TURING 图灵数学·统计学丛书

美国工程院院士力作，MIT等全球众多名校教材！

Introduction to Probability, 2E

概率导论

[美] Dimitri P. Bertsekas
John N. Tsitsiklis 著

郑忠国 童行伟 译

（第2版·修订版）

从直观、自然的角度阐述概率，是理工科学生入门首选。

中国工信出版集团 人民邮电出版社
POSTS & TELECOM PRESS

TURING 图灵数学·统计学丛书

Introduction to Probability, 2E

概率导论

[美] Dimitri P. Bertsekas
John N. Tsitsiklis 著

郑忠国 童行伟 译

（第2版·修订版）

人民邮电出版社
北京

图书在版编目(CIP)数据

概率导论: 第2版: 修订版 /(美)伯特瑟卡斯(Bertsekas, D. P.),(美)齐齐克利斯(Tsitsiklis, J. N.)著; 郑忠国, 童行伟译. —北京: 人民邮电出版社, 2016.1

(图灵数学·统计学丛书)

ISBN 978-7-115-40507-4

Ⅰ. ①概··· Ⅱ. ①伯··· ②齐··· ③郑··· ④童···Ⅲ. ①概率论 Ⅳ. ①O211

中国版本图书馆 CIP 数据核字(2015)第 225786 号

内 容 提 要

本书是在 MIT 开设概率论入门课程的基础上编写的, 内容全面, 例题和习题丰富, 结构层次性强, 能够满足不同读者的需求. 书中介绍了概率模型、离散随机变量和连续随机变量、多元随机变量以及极限理论等概率论基本知识, 还介绍了矩母函数、条件概率的现代定义、独立随机变量的和、最小二乘估计等高级内容.

本书可作为所有高等院校概率论入门的基础教程, 也可作为有关概率论方面的参考书.

◆ 著　[美] Dimitri P. Bertsekas　John N. Tsitsiklis
译　郑忠国　童行伟
责任编辑　朱　巍
特约编辑　江志强
责任印制　杨林杰

◆ 人民邮电出版社出版发行　北京市丰台区成寿寺 11 号
邮编 100164　电子邮件 315@ptpress.com.cn
网址 http://www.ptpress.com.cn
大厂聚鑫印刷有限责任公司印刷

◆ 开本: 700×1000　1/16
印张: 29
字数: 583 千字　2016 年 1 月第 1 版
印数: 1–3 000 册　2016 年 1 月河北第 1 次印刷

著作权合同登记号　图字: 01-2015-6181 号

定价: 79.00 元

读者服务热线: (010)51095186 转 600　印装质量热线: (010) 81055316

反盗版热线: (010)81055315

版 权 声 明

译者简介

郑忠国 北京大学数学科学学院教授、博士生导师, 1965 年北京大学研究生毕业. 长期从事数理统计的教学和科研工作, 研究方向是非参数统计、可靠性统计和统计计算, 发表论文近百篇. 主持完成国家科研项目“不完全数据统计理论及其应用”, 教育部博士点基金项目“应用统计方法研究”和“工业与医学中的应用统计研究”等. 研究项目“随机加权法”获国家教委科技进步二等奖. 出版的教材有《高等统计学》、《概率与统计》(北京大学出版社) 等.

童行伟 北京师范大学数学科学学院副教授、硕士生导师, 2003 年获得北京大学数学科学学院博士学位. 曾在密苏里大学哥伦比亚分校做博士后研究, 多次访问香港各大学和新加坡国立大学. 主要从事生物统计、金融统计的教学和科研工作, 研究方向是生存分析和医学统计.

译　者　序

概率论是研究自然界和人类社会中的随机现象数量规律的数学分支. 概率论的理论和方法与数学的其他分支、自然科学、工程、人文及社会科学各领域相互交叉渗透, 已经成为这些学科中的基本方法. 概率论 (或概率统计) 和高等数学一样, 已经成为我国高等学校各专业普遍设立的一门基础课.

Dimitri P. Bertsekas 和 John N. Tsitsiklis 编写的这本《概率导论》独具特色. 作者用流畅的笔调, 阐述了概率论的基本原理和方法, 同时用大量丰富的例子说明概率论的应用领域的广泛性. 本书在内容上具有一些鲜明的特点. 首先教材的内容丰富, 除了系统地介绍概率论基本原理外, 还包含了随机过程和统计学的内容. 随机过程部分涉及伯努利和泊松过程、马尔可夫过程等内容, 统计学涉及贝叶斯统计和经典统计的主要方法. 本书的内容可以提供两门具有不同特点的一学期课程的材料, 一门是概率论与随机过程, 另一门是概率论与统计推断. 任课教员可以从本书选取相关内容组成相应课程. 本书的另一个特点是它的广泛适应性和理论的完整性. 初学者通过系统学习, 可以掌握概率论和统计学的基本原理; 追求数学严密性的学生, 也可从本书的注解和习题解答中学习到概率统计的严格理论, 了解理论的完整性和逻辑的严密性.

译者曾与本书第一作者有过当面交流的机会. 作者对于中国不断发展的教育科学事业很感兴趣, 乐于看到概率统计在中国教育领域中的地位日益提高, 乐于将本书介绍给中国读者. 本书是麻省理工学院的基础课教材, 是在多年教学的基础上写成的. 作为世界著名高校, 他们的经验值得我们学习, 我们希望本教材的中文版能够对提高我国概率统计教育水平起到积极的作用.

由于译者的学识和中英文水平有限, 译文难免有不妥之处, 欢迎广大读者批评指正.

第 2 版前言

本书对第 1 版进行了重大改动: 对原有材料的编排做了变动, 增加了新的材料, 页数也增加了 25 %. 主要的改动如下.

(a) 统计推断方面增加了两章内容: 一章是贝叶斯统计; 一章是经典统计推断. 这两章的主要内容是介绍基本概念, 并通过例子加深对方法的理解.

(b) 重新安排组织了第 3、第 4 两章的内容, 一方面是为了增加新的内容, 另一方面是为了表达的流畅. 第 1 版中的 4.7 节 (二元正态分布) 已经删去, 但是在本书的网页上还保留着.

(c) 增加了一些例子和习题.

新版的主要目的是为教师提供更多的材料以供他们选材, 特别是提供了统计推断引论的题材. 注意本书第 6~7 章和第 8~9 章在内容上是相互独立的. 另外, 第 5~7 章的内容是不依赖第 4 章的, 第 8~9 章只需要知道 4.2~4.3 节的内容. 因此, 利用本书, 可以提供下列的课程.

(a) 概率论与统计推断引论: 第 1~3 章, 4.2~4.3 节, 第 5 章, 第 8~9 章.

(b) 概率论与随机过程引论: 第 1~3 章, 第 5~7 章, 加上第 4 章少数几节.

我们要对我们的同行表示感谢. 他们对第 1 版的内容提出了宝贵的建议, 同时对新增材料的组织提供了帮助. 特别是 Ed Coffman、Munther Dahleh、Vivek Goyal、Anant Sahai、David Tse、George Verghese、Alan Willsky、John Wyatt 等. 最后, 我们要感谢 Mengdi Wang, 她为新增的两章提供了习题和图表.

Dimitri P. Bertsekas, dimitrib@mit.edu

John N. Tsitsiklis, jnt@mit.edu

2008 年 6 月于麻省剑桥

前言

概率是用计算概括的常识.
—— 拉普拉斯

本书是我们在 MIT 开设的一门概率论入门课程“概率系统分析”的基础上写成的.

选择这门课程的学生来自全校各个科系, 他们背景各异, 且兴趣广泛, 既有刚入学的本科一年级新生也有研究生, 既有学工科的也有学管理的. 为此, 在教学上我们一直力求表达简洁而又不失分析推理的严格. 我们教学的主要目的是培养学生构造和分析概率模型的能力, 希望学生既具备直观理解力又注重数学的准确性.

根据这种精神, 概率论模型中某些很严格的数学推导被简化处理了, 或者只是进行了直观的解释, 免得复杂的证明妨碍了学生对概率论本质的理解. 同时, 有些分析留在每章最后的理论习题部分, 它们用到高等微积分知识. 此外, 为了满足某些专业读者的需要, 我们将某些推理过程中的数学技巧展示在注解中.

本书包含了概率论的基础理论部分 (概率模型、离散随机变量和连续随机变量、多元随机变量以及极限定理), 这些都是概率论入门教材的主要内容. 在第 4~6 章, 也包含了一些较高级的内容, 教师在讲授的过程中可以选择部分内容, 以配合课程大纲的具体需求. 其中第 4 章介绍了矩母函数、条件概率的现代定义、独立随机变量的和、最小二乘估计、二维正态分布等内容; 第 5~6 章较为详细地介绍了伯努利、泊松和马尔可夫过程.

我们在 MIT 开设的 (一学期) 课程中, 讲授了第 1~7 章的几乎全部内容, 只是略去了二维正态分布 (4.7 节) 和连续时间马尔可夫链 (6.5 节) 两部分. 然而, 也可以作如下选择: 略去课本中关于随机过程的全部内容, 这样可使任课教师集中精力介绍概率论的基本概念, 或者增加一些感兴趣的其他材料.

本书的主要省略之处是缺乏对统计学的全面介绍. 我们引入了离散和连续情形下的贝叶斯准则和最小二乘估计, 引入贝叶斯统计理论, 但并不涉及参数估计和非贝叶斯假设检验.

本书的习题可以分成三类.

(a)*理论习题:* 理论习题 (用 * 标明) 是教材的重要组成部分. 具有数学背景的学生会发现这部分内容是由课文自然拓展而来. 我们同时给出了这部分习题的解答. 但是, 善于思考的读者会发现大部分 (特别是前几章的) 习题都能自己独立地做出来.

(b) **课程习题**: 除理论习题外, 书中还包含了难度各异的其他习题. 这些习题是在 MIT 的讨论班上经常研究的题目, 也是 MIT 的学生学习概率论的主要方法之一. 我们希望学生首先独立地做习题, 然后参考标准答案进行核对, 这样可以提高他们的学习能力. 答案公布在教材的网页上: `http://www.athenasc.com/probbook.html`.

(c) **补充习题**: 有很多补充习题并没有印在书上, 但是在本书的网页上可以查到 (且越来越多). 其中许多习题是 MIT 学生的家庭作业和考试题目. 我们希望采用本教材的教师可以同样地利用它们. 这些题目放在网上是公开的, 但是题目的答案是不公开的. 采用本教材的教师可以联系作者得到这些答案.

我们要感谢许多为本书作出贡献的人. 当我们开始在 MIT 接手这门概率论课程的教学任务时, 就开始了写书的计划. 我们的同事 Al Drake 教这门课已经几十年了. 他的课程组织经历了时间的考验, 其经典教材对各个题材均有生动的描述, 还有大量讨论班内容和家庭作业等丰富的材料, 我们十分庆幸自己的工作有这样高的起点. 特别感谢 Al Drake 给我们创造了如此有利的起始条件.

我们也要感谢其他院校的几位同事, 他们有的利用本书的手稿进行教学, 有的阅读过手稿, 并对本书的改进提供了反馈. 我们要特别感谢 Ibrahim Abou Faycal、Gustavo de Veciana、Eugene Feinberg、Bob Gray、Muriel Médard、Jason Papastavrou、Ilya Pollak、David Tse、Terry Wagner 等.

还有 MIT 的助教们, 他们对各阶段的书稿进行了认真的校核, 并丰富和完善了习题和解答. 通过他们与学生的直接交流, 才使得本教材能够适应学生的学习水平.

本书能够为 MIT 的数千学生在其学业生涯之初提供服务, 使我们感到十分欣慰. 在本书的成书过程中, 他们热心反馈书本中的问题和学习心得. 在此感谢他们的反馈与耐心.

最后, 我们还要感谢我们的家人在这个漫长的成书过程中对我们的支持.

Dimitri P. Bertsekas, dimitrib@mit.edu

John N. Tsitsiklis, jnt@mit.edu

2002 年 5 月于麻省剑桥

目　录

第1章　样本空间与概率

“概率”是一个非常有用的概念，它可以从不同的层面来加以解释．先看下面一幅对话场景．

> 一个病人被送进医院，并施以一种急救的药．病人家属为了了解药的疗效，询问了当班的护士．下面是他们之间的一段对话．
>
> 家属：护士小姐，请问这种药有效的概率是多少？
>
> 护士：我希望这种药是有效的，明天就会见分晓．
>
> 家属：是的，但是我想知道这种药有效的概率．
>
> 护士：每个病人的病情是不一样的，看情况发展吧．
>
> 家属：这么说吧，在 100 宗类似的病例中，你认为有多少宗是有效的？
>
> 护士（有些不耐烦）：我已经告诉你了，每个病人的情况是不一样的．这种药，对某些病人是有效的，对另一些病人是无效的．
>
> 家属（继续坚持）：现在请告诉我，如果必须打赌的话，你会押哪一注，这种药是有效还是无效？
>
> 护士（有些惊奇）：那我愿意打赌，对于这位病人，这种药是有效的．
>
> 家属（多少松了一口气）：好吧！我再问你，你是否愿意如此押注：若这药无效，你输掉 2 元钱；若这药有效，你赢 1 元钱？
>
> 护士（有些恼怒）：多么荒谬的想法！你是在浪费我的时间．

在这组对话中，病人家属希望用概率的概念同护士讨论药的疗效这种具有*不确定性*的事件．但是护士的第一反应是对概率这个概念的不认可，或不理解，而家属试图将概率的概念解释得更具体一些．他首先试图将概率解释成偶然事件在多次重复试验中出现的*频率*，这是最通常的解释．例如，我们说一枚两面对称的硬币，在抛掷试验中以 50% 的概率出现正面，这么说实际上是指在多次重复抛掷硬币时，出现正面向上的次数约占一半．但是护士似乎不大愿意接受家属的这种想法，护士的想法不是完全没有道理．如果这种药是第一次在医院里使用，或护士从没有过这方面的经验，那何从谈起治愈的频率呢？

在许多涉及不确定性的事例中，用频率解释是适宜的，然而，也有一些事例不宜用频率解释．比如，有一个学者以 90% 的把握断言《伊里亚特》和《奥德赛》是由同一作者创作的．由于他所讨论的是不可重复的一次性事件，这样的结论只是提供一些主观看法，而与频率无关．所谓概率为 90% 的把握只是学者的*主观信念*．或

许有人认为主观信念是不值得研究的, 至少从数学或科学的观点来看是如此. 但是在实际生活中, 人们面对不确定性的时候, 经常不得不作出抉择. 为了作出正确的或至少保持一致的抉择, 科学和系统地利用他们的主观信念是一个先决条件.

事实上, 一个理智的选择和行动揭示了许多内在的主观概率, 然而在许多场合中, 作出抉择的人自已也没有意识到他们应用了概率推理. 在前面的对话场景中, 病人家属以一种隐蔽的方式试图推断护士的主观信念. 由于护士愿意以 1:1 的赔率打赌这种药是有效的, 那么在护士的主观概念中, 这种药有效的概率至少为 50%. 如果这位护士接受对话最后提出的赔率为 2:1 的赌注的话, 这说明在护士的主观概念中, 这种药有效的概率至少为 2/3.

在此我们不去深究概率推理适用性方面的哲学问题, 而是事先假定概率论在很多方面都具有实用价值, 包括概率只反映主观信念的情形. 概率论在科学、工程、医药、管理等领域中有许多成功应用的事例. 这许多经验证据说明概率论在应用中是一种极其有用的工具.

本书的主要目的是发掘用概率模型描述不确定性的艺术和提高概率推理的能力. 作为第一步, 本章要把概率模型的基础结构及基本性质刻画清楚. 概率是定义在某些试验结果的集合上的. 为此, 我们首先应该对集合论作一简介.

1.1　集　　合

概率论大量应用集合运算. 我们首先引进相关的记号和术语.

将一些研究对象放在一起, 形成集合, 而这些对象就称为集合的*元素*. 设 S 是一个集合, x 是 S 的元素, 我们将元素和集合的这种关系写成 $x\in S$. 若 x 不是 S 的元素, 就写成 $x\notin S$. 一个集合可以没有元素, 这个特殊的集合就称为*空集*, 记作 $\varnothing$.

可用不同的方法刻画一个集合. 若 S 包含有限个元素 $x_1, x_2, \cdots, x_n$, 我们只需将这些元素列在花括弧中:

$$S = \{ x_1, x_2, \cdots, x_n\}.$$

例如, 掷一枚骰子以后的所有可能结果的集合是 $\{1,2,3,4,5,6\}$, 抛一枚硬币的可能结果的集合是 $\{H, T\}$, 其中 H 代表正面向上, T 代表反面向上.

若 S 包含无限多个元素 $x_1, x_2, \cdots$, 但它们可以像正整数那样排成一列, 我们可写成

$$S = \{ x_1, x_2, \cdots\},$$

此时称 S 为*可数无限集*. 例如, 偶数的集合 $\{0, 2, -2, 4, -4, \cdots\}$ 是一个可数无限集.

我们也可以以 x 具有某种性质 P 为条件来刻画一个集合, 记作

$$\{x|\ x \text{ 满足性质 } P\}.$$

例如, 偶数集合可写成 $\{k|\ k/2\text{是整数}\}$. 类似地, 在实数区间 $[0,1]$ 中的数集可表示成 $\{x|\ 0\leqslant x\leqslant 1\}$. 注意, 集合 $\{x|0\leqslant x\leqslant 1\}$ 是一个连续集合, 它不可能排成一列(章后习题中给出了证明概要). 这样的集合是不可数的集合.

若集合 S 的所有元素均为集合 T 的元素, 就称 S 为 T 的子集, 记作 $S\subset T$ 或 $T\supset S$. 若 $S\subset T$ 且 $T\subset S$, 则两个集合相等, 记作 $S=T$. 引入空间的概念是十分必要的. 将我们感兴趣的所有元素放在一起, 形成一个集合, 这个集合称为空间, 记作 Ω. 当 Ω 确定以后, 我们所讨论的集合 S 都是 Ω 的子集.

1.1.1 集合运算

集合 $\{x\in\Omega|\ x\notin S\}$ 称为集合 S 相对于 Ω 的补集, 记作 S^c. 注意 $\Omega^c=\varnothing$.

由属于 S 或属于 T 的元素组成的集合称为 S 和 T 的并, 记为 $S\cup T$. 既属于 S 又属于 T 的元素组成的集合称为 S 和 T 的交, 记成 $S\cap T$. 这些集合可用下列公式表达:

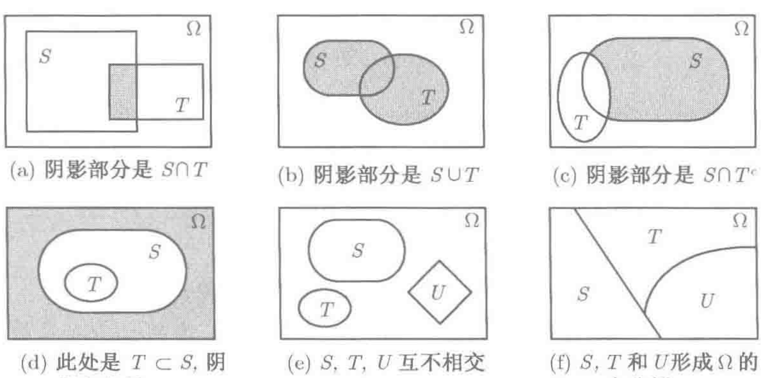
$$S\cup T=\{x|\ x\in S \text{ 或 } x\in T\},$$

$$S\cap T=\{x|\ x\in S \text{ 且 } x\in T\}.$$

有时候我们需要考虑几个甚至无穷个集合的并和交的问题. 例如, 如果每一个正整数 n 都确定一个集合 S_n, 则

$$\bigcup_{n=1}^{\infty} S_n=S_1\cup S_2\cdots=\{x|\ x\in S_n \text{ 对某个 } n \text{ 成立}\},$$

$$\bigcap_{n=1}^{\infty} S_n=S_1\cap S_2\cdots=\{x|\ x\in S_n \text{ 对一切 } n \text{ 成立}\}.$$



(a) 阴影部分是 $S\cap T$　(b) 阴影部分是 $S\cup T$　(c) 阴影部分是 $S\cap T^c$

(d) 此处是 $T\subset S$, 阴影部分是 S^c　(e) S, T, U 互不相交　(f) S, T 和 U 形成 Ω 的一个分割

图 1.1　维恩图的例子

两个集合称为**不相交的**, 如果它们的交集为空集. 更一般地, 几个集合称为**互不相交**的, 如果任何两个集合没有公共元素. 一组集合称为集合 S 的**分割**, 如果这组集合中的集合互不相交, 并且它们的并为 S.

设 x 和 y 为两个研究对象, 我们用 (x,y) 表示 x 和 y 的**有序对**. 我们用 $\mathbf{R}$ 表示实数集合, 用 $\mathbf{R}^2$ 表示实数对的集合, 即二维平面, 用 $\mathbf{R}^3$ 表示三维实数向量的集合 (三维空间). 集合及其运算可用**维恩图**形象化表示, 见图 1.1.

1.1.2 集合的代数

集合运算具有若干性质, 这些运算性质可由运算的定义直接证得, 举例如下:

$$
\begin{aligned}
&S\cup T=T\cup S, && S\cup(T\cup U)=(S\cup T)\cup U,\\
&S\cap(T\cup U)=(S\cap T)\cup(S\cap U), && S\cup(T\cap U)=(S\cup T)\cap(S\cup U),\\
&(S^c)^c=S, && S\cap S^c=\varnothing,\\
&S\cup\Omega=\Omega, && S\cap\Omega=S.
\end{aligned}
$$

下面给出的两个公式就是著名的**德摩根定律**:

$$\left(\bigcup_n S_n\right)^c=\bigcap_n S_n^c,\qquad \left(\bigcap_n S_n\right)^c=\bigcup_n S_n^c.$$

现在证明第一个公式. 设 $x\in(\cup_n S_n)^c$, 这说明 $x\notin\cup_n S_n$, 即对一切 n, $x\notin S_n$. 因而, 对每一个 n, x 属于 S_n 的补集, 即 $x\in\cap_n S_n^c$. 这样, 我们得到 $(\cup_n S_n)^c\subset\cap_n S_n^c$. 反过来包含关系的证明, 只需将我们的论证从后面往前推即可. 而第二个公式的证明完全类似.

1.2 概 率 模 型

概率模型是对不确定现象的数学描述. 为了与本节讨论的基本框架保持一致, 下面列出了它的两个基本构成, 并用图 1.2 做了形象阐释.

概率模型的基本构成

- **样本空间** Ω, 这是一个试验的所有可能结果的集合.
- **概率律**, 概率律为试验结果的集合 A(称为**事件**) 确定一个非负数 $\mathrm{P}(A)$ (称为事件 A 的**概率**). 而这个非负数刻画了我们对事件 A 的认识或所产生的信念的程度. 稍后将指出概率律必须满足的某些性质.