Natural Axiom System of Probability Theory
—— mathematical model of random universe

概率论自然公理系统

随 机 世 界 的 数 学 模 型

熊大国
著

清华大学出版社
http://www.tup.tsinghua.edu.cn

概率论自然公理系统

——随机世界的数学模型

熊大国　著

清 华 大 学 出 版 社

(京)新登字 158 号

内 容 提 要

本著作发展概率论的柯尔莫哥洛夫公理系统，建立现实随机世界的数学模型，形成自然公理系统。

全书共 3 章，其中第 1 章探讨事件和概率的本质以及它们在科学中的巨大作用；第 2 章介绍自然公理系统，给出 6 组公理，这些公理将随机世界抽象成因果空间，并在因果空间中引进两类“图形”——随机试验和概率空间，进一步引入条件概率和独立性等概念；第 3 章介绍随机变量的基本概念，论证它与所建立的公理系统之间的关系。

本书可供从事概率论及相关领域研究工作的人员参考。

书　名：概率论自然公理系统——随机世界的数学模型
作　者：熊大国
出版者：清华大学出版社（北京清华大学学研大厦，邮编 100084）
　　　　http://www.tup.tsinghua.edu.cn
印刷者：清华大学印刷厂
发行者：新华书店总店北京发行所
开　本：850×1168　1/32　印张：6　字数：154 千字
版　次：2000 年 11 月第 1 版　2000 年 11 月第 1 次印刷
书　号：ISBN 7-302-00802-7/O・114
印　数：0001～2000
定　价：14.00 元

序

概率论是研究现实的随机世界，即自然界和人类社会中随机现象的一门数学分支。它的基础是由柯尔莫哥洛夫(Kolmogorov)公理系统(简称为K公理系统)奠定的。在K公理系统中概率空间$(\Omega,\mathscr{F},P)$是一切研究的出发点。其中Ω是某个集合，它的元素称为基本事件；$\mathscr{F}$是Ω上的σ域，它的元素称为事件；P是$\mathscr{F}$上的概率测度，数值$P(A)$是事件A出现可能性的度量。于是，集合论和测度论方法被硬性地搬到概率论中，并用它推演出概率论的理论体系。

这种研究模式至少存在五个缺点。① 它没有建立随机世界的整体形象。作为类比，初等几何建立了自己的整体形象——欧氏空间，使得图形及其度量的研究既形象又生动。试问：概率论为什么不这样做呢？② 由于概率空间被绝对化，从而切断了概率空间内外事件之间的联系，破坏了随机现象的直观性和形象性。③ 无法解释集合论方法为什么能用于研究随机现象。④ 没有讨论条件与概率之间的内在联系，很难揭示条件概率与独立性的本质。事实上，K公理系统引进的条件概率是混乱的。⑤ 概率论的核心内容被模糊。作为类比，不管现代几何学的发展如何多种多样，直观的欧氏几何总以某种形式存在于其他类型的几何之中，并且是产生这些几何的源泉。试问：概率论中的"欧氏几何"存在吗？是什么？

本书提出的概率论自然公理系统克服了上述缺点，发展出一种新的研究模式。自然公理系统把随机世界抽象成因果空间。因果空间中的事件不仅保持随机现象的直观性和形象性，而且获得多种表达形式。概率论中集合论方法是从多种表达式的转换中产生出来的。

从哲学的角度看，事件存在多种表达式是因果律中"因"和"果"在数学中的反映；关于因果律思维方式的数学抽象就是概率论中集合论方法和测度论方法。

因果空间中存在各种各样的"图形"。自然公理系统主要研究两类"图形"——随机试验和概率空间，其中的核心分别是离散型试验和波莱尔(Borel)试验，离散型概率空间和柯尔莫哥洛夫概率空间。在因果空间中不仅可以研究单个的"图形"，而且可以研究不同"图形"之间的关系和进行数学演算。

应用随机试验这个概念，我们揭示出概率、条件概率和独立性的本质。并且发现用拉东—尼科迪姆定理定义的条件概率实际上不是条件概率，而是取了一个不恰当的名称。

现实中的随机现象经过数学处理后几乎都可以用随机变量表达，并且应用柯尔莫哥洛夫概率空间的统计知识可以得到随机变量蕴含的全部统计知识和统计规律。于是，因果空间、离散型试验、波莱尔试验、离散型概率空间、柯尔莫哥洛夫概率空间和随机变量组成了概率论的核心内容。这些核心内容使人们认识到随机世界不是一个杂乱无章、难于捉摸的世界，而是一个条理清楚、按统计规律在不断变化的世界。

由于作者的知识和能力所限，本书又是首次阐明新思想的著作，因此书中必定存在不完善、疏漏、不妥和叙述含糊的地方，恳请读者指正和补充。

书稿完成后得到清华大学出版社的大力支持，使本书能够很快地献给读者。北京科技大学秦明达教授仔细地审阅了全稿，提出许多中肯的、宝贵的意见，使书稿质量大有提高。在写作过程中得到教育科学出版社郑桂泉副编审的许多帮助，加速了书稿的完成。在此，向他们表达最诚挚的谢意。

作　者

1999.11.6

目　　录

绪 论

概率论的研究对象是现实世界中的随机现象（本书称之为随机世界）。柯尔莫哥洛夫（A. H. Kolmogorov）于 1933 年发表专著[1]，建立了概率论的柯尔莫哥洛夫公理系统（简称为 K 公理系统）。K 公理系统为现代概率论的蓬勃发展奠定牢固的基础，成为概率论发展史上的一个里程碑。

对 K 公理系统的批评主要来自贝叶斯（T. Bayes）学派。针对 K 公理系统的缺点和不足，菲纳特（B. de Finetti）提出 7 个讨论题[2]。一些作者建立与 K 公理系统稍有差别的公理系统[3~7]。综述性文献[8]介绍了建立主观概率公理系统的情况。

我们从另一方面加入 K 公理系统的批评者行列，指出如下 7 点不足或可商榷之处。

(1) 概率论公理系统应当完成一个宏伟的目标——建立随机世界的数学模型。

它应当把貌似杂乱无章的随机世界组织成一个既形象又严谨的逻辑体系。即是说，首先找出随机世界中几个最本质的属性，把它们抽象成公理系统，然后对公理系统进行逻辑推理和数学演算，最终推演出概率论的理论体系。目的是发现随机世界中更多更深刻的性质和规律。

但是，K 公理系统没有提出这项任务。它只是研究随机世界的某个局部，并把它抽象成概率空间。概率空间之间的联系是不清楚的。

(2) 概率论为什么没有自身独特的研究场所？

随机现象在现实中如此具体、如此广泛。可是，研究它们的概

率论却没有为自己建筑“住房”——独特的研究场所[①]。这是一种不正常的现象。导致一些数学家认为:概率论成为测度论的一部分。究其原因,来自于在K公理系统中基本事件空间是一个任意的集合,概率空间是在这个集合上形成的一类特殊的测度空间。

(3) K公理系统破坏了随机世界的直观形象:

随机世界	K公理系统
① 必然事件皆相等,即必然事件是惟一的。 ② 任意两个事件都可以进行并、交、差运算。	① 不对。不同的概率空间有不同的必然事件。 ② 不对。只有同一概率空间中的两个事件才能进行并、交、差运算。

由此不难看出,K公理系统人为地切断了概率空间内外随机现象之间的联系,因而缩小了概率论和数理统计的研究范围。

(4) K公理系统中的事件(例如“明天晴”)为什么是基本事件的集合?而由基本事件形成的集合为什么有的是事件?有的又不是?

这个问题相当于菲纳特的第1个讨论题。进而我们可以用更哲理的方式询问以下问题。

(5) 集合论方法为什么会成为研究随机现象的基本方法?

(6) K公理系统中基本事件可以是事件,也可以不是事件。如果不是,试问它是什么?

(7) K公理系统没有探讨条件与概率之间的内在联系。条件概率和独立性的直观背景被掩盖,因而很难把它们推广到一般场合。

这个问题相当于菲纳特的第6个讨论题。

本书完满地解答了上述疑问,把K公理系统发展成概率论的

---

① 自然公理系统提供的“住房”是因果空间。

自然公理系统。自然公理系统把随机世界描述为一个像欧氏空间那样既形象又能进行数学演绎推理的空间——因果空间。表 0.1 用类比的方式显示自然公理系统的结构和意义。

表 0.1　概率论与欧氏几何学[9]类比表

学　科	欧氏几何学	概率论
直观的研究对象	现实世界中物体的形状	现实世界中的随机现象（随机世界）
公理化方法 元素 元素之间的关系	希尔伯特公理系统 点、直线、平面 5 组公理	自然公理系统 随机事件 6 组公理
公理化后的产物 研究对象 研究任务	欧几里得空间 图形 图形的度量 欧氏几何学是研究图形和对图形进行度量的一门学科。它包括： ① 单个图形或单个度量化后图形的性质； ② 不同图形或不同度量化后图形之间的关系。	因果空间 随机试验 概率空间 概率论是研究因果空间中随机试验和对试验进行赋概的一门学科。它包括： ① 单个随机试验或单个概率空间的性质； ② 不同随机试验或不同概率空间之间的关系。

自然公理系统具有两大特点。第 1 个特点是遵守现代化公理系统的要求，是一个形式化的公理系统。

形式化的含义是，首先找到概率论的基本对象——随机事件，把它作为不能定义的元素，并用符号 $A,B,C\cdots$ 表示它们；其次，把元素之间的关系抽象成 6 组公理，并且公理也被符号化；最后，以 6 组公理为基础，对符号进行逻辑推理和数学演算，得到新的符号组合——各种各样的结论，从而形成概率论的理论体系。由于在推理和演算时不必顾虑这些符号的具体意义，因而保证了结论的可

靠性。

第 2 个特点是具备古典公理系统所要求的直观性和形象性。在选择公理时力求简单、直观，的确不需证明就被人们所接受。

概率论有非常具体的研究对象——随机世界。因此，直观形象是概率论产生的源泉，存在的基础。它不仅启迪研究者的思维，帮助思考，而且不断地为概率论和数理统计提供新的研究任务、新的研究方向，并指明发展方向，使概率论充满了生机和活力，生长出许多边缘学科，形成非常庞大的数学分支。

在内容安排上作两点说明：

(1) 我们在引进某组公理之前先扼要地介绍这组公理的直观背景①。

虽然我们生存在随机世界之中，对随机现象积累了许多感性的、直观的知识。但是，这些知识却是零散的、无组织的，多数人仍然认为它是一个杂乱无章、难于捉摸的世界。因此，指明每条公理的直观背景是我们的首要任务。不仅如此，我们用一章的篇幅介绍随机世界的直观形象。目的是把随机世界整理成一个较有条理的、富有规律性的世界，为自然公理系统的出现作思想和舆论的准备②。

(2) 在引进某组公理后立即用它们推导出一些基本的、常用的结论，并给予证明。其作用有 3 方面：① 使内容成为完整的、系统化的理论；② 具备集合论基础知识的读者能够理解自然公理系统；③ 为专家们指明常见的性质和结论在自然公理系统中所占据的位置。

本书是数学基础性专著。阅读本书只须具备基础性的数学知识，但需要有较强的逻辑思维能力。

---

① 建立和培养因果空间的直观形象是自然公理系统的一项不可缺少的任务，是本书的重要组成部分。

② 第 1 章是具有新观点的科普性内容。

第1章 概率论的现实背景

人们在认识随机世界——自然界和人类社会中各种各样的随机现象时积累了大量的简单观察，发现了一些行之有效、久经考验的实际方法。公理化方法就是从这些素材中提取出一些最本质的属性，进行严密精致的逻辑加工，形成概率论的公理系统。然后以公理系统为基础，建立概率论的理论体系，在新高度上认识随机世界。

本章从最重要的素材中提取出两个原理：

概率论原理Ⅰ：随机事件只有通过它和其他事件之间的联系才能被认识。

概率论原理Ⅱ：在合理划定(或给定)的条件下，随机事件出现的可能性大小能够用[0,1]中惟一的实数 p 表达。

原理Ⅰ指导我们建立自然公理系统的前3组公理。原理Ⅱ用于指导建立自然公理系统的后3组公理。

1.1 研究对象

日月星辰，周而复始；雨雪阴晴，气象万千；江河奔流，汇入大海。生物世界，适者生存，繁衍后代，免遭灭绝。战争与和平，政治与经济，科技与文化，文娱与体育，则是万物之灵——人类的创造。

自然界和人类社会永远处在不断变化和发展之中。反映变化和发展的标志是出现各种各样的现象，或者说，发生形形色色的事情。例如：

(1) 太阳从东方升起(用 A_1 表示这个现象，称为事件 A_1)。

(2) 月亮从太空中消失(事件 A_2)。

(3) 两物体之间存在吸引力(事件 A_3)。

(4) 在标准气压下,水加热到 50℃时沸腾(不可能事件 A_4)。

(5) 明天有雨(事件 A_5)。

(6) 明天阴(事件 A_6)。

(7) 明天晴(事件 A_7)。

(8) 明天的最高气温为 ξ℃(事件 A_8)。

(9) 明天的最低气温为 η℃(事件 A_9)。

(10) 某夫妇的胎儿是男孩(事件 A_{10})。

(11) 某夫妇的胎儿是女孩(事件 A_{11})。

(12) 用第 2 代杂交豌豆培育的某植株开白花(事件 A_{12})。

(13) 用第 2 代杂交豌豆培育的某植株开红花(事件 A_{13})。

事件 A_{12}和 A_{13}取自孟德尔(G. Mendel)做过的著名遗传学实验——豌豆杂交实验。

(14) 罐内放入 n 个球,分别标号为 $1,2,\cdots,n$。任意取出一个球,它是第 i 号球(事件 $B_i, i=1,2,\cdots,n$)。

(15) 100 件产品中有 5 件次品。任意取出 3 件,它们全是次品(事件 A_{15})。

(16) 某日进入某商店的人数恰为 5 千人(事件 A_{16})。

(17) 某日进入某商店的人数不超过 5 千人(事件 A_{17})。

(18) 向桌面抛一枚硬币时正面朝上(事件 A_{18})。

(19) 向桌面抛一枚硬币时反面朝上(事件 A_{19})。

(20) 向桌面抛一枚硬币时硬币直立(事件 A_{20})。

(21) 掷一颗骰子时出现 i 点(事件 $C_i, i=1,2,3,4,5,6$)。

(22) 掷一颗骰子时出现偶数点(事件 A_{22})。

(23) 掷两颗骰子时一颗出现 i 点,另一颗出现 j 点(事件 D_{ij}, $i,j=1,2,3,4,5,6$)。

(24) 掷两颗骰子时出现双 6(事件 A_{24})。

(25) 掷两颗骰子时出现的点数之和大于 8(事件 A_{25})。
诸如此类的事件无穷无尽,不可能一一列举。

所有的科学都研究事件。概率论也是研究事件的科学。它和其他科学的区别在于以下 4 个方面:

(1) 概率论只关心事件是否出现。

(2) 概率论不关心事件本身的其他内容。

(3) 概率论通过事件之间的联系来认识事件。

例如,对于事件 A_7,概率论不是通过 A_7 的具体内容(譬如说,出现"烈日当空"等现象)来认识"明天晴",而是通过 A_5 和 A_6 等事件不出现来认识"明天晴"。

(4) 如果事件能够出现,则猜测出现的可能性大小。

现象、事情、或事件是否出现受各种各样的条件(包括种种偶然的,不可预测的因素)支配。概率论把它们称为随机现象,或随机事件,用意是强调它们的出现随"机会"而定。事实上,即使像"太阳从东方升起"这样的事件,其出现也依赖于"机会"。因为当你站在其他星球上,或者站在地球的南北极处进行观察,事件 A_1 就未必出现。

现在能够概括出随机事件的一个特点,我们把它概括成如下原理。

概率论原理 I:随机事件只有通过它与其他事件之间的联系才能被认识。

为了用语上的方便,今后把随机事件简称为事件。

1.2 研究任务

《中国大百科全书·数学》中"概率论"条目明确指出:"概率论:研究随机现象数量规律的数学分支。"数量规律是一种高度概括的语言。依据 1.1 节的讨论,数量规律应当包括随机现象之间的

联系和随机现象出现的规律。

我们把概率论遵循的一个基本信念写成以下原理。

概率论原理Ⅱ:在合理划定(或给定)的条件下,随机事件出现的可能性大小能够用[0,1]中惟一的实数 p 表达。

在这个原理下,随机现象出现的规律就是我们日常生活中的语言:"在给定的条件下,某随机现象以 $\alpha\%$ 的可能性出现。"或者"在给定的条件下,我们以 $\alpha\%$ 的把握断定(或期待)某随机现象出现。"

用 $\mathscr{C}$ 代表合理划定(或给定)的条件;用 A 代表条件 $\mathscr{C}$ 下涉及的某个随机事件;数值 p 改用符号 $P(A|\mathscr{C})$ 表示,称之为条件 $\mathscr{C}$ 下事件 A 的概率。于是,概率论具有非常单纯的任务:求 $P(A|\mathscr{C})$。

为了完成这项任务,必须把条件 $\mathscr{C}$,事件 A,概率 $P(A|\mathscr{C})$ 抽象成能进行逻辑推理和数学演算的对象。这正是概率论自然公理系统所要做的工作。

在结束本节之前我们给出3点注记。

注1 随机性的本质是什么?遵循的基本信念是否合理?概率论拒绝回答这样的问题。它们是哲学探讨的问题。当然,把在基本信念上所建立的概率理论与现实随机世界进行比较,其吻合程度可以间接地检验基本信念是否合理,是否需要修正。

注2 统计规律是概率论中经常使用的语言。按通常的理解,大量随机现象中蕴含的集体规律性被认为是统计规律。或者更明确地叙述为:"就一次试验而言,看不出什么规律,但是'大数次'地重复这个试验,试验的结果又遵循某些规律,这种规律称之为统计规律,这一类试验称为随机试验,试验所代表的现象称之为随机现象。"[10]

注3 在《现代汉语词典》中规律被解释为"事物之间的内在的必然联系。"在投掷硬币一次的实验中我们有结论:"硬币出现正

面的可能性是 50%。"按词典的解释，这个结论不是规律；按注 2 的解释，这个结论也不是统计规律。但是，这个结论的确是随机现象出现的一个规律。为此，本书认为出现规律就是统计规律。

知识一词比规律包含更多的内涵。它通常解释为"人们在研究和改造世界的实践中获得的认识和经验的总和"。涉及随机现象的知识称为统计知识。在本书中把概率论规定为：**概率论是研究随机现象统计知识的数学分支**。即是说，我们把统计知识等同于"数量规律"，它至少包含两项内容：随机现象之间的联系和统计规律。

1.3 概　　率

为了理解概率论原理 Ⅱ，必须探讨概率的含义。这个概念在自然科学和社会科学中已被广泛使用。甚至模式为"明天降雨概率是 60%"的天气预报也已为人们普遍接受，似乎概率这一概念不会再有什么异议。"其实不然，仔细考察一下，就会发现各人的理解和作出的解释往往是不同的，这不能不影响到对于概率论、数理统计的看法和应用的范围。这一现象并不是现在才有的，从概率论的发展历史来看，几乎从它诞生时起，就相伴着争论，只是在不同的时期，争论的出发点，争论的重点有所不同，并且影响到学科的内容和发展趋势。"[11]时至今日，还没有一个为所有概率学者都接受的解释。

作者的观点是：每种解释都从某个侧面发掘出概率的本质，重要的是了解在什么样的条件下某种解释才是合理的、可接受的。事实上，多种解释并存不仅使我们正确地、全面地理解"概率是什么"，而且产生许多方面的应用，甚至得到意料之外的、令人惊喜的发现。

下面介绍 4 种流行的解释。

1.3.1 概率的古典定义

拉普拉斯(P. S. Laplace)首次明确地给出概率的定义(1814年)。他把概率定义为“有利情况”的数目除以“全部可能情况”的总数目。这就是概率的古典定义。在一些场合,古典定义是清楚的、明确的。

例1 向光滑的桌面随意地投掷一枚硬币(实验 $\mathscr{E}_1$)。实验的结果不是出现正面,就是出现反面。因此全部可能的情况为

$$\langle 正面\rangle,\langle 反面\rangle \tag{1.3.1}$$

用 $\mathscr{C}_1$ 表示确定实验的条件,用 $P(\langle 正面\rangle|\mathscr{C}_1)$表示〈正面〉出现的概率。由于正面出现只含一种有利情况,按照古典概率的定义得出

$$P(\langle 正面\rangle|\mathscr{C}_1)=\frac{1}{2} \tag{1.3.2}$$

同理

$$P(\langle 反面\rangle|\mathscr{C}_1)=\frac{1}{2} \tag{1.3.3}$$

例2 向光滑的桌面投掷两枚硬币(实验 $\mathscr{E}_2$)。如果用(α,β)表示第1枚硬币出现 α 面,第2枚出现 β 面。那么全部可能的情况为

$$\langle 正,正\rangle,\langle 正,反\rangle,\langle 反,正\rangle,\langle 反,反\rangle \tag{1.3.4}$$

至少有一个正面出现包含三个有利情况:〈正,正〉,〈正,反〉,〈反,正〉。所以

$$P(至少有一个正面出现\ |\mathscr{C}_2)=\frac{3}{4} \tag{1.3.5}$$

这里 $\mathscr{C}_2$ 表示确定实验 $\mathscr{E}_2$ 的条件。同理

$$P(出现两个反面\ |\mathscr{C}_2)=\frac{1}{4} \tag{1.3.6}$$

于是得到遗传学中的一个重要比例 3∶1[①]。即

---

① 参看本节末的讨论。

$$\frac{P(\text{至少有一个正面出现}\mid\mathscr{C}_2)}{P(\text{出现两个反面}\mid\mathscr{C}_2)}=\frac{3}{1} \tag{1.3.7}$$

例 3 向光滑的桌面投掷一颗均匀的骰子(实验 $\mathscr{E}_3$)。这时全部可能的情况为

$$\langle 1\rangle,\langle 2\rangle,\langle 3\rangle,\langle 4\rangle,\langle 5\rangle,\langle 6\rangle \tag{1.3.8}$$

这里$\langle i\rangle$表示出现 i 点。用 $\mathscr{C}_3$ 表示确定实验 $\mathscr{E}_3$ 的条件,应用概率的古典定义得

$$P(\langle 1\rangle\mid\mathscr{C}_3)=P(\langle 2\rangle\mid\mathscr{C}_3)=\cdots=P(\langle 6\rangle\mid\mathscr{C}_3)=\frac{1}{6} \tag{1.3.9}$$

注意,概率的古典定义只在一些特殊场合下才是合理的,可接受的。例如,在例 3 中条件 $\mathscr{C}_3$ 含有骰子均匀、桌面光滑、投掷随意等诸条件。条件 $\mathscr{C}_3$ 的任何变化都将使 $\mathscr{E}_3$ 不遵守概率的古典定义,因而得不到式(1.3.9)。

1.3.2 几何概率

在概率论刚开始发展的时候人们就已经注意到可能情况的总数目和有利情况的数目都可以是无限的。并且出现一些个别的例子,促使对概率进行另一种解释。

这些个别的例子可以一般化为:设 G 是平面(或直线,或空间)中的一个区域,g 是 G 的一个子区域。现在把一个质点随意地投入 G 中(实验 $\mathscr{E}$)。试问质点落在 g 中的概率多大?

问题的回答产生几何概率。用 $\mathscr{C}$ 表示确定实验 $\mathscr{E}$ 的条件,答案是

$$P(\text{质点落入 } g \text{ 中}\mid\mathscr{C})=\frac{\mu(g)}{\mu(G)} \tag{1.3.10}$$

这里视 G,g 为直线、平面、空间中的区域,把 μ 依次地理解为长度、面积和体积。

布丰(de Buffon)对几何概率进行了详细的讨论(1777 年),计