

国际名校经典教材

[英] 戴维·威廉姆斯 (剑桥大学统计实验室) 著
郑坚坚 译

概率和鞅

Probability with Martingales

学技术大学出版社

国际名校经典教材

[英] 戴维·威廉姆斯 (剑桥大学统计实验室) 著
郑坚坚 译

概率和鞅

Probability with Martingales

中国科学技术大学出版社

安徽省版权局著作权合同登记号:第 12171794 号

Probability with Martingales, first edition by David Williams

first published by Cambridge University Press 1991.

All rights reserved.

This simplified Chinese edition for the People's Republic of China is published by arrangement with Cambridge University Press.

© Cambridge University Press & University of Science and Technology of China Press 2018

This book is in copyright. No reproduction of any part may take place without the written permission of Cambridge University Press and University of Science and Technology of China Press.

This edition is for sale in the People's Republic of China (excluding Hong Kong SAR, Macau SAR and Taiwan Province) only.

此版本仅限在中华人民共和国境内(不包括香港、澳门特别行政区及台湾地区)销售。

图书在版编目(CIP)数据

概率和鞅/(英)戴维·威廉姆斯(David Williams)著;郑坚译. —合肥:中国科学技术大学出版社,2018.3

书名原文:Probability with Martingales

ISBN 978-7-312-04368-0

I. 概… II. ①戴… ②郑… III. 概率论 IV. ①O21 ②O211

中国版本图书馆 CIP 数据核字(2017)第 294420 号

出版 中国科学技术大学出版社
安徽省合肥市金寨路 96 号,230026
<http://press.ustc.edu.cn>
<https://zgkxjstxcbs.tmall.com>
印刷 安徽国文彩印有限公司
发行 中国科学技术大学出版社
经销 全国新华书店
开本 710 mm×1000 mm 1/16
印张 18.25
字数 388 千
版次 2018 年 3 月第 1 版
印次 2018 年 3 月第 1 次印刷
定价 48.00 元

序

请阅读!

本书最重要的章节是 E 章: 练习题 (Exercises). 这是我留给你们做的堪称有趣的事情. 你们可以现在就开始做其中的“EG”练习题, 但需要看一下本序言后文中的“更多关于习题”.

本书 (它基本上是为剑桥大学的三年级学生所开设的课程的讲义) 是一本关于概率的严格理论的尽可能生动 (就我的能力而言) 的入门书. 由于其相当大的篇幅被用来论述鞅, 故它也应该是充满活力的: 请看第 10 章的那些练习题! 当然, 在涉足测度论基础时, 开始总会有深奥与滞重的感觉. 然而必须强调: 测度论的那些就其自身而言最枯燥的内容一旦应用于概率论, 便会变得非常活泼而具有生气. 这不仅仅是因为它得到了应用, 同时也由于其自身的内涵得到了极大的丰富.

你无法回避测度论: 概率论中的一个事件是一个可测集, 一个随机变量是样本空间上的可测函数, 一个随机变量的期望是其关于概率测度的积分, 等等. 当然, 也可以将测度论的一些主要结果作为不证自明的结论放在正文里, 而在附录中给出它们的详细证明; 这也恰好是我在本书中所做的.

测度论就其本身的结构而言是建立在测度的加法法则基础之上的. 概率论则以乘法法则对它加以补充, 后者是刻画独立性的. 事情已经得到改善. 但真正使内容得到充实并充满活力的是我们已能处理许多 σ -代数, 而不是像在测度论中那样, 只涉及一个 σ -代数.

在准备这本书的过程中, 我曾考虑过所有当时我认为多少有一些新意的论题, 但是对于其中的大部分最终我不得不忍痛割爱.

若想对本书中的许多议题有更深入的了解, 可以参考 Billingsley (1979)、Chow 和 Teicher (1978)、Chung (1968)、Kingman 和 Taylor (1966)、Laha 和 Rohatgi (1979) 以及 Neveu (1965) 等人的著作. 至于测度论, 我当初学习它是通过阅读 Dunford 和 Schwartz (1958) 以及 Halmos (1959) 等著作. 在读

完本书之后，你应该去阅读在现今仍属精彩的 Breiman (1968) 的书，而为了在离散鞅论方面获得精辟的指点，你应该去看 Hall 和 Heyde (1980) 的书。

当然，直觉要比测度论知识重要得多，你应该利用每一个机会来磨砺你的直觉。没有比 Aldous (1989) 更好的“磨刀石”了，尽管它是一本非常难读的书。为了开阔概率论的视野并学习如何去思考其相关问题，我强力推荐 Karlin 和 Taylor (1981)、Grimmett 和 Stirzaker (1982)、Hall (1988) 等的著作，还有近期 Grimmett (1989) 关于过滤的极好的书。

更多关于习题. 在编写 E 章（它刚好是由我平时布置给剑桥大学学生的家庭作业题所构成）的时候我考虑了这样一点：像其他任何一本数学书一样，本书也暗含着一大批另外的练习题，而其中的很多题都要比 E 章中的题容易。当然，我更欣赏的还是你们在阅读有关结论之后能自创练习题，并且尝试亲自去证明它们，而不是先去找已有的证明。关于习题还需加以说明的一点是希望你们能够原谅我的地方，诸如我在第 4 章的练习题中使用了数学期望 E ，而 E 的严格定义却要到第 6 章才出现之类。

致谢. 首先我要感谢那些坚持听我的课（它是本书的基础）的学生们，他们出众的才华与素养促使我努力改进本课，使之能与他们的水平相称。感谢那些在我之前就讲授了本课程的教师们，尤其是 David Kendall。感谢 David Tranah 和剑桥大学出版社的其他员工，他们帮助我将该课程变成了这本书。我还要感谢 Ben Garlin、James Norris 和 Chris Rogers，若没有他们，本书会包含更多的错误与含混之处。（本书残存的错误则应由我本人负责。）Helen Rutherford 和我打出了部分文稿，但本书的绝大部分文稿是由 Sarah Shea-Simonds 以娴熟的、堪与 Horowitz（钢琴家——译者注）相比的优美指法打出来的。感谢 Helen，尤其要感谢 Sarah。还要特别感谢我的妻子 Sheila 的全力相助。

但是我的最大感谢（也是你们的，若你们从本书中有所获益的话）应该献给三个其姓名以大写字母出现在本书索引中的人 J. L. 杜布、A. N. 柯尔莫哥洛夫和 P. 莱维：如果没有他们的工作，则当杜布在其著作（1953）中作精彩总结时将没有多少内容可写。

戴维·威廉姆斯

1990 年 10 月于剑桥大学统计实验室

一个术语问题 ——随机变量：是函数还是等价类？

按本书的水平，若将一个随机变量视为样本空间上一个可测函数的等价类，则有关的理论会变得更“漂亮”，两个函数属于相同的等价类当且仅当它们几乎处处相等。那样的话，条件期望映射

$$X \mapsto E(X|\mathcal{G})$$

将成为一个真正定义明确的从 $L^p(\Omega, \mathcal{F}, P)$ 到 $L^p(\Omega, \mathcal{G}, P)$ ($p \geq 1$) 的压缩映射；同时我们也不必总是提及 (随机变量) 不同的版本 (等价类的不同代表元)，且可以避免那些没完没了的“几乎必然”之类的限制条件。

然而我选择了“不漂亮”的方式：首先，我更喜欢处理函数，例如我更喜欢

$$4 + 5 = 2 \pmod{7},$$

而不是

$$[4]_7 + [5]_7 = [2]_7.$$

但是还有一个实质性的原因。我希望本书能吸引读者将来深入了解更加有趣，也更加重要的理论，其中我们的随机过程的参数集合是不可数的 (例如时间参数集合可能是 $[0, \infty)$)。在那里，等价类的构想是不适用的：引进商空间的所谓“聪明”会丧失那种即便是建立基本的、基于连续变化的结果也必须具有的精细与灵敏度，还有其他一些东西，除非你想表演那种扭曲的“柔身术”，然而其结果却很难做到完美。即便这些柔身术允许你建立某些结果，但是要证明它们却仍然要使用真正的函数。所以，其现实性究竟何在？！

符号说明

► 表示重要的事, ►► 非常重要的事, ►►► 诸如鞅收敛定理.

我用 “:=” 来表示 “定义为”. 这一帕斯卡 (Pascal) 符号用起来特别方便, 因为它同时也具有反向的意义.

我使用分析论者 (相对于范畴论者) 的约定:

$$\blacktriangleright \quad \mathbf{N} := \{1, 2, 3, \dots\} \subseteq \{0, 1, 2, \dots\} =: \mathbf{Z}^+.$$

另外众所周知: $\mathbf{R}^+ := [0, \infty)$.

对于一个包含于某全集 S 中的集合 B , I_B 表示 B 的示性函数, 即 $I_B: S \rightarrow \{0, 1\}$ 且

$$I_B(s) := \begin{cases} 1, & \text{若 } s \in B, \\ 0, & \text{否则.} \end{cases}$$

对于 $a, b \in \mathbf{R}$,

$$a \wedge b := \min(a, b), \quad a \vee b := \max(a, b).$$

CF: 特征函数; DF: 分布函数; pdf: 概率密度函数.

σ -代数: $\sigma(\mathcal{C})(1.1)$; $\sigma(Y_\gamma: \gamma \in \mathcal{C})(3.8, 3.13)$. π -系 (1.6); d -系 (A1.2).

a.e.: 几乎处处, 几乎每点 (1.5).

a.s.: 几乎必然 (2.4).

$b\mathcal{S}$: 有界 Σ -可测函数空间 (3.1).

$\mathcal{B}(S)$: S 上的博雷尔 σ -代数, $\mathcal{B} := \mathcal{B}(\mathbf{R})(1.2)$.

$C \bullet X$: 离散随机积分 (10.6).

$d\lambda/d\mu$: 拉东 - 尼可丁导数 (5.14).

dQ/dP : 似然比 (14.13).

$E(X)$: X 的期望 $E(X) := \int_{\Omega} X(\omega)P(d\omega)$ (6.3).

$E(X;F)$: $\int_F X dP$ (6.3).

$E(X/\mathcal{G})$: 条件期望 (9.3).

(E_n, ev) : $\liminf E_n$ (2.8).

$(E_n, \text{i.o.})$: $\limsup E_n$ (2.6).

f_X : X 的概率密度函数 (pdf) (6.12).

$f_{X,Y}$: 联合 pdf (8.3).

$f_{X|Y}$: 条件 pdf (9.6).

F_X : X 的分布函数 (3.9).

\liminf : 关于集合的 (2.8).

\limsup : 关于集合的 (2.6).

$x = \uparrow \lim x_n$: $x_n \uparrow x$, 即满足: $x_n \leq x_{n+1} (\forall n)$ 且 $x_n \rightarrow x$.

\log : 自然对数 (以 e 为底的).

\mathcal{L}_X, Λ_X : X 的分布 (3.9).

\mathcal{L}^p, L^p : 勒贝格空间 (6.7, 6.13).

Leb: 勒贝格测度 (1.8).

$m\Sigma$: Σ -可测函数空间 (3.1).

M^T : 在时间 T 停止的过程 M (10.9).

$\langle M \rangle$: 尖括号过程 (12.12).

$\mu(f)$: f 的关于 μ 的积分 (5.0, 5.2).

$\mu(f; A)$: $\int_A f d\mu$ (5.0, 5.2).

φ_X : X 的特征函数 (第 16 章).

φ : 标准正态分布 $N(0, 1)$ 的 pdf.

Φ : $N(0, 1)$ 的分布函数.

X^T : 在时间 T 停止的 X (10.9).

目 录

序	i
一个术语问题	iii
符号说明	iv
第 0 章 一个分支过程的例子	1
0.0 引言	1
0.1 典型孩子 (下一代) 的个数 X	1
0.2 第 n 代个体的数目 Z_n	2
0.3 利用条件期望	3
0.4 消亡概率 π	4
0.5 停下来思考: 测度	5
0.6 我们的第一个鞅	7
0.7 期望列的敛散性	8
0.8 求 M_∞ 的分布	9
0.9 具体的例子	10

A 部分 基 础

第 1 章 测度空间	14
1.0 引言	14
1.1 代数与 σ -代数的定义	15
1.2 例子: 博雷尔 (Borel) σ -代数, $\mathcal{B}(S), \mathcal{B} = \mathcal{B}(\mathbf{R})$	17
1.3 有关集函数的定义	18
1.4 测度空间的定义	19
1.5 有关测度的定义	19

1.6	引理: 扩张的唯一性, π -系	20
1.7	卡拉泰奥多里 (Carathéodory) 扩张定理	20
1.8	$((0, 1], \mathcal{B}(0, 1])$ 上的勒贝格 (Lebesgue) 测度 Leb	21
1.9	引理: 基本不等式	22
1.10	引理: 测度的单调收敛性	22
1.11	例子与告诫	23
第 2 章	事件	25
2.1	试验的模型: $(\Omega, \mathcal{F}, \mathbf{P})$	25
2.2	直观含义	25
2.3	序偶 (Ω, \mathcal{F}) 的例子	26
2.4	几乎必然 (a.s.)	27
2.5	提醒: \limsup , \liminf , $\downarrow\lim$, 等等	28
2.6	定义: $\limsup E_n, (E_n, \text{i.o.})$	29
2.7	博雷尔 - 肯泰利 (Borel-Cantelli) 第一引理 (BC1)	30
2.8	定义: $\liminf E_n, (E_n, \text{ev})$	30
2.9	练习	31
第 3 章	随机变量	32
3.1	定义: Σ -可测函数, $m\Sigma, (m\Sigma)^+, b\Sigma$	32
3.2	可测性基本命题	33
3.3	引理: 可测函数的和与积为可测	33
3.4	复合函数可测性引理	34
3.5	有关函数列的 \inf, \liminf 等的可测性引理	34
3.6	定义: 随机变量	35
3.7	例子: 掷币	35
3.8	定义: 由 Ω 上的函数族所产生的 σ -代数	36
3.9	定义: 分布, 分布函数	37
3.10	分布函数的性质	37
3.11	具有给定分布函数的随机变量的存在性	38
3.12	具有给定分布函数的随机变量的斯科罗霍德 (Skorokhod) 表示	38
3.13	生成的 σ -代数 (一个讨论)	40
3.14	单调类定理	41
第 4 章	独立性	42
4.1	独立性的定义	42
4.2	π -系引理; 更常见的定义	43

4.3	博雷尔 – 肯泰利第二引理 (BC2)	44
4.4	例	45
4.5	一个有关建模的基本问题	46
4.6	一个掷币模型及其应用	46
4.7	记号: IID 的随机变量 (RV_s)	48
4.8	随机过程: 马尔可夫 (Markov) 链	48
4.9	猴子敲出莎士比亚全集	49
4.10	定义: 尾 σ -代数	50
4.11	定理: 柯尔莫哥洛夫 (Kolmogorov) 0-1 律	51
4.12	练习与告诫	53
第 5 章	积分	54
5.0	符号及其他: $\mu(f) := \int f d\mu, \mu(f; A)$	54
5.1	非负简单函数的积分, SF^+	55
5.2	$\mu(f)$ ($f \in (m\Sigma)^+$) 的定义	56
5.3	单调收敛定理 (MON)	56
5.4	有关函数的法都 (Fatou) 引理 (FATOU)	58
5.5	“线性”	59
5.6	f 的正部与负部	59
5.7	可积函数, $\mathcal{L}^1(S, \Sigma, \mu)$	59
5.8	线性	60
5.9	控制收敛定理 (DOM)	60
5.10	谢菲 (Scheffé) 引理 (SCHEFFÉ)	61
5.11	关于一致可积性	62
5.12	标准机器	62
5.13	子集上的积分	62
5.14	测度 $f\mu$ ($f \in (m\Sigma)^+$)	63
第 6 章	期望	65
6.0	引言	65
6.1	期望的定义	65
6.2	收敛性定理	66
6.3	记号 $E(X; F)$	66
6.4	马尔可夫不等式	67
6.5	非负随机变量和	67
6.6	有关凸函数的詹森 (Jensen) 不等式	68

6.7	\mathcal{L}^p 范数的单调性	69
6.8	施瓦兹 (Schwarz) 不等式	70
6.9	\mathcal{L}^2 空间: 毕达哥拉斯 (Pythagoras) 定理、协方差及其他	71
6.10	\mathcal{L}^p 空间 ($1 \leq p < \infty$) 的完备性	73
6.11	正投影	74
6.12	期望的“初等公式”	76
6.13	从詹森不等式导出的赫尔德 (Hölder) 不等式	77
第 7 章	一个简单的强大数定律	79
7.1	“独立性意味着相乘”(又一例!)	79
7.2	强大数定律 (最初的版本)	80
7.3	切比雪夫 (Chebyshev) 不等式	81
7.4	魏尔斯特拉斯 (Weierstrass) 逼近定理	82
第 8 章	乘积测度	84
8.0	引言和建议	84
8.1	乘积可测结构 $\Sigma_1 \times \Sigma_2$	85
8.2	乘积测度, 富比尼 (Fubini) 定理	86
8.3	联合分布、联合概率密度函数	88
8.4	独立性与乘积测度	89
8.5	$\mathcal{B}(\mathbf{R})^n = \mathcal{B}(\mathbf{R}^n)$	89
8.6	n 重扩张	90
8.7	概率空间的无穷乘积	90
8.8	关于联合分布存在性的技术性注记	91

B 部分 鞅 论

第 9 章	条件期望	94
9.1	一个启发性例子	94
9.2	基本定理与定义 (柯尔莫哥洛夫, 1933)	95
9.3	直观意义	96
9.4	作为最小二乘最优预报的条件期望	96
9.5	定理 9.2 的证明	96
9.6	与传统表示的一致性	98
9.7	条件期望的性质 (一张表)	99
9.8	9.7 节中诸性质的证明	100
9.9	正则条件概率与概率密度函数	102

9.10	在独立性假设下的条件化	103
9.11	对称性的应用 (一个例子)	104
第 10 章	鞅	105
10.1	过滤的空间	105
10.2	适应的过程	105
10.3	鞅, 上鞅, 下鞅	106
10.4	鞅的一些例子	107
10.5	公平与不公平赌博	108
10.6	可料过程, 赌博策略	109
10.7	一个基本原则: 你无法改变这个系统!	109
10.8	停时	110
10.9	停止的上鞅仍是上鞅	111
10.10	杜布 (Doob) 可选停止定理	112
10.11	等待几乎必然要发生的事	113
10.12	简单随机游动的击中时	114
10.13	马尔可夫链的非负上调和函数	116
第 11 章	收敛定理	119
11.1	图说明一切	119
11.2	上穿	120
11.3	杜布上穿引理	121
11.4	推论	121
11.5	杜布“向前”收敛定理	122
11.6	告诫	123
11.7	推论	123
第 12 章	\mathcal{L}^2 中的有界鞅	124
12.0	引言	124
12.1	\mathcal{L}^2 中的鞅: 正交增量性	125
12.2	\mathcal{L}^2 中的零均值独立随机变量之和	126
12.3	随机信号	128
12.4	一个对称化技巧: 扩展样本空间	128
12.5	柯尔莫哥洛夫三级数定理	130
12.6	蔡查罗 (Cesàro) 引理	131
12.7	克罗内克 (Kronecker) 引理	131
12.8	方差约束下的强大数定律	132

12.9	柯尔莫哥洛夫截尾引理	133
12.10	柯尔莫哥洛夫强大数定律 (SLLN).....	134
12.11	杜布分解	135
12.12	尖括号过程 $\langle M \rangle$	136
12.13	$\langle M \rangle_\infty$ 有限时 M 相应的收敛性.....	136
12.14	一个有关 \mathcal{L}^2 中鞅的平凡“强大数定律”	138
12.15	博雷尔 - 肯泰利引理的莱维 (Lévy) 推广	139
12.16	评论	140
第 13 章	一致可积性	141
13.1	“绝对连续”性	141
13.2	定义: 一致可积族.....	142
13.3	一致可积性的两个简单的充分条件	143
13.4	条件期望的一致可积性	143
13.5	依概率收敛	144
13.6	有界收敛定理 (BDD) 的初等证明.....	145
13.7	\mathcal{L}^1 收敛的一个充要条件	146
第 14 章	一致可积 (UI) 鞅	148
14.0	引言.....	148
14.1	一致可积鞅	148
14.2	莱维的“向上”定理.....	149
14.3	柯尔莫哥洛夫 0-1 律的鞅证明.....	150
14.4	莱维的“向下”定理.....	151
14.5	强大数定律的鞅证明.....	152
14.6	杜布的下鞅不等式	153
14.7	重对数律: 特殊情形	154
14.8	有关正态分布的一个标准估计.....	156
14.9	关于指数界的附注; 大偏差理论	157
14.10	赫尔德不等式的一个推论	157
14.11	杜布的 \mathcal{L}^p 不等式	158
14.12	有关“乘积”鞅的角谷 (Kakutani) 定理.....	159
14.13	拉东 - 尼科迪姆 (Radon-Nikodým) 定理.....	160
14.14	拉东 - 尼科迪姆定理与条件期望	164
14.15	似然比与等价测度.....	164
14.16	似然比与条件期望.....	165

14.17	再回到角谷定理; 似然比检验的相容性	165
14.18	有关哈代 (Hardy) 空间的注记及其他 (快速阅读!)	166
第 15 章	应用	169
15.0	引言	169
15.1	一个平凡的鞅表示结果	170
15.2	期权定价; 离散时间的布莱克 - 斯科尔斯 (Black-Scholes) 公式 ...	171
15.3	马比诺吉昂 (Mabinogion) 羊问题	175
15.4	引理 15.3(c) 的证明	177
15.5	(15.3,d) 中结果的证明	179
15.6	条件概率的递归性	180
15.7	有关二元正态分布的贝叶斯 (Bayes) 公式	181
15.8	单个随机变量的含噪观察值	182
15.9	卡尔曼 - 布西 (Kalman-Bucy) 滤波	184
15.10	套紧的马具 (harness)	185
15.11	解开的马具 1	187
15.12	解开的马具 2	187

C 部分 特征函数

第 16 章	特征函数 (CF) 的基本性质	190
16.1	定义	190
16.2	基本性质	191
16.3	特征函数的一些应用	191
16.4	三个关键的结果	192
16.5	原子	193
16.6	莱维反演公式	193
16.7	表	196
第 17 章	弱收敛性	197
17.1	“漂亮”的定义	197
17.2	一个“实用”的公式	198
17.3	斯科罗霍德表示	200
17.4	$\text{Prob}(\mathbf{R})$ 的列紧性	201
17.5	紧致性	202

第 18 章 中心极限定理	203
18.1 莱维收敛定理	203
18.2 记号 o 与 O	205
18.3 一些重要的估计	205
18.4 中心极限定理	207
18.5 例	207
18.6 引理 12.4 的特征函数法证明	209

附 录

A1 章 第 1 章附录	212
A1.1 S^1 的一个不可测子集 A	212
A1.2 d -系	213
A1.3 邓肯 (Dynkin) 引理	213
A1.4 唯一性引理 1.6 的证明	214
A1.5 λ -集: “代数”情形	215
A1.6 外测度	216
A1.7 卡拉泰奥多里引理	217
A1.8 卡拉泰奥多里定理的证明	218
A1.9 $((0, 1], \mathcal{B}(0, 1])$ 上勒贝格测度存在性的证明	219
A1.10 扩张不唯一之例	221
A1.11 测度空间的完备化	222
A1.12 贝尔 (Baire) 范畴定理	223
A3 章 第 3 章附录	225
A3.1 单调类定理 3.14 的证明	225
A3.2 关于生成的 σ -代数的讨论	226
A4 章 第 4 章附录	228
A4.1 柯尔莫哥洛夫重对数律	228
A4.2 斯特拉森 (Strassen) 重对数律	228
A4.3 一个马氏链模型	229
A5 章 第 5 章附录	231
A5.1 双重单调阵列	231
A5.2 引理 1.10(a) 的关键应用	232
A5.3 “积分的唯一性”	233
A5.4 单调收敛定理的证明	234

A9 章 第 9 章附录	235
A9.1 无穷乘积: 把事情说清楚	235
A9.2 (A9.1, e) 的证明	236
A13 章 第 13 章附录	238
A13.1 收敛的方式: 定义	238
A13.2 收敛的方式: 相互间关系	239
A14 章 第 14 章附录	240
A14.1 σ -代数 \mathcal{F}_T (T 为一停时).....	240
A14.2 可选抽样定理 (OST) 的一个特例.....	241
A14.3 杜布关于一致可积鞅的可选抽样定理	242
A14.4 关于一致可积下鞅的结果	243
A16 章 第 16 章附录	244
A16.1 积分号下的微分法	244
E 章 练习题	246
参考文献	264
索引	267
后记	273