普通高等教育“十三五”规划教材
大学本科数学类专业基础课程系列丛书
兰州大学110周年校庆纪念文库

应用随机过程

焦桂梅 编著

科学出版社

普通高等教育"十三五"规划教材
大学本科数学类专业基础课程系列丛书
兰州大学110周年校庆纪念文库

应用随机过程

焦桂梅　编著

科 学 出 版 社
北　京

内 容 简 介

随机过程是研究随机现象的数量规律性的一个数学分支学科，也是构造随机模型的基础理论之一. 本书是在兰州大学数学与统计学院本科生、统计专业硕士生、萃英学院拔尖人才数学班的讲义基础上，经过反复调整、修改而成. 全书共 7 章，主要内容包括基本概念、泊松过程、更新理论、马尔可夫链、离散鞅引论、布朗运动与平稳过程、连续参数马尔可夫链. 本书着重揭示概念的来源与应用背景，用生动的例子刻画随机过程的特性.

本书可作为普通高等院校数学类、统计类以及金融和经济管理类专业本科生的教材，也可作为教师的教学参考书.

图书在版编目(CIP)数据

应用随机过程/焦桂梅编著. —北京：科学出版社，2019.8
(兰州大学 110 周年校庆纪念文库)
ISBN 978-7-03-062015-6

Ⅰ. ①应…　Ⅱ. ①焦…　Ⅲ. ①随机过程　Ⅳ. ①O211.6

中国版本图书馆 CIP 数据核字(2019) 第 161579 号

责任编辑：胡海霞　李　萍／责任校对：杨聪敏
责任印制：张　伟／封面设计：黄华斌

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

北京虎彩文化传播有限公司 印刷
科学出版社发行　各地新华书店经销
*
2019 年 8 月第　一　版　开本：787×1092　1/16
2019 年 9 月第二次印刷　印张：16　3/4
字数：410 000

定价：59.00 元
(如有印装质量问题，我社负责调换)

《兰州大学110周年校庆纪念文库》编委会

序

萃英立根本，昆仑写精神. 2019 年 9 月 17 日，兰州大学将迎来 110 周年校庆. 百十年来，一代代兰大人与国家、民族同呼吸、共命运，屹立西部大地，蕴育时代精英，为世界、为祖国培养了一大批活跃在各行各业的优秀人才，有力地支持了国家特别是祖国西部地区的建设发展.

长期以来，兰州大学始终坚持正确办学方向，落实立德树人根本任务，立足地域特色，发挥科研优势，深度融入参与国家发展战略，主动对接服务地方经济社会发展，“将论文写在中国大地上”，赢得了国内外的广泛认可；熔铸成以“自强不息，独树一帜”为核心的兰大精神，形成了“勤奋，求实，创新”的良好学风，探索走出了一条在西部地区创办高水平大学的成功之路，为中国高校扎根祖国大地创办世界一流大学提供了重要借鉴.

值 110 周年校庆之际，我校策划组织出版《兰州大学 110 周年校庆纪念文库》，旨在展现奋战在教学科研一线的兰大人的家国情怀、理论思考和学术积累. 丛书作者中有致力于教书育人的教学名师，也有在科研一线硕果累累的科学大家，更有长期坚守在教学科研一线、受学生爱戴的“普通”教师. 丛书内容丰富，涵盖理、工、农、医、人文、社科等诸多学科，其中观点颇多见解. 恕我才识单调，难以一一点评. 在此，谨付梓以供学界参考指正.

新时代新起点，所有兰大人将汇聚成推动兰州大学事业蓬勃发展的强大合力. 面向未来，全体兰大人将继续坚守奋斗，以矢志不渝的信念、时不我待的精神、担当奉献的情怀投身中国特色世界一流大学建设，为实现中华民族伟大复兴贡献兰大力量！

是为序.

严纯华

中国科学院院士、兰州大学校长

2019 年 3 月 26 日

前　言

近若干年来, 本书的初稿曾在兰州大学数学与统计学院本科生和统计专业硕士生、萃英学院本科拔尖人才数学班的教学中多次使用, 并根据广大学生的反馈意见, 反复调整、充实和修改. 此次的出版稿就是在这个基础上形成的.

随机过程是研究随机现象的数量规律性的一个数学分支学科, 也是构造随机模型的基础理论之一. 近几十年来, 由于实际问题的需要, 随机过程无论是在理论上还是在实际应用上都有了很大的发展, 它的基础知识、理论和方法, 不仅为数学、概率统计专业所必需, 也在自然科学、工程技术、经济管理和生物信息等领域有广泛的应用. 本书没有从严格的测度论的角度来介绍随机过程, 而是从较为直观、便于理解的角度来处理它, 我们使用实际生活中的许多生动的例子揭示概念的来源、应用背景和随机过程的特性, 寄希望于培养学生对本学科的直觉, 引发读者对随机过程学习和探究的兴趣, 使他们能从概率论的角度思考问题, 提高应用随机过程解决实际问题的能力.

鉴于随机过程这门学科发展迅速, 内容十分丰富, 作为一本大学本科生使用的教材, 不可能面面俱到, 包含其全部内容, 因此, 我们主要介绍随机过程的基础知识、基本方法以及它们在实际研究中的应用, 并配有大量实际问题的例子, 尽量回避与测度论相关的严格证明. 因为条件数学期望是研究随机过程的有力工具, 本书尽可能使用初等的、便于直观理解的方法陈述其定义和重要性质. 相对于同类教材, 本书增加了关于股票期权定价的一节, 因为随着现代金融数学和金融工程理论的发展, 鞅和布朗 (Brown) 运动被广泛应用于一些金融衍生品的研究中.

读者只要具有高等数学和初等概率论的基础知识便可阅读本书. 本书可作为大学本科数学类、统计类以及金融和经济管理类专业本科生的教材, 也可作为教师的教学参考书.

衷心感谢兰州大学中央高校基本科研业务费专项资金资助 (编号 lzujbky-2019-sp17), 兰州大学教务处和兰州大学萃英学院为本书的撰写和出版所给予的建设基金资助和出版基金资助, 也衷心感谢兰州大学数学与统计学院为本书的撰写和出版所给予的热情鼓励和大力支持. 感谢“教育部基础学科拔尖学生培养试验计划”的研究课题“拔尖学生知识积累过程中的能力培养”(编号: 20180706) 基金的资助.

特别感谢郭聿琦教授一直以来的关心、鼓励, 以及对本书出版的热情支持! 感谢科学出版社胡海霞编辑对作者的鼓励和为本书的出版所给予的细致入微的编辑加工! 也向承担助教工作和部分助教工作的硕士研究生陈晓莉、杨敬萍、牛彦、尚琦、邵苗苗、逯瑶瑶、李雨诗、王予涵、高阳等同学致谢, 感谢她 (他) 们提出了若干修订建议, 特别在各章后的习题设置上提出了很多建议, 并承担了书稿的录入和校对工作.

限于作者的水平, 本书的不妥之处在所难免, 欢迎批评指正.

焦桂梅 (兰州大学)

2019 年 3 月 26 日

目录

第1章 基本概念

随机过程通常被视为概率论的动态部分. 在概率论中, 研究的随机现象, 都是在概率空间 $(\Omega, \mathcal{F}, P)$ 上的一个或有限多个随机变量的规律性; 讨论中心极限定理也不过是对随机变量序列的讨论. 但在实际问题中, 我们还需要研究一些随机现象的发展和变化过程, 即随时间不断变化的随机变量, 而且, 所涉及的随机变量通常是无限多个, 这就是随机过程所要研究的对象. 随机过程以概率论作为其主要的基础知识, 为此, 首先简要介绍本书中常用的概率基础知识.

本章乃至全书的讨论要使用一些概率论的语言, 因此, 我们在 1.1 节先回顾概率空间的一些最基本的概念, 并尽量将它们精确化.

1.1 概率空间

概率空间是概率论的基础, 概率的严格定义基于这个概念. 1933 年, 苏联数学家**柯尔莫哥洛夫**(Kolmogorov) 提出了概率论的公理化结构, 使概率论成为严谨的数学分支. 在这个公理化结构中, 称三元总体 $(\Omega, \mathcal{F}, P)$ 为概率空间, 其中 Ω 是样本空间, $\mathcal{F}$ 是事件域, P 是概率, 它们都被认为是预先给定的, 并以此为出发点来讨论各种问题.

概率论是研究**随机现象**数量规律性的科学分支. 随机现象有其偶然性的一面, 但在大量重复随机试验中, 随机事件的出现具有频率稳定性, 即一个随机事件出现的频率常在某个固定的常数附近摆动, 我们称之为统计规律性.

对随机现象的研究必然要联系到对客观事物进行“调查”、“观察”或“实验”, 我们称之为**随机试验** (trial), 随机试验的结果事先不能准确预言, 但具有如下三个特性:

(1) 可以在相同的条件下重复进行;

(2) 每次试验的结果不止一个, 但预先知道试验所有可能出现的结果;

(3) 每次试验前不能确定哪个结果会出现.

把随机试验所有可能出现的结果构成的集合称为该试验的样本空间 (sample space), 记为 Ω, 每个可能的试验结果称为样本点 (sample point), 记为 $\omega_i, i = 1, 2, \cdots$. 由单个的样本点 ω_i 构成的集合称为基本事件, 一个样本点就是一个基本事件. 样本空间 Ω 称为必然事件, 在每次试验中一定会发生; 空集 $\varnothing$ 称为不可能事件, 在每次试验中都不会发生. 样本空间 Ω 的子集 A 由一些基本事件构成, 称为随机事件, 简称事件. 称事件 A 发生, 当且仅当事件 A 所包含的一个样本点出现.

1.1.1 事件的关系与运算

事件是由样本点构成的集合, 因此所有关于集合的关系与运算都可以移植到事件的关系与运算里来, 下面我们逐一说明.

(1) **事件的包含** 若事件 A 发生必然导致事件 B 发生, 则称事件 A 包含于事件 B, 也可称事件 B 包含事件 A, 记作 $A \subset B$. 从集合论的观点看, 就是事件 A 的每一个样本点都包含在事件 B 中.

(2) **事件的相等** 若 $A \subset B$ 且 $B \subset A$, 则称事件 A 与事件 B 相等, 记作 $A = B$. 从集合论的观点看, 就是事件 A 与事件 B 所包含的样本点一模一样.

(3) **事件 A 的逆事件 (对立事件)** 若事件 A 不发生, 记作 $\overline{A}$. 从集合论的观点看, 事件 $\overline{A}$ 表示由所有不包含在 A 中的样本点构成的集合. 显然, A 和 $\overline{A}$ 互为对立事件.

(4) **事件的交 (积)$A \cap B(AB)$** 事件 AB 表示事件 A 与事件 B 同时发生. 从集合论的观点看, 事件 AB 表示同时属于事件 A 与事件 B 的所有样本点构成的集合.

(5) **事件的并 (和)$A \cup B$** $A \cup B$ 表示事件 A 与事件 B 中至少有一个发生. 从集合论的观点看, $A \cup B$ 表示由属于事件 A 或事件 B 的所有样本点构成的集合.

(6) **事件的差 $A - B$** $A - B$ 表示事件 A 发生而事件 B 不发生. 从集合论的观点看, $A - B$ 表示由属于事件 A 而不属于事件 B 的所有样本点构成的集合.

(7) **互不相容事件** 若 $AB = \varnothing$, 则表示事件 A 与事件 B 不可能同时发生, 称事件 A 与事件 B 互不相容. 对互不相容的事件 A 与事件 B, 称它们的并为和.

(8) **有限或可数个事件的交 (积)** $\bigcap\limits_{i=1}^{n} A_i$ 表示有限个事件 $A_1, A_2, \cdots, A_n$ 同时发生, 从集合论的观点看, 事件 $\bigcap\limits_{i=1}^{n} A_i$ 表示同时属于事件 $A_1, A_2, \cdots, A_n$ 的所有样本点构成的集合. $\bigcap\limits_{i=1}^{\infty} A_i$ 表示可数个事件 $A_1, A_2, \cdots, A_n, \cdots$ 同时发生, 从集合论的观点看, 事件 $\bigcap\limits_{i=1}^{\infty} A_i$ 表示同时属于事件 $A_1, A_2, \cdots$ 的所有样本点构成的集合.

(9) **有限或可数个事件的并 (和)** $\bigcup\limits_{i=1}^{n} A_i$ 表示有限个事件 $A_1, A_2, \cdots, A_n$ 至少发生其一, 从集合论的观点看, 事件 $\bigcup\limits_{i=1}^{n} A_i$ 表示由至少属于 $A_i, i = 1, 2, \cdots, n$ 中某个事件的所有样本点构成的集合. $\bigcup\limits_{i=1}^{\infty} A_i$ 表示可数个事件 $A_1, A_2, \cdots, A_n, \cdots$ 至少发生其一, 从集合论的观点看, 事件 $\bigcup\limits_{i=1}^{\infty} A_i$ 表示由至少属于 $A_i, i = 1, 2, \cdots$ 中某个事件的所有样本点构成的集合.

(10) **完备事件组** 对有限个或可数个事件 $A_1, A_2, \cdots, A_n, \cdots$, 如果 $A_i A_j = \varnothing, i \neq j$, 且 $\bigcup\limits_{i=1}^{\infty} A_i = \Omega$, 则称 $A_1, A_2, \cdots, A_n, \cdots$ 为一个完备事件组.

1.1.2 事件的关系与运算的性质

(1) $\varnothing \subset A \subset \Omega$;

(2) $A - B = A\overline{B} = A - AB$;

(3) $\overline{A} = \Omega - A$;

(4) $A \cup B = A + (B - A) = (A - B) + (B - A) + AB$;

(5) $\overline{\overline{A}} = A$.

1.1.3 运算律

(1) 交换律 $A \cup B = B \cup A, A \cap B = B \cap A$;

(2) 结合律 $A\cup(B\cup C)=(A\cup B)\cup C, A\cap(B\cap C)=(A\cap B)\cap C$;

(3) 分配律 $A\cap(B\cup C)=(A\cap B)\cup(A\cap C), A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$;

(4) 德摩根 (De Morgan) 对偶律 $\overline{\bigcup\limits_{i=1}^{\infty} A_i}=\bigcap\limits_{i=1}^{\infty}\overline{A_i}, \overline{\bigcap\limits_{i=1}^{\infty} A_i}=\bigcup\limits_{i=1}^{\infty}\overline{A_i}$.

1.1.4 概率的定义

定义 1.1.1 若用 $\mathcal{F}$ 记事件的全体, 它是由 Ω 的一些子集构成的集类, 满足以下三个条件:

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

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

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

则称 $\mathcal{F}$ 为 σ 域, 亦称为 σ 代数.

定义 1.1.2 若 $\mathcal{F}$ 是由样本空间 Ω 的一些子集构成的一个 σ 域, 则称 $\mathcal{F}$ 为事件域 (event field), $\mathcal{F}$ 中的元素称为事件, Ω 称为必然事件, $\varnothing$ 称为不可能事件.

定义 1.1.3 (概率的公理化定义) 定义在事件域 $\mathcal{F}$ 上的一个集合函数 P 称为概率, 如果它满足以下三个条件:

(1) 非负性 $P(A)\geqslant 0$, 对一切 $A\in\mathcal{F}$;

(2) 规范性 $P(\Omega)=1$;

(3) 可列可加性或完全可加性 若 $A_n\in\mathcal{F}, n=1,2,\cdots$, 且 $A_iA_j=\varnothing\ (i\neq j)$, 则

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

1.1.5 概率的基本性质

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

(2) $0\leqslant P(A)\leqslant 1$.

(3) 概率的有限可加性 对有限个事件 $A_1,A_2,\cdots,A_n$, 如果 $A_iA_j=\varnothing, i\neq j$, 则

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

(4) 对任何事件 A, 有 $P(\overline{A})=1-P(A)$.

(5) 如果 $A\supset B$, 则 $P(A-B)=P(A)-P(B)$ 且 $P(A)\geqslant P(B)$.

更一般地, 有 $P(A-B)=P(A)-P(AB)$.

(6) $P(A\cup B)=P(A)+P(B)-P(AB)$, 因此有

$$P(A\cup B)\leqslant P(A)+P(B),$$

$$P(AB)\geqslant P(A)+P(B)-1.$$

(7) 一般加法公式 若 $A_1,A_2,\cdots,A_n$ 为 n 个事件, 则

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j,i,j=1,2,\cdots,n}^{n} P(A_iA_j)$$
$$+ \sum_{i<j<k,i,j,k=1,2,\cdots,n}^{n} P(A_iA_jA_k) - \cdots + (-1)^{n-1} P(A_1A_2\cdots A_n).$$

(8) 概率的下连续性　若记 $S_n = \sum_{i=1}^{n} A_i$, 则 $S_n \in \mathcal{F}, n = 1,2,\cdots$, 且 $S_n \subset S_{n+1}$, 即 S_n 是 $\mathcal{F}$ 中一个单调不减的集序列, 则有

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

(9) 概率的上连续性　若 $B_n \in \mathcal{F}, n = 1,2,\cdots$, 且 $B_n \supset B_{n+1}$, 即 B_n 是 $\mathcal{F}$ 中一个单调不增的集序列, 则有

$$\lim_{n\to\infty} P(B_n) = P\left(\lim_{n\to\infty} B_n\right) = P\left(\bigcap_{n=1}^{\infty} B_n\right).$$

定理 1.1.1　设 P 为 $\mathcal{F}$ 上满足 $P(\Omega) = 1$ 的非负集合函数, 则 P 具有可列可加性的充要条件是 P 具有有限可加性并且是下连续的.

定义 1.1.4　若 $A_n \in \mathcal{F}, i = 1,2,\cdots$, 所有属于无穷多个集合 A_n 的 ω 的集合事件称为 A_n 的上极限, 记为 $\overline{\lim_{n\to\infty}} A_n$, 可以证明

$$\overline{\lim_{n\to\infty}} A_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n.$$

定义 1.1.5　若 $A_n \in \mathcal{F}, i = 1,2,\cdots$, 除有限个集合 A_i 外, 属于无穷多个集合 A_n 的 ω 的集合事件称为 A_n 的下极限, 记为 $\underline{\lim_{n\to\infty}} A_n$, 它表示 A_n 最多有有限个不发生, 可以证明

$$\underline{\lim_{n\to\infty}} A_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n.$$

$\omega \in \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$ 当且仅当 ω 属于无穷多个 A_n.

类似地, $\omega \in \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n$ 当且仅当存在一个正整数 N, 使得 $\omega \in \bigcap_{n=N}^{\infty} A_n$.

1.2 条件概率及概率的三大重要公式

1.2.1 条件概率

定义 1.2.1　设 $(\Omega, \mathcal{F}, P)$ 是一个概率空间, $A \in \mathcal{F}$, 而且 $P(A) > 0$, 则对任意的 $B \in \mathcal{F}$, 记

$$P(B|A) = \frac{P(AB)}{P(A)}. \tag{1.1}$$

称 $P(B|A)$ 为**在事件 A 发生的条件下事件 B 发生的条件概率**(conditional probability).

如果 $P(A) = 0$, 此时 $P(AB)$ 也必为 0, 因此 $P(B|A)$ 为待定型, 进一步的研究在这里不做介绍. 以后出现 $P(B|A)$ 时, 都假设 $P(A) > 0$.

1.2.2 概率的三大重要公式

1. 乘法公式

若 $P(A)>0$, 则

$$P(AB)=P(A)P(B|A), \tag{1.2}$$

这个公式称为乘法公式, 若 $P(B)>0$, 乘法公式也可以写成

$$P(AB)=P(B)P(A|B). \tag{1.3}$$

乘法公式可以推广到任意 n 个事件的情形:

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1}), \tag{1.4}$$

这里要求 $P(A_1A_2\cdots A_{n-1})>0$.

2. 全概率公式

设事件 $A_1,A_2,\cdots,A_n,\cdots$ 是一个完备事件组, 即 $A_iA_j=\varnothing(i\neq j)$, 且 $\sum\limits_{n=1}^{\infty}A_n=\Omega$. 此时 $B=\sum\limits_{n=1}^{\infty}A_nB$, 则

$$P(B)=\sum_{n=1}^{\infty}P(A_n)P(B|A_n). \tag{1.5}$$

在很多实际问题中, 上式只含有限项.

证明 $P(B)=P(B\Omega)=P\left(B\sum\limits_{n=1}^{\infty}A_n\right)=P\left(\sum\limits_{n=1}^{\infty}BA_n\right)$

$=\sum\limits_{n=1}^{\infty}P(BA_n)=\sum\limits_{n=1}^{\infty}P(A_n)\ P(B|A_n)$. □

3. 贝叶斯 (Bayes) 公式

设事件 $A_1,A_2,\cdots,A_n,\cdots$ 是一个完备事件组, 即 $A_iA_j=\varnothing(i\neq j)$, 若事件 B 能且只能与两两互不相容的事件 $A_1,A_2,\cdots,A_n,\cdots$ 之一同时发生, 且 $\sum\limits_{n=1}^{\infty}A_n=\Omega$, 此时 $B=\sum\limits_{n=1}^{\infty}A_nB$, 则

$$P(A_i|B)=\frac{P(A_i)P(B|A_i)}{\sum\limits_{i=1}^{\infty}P(A_i)P(B|A_i)}. \tag{1.6}$$

证明 $P(A_i|B)=\dfrac{P(A_iB)}{P(B)}=\dfrac{P(A_i)P(B|A_i)}{\sum\limits_{i=1}^{\infty}P(A_i)P(B|A_i)}$. □

1.3 随机变量及分布函数

概率论主要研究能用随机变量描述的随机现象, 而分布函数完整地描述了随机变量的统计规律性, 因此分布函数起着非常重要的作用.

定义 1.3.1　设 $X(\omega)$ 是定义于概率空间 $(\Omega, \mathcal{F}, P)$ 上的单值实函数, 如果对于直线上任一博雷尔 (Borel) 点集 $\mathcal{B}$, 有

$$\{\omega : X(\omega) \in \mathcal{B}\} \in \mathcal{F}, \tag{1.7}$$

则称 $X(\omega)$ 为**随机变量**, 而 $P(X(\omega) \in \mathcal{B})$ 称为随机变量 $X(\omega)$ 的**概率分布**.

特别地, 若取 $\mathcal{B} = (-\infty, x]$, 则有 $P(X(\omega) \leqslant x) \in \mathcal{F}$, 因此 $P(X(\omega) \leqslant x)$ 有定义, 且有

$$P(a < X(\omega) \leqslant b) = P(X(\omega) \leqslant b) - P(X(\omega) \leqslant a). \tag{1.8}$$

定义 1.3.2　称 $F(x) = P(X(\omega) \leqslant x)(-\infty < x < +\infty)$ 为随机变量 $X(\omega)$ 的**分布函数**.

如果有非负可积函数 $f(x)$, 满足 $F(x) = \displaystyle\int_{-\infty}^{x} f(x)\,\mathrm{d}x$, 则称 $f(x)$ 为随机变量 X 或分布函数 $F(x)$ 的**分布密度函数**(density function). 如果 X 具有分布密度函数, 则称 X 为连续型随机变量, 如果 X 最多以正概率取可数多个值, 则称 X 为离散型随机变量.

显然, 若 X 为连续型随机变量, 则 $f(x) = F'(x)$.

由分布函数的性质可知, $f(x)$ 应该满足

$$f(x) \geqslant 0 \quad 且 \quad \int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1. \tag{1.9}$$

定理 1.3.1　分布函数 $F(x)$ 具有下列性质:

(1) 单调性　若 $a < b$, 则 $F(a) \leqslant F(b)$;

(2) $\lim\limits_{x\to-\infty} F(x) = 0$, $\lim\limits_{x\to+\infty} F(x) = 1$;

(3) 右连续性　$F(x+0) = F(x)$.

证明　(1) 因为 $a < b$, 则 $F(b) - F(a) = P(a < X \leqslant b) \geqslant 0$, 故 $F(a) \leqslant F(b)$;

(2) 因为 $F(x)$ 单调有界, 所以 $\lim\limits_{x\to-\infty} F(x)$ 和 $\lim\limits_{x\to+\infty} F(x)$ 都存在, 设事件 $A_n = \{\omega : X(\omega) \leqslant -n\}$, 则事件 $A_n \downarrow$ (表示单调下降趋向于) $\varnothing$, 由概率的上连续性可得

$$F(-\infty) = \lim_{x\to-\infty} F(x) = \lim_{n\to+\infty} F(-n) = \lim_{n\to+\infty} P(A_n) = P(\varnothing) = 0.$$

又设事件 $B_n = \{\omega : X(\omega) \leqslant n\}$, 则事件 $B_n \uparrow$ (表示单调上升趋向于) Ω, 由概率的下连续性可得

$$F(+\infty) = \lim_{x\to+\infty} F(x) = \lim_{n\to+\infty} F(n) = \lim_{n\to+\infty} P(B_n) = P(\Omega) = 1;$$

(3) 因事件 $\{X \leqslant x\} = \bigcap\limits_{n=1}^{\infty} \left\{X \leqslant x + \dfrac{1}{n}\right\}$, 且 $\left\{X \leqslant x + \dfrac{1}{n}\right\} \downarrow$ (关于n), 由概率的上连续性可得

$$\begin{aligned} F(x) &= P(X \leqslant x) = P\left(\bigcap_{n=1}^{\infty} \left\{X \leqslant x + \frac{1}{n}\right\}\right) = P\left(\lim_{n\to+\infty} \left\{X \leqslant x + \frac{1}{n}\right\}\right) \\ &= \lim_{n\to+\infty} P\left(X \leqslant x + \frac{1}{n}\right) = \lim_{n\to+\infty} F\left(x + \frac{1}{n}\right) = F(x+0). \end{aligned}$$

□

有了分布函数, 随机变量 $X(\omega)$ 的许多概率都能方便算出, 例如

$$P(X < b) = F(b-0), \quad P(X > b) = 1 - F(b),$$

$$P(X \geqslant b) = 1 - F(b-0), \quad P(X=b) = F(b) - F(b-0).$$

一般地, 随机变量 $X(\omega)$ 落在任一博雷尔集 B 中的概率可以写成

$$P(X(\omega) \in B) = \int_B \mathrm{d}F(x), \quad 其中 B \in \mathcal{B}.$$

特别地, 若 X 为连续型随机变量, 则

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

若 X 为离散型随机变量, 概率分布律为 $p_k = P(X = x_k), k = 1, 2, \cdots$, 则其概率分布函数可以表示为

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

若 X 为连续型随机变量, 密度函数为 $f(x)$, 则其概率分布函数可以表示为

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t.$$

分布函数是一种分析性质良好的函数, 给定了分布函数, 各种事件的概率都可以用分布函数表示出来, 因此求解各种事件的概率问题就可以转化为数学分析中函数的运算问题, 借助数学分析中的许多结论可以进一步研究概率问题.

1.4 随机向量、随机变量的独立性

1.4.1 随机向量及其分布

在有些随机现象中, 每次试验的结果不能只用一个随机向量来描述, 而要同时用几个随机向量来描述, 例如对于钢的成分, 需要同时指出它的含碳量、含硫量、含磷量等. 这样, 对应于每个样本点 ω, 试验的结果将是一个向量 $(\xi_1(\omega), \xi_2(\omega), \cdots, \xi_n(\omega))$, 这个向量取值于 n 维欧几里得空间 $\mathbb{R}^n$.

定义 1.4.1 若随机变量 $\xi_1(\omega), \xi_2(\omega), \cdots, \xi_n(\omega)$ 定义在同一概率空间 $(\Omega, \mathcal{F}, P)$ 上, 则称

$$\xi(\omega) = (\xi_1(\omega), \xi_2(\omega), \cdots, \xi_n(\omega))$$

构成一个 n **维随机向量**, 亦称 n **维随机变量**. 显然, 一维随机向量即随机变量.

固然可以对随机向量的每个分量分别研究, 但是我们马上就会看到, 把它们作为一个向量, 不但可以研究各个分量的性质, 而且还可以考察它们之间的联系, 对许多问题来说, 这是十分必要的.

对于任意的 n 个实数 $x_1, x_2, \cdots, x_n$,

$$\{\xi_1(\omega) \leqslant x_1, \xi_2(\omega) \leqslant x_2, \cdots, \xi_n(\omega) \leqslant x_n\} = \bigcap_{i=1}^{n} \{\xi_i(\omega) \leqslant x_i\} \in \mathcal{F},$$

亦即对于 $\mathbb{R}^n$ 中的 n 维矩形 $C_n=\prod_{i=1}^{n}(-\infty,x_i]$, 有

$$\{\xi(\omega)\in C_n\in\mathcal{F}\}.$$

利用测度论的方法还可以证明, 若 B_n 为 $\mathbb{R}^n$ 上任一博雷尔点集, 也有

$$\{\xi(\omega)\in B_n\in\mathcal{F}\}.$$

类似于一维的情形, 我们引进如下定义.

定义 1.4.2 称 n 元函数

$$F(x_1,x_2,\cdots,x_n)=P\{\xi_1(\omega)\leqslant x_1,\xi_2(\omega)\leqslant x_2,\cdots,\xi_n(\omega)\leqslant x_n\}$$

为随机向量 $\xi(\omega)=(\xi_1(\omega),\xi_2(\omega),\cdots,\xi_n(\omega))$ 的 **(联合) 分布函数**.

性质 1.4.1 多元分布函数 $F(x_1,x_2,\cdots,x_n)$ 具有下列性质:

(1) 单调性 分布函数 $F(x_1,x_2,\cdots,x_n)$ 对每个变量都是单调非降的;

(2) 右连续性 分布函数 $F(x_1,x_2,\cdots,x_n)$ 对每个变量都是右连续的;

(3) $$\lim_{x_i\to-\infty,i=1,2,\cdots,n}F(x_1,x_2,\cdots,x_n)=0,$$

$$\lim_{x_i\to+\infty,i=1,2,\cdots,n}F(x_1,x_2,\cdots,x_n)=1;$$

(4) 设 $x_i<y_i,i=1,2,\cdots,n$, 则有

$$P(x_1<X_1\leqslant y_1,\cdots,x_n<X_n\leqslant y_n)=\Delta_1\Delta_2\cdots\Delta_nF(x_1,x_2,\cdots,x_n)\geqslant 0,$$

这里 $\Delta_iF(x_1,x_2,\cdots,x_n)=F(x_1,x_2,\cdots,x_{i-1},y_i,x_{i+1},\cdots,x_n)-F(x_1,x_2,\cdots,x_n)$, 表示对第 i 个变量作差分运算. 其中 $n\geqslant 1,x_i\in R,1\leqslant i\leqslant n$.

特别地, 当 $n=2$ 时,

$$F(x_1,x_2)=P(X_1\leqslant x_1,X_2\leqslant x_2),$$

$$\begin{aligned}P(x_1<X_1\leqslant y_1,x_2<X_2\leqslant y_2)&=\Delta_1\Delta_2F(x_1,x_2)\\&=F(y_1,y_2)-F(x_1,y_2)-F(y_1,x_2)+F(x_1,x_2).\end{aligned}$$

一般地, 随机向量 $X=(X_1,X_2,\cdots,X_n)$ 落在任一 σ 域 B_n 中的概率可以写成

$$P(X(\omega)\in B_n)=\int\cdots\int_{B_n}\mathrm{d}F(x_1,x_2,\cdots,x_n),\quad 其中B_n\in\mathcal{B}_n.$$

随机向量 $\xi=(\xi_1,\xi_2,\cdots,\xi_n)$ 或者 $F(x_1,x_2,\cdots,x_n)$ 的边际分布定义为

$$F(x_{k_1},x_{k_2},\cdots,x_{k_m})=F(\infty,\cdots,\infty,x_{k_1},\infty,\cdots,\infty,x_{k_2},\infty,\cdots,\infty,x_{k_m},\infty,\cdots,\infty),$$

这里 $1\leqslant k_1<k_2<\cdots<k_m\leqslant n$.

类似的结论对 n 元情形也成立.

随机向量也有不同类型, 最常见的也是离散型与连续型两类.

在离散情形, 概率分布集中在有限或可列个点上. 重要的多元离散型分布有多项分布与多元超几何分布, 它们分别是二项分布与超几何分布在多元情形的推广.

例 1.4.1 (多项分布) 在试验中, 若每次试验的可能结果为 $A_1, A_2, \cdots, A_r$, 而 $P(A_i) = p_i, i = 1, 2, \cdots, r$, 且 $p_1 + p_2 + \cdots + p_r = 1$, 重复这种试验 n 次, 并假定这些试验是相互独立的, 若以 $\xi_1, \xi_2, \cdots, \xi_r$ 分别记为 $A_1, A_2, \cdots, A_r$ 出现的次数, 则

$$P\{\xi_1 = k_1, \xi_2 = k_2, \cdots, \xi_r = k_r\} = \frac{n!}{k_1!k_2!\cdots k_r!} p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r},$$

这里整数 $k_i \geqslant 0$, 且仅当 $k_1 + k_2 + \cdots + k_r = n$ 时上式才成立, 否则为 0.

例 1.4.2(多元超几何分布) 袋中装有 i 号球 N_i 只, $i = 1, 2, \cdots, r, N_1 + N_2 + \cdots + N_r = N$, 从中随机摸出 n 只, 若以 $\xi_1, \xi_2, \cdots, \xi_r$ 分别记为 $1, 2, \cdots, r$ 号球的出现次数, 则

$$P\{\xi_1 = n_1, \xi_2 = n_2, \cdots, \xi_r = n_r\} = \frac{\dbinom{N_1}{n_1}\dbinom{N_2}{n_2}\cdots\dbinom{N_r}{n_r}}{\dbinom{N}{n}},$$

这里整数 $n_i \geqslant 0$, 且仅当 $n_1 + n_2 + \cdots + n_r = n$ 时上式才成立, 否则为 0.

以上两个分布在抽样中常用, 前者用于有放回情形, 后者则用于不放回情形.

在连续情形, 若 $X = (X_1, X_2, \cdots, X_n)$ 为连续型随机向量, **密度函数**为 $f(x_1, x_2, \cdots, x_n)$, 则其概率分布函数可以表示为

$$F(x_1, x_2, \cdots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(y_1, y_2, \cdots, y_n)\, \mathrm{d}y_1 \cdots \mathrm{d}y_n.$$

密度函数 $f(x_1, x_2, \cdots, x_n)$ 有下列性质:

(1) $f(x_1, x_2, \cdots, x_n) \geqslant 0$;

(2) $\displaystyle\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \cdots, x_n) \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_n = 1$;

(3) $\displaystyle f(x_1, x_2, \cdots, x_n) = \frac{\partial^n F(x_1, x_2, \cdots, x_n)}{\partial x_1 \partial x_2 \cdots \partial x_n}$;

(4) $\displaystyle P(x_1 < X_1 \leqslant y_1, \cdots, x_n < X_n \leqslant y_n) = \int_{x_1}^{y_1} \cdots \int_{x_n}^{y_n} f(x_1, x_2, \cdots, x_n) \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_n$.

例 1.4.3 (均匀分布) 若 G 为 $\mathbb{R}^n$ 中有限区域, 其测度 $S > 0$, 则由密度函数

$$f(x_1, x_2, \cdots, x_n) = \begin{cases} \dfrac{1}{S}, & (x_1, x_2, \cdots, x_n) \in G, \\ 0, & (x_1, x_2, \cdots, x_n) \notin G \end{cases}$$

给出的分布称为 G 上的**均匀分布**.

例 1.4.4 (多元正态分布) 若 $\Sigma = (\Sigma_{ij})$ 是 n 阶正定对称矩阵, 以 $\Sigma^{-1} = (\gamma_{ij})$ 表示 Σ 的逆阵; $\det\Sigma$ 表示 Σ 的行列式的值. $\mu = (\mu_1, \cdots, \mu_n)$ 是任意实值行向量, 则由密度函数

$$p(x_1, x_2, \cdots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det\Sigma)^{\frac{1}{2}}} \exp\left\{ \frac{1}{2} \sum_{j,k=1}^{n} r_{jk}(x_j - \mu_j)(x_k - \mu_k) \right\}$$