概率论

Probability Theory

韩 东 熊德文/编著

科 学 出 版 社

概 率 论

韩 东 熊德文 编著

科 学 出 版 社
北 京

内 容 简 介

“概率论”是高等院校数学和统计专业的基础课程之一．全书共七章，主要内容包括：随机事件及其概率、随机变量及其分布、随机变量的数值特征、多维随机变量及其分布、多维随机变量的数值特征、大数定律与中心极限定理．每章末配有习题，书末附有部分习题参考答案或提示，便于读者学习和检测所学知识．本书着眼于理论联系实际，通过精选例题并结合其他学科的问题介绍概率论的思想、模型、方法和计算，如结合复杂网络讲幂律分布；结合寿命讲 Gamma 分布；结合股价讲对数正态分布；结合风险偏好讲效用期望；结合保险费讲随机变量函数的期望；结合 VaR 讲 p-分位数；结合证券投资组合讲协方差矩阵；结合信息熵最大化讲如何确定概率分布等．本书例题丰富、叙述简洁，所有重要的结论都给出严格证明，其中包括柯尔莫哥洛夫强大数定律、Lindeberg-Feller 中心极限定理以及特征函数序列与分布函数序列之间的关系等．

本书可作为综合性大学、高等师范院校、理工科大学、财经院校本科生概率论课程的教材或参考书，也可作为各专业研究生、教师、科研与工程技术人员的参考书．

图书在版编目(CIP)数据

概率论/韩东，熊德文编著．—北京：科学出版社，2019.3

ISBN 978-7-03-060777-5

I. ①概…　II. ①韩…　②熊…　III. ①概率论　IV. ①O211

中国版本图书馆 CIP 数据核字(2019) 第 043154 号

责任编辑：李静科　李香叶／责任校对：邹慧卿

责任印制：吴兆东／封面设计：陈　敬

科学出版社 出版

北京东黄城根北街 16 号

邮政编码：100717

http://www.sciencep.com

北京建宏印刷有限公司 印刷

科学出版社发行　各地新华书店经销

*

2019 年 3 月第 一 版　开本：720 × 1000 1/16

2019 年 3 月第一次印刷　印张：12 1/2

字数：250 000

定价：58.00 元

(如有印装质量问题，我社负责调换)

前　　言

概率是随机现象 (或随机事件) 出现 (或发生) 的可能性大小的一种度量, 而概率论则是研究随机现象统计规律的一门学科, 更是人们认识和理解随机现象的一种有效工具和方法.

一般认为, 概率论起源于 16、17 世纪一些著名数学家, 如帕斯卡 (B. Pascal)、费马 (P. Fermat) 等探讨赌博中出现的各种概率计算问题. 18—20 世纪, 随着生产实践的需要, 人们提出了大量的概率问题, 这极大地促进了概率论及其应用的发展. 尤其是概率论自身发展的需要, 柯尔莫哥洛夫 (A. N. Kolmogorov) 于 1933 年提出了概率论的公理化体系, 这为概率论的蓬勃发展奠定了坚实的理论基础. 80 多年来, 概率论以其强大的生命力不断丰富和发展, 其理论和方法已被广泛应用于自然界和人类社会的各个领域. 概率论及其相关问题, 无论是理论、模型, 还是方法、计算, 其内容都非常丰富, 因此本书只能有所取舍、有所侧重地来介绍概率论的核心内容.

第 1 章概述概率论的发展简史及其地位、角色和应用.

第 2 章介绍随机事件、古典概型、几何概型、概率的公理化、条件概率、全概率公式与贝叶斯公式.

第 3 章介绍随机变量、概率分布函数、常用的离散型和连续型随机变量及其分布、随机变量函数的分布、随机数生成、离散型和连续型随机变量的模拟.

第 4 章介绍随机变量的数值特征, 主要包括: 数学期望、方差、变异系数、偏度、峰度、k-阶原点矩、k-阶中心矩、中位数、p-分位数、众数.

第 5 章介绍多维随机变量、联合分布、条件分布与条件密度、多维随机变量函数的分布、顺序统计量及其分布.

第 6 章介绍多维随机变量的数值特征, 主要包括: 多维随机变量函数的期望和方差、协方差与协方差矩阵、相关系数与相关系数矩阵、条件期望及其性质、母函数、矩母函数与特征函数、分布函数与特征函数的关系.

第 7 章介绍弱、强大数定律, 经典和一般的中心极限定理, 随机变量序列的 4 种收敛性.

在讲授概率论课程的近 20 年中, 我们曾先后选用过李贤平编著的《概率论基础》, 邓集贤等编著《概率论与数理统计教程》, R. Durrett 编著的 *Elementary Probability for Applications*, 茆诗松等编著的《概率论与数理统计教程》[8], 何书元编著的《概率论》[7], P. Olofsson 编著的 *Probability, Statistics, and Stochastic*

Processes[5] 以及 S. M. Ross 编著的 *A First Course in Probability* 作为概率论这门课的教材. 在借鉴和学习以上教材的基础上, 结合多年来的教学经验和认识, 我们编著了本书.

本书力求做到: ① 以问题为导向, 着重阐明基本概念、模型和方法的来源与背景; ② 强调用概率思想和观点来阐述基本原理与方法; ③ 着眼于理论联系实际, 通过典型的应用实例学习和掌握概率理论与方法的要义.

对于本书中加星号的小节, 教师可以根据需要选讲.

尽管本书曾以讲义的形式使用过 4 年, 但限于作者水平, 书中的内容安排、叙述方式、公式图表恐有不妥, 敬请读者批评指正.

韩　东　熊德文

2018 年 7 月

目　　录

第 1 章 概率论概述

1.1 什么是概率？概率论是什么？

概率 随机现象 (或随机事件) 出现 (或发生) 的可能性大小的一种度量.

概率论 研究随机现象统计规律的一门数学分支学科.

随机现象 在一定条件下并不总是出现相同结果 (每次结果具有一定偶然性) 的现象, 又称偶然现象, 如投硬币、掷骰子等. 与之对应的是确定现象.

统计规律 大量偶然现象所呈现的某种必然性, 它是不依从人的意志为转移的客观规律. 具体地说, 就是各种随机现象出现 (或发生) 的可能性大小的度量, 即概率或概率分布.

例 1.1.1 (高尔顿钉板) 如图 1.1 所示, 自上端放入小球, 让其自由下落, 在下落的过程中碰到钉子时, 从左边落下和从右边落下的可能性相同. 大量落下小球, 就会出现规律性.

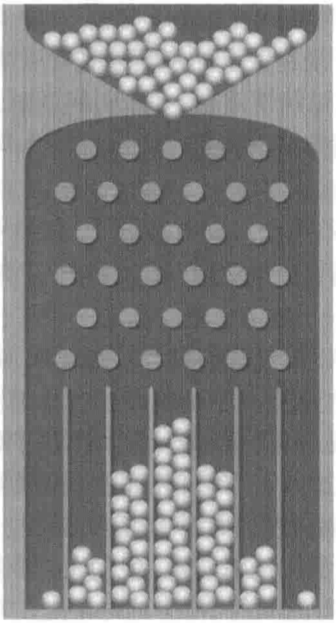

图 1.1

无论是张三还是李四, 试验结果大致相同: 小球的分布像一个“钟型”, 这就是一种统计规律, 是一种不以人的意志为转移的客观规律.

1.2 必然性与偶然性的关系

从哲学上讲, 偶然性和必然性是对应统一的, 必然性总是通过大量的偶然性表现出来, 偶然性是必然性的表现形式和必要补充. 必然性和偶然性在一定条件下相互转化. 也就是说,

(1) 必然性和偶然性是同时存在的;

(2) 必然性存在于偶然性中, 它通过大量的偶然性表现出来;

(3) 偶然性中隐藏着必然性, 它是必然性的补充和表现形式;

(4) 它们在一定条件下可相互转化.

1.3 概率论简史

“好赌似乎是人类的天性”. 追溯概率论, 可以说它起源于赌博问题的概率计算.

据记载, 人类最早的赌博游戏开始于公元前 1500 年, 古埃及人为了忘却饥饿, 经常聚在一起掷一种类似于现在骰子的东西来进行赌博. 15、16 世纪, 意大利数学家卡尔丹、塔塔利亚等研究过赌博问题, 并未引起当时人们的注意.

1654 年左右, 爱好赌博的法国贵族 Méray(梅尔) 向 Pascal(帕斯卡) 提出了两个问题:

(1) 赌金分配问题: 甲乙两个人同时掷骰子, 如果甲先掷出了 3 次“6”点, 或乙先掷出 3 次“4”点, 就赢了全部赌金. 若甲已有 2 次“6”点, 乙有 1 次“4”点, 问如何分配赌金?

(2) 一对骰子抛掷 25 次, 把赌注押到“至少出现一次双‘6’点”是否比“完全不出现双‘6’点”有利?

Pascal 和他的好友 Fermat(费马, 法) 进行通信讨论, 后来, 荷兰数学家、物理学家 Huyghens (惠更斯) 也加入了讨论, 并写了《论赌博中计算》一书.

18、19 世纪, 随着社会的发展和生产实践的需要, 特别是在人口统计、保险、测量、射击等方面提出了大量的概率问题, 促使人们在概率的极限定理 (大数定律和中心极限定理) 等方面进行深入的研究, 这时期先后对概率论发展做出重要贡献的数学家有: Bernoulli(伯努利, 瑞士), Laplace(拉普拉斯, 法), Poisson(泊松, 法), Gauss(高斯, 德).

20 世纪初, 俄罗斯学派做出重要贡献: Chebyshev(切比雪夫不等式, 俄), Markov (马尔可夫过程, 俄).

1933 年, Kolmogorov(柯尔莫哥洛夫, 俄) 提出了概率论的公理化体系, 这不仅部分地回答了 Hilbert 23 个问题中的第 6 个问题:“物理学的公理化”, Hilbert 建议用数学的公理化方法推演出全部物理学, 首先是在概率论和力学领域, 更重要的是, 它为概率论的蓬勃发展和广泛应用奠定了坚实的理论基础.

1.4 概率论的地位、角色和应用

概率论在数学与统计学中的地位与角色见图 1.2.

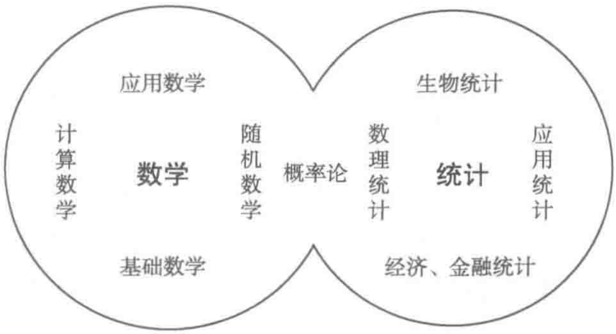

图 1.2

关于概率论的应用主要体现在以下方面:

统计物理: 用概率统计的方法对由大量微观粒子所构成的宏观物体的物理性质及宏观规律做出微观解释. 由微观粒子状态的等概率性假设导出宏观物理性质——热力学性质.

1827 年, 布朗发现花粉在液体中做不规则运动. 1905 年, Einstein(爱因斯坦) 依据分子运动论的原理给出了分子运动的统计规律 (分布), 为证实分子的存在性找到了一种方法, 同时也阐明了布朗运动的根源及其规律性.

量子力学: 薛定谔运用概率波函数得到了薛定谔方程. 根据海森伯测不准原理, 微观粒子的位置与动量不可同时被精确确定, 这时我们需要用概率方法来建模. 在高能物理中, 量子能级分布可用随机矩阵的谱分布来刻画.

天气预报: 在实际生活中, 我们可以用概率方法来预报降水的概率, 描述 PM2.5 浓度的分布.

金融与经济: 1900 年, Bachelier 的博士学位论文 (投机理论) 首次运用布朗运动来描述股价; 保险精算中发生理赔的次数; 在证券投资组合中, 各资产的收益; 期权定价等, 都需要用到概率模型和方法来描述与分析.

医学: 在临床医学中, 各种疾病发病率、疗效率、死亡率等都需要用概率统计的方法进行估计与分析.

军事: 我们可能运用概率模型和方法分析各种武器的命中率、失效率、杀伤力、性能.

社会学: 我们常用出生率、死亡率建立人口数量变化的随机模型, 描述人口变化规律. 通过分析《红楼梦》各章中虚词出现的频率及它们的相互关系来分析、判断《红楼梦》的作者. 我们可以用随机网络的度分布来描述人际关系网络、论文相互引用网络、各公司资金流动网络等.

上述都涉及利用概率方法进行建模与计算.

第 2 章　随机事件及其概率

这一章我们将讨论随机事件及其运算规律, 古典概型和几何概型, 概率的公理化定义及其性质, 条件概率与随机事件的独立性、乘法公式、全概率公式和贝叶斯公式, 这些是学习以后各章的基础.

2.1　随机事件

问题　什么是随机事件? 其数学本质是什么? 有什么运算规律?

2.1.1　样本空间和随机事件

随机试验　对随机现象的观测, 我们称为随机试验.

样本空间　随机试验中的所有可能结果的全体, 记为 Ω.

样本点　随机试验中一种可能的结果, 或者说样本空间 Ω 中的元素 ω.

例 2.1.1　(1) 观察新生婴儿的性别, $\Omega=\{$男孩, 女孩$\}$;

(2) 五一节的天气, $\Omega=\{$下雨, 不下雨$\}$;

(3) 五一节前最后一个交易日的股票价格, $\Omega=\{x;x\geqslant 0\}$,

(4) 本周末光顾上海教育超市的人数, $\Omega=\{0,1,2,\cdots\}$.

随机事件　通俗讲, 随机事件是在观察随机现象中, 可能发生也可能不发生, 而在大量重复试验中具有某种规律性的事件, 简称事件. 从数学上说, 随机事件是样本点的集合, 或者说 Ω 的子集, 记为 A, B, C, $\cdots$.

基本事件　只包含一个样本点的随机事件 (单点集).

不可能事件　不包含任何样本点的随机事件 (空集), $\varnothing$.

必然事件　包含所有样本点的集合, Ω.

$$\boxed{\text{事件 } A \text{ 发生, 当仅当试验结果 } \omega\in A.}$$

例 2.1.2　掷骰子一次, 样本空间为 $\Omega=\{1,2,3,4,5,6\}$. $A=$“点数为奇数”$=\{1,3,5\}$ 为随机事件, 如果试验结果出现点数 1 或 3 或 5, 则 A 发生; 如果结果出现点数 2 或 4 或 6, 则 A 就没有发生.

2.1.2　(随机) 事件的运算

由于事件本质上是集合, 所以事件之间的关系与运算就转化为集合间的关系与运算. 这里我们要着重理解其概率解释与概率含义.

(1) 事件的并 (集合并集): $A \cup B$——A 事件发生或 B 事件发生, A, B 至少有一个发生. 如图 2.1 所示.

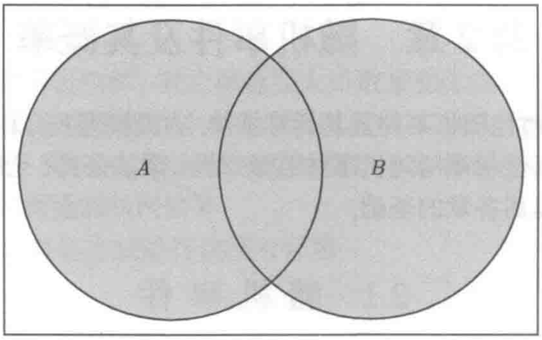

图 2.1

$\bigcup_{i=1}^{n} A_i$ 表示 $A_1, \cdots, A_n$ 中至少有一个发生;

$\bigcup_{i=1}^{\infty} A_i$ 表示 $A_1, \cdots, A_n, \cdots$ 中至少有一个发生.

(2) 事件的交 (集合交集): $A \cap B$——A, B 事件同时发生, 有时也记为 AB. 如图 2.2 所示.

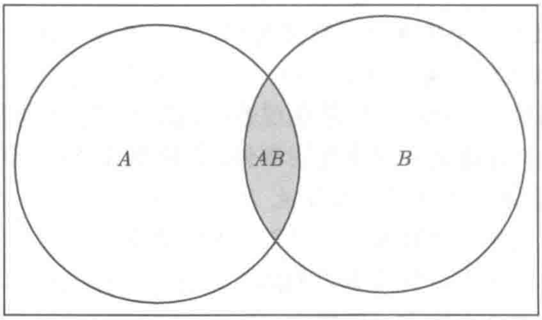

图 2.2

$\bigcap_{i=1}^{n} A_i$ 表示 $A_1, \cdots, A_n$ 同时发生;

$\bigcap_{i=1}^{\infty} A_i$ 表示 $A_1, \cdots, A_n, \cdots$(可列无穷多个) 同时发生.

例 2.1.3　掷骰子一次, 令 $A = \{2, 4, 6\}$, $B = \{点数 \geqslant 4\}$, 则 $A \cap B = \{4, 6\}$.

(3) 事件的差 (集合的差): $B - A = B - (A \cap B)$——B 发生且 A 不发生, 如图 2.3 所示.

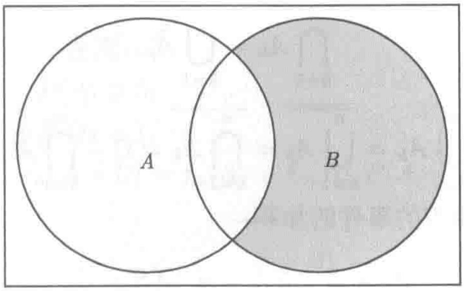

图 2.3

互不相容 (不相交集合) $AB = \varnothing$—— A, B 不能同时发生.

例 2.1.4 掷骰子一次, $A = \{1\}$, $B = \{2\}$, 则 A, B 互不相容, 结果不可能出现点数既为 1 又为 2.

若 $AB = \varnothing$, 则两个事件的并可以写成两个事件的加和

$$A \cup B = A + B.$$

逆事件 (补集合) $\bar{A} = \Omega - A$, 又称为事件 A 的**对立事件**.

对称差 $A \cup B - (A \cap B) = (A - AB) + (B - AB)$. A, B 至少有一个发生, 但 A, B 不能同时发生.

事件运算满足通常的交换律、结合律和分配律:

(1) 交换律: $A \cup B = B \cup A$;

(2) 结合律: $A \cup (B \cup C) = (A \cup B) \cup C$;

(3) 分配律: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;

(4) 对偶律 (De Morgan 律): $\overline{A \cup B} = \bar{A} \cap \bar{B}$,
$\overline{A \cap B} = \bar{A} \cup \bar{B}$.

证明 仅证 (4): $\forall \omega \in \overline{A \cup B} = \Omega - (A \cup B)$, 易知 $\omega \notin (A \cup B) \Rightarrow \omega \notin A$ 且 $\omega \notin B$, 于是, $\omega \in \bar{A}$ 且 $\omega \in \bar{B}$, 即

$$\overline{A \cup B} \subset \bar{A} \cap \bar{B}.$$

同理证 $\overline{A \cup B} \supset \bar{A} \cap \bar{B}$. □

一般地,

$$\overline{\bigcup_{k=1}^{n} A_k} = \bigcap_{k=1}^{n} \bar{A}_k,$$

$$\overline{\bigcap_{k=1}^{n} A_k} = \bigcup_{k=1}^{n} \bar{A}_k,$$

$$\bigcup_{k=1}^{n} A_k = \overline{\overline{\bigcup_{k=1}^{n} A_k}} = \overline{\bigcap_{k=1}^{n} \bar{A}_k} = \Omega - \bigcap_{k=1}^{n} \bar{A}_k.$$

事件并可以写成互不相容的事件的加和:

$$A \cup B = A + (B - A) = A + B\bar{A} = A + (B - AB) = B + (A - AB),$$

多个事件的并也有类似的写法:

$$\begin{aligned}
\bigcup_{k=1}^{n} A_k &= A_1 + (A_2 - A_1) + (A_3 - A_2 - A_1) + \cdots + (A_n - A_{n-1} - A_{n-2} - \cdots - A_1) \\
&= A_1 + A_2\bar{A}_1 + A_3\bar{A}_2\bar{A}_1 + \cdots + A_n\bar{A}_{n-1}\cdots\bar{A}_1 \\
&= A_1 + \sum_{k=2}^{n} A_k\bar{A}_{k-1}\cdots\bar{A}_1.
\end{aligned}$$

2.2 频率及其性质

问题 随机事件发生的概率是一个客观存在的数. 能不能用试验来估计这个数呢?

定义 2.2.1 设 A 为 (随机) 事件, 重复观察 n 次, $n(A)$ 表示在 n 次试验 (或观察) 中 A 事件发生的次数, 则称

$$f_n(A) := \frac{n(A)}{n}$$

为事件 A 发生的 **(相对) 频率**.

例 2.2.2 掷硬币, 历史上很多统计学家都掷过硬币, 结果见表 2.1.

表 2.1　历史上掷硬币的试验结果

试验者	总次数	正面出现次数	频率 (f_n)
De Morgan	2048	1061	0.5181
Buffon	4040	2048	0.5069
Feller	10000	4979	0.4979
Pearson	24000	12012	0.5005

可以看出, 当 n 很大的时候, 正面出现的频率越来越稳定在 0.5 附近.

一般地, 如果事件 A 发生的相对频率的极限值存在, 我们就称这个极限值为 A 发生的概率, 记为 $P(A)$, 即

$$f_n(A) \longrightarrow P(A), \quad n \longrightarrow \infty.$$

更进一步, 我们有如下性质:

(1) $f_n(A) \geqslant 0$, 从而 $P(A) \geqslant 0$;

(2) $f_n(\Omega) = 1$, 从而 $P(\Omega) = 1$;

(3) $A \cap B = \varnothing$, 则 $f_n(A+B) = f_n(A) + f_n(B), P(A+B) = P(A) + P(B)$.

事实上,

$$\begin{aligned} f_n(A+B) &= \frac{n(A+B)}{n} \\ &= \frac{n(A)+n(B)}{n} \\ &= \frac{n(A)}{n} + \frac{n(B)}{n} \\ &\longrightarrow P(A) + P(B). \end{aligned}$$

一般地, 如果 $A_1, \cdots, A_m$ 互不相容, 则 $f_n\left(\sum\limits_{k=1}^{m} A_k\right) = \sum\limits_{k=1}^{m} f_n(A_k)$, 于是概率也具有如下性质:

$$P\left(\sum_{k=1}^{m} A_k\right) = \sum_{k=1}^{m} P(A_k).$$

历史上掷硬币的试验结果 (相对频率) 见表 2.1, 利用大数定律 (见第 7 章) 可以证明:

(1) 当掷硬币是在相同一条件下且每次互不影响, A={出现正面}, 则

$$\lim_{n\longrightarrow\infty} f_n(A) = \lim_{n\longrightarrow\infty} \frac{n(A)}{n} = \frac{1}{2}(\text{几乎处处收敛, 见第 7 章}).$$

(2) 显然, 概率如果都用频率的极限值来确定, 一是试验做不到无穷次, 二是有些事件没法做试验, 三是有时也不需要做试验, 如对均匀对称的骰子, 无须做试验, 也可断定其各点数出现的概率都相等, 都为 $\frac{1}{6}$.

2.3 古典概型

问题 一个口袋有 a 个红球, b 个白球, 不放回地取球多次, 求第 k 次取出红球的概率?

定义 2.3.1 设样本空间 Ω 为有限集, 若每个样本点 (基本事件) 发生的概率相同, 我们就称其为**古典概型**. 对任意 $A \subset \Omega$, 事件 A 发生的概率为

$$P(A) = \frac{|A|}{|\Omega|} = \frac{A\text{ 中包含样本点的个数}}{\Omega\text{ 中包含样本点的个数}},$$

其中 $|A|$ 表示 A 中包含基本事件的个数.