

10堂极简

TEN GREAT
ABOUT CHANCE

概 率 课

关于概率，
你不能不知道的
10个伟大思想

[美] 佩尔西·戴康尼斯
(Persi Diaconis)

/ 著

[美] 布赖恩·斯科姆斯
(Brian Skyrms)

胡小锐 / 译

TEN GREAT IDEAS
ABOUT CHANCE

10 堂极简 概率课

TEN GREAT IDEAS
ABOUT CHANCE

[美] 佩尔西·戴康尼斯
(Persi Diaconis) / 著 胡小锐 / 译
[美] 布赖恩·斯科姆斯
(Brian Skyrms)

图书在版编目 (CIP) 数据

10 堂极简概率课 / (美) 佩尔西·戴康尼斯, (美) 布赖恩·斯科姆斯著; 胡小锐译. --北京: 中信出版社, 2019.4 (2019.8 重印)

书名原文: TEN GREAT IDEAS ABOUT CHANCE
ISBN 978-7-5086-9920-2

I. ①1… II. ①佩… ②布… ③胡… III. ①概率论—普及读物 IV. ①O211-49

中国版本图书馆CIP数据核字(2019)第009811号

Ten Great Ideas About Chance by Persi Disconis & Brian Skyrms

Copyright © 2018 by Princeton University Press

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without permission in writing from the Publisher.

Simplified Chinese translation copyright © 2019 by CITIC Press Corporation

ALL RIGHTS RESERVED

本书仅限中国大陆地区发行销售

10 堂极简概率课

著 者: [美] 佩尔西·戴康尼斯 [美] 布赖恩·斯科姆斯

译 者: 胡小锐

出版发行: 中信出版集团股份有限公司

(北京市朝阳区惠新东街甲4号富盛大厦2座 邮编 100029)

承 印 者: 中国电影出版社印刷厂

开 本: 880mm × 1230mm 1/32

印 张: 9.5 字 数: 200千字

版 次: 2019年4月第1版

印 次: 2019年8月第2次印刷

京权图字: 01-2018-4678

广告经营许可证: 京朝工商广字第8087号

书 号: ISBN 978-7-5086-9920-2

定 价: 49.00元

版权所有·侵权必究

如有印刷、装订问题, 本公司负责调换。

服务热线: 400-600-8099

投稿邮箱: author@citicpub.com

前言

这本书是由我们在斯坦福大学合作教授了约10年的一门课程衍生而来的。这是一门大型的混合性课程，听课的人是本科生或研究生，他们分别来自哲学、统计学和一些交叉学科。随着课程的不断发展，我们越来越相信它的内容应该可以吸引更多的听众。学习这门课的一个先决条件，就是接触过一门概率论或统计学的课程，这本书的读者同样需要满足这个条件。但是，考虑到某些读者可能是在很久以前学过这类课程，我们在书中以附录的形式，对概率论进行了一次简要的复习。

这本书涉及的内容包括历史、概率和哲学。我们不仅介绍了概率论发展过程中的一些伟大思想及其历史，还致力于探索这些思想的哲学意义。一位阅读过本书初稿的读者抱怨说，读到最后，他仍然不了解我们关于概率的哲学观点，原因或许是我们过于中立。这个问题现在已经解决了，你会发现我们是彻头彻尾的贝叶斯学派，是贝叶斯（Thomas Bayes）、拉普拉斯（Pierre-Simon Laplace）、拉姆齐（Frank Ramsey）和菲尼蒂（Bruno de Finetti）的信徒。有人认为贝叶斯学派是与频率学派相对立的，而我们并不否认频率的重要性，或者讨论客观概率的价值。不仅如此，我们还会在合理的置信度框架内统一考虑这些问题。

在这本书的开头，我们与先驱者一起思考，涉及的工具很简单。但到了后半部分，我们将回到当下，不可避免地会接触到一些技术性细节。

为了保证行文简洁流畅，我们将把某些细节内容放到附录中，大家可以根据需要查阅。我们还做了大量注释，以方便读者深入挖掘自己感兴趣的内容。在这本书的最后，我们列出了一份参考书目。此外，脚注也给出了较为详细的解释。

佩尔西·戴康尼斯

布赖恩·斯科姆斯

目录 | CONTENTS

VII 前言

第1课 概率是可以测度的

- 006 概率测度的开始
- 008 帕斯卡和费马
- 012 惠更斯
- 015 伯努利
- 017 小结
- 018 附录 1 帕斯卡和费马
- 021 附录 2 抛硬币的物理学原理
- 025 附录 3 巧合与生日问题

第2课 相关性判断就是概率

- 033 部分 I：赌博与判断概率
- 044 部分 II：效用与判断概率
- 057 小结
- 058 附录 1 条件赌注的相关性
- 060 附录 2 概率运动学

第3课 概率心理学不同于概率逻辑学

- 071 启发法和偏见
- 073 框架
- 076 小结

- 077 附录 1 埃尔斯伯格：有序性还是独立性？
- 079 附录 2 动态一致性与阿莱

第 4 课 频率与概率之间有什么关系？

- 084 雅各布·伯努利与弱大数定律
- 085 伯努利骗局与频率主义
- 088 伯努利骗局与假设检验
- 089 频率学派的中坚力量
- 098 对理想化方法的再思考
- 100 小结

第 5 课 如何用数学方法解决概率问题？

- 103 在数学与现实之间 I
- 104 有限集的概率
- 105 集合的长度与概率
- 109 希尔伯特的第 6 个问题
- 110 柯尔莫哥洛夫的贡献
- 111 把概率论视为数学的一个分支
- 113 把条件概率视为随机变量
- 115 从有限维到无限维
- 116 在数学和现实之间 II
- 117 随机选择的整数？数学的旁白
- 122 柯尔莫哥洛夫对概率空间的有穷性的看法
- 123 小结
- 124 附录 1 复杂集的测度
- 125 附录 2 不可测集

第 6 课 贝叶斯定理如何改变了世界？

- 131 贝叶斯 vs 休谟
- 134 贝叶斯的概率研究
- 137 反演问题与台球桌
- 140 拉普拉斯的玩笑
- 141 广义的拉普拉斯定律
- 144 相容性
- 145 为什么公开发表的研究结果大多是错的？
- 148 贝叶斯、伯努利和频率
- 148 改变世界
- 149 小结
- 150 附录 贝叶斯关于概率和统计学的思考

第 7 课 菲尼蒂定理与可交换概率

- 157 菲尼蒂的论著
- 158 有限可交换序列
- 161 菲尼蒂定理与一般可观测量
- 163 菲尼蒂定理与正态分布
- 165 马尔可夫链
- 166 部分可交换性
- 167 小结
- 168 附录 1 遍历理论——菲尼蒂定理的推广
- 169 附录 2 菲尼蒂可交换定理

第 8 课 如何用图灵机生成随机序列？

- 183 随机数生成器
- 187 随机算法理论
- 190 可计算性

- 197 马丁-洛夫随机序列
- 203 随机性的变化
- 204 小结

第9课 世界的本质是什么？

- 210 玻尔兹曼
- 216 概率、频率和遍历性
- 216 冯·诺依曼和伯克霍夫的遍历性研究
- 219 庞加莱
- 222 遍历性的层次结构
- 223 玻尔兹曼归来
- 224 量子力学
- 225 非定域性
- 229 量子概率归来
- 230 量子混沌
- 233 小结
- 234 附录 量子形而上学：窥视潘多拉的盒子

第10课 如何用概率论解答休谟问题？

- 240 休谟
- 241 康德
- 242 波普尔
- 244 归纳怀疑论的不同等级
- 244 贝叶斯-拉普拉斯
- 248 无知如何量化？
- 250 概率是否存在？
- 251 如果置信度不可交换，会怎么样？
- 252 那些用来描述世界的谓词呢？

254 如何看待不确定性证据呢？

256 小结

附 录 **概率辅导课**

259 符号：把事情记录下来

261 案例：非传递性悖论

263 基本事实：游戏规则

268 随机变量和期望

271 条件期望和鞅

273 案例：波利亚的罐子

275 从离散到连续再到更大空间

276 计算机登场！

277 致 谢

279 注 释

1

第

课

概率是可以测度的



吉罗拉莫·卡尔达诺 (Gerolamo Cardano)

要搞清楚一门学科的本质，认真研究该学科的开创者的想法是一条可行的路径。事实上，某些基础性哲学问题从一开始就是显而易见的。关于概率，我们的第1堂课要介绍的第一个伟大思想是：概率是可以测度的。这个观点的形成时间是16—17世纪，过程为何如此漫长，这个问题至今仍然是一个谜。希腊神话中有命运女神堤喀（Tyche）；德谟克利特（Democritus）及其追随者假设，构建宇宙的所有原子都会受到某种物质偶然性的影响；卢克莱修（Lucretius）在《物性论》（*De Rerum Natura*）中指出，这种偶然性就是原子的偏离；古埃及人和古巴比伦人学会了用指关节骨或骰子玩概率游戏，到了罗马时期，这种游戏流行开来，士兵们通过抽签决定基督斗篷的归属。后来，古希腊学园派怀疑论者将概率视为人生的指南。¹不过，这些时期似乎都没有出现有关概率的定量理论。²

想一想，我们是怎么测量东西的？³以长度为例，我们会先找到一个长度标准，然后计数某个东西包含多少个这样的标准长度。比如，在我们用脚步测量距离时，这个长度标准就是我们的脚。但是，不同的脚有可能得出不同的测量结果，因此，1522年，有人提议改进法定路德（杆）的确定方法。如图1-1所示，当人们从教堂鱼贯而出时，将排成一列的



图 1-1 法定路德的确定

16个人的脚的总长度设定为法定路德。⁴从图1-1可以看出，这些人的脚长度不一，但通过一群人来设定这个长度单位具有明显的平均效应，因此很多人接受了这个方法。不过，当时似乎还没有人明确提出平均数这个概念。

我们有必要指出，这个方法存在哲学上的异议。我们的目的是定义长度，但在用脚长测量距离时，我们已经假定我们采用的长度标准等长。⁵因此，这是一个循环论证的过程。

任何有头脑的人都不会因为这个异议而放弃用脚长测量距离的方法。我们的测量活动就始于此，最终建立并完善了长度的概念。脚的长度因人而异，路德会长短不一，标准米尺的长度在足够高的精度条件下也会各不相同。借助物理学知识，我们可以不断改进长度测量方法。因此，这确实是一个循环论证的过程，但它并不是一个致命的缺陷，反而为我

们指明了一条趋于完善的道路。^①

概率的测度同样如此。在测度之前，我们先要找到（或者制造）同等可能性的情况，然后计数这些情况发生的次数。于是，事件 A 的概率，记作 $P(A)$ ，为

$$P(A) = \frac{\text{事件 } A \text{ 发生的次数}}{\text{所有可能发生的事件的次数}}$$

注意，从上式可知：

1. 概率永远不会是负值；
2. 如果所有可能发生的情况中均包含事件 A ，则 $P(A) = 1$ ；
3. 如果事件 A 和事件 B 不会同时发生，则 $P(A \text{ 或 } B) = P(A) + P(B)$ 。

此外，某个事件不会发生的概率等于1与该事件发生概率的差：

$$P(\text{非 } A) = 1 - P(A)。$$

这个概念虽然十分简单，但如果运用得巧妙得当，就会产生令人惊讶的效果。我们以生日问题为例。如果不考虑闰年，并且假设出生日期的概率均相等，每个人的生日相互独立（即没有双胞胎），那么房间内的所有人中至少有两个人的生日在同一天的概率是多少？如果你以前没有见过这个问题，它的答案肯定会让你大吃一惊。

一群人中有人生日在同一天的概率等于1减去所有人生日均不相

① 促使等概率这个概念逐步完善的道路是什么？继续阅读这本书，你就会找到答案。

同的概率。第二个人与第一个人的生日不同的概率为 $(364/365)$ 。如果前两个人的生日不同，那么第三个人的生日与他们俩都不相同的概率为 $(363/365)$ ，以此类推。因此， N 个人中有人生日相同的概率为

$$1 - \left(\frac{364}{365} \times \frac{363}{365} \times \cdots \times \frac{365 - N + 1}{365} \right)$$

如果你对同额赌注感兴趣，就可以利用上述公式，找到使输赢概率趋近 $1/2$ 的 N 的值。当房间里一共有 23 个人时，生日相同的概率会略高于 $1/2$ 。如果房间里有 50 个人，这个概率就会接近 97%。

人们经常利用生日问题来考虑一些令人吃惊的巧合情况，因此生日问题出现了很多变种。比如，两个美国人的生日相同，而且他们的父亲、祖父和曾祖父生日相同的可能性大到令人吃惊的程度。为帮助大家应对这些问题，本堂课内容的附录部分给出了一些有用的近似值。最后，本书在结尾部分又利用这些近似值，证明了菲尼蒂定理。现在，大家只要知道“等可能情况”这个基本结构应用广泛和深入就可以了。

概率测度的开始

创建等概率情况，最有效的方法莫过于抛掷质地均匀的骰子，或者从洗好的一副牌中抽取扑克牌。概率的测度就是从这里开始的。我们不知道首创者是谁，但数学家、医生、占星家吉罗拉莫·卡尔达诺早在 16 世纪研究赌博游戏时就明明白白地提到过这个概念。⁶ 卡尔达诺有时以赌博为生，因此他对等概率假设非常敏感。此外，他对动过手脚的骰子以及其他作弊手法都了如指掌：“……骰子有时并不诚实，可能是因为它

被打磨过，也可能是因为它被削扁了（这很容易被人看穿），还可能是因为相对应的两面受到挤压而变得扁平了……牌类游戏的作弊手段更是层出不穷。”⁷

17世纪早期，伽利略（Galileo）给他的赞助人托斯卡纳大公爵写了一封简短的信，回答了后者提出的一个关于骰子的问题。公爵认为，通过计算可能情况得出的答案似乎是错误的。投掷三枚骰子时，得到10点和11点的数字组合方式各有6种，9点和12点同样如此。“……但是，众所周知，骰子玩家通过长期观察发现，掷出10点和11点的可能性比9点和12点的可能性更大。”^①这是怎么一回事呢？

伽利略答道，他的赞助人在计算得到9点和10点的可能情况时，把三个3点计作一种可能，把两个3点和一个4点也计作一种可能，这种方法是错误的。伽利略指出，后者涵盖了三种可能的组合，它们彼此之间的不同点就在于是哪枚骰子掷出了4点。

$\langle 4, 3, 3 \rangle$, $\langle 3, 4, 3 \rangle$, $\langle 3, 3, 4 \rangle$ 。

前者的确只有一种可能，即 $\langle 3, 3, 3 \rangle$ 。伽利略完全掌握了排列组合的相关知识，似乎并没有觉得这是什么新鲜事物。

在构建等概率情况时，伽利略和卡尔达诺似乎都隐晦地使用了独立性这个概念。他们认为，对于每一枚骰子，抛掷后得到6个面中的每一个的概率都相等，在抛掷三枚骰子时，得到216个可能结果中的每一个的

① 这个问题的表述有一个奇怪的地方，那就是对长期观察这个表达的解释。观察必须持续进行很长时间。根据伽利略的计算，得到9点的概率是 $25/216$ ，约等于0.116；得到10点的概率是 $27/216$ ，约等于0.125。两者相差0.009，约等于 $1/100$ 。大家可以计算一下需要观察的次数，当作一次练习。