HZ BOOKS
华章教育

WILEY

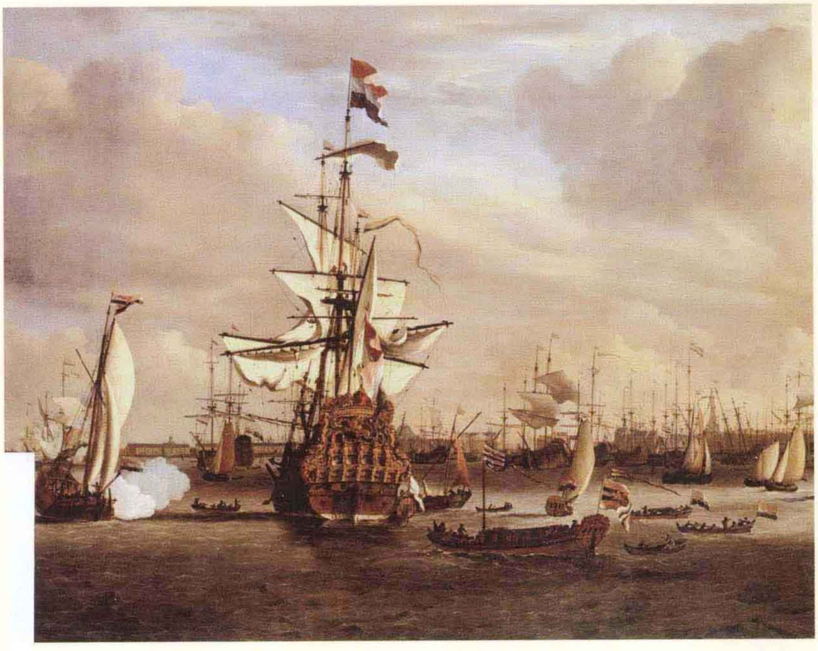
统计学精品译丛

（原书第2版）

随机过程

Stochastic Processes

(Second Edition)



（美）Sheldon M. Ross 著
龚光鲁 译

机械工业出版社
China Machine Press

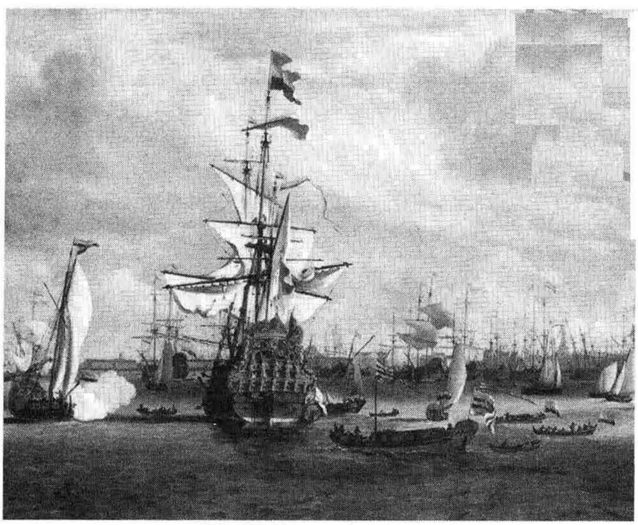
统计

（原书第2版）

随机过程

Stochastic Processes

(Second Edition)



（美）Sheldon M. Ross 著
龚光鲁 译

机械工业出版社
China Machine Press

图书在版编目（CIP）数据

随机过程（原书第 2 版）/（美）罗斯（Ross，S. M.）著；龚光鲁译．—北京：机械工业出版社，2013.7
（统计学精品译丛）
书名原文：Stochastic Processes，Second Edition

ISBN 978-7-111-43029-2

Ⅰ．随…　Ⅱ．①罗…　②龚…　Ⅲ．随机过程　Ⅳ．O211.6

中国版本图书馆 CIP 数据核字（2013）第 136418 号

本书版权登记号：图字：01-2013-0181

本书从概率的角度而不是分析的角度来看待随机过程，书中介绍了随机过程的基本理论，包括 Poisson 过程、Markov 链、鞅、Brown 运动、随机序关系、Poisson 逼近等，并阐明这些理论在各领域的应用．书中有丰富的例子和习题，其中一些需要创造性地运用随机过程知识、系统地解决的实际问题，给读者提供了应用概率研究的实例．

本书是随机过程的入门教材，没有用到测度论，仅以微积分及初等概率论知识为基础，适合作为统计学专业本科生以及其他理工和经管类专业研究生相关课程的教材，更值得相关研究人员和授课教师参考．

机械工业出版社（北京市西城区百万庄大街 22 号　　邮政编码　100037）
责任编辑：明永玲
北京市荣盛彩色印刷有限公司印刷
2013 年 7 月第 1 版第 1 次印刷
186mm×240mm・20.75 印张
标准书号：ISBN 978-7-111-43029-2
定　　价：79.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：（010）88378991　88361066　　　　投稿热线：（010）88379604
购书热线：（010）68326294　88379649　68995259　　读者信箱：hzjsj@hzbook.com

译 者 序

随机过程是研究依赖于时间的随机现象的数学工具. Sheldon M. Ross 的《随机过程》是一本通俗的教材，初版于 1983 年. 它不用概率论基础的严格数学框架（这种框架是建立在测度论基础上的叙证，其长处是绝对的数学严格使得不存在丝毫不清晰与二义性，然而也使得初学者理解概率实质较为困难），用更为直观的概率思想，介绍了随机过程的基本概念、内容和应用. 这本书已为我国概率界熟知并流行，在用随机过程的概念与方法处理应用问题上起了重要的作用.

译者于 1987 年在北京大学概率统计系的研究生班讲授应用随机过程课程时，曾以该书第 1 版作为教材. 如今这些学生已遍及海内外，其中不少已成为各领域的精英. 在 20 世纪 90 年代，该书已由何声武等人译为中文，不幸的是至今没有第 2 版的中文版出现. 声武沉稳敏思，抱负远大，可惜英年早逝.

1995 年，Sheldon M. Ross 的《随机过程（第 2 版）》问世，其中与时俱进地加进了 Gibbs 采样与 Metropolis 采样等可近似地跟踪 Markov 链的路径的方法，以在研究随机现象时能借助高速计算机发展的成果，这些方法已在应用领域中显现其潜力. 在第 2 版中也增加了对较复杂的随机现象的例子的分析.

译者借此机会将第 2 版译出，使读者能品味概率分析的思维方法. 期望读者在学习本书的基础上，按本人的兴趣，能在如下两方面之一得以发展：在金融工程、经济学、生命科学、分子生物学、分子化学、运动目标追踪等方面的应用中做贡献，或在严格的随机过程研究理论方面有突出成果.

龚光鲁

第 2 版前言

第 2 版包括以下的改变：

（i）第 2 章增加了关于复合 Poisson 随机变量的内容，包括能有效地计算矩的一个恒等式，且由该恒等式导出非负整数值复合 Poisson 随机变量的概率质量函数的一个优美的递推方程.

（ii）有关鞅的内容单独作为一章（第 6 章），包括 Azuma 不等式的几节。

（iii）全新的关于 Poisson 逼近的一章（第 10 章），包括给出这些逼近的误差界的 Stein-Chen 方法和一种改进逼近本身的方法.

此外，遍及全书我们还加进了大量例题和习题. 个别章节的增加如下：

在第 1 章，我们新加了关于概率方法、多元正态分布、在图上的随机徘徊和完全匹配问题的例子. 我们也新加了一节关于概率不等式（包括 Chernoff 界）和一小节介绍 Bayes 估计（证明了它们几乎都不是无偏的）. 在此章附录中给出了强大数定律的一个证明.

在第 3 章中给出关于模式和无记忆的最佳硬币投掷策略的崭新例子.

在第 4 章中增加了处理在暂态中停留的平均时间的崭新内容，同时有关于 Gibbs 采样、Metropolis 算法以及在星形图中的平均覆盖时间的崭新例子.

第 5 章包含两性人口增长模型的一个例子.

第 6 章有说明鞅的停止定理的用途的附加例子.

第 7 章包含 Spitzer 等式的新材料，同时用它计算具有 gamma 分布到达间隔和服务时间的单服务线队列中的平均延迟.

第 8 章 Brown 运动已被移至鞅的章节的后面，以便应用鞅来分析 Brown 运动.

第 9 章关于随机序，现在包含相伴随机变量，也包含利用在优惠券收集和装箱问题中的耦合的新的例子.

我们想感谢所有撰写并发送对第 1 版评论的热心人士，特别感谢何声武，Stephen Herschkorn，Robert Kertz，James Matis，Erol Pekoz，Maria Rieders 和 Tomasz Rolski 提出许多有价值的意见.

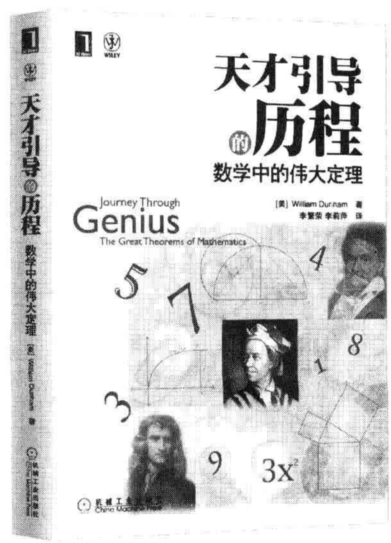
第 1 版前言

这本教材是非测度论的随机过程导论，且至多假定读者具备微积分和初等概率论的知识. 在书中我们试图介绍随机过程的一些理论，显示其在不同领域中的应用，同时也培养学生在思考问题时所需的一些概率直观和洞察力. 我们尽可能从概率的角度而不是分析的角度看待随机过程. 例如，这种尝试引导我们从一条样本路径的观点研究大多数随机过程.

我要感谢 Mark Brown，Cyrus Derman，Shun-Chen Niu，Michael Pinedo 和 Zvi Schechner 提出许多有价值的意见.

Sheldon M. Ross

推荐阅读



天才引导的历程：数学中的伟大定理

作者：William Dunham ISBN：978-7-111-40329-6 定价：45.00元

"Dunham的这本书如此特别，是我以前从未遇到过的……娓娓道来的一个个推理精巧与颇具洞察力的个案，引人入胜。"

——当代美国最著名的科普作家、科幻小说家、文学评论家Isaac Asimov

"把Dunham的这本书推荐给所有热爱探索、思想活跃的人们，不管他们感兴趣的是艺术还是科学，阅读本书都是一次重要的文化体验。"

——英国著名数学教育家暨《上帝掷骰子吗?》的作者Ian Stewart

"如果有一本书让你读了会对数学产生毕生爱好的话，那肯定就是Dunham的这本书。"

——Amazon读者评论

"这门几乎每个人都觉得沉闷、无聊、呆板的学科，在Dunham的笔下充满生机与活力……我是拥有计算机科学学位的外行，但是我喜欢这本书……Dunham巧妙地将数学中的伟大定理编织成数学史，使得本书容易理解，而且我敢说，事实上很有趣味性！本书是一颗珍宝，每一个爱好数学的人都不能与它失之交臂。"

——Amazon读者评论

这是二十年多年来一直畅销不衰的数学大众读物！长期雄踞亚马逊数学大众读物畅销榜榜首！

本书将两千多年的数学发展历程融为十二章内容，每章都包含了三个基本组成部分，即历史背景、人物传记以及在这些"数学杰作"中所表现出的创造性。作者精心挑选了一些杰出的数学家及其所创造的伟大定理，如欧几里得、阿基米德、牛顿和欧拉。这一个个伟大的定理，不仅串起了历史的年轮，更是串起了数学这门学科所涵盖的各个深邃而不乏实用性的领域。当然，这不是一本典型的数学教材，而是一本大众读物，它会让热爱数学的人体会到绝处逢生的喜悦，让讨厌数学的人从此爱上数学。

推荐阅读

书名	书号	定价	出版年	作者
概率统计（英文版·第4版）	978-7-111-38775-6	139	2012	（美）Morris H. DeGroot等
统计模拟（英文版·第5版）	978-7-111-42045-3	59	2013	（美）Sheldon M. Ross
数值分析（英文版·第2版）	978-7-111-38582-0	89	2013	（美）Timothy Sauer
数论概论（英文版·第4版）	978-7-111-38581-3	69	2013	（美）Joseph H. Silverman
数理统计学导论（英文版·第7版）	978-7-111-38580-6	99	2013	（美）Robert V. Hogg等
代数（英文版·第2版）	978-7-111-36701-7	79	2012	（美）Michael Artin
线性代数（英文版·第 8 版）	978-7-111-34199-4	69	2011	（美）Steven J. Leon
商务统计：决策与分析（英文版）	978-7-111-34200-7	119	2011	（美）Robert Stine等
多元数据分析（英文版·第7版）	978-7-111-34198-7	109	2011	（美）Joseph F. Hair, Jr等
统计模型：理论和实践（英文版·第2版）	978-7-111-31797-5	38	2010	（美）David A. Freedman
实分析（英文版·第4版）	978-7-111-31305-2	49	2010	（美）H. L. Royden
概率论教程（英文版·第3版）	978-7-111-30289-6	49	2010	（美）Kai LaiChung
初等数论及其应用（英文版·第6版）	978-7-111-31798-2	89	2010	（美）Kenneth H. Rosen
组合数学（英文版·第5版）	978-7-111-26525-2	49	2009	（美）Richard A. Brualdi
数学建模（英文精编版·第4版）	978-7-111-28249-5	65	2009	（美）FrankR. Giordano
复变函数及应用（英文版·第8版）	978-7-111-25363-1	65	2009	（美）James Ward Brown
数学建模方法与分析（英文版·第3版）	978-7-111-25364-8	49	2008	（美）MarkM. Meerschaert
数论概论（英文版·第3版）	978-7-111-19611-2	52	2006	（美）Joseph H. Silverman
实分析和概率论（英文版·第2版）	978-7-111-19348-7	69	2006	（美） R. M. Dudley
高等微积分（英文版·第2版）	978-7-111-19349-4	76	2006	（美）Patrick M. Fitzpatrick
数学分析原理（英文版·第3版）	978-7-111-13306-3	35	2004	（美）Walter Rudin
实分析与复分析（英文版·第3版）	978-7-111-13305-6	39	2004	（美）Walter Rudin
泛函分析（英文版·第2版）	978-7-111-13415-2	42	2004	（美）Walter Rudin

目录

第 1 章　准备知识

1.1　概率

在概率论中的一个基本概念是随机试验，这种试验的结果不能预先确定. 一个试验所有可能的结果的集合称为此试验的样本空间，而我们将它记为 S.

事件是样本空间的一个子集，如果此试验的结果是这个子集的一个元素，则称这个事件发生了. 我们假定对于样本空间 S 的每个事件 E，定义了一个数 $P(E)$，它满足下述三条公理[⊖].

公理（1）　$0 \leqslant P(E) \leqslant 1$.

公理（2）　$P(S) = 1$.

公理（3）　对于任意相互排斥的事件序列 $E_1, E_2, \cdots$，即对于当 $i \neq j$ 时 $E_iE_j = \varnothing$（此处 $\varnothing$ 是空集合）的事件，有

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$

我们将称 $P(E)$ 为事件 E 的概率.

公理（1），（2）和（3）的一些简单推论如下.

1.1.1　若 $E \subset F$，则 $P(E) \leqslant P(F)$.

1.1.2　$P(E^c) = 1 - P(E)$，其中 E^c 是 E 的补.

1.1.3　$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$，当各个 E_i 相互排斥时.

1.1.4　$P\left(\bigcup_{i=1}^{\infty} E_i\right) \leqslant \sum_{i=1}^{\infty} P(E_i)$.

不等式（1.1.4）是所谓的 Boole 不等式. 1[⊜]

概率函数 P 的一个重要的性质是：它是连续的. 为了更精确地阐述这种性质，我们需要极限事件的概念，下面就来定义它. 如果 $E_n \subset E_{n+1}, n \geqslant 1$，则称事件序列 $\{E_n, n \geqslant 1\}$ 为递增序列；如果 $E_n \supset E_{n+1}, n \geqslant 1$，则称事件序列 $\{E_n, n \geqslant 1\}$ 为递减序列. 如果 $\{E_n, n \geqslant 1\}$ 是事件的一个递增序列，那么我们定义一个新的事件，记为 $\lim\limits_{n\to\infty} E_n$，它定义为

$$\lim_{n\to\infty} E_n = \bigcup_{i=1}^{\infty} E_i\text{，当 } E_n \subset E_{n+1}, n \geqslant 1.$$

同样，如果 $\{E_n, n \geqslant 1\}$ 是事件的一个递减序列，那么 $\lim\limits_{n\to\infty} E_n$ 定义为

$$\lim_{n\to\infty} E_n = \bigcap_{i=1}^{\infty} E_i\text{，当 } E_n \supset E_{n+1}, n \geqslant 1.$$

现在我们可以叙述下述结论.

⊖　事实上，只能对 S 的所谓可测事件定义 $P(E)$，但是我们并不关心这种限制.

⊜　此页码为英文原书页码，与索引页码一致.

命题 1.1.1　如果 $\{E_n, n \geqslant 1\}$ 是事件的一个递增或递减序列，那么

$$\lim_{n\to\infty} P(E_n) = P\left(\lim_{n\to\infty} E_n\right).$$

证明　首先假设 $\{E_n, n \geqslant 1\}$ 是事件的一个递增序列，定义事件 $F_n, n \geqslant 1$ 为

$$F_1 = E_1,$$

$$F_n = E_n\left(\bigcup_1^{n-1} E_i\right)^{\mathrm{c}} = E_n E_{n-1}^{\mathrm{c}}, \quad n > 1.$$

也就是说，F_n 由 E_n 中不属于它之前的任意一个 $E_i, i < n$ 的那些点构成. 容易检查 F_n 是相互排斥的事件，它们使得对于一切 $n \geqslant 1$ 有

2 $$\bigcup_{i=1}^{\infty} F_i = \bigcup_{i=1}^{\infty} E_i \quad 和 \quad \bigcup_{i=1}^{n} F_i = \bigcup_{i=1}^{n} E_i.$$

于是

$$\begin{aligned} P\left(\bigcup_{i=1}^{\infty} E_i\right) &= P\left(\bigcup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} P(F_i) \qquad (由公理(3)) \\ &= \lim_{n\to\infty} \sum_1^n P(F_i) = \lim_{n\to\infty} P\left(\bigcup_1^n F_i\right) \\ &= \lim_{n\to\infty} P\left(\bigcup_1^n E_i\right) = \lim_{n\to\infty} P(E_n), \end{aligned}$$

它证明了当 $\{E_n, n \geqslant 1\}$ 递增时的结论.

如果 $\{E_n, n \geqslant 1\}$ 是一个递减序列，那么 $\{E_n^{\mathrm{c}}, n \geqslant 1\}$ 是一个递增序列，因此

$$P\left(\bigcup_1^{\infty} E_n^{\mathrm{c}}\right) = \lim_{n\to\infty} P(E_n^{\mathrm{c}}).$$

但是，因为 $\bigcup_1^{\infty} E_n^{\mathrm{c}} = \left(\bigcap_1^{\infty} E_n\right)^{\mathrm{c}}$，所以

$$1 - P\left(\bigcap_1^{\infty} E_n\right) = \lim_{n\to\infty}[1 - P(E_n)],$$

或者，等价地

$$P\left(\bigcap_1^{\infty} E_n\right) = \lim_{n\to\infty} P(E_n),$$

这就证明了结果. ◀

3 **例 1.1 (A)**　考虑由能产生同种后代的个体组成的总体. 初始出现的个体数记为 X_0，称为第 0 代的大小. 第 0 代的所有后代构成第 1 代，他们的个数记为 X_1. 一般以 X_n 记第 n 代的大小.

由于 $X_n = 0$ 蕴含 $X_{n+1} = 0$，由此推出 $P\{X_n = 0\}$ 是递增的，于是 $\lim_{n\to\infty} P\{X_n = 0\}$ 存在. 这又意味着什么呢？为了回答这个问题，我们如下使用命题 1.1.1：

$$\lim_{n\to\infty} P\{X_n = 0\} = P\left\{\lim_{n\to\infty}\{X_n = 0\}\right\} = P\left\{\bigcup_n \{X_n = 0\}\right\} = P\{总体终将灭绝\}.$$

即，第 n 代没有个体的极限概率等于总体终将灭绝的概率. ■

命题 1.1.1 可用于证明 Borel-Cantelli 引理.

命题 1.1.2 Borel-Cantelli 引理
以 $E_1, E_2, \cdots$ 记一个事件序列. 如果

$$\sum_{i=1}^{\infty} P(E_i) < \infty,$$

那么

$$P\{\text{无穷多个 } E_i \text{ 发生}\} = 0.$$

证明 有无穷多个 E_i 发生的事件，称为 $\limsup\limits_{i\to\infty} E_i$，可以表示为

$$\limsup_{i\to\infty} E_i = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} E_i.$$

这得自，如果有无穷多个 E_i 发生，那么对于每个 n，$\bigcup\limits_{i=n}^{\infty} E_i$ 都发生，于是 $\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{i=n}^{\infty} E_i$ 发生. 另一方面，如果 $\bigcap\limits_{n=1}^{\infty} \bigcup\limits_{i=n}^{\infty} E_i$ 发生，那么对于每个 n，$\bigcup\limits_{i=n}^{\infty} E_i$ 都发生，于是对于每个 n 至少有一个 E_i ($i \geqslant n$) 发生，因此有无穷多个 E_i 发生. 4

因为 $\bigcup\limits_{i=n}^{\infty} E_i, n \geqslant 1$ 是一个递减的事件序列，由命题 1.1.1 推出

$$P\Big(\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} E_i\Big) = P\Big(\lim_{n\to\infty} \bigcup_{i=n}^{\infty} E_i\Big) = \lim_{n\to\infty} P\Big(\bigcup_{i=n}^{\infty} E_i\Big) \leqslant \lim_{n\to\infty} \sum_{i=n}^{\infty} P(E_i) = 0,$$

于是就证明了结论. ◀

例 1.1 (B) 令 $X_1, X_2, \cdots$ 使

$$P\{X_n = 0\} = 1/n^2 = 1 - P\{X_n = 1\}, \quad n \geqslant 1.$$

如果我们令 $E_n = \{X_n = 0\}$，那么因为 $\sum\limits_{i=n}^{\infty} P(E_i) < \infty$，由 Borel-Cantelli 引理推出，有无穷多个 n 使 X_n 等于 0 的概率等于 0. 因此对于一切充分大的 n，X_n 必须等于 1，所以我们可以得到结论，以概率 1 地有

$$\lim_{n\to\infty} X_n = 1.$$ ■

对于 Borel-Cantelli 引理的逆定理，要求有独立性.

命题 1.1.3 (Borel-Cantelli 引理的逆) 如果 $E_1, E_2, \cdots$ 是独立事件，使得

$$\sum_{n=1}^{\infty} P(E_n) = \infty,$$

那么

$$P\{\text{无穷多个 } E_i \text{ 发生}\} = 1.$$

证明 5

$$P\{\text{无穷多个 } E_i \text{ 发生}\} = P\Big\{\lim_{n\to\infty} \bigcup_{i=n}^{\infty} E_i\Big\} = \lim_{n\to\infty} P\Big(\bigcup_{i=n}^{\infty} E_i\Big) = \lim_{n\to\infty}\Big[1 - P\Big(\bigcap_{i=n}^{\infty} E_i^c\Big)\Big].$$

现在

$$P\Big(\bigcap_{i=n}^{\infty} E_i^c\Big) = \prod_{i=n}^{\infty} P(E_i^c) \qquad \text{(由独立性)}$$

$$
\begin{aligned}
&= \prod_{i=n}^{\infty}(1-P(E_i)) \\
&\leqslant \prod_{i=n}^{\infty} \mathrm{e}^{-P(E_i)} \qquad (\text{由不等式 } 1-x \leqslant \mathrm{e}^{-x}) \\
&= \exp\Big(-\sum_{n}^{\infty} P(E_i)\Big) = 0 \qquad (\text{因为对于一切 } n \text{ 有 } \sum_{i=n}^{\infty} P(E_i) = \infty).
\end{aligned}
$$

由此得出结论. ◀

例 1.1 (C) 令 $X_1, X_2, \cdots$ 独立且使

$$P\{X_n = 0\} = 1/n = 1 - P\{X_n = 1\}, n \geqslant 1.$$

如果令 $E_n = \{X_n = 0\}$，那么因为 $\sum_{n=1}^{\infty} P(E_n) = \infty$，由命题 1.1.3 推出 E_n 无穷多次发生. 又因为 $\sum_{n=1}^{\infty} P(E_n^c) = \infty$，同样推出 E_n^c 也无穷多次发生. 因此，以概率为 1 地有，X_n 无穷多次等于 0，也无穷多次等于 1. 于是，当 $n \to \infty$ 时，以概率为 1 地有，X_n 没有
6 极限值. ■

1.2 随机变量

考虑一个有样本空间 S 的随机环境. 一个随机变量 X 是一个函数，它给 S 中的每一个结果都指定一个实数值. 对于任意实数集合 A，X 假定的值包含于 A 中的概率等于试验的结果包含于 $X^{-1}(A)$ 中的概率. 即

$$P\{X \in A\} = P(X^{-1}(A)),$$

其中 $X^{-1}(A)$ 是使得 $X(s) \in A$ 的一切点 $s \in S$ 组成的事件.

对于任意的实数 x，随机变量 X 的分布函数 F 定义为

$$F(x) = P\{X \leqslant x\} = P\{X \in (-\infty, x]\}.$$

我们将 $1 - F(x)$ 记为 $\overline{F}(x)$，所以

$$\overline{F}(x) = P\{X > x\}.$$

如果一个随机变量 X 可能值的集合是可数的，则称它是离散的随机变量. 对于离散的随机变量，

$$F(x) = \sum_{y \leqslant x} P\{X = y\}.$$

如果存在一个函数 $f(x)$（称为概率密度函数），使得对于一切集合 B 有

$$P\{X \text{ 在 } B \text{ 中}\} = \int_B f(x)\mathrm{d}x,$$

则称随机变量 X 是连续的. 由于 $F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$，由此推出

$$f(x) = \frac{\mathrm{d}}{\mathrm{d}x} F(x).$$

两个随机变量 X 和 Y 的联合分布函数定义为

$$F(x, y) = P\{X \leqslant x, Y \leqslant y\}.$$

X 和 Y 的分布函数，

$$F_X(x) = P\{X \leqslant x\} \quad 和 \quad F_Y(y) = P\{Y \leqslant y\},$$ 7

可以利用概率算子的连续性质由 $F(x,y)$ 得到. 特别地，以 $y_n, n \geqslant 1$ 记趋向 ∞ 的一个递增序列. 那么因为事件序列 $\{X \leqslant x, Y \leqslant y_n\}, n \geqslant 1$ 是递增的，而且

$$\lim_{n \to \infty} \{X \leqslant x, Y \leqslant y_n\} = \bigcup_{n=1}^{\infty} \{X \leqslant x, Y \leqslant y_n\} = \{X \leqslant x\},$$

由连续性质推出

$$\lim_{n \to \infty} P\{X \leqslant x, Y \leqslant y_n\} = P\{X \leqslant x\},$$

或等价地

$$F_X(x) = \lim_{y \to \infty} F(x,y).$$

类似地

$$F_Y(y) = \lim_{x \to \infty} F(x,y).$$

如果对于一切的 x 和 y，有

$$F(x,y) = F_X(x) F_Y(y),$$

则称随机变量 X 和 Y 是独立的.

如果存在一个函数 $f(x,y)$（称为联合概率密度函数），使得对于一切集合 A 和 B 有

$$P\{X 在 A 中, Y 在 B 中\} = \int_A \int_B f(x,y) \mathrm{d}y \mathrm{d}x,$$

则称随机变量 X 和 Y 是联合地连续的，

任意一族随机变量 $X_1, X_2, \cdots, X_n$ 的联合分布定义为

$$F(x_1, \cdots, x_n) = P\{X_1 \leqslant x_1, \cdots, X_n \leqslant x_n\}.$$

而且，如果

$$F(x_1, \cdots, x_n) = F_{X_1}(x_1) F_{X_2}(x_2) \cdots F_{X_n}(x_n),$$

其中

$$F_{X_i}(x_i) = \lim_{\substack{x_j \to \infty \\ j \neq i}} F(x_1, \cdots, x_n).$$

则称这 n 个随机变量是独立的. 8

1.3 期望值

随机变量 X 的期望或均值，记为 $E[X]$，定义为

$$E[X] = \int_{-\infty}^{\infty} x \mathrm{d}F(x) = \begin{cases} \int_{-\infty}^{\infty} x f(x) \mathrm{d}x & 若 X 是连续的 \\ \sum_x x P\{X = x\} & 若 X 是离散的 \end{cases} \tag{1.3.1}$$

如果此积分存在.

方程（1.3.1）同时也定义了 X 的任意函数（例如 $h(X)$）的期望. 因为 $h(X)$ 本身是一个随机变量，由方程（1.3.1）推出

$$E[h(X)] = \int_{-\infty}^{\infty} x \mathrm{d}F_h(x),$$

其中 F_h 是 $h(X)$ 的分布函数. 但是可证明此期望恒等于 $\int_{-\infty}^{\infty} h(x)\mathrm{d}F(x)$. 即

$$E[h(X)] = \int_{-\infty}^{\infty} h(x)\mathrm{d}F(x). \tag{1.3.2}$$

随机变量 X 的*方差*定义为

$$\mathrm{Var}X = E[(X-E[X])^2] = E[X^2] - E^2[X].$$

两个联合地分布的随机变量 X 和 Y 称为不相关的，若它们的*协方差*

$$\mathrm{Cov}(X,Y) = E[(X-EX)(Y-EY)] = E[XY] - E[X]E[Y]$$

是 0. 由此推出独立的随机变量是不相关的. 然而，其逆不一定正确. （读者应构想一个例子.）

期望的一个重要性质是随机变量和的期望等于它们期望的和.

9
$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]. \tag{1.3.3}$$

而方差的相应性质是

$$\mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j}\sum \mathrm{Cov}(X_i,X_j). \tag{1.3.4}$$

例 1.3 (A) 匹配问题　在一次聚会上，n 个人将自己的帽子放到房间中央，经混和后，每人随机地取一个. 我们关心的是取到自己帽子的人数 X 的均值和方差.

为了求解，我们利用表达式

$$X = X_1 + X_2 + \cdots + X_n,$$

其中

$$X_i = \begin{cases} 1 & \text{若第 } i \text{ 个人取了自己的帽子} \\ 0 & \text{其他情形} \end{cases}$$

现在，因为第 i 个人等可能地取 n 个帽子中的任意一个，所以 $P\{X_i = 1\} = 1/n$ ，因此

$$E[X_i] = 1/n, \quad \mathrm{Var}(X_i) = \frac{1}{n}\left(1-\frac{1}{n}\right) = \frac{n-1}{n^2}.$$

再则

$$\mathrm{Cov}(X_i,X_j) = E[X_iX_j] - E[X_i]E[X_j].$$

现在

$$X_iX_j = \begin{cases} 1 & \text{如果参加聚会的第 } i \text{ 个人和第 } j \text{ 个人都取到自己的帽子} \\ 0 & \text{其他情形} \end{cases}$$

于是得到

10
$$E[X_iX_j] = P\{X_i = 1, X_j = 1\} = P\{X_i = 1\}P\{X_j = 1 \mid X_i = 1\} = \frac{1}{n}\frac{1}{n-1}.$$

因此

$$\mathrm{Cov}(X_i,X_j) = \frac{1}{n(n-1)} - \left(\frac{1}{n}\right)^2 = \frac{1}{n^2(n-1)}.$$

所以，由式（1.3.3）和式（1.3.4）得

$$E[X]=1,\quad \operatorname{Var}(X)=\frac{n-1}{n}+2\binom{n}{2}\frac{1}{n^2(n-1)}=1.$$

故匹配数的均值和方差都等于1.（这种结果为什么并不惊人，对此的一个解释，可参见例1.5（f）.）■

例 1.3 (B) 一些概率等式 以 $A_1,A_2,\cdots,A_n$ 记事件，并以

$$I_j=\begin{cases}1 & \text{若 } A_j \text{ 发生}\\ 0 & \text{其他情形}\end{cases}$$

定义示性变量 $I_j, j=1,\cdots,n$. 令

$$N=\sum_{j=1}^{n} I_j,$$

则 N 是事件 $A_j, j=1,\cdots,n$ 中发生的个数. 注意

$$(1-1)^N=\begin{cases}1 & \text{若 } N=0\\ 0 & \text{若 } N>0\end{cases}\tag{1.3.5}$$

可得到一个有用的等式. 而由二项式定理，有

$$(1-1)^N=\sum_{i=0}^{N}\binom{N}{i}(-1)^i=\sum_{i=0}^{n}\binom{N}{i}(-1)^i \qquad \left(\text{因为当 } i>m \text{ 时}\binom{m}{i}=0\right).\tag{1.3.6}$$ 11

因此，若我们令

$$I=\begin{cases}1 & \text{若 } N>0\\ 0 & \text{若 } N=0\end{cases}$$

则由式（1.3.5）和式（1.3.6）可得

$$1-I=\sum_{i=0}^{n}\binom{N}{i}(-1)^i$$

或

$$I=\sum_{i=1}^{n}\binom{N}{i}(-1)^{i+1}.\tag{1.3.7}$$

在式（1.3.7）两边取期望得

$$E[I]=E[N]-E\left[\binom{N}{2}\right]+\cdots+(-1)^{n+1}E\left[\binom{N}{n}\right].\tag{1.3.8}$$

然而

$$E[I]=P\{N>0\}=P\{A_j \text{ 中至少一个发生}\}=P\left(\bigcup_1^n A_i\right)$$

而且

$$E[N]=E\left[\sum_{j=1}^{n} I_j\right]=\sum_{j=1}^{n}P(A_j),$$

$$E\left[\binom{N}{2}\right]=E[A_j \text{ 中成对出现的对的个数}]$$

$$=E\left[\sum\sum_{i<j} I_iI_j\right]=\sum\sum_{i<j}E[I_iI_j]=\sum\sum_{i<j}P(A_iA_j),$$ 12

由于同样的原因，一般地有

$$E\left[\binom{N}{i}\right]=E[A_j\text{ 中以规模为 }i\text{ 的集合出现的集合的个数}]$$
$$=E\left[\sum_{j_1<j_2<\cdots<j_i}\sum\sum I_{j_1}I_{j_2}\cdots I_{j_i}\right]=\sum_{j_1<j_2<\cdots<j_i}\sum\sum P(A_{j_1}A_{j_2}\cdots A_{j_i}).$$

因此式（1.3.8）是熟知的等式

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)-\sum_{i<j}\sum P(A_iA_j)+\sum_{i<j<k}\sum\sum P(A_iA_jA_k)$$
$$-\cdots+(-1)^{n+1}P(A_1A_2\cdots A_n).$$

的一个表述.

其他有用的等式也可用此方法导出. 例如，假设我们想求事件 $A_1,\cdots,A_n$ 中恰有 r 个出现的概率的一个公式. 则定义

$$I_r=\begin{cases}1 & \text{若 } N=r\\ 0 & \text{其他情形}\end{cases}$$

并用等式

$$\binom{N}{r}(1-1)^{N-r}=I_r$$

或

$$I_r=\binom{N}{r}\sum_{i=0}^{N-r}\binom{N-r}{i}(-1)^i=\sum_{i=0}^{n-r}\binom{N}{r}\binom{N-r}{i}(-1)^i=\sum_{i=0}^{n-r}\binom{N}{r+i}\binom{r+i}{r}(-1)^i.$$

在上式两边取期望，导出

13

$$E[I_r]=\sum_{i=0}^{n-r}(-1)^i\binom{r+i}{r}E\left[\binom{N}{r+i}\right]$$

或

$$P\{A_1,\cdots,A_n\text{ 中恰有 }r\text{ 个发生}\}=\sum_{i=0}^{n-r}(-1)^i\binom{r+i}{r}\sum_{j_1<j_2<\cdots<j_{r+i}}\cdots\sum P(A_{j_1}A_{j_2}\cdots A_{j_{r+1}}).\tag{1.3.9}$$

作为式（1.3.9）的一个应用，假设将 m 个球随机地放进 n 个盒子中，每个球等可能地进入 n 个盒子的任意一个，且与其他球的位置独立. 让我们计算恰有 r 个空盒的概率. 令 A_i 表示第 i 个盒子为空这一事件，由式（1.3.9）我们可知

$$P\{\text{恰有 }r\text{ 个盒子为空}\}=\sum_{i=0}^{n-r}(-1)^i\binom{r+i}{r}\binom{n}{r+i}\left(1-\frac{r+i}{n}\right)^m,$$

上式得自：因为和式 $\sum_{j_1<\cdots<j_{r+i}}$ 由 $\binom{n}{r+i}$ 个项组成，且和式中的每项都等于指定的 $r+i$ 个盒子为空的概率. ■

下一个例子将说明*概率方法*是什么. 这是数学家 Paul Erdos 大量使用并普及的方法，这种方法试图首先引进一个概率结构，然后用概率推理去解决确定性问题.