新概率论

自然公理系统中的概率论

熊大国　著

科学出版社

新概率论

自然公理系统中的概率论

熊大国 著

科学出版社

北京

内 容 简 介

本书是在自然公理系统中建立概率论的第一部著作. 本书前五章建立因果空间、随机试验、概率空间、条件概率捆和独立性的理论, 重点介绍离散型、Kolmogorov 型、独立乘积型概率空间, 形成概率论的基础理论.

第 6、8 章论证随机变量、随机向量和宽随机过程是科学实验中子随机局部的数学模型; 应用概率论基础理论介绍因果结构图、各种条件分布函数和独立性, 建立数学期望、方差和协方差等数字特征的知识, 形成随机变量和随机向量的基本理论, 以及随机过程的初步知识; 最后两章介绍两类最重要的统计规律——大数定理和中心极限定理.

本书适用于科技工作者、高等学校教师、研究生和高年级本科生阅读.

图书在版编目(CIP)数据

新概率论: 自然公理系统中的概率论/熊大国著. —北京: 科学出版社, 2016.10

ISBN 978-7-03-050223-0

Ⅰ. ①新… Ⅱ. ①熊… Ⅲ. ① 概率论 Ⅳ. ①O211

中国版本图书馆 CIP 数据核字(2016) 第 250048 号

责任编辑: 陈玉琢 / 责任校对: 张凤琴
责任印制: 张 伟 / 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

北京中石油彩色印刷有限责任公司印刷
科学出版社发行 各地新华书店经销
*
2016 年 10 月第 一 版 开本: 720 × 1000 1/16
2016 年 10 月第一次印刷 印张: 32 3/4
字数: 648 000

定价: 178.00 元

(如有印装质量问题, 我社负责调换)

序

概率论是研究随机现象中因果关系和出现规律 (统计规律是一种量化的出现规律) 的一门数学分支. 概率论有自己独特的思维模式, 要求人们用新思想新方式去看待和思考世界.

现代概率论建立在 Kolmogorov 公理系统的基础上. 在概率论的学习中, 许多初学者感到学习被动, 思维不知如何展开, 概念难理解, 方法难掌握, 习题不会做, 甚至做完题后也不敢说做对了. 这类现象是如此普遍, 使得许多概率论学者认为只有先学习抽象的测度论才能掌握非常现实的概率论, 并感叹“它的基础是一个迷人而又难以捉摸的问题” (H. Wang, *FromMathematicstoPhilosophy*, 转引自《数学与文化》, 北京大学出版社, 1990, 第 99 页).

造成这种局面的原因是什么? 答案可以从 Kolmogorov 的话中找到. 他说: “由于有了公理化的概率论, 就使得我们摆脱了试图寻找一种既具有自然科学的直观确凿性又便于建立形式严整的数学理论的方法来定义概率的诱惑. 这样的定义如我们在几何中把点这样来定义, 即把点作为一个实在物体经过四面八方无数次切削 (而且每次切削均使直径缩小 (例如) 一半) 最后所剩下的东西” (H.Kolmogorov, 概率论, 见《数学 —— 它的内容, 方法和意义 (第 2 卷)》, 科学出版社, 2001, 第 287 页 (数学名著译丛)).

作者建立的概率论自然公理系统实现了 Kolmogorov 想摆脱的诱惑. 自然公理系统把概率论建成“既具有自然科学的直观确凿性又便于建立形式严整的数学理论”, 从而揭示出概率论的“迷人而又难以捉摸的”基础. 美国《数学评论》评论为“Xiong’s approach is novel”; 德国《数学文摘》评论为“The NAS is a certain extension of the KAS” (注: NAS 指自然公理系统; KAS 指 Kolmogorov 公理系统); 国内《外文数字图书馆》评论为“本书 (指参考文献 [2]) 是中国学者熊大国对国际公认的 Kolmogorov 在 1933 年建立的公理系统的一个挑战”.

本书是在自然公理系统基础上建立的概率论, 目标是把概率论的基础部分改造成像初等几何那样既直观形象, 理论又严谨的数学分支. 事实上, 这两门学科之间存在如下表所示的类比.

与欧氏空间和几何图形不同, 随机宇宙和随机局部的数学模型——**因果空间**、**随机试验和概率空间**没有直观的确凿的实体, 而是抽象的数学概念. 因此, 必须用概率论倡导的新思想新方式去认识、理解这些数学概念, 捕捉这些概念代表的“实

体". 我们期待, 随着概率论因果思维模式的普及, 这三个研究对象会像欧氏空间和几何图形 (三角形, 圆, 长方体等) 一样成为人们的常识.

学科	初等几何	概率论
现实世界中的研究对象	物体的形状	随机现象中的因果关系
理论研究的场所	三维欧氏空间	因果空间
理论研究的对象	图形	随机局部 (即事件 σ 域)
研究对象的数学模型	三角形, 圆, 长方体等 (图形的数学模型) 图形的度量	随机试验 (随机局部的因果模型) 概率空间 (随机试验上赋概的因果量化模型)
研究任务	初等几何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的图形的性质 2) 不同图形或不同度量化后的图形之间的关系	概率论是研究随机试验和对随机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

清晰的直觉和逻辑推理是掌握概率论必须学会的两种能力. 为了提高直观能力, 本书用较多的篇幅 (包括绪论、每章的开篇语、标有直观背景或建模的小节等) 讨论现实中的随机现象, 从它们之间的联系中提取研究课题, 探讨解决方法, 获得基本概念和方法的雏形. 在接着的小节中, 我们以直观雏形为依托, 抽象出理论中的研究对象, 建立基本概念和计算方法, 自然地形成概率论的理论系统和建模方法. 其中涉及存在性的一些问题则交给实变函数论和测度论完成. 因此, 本书可以认为是一种 "测度论式的" 概率论, 它把 "先学抽象的测度论再学现实的概率论" 改变为 "先学现实的概率论再学抽象的测度论". 我们期望, 这种处理方法能够使本书成为初等概率论.

本书是新概率论的入门著作, 我们希望较全面地包含基础概率论的内容, 指明主要概念的来龙去脉, 力求做到定理的证明清楚, 例题的解答有理有据. 本书适用于自学, 适当取舍内容后可作为教材使用. 由于作者水平有限, 许多内容又是首次叙述, 书中定有许多不妥之处, 恳请读者批评指正. 我们希望本书能起抛砖引玉的作用, 促使出现更多介绍新概率论的著作和教材.

最后, 感谢北京理工大学和科学出版社对本书出版的支持.

作　者

2015 年 12 月 1 日

目　录

绪论　随机宇宙初探

0.1　随 机 现 象

自然界和人类社会中不断地出现各种各样的现象. 正是这些现象不断地出现和消失, 人们才感觉到自己活着, 体验到时间和空间的存在, 发现自然界在变化, 人类社会在前进.

现象按 “出现” 的特性可以分为三类. 在一定的 (自然形成的或人们规定的) 条件下：① 某现象一定出现；② 某现象一定不出现；③ 某现象可能出现也可能不出现. 前两类称为确定性现象, 第③类称为不确定性现象或随机现象. 例如:

1) 太阳从东方升起;

2) 两物体间存在吸引力 —— 牛顿万有引力定律;

3) 在标准气压下, 温度为 50 ℃的水会沸腾 (一定不出现的现象);

4) 向上抛一颗石子, 石子不落回地面 (一定不出现的现象);

5) 在恒力作用下的质点做等加速度运动 (牛顿第二定律);

等等, 被认为是确定性现象. 随机现象则按下述方式描述：

1) 向桌面扔一枚硬币, 现象 “正面” 可能出现, 也可能不出现；现象 “反面” 可能出现, 也可能不出现.

2) 向桌面掷一颗骰子, 现象 “a 点”(a=1,2,3,4,5,6) 可能出现, 也可能不出现；现象 “偶数”(或 “奇数”) 可能出现, 也可能不出现.

3) 向桌面掷一颗灌铅的骰子, 现象 “a 点”(a=1,2,3,4,5,6) 可能出现, 也可能不出现.

4) 100 件产品中有 5 件次品, 任意地取出 3 件. “3 件皆是次品” 这一现象可能出现, 也可能不出现.

5) 在一次打靶射击中, 现象 “命中靶心” 可能出现, 也可能不出现；现象 “脱靶” 可能出现, 也可能不出现.

6) 服务台 (售票窗口, 电话交换台, 机器维修站, 医疗诊所等) 为顾客服务, 顾客到达后排队等待. “顾客排队的长度超过 a 人 (或件)” 这一现象可能出现, 也可能不出现；“顾客等待时间超过 b 分钟” 这一现象可能出现, 也可能不出现.

7) “某夫妇的胎儿是女性” 这一现象可能出现, 也可能不出现.

8) “某生物 (人, 虎, 树等) 或某产品 (灯泡, 计算机, 拦河大坝等) 的寿命超过 a 年” 这一现象可能出现, 也可能不出现.

9) “明天的最高温度是 ξ °C” 这一现象可能出现, 也可能不出现.

10) “明天晴” 可能出现, 也可能不出现.

等等. 诸如此类的随机现象到处存在, 不胜枚举.

其实, 确定性现象和随机现象之间没有绝对的分界线. 当 “一定的条件” 改变时它们可以相互转化. 例如, 当观察者是地球上的普通居民时, “太阳从东方升起” 是确定性现象. 但是, 当观察者能够到达其他星球或地球的南北极时, “太阳从东方升起” 成为随机性现象. 又如, 向上抛石子的力非常大, “石子不落回地面” 这一现象也可能发生, 也可能不发生, 于是它成为随机现象. 再如, 如果 100 件产品全是次品, 那么 “3 件皆是次品” 成为确定性现象. 至于牛顿的万有引力定律和第二定律, 我们也无法肯定在人类观察范围之外是否正确. 因此, 在广义下所有的现象都是随机现象, 绝对的确定性现象是不存在的.

在概率论中随机现象被称为随机事件, 简称为事件. 事件是概率论中唯一不能定义的基本概念, 它只能进行描述和说明. 概率论认为:

“一个事件被描述清楚, 或 ‘被定义’ 的标志是: 人们能够判定该事件或者处于出现状态, 或者处于不出现状态, 并且二者必居其一. ”

0.2　概率论原理 I

大多数学科只研究某些特定的现象 (例如, 物理现象, 化学现象, 生物现象, 自然现象, 社会现象, 经济现象, 等等). 与它们不同, 概率论研究所有的现象, 但只关心现象一个方面的知识: 它能否出现? 出现的可能性有多大? 这个特点使得概率论成为应用最广泛的学科之一.

详细地说, 概率论和其他学科的区别有如下的四个方面:

1) 概率论只关心事件是否出现;

2) 如果事件能出现, 则询问 (或猜测) 它出现的可能性有多大;

3) 概率论不关心事件本身的其他内容;

4) 概率论通过事件之间的联系来认识事件.

例如, “明天晴” 是一个事件. 我们既可以通过 “烈日当空” “万里无云” 等事件出现来认识 “明天晴”; 又可以通过 “明天阴” “明天下雨” “明天下雪” 等事件不出现来认识 “明天晴”. 我们把这种认识事件的方法总结为:

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, 被理解.

事件之间存在什么样的联系? 怎样把联系抽象为数学运算? 这正是第 1 章应该完成的任务.

0.3 概率论原理 II

人类生存在充满随机事件的自然界和人类社会 (不妨简称为随机宇宙) 之中. 人们在认识随机宇宙时积累了大量的简单观察, 总结出随机事件出现的许多规律. 迈出实质性的一步是: **"用数值把事件出现可能性的大小进行量化"**①. 例如, 我们经常使用如下的判断:

1) 明天降雨的可能性是 80%.

2) 向桌面扔一枚硬币, 正面出现的可能性是 50%.

3) 向桌面扔两枚硬币, "两个正面" 出现的可能性是 25%; "一个正面和一个反面" 出现的可能性是 50%.

4) 向桌面掷一颗骰子, 6 点出现的可能性是 1/6.

5) 向桌面掷两颗骰子, 事件 "双 6" 出现的可能性是 1/36.

6) 某夫妇的胎儿为女性的可能性是 50%.

7) 100 件产品中有 5 件次品, 任意地取出 3 件, "3 件皆为次品" 的可能性很小 (你能说出有多小吗? 概率论能, 它约为 0.0062%. 参看例 3.3.1).

8) 某个灯泡的使用寿命超过 2000 小时的可能性是 80%.

9) 在标准气压下, 温度为 50 ℃的水会沸腾的可能性为 0.

10) 太阳从东方升起的可能性是 100%,

等等. "**用数值量化出现可能性大小**" 是人类在实践中获得的深刻知识. 我们把它总结为概率论原理 II.

概率论原理 II 在合理划定的 (或给定的) 条件下, 某随机事件出现的可能性大小能用 [0,1] 中唯一的实数 p 表达.

我们经常引用原理 II. 为了引用方便, 进行如下处理: 用 $\mathscr{C}$ 表示原理 II 中合理划定的条件; 用 A 表示条件 $\mathscr{C}$ 下可能出现的一个事件; 实数 p 改记为 $P(A|\mathscr{C})$, 称为在条件 $\mathscr{C}$ 下事件 A 的概率. 于是, 概率论有非常单纯的任务:

1) 把 $\mathscr{C}, A, P(A|\mathscr{C})$ 抽象成能进行数学演算的对象;

2) 找出人们关心的条件 $\mathscr{C}$ 和事件 A, 并计算 $P(A|\mathscr{C})$.

至于原理 II 中的 $P(A|\mathscr{C})$ 为什么存在? 为什么唯一? 这是理论无法回答, 只能用实践检验的问题. 幸运的是, 建立在原理 I 和原理 II 上的概率论不仅没有出现矛盾, 而且得出与现实非常吻合的结论, 并发现随机宇宙中蕴含的许多规律, 成为内容庞大、应用广泛的一门数学学科.

① 这一步是一个漫长的历史过程.

0.4　建立理论系统和数学模型

概率论是一门理论学科, 又属于应用数学的范畴. 这个特点使得本学科包含两类工作.

一类工作是建立概率论的理论系统: 构造有价值的研究对象 (如**因果空间**, **随机试验**, **概率空间**, **随机变量**, **随机向量**, **随机过程**, **随机场**等), 发现研究对象中的知识和规律, 以及对象之间的联系, 并把它们组织为系统的、严谨的、便于传授的知识体系.

另一类工作是建立应用问题的数学模型 (称为概率论建模, 简称为建模). 即是说, 把应用问题关心的随机现象抽象成理论中的研究对象 (如随机试验, 概率空间, 随机变量等), 并称这个对象为该问题的数学模型. 然后应用模型中的知识和规律来解释或发现这些随机现象中蕴含的因果知识和出现规律.

一方面, 数学模型存在于理论系统之中, 因此建立理论系统的工作应先于建模工作; 另一方面, 数学模型是连接理论和现实世界的桥梁, 人们不能脱离实际来凭空捏造出概率论的研究对象. 作为一本入门书, 可采用的方法只能是边建模边建立理论系统, 把两项工作交错地进行.

应该指出, 两类工作有完全不同的思维模式. 建立理论系统是演绎性的工作. 即是说, 从公理 (在本书中表现为用定义规定的几个基础概念①) 出发, 应用逻辑推理和数学演算得到各种各样的结论 —— 命题、定理和计算方法等, 形成严谨的知识体系. 在理论体系中所有的问题都是明确的, 答案是准确的和唯一的.

建模工作恰巧相反, 它属于归纳性的工作. 即是说, 建模工作中提出的问题多少有点含糊, 答案也不必唯一. 原因是数学模型和现实对象之间是一种反映和被反映的关系. 为了得到适用的数学模型, 人们必须强调现实对象的某些方面, 扬弃其他方面. 由于强调和扬弃的内容不同, 同样的现实对象可以产生出多种不同的数学模型. 至于哪一个模型更适用、更 "正确", 只能用实践来检验.

由于建立理论系统和建立数学模型有不同的工作方式, 两项工作的内容又是交错地进行研究和讨论, 因此区别正在学习的内容属于哪一类工作对初学者既必要又有益.

0.5　概率论用新思想新方式认识宇宙

0.5.1　随机宇宙 $\mathcal{U}$ 和它的数学模型 —— 因果空间 $(X,\mathcal{U})$

随机宇宙是浩瀚宇宙中古往今来的全体随机现象 ——"事件" 组成的集合. 用

① 重视定义是理解概率论的好方法.

$\mathcal{U}$ 表示这个集合, 即

$$\mathcal{U}=\{A|A\text{ 为事件}\} \tag{0.5.1}$$

概率论原理 I 指出, $\mathcal{U}$ 中事件不是孤立存在的, 它们之间有各种各样的联系, 故称 $\mathcal{U}$ 为事件空间 (见定义 1.2.1). 简言之, 随机宇宙是事件空间 $\mathcal{U}$. 也是集合 $\mathcal{U}$.

事件空间 $\mathcal{U}$ 中有一种联系, 称为包含关系. 它有如下定义.

定义 0.5.1 设事件 $U,V\in\mathcal{U}$. 如果事件 U 出现时 V 必出现, 则称 U 是 V 的子事件, 记为

$$V\supset U \tag{0.5.2}$$

读作 V 包含 U, 或 U 在 V 中. 特别, $V\neq U$ 时称为真包含关系.

哲学中充足理由律说:“任何现象若无出现的原因, 就不会发生. ” 真包含关系式 (0.5.2) 恰好表达出充足理由律:“由于原因事件 U 的出现导致事件 V 发生. ” 即是说, U 是事件 V 发生的原因之一, V 是事件 U 产生的后果之一.

现在, 我们反复地应用充足理由律, 得到一个真包含关系串

$$V\supset U_1\supset U_2\supset\cdots\supset U_\alpha\supset\cdots \tag{0.5.3}$$

其中所有的 V, U_α 皆是可以发生的随机事件, 所有的 $\supset$ 皆是真包含号. 显然, 串 (0.5.3) 没有尾元素. 事实上, 若 (0.5.3) 存在尾元素 U, 那么由充足理由律得出存在事件 U^* 使得 $U\supset U^*$, 与 U 是尾元素矛盾.

试问: 真包含关系串 (0.5.3) 的尽头是什么? 概率论抽象出:“尽头是一个原因点 u”. 由于尽头过于抽象, 不妨认为一个原因点 u 就是一个真包含关系串 (0.5.3). 原因点具有如下性质:

i) 原因点 u 不是事件. 称原因点 u 出现, 如果 (0.5.3) 有一个尾巴出现.

ii) 原因点 u 出现将导致事件 A(存在某个 α 使得 $A\supset U_\alpha$) 出现. 因此称 u 是事件 A 的一个诱因点.

iii) 用 X 表示所有原因点组成的集合, 即

$$X=\{u|u\text{ 是一个原因点}\} \tag{0.5.4}$$

并且 X 中有且仅有一个原因点出现. 称 X 为原因空间.

iv) 对 $\mathcal{U}$ 中任意的事件 A, 引进

$$A=\{u|u\in X,u\text{ 是 }A\text{ 的一个诱因点}\} \tag{0.5.5}$$

于是事件 A 得到另一种表达形式: A 是 X 的子集 —— 一组原因点. 显然, 事件 A 出现当且仅当它的一个诱因点出现.

【建模】 称二元组 $(X,\mathcal{U})$ 为因果空间, 它是随机宇宙 $\mathcal{U}$ 的数学模型. 其中 X 是原因空间, $\mathcal{U}$ 是事件空间. 联系 X 和 $\mathcal{U}$ 的纽带是因果关系式 (0.5.5).

0.5.2 随机局部 $\mathcal{F}$ 和它的因果模型 —— 随机试验 $(\Omega,\mathcal{F})$

常识告诉我们, 人们无法研究因果空间 $(X,\mathcal{U})$ 中的全部事件, 只能研究事件空间 $\mathcal{U}$ 中的一部分事件. 用 $\mathcal{F}$ 表示被关心事件组成的集合, 即

$$\mathcal{F}=\{A|A\in\mathcal{U},A\text{是被关心的事件}\} \tag{0.5.6}$$

今后, 称 $\mathcal{F}$ **为随机局部**. 显然, $\mathcal{F}$ 中事件经过数学演算后产生的事件也是被关心的, 因而是 $\mathcal{F}$ 中元素. 于是, $\mathcal{F}$ 是从 $\mathcal{U}$ 中划出的一个 "相对封闭"(即满足定义 1.2.1 中条件) 的子集. 在集合论中称 $\mathcal{F}$ 为全集 X 中的 σ 域, 在概率论中称 $\mathcal{F}$ 为事件 σ 域, 简称为 σ 域.

随机局部 $\mathcal{F}$ 的相对封闭性保证 $\mathcal{F}$ 像一个 "小型的" 事件空间. 于是, 用 $\mathcal{F}$ 代替 $\mathcal{U}$, 逐句地重复 0.5.1 小节的讨论得出随机局部 $\mathcal{F}$ 的因果模型 —— 随机试验 $(\Omega,\mathcal{F})$. 这里

$$\Omega=\{\omega|\omega\text{是样本原因点}\} \tag{0.5.7}$$

$$\omega\text{是真包含关系串}(0.5.3)\ (\text{其中}\ V,U_\alpha\in\mathcal{F})\ \text{的尽头} \tag{0.5.8}$$

对任意的 $B\in\mathcal{F}$, 如果存在 α 使得 $B\supset U_\alpha$, 则称 ω 是 B 的一个样本诱因点. 令

$$B=\{\omega|\omega\in\Omega,\omega\text{是}\,B\,\text{的样本诱因点}\} \tag{0.5.9}$$

于是, 事件 B 又得到新的表达式 ——Ω 的子集. 直观上, 事件 B 出现当且仅当它的某个样本诱因点出现.

【建模】 称二元组 $(\Omega,\mathcal{F})$ 为随机试验, 它是随机局部 $\mathcal{F}$ 的因果模型, 其中 Ω 是样本原因空间, $\mathcal{F}$ 是事件 σ 域. 联系 Ω 和 $\mathcal{F}$ 的纽带是因果关系式 (0.5.9).

现在发生重要的事情, 由于 (0.5.8) 限定 $V,U_\alpha\in\mathcal{F}$, 使得真包含关系串 (0.5.3) 往往存在尾元素 ω, 并且 ω 是事件 (故 ω 又称为基本事件). 于是, 随机试验把 (0.5.3) 的 "尽头" 具体化明确化. 事实上, 应用中随机试验 $(\Omega,\ \mathcal{F})$ 的样本空间 Ω 往往是能一一列举出元素的集合, 是人们熟悉的集合 (如有限集, 可列集, 实数集, 欧氏空间, 等等).

直观上, 因果空间是一个大容器, 它装载所有的随机试验, 是最大的随机试验; 随机试验则是一个 "小型的" 因果空间.

0.5.3 随机局部 $\mathcal{F}$ 的因果量化模型 —— 概率空间 $(\Omega, \mathcal{F}, P)$

如何得到随机局部 $\mathcal{F}$? 概率论原理 II 提供了方法: 用合理划定的 (或给定的) 条件 $\mathscr{C}$(见 0.3 节). 令

$$\mathcal{F} = \{A|A \text{是条件} \mathscr{C} \text{下能出现的事件}\} \tag{0.5.10}$$

于是 $\mathcal{F}$ 是一个随机局部, 因而存在因果模型 —— 随机试验 $(\Omega, \mathcal{F})$. 不仅如此, 概率论原理 II 还保证, 对任意的 $A \in \mathcal{F}$, 在条件 $\mathscr{C}$ 下事件 A 出现的可能性大小是数值 $P(A|\mathscr{C})$. 故

$$P(A|\mathscr{C}), \quad A \in \mathcal{F}$$

是定义在 $\mathcal{F}$ 上的函数, 称为 $\mathcal{F}$ 上的概率测度, 简称为概率, 简记为 P.

【建模和赋概】 称三元组 $(\Omega, \mathcal{F}, P)$ 为概率空间, 它是随机局部 $\mathcal{F}$ 的因果量化模型.

直观上, Ω 是导致随机局部 $\mathcal{F}$ 中事件出现的全部样本原因点 (或基本事件); $\mathcal{F}$ 是 Ω 中样本原因点出现导致的全部后果; $P(A|\mathscr{C})$ 是在条件 $\mathscr{C}$ 下 $\mathcal{F}$ 中事件 A 出现的可能性大小.

0.5.4 实验 $\mathcal{E}$ 产生的随机局部

如何获得概率论原理 II 中合理划定的 (或给定的) 条件 $\mathscr{C}$? 现实世界中的科学实验是主要的工具.

实验是一种最常见的科学活动. 例如, 预报明天的天气, 检查产品的质量, 投掷一颗骰子, 扔一枚硬币, 等等. 在概率论中, 实验遵守的条件 $\mathscr{C}$ 不仅包含实验装置, 而且包括随机环境中的各种干扰和操作者. 经验表明, 实验 $\mathcal{E}$ 遵守的条件 $\mathscr{C}$ 是概率论原理 II 中合理划定的 (或给定的) 条件. 称 $\mathscr{C}$ 产生的因果量化模型 $(\Omega, \mathcal{F}, P)$ 为实验$\mathcal{E}$的相应的数学模型——概率空间. 于是, 条件$\mathscr{C}$成为理论联系实际的桥梁,

$$\text{实验}\,\mathcal{E}\,(\text{即随机局部}\,\mathcal{F}) \Leftrightarrow \text{条件}\,\mathscr{C} \Leftrightarrow \text{概率空间}(\Omega, \mathcal{F}, P) \tag{0.5.11}$$

由此总结出, 概率论研究实验 $\mathcal{E}$ 的工作包括三方面的内容:

i) 求出实验 $\mathcal{E}$ 的因果模型$(\Omega, \mathcal{F})$和因果量化模型$(\Omega, \mathcal{F}, P)$(建模和赋概部分);

ii) 已知随机试验 $(\Omega, \mathcal{F})$ 和概率空间 $(\Omega, \mathcal{F}, P)$, 寻找和建立其内的因果知识和出现规律 (理论研究部分);

iii) 应用数学模型 $(\Omega, \mathcal{F})$ 和 $(\Omega, \mathcal{F}, P)$, 发现或解释实验 $\mathcal{E}$ 中的因果知识和出现规律 (应用部分).

例 0.5.1 向桌面随意地投掷一颗质料均匀的骰子 (实验 $\mathcal{E}$), 试问将产生什么样的结果.

【确定性方式】 难以回答.

【概率论方式】 解答: 用 $\omega_i(i=1,2,\cdots,6)$ 表示事件 "骰子出现 i 点". 显然, (0.5.8)(其中 V, U_α 是本例实验 $\mathcal{E}$ 产生的事件) 的 "尽头" 是且只能是这六个事件中的一个, 故 $\omega_i(i=1,2,\cdots,6)$ 是基本事件 (即样本原因点). 令

$$\Omega=\{\omega_1,\omega_2,\cdots,\omega_6\}$$

$$\mathcal{F}=\Omega\text{ 的全体子集组成的集合}$$

于是, 实验 $\mathcal{E}$ 划定出随机局部 $\mathcal{F}$, 并且 $\mathcal{F}$ 的因果模型是 $(\Omega,\mathcal{F})$. 注意到骰子均匀和投掷随意, 于是对任意的 $A\in\mathcal{F}$, 赋予事件 A 的概率为 (应用等可能性原理, 见例 3.1.2)

$$P(A)=A\text{ 含基本事件个数}/6$$

即事件 A 出现的可能性大小为 $P(A)\times 100\%$.

最终, 概率论方式给出的答案是: "实验 $\mathcal{E}$ 产生的结果是概率空间 $(\Omega,\mathcal{F},P)$". 答案的直观含义是, 实验 $\mathcal{E}$ 产生的全部结果组成事件 σ 域 $\mathcal{F}$; Ω 是导致 $\mathcal{F}$ 中事件出现的全部样本诱因点; $P(A)(A\in\mathcal{F})$ 是事件 A 出现的可能性大小. 此答案准确地表达出实验结果的全部内容.

0.5.5 实验 $\mathcal{E}$ 中数值指标产生的子随机局部

例 0.5.2 在天气预报 (实验 $\mathcal{E}$) 中, 试预报明天的最高温度?

【确定性方式】 解答: ①粗糙方式: 明天的最高气温是 a ℃; ②精确方式: 明天的最高气温是 a ℃, 误差为 δ ℃. 即是说, 明天的最高气温位于区间 $[a-\delta, a+\delta]$ 之中.

【概率论方式】 解答: ①粗糙方式不确定性的回答: 明天的最高气温是 ξ ℃, ξ 是随着明天出现不同气象条件取不同数值的变量. 简言之, ξ 是随机变量.

②严格方式 (用函数给出的不确定性回答): 认为实验 $\mathcal{E}$ 存在因果量化模型 $(\Omega,\mathcal{F},P)$, 但无法表达出 $\Omega,\mathcal{F},P$. 于是发出预报: "明天的最高气温是随机变量 ξ". 它是定义在 Ω 上的函数

$$\xi(\omega),\quad \omega\in\Omega$$

即是说, 最高气温 ξ 随不同气象条件 ω 的出现取不同的数值.

如何理解严格方式的答案? 第 6 章将论证, 概率空间 $(\Omega,\mathcal{F},P)$ 上的这个数值指标 $\xi(\omega)$ 在随机局部 $\mathcal{F}$ 中划定出反映 "明天冷热情况" 的一个子随机局部 —— 事件 σ 域 $\sigma(\xi)$. 这个子局部 $\sigma(\xi)$ 产生出 "已知的" 因果量化模型 $(\Omega, \sigma(\xi), P)$. 这个 "已知" 的模型蕴含 $\xi(\omega)$ 的全部因果知识和出现规律. 这里, "已知" 的含义

是, 虽然 Ω 仍然未知, 但 $\sigma(\xi)$ 中的事件能一一列举, 用分布函数可以赋予概率测度 $P(A|\mathscr{C}), A \in \sigma(\xi)$. 因此, $\xi(\omega)$ 可以视为子随机局部 $\sigma(\xi)$ 的数学模型.

特别, 在子随机局部 $\sigma(\xi)$ 中事件 B"明天的最高气温位于区间 $[a-\delta, a+\delta]$ 之中" 有符号表达式

$$B = \{\omega \mid a-\delta \leqslant \xi(\omega) \leqslant a+\delta\} \xlongequal{\text{def}} \{a-\delta \leqslant \xi(\omega) \leqslant a+\delta\}$$

并且 $P(B|\mathscr{C}) \approx 1$, 即 B 是大概率事件. 换言之, 我们以大概率 $P(B|\mathscr{C}) \times 100\%$ 的把握断定事件 B 出现, 这个结论与【确定性方式】中②一致.

注: 同样的问题是, 在例 0.5.2 的天气预报 (实验 ε) 中试预报明天的最低气温? 类似的讨论得出, 答案是另一个随机变量 $\eta(\omega), \omega \in \Omega$. 显然, $\eta(\omega) \leqslant \xi(\omega)$; 用两个随机变量 $\xi(\omega)$ 和 $\eta(\omega)$ 能够更全面地反映出 "明天冷热情况" 这一随机局部.

从两个例题看出, 概率论回答问题的工具是概率空间和随机变量 (准确地说, 是随机变量产生的概率空间). 用这种方式给出的答案比确定性方式更符合实际情况 (见例 0.5.2); 而且可以回答确定性方式难以回答的问题 (见例 0.5.1).

《新概率论》把研究对象定位为现实世界中的随机局部和随机变量 (含随机变量族, 如随机向量, 随机过程, 随机场, 等等). 全书共十章, 分为三个板块.

第 1 个板块由前五章组成. 任务是首先建立随机宇宙的数学模型 —— 因果空间; 然后研究随机宇宙中一个又一个的随机局部, 建立它们的数学模型 —— 随机试验和概率空间; 最后探讨数学模型中的因果知识和出现规律, 形成概率论的基础理论.

第 2 个板块由第 6—8 章组成. 任务是应用概率空间这个基本工具研究现实世界中普遍存在的 "随不同机会取不同数值的" 变量 —— 随机变量 (含随机变量族), 建立随机变量和随机向量的基本理论; 并为无限随机变量族 (如随机过程, 随机场等) 的研究做准备工作.

第 3 个板块由第 9, 10 章组成. 第 9 章引进特征函数这一新研究工具. 第 10 章讨论独立随机变量和的极限性质, 得到大数定理和中心极限定理. 这两类极限定理发现, 在大量的随机现象中存在非常隐蔽、非常深刻的出现规律, 因此它们成为古典概率论中的最高成就.