高等院校理工类规划教材

概率论与随机过程

GAILÜLUN YU SUIJI GUOCHENG

主编 周 清 张丽华

北京邮电大学出版社
www.buptpress.com

内 容 简 介

本书共分为两个部分. 第一部分为概率论基础, 包括第1~5章, 其中第1~4章主要介绍了概率空间、可测函数、随机变量及其分布、随机向量变换、条件数学期望、一维和高维随机变量的特征函数等本科阶段尚未或较少涉及的内容; 第5章介绍了在概率论与随机过程中常用的随机变量序列的收敛概论和性质. 第二部分为随机过程基础, 包括第6~10章, 其中第6章介绍了随机过程的基本概念、基本类型以及布朗运动和维纳过程的相关知识; 第7章主要介绍了泊松过程及其生成算法; 第8章除了介绍宽平稳过程的基本概念之外, 还重点讲述了平稳过程相关函数的谱分解; 第9、10章介绍了齐次Markov链、可数齐次Markov过程的基础内容.

本书可以作为对概率论与随机过程理论要求较高的工科研究生的学习用书, 也可以作为一般专业的工科研究生或数学专业本科生"概率论与随机过程"课程的参考书.

图书在版编目(CIP) 数据

概率论与随机过程 / 周清, 张丽华主编. -- 北京: 北京邮电大学出版社, 2023. 7
ISBN 978-7-5635-6937-3

Ⅰ. ①概… Ⅱ. ①周… ②张… Ⅲ. ①概率论②随机过程 Ⅳ. ①O211

中国国家版本馆 CIP 数据核字（2023）第 122013 号

策划编辑: 彭 楠 **责任编辑:** 王晓丹 耿 欢 **责任校对:** 张会良 **封面设计:** 七星博纳

出版发行: 北京邮电大学出版社
社 址: 北京市海淀区西土城路 10 号
邮政编码: 100876
发 行 部: 电话: 010-62282185 传真: 010-62283578
E-mail: publish@bupt.edu.cn
经 销: 各地新华书店
印 刷: 北京虎彩文化传播有限公司
开 本: 787 mm×1 092 mm 1/16
印 张: 14.5
字 数: 350 千字
版 次: 2023 年 7 月第 1 版
印 次: 2023 年 7 月第 1 次印刷

ISBN 978-7-5635-6937-3 定价: 46.00 元

· 如有印装质量问题, 请与北京邮电大学出版社发行部联系 ·

前　　言

由于教学工作的需要, 编者 15 年来一直担任北京邮电大学研究生学位课 “概率论与随机过程” 的教学工作. 本书就是编者在 15 年的教学过程中, 对一直授课使用的王玉孝老师（已故）编写的《概率论与随机过程》教材进行修订而形成的.

“概率论与随机过程” 这门课对通信、信息、计算机、人工智能、管理类专业的研究生来说是一门十分重要的基础课. 通过对本课程的深入学习, 学生不仅可以获得相关知识, 而且可以掌握处理 “ 随机问题” 的方法.

工科 “概率论与随机过程” 的教材与以测度论为基础的近代概率论（包括随机过程）之间有一鸿沟, 填平这一鸿沟或者至少在这一鸿沟上架一座桥梁, 是提高工科研究生概率论与随机过程知识水平必不可少的工作. 本书努力在这方面进行尝试, 这不仅是编者的意愿, 也是北京邮电大学部分研究生导师的共同要求. 本书既具有一定的理论深度, 又不要求学生具有过多的数学基础, 能为学生深入学习相关的数学理论及在各自专业中应用这些理论打下一定的基础.

本书第 1~3 章论述了以测度论为基础的近代概率论的最基本的概念、方法和理论; 第 4 章介绍了在概率论中起重要作用的随机变量特征函数的主要内容; 第 5 章介绍了在概率论与随机过程中常用的随机变量序列的收敛概论和性质; 第 6 章给出了随机过程的基本概念、基本类型以及布朗运动和维纳过程的基础知识; 第 7 章主要介绍了泊松过程及其生成算法; 第 8 章除了介绍宽平稳过程的基本概念之外, 还重点讲述了平稳过程相关函数的谱分解; 第 9、10 章介绍了齐次 Markov 链、可数齐次 Markov 过程的基础内容.

考虑到本书的主要对象不是数学专业的学生, 所以本书部分内容没有给出严格的数学推导, 有些定理也略去了证明, 但指明了出处, 便于学生深入学习.

本书可以作为对概率论与随机过程理论要求较高的工科研究生的学习用书, 也可以作为一般专业的工科研究生或数学专业本科生 “概率论与随机过程” 课程的参考书.

王玉孝老师编写的《概率论与随机过程》教材是本书编写的起点, 本书保留了它的习题及其他部分内容, 在此向已故的王玉孝老师表示深深的敬意.

本书在编写过程中参考了一些文献, 吸取了部分专著的优点, 在此对这些文献和专著的作者表示诚挚的感谢.

本书为北京邮电大学 “十四五” 规划教材（第一批）。孙洪祥教授、郭永江教授、唐碧华教授、杨建奎教授、黎淑兰老师等对本书的编写提出了很好的意见和建议, 研究生李晓楠、王胜安、雷子琦、李钰琪、张梦媛、候惠敏、杨秀琪对本书的编写和校对提供了很大的帮

助, 在此表示衷心的感谢.

北京邮电大学理学院的老师和研究生院的工作人员, 在我的研究生教学工作中给予了很大的支持, 借此机会向他们致谢.

最后, 感谢闵祥伟教授, 感谢我的家人们给予我的爱和支持.

由于编者水平有限, 书中难免有错误和疏漏之处, 敬请读者指正.

周　清

2022 年 12 月 6 日于北京邮电大学理学院

目　录

第 1 章　概率空间 ··· 1
1.1　σ-代数与 Borel 域 ··· 1
1.2　概率空间 ··· 3
1.2.1　概率的公理化定义 ··· 3
1.2.2　概率的性质 ··· 4
1.2.3　条件概率空间与事件的独立性 ··· 5
1.3　习题 ··· 6
第 2 章　可测函数、随机变量及其分布 ··· 8
2.1　可测函数 ··· 8
2.1.1　映射及其逆象 ··· 8
2.1.2　可测函数的定义及判别 ··· 9
2.2　一维随机变量及其分布 ··· 10
2.2.1　随机变量的定义 ··· 10
2.2.2　分布函数及其性质 ··· 12
2.2.3　存在性定理 ··· 13
2.3　高维随机变量及其分布 ··· 13
2.3.1　高维随机变量及其分布 ··· 13
2.3.2　离散型分布和连续型分布 ··· 14
2.3.3　边缘分布 ··· 15
2.4　Borel 函数与随机变量的函数 ··· 16
2.5　随机变量的独立性和条件分布 ··· 17
2.5.1　随机变量的独立性 ··· 17
2.5.2　条件分布 ··· 20
2.6　随机变量函数的分布 ··· 23
2.6.1　一般方法 ··· 23
2.6.2　特殊方法 ··· 27
2.7　习题 ··· 31
第 3 章　随机向量的数字特征 ··· 34
3.1　数字特征、矩母函数 ··· 34
3.1.1　数字特征 ··· 34
3.1.2　Riemann-Stieltjes 积分 ··· 35

3.1.3 关于概率测度的积分 ······ 36
3.1.4 矩母函数 ······ 40
3.1.5 关于数学期望的一些说明 ······ 41
3.2 条件数学期望及其性质 ······ 43
3.2.1 关于事件的条件数学期望 ······ 43
3.2.2 关于随机变量的条件数学期望 ······ 45
3.2.3 关于 σ-代数的条件数学期望 ······ 51
3.3 随机向量的数字特征 ······ 52
3.3.1 均值 (阵) ······ 52
3.3.2 方差阵 ······ 53
3.3.3 协方差阵 ······ 53
3.3.4 条件均值 ······ 54
3.3.5 条件方差阵 ······ 54
3.4 n 维正态随机变量的一个性质 ······ 54
3.5 几个重要的不等式 ······ 56
3.6 习题 ······ 59
第 4 章 随机变量的特征函数 ······ 62
4.1 随机变量的特征函数 ······ 62
4.1.1 随机变量的特征函数 ······ 62
4.1.2 特征函数的性质 ······ 64
4.1.3 逆转公式和唯一性定理 ······ 68
4.2 n 维随机变量的特征函数 ······ 73
4.2.1 n 维随机变量的特征函数的定义 ······ 73
4.2.2 n 维随机变量的特征函数的性质 ······ 75
4.2.3 逆转公式和唯一性定理 ······ 76
4.2.4 随机变量 $X_1, \cdots, X_n$ 独立的充要条件 ······ 77
4.3 n 维正态分布 ······ 78
4.3.1 n 维正态分布的特征函数 ······ 78
4.3.2 n 维随机变量服从正态分布的充要条件 ······ 80
4.3.3 n 维正态随机变量的线性不变性 ······ 81
4.4 习题 ······ 82
第 5 章 极限定理 ······ 85
5.1 随机变量序列的 4 种收敛性 ······ 85
5.1.1 依概率收敛 ······ 85
5.1.2 r 阶收敛 ······ 86
5.1.3 以概率 1 收敛（几乎处处收敛） ······ 88
5.1.4 依分布收敛 ······ 91

5.2 分布函数列弱收敛的条件 ······ 94
5.3 习题 ······ 97
第 6 章 随机过程的概念及其统计特性 ······ 98
6.1 基本概念 ······ 98
6.1.1 随机过程的概念 ······ 98
6.1.2 随机过程的概率分布 ······ 99
6.1.3 随机过程的数字特征 ······ 101
6.1.4 二维随机过程的分布函数和数字特征 ······ 103
6.1.5 复随机过程 ······ 105
6.2 随机过程的基本类型 ······ 106
6.2.1 平稳过程 ······ 106
6.2.2 独立增量过程 ······ 108
6.2.3 高斯 (正态) 过程 ······ 109
6.2.4 Markov 过程 ······ 112
6.3 布朗运动和维纳过程 ······ 112
6.4 习题 ······ 117
第 7 章 泊松过程 ······ 119
7.1 泊松过程的概念 ······ 119
7.2 与齐次泊松过程相关的若干分布 ······ 123
7.2.1 齐次泊松过程到达时间与到达时间间隔的分布 ······ 123
7.2.2 事件发生时刻的条件分布 ······ 126
7.2.3 齐次泊松过程的生成 ······ 129
7.3 泊松过程的推广 ······ 131
7.3.1 非齐次泊松过程的概念 ······ 131
7.3.2 非齐次泊松过程的生成 ······ 134
7.3.3 复合泊松过程 ······ 137
7.4 习题 ······ 139
第 8 章 平稳过程 ······ 141
8.1 平稳过程及其数字特征 ······ 141
8.1.1 平稳过程的概念 ······ 141
8.1.2 相关函数的基本性质 ······ 143
8.2 随机分析 ······ 143
8.2.1 均方收敛 ······ 143
8.2.2 均方连续 ······ 145
8.2.3 均方导数 ······ 147
8.2.4 均方积分 ······ 151
8.3 平稳过程的遍历性定理 ······ 155

8.3.1 遍历性的定义 ······ 155
8.3.2 随机过程具有遍历性的条件 ······ 157
8.4 平稳过程相关函数的谱分解 ······ 158
8.5 习题 ······ 162
第 9 章 Markov 链 ······ 164
9.1 Markov 链的基本概念 ······ 164
9.1.1 Markov 链的基本概念 ······ 164
9.1.2 Markov 链的转移概率 ······ 165
9.1.3 Markov 链的有限维分布 ······ 168
9.2 Markov 链的状态分类 ······ 171
9.2.1 互通和闭集 ······ 171
9.2.2 周期 ······ 173
9.2.3 状态分类 ······ 174
9.2.4 状态分类的判定法 ······ 178
9.3 状态空间的分解 ······ 185
9.3.1 状态空间的分解 ······ 185
9.3.2 不可分闭集 ······ 186
9.3.3 有限链的状态空间 ······ 188
9.3.4 不可分链的状态空间 ······ 189
9.4 平稳分布 ······ 190
9.4.1 $p_{ij}^{(n)}$ 的极限行为 ······ 190
9.4.2 平稳分布 ······ 193
9.5 习题 ······ 202
第 10 章 时间连续、状态离散的 Markov 过程 ······ 206
10.1 Markov 过程与转移函数 ······ 206
10.1.1 基本概念 ······ 206
10.1.2 转移函数的性质与有限维分布 ······ 207
10.2 柯尔莫哥洛夫向前方程和向后方程 ······ 207
10.3 连续时间马氏链的状态分类及例子 ······ 212
10.4 习题 ······ 218
参考文献 ······ 220
附录 主要记号 ······ 221

第 1 章 概率空间

1.1 σ-代数与 Borel 域

概率论是研究随机现象的数量规律的数学分支, 理论严谨, 应用广泛. 在概率论及有关测度理论的讨论中, 我们常常会遇到一些集合类, 其中涉及集合的可列并或者可列交运算, 下面给出下列定义.

定义 1.1.1 设 Ω 是任一非空集合, $\mathscr{F}$ 是由 Ω 的一些子集组成的非空集合类, 如果 $\mathscr{F}$ 满足:

(1) $\Omega \in \mathscr{F}$;

(2) 若 $A \in \mathscr{F}$, 则 $\overline{A} \in \mathscr{F}$;

(3) 若 $A_n \in \mathscr{F}, n=1,2,\cdots$, 则 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{F}$,

则称 $\mathscr{F}$ 是 Ω 上的一个 σ-代数.

由定义 1.1.1 可知: Ω 上的一个 σ-代数是包含 Ω 且对余运算和可列并运算封闭的 Ω 的集合类.

显然, Ω 的一切子集构成的集合类是一 σ-代数, 记作 $\mathscr{F}_\Omega$. 若集合 A 是 Ω 的一个非空真子集, 则集合类 $\{\varnothing, \Omega, A, \overline{A}\}$ 是一 σ-代数. 不难看出, $\{\varnothing, \Omega\}$ 也是一 σ-代数, 一般称为平凡 σ-代数. $\mathscr{F}_\Omega$ 包含 Ω 的任意一个 σ-代数.

由 De Morgan 律, 易得下述结论.

定理 1.1.1 $\mathscr{F}$ 是 Ω 上的一个 σ-代数的充要条件是满足定义 1.1.1 中的 (1), (2) 及 (3)$'$ (若 $A_n \in \mathscr{F}, n=1,2,\cdots$, 则 $\bigcap\limits_{n=1}^{\infty} A_n \in \mathscr{F}$).

定理 1.1.2 Ω 的任意一族 σ-代数的交仍为 Ω 的 σ-代数.

定理 1.1.3 设 $\mathscr{C}$ 是由 Ω 的一些子集组成的非空集合类, 则存在一个唯一的 Ω 的 σ-代数 $\sigma(\mathscr{C})$, 它包含 $\mathscr{C}$ 而且被包含 $\mathscr{C}$ 的任一 σ-代数所包含. $\sigma(\mathscr{C})$ 称为由 $\mathscr{C}$ 生成的 σ-代数或包含 $\mathscr{C}$ 的最小 σ-代数.

证 记 $\mathscr{C}^* = \bigcap\limits_{\mathscr{C}\subset\mathscr{F}} \mathscr{F}$ 表示所有包含 $\mathscr{C}$ 的 σ-代数取交集. 由于 $\mathscr{F} \subset \mathscr{F}_\Omega$, 因而上述表示是有意义的.

由于每一个 $\mathscr{F}$ 都是 σ-代数, 因而 $\mathscr{C}^*$ 是 σ-代数; 由于对于每一个 $\mathscr{F}$, 都有 $\mathscr{F} \supset \mathscr{C}$, 因而 $\mathscr{C} \subset \bigcap\limits_{\mathscr{C}\subset\mathscr{F}} \mathscr{F} = \mathscr{C}^*$, 所以 $\mathscr{C}^*$ 是包含 $\mathscr{C}$ 的 σ-代数.

由于 $\mathscr{C}^*$ 是所有包含 $\mathscr{C}$ 的 σ-代数的交, 因而对任意包含 $\mathscr{C}$ 的 σ-代数 $\mathscr{F}$, 有 $\mathscr{C}^* \subset \mathscr{F}$, 即满足定理条件的 σ-代数存在且唯一, 并且 $\sigma(\mathscr{C}) = \mathscr{C}^*$.

证毕

例 1.1.1 设集合 $\Omega=\{1,2,3,4,5,6\}$, 则定义在 Ω 上的包含 $\{\{2,3\}\}$ 的最小 σ-代数为 $\{\Omega,\varnothing,\{2,3\},\{1,4,5,6\}\}$.

例 1.1.2 设 $\mathbb{R}^{(1)}$ 是一维实数空间, 则 $\mathbb{R}^{(1)}$ 上包含集合 $\{[0,+\infty)\}$ 的最小 σ-代数为 $\{\mathbb{R}^{(1)},\varnothing,(-\infty,0),[0,+\infty)\}$.

在研究随机变量的分布时, 一维实数空间 $\mathbb{R}^{(1)}$ 和 n 维实数空间 $\mathbb{R}^{(n)}$ 上的 Borel 域起着重要的作用.

设 $\Omega=\mathbb{R}^{(1)}$, 考虑由 $\mathbb{R}^{(1)}$ 的一些子集组成的集合类

$$\mathscr{G}=\{(-\infty,a]:a\in\mathbb{R}^{(1)}\},$$

称 $\sigma(\mathscr{G})$ 为 $\mathbb{R}^{(1)}$ 上的 Borel 域, 记作 $\mathscr{B}^{(1)}$, 并称 $\mathscr{B}^{(1)}$ 中的集合为一维 Borel 集.

注 1.1.1 $\mathscr{B}^{(1)}$ 包含一切开区间、闭区间、半开半闭区间、单个实数、可列个实数以及由它们经可列次"并"、"交"及"余"运算而得的集合.

设 $\Omega=\mathbb{R}^{(n)}=\{(x_1,x_2,\cdots,x_n):x_i\in\mathbb{R}^{(1)},i=1,2,\cdots,n\}$ 为 n 维实数空间. 考虑由 $\mathbb{R}^{(n)}$ 的一些子集组成的集合类

$$\mathscr{G}=\left\{\prod_{i=1}^{n}(-\infty,a_i]:a_i\in\mathbb{R}^{(1)},i=1,2,\cdots,n\right\},$$

称 $\sigma(\mathscr{G})$ 为 $\mathbb{R}^{(n)}$ 上的 Borel 域, 记作 $\mathscr{B}^{(n)}$, 并称 $\mathscr{B}^{(n)}$ 中的集合为 n 维 Borel 集, 其中

$$\prod_{i=1}^{n}(-\infty,a_i]=\{(x_1,x_2,\cdots,x_n):-\infty<x_i\leqslant a_i,i=1,2,\cdots,n\}.$$

注 1.1.2 与一维情形类似, $\mathscr{B}^{(n)}$ 中包含一切超矩形、单点集、可列个点以及由它们经可列次"并"、"交"及"余"运算而得的集合.

Borel 域的概念可以推广到所谓的广义实数空间 $\widetilde{\mathbb{R}}^{(1)}$ 和 $\widetilde{\mathbb{R}}^{(n)}$ 上, 其中

$$\widetilde{\mathbb{R}}^{(1)}=[-\infty,+\infty],$$

$$\widetilde{\mathbb{R}}^{(n)}=\left\{(x_1,x_2,\cdots,x_n):x_i\in\widetilde{\mathbb{R}}^{(1)},i=1,2,\cdots,n\right\}.$$

记 $\mathscr{G}=\{[a,b]:a,b\in\widetilde{\mathbb{R}}^{(n)}\}$, 其中 $a=(a_1,\cdots,a_n),b=(b_1,\cdots,b_n),a_i,b_i\in\widetilde{\mathbb{R}}^{(1)},i=1,\cdots,n,[a,b]=\{(x_1,x_2,\cdots,x_n):a_i\leqslant x_i\leqslant b_i,i=1,2,\cdots,n\}$, 则称在 $\Omega=\widetilde{\mathbb{R}}^{(n)}$ 上的包含 $\mathscr{G}$ 的最小 σ-代数 $\sigma(\mathscr{G})$ 为 $\widetilde{\mathbb{R}}^{(n)}$ 上的广义 Borel 域, 记作 $\widetilde{\mathscr{B}}^{(n)}$.

定义 1.1.2 设 Ω 是一非空集合, $\mathscr{F}$ 是 Ω 上的 σ-代数, 称 $(\Omega,\mathscr{F})$ 为可测空间; 若 $A\in\mathscr{F}$, 则称 A 为 $\mathscr{F}$ 可测集或可测集.

1.2 概 率 空 间

1.2.1 概率的公理化定义

古典概型与几何概型中的概率具有共同的性质：非负性、规范性、有限可加性（对古典概型）或可列可加性（对几何概型）．同时, 对于一般的随机试验, 频率也有类似的性质. 由此使人联想到, 可以用这些性质作为一般的概率的定义. 近代概率论的结构正是遵循这种思路建立起来的. 1933 年, 苏联数学家柯尔莫哥洛夫综合了前人的研究成果, 提出了概率论公理化结构, 使概率论成为严谨的数学分支, 这对近几十年概率论的迅速发展起了重要作用.

概率的公理化定义包括三个要素：样本空间、事件域和概率, 下面分别给出定义.

定义 1.2.1 随机试验的可能结果, 称为**样本点**或**基本事件**, 记为 ω, 样本点全体构成**样本空间**, 记为 Ω, 即 $\Omega=\{\omega\}$.

事件 A 一般可以理解为 Ω 的一个子集, 即 $A\subset\Omega$, 它包括若干样本点, 事件 A 发生当且仅当 A 所包含的样本点中有一个发生.

如果把映射 $A\to P(A)$ 看成集合 A 的函数, 那么像普通函数一样, 应当考虑 A 在什么范围内时, $P(A)$ 有定义. 同时为满足分析随机现象的实际需要, 应允许对事件进行必要的运算, 因而事件类不能太小, 至少对某些运算应该是封闭的. 另外, 为了给出每个事件的概率, 并保证对概率有一定的要求 (如可加性等), 事件类不能太大, 否则无法给出一个“兼顾”各方面要求的概率. 这就是运用公理化方法来描述概率模型时所必须考虑的内容.

定义 1.2.2 设 $\mathscr{F}$ 是由样本空间 Ω 的一些子集构成的一个 σ 代数, 则称 $\mathscr{F}$ 为**事件域**, $\mathscr{F}$ 中的元素（集合）称为**事件**, 并称 Ω 为**必然事件**, $\varnothing$ 为**不可能事件**.

此时, 事件的运算与集合的运算完全相同.

在概率公理化结构中, 对于事件域 $\mathscr{F}$ 中的每个元素 A, 都有一个实数 $P(A)$ 与之对应. 这种从集合到实数的映射（记为 P）称为集合函数. 因此, 概率是定义在事件域 $\mathscr{F}$ 上的一个集合函数.

定义 1.2.3 设 P 为定义在事件域 $\mathscr{F}$ 上的一个集合函数, 如果它满足下列三个条件：

(1) 非负性： $\forall A\in\mathscr{F}, P(A)\geqslant 0$;

(2) 规范性： $P(\Omega)=1$;

(3) 可列可加性： 若 $A_i\in\mathscr{F}, i=1,2,\cdots$, 且两两互不相容, 则

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i),$$

则称 P 为**概率测度**, 简称**概率**. 对于 $A\in\mathscr{F}$, 称 $P(A)$ 为**事件 A 的概率**.

定义 1.2.4 三元总体 $(\Omega,\mathscr{F},P)$ 称为**概率空间**, 其中 Ω 为样本空间, $\mathscr{F}$ 是事件域, P 为概率.

在概率论中, 概率空间 $(\Omega,\mathscr{F},P)$ 通常被认为是预先给定的, 我们可以此为出发点讨论各种问题. 至于实际问题中如何选定 Ω、怎样构造 $\mathscr{F}$ 以及怎样给定 P, 则要视具体情况而定.

1.2.2 概率的性质

为了使用方便, 下面把概率的性质列举如下, 略去证明.

定理 1.2.1 设 $(\Omega,\mathscr{F},P)$ 为一概率空间, 则 P 具有下列性质:

(1) $P(\varnothing)=0$;

(2) 有限可加性: 若 $A_i\in\mathscr{F},i=1,2,\cdots$, 且 $A_iA_j=\varnothing(i\neq j)$, 则

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i);$$

(3) $\forall A\in\mathscr{F},P(\overline{A})=1-P(A)$;

(4) 减法及单调性: 如果 $B\subset A$, 则 $P(A-B)=P(A)-P(B)$, $P(B)\leqslant P(A)$;

(5) 一般加法公式与次可加性: 设 A,B 为任意事件, 则

$$P(A\cup B)=P(A)+P(B)-P(AB)\text{(一般加法公式)},$$

$$P(A\cup B)\leqslant P(A)+P(B)\text{(次可加性)};$$

(6) 上、下连续性: 若 $\forall A_n\in\mathscr{F},n=1,2,\cdots$, 且 $A_n\downarrow$, 则

$$\lim_{n\to\infty}P(A_n)=P(\lim_{n\to\infty}A_n)=P\left(\bigcap_{n=1}^{\infty}A_n\right), \tag{1.2.1}$$

若 $\forall A_n\in\mathscr{F},n=1,2,\cdots$, 且 $A_n\uparrow$, 则

$$\lim_{n\to\infty}P(A_n)=P(\lim_{n\to\infty}A_n)=P\left(\bigcup_{n=1}^{\infty}A_n\right). \tag{1.2.2}$$

注 1.2.1 (1) $\forall A$, 因 $A\subset\Omega$, 故 $P(A)\leqslant P(\Omega)=1$.

(2) 一般减法公式: 因 $A-B=A-AB$, $AB\subset A$, 故有

$$P(A-B)=P(A-AB)=P(A)-P(AB).$$

(3) 利用数学归纳法可证得: 对任意自然数 n, 有

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)-\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n)\text{(一般加法公式)},$$

$$P\left(\bigcup_{i=1}^{n}A_i\right)\leqslant\sum_{i=1}^{n}P(A_i)\text{(次可加性)}.$$

从定理 1.2.1 中的 (2) 知, 由可列可加性可以推出有限可加性. 但是一般说来, 由有限可加性并不能推出可列可加性. 那么, 究竟施加什么条件可使逆命题成立呢? 我们有下述定理.

定理 1.2.2 若 $(\Omega, \mathscr{F}, P)$ 为概率空间, 则 P 所具有的可列可加性与 P 同时满足下面的 (1) 和 (2) 是等价的:

(1) 有限可加性;

(2) 在 $\varnothing$ 处上连续.

1.2.3 条件概率空间与事件的独立性

本节在概率空间上讨论问题, 首先给出条件概率空间的概念.

1. 条件概率空间

定理 1.2.3 设 $(\Omega, \mathscr{F}, P)$ 为一已知概率空间, $B \in \mathscr{F}, P(B) > 0$. 若定义 $\mathscr{F}$ 上的集合函数

$$P(A|B) = \frac{P(AB)}{P(B)}, \quad A \in \mathscr{F},$$

则 $P(A|B)$ 是 $\mathscr{F}$ 上的概率, 即 $P(A|B)$ 满足:

(1) $0 \leqslant P(A|B) \leqslant 1, A \in \mathscr{F}$;

(2) $P(\Omega|B) = 1$;

(3) 若 $A_i \in \mathscr{F}, i = 1, 2, \cdots, A_iA_j = \varnothing, i \neq j$, 有

$$P\left(\bigcup_{i=1}^{\infty} A_i|B\right) = \sum_{i=1}^{\infty} P(A_i|B).$$

证明从略.

定义 1.2.5 设 $(\Omega, \mathscr{F}, P)$ 为一已知概率空间, $B \in \mathscr{F}, P(B) > 0$. 记

$$P_B(A) = P(A|B) = \frac{P(AB)}{P(B)}, \quad A \in \mathscr{F},$$

称 $(\Omega, \mathscr{F}, P_B)$ 为 $(\Omega, \mathscr{F}, P)$ 在条件 B 下的条件概率空间.

条件概率空间有下面的性质.

定理 1.2.4 设 $(\Omega, \mathscr{F}, P_B)$ 为 $(\Omega, \mathscr{F}, P)$ 在条件 B 下的条件概率空间. 若记 $\nu(A) = P_B(A), A \in \mathscr{F}$, 又 $C \in \mathscr{F}$, 且 $\nu(C) > 0, (\Omega, \mathscr{F}, \nu)$ 在条件 C 下的条件概率空间为 $(\Omega, \mathscr{F}, \nu_c)$, 则有

$$\nu_c(A) = P_{BC}(A) = P(A|BC), \quad A \in \mathscr{F}.$$

证明从略.

以下假定所涉及的事件类都是在同一概率空间 $(\Omega, \mathscr{F}, P)$ 上的.

2. 事件的独立性

定义 1.2.6 设 $A_1, \cdots, A_n$ 是 n 个事件. 若对任意 $1 \leqslant m \leqslant n$, 任意的 $1 \leqslant k_1 < k_2 < \cdots < k_m \leqslant n$, 都有

$$P\left(\bigcap_{j=1}^{m} A_{k_j}\right) = \prod_{j=1}^{m} P(A_{k_j})(\text{共}2^n - 1 - n\text{个等式}), \tag{1.2.3}$$

则称 $A_1, \cdots, A_n$ 独立.

显然, 若 $A_1, \cdots, A_n$ 独立, 则其中任意 $m(m < n)$ 个也独立, 特别地, 其中任意两个也独立, 但反之不然.

定理 1.2.5 假设 $A_1, \cdots, A_n$ 独立, 若将其中的 $m(1 \leqslant m \leqslant n)$ 个事件换成各自的对立事件, 则所构成的 n 个事件也独立.

证 由对称性和归纳法只需证明: 若 $A_1, \cdots, A_n$ 独立, 则 $\overline{A}_1, A_2, \cdots, A_n$ 也独立. 由于 $A_1, \cdots, A_n$ 独立, 式（1.2.3）中的 $2^n - 1 - n$ 个等式成立, 其中不包含 A_1 的 $2^{n-1} - n$ 个等式仍成立, 因此，只需证明将其他 $2^{n-1} - 1$ 个包含 A_1 的等式中的 A_1 换成 $\overline{A}_1$ 也成立, 这样便证明了 $\overline{A}_1, A_2, \cdots, A_n$ 独立.

事实上, 由概率的可减性和 $A_1, \cdots, A_n$ 独立可得,

$$\begin{aligned} P(\overline{A}_1 A_{j_1} \cdots A_{j_m}) &= P(A_{j_1} \cdots A_{j_m}) - P(A_1 A_{j_1} \cdots A_{j_m}) \\ &= [1 - P(A_1)]P(A_{j_1}) \cdots P(A_{j_m}) = P(\overline{A}_1)P(A_{j_1}) \cdots P(A_{j_m}), \end{aligned}$$

其中 $1 \leqslant m \leqslant n-1, 2 \leqslant j_1 < j_2 < \cdots < j_m \leqslant n$.

证毕

定义 1.2.7 设 $\{A_t, t \in T\}$ 是 $\mathscr{F}$ 中的事件类. 若对 T 的任一有限子集 $\{t_1, \cdots, t_n\}$ $(n \geqslant 2)$, 有

$$P\left(\bigcap_{j=1}^{n} A_{t_j}\right) = \prod_{j=1}^{n} P(A_{t_j}),$$

则称 $A_t, t \in T$ 独立.

事件独立性的概念可以推广到事件类的独立性, 因为需要用到测度论的知识, 这里不做介绍.

1.3 习　　题

1. 证明：$P(A_i) = 0, i = 1, 2, \cdots, n \Longleftrightarrow P\left(\bigcup_{i=1}^{n} A_i\right) = 0$.

2. 证明: $P(B_i) = 1, i = 1, 2, \cdots, n \Longleftrightarrow P\left(\bigcap_{i=1}^{n} B_i\right) = 1$.

3. 证明：包含一切形如 $(-\infty, x)$ 的区间的最小 σ-代数是一维 Borel 域.

4. 求包含集合 A, B 的最小 σ-代数, 其中 $A, B \neq \varnothing$ 或 Ω, $AB \neq \varnothing$, $A \bigcup B \neq \varnothing$, 而且 A, B 互不包含.

5. 若 $\mathscr{G} = \{A_k : A_k \subset \Omega, k = 1, 2, \cdots, \text{且两两不相交}\}$, 试求 $\sigma(\mathscr{G})$.

6. 设 Ω 是一不可列集, $\mathscr{A}$ 是由 Ω 的一切有限子集、可列子集及以有限子集或可列子集为余集的子集所组成的集合类, 试证 $\mathscr{A}$ 是一 σ-代数.

7. 设 $\mathscr{G}$ 是由 Ω 的子集组成的集合类, A 是 Ω 的一个非空子集, 令

$$\mathscr{G} \bigcap A = \left\{A' : A' = B \bigcap A, B \in \mathscr{G}\right\},$$

试证：$\mathscr{G} \bigcap A$ 是以 A 为空间的包含集合类 $\mathscr{G} \bigcap A$ 的最小 σ-代数.

8. 设 A 是 Ω 的一个子集, $\mathscr{G}$ 是由 Ω 的包含 A 的一切子集所组成的集合类, 试问 $\sigma(\mathscr{G})$ 是由哪些子集组成的?

9. 证明下列命题等价:

（1）事件 A 与任一事件 B 独立;

（2）A 与 A 独立;

（3）$P(A) = 1$ 或 0.

第 2 章 可测函数、随机变量及其分布

2.1 可 测 函 数

为了方便讨论, 先给出映射及其逆象的概念和性质, 然后再讨论可测函数的性质.

2.1.1 映射及其逆象

定义 2.1.1 设 Ω, R 为两个非空集合, 若对每一个 $\omega \in \Omega$, 有一个 R 中元素 $x = f(\omega)$ 与之对应, 则称这种对应关系为由 Ω 到 R 上的映射, 记作 $x = f(\omega)$, 并称 f 是由 Ω 到 R 上的映射.

若 f 是由 Ω 到 R 上的映射, B 是 R 中的任一子集, 则称 Ω 的子集 $\{\omega : f(\omega) \in B\}$ 为 B 在 f 下的逆象, 记作 $f^{-1}(B)$.

若 $\mathscr{G}$ 是由 R 的子集组成的集合类, 则称由 Ω 的子集 $f^{-1}(B), B \in \mathscr{G}$ 组成的集合类 $\{f^{-1}(B) : B \in \mathscr{G}\}$ 为 $\mathscr{G}$ 在 f 下的逆象, 记作 $f^{-1}(\mathscr{G})$.

逆象具有下列性质:

(1) $f^{-1}(R) = \Omega, f^{-1}(\varnothing) = \varnothing$.

(2) $f^{-1}(\overline{B}) = \overline{f^{-1}(B)}, B \in R$.

(3) $f^{-1}(B_1 \backslash B_2) = f^{-1}(B_1) \backslash f^{-1}(B_2), B_1, B_2 \subset R$.

(4) $f^{-1}\left(\bigcup\limits_{t \in T} B_t\right) = \bigcup\limits_{t \in T} f^{-1}(B_t), B_t \subset R, t \in T$ (T是任一指标集).

(5) $f^{-1}\left(\bigcap\limits_{t \in T} B_t\right) = \bigcap\limits_{t \in T} f^{-1}(B_t), B_t \subset R, t \in T$ (T是任一指标集).

证明从略.

关于映射有下面的重要结论.

定理 2.1.1 设 f 是由 Ω 到 R 上的映射,

(1) 如果 $\mathscr{B}$ 是 R 上的 σ-代数, 则 $f^{-1}(\mathscr{B})$ 是 Ω 上的 σ-代数;

(2) 如果 $\mathscr{G}$ 是 R 上的任一非空集合类, 则有

$$f^{-1}(\sigma(\mathscr{G})) = \sigma(f^{-1}(\mathscr{G})). \tag{2.1.1}$$

证 (1) 利用逆象的性质以及 $\mathscr{B}$ 是 R 上的 σ-代数便可得到 $f^{-1}(\mathscr{B})$ 是 Ω 上的 σ-代数.

(2) 由 (1) 知 $f^{-1}(\sigma(\mathscr{G}))$ 是 Ω 上的 σ-代数, 且由 $\mathscr{G} \subset \sigma(\mathscr{G})$, 有 $f^{-1}(\mathscr{G}) \subset f^{-1}(\sigma(\mathscr{G}))$, 于是

$$\sigma(f^{-1}(\mathscr{G})) \subset f^{-1}(\sigma(\mathscr{G})).$$

为证（2），只需证明相反的包含关系. 令

$$\mathscr{C}=\{C:C\subset R, f^{-1}(C)\in\sigma(f^{-1}(\mathscr{G}))\},\tag{2.1.2}$$

只需证明 $\mathscr{C}$ 是包含 $\mathscr{G}$ 的 σ-代数, 即有 $\sigma(\mathscr{G})\subset\mathscr{C}$, 但 $f^{-1}(\mathscr{C})\subset\sigma(f^{-1}(\mathscr{C}))$, 所以 $f^{-1}(\sigma(\mathscr{G}))\subset\sigma(f^{-1}(\mathscr{G}))$.

由式 (2.1.2) 知, $f^{-1}(\mathscr{G})\subset\sigma(f^{-1}(\mathscr{G}))$, 所以 $\mathscr{G}\subset\mathscr{C}$. 由于 $\Omega\in\sigma(f^{-1}(\mathscr{G}))$, 而 $f^{-1}(R)=\Omega$, 所以 $R\in\mathscr{C}$. 若 $C\in\mathscr{C}$, $f^{-1}(C)\in\sigma(f^{-1}(\mathscr{G}))$, 那么由于 $\sigma(f^{-1}(\mathscr{G}))$ 是 σ-代数, 有 $f^{-1}(\overline{C})=\overline{f^{-1}(C)}\in\sigma(f^{-1}(\mathscr{G}))$, 所以 $\overline{C}\in\mathscr{C}$. 若 $C_k\in\mathscr{C}, k=1,2,\cdots$, 则 $f^{-1}\left(\bigcup\limits_{k=1}^{\infty}C_k\right)=\bigcup\limits_{k=1}^{\infty}f^{-1}(C_k)\in\sigma(f^{-1}(\mathscr{G}))$, 因而 $\mathscr{C}$ 是一 σ-代数, 从而 $f^{-1}(\sigma(\mathscr{G}))\subset\sigma(f^{-1}(\mathscr{G}))$.

证毕

定义 2.1.2　设 $(\Omega,\mathscr{F}),(R,\mathscr{B})$ 为可测空间, f 是由 Ω 到 R 上的映射, 若对每一个 $B\in\mathscr{B}$, 有 $f^{-1}(B)\in\mathscr{F}$, 则称 f 是从 $(\Omega,\mathscr{F})$ 到 $(R,\mathscr{B})$ 上的可测映射.

2.1.2　可测函数的定义及判别

将可测映射的概念具体化, 可得到可测函数的相应概念和性质.

定义 2.1.3　若 f 是由 $(\Omega,\mathscr{F})$ 到 $(\widetilde{\mathbb{R}}^{(1)},\widetilde{\mathscr{B}}^{(1)})$ 的可测映射, 则称 f 为 $(\Omega,\mathscr{F})$ 上的实可测函数; 若 f 是从 $(\Omega,\mathscr{F})$ 到 $(\widetilde{\mathbb{R}}^{(n)},\widetilde{\mathscr{B}}^{(n)})$ 上的可测映射, 则称 f 为 $(\Omega,\mathscr{F})$ 上的 n 维实可测函数.

若 f_1,f_2 为 $(\Omega,\mathscr{F})$ 上的实可测函数, 则称 $f(\omega)=f_1(\omega)+\mathrm{i}f_2(\omega)$ 是 $(\Omega,\mathscr{F})$ 上的复可测函数.

关于可测函数有下面的结论.

定理 2.1.2　(1) f 是 $(\Omega,\mathscr{F})$ 上的实可测函数的充要条件是对一切 $x\in\widetilde{\mathbb{R}}^{(1)}$,

$$\{\omega:f(\omega)\leqslant x\}\in\mathscr{F};\tag{2.1.3}$$

(2) $f=(f_1,f_2,\cdots,f_n)$ 是 $(\Omega,\mathscr{F})$ 上的 n 维实可测函数的充要条件是对每一个 $k=1,2,\cdots,n$, f_k 是 $(\Omega,\mathscr{F})$ 上的实可测函数.

证　(1) 必要性显然. 为证充分性, 令

$$\mathscr{G}=\{(-\infty,x]:x\in\widetilde{\mathbb{R}}^{(1)}\},$$

则 $\mathscr{G}$ 是 $\widetilde{\mathbb{R}}^{(1)}$ 上的集合类, 而且 $\sigma(\mathscr{G})=\widetilde{\mathscr{B}}^{(1)}$. 于是由式 (2.1.3) 有

$$f^{-1}(\mathscr{G})\subset\mathscr{F}.$$

而 $\mathscr{F}$ 是一 σ-代数, 因而 $\sigma(f^{-1}(\mathscr{G}))\subset\mathscr{F}$. 由式 (2.1.1) 有

$$f^{-1}(\widetilde{\mathscr{B}}^{(1)})=f^{-1}(\sigma(\mathscr{G}))=\sigma(f^{-1}(\mathscr{G}))\subset\mathscr{F}.$$