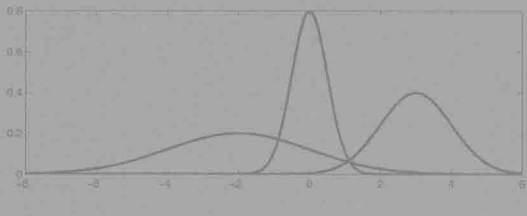
新世纪高等学校教材

数学与应用数学系列教材

概率论

GAILÜLUN

北京师范大学数学科学学院 主 编
李 勇 编 著



$$\varphi_{a,\sigma}(x) \triangleq \frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-a)^2}{2\sigma^2}}$$

北京师范大学出版集团
BEIJING NORMAL UNIVERSITY PUBLISHING GROUP
北京师范大学出版社

新世纪高等学校教材

数学与应用数学系列教材

概率论

GAILÜLUN

北京师范大学数学科学学院 主编

李勇 编著

北京师范大学出版集团
BEIJING NORMAL UNIVERSITY PUBLISHING GROUP
北京师范大学出版社

图书在版编目（CIP）数据

概率论 / 李勇编著. —北京：北京师范大学出版社，2013.8
（新世纪高等学校教材. 数学与应用数学系列教材）
ISBN 978-7-303-16761-6

Ⅰ.①概… Ⅱ.①李… Ⅲ.①概率论-高等学校-教材
Ⅳ.①O211

中国版本图书馆CIP数据核字（2013）第172822号

营销中心电话 010-58802181 58805532
北师大出版社高等教育分社网 http://gaojiao.bnup.com
电子信箱 gaojiao@bnupg.com

出版发行：北京师范大学出版社 www.bnup.com
北京新街口外大街19号
邮政编码：100875
印 刷：北京京师印务有限公司
经 销：全国新华书店
开 本：170 mm × 230 mm
印 张：15
字 数：250千字
版 次：2013年8月第1版
印 次：2013年8月第1次印刷
定 价：25.00元

策划编辑：岳昌庆 责任编辑：岳昌庆 程丽娟
美术编辑：毛 佳 装帧设计：毛 佳
责任校对：李 菡 责任印制：孙文凯

前 言

1915年北京高等师范学校成立数理部，1922年成立数学系．2004年成立北京师范大学数学科学学院．经过近百年的风风雨雨，数学科学学院在学科建设、人才培养和教学实践中积累了丰富的经验．将这些经验落实并贯彻到教材编著中去是大有益处的．

1980年，北京师范大学出版社成立，给教材的出版提供了一个很好的契机．北京师范大学数学科学学院教师编著的多数教材已先后在这里出版．除《北京师范大学现代数学丛书》外，就大学教材而言，共有5种版本．第1种是列出编委会的《高等学校教学用书》，这是在1985年，由我校出版社编写出版了1套(17部)数学系本科生教材和非数学专业高等数学教材．在出版社的大力支持下，这一计划完全实现，满足了当时教学的需要．第2种是未列编委会的《高等学校教学用书》．第3种是《面向21世纪课程教材》．第4种是《北京师范大学现代数学课程教材》．第5种是未标注“高等学校教学用书”，但实际上是高等学校教学用书．在这些教材中，除再次印刷外，已经有多部教材进行了修订或出版了第2版．

2005年5月，李仲来教授汇总了北京师范大学数学科学学院教师在北京师范大学出版社出版的全部著作，由李仲来教授与北京师范大学出版社理科编辑部岳昌庆、王松浦进行了沟通和协商，由北京师范大学数学科学学院主编(李仲来教授负责)，准备对学院教师目前使用的，或北京师范大学出版社已经没有存书的部分教材进行修订后再版, 另有一些教材需要重新编写．计划用几年时间，出版数学与应用数学系列教材、数学教育主干课程系列教材、大学公共课数学系列教材、数学学科硕士研究生系列教材，共4个系列的主要课程教材．

由学院组织和动员全院在职和退休教师之力量，主编出版数学一级学科4个系列的60余部主要课程教材．教材编写涉及面如此之广和数量之大，持续时间之长，这在一所高校数学院系内是为数不多的，其数量在中国数学界列全国第一．经过8年的编写，至今已经出版了50余部教材，原计划的大多数教材已经出版，对于学院来讲，这是一件值得庆贺的大事．现在可以说，数学科学学院和北京师范大学出版社基本上是干成了一件大事．这是很难办成圆满的一件大事．剩下的一些教材在两三年内多数可以出版．若留下缺憾，则需要后人去补充．

从数量上看，按教材系列，出版数学与应用数学系列教材28部、数学教育主干课程系列教材9部、大学公共课数学系列教材7部、数学学科硕士研究生系列教材9部．按出版教材版次，第1版21部、第2版21部、第3版12部．还

出版了3部教辅教材.

从质量上看，7部教材被评为普通高等教育“十一五”国家级规划教材；7部教材被评为普通高等教育本科“十二五”国家级规划教材；7部教材被评为北京市高等教育精品教材；“师范院校数学学科4个系列教材建设”项目获2012年北京师范大学教育教学成果一等奖.

本套教材可供高等院校本科生、教育学院数学系、函授(数学专业)和在职中学教师等使用和参考. 希望使用这些教材的校内外专家学者和广大读者，提出宝贵的修改意见，使其不断改进和完善. (李仲来执笔)

北京师范大学数学科学学院

2013年2月18日

作者的话

时间飞逝，不经意间执教已有31载，期间为北京师范大学数学科学学院本科生讲授“概率论”课程15次以上. 早有出书想法，苦于教学科研任务繁重，无暇顾及. 最后还是在夫人的关心、鼓励和支持下，完成这一愿望.

概率论是研究随机现象的数学分支，是统计学的理论基础，有广泛的理论与应用价值. 本人一直从事概率论与数理统计教学科研工作，期望通过本书展现拙见，与读者分享.

本书试图从整体上再现概率论知识构建过程，展现相关重要知识点的来龙去脉，逐步提高读者的概率知识水平. 第一章主要涉及直到19世纪末的概率知识，以频率为主线介绍概率论基本概念与模型，为后面的概率论公理体系做铺垫. 第二章主要展示现代概率论的基石——概率论公理化知识体系，介绍概率空间的基本性质，使读者概率知识水平上一台阶. 第三章主要介绍随机变量及相关概念，将数学分析工具引入随机现象研究过程，提高概率理论研究效率. 第四章基于数学期望讨论随机变量的数字特征及其概率内涵，引入刻画随机变量分布规律的等价工具特征函数，使读者的理论与应用知识水再上一台阶. 第五章研究随机变量的极限理论，其中的大数定律和中心极限定理成为现代统计学理论与应用的基石之一. 本书配有丰富的练习题，并提供了部分习题答案或证明思路，以帮助读者更好掌握概率知识. 本书还配有名词与符号索引，方便读者在书中查找相关知识内容.

在介绍各个知识点的过程中，本书融合个人一些观点，期望对于读者理解概率知识有所帮助.

对于随机现象，本书明确指出它应该具有频率的稳定性，以使读者明晰随机现象是一种特殊的不确定现象，进而正确理解概率论研究对象. 本书没沿用传统的古典概率计算公式定义古典概型，目的是使读者更好理解古典概型的本质（样本空间有限和等概率），同时又能体会当时概率理论上的不完备之处（严格证明古典概率计算公式需要概率论公理），为介绍概率论公理体系埋下伏笔. 本书花了相当的篇幅介绍概率空间的两个重要组成部件——样本空间和事件域，以使读者体会样本空间和事件域的定义具有灵活性，根据实际问题适当地构造样本空间和事件域，能达事半功倍的效果.

本书没有过分强调古典概率和几何概率的初等计算技巧，而将重点放在利用概率性质解决古典率和几何概率计算问题，使读者通过这些问题巩固概率空间相关知识，体会公理知识体系的威力. 本书通过缩小概率空间的思路引入条件概率，意在强调这是计算条件概率的有效途径之一. 本书花一定篇幅介绍随机试验的独立性，通过乘积概率空间的建立过程，逐步向读者展示问题解决过程，潜移默化地培养他们解决问题的能力.

本书从概率空间数量化角度引入随机变量，以重现随机变量的产生过程，使读者加深对于随机变量、分布及其分布函数概念的理解，熏陶其创造性思维思想. 本书花相当多的篇幅介绍随机变量的构造，通过随机变量的值域引入简单随机变量，避免了经典教科书中简单随机变量表达方式不唯一的情况，以简化随机变量逼近框架，为抽象定义数学期望奠定基础. 在随机变量函数的分布处理上，本书借鉴了何书元教授的想法 [1]，充分利用分布函数的定义与复合函数求导法则，解决其密度函数的计算问题，这种方法思路清晰自然且简单易行，免去读者记忆密度函数计算公式的负担.

基于随机变量的构造，本书借助简单随机变量的特殊定义，将离散型和连续型随机变量的数学期望统一定义，使读者能深入理解数学期望本质，为抽象定义随机变量的数字特征和特征函数奠定基础.

关于极限理论知识，本书包含大量逻辑推理过程，其用意在于查漏补缺，帮助读者利用推理过程复习和巩固前四章的概率知识.

为使读者更好地认识概率极限理论的知识结构，把握推导思路，本书主要的知识点放到各节的前部，而将相应的预备知识作为引理放到各节的后部. 在教师教学过程中，可视学生情况决定是否详细介绍这些引理. 希望读者将基本概念和重要结论的理解与应用放在首位，仅是那些对于理论特别感兴趣者可尝试进一步理解与掌握证明结论的思想方法.

本书基于自编的 $\LaTeX\,2_\varepsilon$ 投影教学课件而成，最初由韩珊珊、张宇静和王锦硕士将课件初步转换为$\LaTeX\,2_\varepsilon$书稿，然后由郑珩、盛瑶、黄勍和黄文贤博士等做了整理和完善，最后李斯名和丛珊等同志对书稿进行了详尽的修改，并配置习题及解答. 在此向上述各位表示感谢.

由于作者水平有限，书中难免会有不妥之处，敬请读者批评指教.

李 勇

2013年2月16日于北京师范大学丽泽区

目 录

第一章　绪论

17世纪中叶，赌博业的发展遇到许多无法解决的公平赌博问题，当时的著名数学家，如Fermat，Pascal，Huygens等都参加了相关问题的讨论，由此引出古典概型. 19世纪末至20世纪初，公理化体系奠定了概率论的严格数学基础，使得概率论作为一个数学分支得以迅速地发展. 本章简述概率论的基本概念、古典概型与几何概型，以及概率研究的早期想法，为第二章介绍概率论公理化体系奠定基础.

§1.1　随机现象及基本概念

本节简要介绍随机现象的基本概念，更加详细的材料可查阅参考文献所列书目（如参考书 [2–4]等）.

§1.1.1　必然现象与随机现象

随机现象是概率论和统计学的最基本概念之一, 与之相关的概念是必然现象和不确定现象, 下面简要介绍这些概念及它们之间的关系.

案例 1.1　*在标准大气压下，纯净的水在100°C会沸腾.*

在案例1.1中，可以把“标准大气压”“纯净的水”和“100°C”看成条件，把“沸腾”和“不沸腾”看成结果. 像这类在给定条件下能预知结果的现象称为**必然现象**.

案例 1.2　*投掷一枚质地均匀的硬币，结果不是正面向上，就是反面向上.*

在案例1.2中，可以把“投掷一枚质地均匀的硬币”看成条件，把“正面向上”和“反面向上”看成是投掷硬币的结果.

我们无法预言投掷的结果，因该结果是由条件之外的一些不确定因素所决定，如硬币出手时的速度和角度、硬币落地点的弹性和光滑程度、气流的

运动等. 像这类在给定条件下不能预知结果的现象称为**不确定现象**.

人们对不确定现象表现出异常的兴趣，总想知道其不确定的原因. 虽然我们不能预知不确定现象的结果，但是却容易得到其结果的范围，我们可以通过重复观测这一现象，了解其内在规律.

人们是通过分析观测结果来研究不确定现象. 例如对于案例1.2，很多学者做了大量的实验，一些结果列入表1中. 这些实验表明：实验次数很大时，出现正面的次数与实验的总次数之比近似于0.5.

表 1 投硬币实验列表

实验者	实验次数	正面次数	正面频率
Buffon	4 040	2 048	0.506 9
De Morgan	4 092	2 048	0.500 5
Feller	10 000	4 979	0.497 9
Pearson	12 000	6 019	0.501 6
Pearson	24 000	12 012	0.500 5
Romannovski	80 640	39 699	0.492 3

在针对某一不确定现象的多次重复实验中，某一结果出现的次数与实验的次数之比称作该结果的**频率**. 表1向我们展示了这样一个规律：随着实验次数的增加，各个结果的频率趋于稳定. 如果一个不确定现象具有这种规律，就称该现象具有**频率稳定性**，并称该现象为**随机现象**. 随机现象十分常见，是人们的重要研究对象，**概率论**就是研究随机现象的一个数学分支.

案例 1.3 Galton**钉板**如图1.1所示，其主体结构是一个上面带有小孔A的箱子，小圆球可以从孔落入箱中. 箱除前面板是透明的玻璃外，其他各面都由木板制成，透过前面板可以看到钉在后面板上的交错排列的钉子，各排相邻两钉子的间距比小圆球的直径稍大. 在箱子的底板上有等距排列的隔板，将箱底隔成10个形状相同的小盒. 从箱顶的小孔中放一个小球，小球最终将落入其中一个盒.

在案例1.3中，给定的条件是“从箱顶的小孔中放一个小球”，由于小球落入箱中后要和每一排的钉子都发生碰撞，才能落到箱底的一个盒中. 而每次碰撞，使得小球等可能地向左或向右落下，因此不能预知小球落到哪个盒中，这是一个不确定现象.

为表示这个不确定现象的所有可能出现的结果，分别将小盒子从左到右编号为$1, 2, \cdots, 10$，用i表示“小球落入第i号箱中”的结果，则所有可能结果

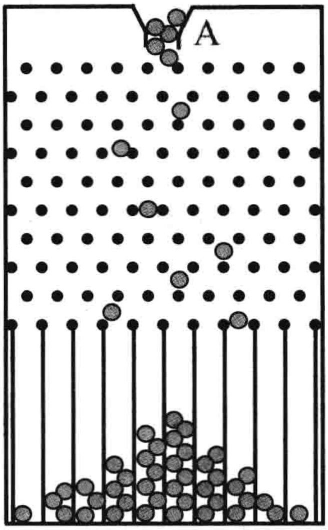


图 1.1 Galton钉板实验

可以表示为

$$\mathscr{F} = \{A : A \subset \Omega\},$$

其中$\Omega = \{1, 2, \cdots, 10\}$.

大量重复Galton钉板实验，可以统计各个结果的频率. 实验结果表明：随着实验的次数增加，它们稳定于特定的常数. 因此我们所关心的现象有频率的稳定性，是随机现象.

如图1.1所示，如果有大量的小球逐一从箱顶的小孔中投入，结果呈现如下规律：落在中间格子中的小球多. 在Galton钉板实验中，所有可能出现的结果都具有频率的稳定性，我们称这类实验为**随机实验**，简称为**实验**.

§1.1.2 样本空间

在实验中，有些结果不需要分解成更小的结果，有些结果则可以分解成其他结果的组合. 如在案例1.3中，结果“小球落入1号盒”是不必再分解的结果；结果“小球落入盒的编号小于3”可以分解为“小球落入1号盒”和“小球落入2号盒”. 为表达方便，将随机实验中无须再分解的结果称为**样本点**，常用ω表示.

定义 1.1.1 全体样本点所组成的集合称为**样本空间**，常用Ω表示.

任何实验的所有结果都可以用样本点表示，并且各个样本点的频率稳定性可以保证所有结果的频率的稳定性.

在案例1.2掷骰子实验中，所有的样本点为“正面”和“反面”，样本空间

$$\Omega = \{\text{“正面”}, \text{“反面”}\};$$

在案例1.3 Galton钉板实验中，所有的样本点为$1, 2, \cdots, 10$，样本空间

$$\Omega = \{1, 2, \cdots, 10\}.$$

例 1.1.1 同时抛两枚硬币的实验的样本空间是什么？

解 用x_i表示第i枚硬币的投掷结果，其中x_1 和x_2 都可代表“正面”或“反面”. 所有的样本点可以表示为$\omega = (x_1, x_2)$，因而样本空间

$$\Omega = \{(x_1, x_2):\ x_i \in \{\text{正面，反面}\}, i = 1, 2\}.$$

∎

需注意：对同一实验，若关心的对象不同，其样本点和样本空间的表达方式也不同. 如在例1.1.1中只关心出现正面的次数，可以用$0, 1$和2 分别代表“出现0个正面”“出现1个正面”和“出现2个正面”，则样本空间变为$\Omega = \{0, 1, 2\}$.

例 1.1.2 某热线服务电话1h内打进电话次数，试给出一个描述此现象的样本空间.

解 用i表示“1h内打进i次电话”，样本空间为$\Omega = \{0, 1, 2, \cdots\}$. ∎

为方便，称由有限个样本点构成的样本空间为**有限样本空间**，称由无穷多个样本点构成的样本空间为**无穷样本空间**. 例1.1.1中的样本空间为有限样本空间，例1.1.2 中的样本空间为无穷样本空间.

§1.1.3 事件及运算

样本点是构成事件的基石，下讨论事件的初步定义及运算.

定义 1.1.2 由样本点组成的集合称为**事件**；一定要发生的事件称为**必然事件**，用Ω表示；一定不发生的事件称为**不可能事件**，用$\varnothing$表示.

事件是一种特殊的集合，结果导致与集合相关的概念有特定的内涵，下将逐步介绍. 为方便，常用大写字母A, B, C等表示事件.

考察随机实验和事件A，如果实验所出现的样本点ω在事件A中，称**事件A发生**，记为$\omega \in A$；实验所出现的样本点ω不在事件A中，称**事件A不发生**，记为$\omega \notin A$.

定义 1.1.3 如果A发生能够推出B发生，称A**包含于**B，或A**被**B**包含**，记为$A \subset B$；也可称为B**包含**A，记为$B \supset A$. 如果A包含B，同时B包含A，称A**等于**B，记为$A = B$.

显然，事件之间的包含（被包含，或等于）关系等价于将它们看作集合之间的包含（被包含，或等于）关系.

定义 1.1.4 对于事件A和B，称

$$A \cup B \triangleq \{\omega:\ \omega \in A \text{ 或} \omega \in B\} \tag{1.1}$$

为A和B之**并运算**；称

$$A \cap B \triangleq \{\omega:\ \omega \in A \text{ 且} \omega \in B\} \tag{1.2}$$

为A和B之**交运算**，简记为AB；称

$$A - B \triangleq \{\omega:\ \omega \in A \text{ 且} \omega \notin B\} \tag{1.3}$$

为A和B之**差运算**，也常记为$A\backslash B$；称

$$\overline{A} \triangleq \{\omega:\ \omega \notin A\} \tag{1.4}$$

为A之**余（补）运算**.

显然，$A\cup B$表示“A和B至少有一个发生”的事件；$A\cap B$表示“A和B同时发生”的事件；$A-B$表示“A和B中仅事件A发生”的事件；$\overline{A}$表示“A不发生”的事件. 事件之间的这些运算有一些性质，如差运算可以用交和补运算表示，补运算可以用差运算表示，详见练习1.1.3.

定义 1.1.5 如果事件A和B满足$A \cap B = \varnothing$，称事件A和B**不相容**.

A和B不相容表示事件A和B不能同时发生，即A和B没有公共的样本点. 可以用平面上的封闭曲线的内部代表事件，进而示意事件之间的关系和运算，即Venn图，如图1.2所示. 下面讨论多个事件的运算.

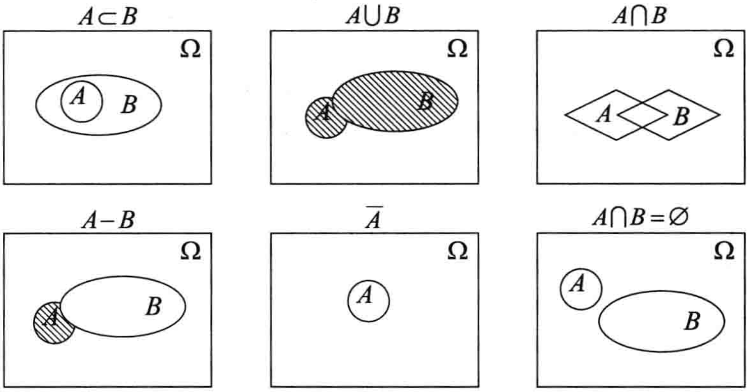


图 1.2 Venn图：示意事件间关系和运算

定义 1.1.6 设$\mathscr{D}$为集合，且对任意$i\in\mathscr{D}$，A_i是事件. 称

$$\bigcup_{i\in\mathscr{D}} A_i=\{\omega:\ \text{存在}i\in\mathscr{D}\ \text{使}\omega\in A_i\} \tag{1.5}$$

为事件类$\{A_i:i\in\mathscr{D}\}$之并；称

$$\bigcap_{i\in\mathscr{D}} A_i=\{\omega:\ \forall i\in\mathscr{D}\ \text{有}\omega\in A_i\} \tag{1.6}$$

为事件类$\{A_i:i\in\mathscr{D}\}$之交.

显然，$\bigcup\limits_{i\in\mathscr{D}} A_i$为事件，它表示“事件类$\{A_i:i\in\mathscr{D}\}$中至少有一个事件发生”的事件；$\bigcap\limits_{i\in\mathscr{D}} A_i$为事件，它表示“事件类$\{A_i:i\in\mathscr{D}\}$中所有事件同时发生”的事件. 当$\mathscr{D}=\{n,n+1,\cdots,m\}$时，记

$$\bigcup_{k=n}^{m} A_k\triangleq\bigcup_{i\in\mathscr{D}} A_i,\quad \bigcap_{k=n}^{m} A_k\triangleq\bigcap_{i\in\mathscr{D}} A_i, \tag{1.7}$$

这里$n\leqslant m$可以是整数或正负无穷，下标k和i可用其他字母；当通过上下文可以判断出指标的变化范围时，可以省略指标集的标注，即

$$\bigcup_{i} A_i\triangleq\bigcup_{i\in\mathscr{D}} A_i,\quad \bigcap_{i} A_i\triangleq\bigcap_{i\in\mathscr{D}} A_i. \tag{1.8}$$

下面讨论**事件列**$\{A_n\}\triangleq\{A_n:n\geqslant 1\}$的极限概念.