

应用随机过程

宋占杰 胡 飞 李晨龙 编著



科学出版社

应用随机过程

宋占杰 胡 飞 李晨龙 编著

科学出版社

北 京

内 容 简 介

本书内容包括:概率论的基本知识、随机过程的基本概念、更新过程、离散时间的 Markov 链、连续时间的 Markov 过程以及随机分析和平稳过程. 本书力求贯彻选材精当和叙述详细的原则,注重说明概念的直观背景和实际意义. 在基本理论和方法上力求通俗易懂、深入浅出、淡化证明. 书中收集许多结合实际问题的数学建模实例,章末配有适当习题,有助于读者把握随机过程精髓.

本书可供高等院校理工类研究生和数学与应用数学专业本科生用作教材,也可作为工程技术管理人员自学参考用书.

图书在版编目(CIP)数据

应用随机过程/宋占杰,胡飞,李晨龙编著.—北京:科学出版社,2019.1
ISBN 978-7-03-060231-2

I. ①应… II. ①宋… ②胡… ③李… III. ①随机过程-高等学校-教材
IV. ①O211.6

中国版本图书馆 CIP 数据核字(2018)第 292125 号

责任编辑:王 静/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 1 月第 一 版 开本:720×1000 B5

2019 年 1 月第一次印刷 印张:13 1/8

字数:265 000

定价:39.00 元

(如有印装质量问题,我社负责调换)

前 言

应用随机过程是研究随时间或空间发生的随机现象的规律性, 建立数学模型并进行随机分析的一门应用学科. 这门学科有重要的应用价值. 因为在自然界, 随机现象远多于确定性现象, 加强随机数学的训练和培养在提高科技工作者能力和素质方面占重要地位.

本书旨在帮助科技人员了解随机过程的基本概念、基本模型和基本计算方法. 电子信息、精密仪器、建筑、化工、运筹学、管理科学、计算机科学及其他工程类专业博士研究生和其他相关科技工作者均可学习本书. 主要内容为: 第 1 章介绍概率论的基本知识; 第 2 章介绍随机过程的基本概念; 第 3 章介绍更新过程; 第 4 章介绍离散时间的 Markov 链; 第 5 章介绍连续时间的 Markov 过程; 第 6 章介绍随机分析; 第 7 章介绍平稳过程.

为增强实用性内容, 所有例题均选自国内外著名大学有代表性的教材, 同时简述了编者团队近十年来在相关方面的最新理论成果, 以使其内容密切联系国际发展前沿, 增加读者的学习兴趣. 为与其他学科内容相衔接并取得一致, 还利用国内现行数学分析中 Fourier 变换与反变换统一了平稳过程自相关函数与谱密度之间的关系.

本书第 1、2 章由宋占杰编写, 第 3~5 章由胡飞编写, 第 6、7 章由李晨龙编写, 由宋占杰统一定稿. 同时本书得到了国家自然科学基金 (编号: 61379014, 91746107) 的支持, 得到了天津大学研究生院创新人才培养项目的支持, 还得到了科学出版社领导及相关人员的热忱帮助. 王勇博士、张硕博士参加了部分校对工作, 在此一并致以衷心的感谢!

编 者

2018 年 1 月

目 录

第 1 章 概率论的基本知识	1
1.1 概率空间	1
1.1.1 事件域	1
1.1.2 概率	2
1.1.3 条件概率	3
1.1.4 事件的独立性	3
1.2 随机变量	4
1.2.1 随机变量及其分布函数	4
1.2.2 随机向量	5
1.2.3 随机变量的独立性	6
1.3 随机变量的数字特征	6
1.4 条件数学期望	9
1.5 概率论中常用的数学变换	16
1.5.1 Newton 插值公式	16
1.5.2 Fourier 变换	17
1.5.3 Laplace 变换	20
1.5.4 随机变量函数的期望	21
1.6 特殊随机变量函数的期望	22
1.6.1 母函数	22
1.6.2 矩母函数	25
1.6.3 特征函数	26
1.6.4 熵	30
1.6.5 偏度与峰度	30
1.7 极限定理	31
1.7.1 分布函数列的弱收敛性	31
1.7.2 四种收敛性	32
1.7.3 中心极限定理	32
1.8 n 维正态分布	34
习题 1	35

第 2 章 随机过程的基本概念	38
2.1 随机过程的定义	38
2.1.1 随机过程引例	38
2.1.2 有限维分布	39
2.2 随机过程的数字特征	40
2.3 几种重要的随机过程	44
2.3.1 正态过程	44
2.3.2 独立增量过程	45
2.3.3 Wiener 过程	45
2.3.4 鞅过程	46
2.4 Poisson 过程	47
2.4.1 Poisson 过程的定义	47
2.4.2 到达时间间隔与等待时间的分布	51
2.4.3 到达时间的条件分布	54
2.4.4 非齐次 Poisson 过程	56
2.4.5 复合 Poisson 过程	58
2.4.6 条件 Poisson 过程	60
习题 2	63
第 3 章 更新过程	67
3.1 更新过程的定义	67
3.1.1 $N(t)$ 的分布与更新函数	67
3.1.2 更新方程及其应用	70
3.2 更新方程与更新定理	73
3.3 更新过程的推广	77
3.3.1 延迟更新过程	77
3.3.2 交替更新过程	77
3.3.3 更新报酬过程	79
习题 3	81
第 4 章 离散时间的 Markov 链	85
4.1 Markov 链的基本概念	85
4.1.1 Markov 链的定义	85
4.1.2 转移概率	85
4.1.3 初始分布与绝对分布	86
4.2 首达时间和状态分类	90
4.3 状态空间的分解	97

4.4 遍历定理与平稳分布	101
4.4.1 遍历定理	101
4.4.2 平稳分布	104
4.5 Markov 链的最优停时	109
习题 4	114
第 5 章 连续时间的 Markov 过程	119
5.1 连续时间 Markov 过程的定义及其基本性质	119
5.2 Kolmogorov 微分方程	120
5.3 平稳分布与遍历性	123
5.4 生灭过程	128
习题 5	134
第 6 章 随机分析	137
6.1 二阶矩过程	137
6.1.1 二阶矩过程及 H 空间的定义	137
6.1.2 均方极限	138
6.2 二阶矩过程的均方微积分	140
6.2.1 均方连续性	140
6.2.2 均方导数	142
6.2.3 均方积分	145
6.2.4 均方导数与均方积分的分布	149
6.3 Itô 随机积分	152
6.3.1 Itô 积分的定义	152
6.3.2 Itô 微分法则	155
6.4 随机常微分方程	157
6.4.1 随机微分方程的均方理论	157
6.4.2 Itô 随机微分方程	159
习题 6	160
第 7 章 平稳过程	163
7.1 平稳过程的概念及性质	163
7.2 平稳过程和相关函数的谱分解	169
7.2.1 相关函数的谱分解	169
7.2.2 平稳过程的谱分解	173
7.3 线性系统中的平稳过程	175
7.3.1 线性时不变系统	176
7.3.2 频率响应函数与脉冲响应函数	177

7.3.3 线性时不变系统对随机输入的响应	180
7.3.4 平稳相关过程与互谱密度	184
7.4 平稳过程的均方遍历性	187
7.4.1 均方遍历性	187
7.5 平稳过程的采样定理	194
习题 7	197
参考文献	200

第1章 概率论的基本知识

1.1 概率空间

概率论中的最基本概念之一是随机试验(random experiment), 它具备以下三个特点:

- (1) 可重复性, 即在相同的条件下可以重复进行;
- (2) 不确定性, 即每次试验的结果不止一个, 且所有可能结果是预先可知的;
- (3) 不可预见性, 即每次试验出现何种结果事先是未知的.

随机试验的所有可能结果 ω 组成的集合称为这个试验的样本空间(sample space), 记为 Ω . Ω 中的元素 ω 称为样本点(sample point), 由样本点构成的单点集合称为基本事件(elementary event), 样本空间 Ω 的子集称为随机事件(random event), 简称事件. 称事件 A 发生, 当且仅当 A 中的一个样本点出现. 样本空间 Ω 也是一个事件, 称为必然事件(certain event). 空集 $\emptyset$ 称为不可能事件(impossible event).

由于事件是集合, 因此集合的运算 (并(union)、交(intersection)、差(difference)、上极限(limit superior)、下极限(limit inferior)、极限(limit) 等) 都适用于事件. 在实际问题中, 我们并非对所有的事件都感兴趣, 常常仅关心某些特定事件及其发生的可能性大小 (概率(probability)), 这样需要引入事件域(field of events) 的概念.

1.1.1 事件域

定义 1.1.1 设 Ω 是样本空间, $\mathcal{F}$ 是由 Ω 的一些子集构成的集类 (族), 如果它满足:

- (i) $\Omega \in \mathcal{F}$;
- (ii) 若 $A \in \mathcal{F}$, 则 $\bar{A} \in \mathcal{F}$;
- (iii) 若 $A_i \in \mathcal{F}, i = 1, 2, \dots$, 则 $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$,

则称 $\mathcal{F}$ 为事件域, $\mathcal{F}$ 中的元素 A 称为事件.

由定义 1.1.1 易知:

- (1) $\emptyset \in \mathcal{F}$;
- (2) 若 $A_i \in \mathcal{F}, i = 1, 2, \dots$, 则 $\bigcup_{i=1}^n A_i \in \mathcal{F}, \bigcap_{i=1}^n A_i \in \mathcal{F}, \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$;
- (3) 若 $A, B \in \mathcal{F}$, 则 $A - B \in \mathcal{F}$.

1.1.2 概率

定义 1.1.2 (公理化定义) 设 Ω 是样本空间, $\mathcal{F}$ 是 Ω 的一个事件域, 定义在 $\mathcal{F}$ 上的实值集函数(set function) $P(\cdot)$ 如果满足:

- (i) $\forall A \in \mathcal{F}$, 有 $P(A) \geq 0$;
- (ii) $P(\Omega) = 1$;
- (iii) 若 $A_i \in \mathcal{F}, i = 1, 2, \dots$, 且 $A_i A_j = \emptyset (i \neq j)$, 有

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

则称 $P(\cdot)$ 为 $\mathcal{F}$ 上的概率, $P(A)$ 为事件 A 的概率, 并且称三元总体 $(\Omega, \mathcal{F}, P)$ 为概率空间(probability space).

本书所指的事件都是指某一个事件域中的事件.

由定义 1.1.2 可知概率 P 具有如下性质.

性质 1 $P(\emptyset) = 0$.

性质 2 若 $A_i \in \mathcal{F}, i = 1, 2, \dots, n$, 且 $A_i A_j = \emptyset (i \neq j)$, 则

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

性质 3 对任意事件 A , 有

$$P(A) = 1 - P(\bar{A}).$$

性质 4 设 A, B 是任意两个事件, 且 $A \subset B$, 则

$$P(B - A) = P(B) - P(A), \quad P(A) \leq P(B).$$

性质 5 对任意事件 $A_i \in \mathcal{F}, i = 1, 2, \dots, n$, 有

$$\begin{aligned} P\left(\bigcup_{i=1}^n A_i\right) &\leq \sum_{i=1}^n P(A_i), \\ P\left(\bigcup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i A_j) \\ &\quad + \sum_{1 \leq i < j < k \leq n} P(A_i A_j A_k) - \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n). \end{aligned}$$

1.1.3 条件概率

定义 1.1.3 设 $(\Omega, \mathcal{F}, P)$ 为概率空间, 事件 $B \in \mathcal{F}$, 且 $P(B) > 0$, 对任意的事件 $A \in \mathcal{F}$, 记

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

则称 $P(A|B)$ 为在事件 B 发生的条件下, 事件 A 发生的**条件概率** (conditional probability).

由条件概率的定义可得:

(1) **乘法公式** (formula of multiplication) 设 $A, B \in \mathcal{F}, P(B) > 0$, 则

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

该公式可以扩展到 n 个事件的情形. 一般地, 若 $A_i \in \mathcal{F}, i = 1, 2, \dots, n$, 且 $P(A_1 A_2 \cdots A_n) > 0$, 则有

$$P(A_1 A_2 \cdots A_n) = P(A_1)P(A_2 | A_1)P(A_3 | A_1 A_2) \cdots P(A_n | A_1 A_2 \cdots A_{n-1}). \quad (1.1)$$

(2) **全概率公式** (total probability formula) 设 $A \in \mathcal{F}, B_i \in \mathcal{F}, i = 1, 2, \dots, n$, $B_i B_j = \emptyset (i \neq j)$, 且 $\bigcup_{i=1}^n B_i = \Omega, P(B_i) > 0$, 则

$$P(A) = \sum_{i=1}^n P(B_i)P(A | B_i). \quad (1.2)$$

(3) **贝叶斯公式** (Bayes formula) 设 $A \in \mathcal{F}, B_i \in \mathcal{F}, i = 1, 2, \dots, n, B_i B_j = \emptyset (i \neq j)$, 且 $\bigcup_{i=1}^n B_i = \Omega, P(B_i) > 0, P(A) > 0$, 则

$$P(B_i | A) = \frac{P(B_i)P(A | B_i)}{\sum_{i=1}^n P(B_i)P(A | B_i)}. \quad (1.3)$$

显然, 由条件概率公式、乘法公式和全概率公式就可以导出**逆概率(贝叶斯)公式** (inverse probability formula). 逆概率公式常用于计算产生该结果的各种可能原因的概率.

1.1.4 事件的独立性

定义 1.1.4 设 $(\Omega, \mathcal{F}, P)$ 为概率空间, $A, B \in \mathcal{F}$, 若

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

则称事件 A, B **相互独立** (mutually independence).

一般地, 设 $A_i \in \mathcal{F}, i = 1, 2, \dots, n$, 若对任意的 $m (2 \leq m \leq n)$ 及任意的 $1 \leq k_1 < k_2 < \dots < k_m \leq n$, 都有

$$P(A_{k_1} A_{k_2} \cdots A_{k_m}) = P(A_{k_1}) P(A_{k_2}) \cdots P(A_{k_m}),$$

则称事件 $A_1, A_2, \dots, A_n$ 是相互独立的. n 个事件独立, 需验证 $2^n - n - 1$ 个等式.

1.2 随机变量

1.2.1 随机变量及其分布函数

随机变量(random variable) 是概率论的主要研究对象, 随机变量的统计规律用分布函数(distribution function) 来描述.

定义 1.2.1 设 $(\Omega, \mathcal{F}, P)$ 是概率空间, $X = X(\omega)$ 是定义在 Ω 上的实值函数(real valued function). 若对任意实数 x , 都有

$$\{\omega : X(\omega) \leq x\} \in \mathcal{F},$$

则称 X 是 $\mathcal{F}$ 上的一个随机变量, 并称

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

为 X 的分布函数.

分布函数 $F(x)$ 具有如下的性质:

- (1) $0 \leq F(x) \leq 1$;
- (2) $F(x)$ 单调不减, 即对任意的 $x_1 < x_2$, 有 $F(x_1) \leq F(x_2)$;
- (3) $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0, F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1$;
- (4) $F(x)$ 右连续, 即任给 $x \in \mathbb{R}$, 有 $F(x+0) = F(x)$.

可以证明, 定义在实数域 $\mathbb{R}$ 上的实值函数 $F(x)$ 满足以上四条性质; 反过来, 如果 $F(x)$ 满足上述四条性质, 则 $F(x)$ 必为定义在某概率空间上某一随机变量 X 的分布函数.

常见的随机变量分为两类: 离散型随机变量(discrete random variable) 和连续型随机变量(continuous random variable).

若 X 的所有可能取值是有限个(finite) 或可列个(countable infinite), 则称 X 为离散型随机变量, 并称

$$P(X = x_i) = p_i, \quad i = 1, 2, \dots$$

为离散型随机变量 X 的概率分布律(distribution law) 或概率分布列.

若 X 的分布函数为 $F(x)$, 且存在非负实值函数 $f(x)$, 使得对任意的实数 x , 都有

$$F(x) = \int_{-\infty}^x f(t) dt$$

成立, 则称 X 为连续型随机变量, 并称 $f(x)$ 为 X 的概率密度函数 (probability density function).

既不是连续型, 也不是离散型的随机变量称为奇异型随机变量 (singular random variable).

1.2.2 随机向量

定义 1.2.2 设 $X_1(\omega), X_2(\omega), \dots, X_n(\omega)$ 是定义在概率空间 $(\Omega, \mathcal{F}, P)$ 上的 n 个随机变量, 则称

$$\mathbf{X}(\omega) = (X_1(\omega), X_2(\omega), \dots, X_n(\omega))$$

为 n 维随机向量 (n -dimensional random vector) (或 n 维随机变量 (n -dimensional random variable)), 简记为 $\mathbf{X} = (X_1, X_2, \dots, X_n)$, 并称 n 元函数

$$F(x_1, x_2, \dots, x_n) = P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\},$$

$$-\infty < x_1, x_2, \dots, x_n < +\infty$$

为 n 维随机变量 $(X_1, X_2, \dots, X_n)$ 的联合分布函数 (joint probability distribution function).

联合分布函数也具有类似于上述一维随机变量分布函数的所有性质.

若随机变量 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ 的所有可能取值只有有限个或可列个, 则称 $\mathbf{X}$ 为 n 维离散型随机变量, 并称

$$p_{i_1, i_2, \dots, i_n} = P\{X_1 = x_{1i_1}, X_2 = x_{2i_2}, \dots, X_n = x_{ni_n}\}, \quad i_1, i_2, \dots, i_n = 1, 2, \dots, n$$

为 n 维离散型随机变量 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ 的联合分布律 (joint probability distribution law).

若随机变量 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ 的分布函数为 $F(x_1, x_2, \dots, x_n)$, 且存在非负实值函数 $f(x_1, x_2, \dots, x_n)$, 使得对任意的实数 $x_1, x_2, \dots, x_n$, 总有

$$F(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(t_1, t_2, \dots, t_n) dt_1 dt_2 \cdots dt_n,$$

则称 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ 为 n 维连续型随机变量, 并称函数 $f(x_1, x_2, \dots, x_n)$ 为 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ 的联合概率密度函数 (joint probability density function).

1.2.3 随机变量的独立性

定义 1.2.3 设 $X_1, X_2, \dots, X_n$ 是定义在概率空间 $(\Omega, \mathcal{F}, P)$ 上的 n 个随机变量, 若对于任意的实数 $x_1, x_2, \dots, x_n$, 都有

$$P\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\} = P\{X_1 \leq x_1\}P\{X_2 \leq x_2\} \cdots P\{X_n \leq x_n\}, \quad (1.4)$$

则称随机变量 $X_1, X_2, \dots, X_n$ 是相互独立的.

(1.4) 式等价于

$$F(x_1, x_2, \dots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2) \cdots F_{X_n}(x_n), \quad (1.5)$$

其中 $F(x_1, x_2, \dots, x_n)$ 是 $(X_1, X_2, \dots, X_n)$ 的联合分布函数.

设 $(X_1, X_2, \dots, X_n)$ 是 n 维离散型随机变量, 则 (1.4) 式等价于

$$\begin{aligned} &P\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\} \\ &= P\{X_1 = x_1\}P\{X_2 = x_2\} \cdots P\{X_n = x_n\}. \end{aligned}$$

设 $(X_1, X_2, \dots, X_n)$ 是 n 维连续型随机变量, $f(x_1, x_2, \dots, x_n)$ 是其联合概率密度函数, 则 (1.5) 式等价于

$$f(x_1, x_2, \dots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2) \cdots f_{X_n}(x_n). \quad (1.6)$$

可以证明, 若随机变量 $X_1, X_2, \dots, X_n$ 相互独立, $g_i(x), i = 1, 2, \dots, n$ 为 Borel 可测函数(Borel measurable function), 则 $g_1(X_1), g_2(X_2), \dots, g_n(X_n)$ 也相互独立.

1.3 随机变量的数字特征

分布函数可以完整地描述随机变量, 但实际问题中求一个随机变量的分布函数多数情况下并非易事, 甚至有时不可能做到; 但对某些实际问题, 我们通常要求知道随机变量的某些概率特征即可, 并非真正需要知道其确切的分布. 本节介绍的数学期望、方差等概念正是刻画随机变量某些方面概率特征的数值指标.

定义 1.3.1 设 $F(x)$ 是随机变量 X 的分布函数, 若 $\int_{-\infty}^{\infty} |x| dF(x) < \infty$, 则称

$$E(X) = \int_{-\infty}^{\infty} x dF(x) \quad (1.7)$$

为 X 的数学期望(mathematical expectation)(或均值(mean value)).

若 X 是离散型随机变量, 其概率分布列为

$$p_i = P\{X = x_i\}, \quad i = 1, 2, \dots, \quad (1.8)$$

且 $\sum_i |x_i|p_i < \infty$, 则 X 的数学期望为

$$E(X) = \sum_i x_i p_i. \quad (1.9)$$

若 X 是连续型随机变量, 其概率密度函数为 $f(x)$, 且 $\int_{-\infty}^{\infty} |x| dF(x) < \infty$, 则 X 的数学期望是

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx. \quad (1.10)$$

定义 1.3.2 设 X 是随机变量, 若 $E(X)$ 存在, 且 $E(X - E(X))^2 < \infty$, 则称

$$D(X) = E(X - E(X))^2 \quad (1.11)$$

为随机变量 X 的方差(variance), 称 $\sqrt{D(X)}$ 为 X 的标准差(standard deviation).

定义 1.3.3 设 X, Y 是随机变量, $E(X^2) < \infty, E(Y^2) < \infty$, 则称

$$\text{cov}(X, Y) = E[(X - E(X))(Y - E(Y))] \quad (1.12)$$

为 X 与 Y 的协方差(covariance), 并且称

$$\rho_{XY} = \frac{\text{cov}(X, Y)}{\sqrt{D(X)}\sqrt{D(Y)}} \quad (D(X) > 0, D(Y) > 0) \quad (1.13)$$

为 X 与 Y 的相关系数(correlation coefficient). 若 $\rho_{XY} > 0$, 则称 X 与 Y 正相关; 若 $\rho_{XY} = 0$, 则称 X 与 Y 不相关; 若 $\rho_{XY} < 0$, 则称 X 与 Y 负相关.

随机变量的数字特征分别具有以下性质:

- (1) $E(C) = C, D(C) = 0$, 其中 C 为常数;
- (2) $E(aX + bY) = aE(X) + bE(Y)$, 其中 a, b 为常数;
- (3) 若 X 与 Y 独立, 则 $E(XY) = E(X)E(Y)$;
- (4) $D(X \pm Y) = D(X) + D(Y) \pm 2\text{cov}(X, Y)$, 特别地, 当 X 与 Y 独立时, $D(aX \pm bY) = a^2 D(X) + b^2 D(Y)$;
- (5) 对任意的实数 x , 有 $D(X) = E(X - E(X))^2 \leq E(X - x)^2$;
- (6) (Schwarz 不等式) 若 $E(X^2) < \infty, E(Y^2) < \infty$, 则 $[E(XY)]^2 < E(X^2)E(Y^2)$;
- (7) $\text{cov}(aX_1 + bX_2, cY_1 + dY_2) = a\text{ccov}(X_1, Y_1) + a\text{dcov}(X_1, Y_2) + b\text{ccov}(X_2, Y_1) + b\text{dcov}(X_2, Y_2)$;

$$(8) |\rho_{XY}| \leq 1;$$

(9) $|\rho_{XY}| = 1$ 的充要条件是 X 与 Y 以概率 1 线性相关 (linearly dependent with probability one), 即存在常数 a, b , 且 $a \neq 0$, 使得

$$P\{Y = aX + b\} = 1;$$

(10) 若 X 与 Y 独立, 则 $\rho_{XY} = 0$.

对于多个不相互独立的随机变量, 其数学期望和方差由下式给出:

$$E\left[\sum_{i=1}^n (a_i X_i)\right] = \sum_{i=1}^n a_i E(X_i);$$

$$D\left[\sum_{i=1}^n (a_i X_i)\right] = \sum_{i=1}^n a_i^2 D(X_i) + 2 \sum_{i < j} a_i a_j \text{cov}(X_i, X_j).$$

例 1.3.1 (Montmort 配对问题) n 个人参加一个聚会, 每个人所戴的帽子外观完全相同. 聚会结束时, n 顶帽子放在一起次序完全混乱, 每人随机取一顶. 求选中自己帽子的人数 X 的数学期望和方差.

解 设

$$X_i = \begin{cases} 1, & \text{第 } i \text{ 个人选中自己的帽子,} \\ 0, & \text{反之,} \end{cases} \quad i = 1, 2, \dots, n,$$

则有

$$X = \sum_{i=1}^n X_i.$$

易知

$$P\{X_i = 1\} = \frac{1}{n}, \quad P\{X_i = 0\} = \frac{n-1}{n}, \quad i = 1, 2, \dots, n.$$

因此

$$E(X_i) = \frac{1}{n} = E(X_i^2), \quad D(X_i) = E(X_i^2) - [E(X_i)]^2 = \frac{1}{n} - \frac{1}{n^2} = \frac{n-1}{n^2}.$$

又因为

$$X_i X_j = \begin{cases} 1, & \text{第 } i \text{ 个人和第 } j \text{ 个人都选中自己的帽子,} \\ 0, & \text{反之,} \end{cases} \quad (1.14)$$

其中 $i \neq j$, 所以

$$\begin{aligned} E(X_i X_j) &= 1 \times P\{X_i X_j = 1\} + 0 \times P\{X_i X_j = 0\} \\ &= P\{X_i = 1\} P\{X_j = 1 | X_i = 1\} = \frac{1}{n(n-1)}, \end{aligned}$$

$$\text{cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = \frac{1}{n^2(n-1)}.$$

因此,

$$\begin{aligned} E(X) &= E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = 1, \\ D\left[\sum_{i=1}^n (a_i X_i)\right] &= \sum_{i=1}^n a_i^2 D(X_i) + 2 \sum_{i < j} a_i a_j \text{cov}(X_i, X_j) \\ &= \frac{n(n-1)}{n^2} + \frac{2C_n^2}{n^2(n-1)} = 1. \end{aligned}$$

上述结果与参加聚会的人数 n 无关, 这是一个很有意思的结论.

因期望的计算公式与物理学中静力矩的计算公式相似, 借用物理学中“矩”的名字, $E(X)$ 称为**一阶矩**(first moment). 其中数学期望和方差作为随机变量最常用的数字特征, 它们都是某种**矩**(moment); 矩是最广泛的一种数字特征, 在概率论与数理统计中占重要地位. 最常用的矩有三种: **原点矩**(origin moment)、**中心矩**(central moment)、**混合矩**(mixed moment).

定义 1.3.4 设 X 是随机变量,

(1) 若 $E(X^k), k = 1, 2, \dots$ 存在, 则称它为 X 的 k 阶**原点矩**(k th origin moment), 简称 k 阶矩. $E|X|^k$ 称为 X 的 k 阶**绝对原点矩**或 k 阶**绝对矩**.

(2) 若 $E[X - E(X)]^k, k = 2, 3, \dots$ 存在, 则称它为 X 的 k 阶**中心矩**. $E|X - E(X)|^k$ 称为 X 的 k 阶**绝对中心矩**.

X 的一阶原点矩即为数学期望, 二阶中心矩即为方差.

注意到

$$|X|^k \leq |X|^{k+1} + 1, \quad k = 1, 2, \dots.$$

事实上, 当 $|X| > 1$ 时, 不等式显然成立. 当 $|X| \leq 1$ 时, $|X|^k \leq 1$, 故 $|X|^k \leq 1 \leq |X|^{k+1} + 1$ 也成立.

由以上不等式可推出

$$E(|X|^k) \leq E(|X|^{k+1} + 1) = E(|X|^{k+1}) + 1.$$

因此, 如果**高阶矩**(high-order moment) $E(|X|^{k+1})$ 有限, 则**低阶矩**(low-order moment) $E(|X|^k)$ 也有限, 从而低于 k 阶的各阶矩都有限.

1.4 条件数学期望

定义 1.4.1 设 (X, Y) 是二维离散型随机变量, 联合分布列是

$$P\{X = x_i, Y = y_j\} = p_{ij}, \quad i, j = 1, 2, \dots,$$