尽管调查者们很少提及，但实际上这句话当然还意味着，每 20 次调查中大概就有 1 次的结果错得离谱）。作为一条经验性规则，容量为 100 的样本所产生的误差幅度，对大多数情况而言都太大；而样本容量为 1000 时的误差幅度差不多是 3%，对大多数情况而言都已足够了。

无论在何时对任何类型的调查或民意测验结果进行评价时，关键都在于我们应该认识到，当这个调查或测验重复进行时，结果会不出意外地发生变化。举例来说，如果真实情况是有 40%的登记选民认同总统的作为，那么 6 次独立调查更可能给出像 37%、39%、39%、40%、42%和 42%这样的结果，而并非所有这 6 次调查都发现总统的支持度为 40%（前面这 6 个数字，实际上是 2006 年 9 月的头两周中，对美国总统任职情况支持度进行的 6 次独立调查结果）[27]。这也就是为什么——同样作为一条经验规则——任何误差幅度之内的差别都应当被忽略的原因。但是，《纽约时报》虽然不会登出《下午两点的职位与薪水中等增长》这样的头条，但在报导政治民意调查时，类似的头条却随处可见。举个例子，2004 年共和党全国大会之后，CNN 发出的头条是《布什的支持度显然有中等程度的反弹》[28]。CNN 的专家们接着就解释道“布什在大会上的支持率反弹看来约为 2 个百分点……潜在选民中声称要选他为总统的比例，

从大会前的50%立刻上升到了之后的52%”。只是到了后面，报导者才指出，调查的误差幅度为正负3.5个百分点，这实际等于说上面的新闻根本毫无意义。显然，“显然”这个词在CNN评论节目中的真正含义是“显然不”。

对于许多调查而言，大于5%的误差幅度都不可接受。但在日常生活中，我们却常常根据远少于必需的数据量来进行判断。人们不可能打100年的职业棒球，或投资100幢公寓楼，或开办100家巧克力曲奇饼公司。因此，在评价他人在以上方面所取得的成功时，我们的判断仅仅只依赖于少数几个数据点。橄榄球队是否应该撒出5000万美元来引诱那个刚度过了一个创纪录赛季的家伙呢？那个想要你钱的股票经纪人能重复她早先的成功的可能性有多高？那个富有的发明者在“海猴”上的成功，是不是意味着他在“隐身金鱼”和“即活蛙”的新点子上成功的几率也挺大呢？（根据记载，他的新点子并不成功。[29]）① 当看到一次成功或失败时，我们所观察到的只是一个数据点，一个钟形曲线上的采样，它仅仅代表了那早已存在的可能性之一。我们无法得知这个单独的观测所代表的，到底是均值本身，还是某个“例外值”；也无法知道这个观测所对应的，到底是我们有把握赌一把的事情，还是某个不大可能再次产生的罕见情况。但至少应当清楚，一个采样点就是一个采样点，相比于简单地将其作为真实情况来接受，我们更应在产生该采样点的标准差或概率分布这个语境下来认识它。这瓶葡萄酒的分数可能是91分，但如果没有足够信息来估计这同一种酒被多次评分或交由其他人评分的过程中可能发生的分数变动的话，91分这个数字就毫无意义。下面这个例子也许能帮助我们理解这一点：几年前，《企鹅版澳大利亚好葡萄酒指南》（*The Penguin Good Australian Wine Guide*）以及

① “海猴”是美国人哈罗德·冯·布朗哈特（Harold von Braunhut）所发明的一种即活生物。海猴实际是一种咸水虾，当被放入水中时就会活过来。

《葡萄酒》（*Wine*）杂志的《澳大利亚葡萄酒年鉴》（*Australian Wine Annual*），都对 1999 年份米切尔顿黑木园雷司令酒进行了点评。《企鹅版》指南按五星制评分给了该酒五颗星的分数，并称之为《企鹅版》年度最佳葡萄酒；而《葡萄酒》杂志则将其列在了所有参评酒的末位，并确信这是 10 年以来最差的酒[30]。正态分布不仅可以帮助我们去理解这样的矛盾，而且还使无数在科学和商业中都得到广泛使用的统计应用得以成为现实——比如，某制药公司要评判某次临床试验的结果是否具有显著性，或某制造商要评判部件的一个样本是否准确反映了缺陷产品所占的比例，或某市场商要决定是否根据一项研究调查的结果行事，这些都是统计应用的例子。

认识到正态分布描述了测量误差的分布，则是棣莫弗取得成果之后数十年的事了。发现这一点的人是德国数学家卡尔·弗里德里希·高斯（Carl Friedrich Gauss）①，他的名字也常常与钟形曲线联系在一起。高斯在解决行星运动问题时开始认识到正态分布的这个性质，至少对于天文测量而言如此。但高斯的“证明”——如他后来所承认的那样——是有问题的[31]。而且，他也未能发现正态分布的许多远为深刻的推论。因此，他将这条定律不显眼地插入了《天体运动理论》（*The Theory of the Motion of Heavenly Bodies Moving about the Sun in Conic Sections*）结尾部分的一节中。定律很可能就此埋没，从而成为那些越来越多的被摈弃的误差定律之一而已。

将正态分布从角落中拉出来的人是拉普拉斯。他于 1810 年接触到高斯的著作，而之前不久，他才刚刚在法兰西科学院宣读了一份备忘，证明了一条称为中心极限的定理，该定理称，大数量独立随机因素总和之取值为任意给定值的概率，服从正态分布。比如，假

---

① 高斯（1777～1855）是德国数学家、物理学家和天文学家，有“数学王子”的美誉，成就遍及数学的各个领域。

设你烤 100 条面包，每次都按生产 1000 克面包的配方来进行。由于随机性，你可能会有时多加一点或少加一点面粉或牛奶，或者在烘烤时多烤掉一点或少烤掉一点水分。最终，这许许多多的因素会使你的面包多个几克或少个几克，但中心极限定理告诉我们，这些面包重量将服从正态分布。阅读高斯的著作时，拉普拉斯立刻意识到，他可以用正态分布来改进他自己的工作，并可给出一个比高斯更好的论述，以支持正态分布的确就是误差定律的说法。拉普拉斯忙不迭地给备忘加了个小尾巴。今天，中心极限定理和大数定律成为了随机性理论中最为著名的两个结果。

为了说明中心极限定理如何解释正态分布确实是正确的误差定律，让我们再来看看丹尼尔·伯努利的弓箭手例子。我就曾扮演过这个弓箭手的角色。那是在一个晚上，在令人愉快的葡萄美酒与成年人陪伴的幕间节目后，我的小儿子尼古拉递给我一张弓和一支箭，问我敢不敢射落放在他头上的苹果。箭头是柔软的泡沫塑料，但先对我可能的失误及其可能性进行一下分析看来还是很合理的。出于显然的原因，我主要考虑垂直方向的误差。误差的一个简单模型如下：每个随机因素——比如瞄准的误差、气流的影响等——都会使我射出的箭在垂直方向上偏离目标，或者高些或者低些，两者出现的可能性相等。如果我走运的话，这些影响因素中大概有一半将使箭朝上偏离，而另一半使之朝下偏离，从而使箭能准确命中目标。但如果我不走运（或者更贴切点说，我儿子不走运）的话，所有这些误差都会使箭朝同一个方向偏离，从而使箭或高或低地远离目标。问题是：这些误差正好相互抵消，或正好累加形成最大误差，或取值在前两者间任意一点处，各自的可能性有多大？这些影响因素形成了一个伯努利过程——就如同扔硬币并问扔出某个特定次数的正面的可能性有多大一样。帕斯卡三角形描述了问题的答案，或者当试验次数很多时，正态分布给出了答案。这也正是中心极限定理所说的内容。（本次射箭挑战的结果是，我既没射中苹果，也没射中儿

子，但却打翻了一杯上好的红酒。）

到 1830 年代，大多数科学家已经开始相信，每个测量值都是一个复合物，受到很多误差源的影响，并因此服从误差定律。由此，误差定律和中心极限定理便使我们能获得对数据及其与物理现实之间关系的崭新且更为深刻的理解。在紧接着的下一个世纪中，对人类社会感兴趣的学者们也掌握了这个想法，并惊奇地发现，个人性格与行为的变化，也常常表现出与测量误差相同的模式。因此，他们试图将误差定律的应用，从物理科学拓展到一门关于世事的新科学之中。

第八章

混沌中的秩序

1960 年代中期，一名 90 多岁高龄的法国妇女让娜·卡尔芒（Jeanne Calment），由于急需维持生活的资金，便和一个 47 岁的律师做了笔交易：她将自己的公寓低价卖给律师以换取每月的生活费，不过，这笔生活费将一直支付到她死的那一天，届时她被抬出去，而律师就搬进来[1]。律师肯定知道卡尔芒女士的寿命已超出法国人均期望寿命达十余年之多。但是，他大概并不清楚贝叶斯定理，所以也不知道与这笔买卖真正有关的，不是她会在一10 年中去世的概率，而是当她已经活到 90 岁时，所剩的期望寿命大概还有 6 年[2]。尽管如此，律师还是觉得很放心，因为他相信，任何一个女人，如果她十来岁时就已经在父亲的店中遇见过文森特·凡·高（Vincent van Gogh）①，那现在的她，一定用不了多久就会和梵高在一起了。（据记载，该女士认为我们的艺术家“邋遢，穿着很差，而且令人生厌”。）

10 年后，律师不出所料地另找了一处住所，因为让娜以良好的健康状况庆祝了她的第 100 个生日。尽管她的期望寿命这时只有大概 2 年，但靠着律师的奉养，她仍然度过了 110 岁生日。此时律师已经 67 岁了。又过了 10 年，律师漫长的等待终于到了头，但却是猜不到的结局。1995 年，律师去世了，而让娜还活着。她的那一天最终于 1997 年 8 月 4 日到来，这时已是 122 岁的高龄了，比律师去

---

① 凡·高（1853～1890）是荷兰画家，后印象派三大巨匠之一。

世时的岁数大了整整 45 岁。

个体的期望寿命——包括其生活本身——是无法预计的。但如果从许多群体采集数据并加以大规模总体分析，规律性的模型就会显现出来。假设你已经无事故地驾驶了 20 年。然后在那个命中注定的下午，你跟妻子和她的娘家人一道在魁北克度假时，岳母突然大叫一声："当心驼鹿!"而你也一打方向盘，撞向了写着同样内容的警示牌。对你而言，这一定是起不会再来一次的很特别的事件。但是，警示牌的存在本身就体现了一种客观需要，即在数以千计的驾驶员群体中，肯定有一定比例的人会碰上驼鹿。实际上，随机行动的人所构成的统计集合，常常表现出一致而可预测的行为，就如同一群有意识地追求目标的人一样。或者如哲学家康德在 1784 年所写的："每个个人，按照自己的偏好，追逐着自己的目标，这目标常与他人的目标相左；但每个个人或民族，都通向一个自然的但对他们每个而言都属未知的目标，就好像追随着某条引领之线一般；所有人都在朝着这个目标前进的道路上出力，哪怕当他们知道目标的存在却对它不加重视时，也一样地在出力。"[3]

例如，根据联邦公路管理局的资料，在美国大约有 2 亿名驾驶员[4]。根据国家公路交通安全局的数据，在最近某年中，这些驾驶员们共行驶了约 2.86 万亿英里[5]，差不多每人 14300 英里。现在，假设全国人民都觉得，在下一年中再次达到同样的总里程数，会是件很好玩的事。那我们来比较一下可以达到这一目标的两种方法。方法 1 中，政府在国家自然科学基金会的超级计算中心之一，建立一个定量配给系统，将每人的目标行驶里程分配下去，以便在满足 2 亿司机各自需要的同时，保持头一年那 14300 英里的平均里程。在方法 2 中，我们告诉司机，不要有任何压力，你想开多远就开多远，不用管头一年到底跑了多少路。如果从来都是步行去酒铺上班的比利·鲍勃叔叔，这次决定要当个霰弹枪推销员，到西得克萨斯跑上 100000 英里，这没有问题；曼哈顿的简表妹总是把大部分路程花在

了在道路清扫日里围着她住的街区团团乱转以找到 1 个停车位上，但现在却结婚去了新泽西，那我们也不用去担心她的里程数是否因此而改变。哪种方法能更接近地达到 14300 英里这个人均年里程数目标呢？方法 1 无法进行试验，不过，有限的汽油配给经历表明，这个方法大概不会很奏效；另一方面，方法 2 却是实际中采用的方法——也就是在下一年中，司机们高兴开多远就开多远，根本不试图去刚好达到某个配额。他们做得如何呢？根据国家公路交通安全局的数据，那一年中，美国驾驶员们跑了 2.88 万亿英里，或者说每人 14400 英里，只比目标多了 100 英里。而且，这 2 亿名驾驶员中，还产生了几乎同样多——相差不到 200 人——的交通死亡人数（42815 人对 42643 人）。

我们总将随机性与无序联系在一起。但尽管 2 亿名驾驶员的生活发生着无法预知的变化，我们却大概很难证明，他们作为整体的行为能比上面那样更加有序。观察投票、购买股票、结婚、被报失踪、写错地址的信件，或是在参加一个本不想首先参加的会议途中被堵在路上之类的情况，或是测量人们腿的长度、脚的尺码、臀部的宽度或是啤酒肚的高度时，都能发现类似的规律性。当 19 世纪的科学家们钻研新近才能获得的社会学数据时，不论何处，生命的混沌似乎总形成可量化可预测的模式。但使人震惊的不仅仅是这些规律性，更包括数据变动的特性。他们发现，社会学数据也常常服从正态分布。

人类特征与行为的差异，以及弓箭手瞄准的误差，两者具有相似的分布情况，这个事实促使 19 世纪的一些科学家开始研究，人类本身之存在的这支箭，它瞄准的目标到底是什么。而更重要的，则是去理解有时使这目标发生改变的那些社会与物理方面的因由。因此，数学统计学这个为帮助科学家分析数据而发展起来的学科，在一个远为不同的领域——对社会本性的研究——之中兴旺发展起来。

至迟到 11 世纪，统计学家们已经开始分析日常生活数据了。当时，征服者威廉（William the Conqueror）派人进行了事实上的史上首次全国普查。威廉的统治始于 1035 年，时年 7 岁的他继承了父亲诺曼底公爵的位子。如其绰号所暗示的，威廉公爵二世喜欢征服，并在 1066 年入侵了英格兰。圣诞那天他送给自己的礼物，便是得以加冕为王。迅速的胜利给他留下了一点小问题：被他征服的到底是什么人？更重要地，从这些新臣民身上，他能收多少税？为了解答这些问题，他向英格兰的每个角落都派去了巡查，记录下每一块土地的大小、所有者及其上的资源[6]。为保证结果正确，他又派了第二批巡查来重复第一批的工作。由于当时的税并非按人头而是按土地面积及用途来收取，因此，巡查们付出了足称得上勇敢的努力，来清数牛、奶牛和猪的数量，而对清理它们排泄物的人，则没有留下什么记录。即使人口数据与税收有关，在中世纪对人最重要的统计数据——寿命与疾病——进行调查，也被认为有违传统基督教的死亡观念。按照基督教教义，将死亡作为思考对象是错误的，而寻找掌管死亡的规律则简直就是渎神。不管人们是死于肺部感染、胃痛，或是由于石头的冲击力超过了头骨的承受力，他或她的真正死因，都被简单地认为是上帝的意旨。花了许多个世纪，这个宿命的态度逐渐退出，而相反的观点产生了出来。根据该观点，对自然和社会的规律性进行研究，并非是挑战上帝的权威，而是试图去了解上帝行事的方式。

这个观念转变的一大步，发生在 16 世纪。当时伦敦市长下令编撰每周一期的“死亡表”，以归总教区教士所记录的洗礼和葬礼。编撰工作在数十年中零星地进行着，但在 1603 年这个瘟疫最为猖獗的年头之一，伦敦创立了周记制度。欧洲大陆的神学家对这些堆满数据的死亡表不屑一顾，视之为典型的英国作派，而且几乎毫无用处。但是，对店主约翰·格朗特（John Graunt）这个英国人来说，它们叙述了一个颇为引人的故事[7]。

格朗特和朋友威廉·配第（William Petty）[1] 被称为统计学的奠基人。纯数学研究者们有时会觉得这个领域十分浅薄，因为它关注的是世俗的实用问题。从这个意义来说，格朗特是一个尤其相称的创建者。与某些发展概率论的业余爱好者　　医生卡尔达诺，法官费马，或是教士贝叶斯　　不同，格朗特是个出售常用小饰物的商人：纽扣，针线，或是其他家居小玩意儿。但格朗特却不仅仅是个纽扣商：他是个有钱的纽扣商。富足使他有足够的闲暇去追求那些与如何把衣服扣在一起之类的问题毫无关系的兴趣与爱好。财富还使他得以和一些当代最伟大的知识分子成为朋友，其中就包括配第。

格朗特从死亡表得出的一个推论与饿死的人数有关。1665 年，这个数字据报称是 45 人，大概只是被处决人数的一倍；相比之下，4808 人死于肺病，1929 人死于斑疹热，2614 人死于牙病和寄生虫病，68596 人死于瘟疫。是什么原因使得伦敦在“乞丐成堆”的时候，却只有这么几个人饿死呢？格朗特认为，普通民众肯定给乞丐施舍了食物，使他们免于饿死。他由此提出，应由国家来提供食物，从而使社会大众不必付出多少，而又能使 17 世纪的伦敦街道，避免出现本该与之相称的叫花子和擦车仔们。格朗特还参加了瘟疫传播方式的两种理论之间的争论。一个理论认为，这种疾病通过污浊的空气传播，而另一个则认为，传播是通过人与人的接触来进行的。格朗特观察了一周又一周的死亡记录，最后认为，数据变动幅度之大，令人无法相信这种改变是完全随机的。在他看来，如果人--人传播理论正确的话，数据变化就应该表现出随机性；而另一方面，由于每周的天气都有不规律的变化，所以他觉得，这些波动的数据与污秽空气理论相一致。后来的结果表明，伦敦城并没有为开设粥场做好准备，而伦敦市民们如果远离丑陋的老鼠而不是肮脏的空气的话，大概会活得更好些。但是格朗特的伟大发现不在于其结论，

---

① 配第（1623～1687）是英国统计学家，英国古典政治经济学创始人。

而在于他认识到，统计数据能使我们洞察那个产生数据的系统。

有时，配第的工作被认为是古典经济学的先驱[8]。配第相信，国家的力量依赖于臣民的数量和性格，并通过它们来得到体现。因此，他运用统计推理方法分析国家大事。他的分析通常都是从统治者的视角出发，而将社会成员当作可随意摆布的物品来对待。他指出，为了对付瘟疫，应该将钱花在预防上，因为拯救生命就等于省下一笔抚养费，而王国就可以从社会为抚养男女直到成年的可观投资中省下一笔。相比于最有利可图的其他投资方式而言，前者的回报也仍然更为丰厚；对待爱尔兰人，配第可就没这么仁慈了。例如，他认为一个英国人的生命，其经济价值要大于一个爱尔兰人的生命，因此，如果把除了几个放牛娃之外的其他爱尔兰人都强迁到英格兰，便可以增加王国的财富。实事求是地说，配第自己的财富就得归功于这些爱尔兰人：作为 1650 年代入侵爱尔兰英军的随军医生，他被指派进行损失评估。评估结论是，他可以大捞一笔后安然脱身。而他也正是这样做的[9]。

如果像配第所相信的那样，人口的数量及增长反映了政府的优劣，那么，缺少测量人口数量的好方法，就会使评价政府质量变得十分困难。格朗特在此问题上展开了他最为著名的计算，特别是对伦敦人口的计算。从死亡表中，格朗特知道了出生人口的数量。而有了出生率的粗略估计，就可以推算出育龄女性的大概数量，进而估计家庭数量，再通过他自己观察到的伦敦城家庭成员的平均数量，就能估计出全城人口。他的计算结果是 384000——而之前人们都觉得是 200 万。格朗特还证明，该城人口增长的主要贡献，来自于边远地区的移民，而非更为缓慢的人口繁衍；此外，尽管仍有瘟疫造成的恐惧，但即使是最严重的疫病所造成的人口减少，也总是在两年之内就得到了恢复。这一结论颇为令人吃惊。此外，格朗特还被认为发布了第一份“寿命表”，时至今日，这个系统化的预期寿命数据排列方式，仍被对人们寿命感兴趣的组织——从人寿保险公司到

世界卫生组织——所使用。下面的寿命表显示了在每 100 人中能活到任意指定寿命的人数。在格朗特所列数据（下表中标着“伦敦，1662”的这一列）之外，我又额外增加了某些国家现在的数据[10]。

年龄	伦敦，1662	阿富汗	莫桑比克	中国	巴西	英国	德国	美国	法国	日本
0	100	100	100	100	100	100	100	100	100	100
6	74	85	97	97	99	100	99	100	100	100
16	40	71	82	96	96	99	99	99	99	100
26	25	67	79	96	95	99	99	98	99	99
36	16	60	67	95	93	98	98	97	98	99
46	10	52	50	93	90	97	97	95	97	98
56	6	43	39	88	84	94	94	92	93	95
66	3	31	29	78	72	87	87	83	86	89
76	1	16	17	55	51	69	71	66	72	77
86	—	4	5	21	23	37	40	38	46	52
96	—	0	0	2	3	8	8	9	11	17

格朗特寿命表的扩展版

1662 年，格朗特在《死亡表的……自然与政治观察》（*Natural and Political Observations … upon the Bills of Mortality*）中发表了他的分析。该书得到了广泛赞誉。一年后，格朗特入选皇家学会，然后在 1666 年，发生了伦敦大火。城市的很大一部分被烧成白地，也烧毁了他的生意。伤害之上又加羞辱，他被指控在大火发生前下令停止供水，从而成了灾难发生的帮凶。实际上，直到火灾发生之后，他才与供水公司有过联系。不管怎样，在这段插曲之后，格朗

特的名字从皇家学会名册上消失了。几年后，格朗特死于黄疸。

很大程度上，正是由于格朗特的成果，法国才在 1667 年与英国站在同一战线上，修法以进行类似死亡表这样的统计调查。其他欧洲国家紧随其后。到 19 世纪，全欧洲的统计学家们，都深陷于如普查数据这样的政府记录之中——“一场数字的大雪崩”[11]。格朗特的遗产在于，他表明通过仔细分析数据的有限样本，可以得到有关人群总体的推论。尽管格朗特与其他人进行了卓绝的努力，以便能利用简单的逻辑来了解数据中的信息，但要解释这些数据中的大部分秘密，还有待于高斯、拉普拉斯和其他 19 世纪及 20 世纪早期的科学家们所发展起来的工具。

“统计”一词源自德语单词 Statistik，是由《比尔菲尔德初等普及教育》（*Bielfield's Elementary Universal Education*）一书的 1770 年翻译版引入英语的。该书写道：“被称为统计的这门科学告诉我们，在已经探知的世界中，所有现代国家的政治组成是如何的。”[12]到 1828 年，该学科已经发生了变化，而诺亚·韦伯斯特（Noah Webster）的《美国词典》（*American Dictionary*）所定义的统计学是“有关社会状态、民族或国家中人们的情况、他们的健康状况、寿命、国内经济、艺术、财产与政治影响力以及国家的状态等事实的集合。”[13]这时，统计学已经包括了拉普拉斯的方法，而他一直都在探索，如何将他的数学分析方法的应用对象，从行星与恒星扩展到日常生活中的问题上来。

正态分布描述了许多现象围绕某中心值发生改变的行为方式，这个中心值就代表了现象的最可能输出值。在其《概率的哲学导论》（*Essai philosophique sur les probabilités*）一书中，拉普拉斯称这门新的数学学科能用于评判法庭证言、预测结婚率、计算保险费等。但当该书最后一版面世时，拉普拉斯已经 60 多岁了，因此，继承与发展其构想的任务，就落到了一名年轻人——1796 年 2 月 22 日出生

于佛兰德的根特（Ghent，Flanders）的阿多夫·奎特雷（Adolphe Quételet）①[14]——身上。

奎特雷并不是因为对社会运转方式感兴趣才开始他的研究的。1819年为他赢得根特的新大学所颁发的第一个科学博士学位的论文，是与圆锥曲线理论有关的几何学课题。然后，他的兴趣就转向了天文学。1820年左右，他积极参与运动，以寻求在他当时任职的布鲁塞尔建立一个新天文台。充满野心的奎特雷，显然将这个天文台视为朝着建立一个科学帝国所迈出的一步。这一步颇为莽撞，因为他对天文学所知不多，而对天文台运作更几乎是一无所知。但他肯定颇有说服力，因为不仅天文台得到了资助，他个人也获得了基金，从而能够用几个月的时间到巴黎去弥补知识上的不足。结果表明投资十分有效，因为奎特雷的比利时皇家天文台直到今天还仍然存在。

在巴黎，奎特雷以其个人的方式，受到了生活之无序的影响，并被拉到了一个完全不同的方向。他与统计学的罗曼史，始于与几位伟大的法国数学家的结识，其中包括拉普拉斯和约瑟夫·傅里叶（Joseph Fourier）②。他师从傅里叶，学习了统计与概率。随后，尽管他学会了如何运作天文台，但却陷入了与另一个追求目标的热恋，这个目标就是运用天文学中的数学工具来处理社会数据。

奎特雷回到布鲁塞尔后，就开始收集和分析人口统计数据，并很快将注意力集中在法国政府于1827年开始发布的犯罪活动记录上。在《论人及其能力的发展》（*Sur l'homme et le développement de ses facultés*）这部发表于1835年的两卷本著作中，奎特雷印上了一张1826年至1831年所报道的法国各年谋杀案数量的表。他注意到，谋杀案的数量相对恒定，而且，每年用火枪、剑、小刀、木棒、

---

① 奎特雷（1796～1874）是比利时天文学家、数学家、统计学家和社会学家。

② 傅里叶（1768～1830）是法国数学家和物理学家，主要贡献有数学中的傅里叶级数和热传导的傅里叶定律等。

石头、砍刺器具、拳脚、绳勒、溺毙和火烧等方式进行的谋杀案，其比例也相对恒定[15]。奎特雷还按年龄、地域、季节、职业以及是否身处医院和监狱，来对死亡率进行分析。他还对酗酒、精神病和犯罪进行了统计学研究。此外，他在巴黎自缢者人数以及比利时 60 多岁老妇配 20 多岁小伙子的婚姻数量上，也发现了描述它们的统计规律性。

统计学家们之前就进行过类似研究，但奎特雷比他们做得更多些：他不仅观察数据的平均值，还仔细研究了它们偏离平均值的方式。不管研究对象为何物，奎特雷都遇到了正态分布：在犯罪、结婚和自杀的倾向上，在美洲印第安人的身高上，乃至在苏格兰士兵的胸围测量值上（从《爱丁堡医学与外科学期刊》（*Edinburgh Medical and Surgical Journal*）的一篇旧文中，他得到了有 5738 个胸围测量数据的样本），都能看到正态分布。在为该书草稿所准备的 100000 名年轻法国人的身高中，他还发现了与正态分布之间的偏离所隐藏着的意味。将应征入伍士兵的数量对身高作图时，钟形曲线被扭曲了：身高刚刚超过 5 英尺 2 英寸的人数过少，恰好小于这个身高的人则又多了些，好像是为了补偿前者的不足似的。奎特雷论述道，这个多出来的约 2200 个“矮子”的差异，是因为造假或说善意的捏造，因为身高不到 5 英尺 2 英寸的人可以免服兵役。

几十年后，伟大的法国数学家儒勒-亨利·庞加莱（Jules-Henri Poincaré），用奎特雷的方法逮到了一个欺骗顾客的面包师。开始时，习惯每天买一条面包的庞加莱，在称量面包后发现，它们的平均重量大约为 950 克，而非广告中所称的 1000 克。他向管理部门投诉了此事。之后，他买到的面包大了些。可他还是觉得面包有什么地方不太对劲。凭着只有著名学者或至少是终身学者才有的耐心，他在接下来的一年中，每天都仔细地称量面包。尽管这些面包现在的平均重量十分接近于 1000 克，但如果这个面包师的确是老老实实地随机挑出一条面包卖给他，那比这个平均重量更重些或更轻些的面包，

其数量应如第七章所说的那样，按误差定律的钟形曲线逐渐减少。而庞加莱发现，他的面包里偏轻的太少，而偏重的则相应过多。庞加莱由此得出结论，即面包师其实并没有停止制作缺斤少两的面包，而只不过总拿手头最大的一条面包来安抚他罢了。警察再次造访了骗人的面包师，据报道所言，他表现了应有的震惊，并可想而知地同意改变自己的行事[16]。

奎特雷无意中得到了一个有用的发现：随机性的模式是如此值得信赖，以至于在某些社会学数据中，对该模式的违背，可作为指证坏事的证据。如今，这种分析被应用于奎特雷的时代所不可能处理的大量数据上。实际上，在近几年中，类似的统计侦察术正逐渐流行，并创造了一个称为法律经济学的新领域，作为其最著名案例的统计研究，证明一些公司将其优先认股权的获准时间谎称为更早些的某个时候。目的很简单：公司通过承认优先认股权——即在事后以认股权获准时的价格购买股票的权利——来激励管理人员努力帮助公司股价上涨。如果认股权获准时间被提前到股价特别低的时候，管理人员的利润就会应声大涨。这是个聪明的点子，但如果秘密进行就违反了证券法。同时，它也留下了统计学上的“指纹”，并最终导致了对一打大公司的这类行为所进行的调查[17]。在一个不那么为人所知的例子里，沃顿学院的经济学家贾斯廷·沃尔弗斯(Justin Wolfers)，在大约 70000 场大学篮球比赛结果中发现了做假的证据[18]。

沃尔弗斯在比较拉斯维加斯赌球的得分差与球赛真实结果时，发现了这个不正常的地方。对于被看好的球队，赌球庄家会给出一个球赛得分差，来使得下注于两队的人数基本相等。比方说，假设大家都认为加州理工（Caltech）篮球队比加州大学洛杉矶分校(UCLA）篮球队要强（对于大学篮球赛的球迷们来说，是的，在 1950 年代就的确如此），那么为了避免赌球双方获胜的机会不等，庄家们只会在加州理工胜出 UCLA——比方说——13 分以上时，才会

付钱给那些押宝在加州理工的人，从而使得双方输赢的机会大致均等。

尽管这个分差由庄家设定，但实际上，却是由广大赌徒们确定的，因为庄家会根据赌徒下注的情况来调整分差，从而形成需求上的平衡（庄家通过从下注的钱中抽取抽头来赚钱，因此，他们要的是两方所下赌注相等，从而无论比赛结果如何，庄家都稳赚不赔）。经济学家们使用一个称为预测误差的数字，来衡量赌徒们对两支球队评估的准确程度，这个预测误差等于被看好球队的真实胜出分差与赌博市场所设定的分差之间的差数。预测误差作为一种误差，其分布毫不奇怪地服从正态分布。沃尔弗斯发现，预测误差的均值为0，即得分差既不倾向于过高评价球队，也不倾向于过低评价它们；其标准差为 10.9 分，也就是说，在大约 2/3 的比赛中，赌场设置的得分差与真实比分差之间，相差不会超过 10.9 分（在一项对职业橄榄球赛开展的研究中，也发现了类似的结果，其均值为 0，标准差为 13.9 分）[19]。

在检查了那些参赛队强弱分明的比赛构成的子集后，沃尔弗斯发现了令人吃惊的事情：这些比赛中，被看好的队伍胜出的比分恰恰超过得分差的场次过分地少，而恰恰不到得分差的场次却无法解释地多了许多。又是奎特雷异常。与奎特雷和庞加莱一样，沃尔弗斯的结论是有人打假球。他是这样分析的：即使对于顶尖球员而言，要保证球队胜出的分数比得分差更多，也是十分困难的；但如果一支队伍明显强于对手，那么球员便可以在不危及球队胜利的前提下，很容易地通过马虎比赛来保证球队不会赢得比得分差更多。因此，如果不道德的赌博者希望改变赌博的结果，而又不要求球员故意输掉比赛的话，结果就恰好是沃尔弗斯所发现的这个与正常分布间的偏离。那么，沃尔弗斯的结果是否证明，在一定比例的大学篮球赛中，球员们因受贿而故意压低得分呢？不是。但正如沃尔弗斯所说的：“我们不能认为法庭上发生的事情，就如实反映了拉斯维加斯发

生的事情。”有意思的是，在一次由美国大学体育联合会（National Collegiate Atheletic Association）所进行的调查中，1.5%的运动员承认知道有队友“因为收了钱而故意发挥得差些”[20]。

奎特雷并不打算在法庭上应用他的想法。他有个更宏伟的计划：利用正态分布来说明人与社会的本性。他写道，如果给一尊雕像做1000个复制品，那由于测量和手艺的误差，这些复制品将各有不同，而这个不同将服从误差定律；由此推理，如果人们的物质特性也服从同样的定律，则必然是因为我们也同样是某个原型的不完美复制品。奎特雷称这个原型为“平均人”（*l'homme moyen*）。他还感到，对人类行为而言，也同样存在着一个模板。一家大百货公司的经理可能无法知道，那个好像吸毒后飘飘欲仙的出纳，会不会把他正嗅着的那瓶半盎司香奈尔 Allure 香水揣到自己兜里去；但下面这个预测却是靠得住的：零售业中，库存商品的年损失率相当稳定地维持在1.6%左右，而同样稳定的是，45%～48%的损失是员工盗窃所造成[21]。犯罪　　奎特雷写道　　就“像一笔以令人悚然的规律性来支付的预算”[22]。

奎特雷认识到，不同的文化下，“平均人”也会不同，而且随着社会环境的改变，它也随之而变。实际上，研究“平均人”之改变及造成改变的原因，才是奎特雷最大的野心。“人遵循特定的定律出生、成长和死亡。”他这样写道，而这些定律“从未得到过研究”[23]。牛顿认识了一套宇宙定律，并将其公式化，从而成为现代物理学之父。以牛顿为榜样的奎特雷，则希望创造一门新的“社会物理学”，来描述人类行为的定律。在奎特雷的类比中，正如一个物体在不受干扰时保持其运动状态一般，人类群体的行为在社会环境保持不变时，也同样保持恒定。而且，正如牛顿描述了物理力如何使物体偏离其直线路径一样，奎特雷期望找到人类行为定律，以描述社会力量如何改变社会的特性。举例来说，奎特雷认为财富的巨

大悬殊以及价格的巨大波动，对于犯罪和社会不安负有责任，而一个稳定的犯罪率水平则代表了一种平衡状态，当这个平衡状态之后的影响因素发生改变时，平衡状态也随之改变。2001 年 9·11 事件后的几个月，就是这个社会平衡变化的生动例子。当时，旅行者们由于害怕搭乘飞机而突然纷纷转为驾车出行。这个害怕就此转变成为比前一年同期额外多出的约 1000 名车祸丧生者　　9·11 事件背后所隐藏的伤亡[24]。

但是，相信这样一门社会物理学的存在是一回事，对之进行定义则是另一回事。奎特雷认识到，通过将人们置于大量的实验环境中，并测量其行为，可以对一门真正科学中的理论加以检视。由于这个方法实际不可行，所以他相信社会科学更类似天文学而非物理学，其理论的洞察需要通过被动的观测来获得。因此，为了发现社会物理学的定律，他对“平均人”随时间和文化所发生的变化进行了研究。

人们很好地接受了奎特雷的观点，特别是在法国和英国。一名生理学家甚至从一个常常为不同国籍的人们所光顾的火车站洗手间中收集了尿液，以确定“平均欧洲人尿液”的性质[25]。在英国，奎特雷最热情的信徒，是名叫亨利·托马斯·巴克尔（Henry Thomas Buckle）的富有的棋手和历史学家，而他最为人所知的，就是其野心勃勃的多卷本著作《英格兰文明史》（*History of Civilization in England*）。不幸的是，在 1861 年他 40 岁时，巴克尔在大马士革旅行期间患上了斑疹伤寒。一名当地的医生前来为他治疗，但他拒绝了这名法国医生，并因此病亡。巴克尔未能完成其著作。不过，他的确完成了头两卷，而第一卷中就以统计观点来阐述历史。它们根基于奎特雷的成果，并立刻大获成功。整个欧洲都在读这本书。它被译为了法文、德文和俄文。达尔文（Charles Darwin）读过它，阿

弗雷德·罗素·华莱士（Alfred Russel Wallace)① 读过它，而陀思妥耶夫斯基（Dostoyevsky)② 则读了两遍[26]。

尽管该书大受欢迎，历史却证明，奎特雷的数学比他的社会物理学更为合理。一方面，并非所有的社会事件都服从正态分布，特别是在金融领域中。比如，如果电影票房按正态分布，那大多数电影赚到的钱都该落在某个平均值附近，而有 2/3 的电影票房，将落在这个平均值周围一个标准差的范围内。但在电影业中，20%的电影带来了 80%的票房收入。这类由热点驱动的产业，尽管完全无法预测，却遵循了一个完全不同的分布，其中均值和标准差的概念毫无意义，因为根本就没有什么“典型”的票房表现，而那些类似百万票房的例外，在一般产业中可能几个世纪才会出现一次，在这里却隔上几年就会出现[27]。

相比于奎特雷忽略了其他类型概率分布，更为严重的是他未能在寻找所希望的定律和驱动力的道路上前进多少。因此，他对于社会物理学的直接影响最终不过是中等程度而已。不过，他的遗产却无可否认且意义深远。它并不在于社会科学，而在于所谓的“硬”科学。在这里，通过大量随机事件来理解其中之秩序的方法，启发了许多的学者，产生了革命性的成果，并改变了生物学和物理学的思维方式。

将统计学思想引入生物学的，是达尔文的大表兄。弗朗西斯·高尔顿（Francis Galton)③ 是一个悠闲的人。他在 1840 年进入了剑桥三一学院（Trinity College，Cambridge)[28]，先是学习医学，

---

① 阿弗雷德·华莱士（1823～1913）是英国自然学家、人类学家和生物学家，他独立于达尔文提出了自然选择理论。

② 陀思妥耶夫斯基（1821～1881）是俄国文学家，代表作有《罪与罚》。

③ 弗朗西斯·高尔顿（1822～1911）是英国人类学家、优生学家、发明家和统计学家。

随后就接受了达尔文的建议，改学数学。当父亲去世时，22 岁的他继承了一笔可观的财产。由于不再需要靠工作来糊口，他成了一名业余科学家。他着迷于测量。他测量人们头颅的尺寸、鼻子的大小、四肢的长短；他测量人们在听课时坐立不安的次数；他测量那些在路上看到的女孩的吸引力（伦敦姑娘们得分最高，而阿伯丁女孩最低）；他测量人们指纹的特征，并最终使苏格兰场于 1901 年采纳了指纹识别；他甚至还测量了君主和教士的寿命，它们与其他职业人群的寿命相差不多，因此他的结论是，祈祷其实没带来什么好处。

高尔顿在 1869 年的《遗传的天才》（*Hereditary Genius*）一书中写道，任何身高落在给定范围内的人数占总人口的比例，应保持恒定而不随时间而改变，而且身高与所有其他的物理特征——头颅周长、大脑大小、大脑灰质重量、大脑纤维数量等——都服从于正态分布。他相信人的性格也由遗传所决定，并如同生理特征一样，以某种方式服从于正态分布。因此，按高尔顿的话说，人“作为社会之基本单元”，并不“具有相同的在投票及其他方面能力上的平等价值”[29]。相反地，他断言每 1 百万名男人中，约有 250 人继承了某个方面的特殊才能，并因此在该领域内出类拔萃（在他那个年代，妇女一般不参加工作，因此他没有对妇女进行类似的分析）。高尔顿根据这些想法创建了一门新的研究领域，并称之为优生学，这个词来自于希腊词 eu（好）和 genos（出生）。对于其后许多年中不同的人而言，优生学意味着不同的东西。这个术语和观点被纳粹所采用，但没有证据表明，高尔顿会赞同德国人的杀人方案。相反地，他希望找到一种办法，通过选择性繁殖来改善人类状况。

高尔顿关于成功所给出的简单因果解释很有诱惑性，而第九章的大部分内容就是为了解释其缘由。但在第十章中将看到，具有任何程度之复杂性的任务，其完成过程中都充满了无数可预见的和偶然性的障碍，因此，能力和成就之间的联系，远不如高尔顿所解释地那样直接。实际上，心理学家在近年发现，就成功而言，面对困

难毫不退缩的品质，至少与天分同等重要[30]。这也就是专家们常常提及的“十年规则”，即对于多数事情而言，要能够做得相当成功，就至少需要 10 年的刻苦、实践和奋斗。勤奋与机遇跟天分同等重要，这看来也许令人沮丧，但我却觉得这更应该使人振奋，因为我们虽然不能控制自身的基因组成，但努力的程度却完全取决于自己。机遇的影响在一定程度上也能得以控制：只要反复尝试，就能提高成功的机会。

无论优生学的好坏如何，高尔顿在遗传方面进行的研究，导致了现代统计学两个核心数学概念的发现。其中之一产生于 1875 年，当时，他把几包甜豌豆荚分给了 7 位朋友，每人都收到了大小和重量均匀的种子，然后将种出的下一代种子还给高尔顿。高尔顿对这些种子进行测量后注意到，大个头种子产生出的后代，其直径的中值比父代种子小，而小个头种子产生的后代则比父代更大。后来，从他在伦敦所建实验室获得的数据中，他发现人类父母和子女身高上也存在同样的情况。他将这种现象——在相关联的测量中，如果一个被测量远离均值，那么另一个将会更接近于均值——称为回归均值。

高尔顿很快意识到，没有回归行为的过程，将不可避免地失控。举例来说，假设高个子爸爸的儿子，平均而言与父亲一样高。既然身高各有不同，那么，某些儿子就会比爸爸更高。到了下一代，假定这些更高的儿子们，他们的儿子，即开始那群人的孙子，平均而言也跟他们的父亲一样高，那么其中又有某些人同样会高过他们的父亲。这样一代又一代后，人类中最高的人将变得越来越高。由于回归，这种情况并没有发生。同样的论述也可用于天生的智力、艺术才能或是打高尔夫球的技术。因此，个头很高的父母不应该指望孩子会长得同样高，非常聪明的父母不应该指望孩子会同样聪明，而那些毕加索们和泰格·伍兹们也不应该指望孩子能取得跟他们一样的成就。另一方面，非常矮的父母可以期待更高些的后代，而我

们这些并不聪明或不能画画的人，则有理由指望这种无能将在下一代中得到改善。

高尔顿用广告将被试者吸引到实验室中，然后对他们进行了一系列身高、体重乃至特定骨骼尺寸的测量。他希望发现一种根据父母的测量值来预测孩子对应的测量值的方法。高尔顿所画的图之一种，是将父母的身高对孩子的身高来作图。打个比方，如果父母和孩子的身高总是相等，那么就应该得到一条 45°倾角的完美直线。如果这种关系对平均值成立，但单个数据点可能发生波动，那么在图中就应表现为某些点落在这条直线上方，而另一些落在下方。以这种方式，高尔顿的图不但显示了父母和子女身高间的普遍关系，还显示了这个关系得到保持情况。这就是高尔顿对统计学所做的另一大贡献：他定义了一种数学指标，来描述类似关系的一致性。他称该指标为相关系数。

相关系数是取值于－1 到 1 之间的一个数。如果相关系数很接近于 1 或－1，就表明两个变量之间是线性关系；相关系数为 0 则表示两者无关。比方说，假设数据表明，每周吃一次麦当劳最新推出的 1000 卡（1 卡≈4.18 焦）套餐，一年下来体重会增加 10 磅；如果吃两次，则会增加 20 磅；如此等等。这时的相关系数就为 1。如果由于某种原因，吃这个套餐的人会减掉上述体重，那么相关系数就是－1。如果增加和减少的体重，其取值五花八门，而并不依赖于吃了几次套餐的话，相关系数就为 0。现在，相关系数已经成为统计学中使用最为广泛的概念之一。人们用它来评估诸如吸烟量和癌症发生、恒星距地球的距离和它们远离的速度、学生在标准化考试中所得分数和他们家庭收入这样的关系。

高尔顿的成果十分重要，这不仅仅是因为它所具有的直接重要性，还因为它促成了随后几十年中统计学所取得的成绩。在这段时期内，统计学领域迅速发展成为一门成熟的学科。最重要的进展之一，是由高尔顿的信徒皮尔森完成的。在本章早些时候，我曾给出

了许多服从正态分布的数据类型。但对于有限数据集而言，它总无法与正态分布曲线完美匹配。在早期的统计学中，科学家们简单地将数据作图，然后观察所得曲线的形状，来判断这些数据是否按正态分布。但应该如何量化这个曲线匹配的精确程度呢？皮尔森发明了一种称为 χ^2 检验的方法，可以确定一个数据集是否确实服从我们所认为的分布。1892 年，他在蒙特卡罗演示了这个检验方法。他所进行的测试可以看作是贾格之事迹的严格复制品[31]。如贾格所碰到的情形一样，在皮尔森的测试中，轮盘赌开出的数字并不服从所期望的分布，即当轮子确实产生随机数时所应服从的分布。另一次测试中，皮尔森观察了在将 12 枚骰子扔出 26306 次时，每次中扔出的 5 点和 6 点的骰子的数量。他发现这个分布同样不满足使用完美骰子进行公平赌赛时所应得到的分布——在公平的情况下，一枚骰子扔出 5 点或 6 点的概率是 1/3，或说 0.3333；实际的分布却与此概率为 0.3377 时的分布相一致——换句话说，骰子是歪的。对于这个轮盘赌而言，可能是有人在暗中操纵，但骰子的问题则很可能是制造过程所造成的波动，正如我的朋友莫希所强调的，这种波动总是存在。

今天，χ^2 检验得到了广泛运用。假设现在不是对骰子、而是对 3 盒麦片进行检验，以观察消费者对它们的喜好程度。如果消费者们没有特殊的偏好，那么可以期望，每盒麦片都会被近 1/3 的消费者所选购。如我们已经看到的，真实结果很难如此均匀地分布。使用 χ^2 检验就能确定，得票最多的那盒麦片，它的胜出的确是由于消费者的偏爱而非碰巧的可能性有多大。类似地，假设某制药公司的研究者们，在实验中测试两种用于防止急性器官移植排斥反应的治疗方法，那他们便可通过 χ^2 检验，来确定所得结果中是否确实存在显著性差异。再或者在一个新的汽车租赁直销点开张之前，汽车租赁公司的首席财务官（CFO）认为会有 25%的顾客想租用微型汽车，50%的要小型汽车，还分别有 12.5%的需要中型或其他类型的车辆。

当租赁数据逐渐累积起来后，χ^2 检验就能帮助这位 CFO 来迅速判断，到底是假设正确，还是说新店是个偏离典型的例外，而公司此时最好将各种车型的比例进行相应调整。

通过高尔顿，奎特雷的工作融入了生物科学。但奎特雷同样帮助了物理科学中一项革命的发生：统计物理学创建者中的两位——詹姆斯·克拉克·麦克斯韦（James Clerk Maxwell）① 和路德维希·玻尔兹曼（Ludwig Boltzmann）② ——都从奎特雷的理论中获得了灵感（跟达尔文和陀思妥耶夫斯基一样，他们是在巴克尔的书中读到这些内容的）。不管怎样，如果 5738 名苏格兰士兵的胸围能很好地按正态曲线分布，而 2 亿名驾驶员每年的驾驶里程之间的相差可以小到 100 英里，那么也并非必须具有爱因斯坦的天才，才能猜到一升气体那 1 万亿上下的分子也会表现出某些有趣的规律性。但要使科学界最终确信这个物理学新方法的必要性，却的确需要一位爱因斯坦。阿尔伯特·爱因斯坦在 1905 年——即他发表第一篇相对论论文的那一年——做到了这一点。尽管不为流行文化所熟知，但爱因斯坦 1905 年关于统计物理学的文章，与其相对论的论文有着同样的革命性。实际上，这是他在科学界被引用次数最多的论文[32]。

爱因斯坦 1905 年的统计物理学论文，其目的是要解释一种被称为布朗运动（Brownian motion）的现象。该过程以罗伯特·布朗（Robert Brown）的名字命名。此人是植物学家，使用显微镜的世界级专家，而且被认为是第一个对细胞核作出清晰描述的人。布朗不懈努力追求的生活目标，是通过观察来发现生命力的源泉。在当时，

---

① 麦克斯韦（1831～1879）是英国物理学家，主要贡献在于电磁场理论，其他贡献还包括统计物理学、光学、力学、弹性理论等。

② 玻尔兹曼（1844～1906）是奥地利物理学家，主要贡献包括气体分子运动理论、统计力学和热力学等。

这个生命力据信是一种神秘的作用力，能将那些活物之所以成为活物的性质赋予事物。布朗的这项任务注定将以失败告终，但在 1827 年 6 月的某一天，他觉得他成功了。

透过目镜，布朗注意到，他所观察的花粉迸裂所释放出的微粒似乎在运动[33]。尽管花粉是一种生命之源，其本身却不是有生命的物体。但不管布朗盯上多久，运动却从不曾停止，就好像这些微粒拥有某种神秘的能量似的。这个运动似乎没有特定的目的地；实际上，它看起来完全随机。无比激动的布朗首先得出的结论，是他已经找到了所追寻的东西，因为这种神秘的能量，除了是为生命提供动力的能量之外，还能是什么呢？

接下来的一个月里，布朗勤勉地进行了一系列实验，并发现，对于所有他能搞到手的各色有机小颗粒——诸如分解了的小牛肉纤维、“被伦敦的灰尘染黑的”蜘蛛网、甚至他自己的黏液等，当它们悬浮于水中或——有时——于杜松子酒中时，都能观察到同样的运动。然后，他期盼的解释遭到致命一击：当他用石棉、铜、铋、锑和锰之类的无机小颗粒进行实验时，观察到了同样的运动。此时他认识到，所观察到的运动根本与生命无关。事后将证明，布朗运动的真实原因，与产生奎特雷所注意到的人类行为规律性的力量一样，不是一个实在的物理力，而是产生于随机模式的虚拟力。不幸的是，当对他所发现的现象的正确解释出现时，布朗已不在人世了。

理解布朗运动的基础，是玻尔兹曼、麦克斯韦与其他人在布朗的发现之后数十年间打下的。受奎特雷的启发，他们建立了统计物理学这一新领域，运用概率与统计的数学思想，来解释构成流体的（当时仍然是假想的）原子之运动是如何造就流体性质的。但之后几十年中，这些想法却未能得到发展。某些科学家对该理论所使用的数学工具心存疑虑；其他人则因为从未见过或根本不相信有人能见到这个所谓的原子而反对该理论。但大多数物理学家都是实用主义者，因此理论未被接受的最主要障碍在于，尽管它能导出若干已知

的流体定律，但却没有做出预测。事情就这么耽搁到了 1905 年，这时麦克斯韦过世已久，而沮丧的玻尔兹曼也快要自杀了。爱因斯坦使用这个新生的理论，通过详尽的数学描述，解释了布朗运动的准确机制[34]。统计学方法对于物理学的必要性，从此不再存有疑问，而物质由原子和分子组成的观点，将成为大多数现代技术的基础，同时也是物理学史上最为重要的思想之一。

如将在第十章中所见，流体中分子的随机运动，可视为我们自己生命之路的一个类比，因此，爱因斯坦的工作值得我们多花点时间来仔细看看。在原子论描绘的图景中，水分子的基本运动完全混乱无章。分子先是沿着这条路线运动，然后沿着那条路线飞走；它将保持直线运动，直至撞上另一个分子而改变自己的运行方向。前言中已经提到过，这类路径——在路径上的某些点上，运动方向发生随机改变——常常被称为醉汉的脚步，至于这个名字的来由，对于任何曾享用了过多一点马提尼酒的人而言，都非常显然（有时，更冷静的数学家和科学家们称之为随机行走）。按原子论的预言，如果漂浮于液体中的颗粒一直不断地被液体分子随机地轰击，那么可以预计，颗粒将因为这些碰撞，一会儿朝这边一会儿朝那边地动来动去。但布朗运动的这幅图景中存在两个问题：首先，对于可见的悬浮颗粒而言，分子实在是太小了，无法推动颗粒移动；其次，分子的碰撞发生得远比所观察到的颗粒晃动要来得频繁。爱因斯坦天才想法中的一点，就是认识到这两者其实会相互抵消：尽管碰撞发生得非常频繁，但由于分子如此之轻，所以这些频繁而独立的碰撞不能产生可察觉的效果。只有在某些时候，纯粹的运气使得某个方向上的撞击具有了优势——类似于分子世界中的马立斯纪录之年——才会发生足以使人注意的晃动。爱因斯坦在计算中发现，尽管有着分子层面上的混沌，但在分子的参数如大小、数量和速度等，以及可观察的颗粒晃动频率与幅度之间，存在着可预测的关系。有史以来头一遭，新的可观测的结果与统计物理学，被爱因斯坦联系

在了一起。虽然听起来好像主要是纯技术上的成就，但恰恰相反，这是一个伟大原则的胜利：许多我们感知的自然界秩序，在其掩饰之下，都是看不见的无序；而这些秩序也只有通过随机性的法则才能理解。正如爱因斯坦所言："能认识到那些看来与直接而可见的真理颇有出入的现象，其实可以被统一为一体，这种感觉无比美好。"[35]

在爱因斯坦的数学分析中，正态分布再次扮演了核心角色，并达到了科学史上一个新的辉煌地位。醉汉的脚步也就此确立了最具基础性、并很快得到最多研究的自然过程之一的地位。等到各领域的科学家接受了统计学方法的正统性后，在几乎所有的研究领域中，他们都发现了醉汉脚步的指纹——在清除非洲丛林以消灭蚊虫的研究中，在制造尼龙的化学中，在塑料的形成中，在自由量子粒子的运动中，在股票价格的变化中，甚至在研究无数年代内智力的进化中，都是如此。在第十章中，我们将看一看随机性对我们自身的生命之路所造成的影响。但如即将看到的那样，尽管在随机变化中，存在着有规律的模式，模式却并非总是有意义的。当模式中存在意义时，我们应该去发现它；但同样重要的，则是当模式实际并无意义时，就不该试图去抽取出"意义"。要避免随机模式"有意义"的错觉，是一件困难的工作。而这就是下一章的主题。

第九章

模式的错觉与错觉的模式

1848 年，两个十来岁的小女孩玛格丽特和凯特·福克斯，听到了一种无法解释的、如同敲打或移动家具的声音。凑巧的是，她们的住处还背着个鬼屋的名声。据故事所述[1]，凯特向声源发起了挑战，要求它重复她弯折手指时关节发出的噼啪声，并突然要它说出她的名字。声源回应了这两个挑战。随后几天里，在母亲和几个邻居的帮助下，两姐妹找到了一种“密码”与这个敲击者（我可并非有意语带双关）① 沟通。她们的最终结论是，这敲击声发自于一个小贩的鬼魂，几年前，他在现在她们所住的屋内被人谋杀。伴随着这个故事，现代招魂说——它认为死者能与生者沟通交流　　就此诞生。到 1850 年代初，被称为“敲桌子”的这种灵魂接触方法，以及技出同门的“移桌子”和“翻桌子”，在欧美风靡一时。这些通灵法由一群人参与，他们围坐在桌旁，手放在桌上，然后就是等待。在“敲桌子”中，一段时间的等待后，就会听到一声敲打声；而在“移桌子”和“翻桌子”中，桌子则会开始翘起或移动，有时甚至拽着坐在桌旁的人一起移动。有人还画下了那些严肃的、穿着长到大腿中部的夹克的大胡子男人们，以及穿着带裙撑长裙的热衷的妇女们，当他们的手跟随桌子移来移去时，眼睛因惊奇而瞪得老大。

“移桌子”变得如此流行，使得科学家们在 1853 年夏天时开始对之进行研究。一组医生注意到，在那段静坐的时间里，一种无意

---

① 原文中敲击者（rapper）一词又有“说唱歌手”的意思。

识的共识似乎逐渐成形，以决定桌子移动的方向[2]。他们发现，如果分散在座者的注意力以阻止共同期望形成的话，桌子就不会动。在另一项实验中，他们成功构造出一种情况，使半数参与者希望桌子往左移，而另一半参与者则希望它往右移，从而也使得桌子不发生移动。他们的结论是："桌子的移动是由于肌肉反应、主要是无意识训练所形成的反应而造成的。"给这个问题画上句号的研究，是由物理学家迈克尔·法拉第（Michael Faraday）① ——电磁理论的创始人之一，电机的发明者，历史上最伟大的实验科学家之一——所完成的[3]。法拉第首先发现，哪怕桌旁只坐了一个人，桌子移动的现象也仍会发生。然后，在招募了"非常高尚"而且事业有成的"移桌者"后，他进行了一系列复杂而富有创造性的实验，证明参与者双手的移动发生于桌子移动之前。他进而设计了一个指示器，一旦上述情况发生，就立刻向被试者报警。他发现，"一旦（指示器）……放到最认真的（被试者）之前……（这种错觉的）力量就消失了；而原因只能是，所有参与者都被迫意识到了他们真正在做着的事情。"[4]

和那些医生们一样，法拉第的结论是，这些坐着的参与者们无意识地拉动或推动着桌子。运动也许是由于不耐烦而随机发生的，然后在某个时刻，这些参与者们在随机运动中感知到了某个模式，而当被试者们的手跟随这个想象中的桌子移动的带领时，该模式便成为了某种自我实现的期望。法拉第继续写道，他设置的指示器，其价值就在于"它所具有的纠正这些转桌者意识的能力"[5]。法拉第认识到，人类感知并不是实在的直接结果，而是想象活动的结果。[6]

感知需要想象，因为生活中所遇到的数据从来都是不完整且模棱两可的。举例来说，多数人都认为，没有什么证据比亲眼所见更加可靠，而在法庭上，也没有什么比目击证言更受尊重的了。但如

---

① 法拉第（1791～1867）是英国物理学家和化学家。

果在法庭上展示一段成像质量与人类视网膜上形成的未经处理的数据完全一样的视频，那法官可能就会想，你到底在拖延些什么。首先，这段影像中会有一个盲点，即视觉神经与视网膜的连接点。其次，在视野中，唯一具有足够高分辨率的区域，不过是围绕视网膜中心大约 1 度视角的范围，这个范围大致相当于将手臂伸直时大拇指所对应的宽度。在这个区域之外，分辨率急剧下降。作为补偿，我们不停地移动眼球，以使成像分辨率高的区域能落在希望看到的场景的不同部分上。因此，送往大脑的原始数据，是一幅晃来晃去、酩酊大醉的图像，里面还有一个洞。幸运的是，大脑会对这些数据进行处理，将双眼传来的输入加以组合，并在相邻位置上的视觉特征相似且可内插值的假设下，来填充图像中的空隙[7]。最后的结果——至少在年老、受伤、疾病或过量的麦泰酒造成严重损伤之前——就是一个快快活活却承受着他/她的视觉既敏锐又清晰之错觉的人类。

我们还利用想象取巧地填补非视觉数据模式中的空当。如同视觉输入一样，我们根据不确定和不完整的信息，来作出结论并进行判断；而完成模式分析后的结论是，我们所得的“画面”清晰又准确。但真是如此吗？

在防止辨识出虚假模式的方向上，科学家们已经开始了迈进。通过发展统计分析方法，能判断一组观测值是否能对某个假设提供足够好的支持，或相反地，它们那表面上对假设的支持，是否完全是偶然性所致。比如，当物理学家们希望确定超级对撞机所得数据的显著性时，他们不是盯着得到的曲线来寻找噪声中冒出的尖峰，而是使用数学技巧来进行判断。这类技巧之一的显著性检验，是由 20 世纪最伟大统计学家之一的费希尔（R. A. Fisher）（同样以其无法控制的火爆脾气而著称，与同道的统计学先驱皮尔森长期不和，甚至在 1936 年皮尔森死后很久，还一直对这个死对头进行攻击）于 1920 年代发展起来的。

为了说明费希尔的思路，设想在一项超感官知觉研究中，参与的学生要预测扔硬币的结果。如果观察发现她差不多每次都能正确预测出结果，我们就可能形成假设，即她确实具有这方面的能力，比如通过意念力来预测。另一方面，如果她能正确预测出差不多半数的结果，那么这个数据支持的假设就是她只不过在瞎猜而已。但如果结果落在两者之间，或者根本就没有足够多的数据时，又该如何呢？在接受与拒绝这两个对立的假设之间，决策的分界线应该画在哪儿呢？这就是显著性检验所做的事情：它是一个规范的计算方法，用来计算如果所检验的假设为真时，我们能观测到所得数据的概率。如果这个概率小，我们就拒绝假设；反之则接受假设。

让我们举个例子。假设我们都是怀疑论者，并且这名学生不能完全准确地预测出扔硬币的结果。再假设在一次试验中，她正确预测了若干次扔硬币的结果。第四章中的分析方法可以算出她仅仅靠运气就能达到该预测水平的概率。如果她频繁地猜出准确结果，而纯靠运气做到这一点的可能性仅为3%时，我们就应当拒绝认为她完全是在瞎猜。用显著性检验的行话来说，拒绝该假设的显著性水平为3%，即由于纯偶然性而得出错误判断的可能性最多为3%。3%的显著性水平算是相当令人印象深刻了，因此媒体很可能会将这位学生的表演作为精神力存在的证据而大加报道。尽管如此，不相信精神力的人们则可能会仍持保留态度。

这个例子说明了一个重要问题：即使数据的显著性达到了诸如3%的水平，在对100名不具有精神力的人进行精神力测试——或者对100种无效药物进行有效性测试——时，我们也应该预计到，总会有那么几个人表现得好像确实具有精神力，或总有几种无效药物表现得好像确实有效一样。政治民调或医学研究——特别在数据量较小时——有时会与之前的调查或研究结果相抵触，而这就是原因之一。虽然如此，显著性检验与其他统计方法对科学家提供了很好的帮助，特别是在大规模的受控研究中。不过，我们平常并不会进

行这样的研究，也不会下意识就能用统计分析来加以处理。我们依靠的是本能。如果我的维京牌炉灶存在缺陷，而凑巧一个熟人也告诉了我同样的事情，那我就会告诫朋友们不要买这个牌子的炉子；如果在好几次乘坐联合航空（United Airlines）班机时，同机的乘客似乎总比最近坐过的其他航空公司航班的乘客更为乖戾的话，我就会避免继续乘坐联航的班机。这些情况中，数据量并不大，但我们内在的本能却找到了模式。

这些模式有时有意义，有时没有。不论是哪种情况，我们对生活中各种模式的感知既高度可信却也高度主观。这一事实有着深远的含义。它隐含了某种相对性。正如法拉第所发现的，在某些情况下，现实只存在于信者的眼中。比如，2006 年《新英格兰医学学报》（*New England Journal of Medicine*）发表了一项耗资 1250 万美元、在膝关节炎登记患者中开展的研究。研究表明，额外的葡萄糖胺与软骨素的组合营养成分，并不比安慰剂更有助于减轻关节炎的痛苦。尽管如此，某著名医生仍然难以放弃这些营养成分确实有效的感觉，因此，他在一个全国性广播节目中重申了该治疗方法的潜在好处，并不再对该研究进行分析。他指出："我妻子的一位医生养了只猫，她说如果这只猫不吃一点葡萄糖胺加软骨素硫酸盐的话，早上就根本站不起来。"[8]

看得更仔细些就能发现，现代社会中的许多假设，都像"移桌子"一样，建立在共享的错觉之上。在第八章中，我们所关注的，是随机事件中展现的令人吃惊的规律性；与之不同，接下来的内容将从相反的方向，来看看那些似乎具有确定原因的事件模式，其实可能纯粹是偶然性的产物。

寻找模式并赋予其意义是人类的天性。卡尼曼和特沃斯基分析了在评价数据模式及面临不确定性进行决策时所使用的许多简化方法。他们称这些简化方法为"启发式方法"。一般情况下，启发式方

法十分有用，但正如我们处理视觉信息的方式有时会带来视觉错觉一样，启发式方法有时也会导致系统性错误。卡尼曼和特沃斯基称这种错误为“偏误”。我们都使用启发式方法，同时也都受到偏误的影响。不过，尽管视觉错觉很少与日常体验发生关系，认知偏误却在决策中扮演着重要的角色。因此，20 世纪晚期兴起了一股研究人类思维如何认知随机性的潮流。研究者们的结论是，“人们对于随机性只有很糟糕的概念：他们对随机性视而不见，而当他们试图去产生随机性时，也无法做到。”[9] 更糟的是，我们总是照例地误判了机遇在生活中的角色，并总是做出与我们的最大利益确不相符的决定。[10]

设想有一个事件序列。这些事件可以是季度赢利，或是通过因特网约会服务获得的一串或好或糟的约会。不管是哪种情况，序列越长，或者对更多的这种序列进行观察，就越有可能发现所有那些想象得到的模式——而这些“模式”完全是凑巧的结果，根本就不需要有个什么“原因”来造成那一连串或好或糟的季度、或好或糟的约会。数学家斯潘塞-布朗（George Spencer-Brown）充分展示了这一点。他指出，在 $10^{1000007}$ 个 0 或 1 构成的随机序列中，我们应当预计到，至少会出现 10 个互不重叠的、由 100 万个连续的 0 所构成的子序列[11]。设想一个可怜的家伙，正好就撞进了这些子序列中的一个，并试图利用随机数来进行科学研究，那么他的软件就会连续产生 5 个 0，然后是 10 个，然后 20 个、1000 个、10000 个、100000 个、500000 个。他是否应该将这个软件退货呢？当一名科学家打开一本新买的随机数字的书，却发现所有的数字都是 0 的时候，他会如何反应呢？斯潘塞-布朗的要点在于，一个过程本身是随机的，并不同于这个过程产生的结果看起来是随机的。苹果公司最初在 iPod 音乐播放器上使用的随机播放歌曲的软件，就碰到了这个问题：真正的随机性有时会造成重复，但如果使用者连着听到同一首歌或同一个人唱的歌，他们就会相信这个歌曲的随机排列实际并非随机。

因此，公司进行了修改，通过使这个软件“不那么随机，来让人觉得它更随机”［苹果公司创始人斯蒂夫·乔布斯（Steve Jobs）语］。[12]

对于随机模式感知进行的最早期的思索之一，出自哲学家汉斯·赖欣巴哈（Hans Reichenbach）①。他在1934年写道，未受过概率论训练的人们难以识别出随机事件序列[13]。看看下面这个扔200次硬币所得到的结果序列，其中X表示反面朝上，而O表示正面朝上：OOOOXXXXOOOXXXOOOOXXOOXOOOXXXOOXXOOOXXXXOOOXOOXOXOOOOOXOOXOOOOOXXOOXXXXOXXOXOXXXXOOOXXOOXXOXOOXXXOOXOOXOXOXXXOXOOOXOXOOOOXXXXOOOXXOOXOXXOOOXOOOXXOXOOXXOOOOXOOXXXXOOOOXXXOOOXOOOXXXXXXXOOXXXOOXOOXOOOOOXXXX。

我们可以很容易地找到数据中的模式——比如，开始时跟在4个X后面的4个O，以及快结束时出现的连续6个X。数学计算表明，这类模式很可能会出现在200次扔硬币的结果序列中，但仍有许多人对之感到吃惊。如果这些X和O的字符串代表的不是扔硬币的结果，而是对我们产生影响的事情，那么人们就会为它们寻找有意义的解释。如果一串X代表股市连续下跌的日子，人们就会相信那些宣称股市即将崩溃的专家观点；如果一串O代表你最爱的运动明星连续获得的好成绩，那么那些大夸该运动员“稳定表现”的话语，听起来就十分令人信服；而如果像已经看到的那样，这些X或O代表的是派拉蒙或哥伦比亚电影公司的失败之作时，那电影业界的小报所称的某人能够或不能够把握全世界电影观众脉搏这样的话语，就足以让每个人都点头称是。

学术界及作家们已经投入了大量精力，来研究金融市场中的随机成功事例的模式。比如，许多证据表明，股票的表现是随机

① 赖欣巴哈（1891～1953）是科学哲学家和教育家，生于德国，卒于美国。

的——或者它是如此接近于随机，故尔如果没有内幕消息，或者在进行交易或账户管理时要额外出钱的话，人们根本就无法从随机变动中获取任何利益[14]。尽管如此，华尔街长期以来却对分析大师们有着一种不成文的信仰。1990 年代末，这些分析师的平均年薪约为 300 万美元[15]。他们做得怎样呢？1995 年的一项研究显示，*Barron's* 杂志每年都会邀请 8～12 名薪金最高的“华尔街超级明星”，举行圆桌会议来推荐股票。但这些股票的收益，也仅仅只是和市场的平均回报率相当[16]。1987 年和 1997 年所进行的研究表明，预言家们在《华尔街一周》（*Wall Street Week*）这个电视节目上所推荐的股票，表现还要差得多：它们远远落后于市场平均盈利水平[17]。哈佛经济研究所（Harvard Institute of Economic Research）的一名研究人员对 153 条时事通讯进行了分析，发现“并没有显著证据证明挑选绩优股能力的存在”[18]。

仅仅靠偶然因素，某些分析师或基金也总能表现出令人印象深刻的成功模式。许多研究都表明，这些市场中以往的成功事例，并非是未来成功的一个良好指标——即这些成功很大程度上不过是走运而已。尽管如此，大多数人还是觉得，他们的股票经纪人或运营基金专家们所提供的建议是值得花钱的。许多人，甚至包括聪明的投资者，都因此购买了管理费过高的基金。实际上有这么一项实验，一群来自沃顿商学院（Wharton business school）聪明的学生们，得到了 10000 美元的虚拟资金，以及反映了标准普尔 500 指数（S&P 500）的 4 个指数型基金的汇总报告。绝大多数的学生都未能以最低费用来选择基金[19]。由于哪怕是多付 1%的年费，也可能使你多年后的退休基金缩减 1/3 甚至 1/2，因此，这些聪明学生的表现一点也不聪明。

当然，如斯潘塞-布朗的例子所示，如果我们观察足够长的时间，就肯定会碰到某个家伙，他仅仅靠好运也能做出令人称奇的成功预测。相比于包含 $10^{1000007}$ 个随机数字的数学例子，如果我们更喜

欢现实世界中的例子的话，就来看看专栏作家莱昂纳多·科佩特（Leonard Koppett）吧[20]。1978年，科佩特公开了一个系统，据他本人声称，在每年1月底，该系统能确定当年股市的涨跌。这个系统对过去11年的股市都做出了正确的预测[21]。当然，在事后对股市涨跌进行“预测”，对于股票选择系统而言是很容易的；真正的考验在于，它是否能预测出股市未来的发展。科佩特的系统同样经受住了后一个考验：按道琼斯工业平均指数来衡量，该系统从1979年到1989年，连续11年都预测正确；1990年预测失败，但之后直到1998年，又连续给出了正确的答案。虽然科佩特在19年预测中命中了18次，我却很有把握地断定，这个稳定表现根本与技能无关，不管这技能到底是什么。为什么呢？因为科佩特是《体育新闻》（*Sporting News*）的专栏作家，而他的系统则是根据超级碗（Super Bowl）（职业橄榄球的冠亚军决赛）的比赛结果来进行预测的。如果在比赛中获胜的是（前）全国橄榄球联盟（National Football League）的球队，就预测股市上涨；而如果获胜的是（前）全美橄榄球联盟（American Football League）的球队，那就预测股市下跌。知道这一点的人中，大概没几个会辩解说科佩特的成绩不是靠运气得来的。但如果他换个身份——而且也不公开那个方法——那人们大概就会欢呼他为自查尔斯·道（Charles H. Dow）以来最聪明的分析师。

相对于科佩特的故事，现在来看一个确有着恰当身份的名叫比尔·米勒（Bill Miller）的家伙的故事。与科佩特19次中猜中18次的成功不同，多年中，米勒一直保持着连胜，这个连胜完全可以与乔·迪马吉奥（Joe DiMaggio）的连续56场比赛都有成功的击球，或肯·杰宁斯（Ken Jennings）在智力问答节目《危险!》（*Jeopardy!*）中连续74次获胜的纪录相比拟。但至少从一个角度而言，这个比较并非十分贴切：米勒的连胜在一年中所赚的钱，比后两位先生一辈子赚到的钱还要多。因为米勒是美盛价值信托基金

(Legg Mason Value Trust Fund) 的唯一经理，而在其 15 个连续成功的年头中，基金表现每年都好于标准普尔 500 指数。

由于这个成就，米勒被《金钱》(*Money*) 杂志誉为“1990 年代最伟大的资金经理”，而晨星 (Morningstar) 则称之为“年代基金经理”，*SmartMoney* 杂志也在 2001 年、2003 年、2004 年、2005 年和 2006 年将其选为对投资最有影响力的 30 人之一[22]。在米勒连胜的第 14 年时，CNNMoney 网站曾引用某分析师的话，称纯靠运气而获得连续 14 年成功的机会是 1/372529 (之后我们会对此再多谈一些)[23]。

学术界把这种将随机重复归因于超优异表现的错误认识称为“热手谬误”。热手谬误的大多数研究工作开展于体育运动领域，因为在体育运动中可以很容易地定义和衡量表现；同时，体育比赛的规则清晰而确定，数据充分而公开，感兴趣的情况也能一再复现，更不用说这个课题还另给了学者们一个在参与比赛的同时还能假装是在工作的方法。

对热手谬误的兴趣产生于 1985 年前后，更具体点，它出现在特沃斯基与同事的一篇《认知心理学》(*Cognitive Psychology*) 期刊论文中[24]。在名为《篮球运动中的热手：论对随机序列的错误感知》的文章中，特沃斯基和同事们研究了大量的篮球比赛统计数据。当然，球员天赋各不相同：有的投篮命中率为 50%，有的更多些，有的更少些；每名球员也都有过一段时间中手热或手冷的经历。文章作者的问题是，假设每一次投篮的中与不中是随机决定的话，此时应观察到的手热期或手冷期的次数及长度，与实际的次数及长度相比是怎样的？也就是说，如果球员们并不是在投篮，而是在扔一枚能够反映其命中率的不完美硬币时，所得的结果是怎样的？研究者们发现，尽管存在着连续命中或连续投失的情况，费城 76 人 (Philadelphia 76ers) 的投篮，波士顿凯尔特人 (Boston Celtics) 的罚篮，以及受控实验中康奈尔大学 (Cornell University) 男子和女

子篮球队的投篮，都未显示出任何存在非随机行为的证据。

对运动员表现“稳定性”的一个具体而直接的衡量标准，就是在上一次投篮成功（即投中一球）的前提下，本次投篮也成功的条件概率。对于发挥稳定的球员，这个条件概率应该比他/她的总成功率要高。但论文作者们发现，对于每名球员，紧跟在一次成功之后的成功，与紧跟在一次失败（即投篮不中）之后的成功，两者的可能性相等。

特沃斯基的论文发表几年后，诺贝尔奖得主、物理学家普塞尔决定也来研究一下棒球运动中稳定表现的本质[25]。如第一章中所提到的，他的发现——用他的哈佛同事古尔德的话来说　就是除了迪马吉奥连续 56 场都能击中球这个例外，“在棒球运动中，从没发生过什么比扔硬币模型预测所得频率更常发生的事情”。甚至棒球大联盟（Major League Baseball）的巴尔的摩金莺队在 1988 年赛季开始遭遇的 21 场连败，也在预测的频率之内。相比于好的球员和球队，糟糕的球员和球队会出现持续更久的连续失败，而且出现连败的频率也更高些；而相比于差一些的对手，好的球员和球队会出现更长的连胜，同时出现连胜的频率也更高些。但这是由于他们平均的败率或胜率要高一些。该平均比率越高，随机产生的连败或连胜就越长且越频繁。我们只需要理解扔硬币是怎么回事，就可以理解上面的情况了。

那么米勒的连续成功又是怎么回事呢？如果知道其他几个统计数字，就不会那样震惊于米勒的稳定表现是某个随机过程产物这一事实了。比如，在 2004 年，米勒的基金赢利不过是 12%不到一点，而标准普尔中的股票平均赢利为 15%多[26]。这么听起来，这一年标准普尔应该击败了米勒，但实际上，2004 年被归入了米勒“获胜”的一栏，因为标准普尔 500 指数不是股票价格的简单平均，而是按各公司资本额对其股票加权后的均值。米勒的基金比标准普尔股票的简单平均要差，但却比加权平均要好。事实上，在其连续成功那

些年的30来个“12个月时间段”中，他做得还不如加权平均好，但这些时间段都并非正好是整日历年，而他的连续成功是按1月1日到12月31日的日历年来计算的[27]。因此从某种意义上来说，他的连续成功从一开始就有些虚假的味道，因为这个成功的定义方式正好对他奏效了。

但如何才能将这些新发现，与不利于米勒的那个1/372529的概率协调起来呢？在讨论米勒2003年的表现时，为《融会贯通的观察家》（*The Consilient Observer*）时事通讯［由瑞士信贷第一波士顿公司（Credit Suisse-First Boston）出版］撰稿的作家们称，“在过去40年中，没有其他任何一支基金的表现，能连续12年优于市场”。他们提出了一个问题：纯靠运气做到这一点的概率是多少？接着，他们给出了3个估计值（这一年是2003年，而他们考虑的，是一支基金在区区12个年头中连续击败市场的概率）：1/4096，1/477000，以及22亿分之1[28]。借用爱因斯坦的话说，如果他们估计正确，那么只需要一个值就够了。真正的概率是多少呢？大致等于3/4，或说75%。这个差别实在有些悬殊，因此我最好来解释一下。

得出那些极低概率值的人，在某个意义下是对的：如果你正好在1991年开始时正好挑出了米勒这么个人，而完全出于偶然，这个你挑选的人不多也不少地在下面15年中击败了市场，那这个可能性的分母就的确是个天文数字。这一概率也是在15年中每年扔一次硬币并次次都扔出正面的概率。但正如对马立斯的本垒打分析中那样，这并不是真正与问题有关的概率，因为除了米勒之外，还有着数以千计的基金经理（目前有6000多名），同时，还有许多个15年的时间段可以用来完成这一丰功伟业。因此，真正相关的问题是，如果几千人每年都扔一次硬币，并且一直这样做了好几十年，那么其中的某一人，在某个15年时间段中，能够全部扔出正面的机会是多大？这个值就远大于简单地连续扔出15次正面的机会。

为了使解释更具体些，假设有1000名基金经理——这肯定是个

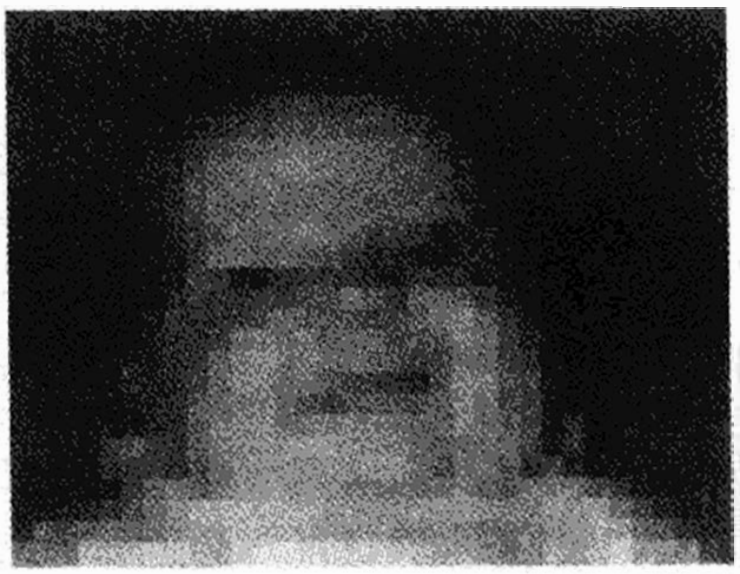
低估的数字——从 1991 年（米勒的连胜开始的那一年）开始，每年都扔一次硬币。第一年，这些人中的大约一半会扔出正面；过两年后，大概有 1/4 的人会扔出两次正面；过 3 年后，大概有 1/8 的人扔出 3 次正面；依此类推。那些扔出反面的人逐渐退出了游戏。这并不会影响我们的分析，因为他们已经失败了。在 15 年后，某个特定的扔硬币者一直扔出正面的概率为 1/32768。但是，从 1991 年开始扔硬币的这 1000 人中的某一个能一直扔出正面的机会则大得多，大约为 3%。最后，我们没有理由只考虑那些从 1991 年开始扔硬币的人——这些经理们可以从 1990 年或 1970 年或现代基金时代中的任意一年开始扔硬币。既然《融会贯通的观察家》的作者们用的是 40 年这个值，我就来算一算，在过去 40 年中，某个经理能在某个 15 年时间段内，每年都击败市场的概率。如此一来，机会就增加到了之前所给的值，即差不多 3/4。因此，与其说我对米勒的连胜感到惊讶，倒不如说如果没有人做到这一点，我们就可以合理地抱怨所有这些拿着高薪的经理们，他们的表现实在比瞎碰运气还要来得差啊！

至此我引用了一些体育和金融领域中热手谬误的例子。但是，在生活中的方方面面，都能碰到连续的成功或失败，或其他特定的成功与失败模式。有时成功是主流，有时则是失败。不管是哪种方式，生活中重要的是将眼光放长远，并能理解不管是连续的成功或失败还是其他看似不像随机序列的模式，都确实可能是纯运气的结果。同样重要的是，当评价他人时，我们应该认识到，在一大群人中，如果没有人经历过长时期的连续成功或失败，这才真是很奇怪的一件事。

没有人会相信科佩特的成功，也没有人会相信一个扔硬币的家伙。但的确有许多人相信米勒。在这个个案中，尽管我所作的这类分析似乎被那些为媒体所引用的许多观察家们所遗忘，但对于从学术角度来研究华尔街的人来说，却根本不是什么新闻。比如，诺贝

尔奖获得者、经济学家默顿·米勒（Merton Miller）（跟比尔·米勒没啥关系）写道："如果有 10000 个人来观察股票市场，并试图挑出最赚钱的那些股票，那么，这 10000 个人中总有一个能成功，而他所靠的仅仅只是运气而已。这就是所发生的一切。这是场游戏，是个机遇性操作；人们以为他们在有目的地做着些什么，但实际却并非如此。"[29]每个人都必须根据所处的环境来得出自己的结论，但了解了随机性的运作方式，这个结论至少不会太天真。

前面讨论了我们是怎样被那些在一段时间中产生的随机序列里的模式所愚弄的。但空间的随机模式同样具有误导性。科学家们知道，要揭示数据的含义，最清晰的方式之一，就是通过某种图像或图形来显示它们。当观察以此种方式呈现的数据时，那些很可能被忽略的有意义的关系会变得更明显。但这个做法的代价是，有时我们也会感知到一些实际没有意义的模式。我们的大脑就是被创造得如此行事的——吸收数据，填充空隙，然后寻找模式。举例而言，看看下面这幅图中灰色方块的排列。



照片引自 Frank H. Durgin, "The Tinkerbell Effect," Journal of Consciousness Studies, 9, nos. 5～6, May to June 2002.

这张图看上去并不像一个名副其实的人，但我们可以从这个模式中得到足够的信息，如果看到画像中那名婴儿的本人，我们很可能会认出他来。如果你拿着这本书并伸直手臂，然后斜向看过去的话，甚至可能察觉不到图像中的不完美之处。现在，让我们看看下面的 X 和 O 组成的模式：

OOOOXXXXOOOXXXOOOOXXOOXOOOXXXOOXXOOOXXXX
OOOXOOXOXOOOOOXOOXOOOOOXXOOXXXOXXOXOXXXX
OOOXXOOXXOXOOXXXOOXOOXOXOXXXOXOOOXOXOOOOX
XXXOOOXXOOXOXXOOOXOOOXXOXOOXXOOOOXOOXXXX
OOOOXXXOOOXOOOXXXXXXOOXXXOOXOOXOOOOOXXXX

在其中可以看到矩形的聚集体，特别是在角落上。我将它们加粗表示。如果这些 X 和 O 代表着我们感兴趣的事件，那我们可能会想，这些聚集是不是意味着什么。但给它们赋予任何意义都将是一种误读，因为这些数据与早先那 200 个随机的 X 和 O 的序列完全相同，只不过现在按 5 行 40 列来排列，并选出了某些部分来加粗显示而已。

这个问题在二战即将结束时引起了许多关注。当时，德军的 V2 火箭开始倾泻到伦敦城里。这些火箭非常可怕，其飞行速度是音速的 5 倍，因此当人们听到它们飞过来时，火箭其实早已落下。很快，报纸就公开了火箭落点分布图，其中似乎可以看到并非随机而是有目的性的成簇的落点。对于某些观察家而言，这一簇簇落点体现了火箭飞行轨迹的控制精度。由于火箭飞行距离很长，因此，这些模式似乎表明，德国人的科技水平已经达到了一个大家做梦都想不到的先进程度。公众怀疑那些躲过了 V2 攻击的区域，就是德国间谍的所在之处。军方高层则担心德国人将关键的军事地点作为目标，从而造成毁灭性的后果。

1946年，上述轰炸数据的数学分析结果发表于《精算学会杂志》（*Journal of the Institute of Actuaries*）。作者克拉克（R. D. Clarke）将所研究的区域分为边长0.5千米的576个正方形小格子。这些格子中的229个从未遭到轰击，而另一方面，尽管格子的面积很小，却有8个格子遭到了4到5次轰击。不过，克拉克的分析表明，跟前面那个扔硬币的数据一样，这些数据的总体分布模式符合随机分布[30]。

类似问题常出现于关于癌症集中发作的报告中。如果将任何城市或国家分成小格，然后将癌症病例随机分布于其中，那么有些格子中的病例数将低于平均值，而有些则高于平均值。实际上，按照加州健康部环境与职业疾病控制分部（Division of Environmental and Occupational Disease Control in California's Department of Health）主管雷蒙德·里查德·纽特拉（Raymond Richard Neutra）的观点，从加州5000个普查区域的常见癌症登记数据——记录几十种不同癌症本地发病率的数据库——中，可以期望发现2750个具有统计显著性但却是随机的某种癌症的高发区[31]。如果对足够多的格子进行观察，就能发现，有些区域的癌症发病率，比正常水平高出了许多倍。

如果将癌症病例分配好再来画格子，那得到的景象会让人感觉更糟。这时的结果称为神枪手效应。这个名字源于一个造假的家伙，他的枪法比别人都要好，因为他先朝一张白纸开枪，然后再在纸上画出靶子。不幸的是，这种事情在实际中经常发生：某些市民首先注意到一些邻居得了癌症，然后他们再来确定问题区域的边界。多亏因特网提供了广泛且能够获得的数据，近日就有人在寻找美国境内的这类团簇。毫不意外地，这些团簇被不断找到。但癌症的发展需要不断地变异，这意味着需要长时间暴露于致癌物质和/或存在着高浓度的致癌物。如果这些高发区团簇确系环境因素所引发，并在受害者从受影响地区搬走之前就登场亮相，那可真是相当走运了。

按纽特拉的说法，要产生这一类型的、经常害得流行病学家被叫来进行调查的癌症发病团簇，人群将不得不暴露于高浓度的致癌物质中，而这样的高浓度通常仅见于接受化疗的病人体内，或是某些特定的工作环境中——远比人们在受污染的住所周围或学校中所遇到的浓度要高得多。但人们仍然不愿接受这些团簇是因为随机波动而成的解释，因此，州政府的健康部门每年都会收到数以千计的居民点癌症集中发作的报告，而结果则是成百的巨细靡遗的分析报告。但没有一个报告能令人信服地给出隐藏的致癌环境因素。明尼苏达州健康部流行病学家阿兰·本德尔（Alan Bender）说：这些研究“绝对、完全、彻底地是对纳税人金钱的一种浪费”[32]。

我们至此考察了若干随机模式愚弄人们的方式。但心理学家并不满足于仅仅对这些误解进行研究与分类，他们还研究了人们之所以会受到愚弄的原因。现在，我们把注意力转向这些原因中的某一些。

人们喜欢对周围的环境施加控制。有些人灌了半瓶苏格兰威士忌后还照样开车，却在飞机遇上一点小小颠簸时就吓得举止失常，原因就在于此。这种控制欲并非毫无缘由，因为自我控制的感觉，已经与自我感知和自尊融为一体。实际上，我们能做的对自己最有好处的事情，就是找到种种方式来控制自己的生活——或至少找到方式来让我们觉得的确控制了自己的生活。心理学家布鲁诺·贝特尔海姆（Bruno Bettelheim）就观察到，要在纳粹集中营中幸存下来，“所依靠的是这样一种能力，即无论周围环境看来如何具有压迫性，都能进行打算和安排，以保留一些独立行动的区域，并对生活中的某些重要方面保持控制。”[33]后来的研究表明，无由来的无助感以及控制力的缺乏，与情绪紧张和疾病发生这两者都有关系。在一项研究中，野鼠被突然剥夺了所有对环境的控制权。它们很快便不再奋力求生，并就此死亡[34]。另一项研究中，一群被试者被告知要

进行一组重要的测试，而哪怕是根本毫无目的地控制这些测试之进行顺序的能力，也降低了被试者的焦虑感[35]。

控制心理学先驱之一的心理学家兼业余画家爱伦·朗格（Ellen Langer）是哈佛大学的一名教授。多年前还在耶鲁时，她与人合作，对控制感在看护住家病人的老年人身上造成的影响进行了研究[36]。研究者告诉一组看护，他们可以决定房间的布置，还可以挑选一株植物来照料；另一组的房间则已经安排好，植物也已经选好并有人打点。在几周后，根据事先定义好的快乐度打分时，能对环境施加控制的那一组得分更高。令人不安的是，18 个月后，跟进研究令研究者们深为震惊：没有控制权的那一组老年人，死亡率高达 30%，而有控制权的那一组，其死亡率才不过 15%[37]。

人类对控制权的需要，与讨论随机模式有何关系呢？因为如果事情是随机的，那我们就没有做主；而如果我们做主的话，事情就不会是随机的。因此，在自觉掌握控制权的需要与认识随机性的能力之间，存在着根本的冲突，同时也是错误解释随机事件的主要原因之一。实际上，诱使人们将幸运误认为是能力，或将无目的的行动误认为是在进行控制，这是心理学研究者最容易做到的事情之一。如果要人们按一个无用的按钮来控制灯的闪烁，他们就会相信他们的确做到了这一点，哪怕这些灯的闪烁是完全随机的[38]。给人们看一圈随机闪烁着的灯，并告诉他们，如果集中精神就能使闪烁沿顺时针方向发生，接着，被试者就会被自己竟然真能够做到这一点而惊得目瞪口呆。或者可以让两个小组在类似的情况下相互竞争——其中一组努力让闪烁按顺时针进行，而另一组则要让它按逆时针发生——而这两组人同时都觉得闪烁方向与各自所希望的方向一致[39]。

朗格一次又一次地证明，这种对掌握控制权之感觉的需要，是如何影响了对随机事件的准确认识的。在她的一项研究中可以发现，相对于一个充满自信的对手，参与者在与一个神经兮兮、笨手笨脚

的对手竞争时，对成功更有信心，尽管他们所参加的纸牌游戏以及获胜的可能，都完全由运气所决定[40]。在另一项研究中，她要求一组聪明而且受过良好教育的耶鲁本科生，对 30 次扔硬币的随机结果加以预测[41]。实验设计者们秘密操纵扔硬币的结果，使每个学生都正好猜对一半，并通过安排使某些学生在开始时连续猜中结果。扔硬币结束后，研究者让学生们做了个小测验，以了解他们如何评价自己猜测的能力。许多学生的答案显示，他们似乎认为猜测扔硬币的结果也是种可以培养锻炼的技能。1/4 的学生称，他们的表现被分心的事情所影响。40%的学生觉得其表现能通过训练得到提高。当被直接要求对他们猜测扔硬币结果的能力打分时，开始时连续猜中的学生给自己的打分要更高些，尽管对所有被试者而言，他们猜对的次数都是一样的。

在另一个聪明的实验中，朗格设计了一个彩票抽奖游戏，参与的志愿者每人都得到了一张可交易的卡片，其上印着一个游戏者的相片[42]。袋子里放着另一张卡片，并且与分发下去的某张卡片一模一样，而那张卡片的持有者就是赢家。游戏者们被分为两组，其中一组可以自行挑选卡片，而另一组的卡片则是随机分发的。开奖前，每个参与者都可以出售其卡片。显然，不管参与者的卡片是自行挑选的，还是随机分发的，其获胜几率都不会受到影响。但相比于随机分配得到卡片的参与者，那些自行挑选卡片的人，其卡片的售价是前者的 4 倍。

朗格实验中的被试者都“知道”——至少理智上知道——他们所参与的游戏是随机的。例如当被问及时，没有一个参与者称自己相信自行挑选卡片会改变获胜的概率。但他们的行为却好像他们确信会有所改变一样。或者如朗格所写道的：“人们可能口头上会认同偶然性的概念，但他们却按偶然事件能被控制一样来行事。”[43]

随机性在真实生活中的角色，远不像在朗格的实验中那样明显，而我们对结果以及控制结果的能力则更为投入。因此，在真实生活

中抗拒这种控制的错觉就越显困难。

在经历了一段时期的成功或失败的组织机构中，常能见到这种错觉的一个表现：成功或失败很容易被归因于头头的身上，而不是归因于运气以及造成机构目前总体状况的无数环境因素。在体育运动中，这一点尤为明显。如前言中提到的，如果球员们一两年内的表现很糟糕，那么被炒的将是教练。大公司的业务庞杂，而且很大程度上受到不可预料的市场力量的影响，这种情况下，高层管理者的聪明才智与公司的业绩之间的关系，就更不那么直接，因此，条件反射式地解雇高管，其效果一点也不比解雇教练来得好。例如，哥伦比亚大学（Columbia University）和哈佛的研究者们最近就对许多公司进行了研究。这些公司的规章制度使它们很容易受到股东的影响，因此，困难时期的应对之策往往就是更换管理层[44]。研究者们发现，在解雇了管理人员后的 3 年中，平均而言，公司的业绩（以赢利计算）并无改进。不管 CEO 们的个人能力上有何差别，这些能力总是被系统中不可控因素的影响所压倒，正如音乐家的演奏水平间的差别，在充满噪声的广播中会变得不明显一样。但在决定补救措施的时候，公司董事会却常常表现得好像 CEO 是唯一紧要的人似的。

研究还表明，如果在随机型任务的结果最终揭晓之前，还进行了一段时间的战略安排（那开不到头的会议），或当任务需要主动参与时（那些呆在办公室里的漫长时间），或当竞争存在时（从来就没有竞争这回事，不是么?），对于偶然事件的可控错觉，在金融、体育以及——特别是——商业环境中进一步强化。与可控错觉做斗争的第一步，就是要认识到它的存在。但即使做到了这一点，前路仍然艰苦，因为正如接下来将看到的，一旦我们认为我们发现了一个模式，这个念头就不会被轻易放弃。

假设我制订了一条规则，用于构造由 3 个数字组成的一个序列，而且 {2，4，6} 这个序列符合规则，那你是否能猜出这条规则是什

么？单单一个由 3 个数字组成的集合，并不足以让游戏继续下去。因此，设想现在由你来告诉我一些其他的 3 数字序列，而我则告诉你它们是否满足规则。请稍微停顿一下，想一些这样的数字序列出来　　相比于和人的互动，阅读的好处就在于，作者可以在书中表现无限的耐心。

现在你已经考虑好了你的策略，而如果你跟大多数人一样的话，我敢说你设想的差不多就是｛4，6，8｝或｛8，10，12｝或｛20，24，30｝这一类的序列。是的，这些序列都满足我的规则。那么这个规则到底是什么呢？大多数人在列举了若干类似的测试例子后，会变得越来越有信心，并得出结论认为，规则就是序列由递增偶数所构成。实际上，我的规则不过是要求序列必须由递增的数字构成。例如序列｛1，2，3｝也满足规则。这些数字并不必为偶数。你设想的序列是否能揭示这一点呢？

一旦被错觉所掌控　　或就本例而言，当有了个新想法时——我们通常不是去寻找方法来证明这个想法错误，而是试图去证明它正确。心理学家称之为“确认偏误”。对于避免对随机性进行错误解释而言，它正是一个主要障碍。在上例中，大多数人立刻认识到，我给出的序列由递增的偶数构成。然后为了证实猜测，他们测试了许多这一类型的序列。但很少有人能更快捷地找到答案——通过测试包含奇数的序列来证伪他们的猜想[45]。如哲学家弗朗西斯·培根(Francis Bacon)① 于 1620 年所写的：“人类的理解力一旦采纳了某个想法，就会收集任何能证实该想法的例子，哪怕反例可能数量更多而且更有分量，但人类的理解力要么不去注意它们，要么就干脆拒绝接受它们，以此保证所采纳的观点仍能维持那不可动摇的地位。”[46]

更糟的是，我们不仅有偏向性地寻找证实预设观念的证据，而

① 培根（1561～1626）是英国哲学家、思想家、作家和科学家。

且还会对有歧义的证据进行偏向于我们想法的解释。这是个大问题，因为数据常常都具有歧义。因此，通过忽略某些模式并强调其他一些模式，聪明的大脑可以强化它的信念，即使在缺乏确有说服力的数据时也是如此。比方说，如果我们根据脆弱的证据认为新邻居是个不友好的家伙，那么他今后的任何可以用“不友好”来加以解释的行为，就会留在我们的脑海中，而其他的行为却被很容易地遗忘。或者如果我们相信某个政治家，那么当她做得好时，功劳当然是她的，而当她做得不好时，该责怪的就是大环境或她的反对派，而不管怎样，都只是在强化我们开始时的想法。

一项研究生动地展现了确认偏误的效果。研究者们召集了一群本科生，其中某些人支持死刑，另一些人则反对[47]。然后，研究者给所有学生提供了一组相同的、有关死刑实际效果的学术研究成果。其中的一半成果支持死刑具有威慑力的观点，而另一半成果则正好相反。研究者们还给予被试者若干线索，表明这些研究都存在若干弱点。然后，研究者要求这些学生独自给各项研究的质量评分，并说明他们对死刑的态度是否以及在多大程度上受到这些研究的影响。参与者给支持其初始观点的研究都打了更高分，哪怕两方面的研究都是有意用同一方法进行的。最后，尽管每个人都阅读了所有这些相同的研究报告，开始时支持和反对死刑的人，都宣称这些材料加强了他们的信念。数据不但没有说服任何人，反而激化了观点的两极分化。因此，即使是随机模式，如果它与我们的预设观念有关的话，也能被解释为具有说服力的证据。

确认偏误在现实世界中造成了许多不幸的后果。如果一名教师开始时相信某个学生比另一个要聪明，那他就会有选择性地将注意力集中在倾向于证实这一猜想的证据上[48]。如果一名雇主对一个符合其预想的应聘者进行面试，那他通常会迅速形成第一印象，然后将剩下的面试时间用来寻找支持这个印象的信息[49]。如果临床咨询顾问被提前告知某受访者生性好斗，那么他们将倾向于得出受访者

确实好斗的结论，哪怕该受访者实际并不比普通人更加好斗[50]。如果人们对少数族裔的行为进行解释的话，他们也常根据一个预设的“典型”来加以解释[51]。

人类的大脑已进化到能高效进行模式识别的水平。但正如确认偏误所显示的，我们集中于发现和确认这些模式，而不是将错误结论降到最少。但无需感到悲观，因为克服偏见仍然是可能的。仅仅认识到偶然事件也能产生模式，是一个开始；如果能学会对我们的观点和理论提出疑问，则又是一大步。最后，我们应该学会用同样多的时间，不但去寻找我们之所以正确的原因，更要去寻找证明我们错误的证据。

我们这段穿越随机性的旅程，现在已经快到尽头了。从简单的规则出发，我们了解了这些规则如何体现于复杂系统中。那在我们自身的命运这个最为重要的复杂系统中，偶然性扮演的角色到底有多重要呢？这是个难题，其中包含了到此为止我们所考虑过的大部分内容。虽然我并不指望能给出完整答案，但却希望能使这个问题变得更为清晰。从下一章的标题，同时也是本书的书名中，可以很明显地看出我的结论：“醉汉的脚步”。

第十章

醉汉的脚步

1814 年，差不多在牛顿物理学成功的巅峰时刻，拉普拉斯这样写道：

“假如一个智能体，在一个给定的时刻，知道了所有使世界运行的力，以及世界的每一个组成体的位置；进一步地，如果该智能体足够强大到对这些数据进行分析，它便可以用同一个方程将宇宙中最大之天体及最小之原子的运动皆囊括其内；对于这个智能体而言，没有什么是不确定的，而未来就如同过去一样，呈现于它的眼前。”[1]

拉普拉斯表达的观点叫“决定论”：世界的当前状态，精确地决定了未来发展的方式。

在日常生活中，决定论隐含着这样一个世界，其中，个人素质及任何给定形势或环境的性质，将直接而毫不含糊地带来精准的后果。这是一个有序的世界，其中任何事情都能被预见、计算和预测得到。但拉普拉斯的梦想要能成真，必须满足几个条件。首先，自然定律必须给定一个确切的未来，而且我们必须知道这些定律；其次，我们必须获得那些完全描述了所感兴趣的系统的数据，而且不允许出现任何不可预见的影响；最后，必须具有足够的智慧或计算能力，以根据给定的描述现在的数据，计算出定律所确定的未来是何等模样。在本书中，我们已经审视了许多有助于理解随机现象的

概念。一路走来，我们已经对生活中的许多具体情况有了深入了解。但我们还需要看到庐山的全貌，即随机性到底在多大程度上造就了我们在生活中所处的位置，而我们又能多好地预测前进的方向。

从文艺复兴晚期到维多利亚时代的这段时期中对世事所进行的研究中，许多学者对决定论都持有与拉普拉斯相同的信念。如同高尔顿一样，他们觉得人的生命历程由个人素质所严格确定，或者像奎特雷一样，他们相信社会的未来可以预计。他们常常受到牛顿物理学成功的启发，相信人类行为能被可靠地预言，就像预言其他自然现象一样。对他们而言似乎再合理不过的是，日常世界的未来就像行星轨道一般，应当能够由当前的事件状态所严格确定。

1960 年代，一位名叫爱德华·洛伦兹（Edward Lorenz）的气象学家，试图利用当时的最新科技——一台原始计算机——在气象研究这个有限的领域内实践拉普拉斯的方案，也就是说，如果洛伦兹将他那理想化地球在某给定时刻的大气状况数据，输入那台噪声不断的计算机的话，计算机就会使用已知的气象学定律，计算和打印出一行行代表未来天气状况的数字。

一天，洛伦兹决定将某个仿真的模拟时间延长些。他打算抄条近路，以运行到一半时的结果作为初始条件来开始新的仿真，而不是重复整个计算过程。为了做到这一点，他用之前打印出的仿真结果作为新仿真的初始数据，并期望计算机能再度生成先前仿真结果的剩余部分，并在此基础上给出对更遥远将来的仿真结果。但和所期待的不同，他注意到一些很奇怪的事情：新仿真中的天气演化，与旧仿真的结果完全不同。新仿真没能重复旧仿真余下的结果，反而与之严重偏离了。很快，他就意识到原因何在：在计算机内存中，数据存储精度是到小数点后 6 位，而在打印结果时，只保留了小数点后 3 位。其后果就是，他用来作为新仿真初始条件的数据，与旧仿真的对应数据之间，存在微小的差别。例如，0.293416 这个值被打印出来时，就变成了 0.293。

科学家通常假设，如果某个系统的初始条件发生很小的变化，那么系统的演化过程也同样只会发生很小的变化。不管怎样，收集天气数据的卫星在测量各种参数时，只能达到小数点后两位或三位的精度，因此，它们不能发觉像 0.293416 和 0.293 之间这样细微的差别。但洛伦兹发现，这微小的差别，却可以导致结果的巨大改变[2]。这种现象称为“蝴蝶效应”（butterfly effect），其名称的来由，是指一只蝴蝶舞动翅膀所造成的大气状况的些微改变，也可能在之后的全球天气模式中造成巨大影响。这个观点听起来十分荒谬——这就等于说某天早上多喝的一杯咖啡，将对生活产生深远的影响。但事实上，这种事的确会发生——比如说，这额外的时间使你在火车站中与未来的妻子偶遇，或者使你刚好避开一辆加速闯红灯的汽车。实际上，洛伦兹的故事本身，就是蝴蝶效应的一个例子：如果他没有做出这个走捷径延长计算的小小决定，他就不会发现将导致一个全新数学领域之诞生的蝴蝶效应。回顾生命中那些重大事件的细节时，并不难发现这类看似无足轻重却导致巨大改变的随机事件。

世事的决定论无法满足那些拉普拉斯暗中所指的可预测性所需要的条件，原因有以下几点。首先，就目前所知，人类社会不像物理学那样，受确定而基本的定律主宰。相反，如卡尼曼和特沃斯基所一再证明的，人们的行为不仅不可预测，而且（从行为违背自身最大利益的意义上来说）还常常是非理性的；其次，即使能够如奎特雷所希望的那样，发现世事之定律，也不可能精确获知或控制生活中的局面。即跟洛伦兹一样，我们无法得到预测所需的精确数据；第三，人的事情是如此复杂，因此，即使确实了解了这些定律，获取了这些数据，能否完成必要的计算还是颇值得怀疑的。这些原因造成的结果就是，对于人类经验而言，决定论是一个很糟糕的模型。

或如诺贝尔奖获得者马克思·玻恩（Max Born）① 所言："相比于因果性，偶然性是一个更为基本的概念。"[3]

醉汉的脚步是随机过程的科学研究中所使用的原型。在生活中，它同样也提供了合适的模型，因为就像悬浮在布朗流体中的花粉粒那样，我们也不断地被随机事件所推动，先往这个方向，然后又朝着那个方向。如此一来，尽管在社会学数据中可以发现统计规律，但对于具体的个人而言，未来却无法预测。对于具体的成就、工作、朋友、经济状况等，我们应归功于机遇的比例，比许多人所认识到的还要大。接着，我将进一步说明，除了现实生活中最简单的情势之外，我们无法躲开那些不可预见或预测的力量，而这些随机力量以及我们相应的反应，造就了我们大部分的具体的生命之路。作为开始，我将探察一个明显与上述观点相矛盾的问题，即：如果未来果真是一团混沌而无法预测，那为什么很多事情，在事后看来却常常似乎应该能够被预见到似的？

1941 年秋，在日本人攻击珍珠港的几个月前，东京的情报机关给在檀香山（Honolulu）的间谍发去了一个令人不安的命令[4]。美军截获了这个命令，并把它送到了海军情报办公室（Office of Naval Intelligence）。在一套官僚作风严重的处理程序后，情报被解密和翻译，并于 10 月 9 日送达华盛顿。消息要求檀香山的日本情报人员将珍珠港分为 5 个区域，并按区域报告在港舰艇，特别是战列舰、驱逐舰与航空母舰以及同一船坞停靠多艘舰船的情况。几周后，发生了另一件古怪的事情：美国监听人员再也无法跟踪到日军第一、第二舰队所有已知航母的无线电通讯，对它们的位置也一无所知。然后，到 12 月初，位于夏威夷的第 14 海军辖区战斗情报单位报告，称日本人在一个月中第二次更换了舰艇呼号。呼号——如 WCBS 或

---

① 玻恩（1882～1970）是德/英双重国籍的物理学家，量子力学的重要奠基人之一。

KNPR——用于标识无线电通讯源，在战时，它们不但能使友方、也能使敌方发现舰只的身份，因此这些呼号被周期性地更换。日军的习惯是每 6 个月或更长时间更换一次呼号，而在 30 天中更换两次呼号，则被视为了“大规模主动行动准备工作中的一个步骤”。这一改变使随后几天中日军航母和潜艇的行踪难以发现，从而让前面那个无线电静默的问题更令人感到困惑。

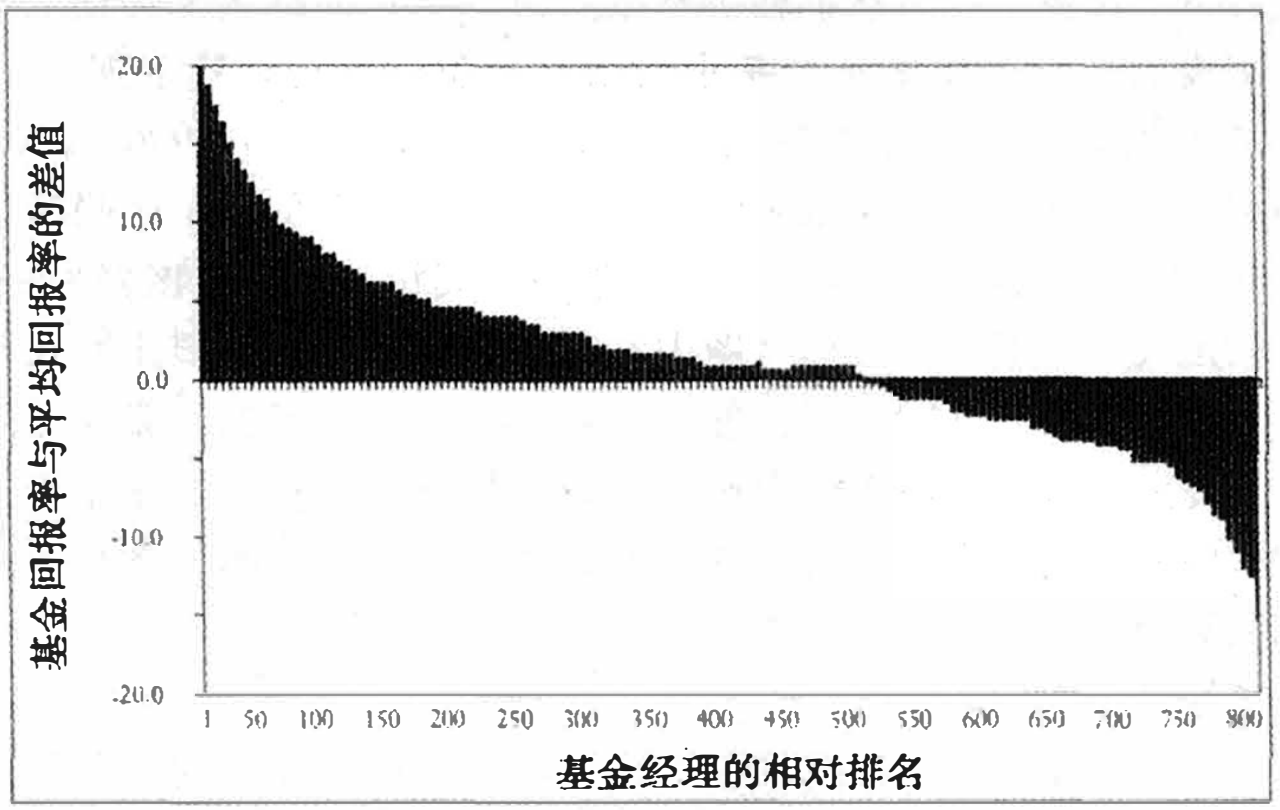
两天后，发送到日本驻香港、新加坡、巴达维亚、马尼拉、华盛顿和伦敦外交与领事机构的消息被一一破解和翻译。这些消息要求外交官员立即销毁密码本，并烧掉其他所有重要的机密与秘密文件。这段时间里，FBI 也截获到一个由日本驻夏威夷领事馆的厨师打给身在檀香山的某人的电话，厨师兴奋地告诉对方，官员们正在烧毁所有重要文件。陆军主要情报单位的助理负责人比克内尔（George W. Bicknell）中校带着一份截获的消息，来到了当时正准备与陆军夏威夷部负责人共进晚餐的顶头上司面前。这时已是 12 月 6 日星期六——即攻击的前一天——下午接近傍晚的时候了。这些事后看来如此有预示性的事情，为何未能让那些与此事利益攸关的人意识到敌军的进攻正在逼近呢？

在一个复杂的事件链条中，每个事件的发展都带有一定的不确定性，这时，在过去与将来之间，就存在着一个根本性的不对称。自从玻尔兹曼对形成流体性质的分子过程进行统计分析（见第八章）以来，这个不对称性就一直是科学研究的对象。设想浮在一杯水中的染料分子。这个分子将如同布朗的花粉微粒那样，踏着醉汉的步伐前进。但即使是这种无目的性的运动，也会在某个方向上产生位移。如果等上比方说 3 个小时，这个分子通常将从最初位置移开约 1 英寸。假设在某个时间点上，分子移动到了某个显眼的位置，并最终吸引了我们的注意，那正如许多人在珍珠港事件后所做的那样，我们可能会去寻找这个出乎意料的事件得以发生的原因。现在假定我们能够深挖这个分子的既往，跟踪其所有的碰撞记录，那就的确

可以发现，这个或那个水分子的碰撞，推动着染料分子沿着它那曲曲折折的道路前进。换句话说，我们在事后可以清楚地解释，为什么这个染料分子的过去是以这种确定的方式发展的。但水中还有许多其他的水分子，它们也有可能与染料分子相互作用。要在实际发生之前预测染料分子的路线，需要计算所有那些具有潜在重要性的水分子的路线与相互作用，而这需要几乎无法想象的巨量数学计算。相比于理解过去所需的那张碰撞清单，前者的计算规模和难度都要高出许多。也就是说，要在实际发生前预测染料分子的运动，几乎是不可能的，哪怕这个运动相对而言在事后更容易理解一些。

这个根本的不对称性，就是日常生活中存在许多其发生看来似乎十分明显但却无法预测的事情的原因。这也就是为什么天气预报员能告诉你，因为 3 天前冷空气前锋这么运动而昨天暖空气前锋那么移动，所以你那浪漫的室外婚礼就碰到了下雨，但在了解这些冷暖空气将如何在之后的 3 天中运动并提醒你为婚礼准备一顶大帐篷的这个问题上，同样的天气预报员做得就远不如前者成功了。或者来看看国际象棋比赛的情况。与纸牌不同，国际象棋中并没有直接的随机因素，但其中同样存在不确定性，因为比赛双方都不能确信地知道，对手的下一步将怎么走。如果双方都是高手，那么在大多数回合中，还有可能预见到接下来的几步棋会怎么走，但要再看远些，不确定性就会混入其中，使得任何人都无法以任何确信度来预言比赛的发展。另一方面，我们通常很容易在事后指出每名棋手下出每步棋的原因。国际象棋同样也是一个过往易懂而未来难测的随机过程。

上述情况对股票市场同样成立。考虑一下基金的业绩表现。如第九章中提到的，选择基金时，我们常常会参考其过去的表现。确实，在过去的表现中能很容易地发现干净漂亮的模式。例如，下图就描绘了 800 支基金在 1991～1995 年的五年间的表现。



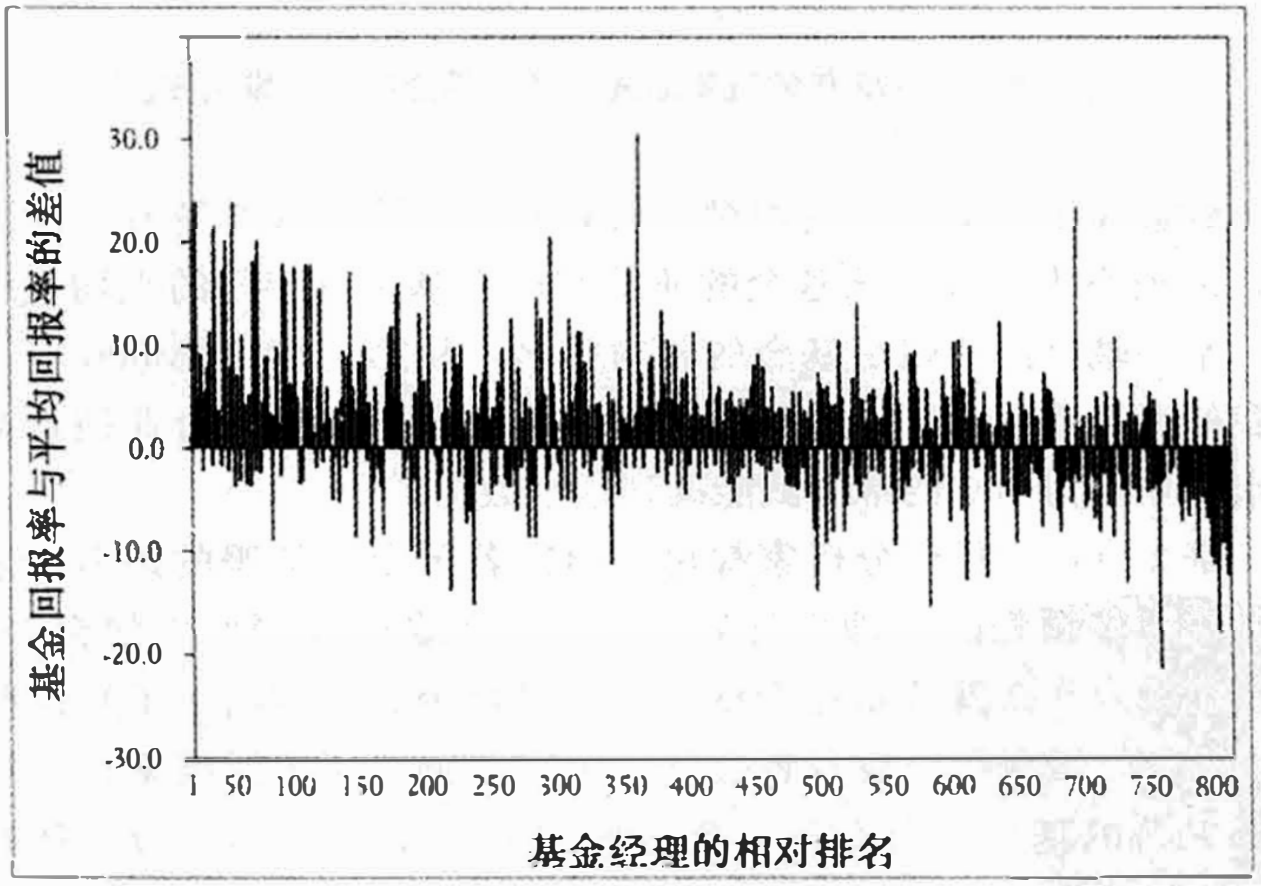
在 1991～1995 年的五年期间，最佳基金的业绩排名图。

纵轴标出的是各基金相对于所有基金平均收入的盈亏，换而言之，回报率为 0 表示该基金的业绩与所有基金五年间的平均业绩持平。水平轴标出的则是基金的相对排名，从第 1 名到第 800 名。如果我们要看看　　比方说——这五年中排名 100 的基金业绩，那么在水平坐标为 100 的地方就能读到这个数据。

毫无疑问，任何分析家都可以给出若干令人信服的理由，来说明图中那些领先的经理们为何会成功，垫底的经理们为何会失败，以及曲线为什么就该是这个样子。不管是否会花时间去了解这些分析的细节，没有哪个投资者会选一个在过去五年中回报率比平均水平低 10%的基金，而不是一个比平均水平高 10%的基金。观察过去，我们很容易构造出这样漂亮的图形和直截了当的解释，但这个充满逻辑性的图景，却不过是马后炮式的错觉，它与预测未来几乎没什么关系。例如，在第 200 页的图中，我将同样的 800 个基金再

次进行了比较，其中水平轴仍然按之前五年中的排名来排列，而纵轴则表示它们在接下来五年中的表现。也就是说，我根据 1991～1995 年的表现对基金进行排名，但显示的却是它们在 1996～2000 年的回报率。如果过去是未来的一个不错的指示器的话，这些基金在 1991～1995 年间以及 1996～2000 年间，其相对表现应该多少一致。也就是说，如果胜利者们（图中左方的那些基金）依旧比别人干得好，而失败者们（图中右方的那些基金）依旧干得差，那么这张图应该与上一张图差不多完全相同。但正如图中所示，将存在于过去的秩序外推到将来时，这个秩序就解体了，而新图形最终看来就如同随机噪声一般。

偶然性在投机以及如股权基金经理米勒这类人的成功中所扮演



1991～1995 年期间的最佳基金在 1996～2000 年间的业绩。

的角色，被人们系统性地遗漏了，而我们却总是毫无根据地相信，过去的错误必定是无知或无能的结果，并可以通过进一步学习以及提高洞察力来加以纠正。这也就是为什么——举例来说——在2007年的春天，当美林证券的股价为每股95美元上下时，其CEO奥尼尔（E. Stanley O'Neal）被认为是冒险的天才，而到了2007年秋天证券市场崩溃后，他就被嘲笑成不计后果的鲁莽牛仔——并很快被解雇。我们条件反射式地对商业显贵、政治家与演员中的超级巨星以及任何乘坐私人飞机飞来飞去的家伙们表示尊敬，好像他们的成就必然反映了其独有的素质，而且这些素质不是那些被迫吃着经济航班提供的食物的人所拥有的。而我们对那些声称以往记录代表了职业水平的人们——政治领域的专家、金融专家、商业顾问——所做的过分精确的预测，也投入了太多的信心。

我熟知的一家大型出版公司，曾在制订其教育软件部门的一年、三年与五年计划时，碰到了极大的困难。公司有高薪聘请的顾问，有冗长的市场调研会议，有熬夜的金融分析讨论，以及长时间的外出召开的午后会议。终于，预感变成了公式，给出的结果精确到小数点后好几位，而疯狂的猜想则成了颇有可能实现的结果。当头一年某些产品卖得不如预期那么好时，或者当另一些产品卖得好过预期时，总能找到之所以好或不好的理由，并归功或归咎于相应人等，好像那些预期真有什么意义似的。第二年，两名竞争对手发起了一系列的未预见到的价格战，再过一年，教育软件市场彻底崩溃了。当不确定因素搅进来时，那个三年计划根本就没有成功的机会。而精雕细琢且精确有如钻石的五年计划，则免掉了与实际结果进行比较的麻烦，因为五年之后，那个部门的差不多所有人都已经跳了槽。

那些职业就是研究过去的历史学家们，对于事件按可预测的方式来发展的观念，抱持着与科学家们同样的警惕态度。实际上，必然性错觉在历史研究中有着严重的后果，这也使得它成为了极少几

件保守主义与社会主义历史学家①们能达成一致的事情之一。例如社会主义历史学家陶尼（Richard Henry Tawney）就说道："历史学家们……通过突出获胜的力量，并将那些被它们所吞噬的力量扔到角落中……来造成一种必然性的表象。"[5]而曾被里根（Ronald Reagan）授予总统自由勋章的历史学家沃尔斯泰特（Roberta Wohlstetter）则这样说道："当然，各种信号在事后看来，总是水晶般地清晰；我们可以看到它究竟预示的是怎样的灾难……但在事前，它却晦涩朦胧，并包含着相互矛盾的含义。"[6]

从某种意义上来说，世事难料的观念从来都被包裹于"事后诸葛亮总是看得最清楚"这样的谚语中，但人们的行事却总是如同这句话根本不对似的。例如，每一个悲剧发生后，"你们早该知道它会发生"这样的问责游戏就会在政府中上演。在珍珠港（以及9·11攻击）的例子中，回过头来看那些攻击前所发生的事情，它们似乎都指向了一个显然的方向。但正如在染料分子、天气与国际象棋的例子中那样，如果在事前跟踪事情的发展，这种必然性的感觉很快就消失无踪了。首先，除我所引用的情报之外，还有着数量巨大的无用情报，每周都有成堆的这类新消息或报告，有时带有警告意味，有时显得颇为神秘，但后来都被证明是误导或根本不重要。即使将注意力集中在事后证明具有重要性的报告上，在攻击发生前，每份这样的报告也都存在着其他的合理解释，而并不能说明针对珍珠港的奇袭正在进行中。例如，那个将珍珠港分为5个区域的要求，与其他发往驻巴拿马、温哥华、旧金山和俄勒冈州波特兰的日本情报人员的要求相类似。至于失去无线电接触也并非前所未闻，这种情况过去也常常只意味着战舰正停泊在本土港口，它们之间的通讯是利用陆上电缆来以电报进行的。更重要的是，即使相信战争的扩大

---

① 这里的"社会主义"并不是我们熟知的马克思主义中所提的社会主义，而是指类似于英国工党那样的"社会民主主义"一类的左翼政治思潮。

迫在眉睫，许多信息却表明攻击会发生在别处——例如菲律宾群岛、马来半岛或关岛。在珍珠港的例子中，使注意力发生偏移的因素，肯定没有染料分子所碰到的水分子多，但已足够使得未来的清晰图景变得模糊不清。

在珍珠港事件后，美国国会的 7 个委员会就投身于钻研之中，以求发现军方之所以忽略了所有那些表明攻击即将到来的“信号”的原因。作为军方一员的陆军参谋长马歇尔（George Marshall），由于 1941 年 5 月致罗斯福（Franklin Roosevelt）总统的一份备忘，而招致了强烈的批评。该备忘称“由于其防御工事、驻军以及自然条件，瓦胡岛据信是世界上最为坚固的堡垒”，并再次使总统确信，如果敌人来袭，敌方将“在距目标 200 英里的范围内……被形形色色的弹雨”所阻截。马歇尔将军不是笨蛋，不过他手上也没有水晶球。随机性研究告诉我们，如水晶球般清澈的事件场景是可能的，但不幸的是，仅仅在事后才如此。因此，我们相信我们知道为什么一部影片表现不错，为什么一名候选人赢得了选举，为什么有一场风暴来袭，为什么一支股票价格下跌，为什么一支足球队输了比赛，为什么一种新产品在市场上大败，或为什么一种疾病开始恶化；但这种技能，对于预测何时一部影片会表现不错、一名候选人会赢得选举、一场风暴会来袭、一支股票价格会下跌、一支足球队会输掉比赛、一种新产品会在市场上遭遇失败或一种疾病会开始恶化等问题而言，则几乎毫无用处。就预测而言，这个技能实际上啥都不是。

我们可以很容易地编出故事来解释过去或者对不确定的未来发展变得充满信心。我们不必因为这些做法中存在陷阱而根本不予采用。不过，我们可以努力使自己对直觉的错误产生免疫力。我们可以学会用怀疑的眼光去看待解释和预言；我们可以更注重对事件做出反应的素质，如灵活性、信心、勇气与坚毅，而非依赖于对事件发生的预测能力；我们可以更看重直接的第一印象，而非看重那些大肆张扬的当年之勇。如此就能避免在机械决定论的框架内形成

判断。

1979 年 3 月，发生了另一连串的事件，其出乎意料之处也因此而出了名。这一次是在宾夕法尼亚的一个核电站中[7]。事情的结果是核反应堆芯部分融化，并可能向环境释放出警戒水平剂量的辐射。灾难是这样开始的：有大约一杯左右的水，从一个称为精处理器的水过滤器，通过密封处漏水的地方渗漏出来，并流人了驱动电厂某些仪器的一个气动系统中，导致了两个阀门故障。这些出故障的阀门切断了电厂蒸汽发电机组的冷水供应——而该蒸汽发电机组正是负责带走堆芯处核反应所产生热量的系统。然后紧急水泵被启动，但它的两路管道上各有一个阀门，在两天前的维修工作后，仍然被设置在关闭状态。因此，水泵只是徒劳地将水打进了死胡同。此外，某个减压阀门未能正常工作，而本应在控制室显示该阀门有问题的仪表也同样发生了故障。

独立看待这些故障中的任何一个时，它们都属于常见也可以接受的情况。精处理器故障在这个电厂里不是什么稀罕事，通常也不是什么大问题；核电站中有成百的阀门有规律地打开和关闭，其中一些被设置在错误的状态，这既不罕见，也不值得太多担心；而那个减压阀的不可靠，则本来就众所周知，而且，至少在 11 个其他的电站中，也不时发生故障，却没有造成严重后果。但这些故障凑到一起后，整个电站看起来就好像是由奇斯通警察（Keystone Kops）① 在打理一样。因此，紧随三里岛（Three Mile Island）上这场事故之后的，就是众多的调查和非难，以及一个完全不同类型的后果。导致事故的那一连串事件，使耶鲁大学的社会学家查尔斯·培罗（Charles Perrow）创建了一个新的意外事故理论，其中就给出了本

① 奇斯通警察是无声电影时代的一个系列喜剧的主角，该剧描述了一帮无能透顶的警察。

章的中心论点：应当预计到，复杂系统中（我将生活也归入其中）那些通常被忽略不计的次要因素，会由于偶然性而在某些时候导致重大事件的发生。[8]

在其理论中，培罗认识到现代系统都是由数以千计的部分所组成，其中包括可能出错的人类决策者。这些组成部分以无法单独跟踪与预测的方式——就如同拉普拉斯的原子一样——相互联系在一起。但可以打赌，正如那踏着醉汉步伐行走的原子最终将到达某个位置一样，事故也将不可避免地发生。培罗这个被称为事故常态理论的学说，描述了这种情况是如何发生的，即在没有明确的原因、没有公司或政府调查组所希望找到的明显错误和无能笨蛋的情况下，事故是如何得以发生的。尽管事故常态理论研究的是事情有时必然出错的原因，但它同样能反过来解释事情有时必然成功的原因。因为在复杂过程中，不管失败了多少次，如果一直去尝试，那么最终成功的机会还是不小。实际上，如 W·布里安·亚瑟（W. Brian Arthur）这样的经济学家就认为，许多次要因素的同时发生，甚至可以使没有特别优势的公司能够压倒竞争对手。“在现实世界中，”他写道，“如果几个规模相近的公司一同进入市场，那么，小小的偶然事件——意外的订单、与购买者偶然的碰面、管理上的突发奇想——将可能决定哪些公司能更早获得销售收入，并通过日积月累，最终成为市场的支配力量。经济活动……由单独的业务（所决定），这些业务如此之小，以至于无法预见。但这些小小的‘随机’事件能（积）累起来，并随着时间的流逝，通过正反馈而得到放大。”[9]

社会学研究者们也注意到了同样的现象。比如，一群研究者对社会学家称之为文化产业——书籍、电影、艺术、音乐——领域的消费者购买习惯进行了研究。这些领域中的传统市场营销思路认为，预测消费者的偏好能带来成功。按照这种观点，管理者们最有成效的使用时间方式，应该是去研究到底是什么东西让斯蒂芬·金（Stephen King）、麦当娜或布鲁斯·威利斯（Bruce Willis）这类人

的粉丝们如此着迷。他们对过去进行了研究，并且——如我所说的——没费什么力气就找出了成功的原因，不论他们希望解释的成功到底是什么东西。然后，他们希望能复制这些成功。

这就是市场的决定论观点，在这种观点下，成功主要取决于某个人或产品的内在素质。但我们还能以另一种方式——一种非决定论的方式——来看待成功。在后一种观点下，世界上还有着许多高质量但不为人知的书籍、歌手、演员，而真正能让这一个或那一个冒出头来的，则很大程度上是随机性与各种次要因素合谋的结果，也就是运气。根据这种观点，那些传统的行政管理人员不过是在玩轮盘赌而已。

多亏了因特网，这个观点已经得到了验证。进行验证的研究者们将注意力放在了销售额主要来自因特网的音乐市场。为进行研究，他们雇用了 14341 名参与者，让他们听一些他们从未听说过的乐队所演唱的 48 首歌曲，对它们打分，并且——如果愿意的话——下载这些歌曲[10]。有些参与者能够看到这些歌曲的流行度数据——即在其他参与者中，有多少人下载了这些歌曲。这些参与者被分到 8 个相互分离的“世界”中，并且只能看到同一世界中的参与者下载歌曲的情况。每个乐队在各个世界中的初始下载量都为 0，之后每个世界就独立地演变。此外还有一个第九组，其中的参与者看不到任何数据。研究者们用这最后一组隔绝的听众们所给出的歌曲流行度，来定义每首歌的“固有质量”——即在没有外界影响的情况下，这些歌曲所具有的吸引力。

如果决定论的世界观是正确的话，在前 8 个世界中占优的，应该都是同一些歌曲，而且这 8 个世界中的歌曲排行，应该与第九组的孤立听众们所给出的固有质量相一致。但研究者们所发现的却恰好相反：各首歌曲在不同世界中的受欢迎程度天差地别，而固有质量相近的不同歌曲，其受欢迎程度也大相径庭。比如，“52metro”乐队所演唱的“Lockdown”，其固有质量在 48 首歌中排行第 26，但

在一个世界中，它的排行为第 1，而在另一个世界中，排行为第 40 名。在这个实验中，如果这首或那首歌曲，偶然地在早期获得了下载量优势，那么这首歌曲颇受欢迎的表象，就会进一步影响后来的购买者。在电影产业中，这个现象也广为人知：电影观众如果事前听说某部电影不错的话，他们通常会对其表示出更多的喜好。在这个例子中，小的随机影响产生了滚雪球效应，使歌曲的未来天差地别。我们又一次看到了蝴蝶效应。

通过生活中细致入微的观察，同样能发现，在许多重大事件中，如果没有那些偶遇的人、偶然到来的工作机会以及各种各样小因素的随机汇合，结果可能会大相径庭。设想这样一名演员，从 1970 年代后期开始的 7 年中，一直住在曼哈顿第 59 大街一幢没电梯的公寓 6 楼，为着有朝一日闯出个名头而努力着。他从事那些没法打人百老汇——有时根本还差得远——的戏剧工作，并出演电视广告节目，总之是用尽一切办法来引人注意，以确立自己的职业生涯，并赚到足够的钱能偶尔到饭店享用一餐烤牛排，而不必在买单之前偷偷摸摸地溜走。但如同许多赶超偶像的人一样，不管这名志向远大的演员如何努力地去选择正确的角色、作出正确的职业选择并在交际中突出自己，他最可靠的角色，却仍然是在另一份工作中所扮演的那一个　　酒吧招待。之后在 1984 年夏季的一天，他飞到了洛杉矶，要么是去观看奥林匹克运动会（如果你相信他的传记作家的话），要么是去拜访一位女性朋友（如果你相信《纽约时报》的话）。不管真实情况是哪一个，有一件事却是非常清楚的：拜访西海岸的这个决定，跟演出没什么关系，跟爱情——至少是对体育的热爱　　倒是颇有瓜葛。但后来发生的事情却证明，这是他所有职业决策中最好的一个，而且还很可能是他这辈子做过的最好决定。

这名演员名叫布鲁斯·威利斯。在洛杉矶时，一名星探建议他去参加一些电视试演[11]，包括一个已进入最后选角阶段的节目的试演。制作方已经有了一个最终入围者的名单，但是，在好莱坞，没

有什么东西是“最终”的，除非合同上的墨迹都干透了，官司也已经了结了。威利斯得到了试演机会，并最终获得了大卫·艾狄生（David Addison）这个角色——ABC的新节目*Moonlighting*中希碧尔·谢菲德（Cybill Shepherd）的搭档男领队。

颇具诱惑力的想法，是认为威利斯显然比本来排名榜首的X先生更适合这个角色，至于接下来的事情就是——如他们所说的——历史了。我们已经知道*Moonlighting*和威利斯都取得了巨大的成功，所以，很难想象好莱坞那帮拍板的家伙在看到威利斯时，除了烧掉笨头笨脑的长靴来庆祝新发现的耀眼新星，并把那张过时的最终入围者名单扔到火里之外，还能做些什么其他的事情。但在确定演员阶段实际发生的事情，则更像是让孩子们出去买一加仑冰激凌，结果有两个想要草莓味，而第三个想要三层巧克力芙奇核桃仁味时的情形。电视网的管理人员们为X先生而战，他们觉得威利斯看起来并不像一个严肃的领队。*Moonlighting*的制作总监格伦·卡伦（Glenn Caron）则支持威利斯。前者在事后很容易被看作无知的小丑。就我的经验而言，电视节目制作人常常都持这种意见，特别是当这些评论不会被搞管理的人听到时就更加如此。但在选择这个想法之前，考虑考虑下面的问题：电视观众在开始时都认同管理人员觉得威利斯普普通通的评价。*Moonlighting*于1985年3月初次亮相时，收视率很低；在整个第一季中，表现也不过平庸而已。只是到了第二季，观众们才改变了想法，然后节目就炙手可热了。在他突然成为一名明星之前，威利斯的吸引力以及成功一直都理所当然地无法预见。人们现在也许会把账记在疯狂的好莱坞头上，但威利斯那通往成功的醉汉之路，却根本不是什么不寻常的事情。许多成功人士所经历的，正是这样一条被许多随机影响及意料之外的结果一段段隔开的道路，不仅仅职业生涯如此，他们的爱情、嗜好及友谊也是如此。实际上，这种情况更像是规律而非例外。

最近有一次观看深夜电视节目时，在一个访谈中出现了另一位

明星，虽然他不是来自于娱乐圈。他的名字叫比尔·盖茨（Bill Gates）。尽管访谈主持人以其嘲讽的风格而著称，但面对着盖茨，他却似乎颇不寻常地恭敬。甚至连下面的听众，似乎也在对盖茨大抛媚眼。之所以如此，当然是因为盖茨已连续 13 年被《福布斯》（*Forbes*）杂志评为全世界最富有的人。实际上，自创立微软公司以来，盖茨每一秒钟都要赚 100 多美元。因此，当被问及对互动电视持何种观点时，每个人都怀着极大的期待等待着他的话语。但他的回答十分寻常，也并不比其他一打计算机专家嘴中说出的回答更富有创意、更独特或更有洞察力。这不禁使我们思忖：到底是因为盖茨像神一样，所以才能每秒钟赚 100 美元呢，还是因为他每秒钟赚 100 美元所以才像神一样？

1980 年，一群从事于某秘密个人计算机制造项目的 IBM 雇员们，为此飞到西雅图来会见这位年轻的计算机企业家。盖茨当时开办了一家小公司，而 IBM 需要为其计划中的“家庭计算机”寻找一个被称为“操作系统”的程序。对后来所发生的事情，各方的回忆有所不同，但大体是这样的[12]：盖茨说他无法提供这个操作系统，并建议 IBM 的人去找一位著名的软件程序员——数码研究公司（DRI）的加里·基尔代尔（Gary Kildall）。IBM 与基尔代尔之间的商谈并不顺利，原因之一，是当 IBM 的人来到 DRI 的办公室时，基尔代尔当时的妻子，公司的业务经理，拒绝与 IBM 签署保密协议。IBM 的使者们再次造访，这一次，基尔代尔与他们见了面。没有人确切知道在这次会面中发生了什么，但即使真做了一笔非正式的交易，交易也没能继续下去。大约在这时，IBM 的一名雇员杰克·山姆士（Jack Sams）又碰到了盖茨。两人当时都知道，还有另一个可用的操作系统，它是——看你问的是哪一个人——根基于基尔代尔的系统，或是在基尔代尔系统的启发之下做出来的。按山姆士的话，当时盖茨说道：“你想自己去弄……（那个操作系统）呢，还是想要我帮你弄到？”山姆士显然不想被牵连进去，因此他说：“不管用什

么方法，你去搞到它。”盖茨做到了，代价是50000美元（或根据某些记述，还要再多一点点），作了一些改动，并给它重新起了个名字叫DOS（磁盘操作系统）。IBM显然对自己新点子的潜力没什么信心，因此它以低廉的软件拷贝使用费从盖茨处获得了许可，而盖茨则能保留其版权。DOS并不比——比方说——苹果公司的Macintosh操作系统更好，而且还有很多人——包括大多数计算机专业人士——会说它比后者差得多。但增长中的IBM用户群，鼓励了软件开发者为DOS编制应用软件，从而又刺激了潜在用户来购买IBM的机器，再反过来又激励了软件开发人员为DOS编制程序。换而言之，正如亚瑟说的那样，人们购买IBM的机器，是因为人们正在购买IBM的机器。在计算机企业界这个流体世界中，盖茨成为了脱颖而出的那个分子。但要不是基尔代尔的不合作，IBM的缺乏远见，以及山姆士与盖茨的再度会面，那不管盖茨有多少先见之明或商业敏感性，他可能都只会成为另一个软件企业家而非世界上最富有的人。他的观点看来不过如同另一个软件企业家的观点而已，而这大概就是原因所在。

我们的社会能很快将富人变成英雄，把穷人变成狗熊。这也就是为什么房地产大亨唐纳德·特朗普（Donald Trump）在他的广场酒店破产而赌场帝国破产两次（如果一个合作伙伴在1994年向其赌博公司投资10000美元的话，13年后，他将只能带着剩下的636美元离开）[13]之后，却仍然敢于在一个颇为成功的电视节目中以明星的面目出现，并对那些雄心壮志的年轻人的商业触觉品头论足。

显然，以与其财富成比例的方式来赋予人们才智，很可能是个错误。我们无法看到某人的潜能，而只能看到其结果，因此，我们常常认为结果必然反映能力，并因而对这个人形成错误的判断。生活的事故常态理论证明，行动和收获之间的联系虽然并非随机，但随机影响却与我们本身的素质及行动同等重要。

许多人情感上都不愿接受随机影响非常重要的这个看法，哪怕

他们从理智上能够理解事实的确就是如此。如果人们低估了机遇在那些巨头的职业生涯中所扮演的角色，那对于机遇在最失败者的生活中所扮演的角色，他们是否同样没有给予足够重视呢？这个问题在 1960 年代促使社会心理学家梅尔文·勒纳（Melvin Lerner）去探察社会对穷人所持的负面态度[14]。勒纳认识到，“如果人们相信在其作为和收获之间存在某种随机联系的话，就没有几个人会去付出长期的努力”[15]，因此他的结论就是，“为了保持他们自己的神经正常”，人们高估了由成功所能推测出的能力[16]，即我们倾向于认为，成名的电影明星比奋斗中的电影明星更有天分，而世界上最富有的人也必然是最聪明的人。

我们也许不觉得自己会根据人们的收入或成功的外在表现来衡量他们，但即使知道一个人的薪水完全由随机所确定，还是有许多人无法避免直觉地认为薪水与价值相关。勒纳考察了这个问题。他将被试者们安排在一个小小的暗室中，面对一面单面透光的镜子坐着[17]。从座位上，这些观察者们能够注视一个放了一张桌子两把椅子且照明良好的小房间。实验者诱导这些观察者们相信，两名工人——汤姆和比尔——将很快进入房间，并一起花 15 分钟来还原若干被打乱的单词。观察窗前的窗帘此时是拉上的，而且勒纳告诉观察者们，窗帘将一直保持关闭，因为如果他们能听到但不能看到这两名工人的话，就不会受外表的影响，从而实验就可以更好地进行。勒纳还告诉他们，由于经费有限，他只能随机挑出一名工人来付给酬劳。勒纳离开房间后，助手就打开开关，开始播放一卷录音带。这些观察者们相信，他们正在听着汤姆和比尔进入窗帘后的那个房间，并开始工作。实际上，他们只是在听汤姆和比尔朗读一个固定剧本的录音，这个剧本经过特别设计，使得从多个客观评判标准来看，他们两人都显得同样擅长于实验中的任务。之后，对此毫不知情的观察者们被要求对汤姆和比尔的勤劳程度、创造性和成功程度

加以评分。当汤姆被随机选为报酬获得者时，大约90%的观察者评判他为任务的完成做出了更多贡献；当比尔拿到钱时，则有大约70%的观察者持同样观点。尽管汤姆和比尔的表现相当，而且观察者们也知道报酬是随机发放的，他们仍然认为得到报酬的工人比那个忙了半天却什么也没得到的人要好。哎，正如所有为体现成功而精心打扮的人所知道的那样，我们都太容易被某人赚到的钱所愚弄了！

一系列相关研究也从这两名工人本身的视角探究了该效应[18]。人人都知道，具有适当的社会与学术资历、拥有响亮头衔及丰厚收入的老板们，常常会更高地评价自己而非部下的观点。研究者们所思索的是，那些纯靠运气而赚到更多钱的人，是不是也是如此？哪怕那些“成功”是不劳而获，是否也慢慢向他们灌输一种高人一等的感觉？为了找到答案，研究者让若干对志愿者在多种无意义的任务中相互配合。例如在一项任务中，实验将短暂地显示一幅黑白图像，而被试者们需要确定，图像的上部和下部中，哪个部分的白色区域比例更大。每项任务开始前，一名被试者被随机挑出，并获得比另一名被试者明显更多的报酬。如果被试者并不知道他们各自获得报酬的情况，他们之间的协作就颇为融洽。可当他们知道了各自的报酬是多少时，报酬更高的那名被试者就对其搭档的意见表现出更强烈的抵触态度。即使是报酬上的随机差别，也导致了以报酬推断技能的反向推理，并因此形成了影响力上的不平等。这是个人及办公室动力学中不能忽视的部分。

不过更接近于勒纳最初的动机的，是上述问题的反面。勒纳与同事一起问的问题是，人们是否倾向于认为，那些不成功或正在承受痛苦的人，确实是活该如此[19]。研究中，若干组女大学生汇集在接待室中。几分钟后，其中一人被挑中并离开房间。这名离开的学生——下面称之为受害者——实际上并非被试者，而是实验者们在房间中所埋伏下的工作人员。剩下的被试者们被告知，她们将观察

受害者，看她进行一项学习任务，每次她做出错误回答，就会遭到一次电击。实验者调整了一些据称可以控制电击强度的旋钮，然后打开电视监视器。被试者们就这样看着受害者走进了隔壁的房间，被绑上了一个“电击装置”，并开始学习一对对毫无意义的音节。

在学习期间，受害人由于回答错误而遭到了几次看来颇为痛苦的电击，并且哀呼连连。实际上受害人不过是在演戏，而监视器所播放的，也只是一盘预先录制好的录像带。不出所料，大多数的观察者开始时都表示，她们对同伴所遭受的不公正的磨难感到不快。但随着实验继续进行，她们对受害者的同情被慢慢侵蚀。最终，这些无力提供帮助的观察者们开始将不公平的批评加诸于受害者。受害者遭受的痛苦越大，观察者们对她的观感就越差。正如勒纳所预计的，观察者们需要根据因和果来理解所面对的局面。为了确保不是别的什么因素才造成这个实验结果，实验者们又在其他若干组被试者上重复了实验，并告诉这些被试者，受害人将会因其所受的苦难而得到补偿；换句话说，除了相信受害者得到了“公平”对待这一点外，被试者所经历的实验其余部分都与前面完全相同。这一次，观察者们没有形成恶评受害人的倾向。很不幸地，对那些来自社会底层的人们持有偏见，似乎是一种下意识的行为。

我们看漏了随机性在生活中的影响，因为在评价这个世界时，我们倾向于看到自己所希望看到的东西。我们实际是通过成功的程度来定义才能的高低，然后再通过才能与成功之间的相关性来进一步强化这种因果关系。尽管一个非常成功的人与一个不那么成功的人，他们的能力有时基本没有差别，但我们看待他们的方式通常却大不相同，其原因就在于此。在 *Moonlighting* 之前，如果年轻的酒吧招待威利斯跟你说他希望成为电影明星，你肯定不会想：噢！有机会与这样一位魅力超凡的未来名流单独聊聊天，我可真是撞了大运了！相反地，你更可能产生类似如下的想法：嗯，好吧，现在苦艾酒别喝过头就好。但节目大获成功后的那一天，突然间每个人都

把威利斯视为一个明星，一个具有某种特殊能力来抓住观众的心灵与想象的家伙。

期望所具有的这种威力，在数年前由心理学家罗森汉（David L. Rosenhan）进行的一项大胆的实验中，得到了戏剧性的展示[20]。实验中的 8 名“伪病人”被分别指派到多家医院，并在医院接诊处抱怨称他们听到了奇怪的声音。这些伪病人的职业多种多样：3 名心理学家，1 名精神病学家，1 名儿科医生，1 名学生，1 名画家，以及 1 位家庭主妇。除了所声称的症状以及虚假的名字和职业之外，他们都完全如实地描述了平时的生活。有些被试者后来表示，他们十分信任心理健康体系那钟表般精准的工作，因此他们曾担心，医院会立刻察觉他们显然清醒而正常的头脑，这可是很令人难堪的。但他们完全多虑了。除一名被试者外，其余所有人都被许可进行精神分裂症的诊断；而即使是这一名例外，也得到了躁狂抑郁症的诊断许可。

得到接诊后，被试者们不再虚拟任何不正常的症状，并表示那些怪声都已消失。根据罗森汉预先所给的指示，他们接下来就等着医院的工作人员来发现，他们其实并没有发疯。但无人注意到这一点。相反地，医院工作人员透过精神病的眼镜，来解释这些伪病人的行为。当看到某位病人写日记时，他的护理记录中就加上了一条“病人忙于书写行为”，并将书写视为精神疾病的一个标志。而当另一名病人因遭到服务人员的虐待而爆发时，这个行为也同样被认为是其疾病的一部分。甚至连在自助餐厅开放之前就到那里等着吃中饭这样的行为，也被看成是精神病的症状。对正式医疗诊断一无所知的其他病人，常常用“你没有发疯。你是个记者……你在对这家医院进行暗访”这样的话语来挑衅参加实验的伪病人们；但伪病人的医生们所做的记录则类似于“该名 39 岁白人男性……表现出了长期以来在密切关系中明显存在的正反感情并存的情况，这种情况自其童年早期就已经开始：与其母亲的温暖的关系在成年时逐渐冷淡；

而与其父亲疏远的关系则被描述为非常热切。”

好消息是，尽管有着诸如书写和提前等吃中饭等可疑习惯，这些伪病人还是被判定为对自身及他人不具威胁性，从而平均在医院呆了 19 天后，他们都被放了出来。这些医院完全没有察觉到进行中的诡计，而且在事后被告知真实情况后，都否认这种事情有发生的可能。

我们很容易沦为期望的受害者，却也同样容易去利用这一点。这也就是为什么那些在好莱坞拼搏的人们，却要费尽力气使自己看起来并不在奋斗；为什么医生们都穿着白大褂并在墙上挂满各种各样的执照和学位证书；为什么二手车经销商宁可去修理车辆表面的污迹，也不愿把钱花在引擎上；为什么通常老师批改作业时，对完全相同的作业，给“优秀”学生的分数要比“差”学生的分数更高一些[21]。搞市场的同样知道这一点，并通过广告宣传设计，来诱导并进而利用我们的期望。伏特加市场就是此种做法颇为行之有效的一个舞台。伏特加是一种中性的烈酒，按照美国政府的定义，其蒸馏过程的目的是使其“没有可区分的性质、气味、味道或颜色”。因此，多数美国伏特加都不是由那些充满激情、身穿法兰绒衬衫、创造各种葡萄酒的人所发明，而是由诸如农业化学品供应商阿切尔·丹尼尔斯·米德兰（Archer Daniels Midland）这样的业界巨头所创造。而伏特加蒸馏工的工作，也不是通过酿造过程管理来在酒中加入微妙的气味，而是往上面那些大公司提供的、190 标准酒精强度工业级蒸馏酒精深层抽取液中掺水，并尽可能地去除其他味道。但是，通过大规模的幻觉制造，伏特加生产商们成功创造出了对于伏特加酒之间差异的强烈期望，而结果就是，人们相信不同品牌的伏特加——按其本身定义而言应该是没有可相互区别的性质——之间确实存在巨大差异，而且还愿意为这些差别付出大把的钞票。为了不让自己被当作一个没有品味的乡巴佬而踢开，我很乐于指出，这些胡言乱语其实是有办法检验的。我们可以排出一系列不同的伏特加，

再请一帮子伏特加品酒专家，进行一次盲试。实际上，《纽约时报》就这么做了[22]。去掉标签之后，如 Grey Goose 和 Ketel One 这样高品位的伏特加，其表现就不怎么好了。事实上，相比于传统看法，检验结果看来是随机的。而且，在被品尝的 21 种伏特加中，价格最便宜的那个牌子——Smirnoff——得分最高。如果判断能与期待隔离开来，并完全根据数据来得出的话，我们对世界的评价将大为不同。

几年前，伦敦的《星期日时报》（*The Sunday Times*）进行了一项实验。编辑们把获得布克奖（the Booker Prize）——当代小说领域最有声望与影响力的世界性奖项之一——的两部小说的最初几章，用打字机打了出来，并投稿到 20 个主要的出版商与中介人处[23]。两部小说之一是诺贝尔文学奖获得者奈保尔（V. S. Naipaul）的《在自由的国度》（*In a Free State*），另一部则是斯坦利·米德尔顿（Stanley Middleton）的《假日》（*Holiday*）。我们有把握相信，要是收到稿件的人知道稿件来历的话，他们都会为这两部广受赞誉的小说堆砌起赞美之辞。但这些投稿的小说片断被弄得好像是某些仍在奋斗的作家的作品一般，让出版商和中介们都无法看出稿件的真实来源。这两部非常成功的小说，其命运如何呢？所有的回复都是退稿，只有一名伦敦的文学中介人例外，他对米德尔顿的小说表示了一定的兴趣。同样是这名中介，对于奈保尔的小说写道："我们……认为作品很具原创性，但我们恐怕没有足够的激情来更进一步了。"

作家斯蒂芬·金无意中也进行了一次类似的实验。由于担心公众不能如他写书般迅速地接受他的作品，他以理查德·巴赫曼（Richard Bachman）的笔名完成了一个系列小说。根据销售数据来看，即使是斯蒂芬·金，当他没有了这个名字之后，也不再是斯蒂芬·金了（当作者的真实姓名终于泄露后，该书的销售量明显上升）。可悲的是，金所没有进行的实验，恰恰是相反的那一个：将那

些奋斗作家所写的很棒但未出版的小说，以他的名字包装后，再来看看销量如何。但既然连斯蒂芬·金在没有这个名字时也不再是斯蒂芬·金，那么当我们这些“其他人”的作品得到的待遇不是那么“斯蒂芬·金”式的时候，我们也许可以安慰自己，作品质量上的差别大概并不像某些人希望我们相信的那样悬殊。

几年前，我曾在加州理工拥有一间办公室，就在一位名叫约翰·施瓦兹（John Schwarz）的物理学家的办公室过来拐角的地方。他几乎不为人知，而且差不多是单枪匹马地维持着一个名为“弦理论”的没人相信的理论，该理论预言空间具有比所观察到的 3 维还要多得多的维数。他忍受了整整 10 年的奚落。然后在某一天，他与一名同事取得了技术上的突破，并由于一些跟本书关系不大的原因，那些多余的维数突然间变得可接受了。之后，弦理论成为了物理学最大的热点。如今，约翰被看作物理学界杰出的元老之一。但如果他在那些灰色的年头中屈服了，他就会成为托马斯·爱迪生（Thomas Edison）所说的“生活中的许多失败，乃是由于很多人没有认识到，在他们放弃的时候，他们是多么接近于成功”[24]这句话的活证据。

我所认识的另一名物理学家，也有着与约翰惊人相似的经历。他碰巧就是约翰在加州大学伯克利分校的博士生导师。这位被认为是他那一代中最为出众的科学家之一的物理学家，是一个被称为 S 矩阵理论的研究领域的带头人。与约翰一样，他十分固执，并在其他人都已放弃的情况下，仍然坚持对该理论进行了多年的研究。但与约翰不同，他最终未能成功。正是由于没有成功，他被许多人视为怪人，并以此身份结束了职业生涯。但我的意见是，他与约翰都是非常聪明的物理学家，也都有勇气在出现重大突破的前景还不明朗时，仍然继续在一个过时的理论上进行研究。正如对作家应该由他们写作的内容而非作品的销量来加以评判一样，对物理学家——以及所有努力奋斗的人——都更应该以能力来评判，而非更多地看

他们是否获得了成功。

将能力与成功绑在一起的绳子，不但很松，而且还弹性十足。我们可以容易地在成功的书中发现其优点，也可以在未出版的手稿、廉价的伏特加或仍然在某个领域中艰苦奋斗的人身上发现其缺点；我们很容易相信那些奏效的点子都是好点子，那些成功的计划都是设计良好的计划，而没有奏效的点子与计划则根本在构思上就有毛病；我们也很容易从那些最为成功的人中制造出英雄，而对那些最不成功的人报以轻蔑的一瞥。但能力并不保证定能获得成就，而成就也并非与能力成比例。因此，重要之处在于不要忘记这个成功方程式中的另一项——偶然性所扮演的角色。

将某个领域中最为成功的人视作超级英雄，这并不是悲剧；但如果我们的放弃乃是由于对专家评判或市场业绩的信任，而非由于对自身的信念，如图尔在他那部后来成为畅销书的《笨伯联盟》手稿被一次次退稿后自杀那样，这就的确是一个悲剧。因此，在试图根据某人的成功程度而对他/她进行评价时，我会提醒自己，如果他们有机会从头再来，斯蒂芬·金也许不过是一个理查德·巴赫曼，奈保尔也不过是另一个还在努力拼搏的作家，而在那人群中，还游荡着许许多多与比尔·盖茨、布鲁斯·威利斯和罗杰·马立斯同样了不起的人，只不过命运没有赐给他们那些对路的划时代产品或电视节目或体育赛季。我所学到的，首先就是一直向前，因为最好的消息就是，既然偶然性确实扮演了某个角色，那通往成功的要素之一就在我们掌握之中了：我们上垒击球的次数，我们获得机会的次数，我们把握机会的次数。因为即使是一枚更容易掷出失败结果的硬币，有时也会掷出成功。或者如 IBM 先驱托马斯·华生（Thomas Watson）所说："想要成功，就把你的失败速度加倍。"

我已试着在本书中展现随机性的基本概念，说明它们如何运用于与世事之中，并表明了我的观点，即在解释事件时，以及在期望与决策中，随机性的作用很大程度上被忽视。哪怕只是认识到随机

过程在生活中扮演的无处不在的角色，也可以让我们豁然开朗。但随机过程理论的真正力量，在于一旦理解了随机过程的本性，我们就会改变对于周围事物的认识方式。

心理学家罗森汉写道："一个人一旦不正常，他所有的举止和特征就都会被挂上'不正常'的标签。"[25]这句话对演员以及许多其他成败的标签也同样适用。我们通过结果来评判人和动机，我们希望事情的发生乃是由于良好且可以理解的原因。但我们所看到的清晰的必然性，通常不过是错觉罢了。我相信，在面对着不确定性时，我们能够重新组织我们的思维方式。出于这种信念，我写下了这本书。我们可以提高决策技巧，并克服某些能导致不当判断与选择的偏见。我们应该不以成败论英雄地了解他人的能力与各种情况，并学会通过所有可能结果的分布而非实际所得的结果，来判断决策的优劣。

母亲总是告诫我，不要以为自己可以预测或控制未来。有一次，她讲述了使她转向这个信念的那件事。这事儿与她的姐姐萨比娜有关，尽管距母亲上次见到她已有 65 年，但母亲却仍然常常提到她。母亲当时 15 岁，以一种弟弟妹妹们对哥哥姐姐有时所怀有的崇敬之情，崇拜着她的姐姐。那时纳粹已经侵入了波兰，出身贫民区的父亲加入了地下抵抗组织，并如我所提到过的，最终被关进了布痕瓦尔德集中营。当时还不认识父亲的母亲则出身于富人区，并被关到了一个劳动改造营。在那里，她被指派为护士助手去照料斑疹伤寒患者。食物十分稀缺，不可预料的死亡却总在近旁。为了在那危险无时不在的环境中保护我母亲，萨比娜采取了一个计划。她的一名朋友是犹太警察成员，这些警察按照德国人的命令维持集中营秩序，并通常都遭到同集中营的其他人的蔑视。萨比娜的朋友向她求婚——只是名义上的婚姻——以使萨比娜能得到他的职位所能提供的保护。考虑到我母亲也可能沾光，萨比娜接受了求婚。然后事情就发生了。德国人开始讨厌这些犹太警察了，于是，他们将若干犹

太警察官员连同配偶一起送进了毒气室——其中就包括萨比娜和她的丈夫。相比于与萨比娜相处的日子，我母亲现在已经活过了更长的岁月，但萨比娜的死仍然纠缠着她。母亲担心一旦她自己也过世，那萨比娜留下的痕迹就不复存在了。对她而言，这个故事说明计划根本毫无意义。我并不赞同这一点。我相信计划十分重要，如果我们是睁着眼睛做计划的话。但更重要的是，母亲的经历告诉我，我们应当看到我们所得的好运，并对此心怀感恩，进而认识到随机事件在成功中所占的分量。这件事还教会我去接受那些令人悲痛的偶然事件。最重要的是，它教会了我去感恩，庆幸没有碰到坏运气，没有碰到那些可能击倒我们的事，没有经历疾病、战争、饥荒，以及那些没有——或者说尚且没有——落到我们头上的意外。

致 谢

如果您一直读到了这个地方，那我就假定您喜欢本书。对于本书的优点，我愿声明本人拥有全部的功劳。不过，正如尼克松（Nixon）曾说过的："这是不对的。"因此，我愿在此列出那些以其时间、知识、天分和/或耐心来帮助我创作出这样一本比我独力所能写出的任何作品都要好的书。首先，唐娜·斯科特（Donna Scott）、马克·希勒里（Mark Hillery）和马特·科斯特洛（Matt Costello）一直给予我鼓励。特别是马克，他让我写一本关于熵的书，而当我将许多与之相同的思路应用于日常世界时，他报以我耐心的倾听（和阅读）。我的中介人苏珊·金斯伯格（Susan Ginsburg）从来也不希望我写关于熵的东西，但如马克一样，她也是从不动摇的建设性意见和鼓励的源泉。我的朋友茱蒂丝·克劳斯代尔（Judith Croasdell）一直都支持我，而且在我的请求下，做出了一两件非常了不起的事情。编辑爱德华·卡斯滕迈尔（Edward Kastenmeier）从不厌烦我把他扯进关于文章风格以及几乎每句话的内容的冗长讨论之中，或者更可能的是，他出于礼貌而没有向我抱怨罢了。我还欠爱德华的同事马提·阿什尔（Marty Asher）、丹·福兰克（Dan Frank）和提姆·奥康内尔（Tim O'Connell）一个人情，他们和爱德华一道养育了这本书，并帮助了写作；还有詹妮斯·戈尔德克朗（Janice Goldklang）、美智子·克拉克（Michiko Clark）、克里斯·吉莱斯皮（Chris Gillespie）、凯斯·戈尔德史密斯（Keith Goldsmith）、詹姆斯·金伯（James Kimball）和瓦尼莎·施耐德（Vannessa Schnei-

der)，正是他们不倦的努力，帮助我将本书献给您面前。

技术方面，拉里·戈尔德斯泰恩（Larry Goldstein）和特德·希尔（Ted Hill）在许多有趣而刺激的数学争辩和讨论中启发了我，并对我的手稿给予了无价的反馈意见。弗雷德·罗斯（Fred Rose）则似乎纯粹为了有时间来向我提供有关金融市场运作方面的建议，而放下了《华尔街时报》的工作。莱尔·朗（Lyle Long）将他那了不起的专业能力用在了数据分析上，帮助创作了有关基金经理表现的那些图。克里斯托弗·科赫（Christof Koch）则将我迎入了加州理工他的实验室中，使我在令人激动并散布于书中的神经科学新进展上大开眼界。许多其他朋友和同事阅读了若干章节，有时还不止一稿；还有的则提供了有益的建议或信息。他们包括杰德·布赫瓦尔德（Jed Buchwald）、林恩·寇克斯（Lynne Cox）、理查德·切维顿（Richard Cheverton）、吕蓓卡·福斯特（Rebecca Forster）、米里安·古德曼（Miriam Goodman）、凯瑟琳·基弗（Catherine Keefe）、杰夫·麦考维亚克（Jeff Mackowiak）、辛迪·梅尔（Cindy Mayer）、派翠西亚·麦克福尔（Patricia MaFall）、安迪·梅斯勒（Andy Meisler）、斯蒂夫·蒙洛迪诺（Steve Mlodinow）、菲尔·里德（Phil Reed）、塞斯·罗伯茨（Seth Roberts）、劳拉·萨阿里（Laura Saari）、马特·萨尔加尼克（Matt Salganik）、马丁·史密斯（Martin Smith）、斯蒂夫·托马斯（Steve Thomas）、黛安·特纳（Diane Turner）和杰里·韦伯曼（Jerry Webman）。感谢你们。最后，我还欠我的家人们——唐娜（Donna）、阿列克赛、尼古拉、奥丽维亚和我母亲伊莱娜（Irene）——一个深深的谢意，我常常占用了你们的时间来改进或至少是使我能专注于写作。

注 释

前言

1. Stanley Meisler, "First in 1763: Spain Lottery—Not Even War Stops It," Los Angeles Times, 1977. 12. 30.

2. 关于篮球，请参见 Michael Patrick Allen, Sharon K. Panian, and Roy E. Lotz, "Managerial Succession and Organizational Performance: A Recalcitrant Problem Revisited," Administrative Science Quarterly, 24, no. 2, 1979 年 7 月：167～180；关于橄榄球，请参见 M. Craig Brown, "Administrative Succession and Organizational Perfromance: The Succession Effect," Administrative Science Quarterly, 27, no. 1, 1982 年 3 月：1～16；关于棒球，请参见 Oscar Grusky, "Managerial Succession and Organizational Effectiveness," American Journal of Sociology, 69, no. 1, 1963 年 7 月：21～31，以及 William A. Gamson and Norman A. Scotch, "Scapegoating in Baseball," American Journal of Sociology, 70, no. 1, 1964 年 7 月：69～72；关于足球，请参见 Ruud H. Koning, "An Econometric Evaluation of the Effect of Firing a Coach on Team Performance," Applied Economics, 35, no. 5, 2003 年 3 月：555～564.

3. James Surowiecki, The Wisdom of Crowds (New York: Doubleday, 2004), pp. 218～219.

4. Armen Alchian, "Uncertainty, Evolution, and Economic Theory," Journal of Political Economy, 58, no. 3, 1950 年 6

月：213.

第一章　透过随机性的目镜凝视

1. Kerstin Preuschoff, Peter Bossaerts, and Steven R. Quartz, "Neural Differentiation of Expected Reward and Risk in Human Subcortical Structures," Neuron, 51, 2006. 8. 3.: 381～390.

2. Benedetto De Martino et al., "Frames, Biases, and Rational Decision-Making in the Human Brain," Science, 313, 2006. 8. 4.: 684～687.

3. George Wolford, Michael B. Miller, and Michael Gazzaniga, "The Left Hemisphere's Role in Hypothesis Formation," Journal of Neuroscience, 20: RC64, 2000: 1～4.

4. Bertrand Russel, An Inquiry into Meaning and Truth, 1950. Oxford: Routledge 1996 年重印, p. 15.

5. Matt Johnson and Tom Hundt, "Hog Industry Targets State for Good Reason," Vernon County (Wisconsin) Broadcaster, 2007. 7. 17.

6. Kevin McKean, "Decisions, Decisions," Discover, 1985 年 6 月：pp. 22～31.

7. David Oshinsky, "No Thanks, Mr. Nabokov," New York Times Book Review, 2007. 9. 9.

8. 关于这些手稿被退稿的次数，各个报道之间略有出入。

9. William Goldman, Adventures in the Screen Trade, New York: Warner Books, 1983, p. 41.

10. 请见 Arthur De Vany, Hollywood Economics, Abington, U. K.: Routledge, 2004.

11. William Feller, An Introduction to Probability Theory and Its Applications, 第 2 版, New York: John Wiley and Sons, 1957, p. 68. 请注意，为简洁起见，如果双方打成平手时，Feller 定义在上

一轮中领先的选手仍然保持领先地位。

12. Leonard Mlodinow, "Meet Hollywood's Latest Genius," Los Angeles Times Magazine, 2006. 7. 2.

13. Dave McNary, "Par Goes for Grey Matter," Variety, 2005. 1. 2.

14. Ronald Grover, "Paramount's Cold Snap: The Heat Is On," BusinessWeek, 2003. 11. 21.

15. Dave McNary, "Parting Gifts: Old Regime's Pics Fuel Paramount Rebound," Variety, 2005. 8. 16.

16. Anita M. Busch, "Canton Inks Prod'n Pact at Warner's," Variety, 1997. 8. 7.

17. "The Making of a Hero," Time, 1961. 9. 29, p. 72. 这位老前辈是罗杰斯·霍恩斯比。

18. "Mickey Mantle and Roger Maris: The Photographics Essay," Life, 1961. 8. 18, p. 62.

19. 我为那些不知道棒球的读者解释一下：本垒是一块厚橡胶皮，它嵌在地面里，而球员站在它前头来击球。那些知道棒球规则的读者请注意，我将投手投出 4 个坏球后将击球手保送上垒的情况，也包括在了我所说的机会之中。如果只使用正式的击球机会数来重新计算，结果也基本上差不多。

20. 请参见 Stephen Jay Gould, "The Streak of Streaks," New York Review of Books, 1988. 8. 18., pp. 8～12（我们将在后面以更多的细节来谈论他们的工作）. 关于体育中的扔硬币模型的更有说服力且数学方面更为详细的分析，出现在 Charles M. Grinstead、William P. Peterson 和 J. Laurie Snell 正在撰写的一本初步定名为 Fat Chances 的书的第二章中：www. math. dartmouth. edu/～ prob/prob/NEW/bestofchance. pdf.

第二章　真理与半真理的法则

1. Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., Judgment under Uncertainty: Heuristics and Biases, Cambridge: Cambridge University Press, 1982, pp. 90～98.

2. Amos Tversky and Daniel Kahneman, "Extensional versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment," Psychological Review, 90, no. 4, 1983. 10.: 293～315.

3. Craig R. Fox and Richard Birke, "Forecasting Trial Outcomes: Lawyers Assign Higher Probabilities to Possibilities that Are Described in Greater Detail," Law and Human Behavior, 26, no. 2, 2002. 4.: 159～173.

4. Plato, The Dialogues of Plato, Benjamin Jowett 英译本, Boston: Colonial Press, 1899, p. 116.

5. Plato, Theaetetus, Whitefish, Mont: Kessinger, 2004, p. 25.

6. Amos Tversky and Daniel Kahneman, "Availability: A Heuristic for Judging Frequency and Probability," Cognitive Psychology, 5, 1973: 207～232.

7. Reid Hastie and Robyn M. Dawes, Rational Choice in an Uncertain World: The Psychology and judgement of Decision Making, Thousand Oaks, Calif.: Sage, 2001, p. 87.

8. Robert M. Reyes, William C. Thompson, and Gordon H. Bower, "Judgmental Biases Resulting from Differing Availabilities of Arguments," Journal of Personality and Social Psychology, 39, no. 1, 1980: 2～12.

9. Robert Kaplan, The Nothing That Is: A Natural History of Zero, London: Oxford University Press, 1999, pp. 15～17.

10. Cicero, 转引自 Morris Kline, Mathematical Thought from Ancient to Modern Times, London: Oxford University Press, 1972, 1: 179.

11. Morris Kline, Mathematics in Western Culture, London: Oxford University Press, 1953, p. 86.

12. Kline, Mathematical Thought, pp. 178～179.

13. Cicero, 转引自 Warren Weaver, Lady Luck, Mineola, N. Y.: Dover Publications, 1982, p. 53.

14. Cicero, 转引自 F. N. David, Gods, Games and Gambling: A History of Probability and Statistical Ideas, Nineola, N. Y.: Dover Publications, 1998, pp. 24～26.

15. Cicero, 转引自 Bart K. Holland, What Are the Chances? Voodoo Deaths, Office Gossip, and Other Adventures in Probability, Baltimore: Johns Hopkins University Press, 2002, p. 24.

16. Cicero, 转引自 Bart K. Holland, What Are the Chances? Voodoo Deaths, Office Gossip, and Other Adventures in Probability, Baltimore: Johns Hopkins University Press, 2002, p. 25.

17. James Franklin, The Science of Conjecture: Evidence and Probability before Pascal, Baltimore: Johns Hopkins University Press, 2001, pp. 4, 8.

18. James Franklin, The Science of Conjecture: Evidence and Probability before Pascal, Baltimore: Johns Hopkins University Press, 2001, p. 13.

19. James Franklin, The Science of Conjecture: Evidence and Probability before Pascal, Baltimore: Johns Hopkins University Press, 2001, p. 14.

20. William C. Thompson, Franco Taroni, and Colin G. G. Aitken, "How the Probability of a False Positive Affects the Value of DNA Evidence," Journal of Forensic Sciences, 48, no. 1, 2003. 1.: 1～8.

21. William C. Thompson, Franco Taroni, and Colin G. G. Ait-

ken, "How the Probability of a False Positive Affects the Value of DNA Evidence," Journal of Forensic Sciences, 48, no. 1, 2003. p. 2. 这个故事还被复述于 Bill Braun, "Lawyers Seek to Overturn Rape Conviction," Tulsa World, 1996. 11. 22. 。另见 www. innocenceproject. org. (达拉谟于 1997 年被释放。)

22. People v. Collins, 68 Calif. 2d319, 438 P. 2d 33, 66 Cal. Rptr. 497, 1968.

23. Thomas Lyon, 私人通信。

第三章 寻找穿越可能性空间之路

1. Alan Wykes, Doctor Cardano: Physician Extraordinary, London: Frederick Muller, 1969. 另见 Oystein Ore, Cardano: The Gambling Scholar, 以及由 Sydney Henry Gould 所译的卡尔达诺的《论赌博》, Princeton, N. J. : Princeton University Press, 1953.

2. Marilyn vos Savant, "Ask Marilyn," Parade, 1990. 9. 9.

3. Bruce D. Burns and Mareike Wieth, "Causality and Reasoning: The Monty Hall Dilemma," in Proceedings of the Twenty-fifth Annual Meeting of the Cognitive Science Society, ed. R. Alterman and D. Kirsh, Hillsdale, N. J. : Lawrence Erlbaum Associates, 2003, p. 198.

4. National Science Board, Science and Engineering Indicators—2002, Arlington, Va. : National Science Foundation, 2002. http://www. nsf. gov/statistics/seind02/. 请参见第 2 卷第 7 章表 7～10。

5. Gary P. Posner, "Nation's Mathematicians Guilty of Innumeracy," Skeptical Inquirer, 15, no. 4, 1991 年夏。

6. Bruce Schechter, My Brain Is Open: The Mathematical Journeys of Paul Erdös, New York: Touchstone, 1998, pp. 107～109.

7. Bruce Schechter, My Brain Is Open: The Mathematical Jour-

neys of Paul Erdös, New York: Touchstone, 1998, pp. 189～190, 196～197.

8. John Tierney, "Behind Monty's Doors: Puzzle, Debate and Answer?" New York Times, 1991. 7. 21.

9. Robert S. Gottfried, The Black Death: Natural and Human Disaster in Medieval Europe, New York: Free Press, 1985.

10. Gerolamo Cardano, 转引自 Wykes, Doctor Cardano, p. 18.

11. Kline, Mathematical Thought, pp. 184～185, 259～260.

12. "Oprah's New Shape: How She Got It," O, the Oprah Magazine, 2003. 1.

13. Lorraine J. Daston, Classical Probability in the Enlightenment, Princeton, N. J.: Princeton University Press, 1998, p. 97.

14. Marilyn vos Savant; "Ask Marilyn", Parade, 1996. 3. 3., p. 14.

15. 车子有 4 个轮胎，因此，令 RF 表示"右前轮"，LF 表示"左前轮"，RR 表示"右后轮"，LR 表示"左后轮"，则该两名学生所给的答案有 16 种可能的组合。如果把学生 1 的答案写在前面而学生 2 的答案写在后面，则这些可能的答案组合为（RF，RF）、（RF，LF）、（RF，RR）、（RF，LR）、（LF，RF）、（LF，LF）、（LF，RR）、（LF，LR）、（RR，RF）、（RR，LF）、（RR，RR）、（RR，LR）、（LR，RF）、（LR，LF）、（LR，RR）和（LR，LR）。这 16 个组合中，有 4 个是相互吻合的：（RF，RF）、（LF，LF）、（RR，RR）和（LR，LR）。因此所求的可能性为 4/16，或说 1/4。

16. Martin Gardner, "Mathematical Games", Scientific American, 1959. 10., pp. 180～182.

17. Jerome Cardan, The Book of My Life: De Vita Propia Liber, Jean Stoner 英译本, Whitefish, Mont.: Kessinger, 2004, p. 35.

18. Cardano, 转引自 Wykes, Doctor Cardano, p. 57.

19. Cardano，转引自 Wykes，Doctor Cardano，p. 57.

20. Cardano，转引自 Wykes，Doctor Cardano，p. 172.

第四章　追寻通往成功的路径

1. Bergt Ankarloo and Stuart Clark，eds.，Witchcraft and Magic in Europe：The Period of the Witch Trials，Philadelphia：University of Pennsylvania Press，2002，pp. 99～104.

2. Meghan Collins，"Traders Ward Off Evil Spirits，" 2003. 10. 31.，http://www. CNNMoney. com/2003/10/28/markets _ trader _ superstition/index. htm.

3. Henk Tijms，Understanding Probability：Chance Rules in Everyday Life，Cambridge：Cambridge University Press，2004，p. 16.

4. Henk Tijms，Understanding Probability：Chance Rules in Everyday Life，Cambridge：Cambridge University Press，2004，p. 80.

5. David，Gods，Games and Gambling，p. 65.

6. Blaise Pascal，转引自 Jean Steinmann，Pascal，Martin Turnell 英译本，New York：Harcourt，Brace & World，1962，p. 72.

7. Gilberte Pascal，转引自 Morris Bishop，Pascal：The Life of a Genius，1936，New York：Greenwood Press，1968 年重印，p. 47.

8. Gilberte Pascal，转引自 Morris Bishop，Pascal：The Life of a Genius，1936，New York：Greenwood Press，1968 年重印，p. 137.

9. Gilberte Pascal，转引自 Morris Bishop，Pascal：The Life of a Genius，1936，New York：Greenwood Press，1968 年重印，p. 135.

10. 请参见 A. W. F. Edwards，Pascal's Arithmatical Triangle：The Story of a Mathematical Idea，Baltimore：Johns Hopkins University Press，2002.

11. Blaise Pascal，转引自 Herbert Westren Turnbull，The Great Mathematicians，New York：New York University Press，1961，p. 131.

12. Blaise Pascal，转引自 Bishop，Pascal，p. 196.

13. Blaise Pascal，转引自 David，Gods，Games and Gambling，p. 252.

14. Bruce Martin，“Coincidences：Remarkable or Random?” Skeptical Inquirer，22，no. 5，1998. 9. /10.

15. Holland，What Are the Chances? pp. 86～89.

第五章　针锋相对的大数定律与小数定律

1. Tijms，Understanding Probability，p. 53.

2. Scott Kinney，“Judge Sentences Kevin L. Lawrence to 20 Years Prison in Znetix/HMC Stock Scam，” Washington State Department of Financial Institutions 的新闻发布会，2003. 11. 25. http://www. dfi. wa. gov/sd/kevin _ laurence _ sentence. htm.

3. 与 Darrell Dorrell 的访谈，2005. 8. 1.

4. Lee Berton，“He’s Got Their Numbers：Scholar Uses Math to Foil Financial Fraud，” Wall Street Journal，1995. 7. 10.

5. Charles Sanders Peirce，Max Harold Fisch，and Christian J. W. Kloesel，Writings of Charles S. Peirce：A Chronological Edition，Bloomington：Indiana University Press，1982，p. 427.

6. Rand Corporation，A Million Random Digits with 100 000 Normal Deviates，1955，Santa Monica，Calif. ：Rand，2001 年重印，pp. ix～x. 另见 Lola L. Lopes，“Doing the Impossible：A Note on Induction and the Experience of Randomness，” Journal of Experimental Psychology：Learning，Memory，and Cognition，8，no. 6，1982. 11. ：626～636.

7. 关于约瑟夫·贾格（有时被拼写为贾格斯）的记载出自 John Grochowski，“House Has a Built-in Edge When Roulette Wheel Spins，” Chicago Sun-Times，1997. 2. 21.

8. 有关伯努利家族和雅可布生平的细节，请参见 E. S. Pearson，

ed. The History of Statistics in the 17th and 18th Centuries against the Changing Background of Intellectual, Scientific and Religious Thought: Lectures by Karl Pearson Given at University College, London, during the Academic Sessions 1921～1933, New York: Macmillan, 1978, pp. 221～237; J. O. Fleckenstein, "Johann und Jakob Bernoulli," in Elemente der Mathematik, Beihefte zur Zeitschrift, no. 6, Basel, 1949; 及 Stephen Stigler, "The Bernoullis of Basel," Journal of Econometrics, 75, no. 1, 1996: 7～13.

9. 转引自 Pearson, The History of Statistics in the 17th and 18th Centuries, p. 224.

10. Stephen Stigler, The History of Statistics: The Measurement of Uncertainty before 1900, Cambridge, Mass.: Harvard University Press, 1986, p. 65.

11. Pearson, The History of Statistics in the 17th and 18th Centuries, p. 226.

12. William H. Cropper, The Great Physicists: The Life and Times of Leading Physicists from Galileo to Hawking, London: Oxford University Press, 2001, p. 31.

13. Johann Bernoulli, 转引自 Pearson, The History of Statistics in the 17th and 18th Centuries, p. 232.

14. 当然，这依赖于你把什么标记为"现代概念"。我所使用的是 Hankel 于 1871 年所著的关于这个问题的历史书中采用的定义，该定义的极详细描述见 Gert Schubring, Conflicts between Generalization, Rigor, and Intuition: Number Concepts Underlying the Development of Analysis in 17th～19th Century France and Germany, New York: Springer, 2005, pp. 22～32.

15. David Freedman, Robert Pisani, and Roger Purves, Statistics, 第 3 版, New York: W. W. Norton, 1998, pp. 274～275.

16. 有关哈金的引文出自 Ian Hacking，The Emergence of Probability，Cambridge：Cambridge University Press，1975，p. 143. 有关伯努利的引文出自 David，Gods，Games and Gambling，p. 136.

17. 对于伯努利所真正证明的内容的一个讨论，请见 Stigler，The History of Statistics，pp. 63～78；及 Ian Hacking，The Emergence of Probablity，pp. 155～165.

18. Amos Tversky and Daniel Kahneman，"Belief in the Law of Small Numbers，" Psychological Bulletin，76，no. 2，1971：105～110.

19. Jakob Bernoulli，转引自 L. E. Maistrov，Probability Theory：A Historical Sketch，Samuel Kotz 英译本，New York：Academic Press，1974，p. 68.

20. Stigler，The History of Statistiss，p. 77.

21. E. T. Bell，Men of Mathematics，New York：Simon & Schuster，1937，p. 134.

第六章　假阳性与好错误

1. 关于这名哈佛学生的记载出自 Hastie and Dawes，Rational Choice in an Uncertain World，pp. 320～321.

2. 纽约城市大学亨特学院物理系的 Mark Hillery 曾告诉我该问题的一个变体，而他是在一次去斯洛伐克布拉迪斯拉发的旅途中听说的。

3. 引自 Stigler，The History of Statistics，p. 123.

4. 引自 Stigler，The History of Statistics，pp. 121～131.

5. 美国社会福利局，"Popular Baby Names：Popular Names by Birth Year；Popularity in 1935，" http：//www. ssa. gov/cgi-bin/popularnames. cgi.

6. 传染性疾病中心 HIV/AIDS 部，HIV/AIDS Surveillance Report，Atlanta：Centers for Disease Control，1990. 1. 我引用其中给

出的数据进行了统计，不过计算中仍然需要某些估计，特别是我所引用的数据给出的是艾滋病发病数量而非 HIV 感染数量。不过对于说明概念的目的而言，这个统计也足以接受了。

7. 更准确点来说，当 B 发生时 A 也将发生的概率，等于当 A 发生时 B 也将发生的概率，再乘上一个校正因子，它等于 A 发生概率与 B 发生概率之比。

8. Gerd Gigerenzer, Calculated Risks: How to Know When Numbers Deceive You, New York: Simon & Schuster, 2002, pp. 40～44.

9. Donald A. Barry and Lee Ann Chastain, "Inferences About Testosterone Abuse Among Athletes," Chance, 17. no. 2, 2004: 5～8.

10. John Batt, Stolen Innocence, London: Ebury Press, 2005.

11. Stephen J. Watkins, "Conviction by Mathematical Error? Doctors and Lawyers Should Get Probability Theory Right," BMJ, 320, 2000. 1. 1. : 2～3.

12. "Royal Statistical Society Concerned by Issues Raised in Sally Clark Case,''伦敦，皇家统计学会，2001. 10. 23. 新闻发布会；http://www. rss. org. uk/PDF/RSS% 20Statement% 20regarding% 20statistical% 20 issues% 20in% 20the% 20Sally% 20Clark% 20case,%20October%2023rd%202001. pdf.

13. Ray Hill, "Multiple Sudden Infant Deaths - Coincidence or beyond Coincidence?" Paediatric and Perinatal Epidemiology, 18, no. 5, 2004. 9. : 320～326.

14. 转引自 Alan Dershowitz, Reasonable Doubts: The Criminal Justice System and the O. J. Simpson Case, New York: Simon & Schuster, 1996, p. 101.

15. 联邦调查局，"Uniform Crime Reports," http://www. fbi.

gov/ucr/ucr. htm.

16. Alan Dershowitz, The Best Defense, New York: Vintage, 1983, p. xix.

17. Pierre-Simon de Laplace, 转引自 James Newman, ed., The World of Mathematics, Mineola, N. Y.: Dover Publications, 1956: 2: 1323.

第七章 测量与误差定律

1. Sarah Kershaw and Eli Sanders, "Recounts and Partisan Bickering Bring Election Fatigue to Washington Voters," New York Times, 2004. 12. 26.; 与 Timothy Egan, "Trial for Governor's Seat Set to Start in Washington," New York Times, 2005. 5. 23.

2. Jed Z. Buchwald, "Discrepant Measurements and Experimental Knowledge in the Early Modern Era," Archive for History of Exact Sciences, 60, no. 6, 2006. 11.: 565～649.

3. Eugene Frankel, "J. B. Biot and the Mathematization of Experimental Physics in Napoleonic France," in Historical Studies in the Physical Sciences, ed. Russel McCormmach, Princeton, N. J.: Princeton University Press, 1977.

4. Charles Coulston Gillispie, ed., Dictionary of Scientific Biography, New York: Charles Scribner's Sons, 1981, p. 85.

5. 关于雷达枪测量误差的一个讨论，请见 Nicole Weisensee Egan, "Takin' Aim at Radar Guns," Philadelphia Daily News, 2004. 3. 9.

6. Charles T. Clotfelter and Jacob L. Vigdor, "Retaking the SAT"（工作论文 SAN01～20, Terry Sanford Institute of Public Policy, 杜克大学, Durham, N. C., 2001. 7.）.

7. Eduardo Porter, "Jobs and Wages Increased Modestly Last Month," New York Times, 2006. 9. 2.

8. Gene Epstein 于“数学魔术家”，On the Media，WNYC 广播，2006 年 8 月 25 日播放.

9. Legene Quesenberry et al.，“Assessment of the Writing Component within a University General Education Program,” 2000. 11. 1.；http://wac. colostate. edu/aw/articles/quesenberry2000/quesenberry2000. pdf.

10. Kevin Saunders，“Report to the Iowa State University Steering Committee on the Assessment of ISU Comm-English 105 Course Essays,” 2004. 9.； www. iastate. edu/～isucomm/InYears/ISUcomm-essays. pdf（2005 年浏览，网站现已停止运行）.

11. University of Texas，Office of Admissions，“Inter-rate Reliability of Holistic Measures Used in the Freshman Admissions Process of the University of Texas at Austin,” 2005. 2. 22.；http://www. utexas. edu/student/admissions/research/Inter-rateReliability2005. pdf.

12. Emily J. Shaw and Glenn B. Milewski，“Consistency and Reliability in the Individualized Review of College Applicants,” College Board，Office of Research and Development，Research Notes RN—20，2004. 10.：3；http：//www. collegeboard. com/research/pdf/RN—20. pdf.

13. Gary Rivlin，“In Vino Veritas,” New York Times，2006. 8. 13.

14. William James，The Principles of Psychology，New York：Henry Holt，1890，p. 509.

15. Robert Frank and Jennifer Byram，“Taste-Smell Interactions Are Tastant and Odorant Dependent,” Chemical Senses，13，1988：445～455.

16. A. Rapp，“Natural Flavours of Wine：Correlation between Intrumental Analysis and Sensory Perception,” Fresenius’ Journal of Analytic Chemistry，337，no. 7，1990. 1.：777～785.

17. D. Laing and W. Francis, "The Capacity of Humans to Identify Odors in Mixtures,' Physiology and Behavior, 46, no. 5, 1989. 11. : 809～814; 与 D. Laing et al., "The Limited Capacity of Humans to Identify the Components of Taste Mixtures and Taste-Odour Mixtures," Perception, 31, no. 5, 2002: 617～635.

18. 关于玫瑰红的研究，请见 Rose M. Pangborn, Harold W. Berg, and Brenda Hansen, "The Influence of Color on Discrimination of Sweetness in Dry Table-Wine," American Journal of Psychology, 76, no. 3, 1963. 9. : 492～495. 关于葡萄花青素染料的研究，请见 G. Morrot, F. Brochet, and D. Dubourdieu, "The Color of Odors," Brain and Language, 79, no. 2, 2001. 11. : 309～320.

19. Hilke Plassman, John O'Doherty, Baba Shia, and Antonio Rongel, "Marketing Actions Can Modulate Neural Representations of Experienced Pleasantness," Proceedings of the National Academy of Sciences, 2008. 1. 14. ; http://www. pnas. org.

20. M. E. Woolfolk, W. Castellan, and C. Brooks, "Pepsi versus Coke: Labels, Not Tastes, Prevail," Psychological Reports, 52, 1983: 185～186.

21. M. Bende and S. Nordin, "Perceptual Learning in Olfaction: Professional Wine Tasters Versus Controls," Physiology and Behavior, 62, no. 5, 1997. 11. : 1065～1070.

22. Gregg E. A. Solomon, "Psychology of Novice and Expert Wine Talk," American Journal of Psychology, 103, no. 4, 1990 年冬: 495～517.

23. Rivlin, "In Vino Veritas."

24. Rivlin, "In Vino Veritas."

25. Hal Stern, "On the Probability of Winning a Football Game," American Statistician, 45, no. 3, 1991. 8. : 179～182.

26. 图片来自 Index Funds Advisors，“Index Funds. com：Take the Risk Capacity Survey，” http://www. indexfunds3. com/step3page2. php，其中注明出处为 Walter Good and Roy Hermansen，Index Your Way to Investment Success，New York：New York Insitute of Finance，1997. 这 300 名基金经理的数据取自 Morningstar Principia 数据库，并将其中 10 年（1987～1996）的数据列为表格。

27. Polling Report，“President Bush—Overall Job Rating，” http://pollingreport. com/BushJob. htm.

28. “Poll：Bush Apparently Gets Modest Bounce，” CNN，2004. 9. 8.，http://www. cnn. com/2004/ALLPOLITICS/09/06/presidential. poll/index. html.

29. “Harold von Braunhut，” Telegraph，2003. 12. 23.；http://www. telegraph. co. uk/news/main. jhtml? xml＝/news/2003/12/24/db2403. xml.

30. James J. Fogarty，“Why Is Expert Opinion on Wine Valueless?” 西澳大利亚大学经济学系研讨论文 02. 17，珀斯，2001.

31. Stigler，The History of Statistics，p. 143.

第八章　混沌中的秩序

1. Holland，What Are the Chances? p. 51.

2. 这只是根据更近期的美国统计数据做出的近似，请见 U. S. Social Security Administration，“Actuarial Publications：Period Life Table.” 最近的表单可在如下网址得到：http://www. ssa. gov/OACT/STATS/table4c6. html.

3. Immanuel Kant，转引自 Theodore Porter，The Rise of Statistical Thinking：1820～1900，Princeton，N. J.：Princeton University Press，1988，p. 51.

4. U. S. Department of Transportation，Federal Highway Ad-

ministration，“Licensed Drivers，Vehicle Registrations and Resident Population，” http://www.fhwa.dot.gov/policy/ohim/hs03/htm/dlchrt.htm.

5. U.S. Department of Transportation，Research and Innovative Technology Administration，Bureau of Transportation Statistics，“Motor Vehicle Safety Data，” http://www.bts.gov/publications/national_tranportation_statistics/2002/html/table_02_17.html.

6. “The Domesday Book，” History Magazine，2001.10/11.

7. 格朗特的故事请见 Hacking，The Emergence of Probability，pp. 103～109；David，Gods，Games and Gambling，pp. 98～109；以及 Newman，The World of Mathematics，3：1416～1418.

8. Hacking，The Emergence of Probablity，p. 102.

9. Theodore Porter，The Rise of Statistical Thinking，p. 19.

10. 格朗特的原始表单请见 Hacking，The Emergence of Probability，p. 108. 目前的数据请见 World Health Organization，“Life Tables for WHO Member States，” http://www.who.int/whosis/database/life_tables/life_tables.cfm. 文中所引的表单，其数据取自节略的原始表单，并进行了取整。

11. Ian Hacking，The Taming of Chance，Cambridge：Cambridge University Press，1990，p. vii.

12. H. A. David，“First (?) Occurrence of Common Terms in Statistics and Probability，” in Annotated Readings in the History of Statistics，ed. H. A. David and A. W. F. Edwards，New York：Springer，2001，附录 B 与 pp. 219～228.

13. Noah Webster，American Dictionary of the English Language，1828，第 1 版的临摹本，Chesapeake，Va.：Foundation for American Christian Education，1967.

14. 有关奎特雷的材料主要来自 Stigler，The History of Statis-

tics, pp. 161～220; Stephen Stigler, Statistics on the Table: The History of Statistical Concepts and Methods, Cambridge, Mass.: Harvard University Press, 1999, pp. 51～66; 以及 Theodore Porter, The Rise of Statistical Thinking, pp. 100～109。

15. Louis Menand, The Metaphysical Club, New York: Farrar, Straus & Giroux, 2001, p. 187.

16. Holland, What Are the Chances? pp. 41～42.

17. David Yermack, "Good Timing: CEO Stock Option Awards and Company News Announcements," Journal of Finance, 52, no. 2, 1997. 6.: 449～476; 与 Erik Lie, "On the Timing of CEO Stock Option Awards," Management Science, 51, no. 5, 2005. 5.: 802～812. 另见 Charles Forelle and James Bandler, "The Perfect Payday—Some CEOs Reap Millions by Landing Stock Options When They Are Most Valuable: Luck or Something Else?" Wall Street Journal, 2006. 3. 18.

18. Justin Wolfers, "Point Shaving: Corruption in NCAA Basketball," American Economic Review, 96, no. 2, 2006. 5.: 279～283.

19. Stern, "On the Probability of Winning a Football Game."

20. David Leonhardt, "Sad Suspicions about Scores in Basketball," New York Times, 2006. 3. 8.

21. Richard C. Hollinger et al., National Retail Security Survey: Final Report, Gainsville: Security Research Project, Department of Sociology and Center for Studies in Criminal Law, University of Florida, 2002～2006.

22. Adolphe Quételet, 转引自 Theodore Porter, The Rise of Statistical Thinking, p. 54.

23. Quételet, 转引自 Menand, The Metaphysical Club, p. 187.

24. Jeffrey Kluger, "Why We Worry about the Things We

Shouldn't ... and Ignore the Thinks We Should," Time, 2006. 12. 4., pp. 65～71.

25. Gerd Gigerenzer, Empire of Chance: How Probability Changed Science and Everyday Life, Cambridge: Cambridge University Press, 1989, p. 129.

26. Menand, The Metaphysical Club, p. 193.

27. De Vany, Hollywood Economics; 请见第 IV 部分" A Business of Extremes."

28. 请见 Derek William Forrest, Francis Galton: The Life and Work of a Victorian Genius, New York: Taplinger, 1974; Jeffrey M. Stanton, "Galton, Pearson, and the Peas: A Brief History of Linear Regression for Statistics Instructors," Journal of Statistics Education, 9, no. 3, 2001; 以及 Theodore Porter, The Rise of Statistical Thinking, pp. 129～146.

29. Francis Galton, 转引自 Theodore Porter, The Rise of Statistical Thinking, p. 130.

30. Peter Doskoch, "The Winning Edge," Psychology Today, 2005. 11/12., pp. 44～52.

31. Deborah J. Bennett, Randomness, Cambridge, Mass.: Harvard University Press, 1998, p. 123.

32. Abraham Pais, The Science and Life of Albert Einstein, London: Oxford University Press, 1982, p. 17; 另见 89 页上的讨论.

33. 关于布朗与布朗运动的历史，请见 D. J. Mabberley, Jupiter Botanicus: Robert Brown of the British Museum, Braunschweigh, Germany, and London: Verlag von J. Cramer/Natural History Museum, 1985; Brian J. Ford, "Brownian Movement in Clarkia Pollen: A Reprise of the First Observations," Microscope, 40, no. 4, 1992: 235～241; 以及 Stephen Brush, "A History of Random Process. I. Brownian

Movement from Brown to Perrin," Archive for History of Exact Sciences, 5, no. 34, 1968.

34. Pais, Albert Einstein, pp. 88～100.

35. Albert Einstein, 转引自 Ronald William Clark, Einstein: The Life and Times, New York: HarperCollins, 1984, p. 77.

第九章　模式的错觉与错觉的模式

1. 请见 Authur Conan Doyle, The History of Spitirualism, New York: G. H. Doran, 1926; 与 R. L. Moore, In Search of White Crows: Spiritualism, Parapsychology, and American Culture, London: Oxford University Press, 1977.

2. Ray Hyman, "Parapsychological Research: A Tutorial Review and Critical Appraisal," Proceedings of the IEEE, 74, no. 6, 1986. 6.: 823～849.

3. Michael Faraday, "Experimental Investigation of Table-Moving," Athenaeum, 1853. 7. 2., pp. 801～803.

4. Michael Faraday, 转引自 Hyman, "Parapsychological Research," p. 826.

5. Faraday, 转引自 Hyman, "Parapsychological Research," p. 826.

6. 请见 Frank H. Durgin, "The Tinkerbell Effect: Motion Perception and Illusion," Journal of Consciousness Studies, 9, nos. 5～6, 2002. 5～6.: 88～101.

7. Christof Koch, The Quest for Consciousness: A Neurobiological Approach, Englewood, Colo.: Roberts, 2004, pp. 51～54.

8. 该项研究为 D. O. Clegg et al., "Glucosamine, Chondroitin Sulfate, and the Two in Combination for Painful Knee Osteoarthritis," New England Journal of Medicine, 354, no. 8, 2006. 2.: 795～808. 该访谈节目为 "Slate's Medical Examiner: Doubts on Supple-

ments," Day to Day, NPR广播, 2006. 3. 13.

9. 请见 Paul Slovic, Howard Kunreuther, and Gilbert F. White, "Decision Processes, Rationality, and Adjustment to Natural Hazards," in Natural Hazards: Local, National, and Global, ed. G. F. White, London: Oxford University Press, 1974; 另见 Willem A. Wagenaar, "Generation of Random Sequences by Human Subjects: A Critical Survey of Literature," Psychological Bulletin, 77, no. 1, 1972. 1.: 65～72.

10. 请见 Hastie and Dawes, Rational Choice in an Uncertain World, pp. 19～23.

11. George Spencer-Brown, Probability and Scientific Inference, London: Longmans, Green, 1957, pp. 55～56. 实际上，10这个值是一个严重的低估。

12. Janet Maslin, "His Heart Belongs to (Adorable) iPod," New York Times, 2006. 10. 19.

13. Hans Reichenbach, The Theory of Probability, E. Hutton and M. Reichenbach 英译本, Berkeley: University of California Press, 1934.

14. 详述该观点的经典文本是 Burton G. Malkiel, A Random Walk Down Wall Street, 现在已被重新修订为第8版, New York: W. W. Norton, 2003.

15. John R. Nofsinger, Investment Blunders of the Rich and Famous and What You Can Learn from Them, Upper Saddle River, N. J.: Prentice Hall, Financial Times, 2003, p. 62.

16. Hemang Desai and Prem C. Jain, "An Analysis of the Recommendations of the 'Superstar' Money Managers at Barron's Annual Roundtable," Journal of Finance, 50, no. 4, 1995. 9.: 1257～1273.

17. Jess Beltz and Robert Jennings, "Wall $treet Week with

Louis Rukeyser's Recommendations: Trading Activity and Performance," Review of Financial Economics, 6, no. 1, 1997: 15～27; 与 Robert A. Pari, "Wall $ treet Week Recommendations: Yes or No?" Journal of Portfolio Management, 14, no. 1, 1987: 74～76.

18. Andrew Metrick, "Performance Evaluation with Transactions Data: The Stock Selection of Investment Newsletters", Journal of Finance, 54, no. 5, 1999. 10. : 1743～1775; 与" The Equity Performance of Investment Newsletters", 哈佛经济研究所研讨论文 no. 1805, Cambridge, Mass. , 1997. 11.

19. James J. Choi, David Laibson, and Brigitte Madrian, "Why Does the Law of One Price Fail? And Experiment on Index Mutual Funds," 美国经济研究局研讨论文 no. W12261, Cambridge, Mass. , 2006. 5. 4.

20. Leonard Koppett, "Carrying Statistics to Extremes," Sporting News, 1978. 2. 11.

21. 根据某些定义，科佩特的系统会被认为在 1970 年预测失败；而根据另一些定义，它在当年的预测仍然是成功的. 请见 CHANCE News 13. 04, 2004. 4. 18 ～ 6. 7, http://www. dartmouth. edu/～chance/chance _ news/recent _ news/chance _ news _ 13. 04. html.

22. 正如在 Legg Mason Capital Management 的 Web 网页上被吹捧的那样: http://www. leggmasoncapmgmt. com/awards. htm.

23. Lisa Gibbs, "Miller: He Did It Again," CNNMoney, 2004. 1. 11. , http://money. cnn. com/2004/01/07/funds/ultimateguide _ billmiller _ 0204.

24. Thomas R. Gilovich, Robert Vallone, and Amos Tversky, "The Hot Hand in Basketball: On the Misperception of Random Sequences," Cognitive Psychology, 17, no. 3, 1985. 7. : 295～314.

25. 普塞尔的研究被讨论于 Gould, "The Streak of Streaks."

26. Mark Hulbert, "Not All Stocks Are Created Equal," www. MarketWatch. com, 2005. 1. 18. , 2005 年 3 月浏览, 网站现已停止运行.

27. Kunal Kapoor, "A Look at Who's Chasing Bill Miller's Streak," Morningstart, 2004. 12. 30. , http://www. morningstar. com.

28. Michael Mauboussin and Cristen Bartholdson, "On Streaks: Perception, Probability, and Skill," Consilient Observer, 瑞士信贷第一波士顿公司, 2, no. 8, 2003. 4. 22.

29. Merton Miller 于 "Trillion Dollar Bet," NOVA, PBS 广播, 2000. 2. 8.

30. R. D. Clarke, "An Application of the Poisson Distribution," Journal of the Institute of Actuaries, 72, 1946: 48.

31. Atul Gawande, "The Cancer Cluster Myth," The New Yorker, 1998. 2. 28. , pp. 34～37.

32. Atul Gawande, "The Cancer Cluster Myth," The New Yorker, 1998. 2. 28. , pp. 34～37.

33. Bruno Bettelheim, "Individual and Mass Behavior in Extreme Situations," Journal of Abnormal and Social Psychology, 38, 1943: 417～452.

34. Curt P. Richter, "On the Phenomenon of Sudden Death in Animals and Man," Psychosomatic Medicine, 19, 1957: 191～198.

35. E. Stotland and A. Blumenthal, "The Reduction of Anxiety as a Result of the Expectation of Making a Choice," Canadian Review of Psychology, 18, 1964: 139～145.

36. Ellen Langer and Judith Rodin, "The Effects of Choice and Enhanced Personal Responsibility for the Aged: A Field Experiment in an Insitutional Setting," Journal of Personality and Social Psychology, 34, no. 2, 1976: 191～198.

37. Ellen Langer and Judith Rodin, "Long-Term Effects of a Control-Relevant Intervention with the Instutionalized Aged," Journal of Personality and Social Psychology, 35, no. 12, 1977: 897～902.

38. L. B. Alloy and L. Y. Abramson, "Judgement of Contingency in Depressed and Nondepressed Students: Sadder but Wiser?" Journal of Experimental Psychology: General, 108, no. 4, 1979. 12.: 441～485.

39. Durgin, "The Tinkerbell Effect."

40. Ellen Langer, "The Illusion of Control," Journal of Personality and Social Psychology, 32, no. 2, 1975: 311～328.

41. Ellen Langer and Jane Roth, "Heads I Win, Tails It's Chance: The Illusion of Control as a Function of Outcomes in a Purely Chance Task," Journal of Personality and Social Psychology, 32, no. 6, 1975: 951～955.

42. Langer, "The Illusion of Control."

43. Langer, "The Illusion of Control." p. 311.

44. Raymond Fisman, Rakesh Khurana, and Matthew Rhodes-Kropf, "Governance and CEO Turnover: Do Something or Do the Right Thing?" 哈佛商学院工作论文 no. 05～066, Cambridge, Mass., 2005. 4.

45. P. C. Wason, "Reasoning about a Rule," Quarterly Journal of Experimental Psychology, 20, 1968: 273～281.

46. Francis Bacon, Novum Organon, P. Urbach and J. Gibson 英译本, Chicago: Open Court, 1994, p. 57 (最初出版于 1620 年).

47. Charles G. Lord, Lee Ross, and Mark Lepper, "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence," Journal of Personality and Social Psychology, 37, no. 11, 1979: 2098～2109.

48. Matthew Rabin, "Psychology and Economics", 加州大学伯克利分校白皮书, 1996. 9. 28.

49. E. C. Webster, Decision Making in the Employment Interview, Montreal: Industrial Relations Centre, McGill University, 1964.

50. Beth E. Haverkamp, "Confirmatory Bias in Hypothesis Testing for Client-Identified and Counselor Self-generated Hypotheses," Journal of Counseling Psychology, 40, no. 3, 1993. 7. : 303～315.

51. David L. Hamilton and Terrence L. Rose, "Illusory Correlation and the Maintenance of Stereotypic Beliefs," Journal of Personality and Social Psychology, 39, no. 5, 1980: 832～845; Galen V. Bodenhausen and Robert S. Wyer, "Effects of Stereotypes on Decision Making and Information-Processing Strategies," Journal of Personality and Social Psychology, 48, no. 2, 1985; 以及 C. Stangor and D. N. Ruble, "Strength of Expectancies and Memory for Social Information: What We Remember Depends on How Much We Know," Journal of Experimental Social Psychology, 25, no. 1, 1989: 18～35.

第十章 醉汉的脚步

1. Pierre-Simon de Laplace, 转引自 Stigler, Statistics on the Table, p. 56.

2. James Gleick, Chaos: Making a New Science, New York: Penguin, 1987; 请见第一章。

3. Max Born, Natural Philosophy of Cause and Chance, Oxford: Clarendon Press, 1948, p. 47. 玻恩谈的是一般意义的自然，但特别谈论的是量子理论。

4. 关于珍珠港事件的分析出自 Roberta Wohlstetter, Pearl Harbor: Warning and Decision, Palo Alto, Calif. : Stanford University

Press，1962。

5. Richard Henry Tawney, The Agrarian Problem in the Sixteenth Century, 1912; New York: Burt Franklin, 1961 年重印本。

6. Wohlstetter, Pearl Harbor, p. 387.

7. 三里岛上所发生的事情之描述出自 Charles Perrow, Normal Accidents: Living with High-Risk Technologies, Princeton, N. J.: Princeton University Press, 1999; 与 U. S. Nuclear Regulatory Commission, "Fact Sheet on the Three Mile Island Accident," http://www.nrc.gov/reading-rm/doc-collections/fact-sheets/3mile-isle.html.

8. Perrow, Normal Accidents.

9. W. Brian Arthur, "Positive Feedbacks in the Economy," Scientific American, 1990. 2., pp. 92～99.

10. Matthew J. Salganik, Peter Sheridan Dodds, and Duncan J. Watts, "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market," Science, 311, 2006. 2. 10.; 与 Duncan J. Watts, "Is Justin Timberlake a Product of Cumulative Advantage?" New York Times Magazine, 2007. 4. 15.

11. Mlodinow, "Meet Hollywood's Latest Genius."

12. John Steele Gordon and Michael Maiello, "Pioneers Die Broke," Forbes, 2002. 12. 23; 与 "The Man Who Could Have Been Bill Gates," BusinessWeek, 2004. 10. 25.

13. Floyd Norris, "Trump Deal Fails, and Shares Fall Again," New York Times, 2007. 7. 6.

14. Melvin J. Lerner and Leo Montada, "An Overview: Advances in Belief in a Just World Theory and Methods," in Responses to Victimizations and Belief in a Just World, ed. Leo Montada and Melvin J. Lerner, New York: Plenum Press, 1998, pp. 1～7.

15. Melvin J. Lerner, "Evaluation of Performance as a Function

of Performer's Reward and Attractiveness," Journal of Personality and Social Psychology, 1, 1965: 355～360.

16. Melvin J. Lerner and C. H. Simmons, "Observer's Reactions to the 'Innocent Victim': Compassion or Rejection?" Journal of Personality and Social Psychology, 4, 1966: 203～210.

17. Lerner, "Evaluation of Performance as a Function of Performer's Reward and Attractiveness."

18. Wendy Jean Harrod, "Expectations from Unequal Rewards," Social Psychology Quarterly, 43, no. 1, 1980. 3.: 126～130; Penni A. Stewart and James C. Moore Jr., "Wage Disparities and Performance Expectations," Social Psychology Quarterly, 55, no. 1, 1992. 3.: 78～85; 以及 Karen S. Cook, "Expectations, Evaluations and Equity," American Sociology Review, 40, no. 3, 1975. 6.: 372～388.

19. Lerner and Simmons, "Observer's Reactions to the 'Innocent Victim.'"

20. David L. Rosenhan, "On Being Sane in Insane Places," Science, 179, 1973. 1. 19.: 237～255.

21. Elisha Y. Babad, "Some Correlates of Teachers' Expectancy Bias," American Educational Research Journal, 22, no. 2, 1985 年夏: 175～183.

22. Eric Asimov, "Spirits of the Times: A Humble Old Label Ices Its Rivals," New York Times, 2005. 1. 26.

23. Jonathan Calvert and Will Iredale, "Publishers Toss Booker Winners into the Reject Pile," London Sunday Times, 2006. 1. 1.

24. Peter Doskoch, "The Winning Edge," Psychology Today, 2005. 11/12., p. 44.

25. Rosenhan, "On Being Sane in Insane Places," p. 243.

醉汉的脚步

The Drunkard's Walk: How Randomness Rules Our Lives



你知道吗，在去买彩票的路上因车祸身亡的可能性，是彩票中奖的可能性的两倍！打破了贝比·鲁斯本垒打纪录的奇人罗杰·马立斯，也非常可能仅仅是幸运而非伟大！一种葡萄酒被某本刊物给予了五颗星的最高评分，却被另一本刊物评为一个年代中最差的葡萄酒，这是怎么回事？

在这本颠覆常识又具有启蒙性的书中，列纳德·蒙洛迪诺通过解开偶然性的真实本性，以及导致我们误判周遭世界的那些心理错觉，生动地展示了什么是真正有意义的东西，而我们又如何才能在一个更深层次真理的基础上，来进行我们的决策。

本书带给你的，不仅是在随机性、偶然性和概率中的一次漫游，还是一个看待世界的全新视角。它同时提醒着我们，生活中的许多事情，大致就如同刚在酒吧待了一夜的家伙那蹒跚的步履一般难以预测。

责任编辑／吴炜 孙桂均
文字编辑／陈一心
整体设计／刘苏斌

ISBN 978-7-5357-6225-2
9 "787535" 762252" >
定价：35.00 元