

概率论引论

复旦大学 谢践生 编著

(2022 年 8 月校稿版)

序

与欧几里得几何的发展历史比起来，概率论的诞生和发展要晚得多，直到17世纪才有了Fermat和Pascal的书信讨论及其Huygens关于概率论的十几页的一本书。至于建立起概率理论公理化体系，最终使数学家们能够站在一个统一而无歧义的测度论的讨论平台，则更要等到1933年Kolmogorov的奠基性的著作了。我想主要原因是概率事件的概率和随机变量的期望的确是更抽象一些的吧；实际上，与理解长度和面积等物理度量这件事比起来，理解概率、期望等概念还是要更加高上一个层次。但不管我们是否能看清楚、想明白，概率的规则总是支配着我们的现实世界，无时不在，无处不在。

而概率论近期有一个非常迅速而突出的发展，固然是因为其在金融领域、数据科学、人工智能科学的大规模的应用和由此而引起的金融风险的度量的迫切需要，也由于许多挑战性的数学问题的提出和解决。

谢践生教授的《概率论引论》是一个很接地气，而同时又保持了概率论公理系统内在逻辑的严密性的本科生概率论教材，对于培养本科生、研究生的数学素质，也是一本难得的好教材。我本人在读这本教材的同时已感到受益匪浅，可见作者在成书的过程中所注入的心血。本书的一个重要特点是给出了大量的与我们现实世界有紧密联系的意义深刻的例子，这对于读者学习、应用和发展概率理论将会起到很大的作用。

概率论的一些深刻而重要的结果在初学时常会引起一些概念上的迷惑。谢践生教授在度过他的概率理论初学年代那么多年后，还能清楚地追忆起那同学少年的迷茫，并将之提取出来而做透彻的分析，使本书的读者能够更早更快地走出迷津，这也是此书的一个难能可贵之处。

《概率论引论》的另一个重要特点是：作者在讲解到一些概率统计理论的重要结果的同时，也对相应结果的发现者，或相关领域的开创者本人予以介绍。我很赞同这种风格，它大幅度地拉近了那些看似很抽象的数学定理与初涉概率论领域的读者之间的距离，真可能就点通了那位读者的灵犀。

我从小就喜欢爬山；后来发现，不少数学家都有这个爱好。如果你准备爬

概率论这座山，你最好借助一张地图或一本地图册，以免迷失在大山里，而谢践生的《概率论引论》就是一本很好的地图册兼导游书。

彭实戈

2022 年 4 月 11 日

前言与导读

本讲义是我在 2021 年秋季学期为复旦大学数学科学学院的本科生主讲概率论荣誉课程的授课过程中逐渐整理而来。在此之前，我常年从事学院的多门概率论专业课程的教学工作，因此讲义中毫无疑问有众多内容来源于那些教学工作中积累的素材。特别是在开设本科生概率论荣誉课程之前，我刚好差不多完成了自己编撰的《现代概率论基础》课程的电子讲义的修订工作，这为本讲义的出现做了比较充分的准备。

从我本科阶段开始概率论方向的相关课程学习，直到自己成为教师来教授本科生概率论，所使用的教材总是或多或少让我感觉不舒服¹，有众多迷思：几乎所有的初等概率论教材对于为何（及为何如此模式地）引入相关的概念缺乏必要且令人信服的解说。比如说，为何把两事件的独立性定义为 $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ ？为何大多数教材讲完概率论的几条公理后，后续介绍随机变量时不再明确说明随机变量赖以生存的样本空间了？一个取定了的概率空间怎么可能容纳源源不断在讨论的各式各样的随机变量？在实务工作中能用频率来代替概率的道理是什么？其他数学领域中的数学结构如线性空间、群、环、域、流形等有同构的概念，为啥概率论中概率空间作为一种概率结构不好好讲讲它的同构问题？为什么数学期望、条件数学期望按照大多数初等概率论教材中的模式定义？明明到了概率论的高级课程中数学期望、条件数学期望又有另一套抽象定义，这两套定义怎么联系起来？初等概率论中，经典条件概率定义为 $\mathbb{P}(B|A) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$ ，它与条件数学期望又有何联系？另外，数学期望、条件数学期望定义出来到底是为了解决什么实际问题？总不能说成是概率论学者们自娱自乐的数字计算游戏吧？在概率论中经常提分布律与条件分布律，那么到底什么是分布律与条件分布律？类似的问题还有很多。这一系列问题在我所见到的国内外的为初学者而准备的初等概率论教材中很少有给出系统讨论的；那

¹显然，编写本讲义的过程中，作者也从前人所编教材中吸取了它们的一些长处，借鉴了其中的一些有趣的例子与论述等；天道轮回，限于编者的个人能力，本讲义毫无疑问同样可能有一些地方会给读者不舒服的阅读体验，那时可能就是新一代的读者变编者而进一步改进教材的时机。

么作为教师，又应当如何从整体上理解、并在实际教学工作中安排和讲述初等概率论的有关内容呢？这些“为何”及“如何”的问题我最终从概率论的高级课程的教材等材料的字里行间揣摩到了我自己认可的答案，这也是本讲义的写作动机。

带着这些“为何”与“如何”的问题（起初一大部分问题已在前期的学习和科研路上摸索出自己的回答），我在复旦多年的本科生概率论教学实践中，几乎每一年的讲法都与上一年有所差别，其目的就是在探索一种自己满意的（初等）概率论教学内容安排体系，使得学生们再次遇到我当初的困惑场景时，有关疑惑在这个教学体系下能够自然而然地如冬雪春融。这是一个教学的迭代、优化过程；这个过程反映在电子讲义上，我的《现代概率论基础》电子讲义草稿始于2012年，直到2021年才达到我基本满意的状态，但仍需少许修订，近乎十年磨一剑；如果加上前期的教学磨练，则远超十年。这是一个把过往大多数初等概率论教材的教学内容次序推倒之后逐步重新构建的过程；当然概率论的理论还是那些理论，改变的是它们的出场次序以及介绍的模式；无疑，这需要勇气、决心与大量的付出。这也是一个痛并快乐的过程。痛，除了来源于这个探索过程中自身的精神折磨，还来源于外界。众所周知，在较长时期里，国内学界的生态状况是重视科研远远重于教学的情况；近些年虽然管理层从文件精神上要求重视教学，但因为教学上投入的时间和心血难量化比较等原因，从政策的实际落实角度来看，从我的很多院校的兢兢业业服务于教学一线的教师朋友们的吐槽反馈看，实际状况并没有得到多少改善。鉴于有关“游戏规则”，朋友们互相劝慰，在教学中的投入还是对得起自己的良心就算了；教材之类的就等老了、退休有精力再说，毕竟基础学科专业教材的编写费时费力，即使出版后得到好评也就那样。这些酸楚我都能理解，毕竟我们都是这个体系中的感同身受者；我也没有放弃科研，但我更享受科研与教学的相互融合与促进：融入科研体悟的教学和通过教学促进科研这两种快乐的体会我都真真切切地感受过。真正的痛是在教学过程中的不被学生理解，尽管这种状况不多。在我多年的教学过程中，有两次被学生投诉的历史（其中一次投诉原因让人哭笑不得，在此也不细说），还好两次投诉院系领导通融理解，没有最终升级成教学事故而进一步造成悲剧；也曾有同学询问，“为啥课程中要讲那么难的内容，反正我们也学不懂，最后考试也不考”云云，让我略感愤慨而联想到如今的教学管理又有点无言以对。我由于主要为本院的学生讲授专业课程，总体来说师生对数学的价值取向与品味趋同，我的授课方式还是能被大多数同学理解；身边一些为外院系开设数学课程的同事、朋友在教学中的类似故事更多。这里把这些事情写出来，只为呼吁同学们理解那些在教学一线兢兢业业服务的老师

们，特别是青年教师。教学经验是在教学实践的磨练中逐步提高的；我清清楚楚记得自己在职业生涯的早期有过一次挂黑板的经历；至于授课过程中的口误或笔误，应该说只要继续从事教学我就无法向各位保证能避免。同时，我也呼吁管理层对于学生们的教学投诉审慎应对与处理。

除了中学阶段的初步积累，我的一点点物理学理论底子始于北大的理科实验班的“攀登计划”，现如今似乎已改名为“元培实验班”或“元培计划”。后来在跟从钱敏先生、并在刘培东教授实际指导下从事随机动力系统研究的过程中，逐步积攒了点统计力学、热力学的理论基础，也因受钱敏先生影响对物理学发生更多的兴趣。到复旦大学开始概率论课程的教学工作后，为了更好地讲授这门课程，这才又逐渐开始去追寻概率论发展的历史。在反复阅读、思考的过程中注意到 Hilbert（希尔伯特）在他倡导的公理化运动中，把概率和力学的公理化置于物理学的公理化的范畴内，大受震撼；但直到 2021 年秋季的某天，才突然感觉自己摸索到了完成前述重建我心目中的初等概率论教学体系任务的钥匙：把概率论中所有的概念和理论置于概率空间是对现实中的随机现象进行的概率建模这一构想下来进行论述；当然有关论述最好能与学科发展的史实交叉融合。如今的讲义就是在这样的指导精神下组织相关材料，逐步展开有关理论论述的。个人感觉这种组织材料的方式可以给读者呈现概率理论是如何从模糊的直观中萌芽、生长的。Newton（牛顿）本人发展他的力学理论体系似乎也是如此；有关史料给我的感觉是：他对微积分理论的创建，只是在力学理论构建过程中出于解决有关问题的需要，在前人如 Archimedes（阿基米德）、Galileo（伽利略）、Pascal（帕斯卡）等的工作基础上顺手而为而已，在职业生涯的早期他自己也并没觉得有多了不起，相反他更看重基于直观和简单物理现象来发展力学理论体系的哲学思想及与之对应的数学原理²；到了职业生涯的后期他才被迫卷入微积分发明的优先权之争。

在此我们介绍一下全书的框架结构。

在前述“概率建模”的观点和指导精神下，考虑到教学的具体安排，我们把全书正文内容分成了 12 章。

第 1 章用于引出概率论关心的随机现象与随机事件，并简单回顾概率论的发展历史。第 2 章论述了初学者最容易理解和接受的两类满足等可能性假设的概率模型：古典概率模型和几何概率模型；注意到在几何概率模型中，并不是所有的事件都能谈论概率，因此在古典概率模型和几何概率模型的基础上升华出公理化的概率模型也就不难理解，尽管这一点在概率论学科的实际发展进程上整整耗费了 100 多年：从 Laplace（拉普拉斯）在 1812 年发表《分析概率

²Newton 的经典著作是《自然哲学的数学原理》（*Philosophiae Naturalis Principia Mathematica*）。

论》开始，到 Kolmogorov（柯尔莫哥洛夫）在 1933 年发表德文著作《概率论基础》，期间耗费了 121 年。

第³章是全书中一个极其特殊的章节，介绍了经典条件概率与事件独立性，是编者的“概率建模”观点在全书的第一次重要的落实³。我们基于古典概率模型，导出了古典概率模型中的条件概率的概念，并把这一概念外推至一般的公理化概率模型；同样基于古典概率模型讨论了如何为使用两套赌博道具的复杂赌博进行概率建模，由此导出了公理化概率论框架下概率空间的乘积空间这一概念，并在此基础上引出了多个事件的相互独立性这一概念。这也很自然地回答了《测度论》中为何要讨论乘积可测空间与乘积测度空间、为何要探讨测度扩张问题，同时乘积空间也是初等概率论的后续课程中有关随机过程的轨道构造等理论探讨的必备手段；我们的讲义作为初学者的教材，在这些地方只能简单引用《测度论》等的相关结论，而把这些结论的略细致论述统一放在了下一章，即第⁴章⁴。

概率论的发展（特别是有关随机变量及其数学期望的公理化定义等内容）受到同时期的其他多个数学分支发展的影响与推动：例如微积分、测度论，特别是 Lebesgue（勒贝格）积分理论等的发展对概率论有深远的影响。为此我们在全书中插入了另一个特殊的、分析味道浓重的章节作为第⁴章，用于介绍二十世纪才发展起来的 Lebesgue 积分的理论；除了传统的欧氏空间中的 Lebesgue 积分理论，里面顺带介绍了一般测度的抽象 Lebesgue 积分的理论。其中，欧氏空间中的 Lebesgue 积分理论主要是为连续型随机变量的密度函数的定义、分布律的计算等做铺垫，一般测度的抽象 Lebesgue 积分理论主要是为数学期望、条件数学期望的定义做技术准备。

初等概率论中引进随机变量的一大目的是为了在概率建模中更方便、并且更精确地描述有关复杂的随机现象与随机事件，而这些随机现象与随机事件的概率计算、独立性的判断等都可以归结到随机变量的分布律的演算；在随机变量概念引入后也很自然有随机变量的相互独立性、随机变量的“平均值”与“条件平均值”（数学期望与条件数学期望）等直观概念，这些由随机变量诱导出来的概念在计算公式上也同样可以归结到随机变量的分布律的演算。因此随机变量及其分布律是初等概率论中的一个核心内容。综合考虑全书理论陈述的条理性和教学的便利性，我们把随机变量的介绍拆分成三个不连贯的章

³但毫无疑问并不是整个概率论发展历史中的第一次：我个人觉得古典概率模型的建立可以认为是第一次虽然粗浅但也完美的“概率建模”，它第一次把模糊的自然语言中的概率在最容易理解的情况下予以了数学的精确化定义。

⁴这类知识的教学似乎置于《测度论》之类的课程更为适合，在本课程的教学由于教学时间等原因建议只作介绍而不展开讨论。

节进行：第[五]章先介绍随机变量及其分布律和随机变量的独立性的概念，之后介绍了随机变量中最容易理解的离散型随机变量的概念以及常见的离散型分布；第[四]章先介绍连续型随机变量及其密度函数的概念，之后介绍了概率微元法用于解决（光滑可逆或分片光滑可逆）变换下连续型随机变量的密度函数的计算问题，最后介绍了常见连续型分布及它们的密度函数；第[九]章介绍了一个奇异型随机变量的例子并讨论了随机变量的分布分类的问题，此外还讨论了一般分布函数的性质、随机变量的实现问题以及次序统计量的一些初步结果。在第[五]章，我们指出，随机变量的定义应当在可测函数的认同基础上做一个“补丁程序”，这与我们“概率建模”的精神是一致的。

在第[六]、[七]章，我们分别介绍了概率论中两个重要的工具：数学期望与条件数学期望。在第[六]章的论述中，我们通过回顾历史上 Huygens（惠更斯）提出的期望收益的概念，给出严谨的数学期望的定义，并通过引入保测映射的观点最终给出数学期望的计算公式。第[三]章定义的经典条件概率结合第[六]章刚刚定义的数学期望概念，在第[七]章中进一步发挥，导出了经典条件概率对应的条件数学期望，这个概念进一步拓广成更一般的条件数学期望的概念，由此引出更抽象的条件概率的定义；特别地，我们进一步可以探讨所谓的条件分布律，基于此我们最终给出了概率论中常用的一个随机变量关于另一个随机变量的条件数学期望的计算公式；这些内容请参见第[七]章的有关论述。

以上章节中多处概念的论述都遵循了物理学或数学中拓广概念的原则，并且介绍这些概念时都是尽力先从最容易理解的情形开始的。

本书正文最后三章是全书的高潮。在介绍了随机变量的几种收敛性概念后，第[十]章基于 Chebyshev（契比雪夫）不等式系统介绍了 Borel-Cantelli 第一引理和第二引理等论证强大数律的重要工具，进而讨论了强弱大数律，这是对本书第[一]章中“频率稳定性”现象的呼应；另外，该章中 Chebyshev 不等式的一个推论也给出了前述随机变量概念的“补丁程序”的一种有效率的落实方法。第[十一]章通过介绍特征函数理论讨论了中心极限定理。为了呼应我们的“概率建模”的观点，在正文最后一章即第[十二]章⁵中，我们分两节给出了两部分的讨论：（I）公理化概率论观点下“经验的总结与外推”；（II）基于大样本理论的对模型的检验。在（I）部分中，我们首先用两个简单模型讨论了第一章习题中介绍的赌徒佯谬与热手效应（涉及“经验的总结”与“经验的外推”）；之后讨论了在概率论观点下的“经验的总结”，见定理[12.1.1]，其中涉及的极限分布律恰好是热力学中的 Gibbs（吉布斯）态。在（II）部分中，我们给出了可以借助大样本数据来检验模型准确性的如下几个定理：

⁵该章（第[十二]章）建议作为阅读材料，在实际教学中仅粗略介绍即可。

- (1) 作为统计学基本定理的 Glivenko-Cantelli 定理。实务上我们可以基于这个理论，利用大样本数据做 P-P 图或 Q-Q 图，形象（或者说“定性”）地检验样本是否来自某个具体的分布；
- (2) Kolmogorov 定理。这个定理由于其证明需要更高级的技巧，在本书中就仅在附录 D 中提供了在初等概率论范围内能大致理解的一个粗线条的证明框架。实务上我们可以基于这个定理，构建 Kolmogorov 检验，“定量”地检验大样本是否来自于某个分布函数连续分布；
- (3) Pearson 定理。实务上，基于 Pearson 定理，我们可以构建拟合优度检验，“定量”地检验大样本是否来自于某个简单的离散型分布；
- (4) Pearson-Fisher 定理。实务上，基于这个定理，我们可以构建关于独立性的列联表检验，基于大样本数据“定量”地判断“二维”数据之间是否具有独立性。

当然，以上四个定理只做到了对初等概率论中最重要的随机变量的分布律与独立性的检验。虽然在概率与统计等方向学者们几百年的研究积累下，现代概率论与现代统计学的理论已经非常丰富，发展得比较成熟，但贴近我们关心的一些更复杂的随机现象的概率建模理论以及对这样的概率模型的检验理论这两个方面仍然可能有待读者们去探索和发展。正如分析学中极限概念的引入与探讨是初等数学与高等数学之间的分水岭，我们特地以第四、十一章节作为初等概率论课程的两个收尾章节，用以介绍在整个概率论中都占据核心地位的两个经典理论与研究范式——大数律和中心极限定理。而第十二章，最后一个收尾章节，则可视作我们概率建模观点下对数理统计课程和“高等”概率论课程的广告：除了该章节介绍的模型检验的理论，更多的对模型的检验问题留待读者在《统计学》的有关课程中去学习和研究；但另一方面，我们也不难看到，要论证上述 Kolmogorov 定理，需要随机过程的语言以及泛函中心极限定理等更高端的概率理论，而在定理12.1.1的证明过程中，大偏差理论起到了核心的作用。因此这也是对随机过程、大偏差等“高等”概率论课程的广告。

为了全书的相对自封性，我们还准备了四个附录章节：前两个附录章节用来补充探讨全书中因教学安排的原因未能详尽展开的一些证明细节；第三个附录章节用来继续讨论在特殊场合能刻画随机变量的分布律的两种统计特征——矩和分布函数/分布测度的 Laplace 变换。最后一个章节用来提供前述 Kolmogorov 定理的粗线条的证明框架，其中顺带介绍了泛函中心极限定理等密切关联的内容。

在本书各章编者都尽力配备了相应的习题，部分较难习题提供了提示信息；这些习题大部分来源于众多参考文献，少部分是自编习题，另有少部分来源于知乎社区等网络渠道⁶。全书行文中，除了通常的人物信息方面的注记（我们使用“🔔 注”标识）和对有关理论结果的注记（我们使用“🔔 注记”标识）外，书中定理、引理、推论也各自使用了不同的符号标识，依次为☞、♠、♦。在本书的最后还准备了一个基于分类思想（而不是完全基于拼音顺序）的索引。

编者希望这样的论述方式和结构安排能够激发初学者学习概率论的热情与兴趣，提高学习的效率。当然，这样一来，全书虽有主线，但内容还是比较繁杂的⁷，对于学生的自学和教师的教学都提出了挑战。本书对学生的数学基础有较高的要求，相对而言更适合专业方向的学生学习。我的建议是：学生初次学习时要学会抓住主线，忽略一些对数学基础要求较高的证明细节，在学习概率论发展历史和概率理论体系的构建方法的过程中去学习和掌握有关理论；在数学基础提高后宜再次学习，补上当初忽略的那些理论细节。对教师的具体授课而言，就需要教师根据实际需要调整教学内容。在我自己的实际教学中，第十二章的内容通常不专门讲述，留给学生自行探索；第四章内容虽然也不专门讲述，但其中几个分析的大结果（单调收敛定理、Fatou 引理、Lebesgue 控制收敛定理，特别是 Fubini 定理）穿插于有关概率问题中展示其应用，让学生们在概率论的应用中熟悉和掌握这些大结果；其他一些章节（如第七、九、十、十一章）的内容也是根据课程时间进度和学生对讲义内容的掌握情况有所筛选地进行课程的讲解。

很庆幸一路走来获得了一些经济援助、结识了一些尊敬可爱的师友，温暖了内心；因为内心有爱，这才能在如今的教学岗位上做到持续用心与坚守底线。我大学阶段是一名来自农村的贫困生，幸得“晨兴助学金”资助，又得彼时舍友与同学的照顾介绍勤工俭学机会，稍解彼时困顿。彼时的老师们在生活上大多看淡物质享受与物质利益、却又能理解并在力所能及时纾解学生们生活中的困苦；在科研上他们重视对未知的探索过程中的严谨与求真、求实；在教学上则一丝不苟、兢兢业业，讲到自己熟悉领域时神采飞扬，与学生的讨论中平等相处。所有这些风貌让我产生了以后要成为这样的教师的冲动。研究生阶段钱敏平教授开朗、乐观的个性感染了作为学生的我，并从她那里学到了“敝帚自珍”，珍视自己于黑暗中摸索出的各种收获；而刘培东教授在指导我论文的过

⁶书中全部人物图片、部分其他图片资源也来源于各种网络渠道，特作此说明并向分享方（万维百科、知乎社区等）致谢。

⁷但这是概率论学科本身特点的一部分，而不是作者以难为学生取乐而故意如此编写本讲义。王小波说，“人生在世，就如一本打开的书，我们更希望这本书的主题始终如一，不希望它在中途改变题目”；编者也希望全书没有偏离主题。

程中表现出来的严谨与细致让我暗下决心，日后亦当如是。钱敏先生则被我自己视作精神教父，其为人、行事及在教育和科研上的表现与相关思想时常被我用来回忆观照己身；斯人已逝，唯记忆永存心底，怀念时不时涌现脑海。勇士屠恶龙的童话小说中，勇士成为英雄后由于种种原因成为了另一条恶龙；生活中的挫折、苟且多了，你也可能忘记你的诗与远方。我党提出要“不忘初心”；而我也时常心心念念提醒自己，至少不要成长为自己当初厌弃的模样。说实话，走上科研和教育的道路，对于在老家农村的老母及兄姐是有无数欠疚的，因为只能成为一个无法承担更多家庭责任的自了汉了。感恩家人的理解，感恩从事慈善事业的各界人士，感恩一路走来师友的教育、提携、理解与帮助！

讲义中的内容借鉴了众多师长的材料，如我的同学王岩华向我分享的何书元老师的《概率论》PPT 资料，我自己学生时代向钱敏平（《概率论》）、陈家鼎（《测度论》）、程士宏（《高等概率论》）等教授学习概率论相关课程的笔记。在复旦的教学过程中较长时期参考了我院同事应坚刚教授的《概率论》（与何萍教授合著），其他参考过的教材资料此处就不一一细说，见参考文献。而浙大的朋友赵敏智老师提供的分享以及与她的相关讨论尤为珍贵。在书稿临近完成时，通过与钱紘教授交流讨论有关学术问题，我获益很多，直接导致了第 15 章中第一节内容的出现。我的老师钱敏平教授及我的师兄刘勇教授都为本讲义提出了不少重要的建设性的修改意见；还有很多朋友及参加我课程的多位同学指出了草稿中的一些细节性的打印错误等，此处就不一一列出。在本讲义最终定稿前，作者特别邀请了戴捷博士帮忙审阅，他的诸多修改建议进一步提高了讲义的严谨性和可读性；期间我的同学黄涛教授应我统计学历史方面的咨询，特地为我提供了文献 [15, 26, 27, 29, 30, 51, 52, 53, 57]。特在此向所有这些良师益友致谢！

对于这个讲义来说，如果没有教学的讲台，没有讲台下学生们的宽容与理解，或许它就不会诞生、即使诞生也可能夭折。因此也感谢学院提供了这个教学舞台，感谢 2021 年秋季学期参加我课程的所有同学！

虽经多次修改才定稿，本讲义无疑仍然会包含一些错误或疏漏不当之处，敬请各位同学、同行持续批评指正；也欢迎你们提出写作上的建议，分享你们觉得有趣的案例资料等，以利适当机会继续修订。

谨以此作为讲义前言与导读。

2021 年 11 月 13 日初稿、2022 年 3 月 08 日二稿于复旦光华楼

2022 年 8 月 30 日定稿于上海宝山

目录

序	iii
前言与导读	v
第一章 引言：随机现象与概率论	1
1.1 随机现象与随机事件	1
1.2 频率与概率	3
1.3 概率论简史	5
习题 1	18
第二章 从古典概率模型、几何概率模型到概率论的公理	23
2.1 事件与集合	23
2.2 古典概率模型	27
2.2.1 古典概率模型简介	28
2.2.2 计数方法简介	29
2.2.3 古典概率模型应用实例：(1) Polyá 坛子模型	30
2.2.4 古典概率模型应用实例：(2) 同生日问题	32
2.2.5 古典概率模型应用实例：(3) 唱票问题与反射原理	34
2.2.6 古典概率模型应用实例：(4) 图论中的染色问题	35
2.3 几何概率模型	37
2.3.1 几何概率模型简介	37
2.3.2 几何概率模型应用实例	38
2.3.3 Bertrand 悖论	40
2.4 公理化的概率模型及其性质	42
2.4.1 概率的公理	42
2.4.2 概率的基本性质	44

2.4.3 Jordan 公式	45
习题 2	47
第三章 经典条件概率与事件独立性	55
3.1 经典条件概率	55
3.1.1 条件概率空间与经典条件概率的定义	55
3.1.2 乘法公式	59
3.1.3 全概率公式	61
3.1.4 Bayes 公式	68
3.2 乘积概率空间与事件独立性	74
3.2.1 乘积概率空间	74
3.2.2 事件独立性的定义	77
3.2.3 重复实验与条件实验	79
3.3 讨论: 事件的概率与条件概率及其概率值大小	81
3.3.1 易混淆概念: 特定语境中的概率与条件概率	82
3.3.2 零概率、小概率与大概率	83
习题 3	85
第四章 Lebesgue 积分理论简介	93
4.1 可测集与可测函数	93
4.1.1 Borel 可测集与一般可测集	93
4.1.2 Borel 可测函数与一般的可测映射	94
4.2 Lebesgue 测度与一般可测空间中的非负测度	96
4.2.1 测度、符号测度、预测度与外测度的定义	96
4.2.2 符号测度的 Hahn 分解与 Jordan 分解	98
4.2.3 Carathéodory 扩张与 Lebesgue 测度	101
4.3 欧氏空间中的 Lebesgue 积分	107
4.3.1 Lebesgue 积分的定义	107
4.3.2 Lebesgue 积分与极限的交换	109
4.3.3 Lebesgue 积分与 Riemann 积分	112
4.3.4 微积分基本定理的探讨	116
4.3.5 重积分与累次积分—Fubini 定理	118
4.4 抽象测度的 Lebesgue 积分	120
4.4.1 抽象测度的 Lebesgue 积分的定义	120
4.4.2 抽象测度的 Lebesgue 积分的极限交换问题	121

4.4.3 微积分基本定理的推广: Radon-Nikodym 定理	122
习题 4	126
第五章 随机变量及其分布律 (I)	129
5.1 随机变量、随机向量及其分布律	130
5.1.1 随机变量与随机向量的定义	130
5.1.2 随机变量/向量之间的相互独立性	134
5.2 离散型分布	136
5.2.1 离散型随机变量的定义	136
5.2.2 常见离散型分布	138
5.2.3 应用 (I): 真假随机性的判断	142
5.2.4 应用 (II): 两个极大似然估计的例子	145
习题 5	146
第六章 数学期望与分布律	149
6.1 数学期望的定义	149
6.2 数学期望的性质与简单应用	152
6.2.1 数学期望的性质	152
6.2.2 数学期望的简单应用	153
6.3 基于分布律的数学期望计算公式	156
6.3.1 数学期望的计算公式 (I)	156
6.3.2 概率空间的同构及其应用	159
6.3.3 数学期望的计算公式 (II)	161
6.4 方差、协方差与独立性	161
习题 6	164
第七章 条件数学期望与条件分布律	169
7.1 条件数学期望的定义	169
7.1.1 数学期望的一个性质	170
7.1.2 从经典条件概率到“经典”条件数学期望	170
7.1.3 抽象的条件数学期望的定义	173
7.1.4 从条件数学期望 $\mathbb{E}[\cdot \mathcal{G}]$ 到条件概率 $\mathbb{P}(\cdot \mathcal{G})$	176
7.1.5 抽象条件数学期望 $\mathbb{E}[Y \mathcal{G}]$ 的计算	177
7.2 条件数学期望的性质	177
7.3 条件数学期望的积分变换公式	179

7.4 条件分布律及条件数学期望的计算公式	180
7.5 条件数学期望与独立性	184
7.6 条件数学期望的应用	185
习题 7	187
第八章 随机变量及其分布律 (II)	191
8.1 连续型分布与密度函数的定义	191
8.2 连续型随机变量的数学期望计算公式	193
8.3 边缘密度、条件密度与条件分布函数	193
8.4 概率微元法	195
8.4.1 光滑可逆变换情形	195
8.4.2 分片光滑可逆变换情形	196
8.5 常见连续型分布	198
习题 8	214
第九章 随机变量及其分布律 (III)	219
9.1 一个奇异型分布的例子及随机变量的分类	219
9.2 分布函数的性质与随机变量的实现	221
9.2.1 分布函数的性质	221
9.2.2 随机数发生器的构造	223
9.2.3 次序统计量	225
习题 9	228
第十章 随机变量列的收敛与大数律	233
10.1 Chebyshev 不等式	233
10.2 各种收敛性的定义	236
10.3 几种收敛之间的关系	237
10.3.1 L^p -收敛、依概率收敛与几乎处处收敛之间的关系	237
10.3.2 阅读材料: 随机序与随机控制收敛定理	240
10.3.3 依概率收敛与依分布收敛之间的关系	242
10.4 大数律简介	243
10.4.1 Borel-Cantelli 引理与 Borel-Cantelli 第二引理	243
10.4.2 从 Bernoulli 弱大数律到 Borel 强大数律	249
10.4.3 从 Khintchine 弱大数律到 Kolmogorov 强大数律	251
10.5 应用举例	255

习题 10	259
第十一章 随机变量的特征函数与中心极限定理	269
11.1 特征函数与分布函数	269
11.2 特征函数的三大定理	272
11.2.1 特征函数的唯一性定理	272
11.2.2 特征函数的连续性定理	276
11.2.3 特征函数的刻画定理	281
11.3 随机变量对称化方法简介	281
11.4 中心极限定理	283
习题 11	287
第十二章 概率建模的有关讨论与检验	297
12.1 概率论视角下的“经验总结与外推”	298
12.1.1 关于赌徒佯谬与热手效应的讨论	298
12.1.2 Gibbs 分布与 Gibbs 条件概率	301
12.2 模型的检验问题	306
12.2.1 对简单样本总体的分布律的检验	306
12.2.2 对独立性的检验	310
习题 12	317
附录 A 单调类定理及其应用	321
A.1 集合类简介	321
A.2 单调类定理	324
A.3 乘积概率测度的存在唯一性	325
A.4 定理 5.1.1 的证明: (5.3) $\Rightarrow$ (5.5) 部分	326
习题 A	327
附录 B 数学期望、条件数学期望的一些性质的证明	329
B.1 数学期望的一些性质的证明	329
B.1.1 数学期望的线性性质	329
B.1.2 Jensen 不等式与矩不等式的证明	330
B.1.3 Hölder 不等式与 Minkowski 不等式的证明	331
B.2 条件数学期望的一些性质的证明	332
B.2.1 全期望公式	332
B.2.2 条件数学期望 $\mathbb{E}[\cdot \mathcal{G}]$ 的单调性与 $\mathcal{G}$ -线性性质	332

B.2.3 条件数学期望 $\mathbb{E}[\cdot G]$ 的性质 (CE6) 的证明	333
B.2.4 条件 Jensen 不等式的证明	334
习题 B	334
附录 C 矩问题与 Laplace 变换	337
C.1 矩问题	337
C.1.1 Hausdorff 矩问题	338
C.1.2 Hamburger 矩问题: 可解的条件	338
C.1.3 Hamburger 矩问题: 解的唯一性与不唯一性	340
C.2 Laplace 变换	342
C.2.1 Laplace 变换的定义与基本性质	342
C.2.2 Laplace 变换的唯一性定理	345
C.2.3 Laplace 变换的连续性定理	346
C.2.4 Laplace 变换的刻画定理	346
C.2.5 Laplace 变换的 Tauber 定理	347
习题 C	351
附录 D Kolmogorov 定理的“证明”	355
D.1 Brown 运动、Brown 桥及其特征函数刻画	355
D.2 概率测度的弱收敛与泛函中心极限定理	357
D.2.1 欧氏空间与一般度量空间中概率测度的弱收敛	357
D.2.2 相对紧与胎紧: Prokhorov 定理	359
D.2.3 泛函中心极限定理	360
D.3 Kolmogorov 定理的“证明”	361
习题 D	362
参考文献	365
索引	373

第一章 引言：随机现象与概率论

现代人衣食住行的方方面面充斥着“概率”、“几率”、“机会”、“可能性”、“随机”等字眼。每天早晨起床，天气预报会告知你，今天是晴天还是阴天，降水概率是多少；学生们期末考试后会估计自己获得 A 类成绩的几率有多大；求职的时候，人们也会估计自己被录用的机会有多大；在农作物收获前，农民们有时会提前估计自己农田高产的可能性有多高。走在大街上，当你和朋友不知道去哪家饭店吃饭时，你可能会掏出一枚硬币，让它“随机”地为你挑选一家。我们自然语言中的“概率”、“几率”、“机会”、“可能性”通常认为是同义词¹。那么到底什么是概率？在本章，我们无法完整回答这个问题。但我们希望学完本书，读者对此问题已有自己的答案与体悟。

几乎所有的概率论教科书都告诉我们，概率论是研究随机现象的数量规律的数学分支学科。那么什么是随机现象？我们通常所说的某个随机事件的概率又是什么含义呢？

1.1 随机现象与随机事件

在自然界和社会生活中，我们经常遇到两类现象：一类是确定性现象，它的结果是单一而确定的；一类是不确定性现象（也称为随机现象、偶然现象），它的结果是不唯一的，并且通常每种结果都有可能发生。

例 1.1. 以下一些现象或事件是确定的（即一定发生）：

- （在地球上观察）太阳东升西落；
- （地球上的）学校操场上向上抛出的帽子最终会掉落；
- （在地球上，常温常压条件下）人受伤、血管裂开了会流血。

¹在英文、法文等外语中，“probability”（概率，法文是“probabilité”）一词的拉丁词根是“probare”和“habilis”。前者的意思是“去试验、去证明，或者是去批准”；后者表示“才能、技术和能力”。在后来的使用中，“probable”一词才渐渐有了“可能的、合理的”这一层意思，这才跟随机性产生了联系。