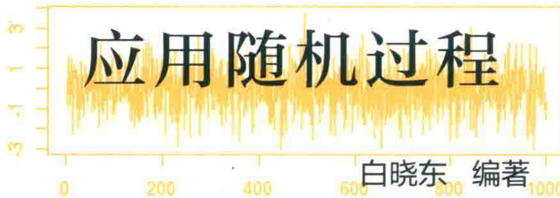
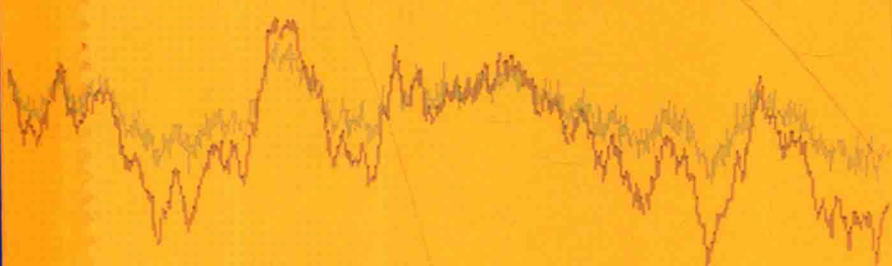
普通高等教育“十三五”规划教材

应用随机过程



白晓东 编著



清华大学出版社

应用随机过程

白晓东 编著

清华大学出版社
北京

内容简介

本书主要介绍了现代应用随机过程理论中部分经典的理论，主要内容包括预备知识、随机过程的基本概念、泊松过程、布朗运动、条件数学期望与鞅、更新过程、马尔可夫链、随机积分与随机微分方程以及它们在破产理论和金融衍生产品定价方面的应用.

本书选材精简实用，内容安排得当，论述简洁明澈，语言平易流畅，具有很好的可读性. 此外，每小节之后都配有精选的习题，便于掌握和巩固知识.

本书可以作为高等院校统计、经济、金融、管理以及理工科各相关专业的高年级本科生学习随机过程的教材或教学参考书，也可作为有关专业硕士研究生的教材和教学参考书，对广大从事与随机现象相关工作的科技工作者也具有参考价值.

图书在版编目(CIP)数据

应用随机过程/白晓东编著.—北京：清华大学出版社，2018

ISBN 978-7-302-50734-5

Ⅰ. ①应…　Ⅱ. ①白…　Ⅲ. ①随机过程　Ⅳ. ①O211.6

中国版本图书馆 CIP 数据核字(2018)第 172055 号

责任编辑：刘　颖
封面设计：傅瑞学
责任校对：刘玉霞
责任印制：刘海龙

出版发行：清华大学出版社
　网　　址：http://www.tup.com.cn，http://www.wqbook.com
　地　　址：北京清华大学学研大厦 A 座　　邮　　编：100084
　社 总 机：010-62770175　　邮　　购：010-62786544
　投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
　质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：三河市吉祥印务有限公司
经　　销：全国新华书店
开　　本：170mm×240mm　　印　张：12.75　　字　数：302 千字
版　　次：2018 年 9 月第 1 版　　印　次：2018 年 9 月第 1 次印刷
定　　价：32.00 元

产品编号：076487-01

前　言

随机过程是一族随机事件动态关系的定量描述. 作为概率论的延伸和发展, 随机过程论与数学、物理等许多学科的分支有着密切的联系, 并已广泛应用于物理、化学、生物、气象、天文、经济、金融、运筹决策、安全科学、人口理论、可靠性及计算机科学等诸多领域, 已经成为自然科学、工程科学和社会科学各领域研究随机现象的重要工具.

目前, 国内外有关随机过程的教材已有很多, 其中一些需要读者具备实变函数、泛函分析和高等概率论的基础知识, 其主要阅读对象是数学系概率论与数理统计专业和精英型的统计学专业的学生; 另一些则只需读者具备高等数学、线性代数和初等概率论的知识, 其主要阅读对象是侧重应用的统计学类各专业和其他理工科及经管类相关专业的学生. 随着我国招生制度的变化, 大部分高校的统计学及其相关专业的培养目标逐步转为复合应用型人才的培养. 在人才培养过程中, 既强调具备一定的理论基础, 又强调应用能力的提高. 显然那些需要过多数学基础, 讲述过于抽象的教材和重点放在直观理解、叙述不严谨的教材都已经不能适应这一变化.

为适应培养要求的转变, 满足更多专业学生的学习需求, 本书在借鉴国内外相关优秀教材的基础上, 着重突出三个特色. 第一是内容精简实用. 本书在保留了几类常用的经典随机过程的基础上, 力求讲清楚最基本的概念、性质和方法, 去掉过于专门化的讨论, 使得绝大多数内容能够在课程规定的学时内完成. 第二是内容的安排上, 尽量在保持理论体系的同时将难点分散; 同时, 兼顾不同专业讲授时选题的需要. 例如, 我们在 1.2 节讲完随机过程的基本概念之后, 在 1.3 节和 1.4 节就安排了两类重要的平稳独立增量过程 —— 泊松过程和布朗运动. 目的是, 一方面让学生尽快感受具体随机过程的研究方法; 另一方面将难点分散, 便于学生掌握. 又如, 我们将条件数学期望和鞅论合在一起放在第 2 章, 这是考虑到, 一方面大多数初等概率论对条件数学期望的讲解偏少, 甚至忽略, 需要专门的篇幅讲解; 另一方面这两块知识联系紧密, 便于学生迅速将条件数学期望的知识应用于鞅论, 从而更好地掌握这两部分知识. 再如, 我们在第 6 章专门安排了随机过程在保险和金融中的两个应用, 便于需要这部分知识的专业去选题. 其实, 6.1 节破产理论, 在讲完第 3 章更新过程之后就可以讲授了. 第三是力求增

强可读性. 教材的可读性体现在, 要有严谨的理论, 准确的语言, 明晰的脉络, 简练、生动而明达的表述, 不能让读者感到冗繁与艰涩. 诚然, 上述种种是本教材努力追求的.

本书主要包括以下内容. 第 1 章在总结复习概率论相关知识的基础上, 介绍随机过程的基本概念, 同时较为系统地介绍两类重要的平稳独立增量过程——泊松过程和布朗运动. 第 2 章分别介绍条件数学期望和鞅论的知识以及它们简单的应用. 第 3 章介绍更新过程的概念、性质、更新方程、更新定理以及更新过程的几个推广. 第 4 章介绍离散时间马尔可夫链的概念、性质、状态的分类、状态空间的分解、极限定理、平稳分布以及连续时间马尔可夫链的概念和基本性质. 第 5 章介绍伊藤积分的概念和性质、伊藤积分过程、伊藤公式以及随机微分方程. 第 6 章简要地介绍随机过程在破产理论和金融衍生产品定价方面的应用. 学习本书的内容, 只需要具备高等数学、线性代数和初等概率论的基础知识就可以了. 此外, 本书在每小节之后还配备了一定数量的习题. 目的是通过这些习题的演练, 希望读者尽快掌握相应章节的基本理论和方法.

本书可以作为高等院校统计、经济、金融、管理以及理工科各相关专业的高年级本科生学习随机过程的教材或教学参考书, 也可作为有关专业硕士研究生的教材和教学参考书, 对广大从事与随机现象相关工作的科技工作者也具有参考价值.

本书在写作过程中参考了国内外许多优秀的教材和论著, 在此向他们表示感谢和敬意. 本书能够及时出版, 还要感谢清华大学出版社刘颖编审的大力支持和帮助. 本书内容在大连民族大学统计学专业、数学与应用数学专业以及信息与计算科学专业讲授多次, 感谢同学们对课程内容的浓厚兴趣和热烈讨论, 同时纠正了一些打印错误.

白晓东

baixd_dlnu@163.com

2018 年 5 月

目　　录

第 1 章　引论 1

1.1　预备知识 1

1.1.1　概率空间 1

1.1.2　随机变量 5

1.1.3　黎曼-斯蒂尔切斯积分 10

1.1.4　数字特征 13

1.1.5　矩母函数、特征函数 15

1.1.6　几个重要的极限定理 17

习题 1.1 19

1.2　随机过程的基本概念 19

1.2.1　随机过程的定义 19

1.2.2　随机过程的有限维分布族和数字特征 21

1.2.3　平稳过程 23

1.2.4　独立增量过程 23

习题 1.2 24

1.3　泊松过程 25

1.3.1　泊松过程的概念 25

1.3.2　指数流与泊松过程 29

1.3.3　指数流的条件分布 33

1.3.4　剩余寿命与年龄 35

1.3.5　非时齐泊松过程 37

习题 1.3 40

1.4 布朗运动 42
1.4.1 布朗运动的概念 42
1.4.2 布朗运动轨道的性质 45
1.4.3 首中时 47
1.4.4 布朗运动的几种变化 48
习题 1.4 52

第 2 章 条件数学期望与鞅 53
2.1 条件数学期望的概念 53
2.1.1 离散型随机变量的条件数学期望 53
2.1.2 连续型随机变量的条件数学期望 54
2.1.3 一般随机变量的条件数学期望 56
习题 2.1 57
2.2 条件数学期望的基本性质与应用 58
2.2.1 条件数学期望的基本性质 58
2.2.2 复合泊松过程 59
2.2.3 条件泊松过程 60
2.2.4 反正弦律 62
2.2.5 其他例子 63
习题 2.2 64
2.3 鞅的基本概念 65
2.3.1 鞅的概念与举例 66
2.3.2 上鞅与下鞅 70
2.3.3 鞅的分解定理 72
2.3.4 关于鞅的两个不等式 74
习题 2.3 75
2.4 停时与停时定理 76
2.4.1 停时的概念 77
2.4.2 停时定理 78
*2.4.3 停时定理的补充 82
习题 2.4 83

2.5 鞅收敛定理 85
2.5.1 上穿不等式 85
2.5.2 鞅收敛定理 86
习题 2.5 88
2.6 连续鞅初步 89
习题 2.6 91

第 3 章 更新过程 92
3.1 更新过程的概念 92
3.1.1 更新过程的定义 92
3.1.2 更新次数的极限 93
3.1.3 卷积及其性质 95
3.1.4 更新函数及其基本性质 96
习题 3.1 97
3.2 更新方程和更新定理 98
3.2.1 更新方程及其基本性质 98
3.2.2 更新定理 102
习题 3.2 106
3.3 更新过程的推广 107
3.3.1 交替更新过程 107
3.3.2 延迟更新过程 109
3.3.3 更新回报过程 109
习题 3.3 111

第 4 章 马尔可夫链 113
4.1 马尔可夫链及其转移概率 113
4.1.1 基本概念 113
4.1.2 查普曼-柯尔莫戈洛夫方程 115
习题 4.1 118
4.2 状态的分类及其性质 119
4.2.1 互通 119

4.2.2 常返与非常返状态 120
4.2.3 正常返和零常返状态 124
4.2.4 周期与遍历状态 125
习题 4.2......127
4.3 状态空间的分解 129
4.3.1 闭集 129
4.3.2 分解定理 130
习题 4.3......132
4.4 极限定理与平稳分布 135
4.4.1 极限定理 135
4.4.2 平稳分布 136
习题 4.4......140
4.5 连续时间马尔可夫链 142
4.5.1 概念和基本性质 142
4.5.2 转移概率的性质 143
4.5.3 柯尔莫戈洛夫向前—向后微分方程 145
习题 4.5......148

第 5 章 随机积分与随机微分方程......149
5.1 伊藤积分的定义 149
5.1.1 简单过程的伊藤积分 149
5.1.2 适应过程的伊藤积分 152
习题 5.1......155
5.2 伊藤积分过程 155
5.2.1 伊藤积分的鞅性 155
5.2.2 伊藤积分的二次变差和协变差 156
5.2.3 伊藤积分与高斯过程 158
习题 5.2......159
5.3 伊藤公式 159
5.3.1 关于布朗运动的伊藤公式 159

5.3.2 伊藤过程与随机微分 162
5.3.3 关于伊藤过程的伊藤公式 165
习题 5.3 169
5.4 随机微分方程 171
5.4.1 随机微分方程的定义 171
5.4.2 随机指数和对数 173
5.4.3 线性随机微分方程的解 176
5.4.4 随机微分方程解的存在唯一性 177
习题 5.4 178

第 6 章 随机过程在金融保险中的应用举例 179
6.1 破产理论 179
6.1.1 风险过程与破产概率的相关概念 179
6.1.2 安全负荷与调节系数 182
6.1.3 破产概率的估计 184
习题 6.1 188
6.2 金融衍生产品的定价 188
6.2.1 金融术语和基本假定 189
6.2.2 定价方法 190
习题 6.2 192

参考文献 193

第1章 引　　论

学习目标与要求

1. 回顾概率论的基本概念以及相关性质.
2. 理解随机过程的概念、平稳性、独立增量性和平稳增量性.
3. 理解泊松过程的有关概念, 并掌握它的重要性质.
4. 理解布朗运动的有关概念, 并掌握它的重要性质.

1.1 预备知识

1.1.1 概率空间

在概率论中, 通常把按照一定的想法去做的事情称为试验. 一个试验, 若它的结果预先无法确定, 则称之为**随机试验**, 简称为**试验**. 试验的每个可能结果称为**样本点**, 样本点的集合称为**样本空间**. 在本书中, 用 Ω 表示样本空间, 用 ω 表示样本点, 于是

$$\Omega=\{\omega:\omega\ \text{是试验的样本点}\}.$$

样本空间 Ω 中的样本点 ω 也称为**基本事件**. 样本空间的子集, 也即由基本事件构成的集合称为**事件**, 通常用大写英文字母 A,B,C 等表示. 样本空间 Ω 称为**必然事件**, 空集 $\varnothing$ 称为**不可能事件**. 由样本空间 Ω 中的若干子集构成的集合称为 Ω 的**集类**, 用花写的字母 $\mathcal{A},\mathcal{B},\mathcal{C}$ 等表示. 显然, 样本空间 Ω 的集类就是由一些事件构成的集合.

在实际问题中, 人们通常不是对样本空间的所有子集都感兴趣, 而是只关心某些事件及其发生的可能性大小. 为了方便地在人们感兴趣的事件上定义概率, 我们引入如下概念.

定义 1.1 设 $\mathcal{F}$ 是样本空间 Ω 的集类. 如果满足:

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

(2) 若 $A \in \mathcal{F}$, 则 $A^{c} = \Omega - A \in \mathcal{F}$;

(3) 若 $A_n \in \mathcal{F}, n \in \mathbb{N}$, 则 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{F}$,

则称 $\mathcal{F}$ 为样本空间 Ω 的一个 **σ 域** (或 **σ 代数**). 将 $(\Omega, \mathcal{F})$ 称为**可测空间**. $\mathcal{F}$ 中的元素便是人们感兴趣的事件.

容易验证, 若 $\mathcal{F}$ 是样本空间 Ω 的一个 σ 域, 则有: (1) $\varnothing \in \mathcal{F}$; (2) 若 $A, B \in \mathcal{F}$, 则 $A - B \in \mathcal{F}$; (3) 若 $A_n \in \mathcal{F}, n \in \mathbb{N}$, 则 $\bigcap\limits_{n=1}^{\infty} A_n \in \mathcal{F}$.

定义 1.2 设 $\mathcal{A}$ 为样本空间 Ω 的集类, 称一切包含 $\mathcal{A}$ 的 σ 域的交集为**由 $\mathcal{A}$ 生成的 σ 域**, 或称为**包含 $\mathcal{A}$ 的最小 σ 域**, 记为 $\sigma(\mathcal{A})$.

例 1.1 容易看出, 样本空间 Ω 的集类 $\mathcal{A}_0 = \{\varnothing, \Omega\}$, $\mathcal{A}_1 = \{\varnothing, A, A^{c}, \Omega\}$ 和 $\mathcal{A}_2 = \{A: A \subseteq \Omega\}$ 都是 σ 域, 而集类 $\mathcal{C} = \{\varnothing, A, \Omega\}$ 不是 σ 域. 由 $\mathcal{C}$ 生成的 σ 域为 $\sigma(\mathcal{C}) = \{\varnothing, A, A^{c}, \Omega\}$.

定义 1.3 设 $\Omega = \mathbb{R}$, 集类 $\mathcal{A} = \{(-\infty, x) : x \in \mathbb{R}\}$, 则称由 $\mathcal{A}$ 生成的最小 σ 域 $\sigma(\mathcal{A})$ 为 $\mathbb{R}$ 上的 **Borel σ 域**, 记为 $\mathcal{B}(\mathbb{R})$. $\mathcal{B}(\mathbb{R})$ 中的元素称为 **Borel 集合**. 类似地, 可定义 $\mathbb{R}^n$ 上的 Borel σ 域 $\mathcal{B}(\mathbb{R}^n)$.

定义 1.4 设 $(\Omega, \mathcal{F})$ 是一个可测空间, $P(\cdot)$ 是一个定义在 $\mathcal{F}$ 上的集函数. 如果 $P(\cdot)$ 满足以下条件:

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

(2) 完全性: $P(\Omega) = 1$;

(3) 可列可加性: 对两两互不相容的事件 $A_i \in \mathcal{F}, i = 1, 2, \cdots$ (即当 $i \neq j$ 时, $A_i \cap A_j = \varnothing$), 有

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

那么称 P 是 $(\Omega, \mathcal{F})$ 上的一个**概率测度**, 简称**概率**. 称 $(\Omega, \mathcal{F}, P)$ 为**概率空间**. 称 $P(A)$ 为事件 A 的概率. 本书约定概率为零的事件的任何子集都属于 $\mathcal{F}$. 满足这样约定的概率空间通常称为**完备的概率空间**.

概率具有如下基本性质:

(1) $P(\varnothing) = 0, P(A^{c}) = 1 - P(A), \forall A \in \mathcal{F}$.

(2) 有限可加性: 如果 $A_i \in \mathcal{F}, i = 1, 2, \cdots, n$, 且两两互不相容, 那么

$$P\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} P(A_i).$$

(3) 单调性: 如果 $A, B \in \mathcal{F}$, 且 $A \subseteq B$, 那么 $P(A) \leqslant P(B)$.

(4) 进出公式: 如果 $A_i \in \mathcal{F}, i=1,2,\cdots,n$, 那么

$$P\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} P(A_i) - \sum_{1\leqslant i<j\leqslant n} P(A_iA_j) + \sum_{1\leqslant i<j<k\leqslant n} P(A_iA_jA_k) - \cdots + (-1)^{n+1}P(A_1A_2\cdots A_n).$$

(5) 次可加性: 如果 $A_i \in \mathcal{F}, i=1,2,\cdots,n$, 那么

$$P\Big(\bigcup_{i=1}^{n} A_i\Big) \leqslant \sum_{i=1}^{n} P(A_i).$$

下面介绍概率的连续性. 为此引入事件序列的极限.

定义 1.5 若一个事件序列 $\{A_n, n\geqslant 1\} \subseteq \mathcal{F}$ 满足 $A_n \subseteq A_{n+1}$ (或 $A_n \supseteq A_{n+1}$), $n \geqslant 1$, 则称该事件序列 $\{A_n, n\geqslant 1\}$ 为**单调递增事件序列**(或**单调递减事件序列**). 若 $\{A_n, n\geqslant 1\}$ 是一个单调递增事件序列, 则定义 $\lim\limits_{n\to\infty} A_n = \bigcup\limits_{n=1}^{\infty} A_n$. 若 $\{A_n, n\geqslant 1\}$ 是一个单调递减事件序列, 则定义 $\lim\limits_{n\to\infty} A_n = \bigcap\limits_{n=1}^{\infty} A_n$. 若 $\{A_n, n\geqslant 1\}$ 是一个事件序列, 则定义 $\limsup\limits_{n\to\infty} A_n = \bigcap\limits_{n=1}^{\infty}\bigcup\limits_{i=n}^{\infty} A_i$.

(6) 概率的连续性: 如果$\{A_n, n\geqslant 1\}$是一个单调递增(或单调递减)事件序列, 那么

$$\lim_{n\to\infty} P(A_n) = P\big(\lim_{n\to\infty} A_n\big).$$

证明 设 $\{A_n, n\geqslant 1\}$ 是一个单调递增事件序列, 并令

$$B_1 = A_1, \quad B_n = A_n - A_{n-1}, \quad n \geqslant 2.$$

显然 $\{B_n, n\geqslant 1\}$ 两两互不相容, 而且对于每个 $n \geqslant 1$, 有 $\bigcup\limits_{k=1}^{n} A_k = \bigcup\limits_{k=1}^{n} B_k$, 从而 $\bigcup\limits_{n=1}^{\infty} A_n = \bigcup\limits_{n=1}^{\infty} B_n$. 于是得到

$$\begin{aligned}P\big(\lim_{n\to\infty} A_n\big) &= P\Big(\bigcup_{n=1}^{\infty} A_n\Big) = P\Big(\bigcup_{n=1}^{\infty} B_n\Big) = \sum_{n=1}^{\infty} P(B_n)\\ &= \lim_{n\to\infty}\sum_{k=1}^{n} P(B_k) = \lim_{n\to\infty} P\Big(\bigcup_{k=1}^{n} B_k\Big)\\ &= \lim_{n\to\infty} P\Big(\bigcup_{k=1}^{n} A_k\Big) = \lim_{n\to\infty} P(A_n).\end{aligned}$$

另一种情况的证明, 留给读者练习.　　证毕

(7) Borel-Cantelli 引理: 如果 $\{A_n, n \geqslant 1\}$ 是一个事件序列, 且满足 $\sum\limits_{n=1}^{\infty} P(A_n) < \infty$, 那么

$$P\Big(\limsup_{n\to\infty} A_n\Big) = 0.$$

证明　显然 $\bigcup\limits_{k=n}^{\infty} A_k (n = 1, 2, \cdots)$, 是关于 n 的单调递减事件序列. 根据性质 (5) 和性质 (6) 得

$$0 \leqslant P\Big(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\Big) = P\Big(\lim_{n\to\infty} \bigcup_{k=n}^{\infty} A_k\Big) = \lim_{n\to\infty} P\Big(\bigcup_{k=n}^{\infty} A_k\Big) \leqslant \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0.$$

证毕

事件的独立性是事件间的一种重要关系. 下面介绍事件的独立性概念.

定义 1.6　称两个事件 $A, B \in \mathcal{F}$ **相互独立**, 如果满足

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

称三个事件 $A, B, C \in \mathcal{F}$ **相互独立**, 如果满足

$$P(AB) = P(A)P(B), \quad P(AC) = P(A)P(C), \quad P(BC) = P(B)P(C)$$

和

$$P(ABC) = P(A)P(B)P(C).$$

一般地, 称 n 个事件 $A_1, A_2, \cdots, A_n \in \mathcal{F}$ **相互独立**, 如果对于其中任意 k 个事件 $A_{i_1}, A_{i_2}, \cdots, A_{i_k}$ $(1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n)$, 满足

$$P(A_{i_1} A_{i_2} \cdots A_{i_k}) = P(A_{i_1}) P(A_{i_2}) \cdots P(A_{i_k}).$$

称事件序列 $A_1, A_2, \cdots, A_n, \cdots \in \mathcal{F}$ **相互独立**, 如果任取其中有限个均相互独立.

容易证明, 如果事件 $A_1, A_2, \cdots, A_n$ 相互独立, 且令 $\mathcal{F}_1 = \sigma(A_k, 1 \leqslant k \leqslant m)$, $\mathcal{F}_2 = \sigma(A_k, m+1 \leqslant k \leqslant n)$, 那么对于 $B_1 \in \mathcal{F}_1$ 和 $B_2 \in \mathcal{F}_2$, 有 B_1 和 B_2 相互独立.

(8) 如果事件序列 $A_1, A_2, \cdots, A_n, \cdots$ 相互独立, 且 $\sum\limits_{n=1}^{\infty} P(A_n) = \infty$, 那么

$$P\Big(\limsup_{n\to\infty} A_n\Big) = 1.$$

性质 (8) 的证明留给读者.

条件概率是概率论中的重要概念, 用途广泛. 下面介绍一下这个概念.

定义 1.7 设 A 是一个事件, 且 $P(A)>0$, 则称 $P(AB)/P(A)$ 为事件 A 发生下事件 B 的**条件概率**, 记为 $P(B|A)$, 即

$$P(B|A)=\frac{P(AB)}{P(A)}.$$

由条件概率的定义, 易得 $P(AB)=P(B|A)P(A)$, 这称为**乘法公式**. 进一步, 可得更一般的乘法公式

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)\cdots P(A_n|A_1A_2\cdots A_{n-1}).$$

借助于条件概率的概念, 容易得到下列全概率公式和贝叶斯公式. 设 $A_1,A_2,\cdots$ 两两互不相容, 且 $\bigcup\limits_n A_n=\Omega$, 则称 $\{A_n\}$ 为 Ω 的一个**分割**.

(9) 全概率公式: 如果事件 $\{A_n\}$ 为 Ω 的一个分割, 且 $P(A_n)>0$, 那么 $\forall B\in\mathcal{F}$,

$$P(B)=\sum_n P(A_n)P(B|A_n).$$

(10) 贝叶斯公式: 如果事件 $\{A_n\}$ 为 Ω 的一个分割, 且 $P(A_n)>0$, 那么对于事件 $P(A)>0$,

$$P(A_n|A)=\frac{P(A_n)P(A|A_n)}{\sum\limits_k P(A_k)P(A|A_k)},\quad n\geqslant 1.$$

容易看出, 如果记 $P_A(\cdot)=P(\cdot|A)$, 那么 $P_A(\cdot)$ 也是可测空间 $(\Omega,\mathcal{F})$ 上的一个概率测度, 从而 $(\Omega,\mathcal{F},P_A(\cdot))$ 也构成概率空间, 而且对于任意的 $A,B,C\in\mathcal{F}$, 当 $P(AB)>0$ 时, 有 $P_A(C|B)=P(C|AB)$; 并且 $P(C|AB)=P(C|A)$ 与 $P(BC|A)=P(B|A)P(C|A)$ 等价. 同样全概率公式也成立, 即如果 $\{B_n\}$ 是 Ω 的一个分割, 那么对于 $C\subseteq\bigcup\limits_n B_n$, 有

$$P_A(C)=\sum_n P_A(B_n)P_A(C|B_n),$$

也即

$$P(C|A)=\sum_n P(B_n|A)P(C|AB_n).$$

1.1.2 随机变量

定义 1.8 设 $(\Omega,\mathcal{F},P)$ 是一个概率空间, X 是定义在 Ω 上, 取值于实数集 $\mathbb{R}$ 的函数. 如果对任意实数 $x\in\mathbb{R}$, $\{\omega:X(\omega)\leqslant x\}\in\mathcal{F}$, 那么称 $X(\omega)$ 是概率空间 $(\Omega,\mathcal{F},P)$ 上的**随机变量**, 简称为随机变量. 函数

$$F(x)=P(\{\omega:X(\omega)\leqslant x\}),\quad -\infty<x<\infty$$

称为随机变量 X 的**分布函数**.

需要说明的是, 定义中样本点 ω 的集合 $\{\omega : X(\omega) \leqslant x\}$ 是一个事件, 一般简记为 $\{X \leqslant x\}$ 或 $\{X \in (-\infty, x]\}$. 容易验证, 对于任意实数 x, $\{X \leqslant x\}, \{X \geqslant x\}, \{X < x\}$ 和 $\{X > x\}$ 中只要有一个属于 σ 域 $\mathcal{F}$, 那么其余的就都属于 $\mathcal{F}$; 而且对于两个随机变量 X 和 Y 而言, $\{X < Y\}, \{X \leqslant Y\}, \{X = Y\}$ 和 $\{X \neq Y\}$ 也都属于 $\mathcal{F}$. 给定随机变量 X, 称包含所有形如 $\{X \leqslant x\}(x \in \mathbb{R})$ 的最小 σ 域为**由随机变量 X 生成的 σ 域** (**或代数**), 记为 $\sigma(X)$. 类似地, 可以定义由随机变量 $X_1, X_2, \cdots, X_n$ 生成的 σ 域 (或代数).

设 X 和 Y 是同一个概率空间 $(\Omega, \mathcal{F}, P)$ 上的两个随机变量, 若 $P(\{X \neq Y\}) = 0$ 成立, 则称它们是**几乎必然相等**, 记为 $X = Y$, a.s.. 对于随机变量 X 和 Y 而言, $X \pm Y$ 和 XY 也都是随机变量. 不难证明, 如果 $\{X_n\}$ 是一个随机变量序列, 那么 $\sup_n X_n, \inf_n X_n, \limsup\limits_{n \to \infty} X_n$ 和 $\liminf\limits_{n \to \infty} X_n$ 都是随机变量.

一般地, 常用的随机变量有两种类型: 离散型随机变量和连续型随机变量.

定义 1.9 若随机变量 X 的可能取值构成一个有限集或可列集, 则称 X 是一个**离散型随机变量**. 若随机变量 X 的可能取值充满数轴上的一个区间 (a, b), 则称其为**连续型随机变量**, 其中 a 可以是 $-\infty$, b 可以是 ∞.

对于一个离散型随机变量 X 而言, 如果 X 的所有可能取值是 $x_1, x_2, \cdots, x_n, \cdots$, 那么 X 取 x_i 的概率

$$p_i = P(\{X = x_i\}), \quad i = 1, 2, \cdots, n, \cdots$$

称为**概率分布列**或简称**分布列**, 其分布函数为

$$F(x) = \sum_{x_k \leqslant x} p_k.$$

对于一个连续型随机变量 X 而言, 存在实数轴上的一个非负可积函数 $f(x)$, 使得其分布函数 $F(x)$ 可表示为

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t, \quad \forall x \in \mathbb{R}.$$

此时, 称非负可积函数 $f(x)$ 为 X 的**概率密度函数**, 简称为**密度函数**或**密度**. 密度函数 $f(x)$ 的值虽然不是概率, 但是乘以微分元 $\mathrm{d}x$ 就可得小区间 $(x, x + \mathrm{d}x)$ 上概率的近似值, 即

$$f(x)\mathrm{d}x \approx P(x < X < x + \mathrm{d}x).$$

将相邻微分元累积起来就得到 X 在 (a,b) 上取值的积分, 此积分值就是 X 在 (a,b) 上取值的概率, 即

$$\int_a^b f(x)\mathrm{d}x = P(a < X < b).$$

更一般地, 对于 $(-\infty,\infty)$ 的子集 A, 有

$$P(X \in A) = \int_A f(x)\mathrm{d}x.$$

定义 1.10 如果 $X_1, X_2, \cdots, X_n$ 都是随机变量, 那么称 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ 是 n **维随机向 (变) 量**. 称 $\mathbb{R}^n$ 上的 n 元函数

$$F(\boldsymbol{x}) = F(x_1, x_2, \cdots, x_n) = P(X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_n \leqslant x_n)$$

为 $\boldsymbol{X}$ 的 n **维联合分布函数**或 n **维分布函数**. 如果有 $\mathbb{R}^n$ 上的非负函数 $f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_n)$, 使得对于 $\mathbb{R}^n$ 的任何子立方体

$$D = \{(x_1, x_2, \cdots, x_n) : a_i < x_i \leqslant b_i,\ 1 \leqslant i \leqslant n\},$$

有

$$P(\boldsymbol{X} \in D) = \int_D f(\boldsymbol{x})\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n, \tag{1.1}$$

那么称 $\boldsymbol{X}$ 是连续型随机向量, 称函数 $f(\boldsymbol{x})$ 为 $\boldsymbol{X}$ 的**概率密度函数**, 简称为**密度函数**或**密度**.

事实上, 可以证明对于 $\mathbb{R}^n$ 的任何子区域 D, (1.1) 式仍然成立. 而且显然

$$F(\boldsymbol{x}) = \int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n} f(s_1, s_2, \cdots, s_n)\mathrm{d}s_n\cdots\mathrm{d}s_1.$$

若 $F(\boldsymbol{x})$ 在 $\mathbb{R}^n$ 的开区域 D 中有连续的 n 阶混合偏导数, 且 $P(\boldsymbol{X} \in D) = 1$, 则

$$f(\boldsymbol{x}) = \begin{cases} \dfrac{\partial^n F(\boldsymbol{x})}{\partial x_n\cdots\partial x_1}, & \boldsymbol{x} \in D, \\ 0, & \text{其他}. \end{cases}$$

在后面章节里, 要经常用到一个随机向量变换的重要公式, 这就是下面的定理.

定理 1.1 设随机向量 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ 的概率密度函数是 $f(\boldsymbol{x}) = f(x_1, x_2, \cdots, x_n)$, $Y_i = g_i(\boldsymbol{X})(i = 1, 2, \cdots, n)$ 是 $\boldsymbol{X}$ 的函数, 且存在唯一的反函数