刘嘉焜 编著

应用随机过程

科学出版社

应用随机过程

刘嘉焜　编著

科学出版社

内 容 简 介

本书是为高等院校非数学专业高年级学生和研究生编写的教材. 内容包括随机过程基本概念、马氏过程、平稳过程和随机分析、随机微分方程等. 本书注重概念的直观背景、逻辑推导和实际应用三个方面,包含许多实际问题的例子,每章后面有习题,有助于读者学习和理解本书的内容.

本书的读者对象为高等院校物理、化学、生物、工程、管理、经济和金融等专业大学生、研究生和教师.

图书在版编目(CIP)数据

应用随机过程/刘嘉焜编著. -北京:科学出版社,2000

ISBN 7-03-008067-X

Ⅰ. 应… Ⅱ. 刘… Ⅲ. 随机过程-高等学校-教材 Ⅳ. O211.6

中国版本图书馆 CIP 数据核字 (1999) 第 67384 号

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

科地亚印刷厂 印刷

新华书店北京发行所发行　各地新华书店经售

*

2000 年 3 月第 一 版　开本: 787×1092　1/16

2000 年 3 月第一次印刷　印张: 16½

印数: 1-3 000　字数: 381 000

定价: 29.00 元

前　言

在当代任何学科的研究中，无论是自然科学还是社会科学，定性的与经验的理论被定量的与数学的理论所代替或补充已成为重要的方向之一．而其中一种主要方法是建立函数满足的一类方程或方程组来描述现实过程的一种假设的，可以产生一组理论结果的机制或模型．在构造数学模型时，按照所用的假设机制或模型的因果性质可分为确定性与随机性的两类.

随机过程论是研究随机现象的数量规律性的数学理论分支，是构造随机模型的基础理论之一．自从人们发现实际中几乎一切可察现象都具有随机性以后，随机过程的方法已经广泛地应用于科学与工程技术的所有领域了.

本书是为非数学专业的读者们编写的．可供自然科学、工程技术、经济管理各专业高年级学生、研究生及有关的实际工作者学习或参考．内容包括随机过程的基本概念，Markov 过程，平稳过程和随机分析与随机微分方程的初步介绍．在叙述时注意从数学概念的直观背景，形式逻辑的推导及实际应用三个方面去讨论．了解数学概念的直观意义有助于深入掌握概念的本质，它常常是理论的先导，并为逻辑推演提供思路和方法.

本书有较多的实际问题的例，较详细地介绍了构造随机模型的方法，以利于读者自学，有助于提高解决实际问题的能力．除第一章预备知识外，每章后面附有习题，完成这些习题是学习的重要环节.

本书定义、命题和例统一编号．同一章内公式 (4.3) 表示第 4 节第 3 个公式，命题 1.5 表示本章第 1 节命题 5．引用另一章的公式或命题时，前面加上章号，如 (2.2.4) 表示第二章第 2 节第 4 个公式.

这里作者特别感谢天津大学研究生院培养处的各位同仁，他们鼓励并帮助作者完成了这本书．作者还要感谢科学出版社林鹏副总编，他对本书的出版给予很大的支持与帮助．本书曾在天津大学及其它一些院校作为教材为应用数学专业及其它专业的研究生讲授多次．很多学生提出了有益的建议，其中吕泽林，李虹霖，周聿飞等人改正了书稿中的一些错误并打印了全书．没有他们的帮助，此书的出版也是不可能的.

由于作者水平有限，难免有疏漏与错误，不当之处请读者指正.

作　者

1999.9 于天津大学

目　录

第一章　预备知识 1

§1.1　概率空间 1

§1.2　随机变量 5

§1.3　随机变量的数字特征 10

§1.4　概率论中常用的几个变换 14

§1.5　条件期望 19

§1.6　随机变量的收敛性及极限定理 24

§1.6.1　分布函数列的弱收敛性 24

§1.6.2　随机变量的四种收敛性 25

§1.6.3　极限定理 27

第二章　随机过程的基本概念 32

§2.1　随机过程的定义 32

§2.2　正态过程 39

§2.3　Poisson 过程 47

§2.3.1　Poisson 过程的定义 47

§2.3.2　到达时间间隔与等待时间的分布 49

§2.3.3　非齐次 Poisson 过程 53

§2.3.4　复合 Poisson 过程 54

§2.3.5　条件 Poisson 过程 57

§2.4　更新过程 58

§2.4.1　引言 58

§2.4.2　$N(t)$ 的分布与更新函数 59

§2.4.3　极限定理与停时 61

§2.4.4　更新定理及其应用 65

§2.4.5　延迟更新过程 68

§2.4.6　有酬更新过程 71

§2.5　习题 73

第三章　Markov 过程 77

§3.1　可数状态 Markov 链 77

§3.1.1　定义与基本性质 77

§3.1.2　首达时间和状态分类 84

§3.1.3　闭集与状态空间的分解 90

§3.1.4　遍历定理 93

§3.1.5　平稳分布 97

§3.2 跳跃型 Markov 过程 105
§3.2.1 跳跃型 Markov 过程的定义 106
§3.2.2 Kolmogorov-Feller 积微分方程 108
§3.2.3 状态空间可数的齐次 (跳跃型)Markov 过程 111
§3.2.4 $p_{ij}(t)$ 的遍历性质 117
§3.3 扩散过程 125
§3.3.1 扩散过程的定义 125
§3.3.2 Kolmogorov 方程 127
§3.3.3 离散过程的扩散方程表示 134
§3.4 习题 137

第四章 随机分析与随机微分方程 **143**
§4.1 二阶矩过程和二阶矩随机变量空间 H 143
§4.1.1 二阶矩过程 143
§4.1.2 二阶矩随机变量空间 H 144
§4.1.3 均方极限的性质 147
§4.2 二阶矩过程的均方微积分 150
§4.2.1 均方连续性 150
§4.2.2 均方导数 151
§4.2.3 均方积分 155
§4.2.4 普通函数关于正交增量过程的积分 159
§4.2.5 均方导数与均方积分的分布 164
§4.3 Ito 积分 170
§4.3.1 Wiener-Einstein 过程及其形式导数 171
§4.3.2 Ito 积分的定义 171
§4.3.3 Ito 积分的性质 176
§4.3.4 Ito 微分法则 177
§4.4 随机常微分方程 179
§4.4.1 随机微分方程的均方理论 179
§4.4.2 Ito 随机微分方程 182
§4.5 习题 189

第五章 平稳过程 **192**
§5.1 平稳过程的基本概念 192
§5.1.1 平稳过程的定义 192
§5.1.2 平稳过程的性质 198
§5.1.3 平稳正态 Markov 过程 200
§5.2 平稳过程和相关函数的谱分解 204
§5.2.1 相关函数的谱分解 204

§5.2.2　平稳过程的谱分解 213
§5.2.3　平稳过程的线性运算 222
§5.3　均方遍历性 225
§5.3.1　平稳过程均方遍历性的基本概念 225
§5.3.2　平稳过程的遍历性定理 226
§5.4　线性系统中的平稳过程 230
§5.4.1　线性时不变系统 231
§5.4.2　输入为平稳过程的情形 234
§5.4.3　平稳相关过程和互谱函数 238
§5.5　平稳过程的采样定理 242
§5.5.1　采样定理 242
§5.5.2　白噪声 244
§5.6　平稳时间序列的线性预测 246
§5.7　习题 251

参考文献　254

第一章 预备知识

§1.1 概率空间

概率论中的一个基本概念是随机试验, 它是指其结果不能事先确定且在相同条件下可以重复进行的那种试验. 注意区别这里的试验与物理、化学中的实验. 以后就简称随机试验为试验. 一个试验所有可能结果的全体称为样本空间, 记为 Ω, 试验的一个结果称为样本点, 记为 ω. 即 $\Omega=\{\omega全体\}$.

样本空间的一个子集称为随机事件, 简称事件. 说事件 A 发生, 当且仅当 A 中的一个样本点发生. 事件的运算与集合论中集合的运算是一致的. 例如两个事件 A,B 至少有一个发生也是事件, 称为事件 A 与事件 B 的并, 记作 $A\cup B$; A 不发生是事件 A 的对立事件记作 $\bar{A}$, 分别对应于 A 与 B 的并集与集合 A 的余集. 根据实际情况, 我们并不总是有兴趣研究 Ω 的一切子集 (事件), 而仅对 Ω 中的某些事件类感兴趣. 于是我们得到事件域的概念. 如果 $\mathcal{F}$ 是 Ω 中某些事件的集, 它满足:

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

(ii) 若 $A\in\mathcal{F}$, 则 $\bar{A}\in\mathcal{F}$;

(iii) 若 $A_i\in\mathcal{F},i=1,2,\ldots$, 则 $\cup_{i=1}^{\infty}A_i\in\mathcal{F}$,

就称 $\mathcal{F}$ 为事件域. 以下论及的事件都是指某一事件域中的.

对 $\mathcal{F}$ 中每一事件 A, 都有一个实数 $P(A)$ 与之对应, 它满足

(i) $0\le P(A)\le 1$;

(ii) $P(\Omega)=1$;

(iii) 若 $A_i\cap A_j=\emptyset,i\ne j,i,j=1,2,\ldots$, 则 $P(\cup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}P(A_i)$.

我们称 $P(A)$ 为事件 A 的概率. 关于 P 的公理 (i),(ii),(iii) 有以下推论:

推论 1.1.1 若 $E\subset F$, 则 $P(E)\le P(F)$;

推论 1.1.2 $P(\bar{E})=1-P(E)$;

推论 1.1.3 若 $E_i\cap E_j=\emptyset,i\ne j$, 则 $P(\cup_{i=1}^{n}E_i)=\sum_{i=1}^{n}P(E_i)$;

推论 1.1.4 $P(\cup_{i=1}^{\infty}E_i)\le\sum_{i=1}^{\infty}P(E_i)$.

三元总体 $(\Omega,\mathcal{F},P)$ 称为概率空间. 事件域的引进使我们的模型有更大的灵活性. 在实际问题中可根据问题的性质选定 $\mathcal{F}$, 比如 Ω 为有穷集或为可列集时, 一般选 $\mathcal{F}$ 为 Ω 一切子集组成的集合; 在几何概率问题中选 $\mathcal{F}$ 为 Lebesgue 可测集. 以下如不预先说明, 所有的事件都是指某一 $\mathcal{F}$ 中的事件.

$P(\cdot)$ 为以 $\mathcal{F}$ 为定义域的集函数. P 的一个重要性质是连续性, 为描述 P 的连续性, 我们引进极限事件的概念. 一列事件 $A_n,n\ge 1$ 称为递增的事件列, 如果 $A_n\subset A_{n+1},n\ge 1$; 称为递减的事件列, 如果 $A_n\supset A_{n+1},n\ge 1$. 如 $\{E_n,n\ge 1\}$ 是一递增 (减) 的事件列, 我们称 $\cup_{i=1}^{\infty}E_i(\cap_{i=1}^{\infty}E_i)$ 为极限事件, 记为 $\lim_{n\to\infty}E_n$, 即 $\lim_{n\to\infty}E_n=\cup_{i=1}^{\infty}E_i(\cap_{i=1}^{\infty}E_i)$, 这里 $E_n\subset E_{n+1}(E_n\supset E_{n+1}),n\ge 1$. 对满足 (i)(ii)(iii) 的 P, 我们有

命题 1.1.5 (P 的连续性) 若 $\{E_n, n \geq 1\}$ 是递增或递减的事件列, 则

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

证明 先假设 $\{E_n, n \geq 1\}$ 是递增的事件列. 定义事件 $F_n, n \geq 1$, $F_1 = E_1, F_n = E_n\overline{(\cup_{i=1}^{n-1} E_i)} = E_n\overline{E_{n-1}}, n > 1$, 显然有 $F_iF_j = \emptyset, i \neq j$ 且对 $n \geq 1$ 有 $\cup_{i=1}^{\infty} F_i = \cup_{i=1}^{\infty} E_i$, 故

$$\begin{aligned}
P(\cup_{i=1}^{\infty} E_i) &= P(\cup_{i=1}^{\infty} F_i) \\
&= \sum_{i=1}^{\infty} P(F_i) \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} P(F_i) \\
&= \lim_{n\to\infty} P(\cup_{i=1}^{n} F_i) \\
&= \lim_{n\to\infty} P(\cup_{i=1}^{n} E_i) \\
&= \lim_{n\to\infty} P(E_n).
\end{aligned}$$

若 $\{E_n, n \geq 1\}$ 是递减的, 那么 $\{\overline{E_n}, n \geq 1\}$ 是递增的, 由上面证明有

$$P(\cup_{i=1}^{\infty} \overline{E_i}) = \lim_{n\to\infty} P(\overline{E_n}).$$

由 De Morgan 定律以及推论 (1.1.2), 上式成为

$$P(\overline{\cap_{i=1}^{\infty} E_i}) = \lim_{n\to\infty} P(\overline{E_n}),$$

即

$$1 - P(\cap_{i=1}^{\infty} E_i) = \lim_{n\to\infty} [1 - P(E_n)],$$

或

$$P(\cap_{i=1}^{\infty} E_i) = \lim_{n\to\infty} P(E_n). \quad \blacksquare$$

例 1.1.6 考虑生物群体模型.

某个群体由同类个体组成, 每个个体可产生同类的后代. 设开始的个体数目为 X_0, 即第 0 代个体的个数是 X_0, 第 0 代的后代是第 1 代, 其个数为 X_1. 一般地, 第 n 代的个数记为 $X_n, n = 1, 2, \ldots$.

由于 $X_n = 0$ 蕴涵 $X_{n+1} = 0$, 故 $\{X_n = 0\}$ 是递增的事件列, 所以极限 $\lim_{n\to\infty} P(X_n = 0)$ 必然存在. 这一极限表示什么呢? 由命题 1.1.5,

$$\begin{aligned}
\lim_{n\to\infty} P(X_n = 0) &= P(\lim_{n\to\infty} \{X_n = 0\}) \\
&= P(\cup_{n=0}^{\infty} (X_n = 0)) \\
&= P(\text{群体终于灭绝}),
\end{aligned}$$

即这一极限表示群体终于灭绝的概率.

对任一事件列 $\{E_n, n \geq 1\}$, "无穷多个 E_n 出现" 也是一个事件, 称为上极限事件, 记作 $\overline{\lim}_{n\to\infty} E_n$ 或 $\limsup_{n\to\infty} E_n$; "至多只有有穷多个 E_n 不出现" 或者 "几乎一切 E_n 都出现" 也是一个事件, 称为下极限事件, 记作 $\underline{\lim}_{n\to\infty} E_n$ 或 $\liminf_{n\to\infty} E_n$. 我们有

$$\limsup_{n\to\infty} E_n = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} E_k,$$

$$\liminf_{n\to\infty} E_n = \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} E_k.$$

事实上, 若无穷多个 E_k 出现, 那么对任意一个 n,$\cup_{k=n}^{\infty} E_k$ 都出现, 也就是说 $\cap_{n=1}^{\infty}\cup_{k=n}^{\infty} E_k$ 出现, 即 $\limsup_{n\to\infty} E_n \subset \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} E_k$; 反之, 若 $\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} E_k$ 出现, 那么对每一个 n, $\cup_{k=n}^{\infty} E_k$ 都出现, 这样对每一个 n, 至少有一个 $E_k, k \geq n$ 出现, 亦即有无穷多个 E_k 出现, 即 $\limsup_{n\to\infty} E_n \supset \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} E_k$. 第二个式子的证明完全类似.

命题 1.1.7 Borel–Cantelli 引理

设 $E_1, E_2, \ldots$ 为一列事件, 若 $\sum_{i=1}^{\infty} P(E_i) < \infty$, 则

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

证明 容易看出 $\cup_{k=n}^{\infty} E_k, n \geq 1$, 是一列递减的事件, 由命题 (1.1.5) 可知

$$\begin{aligned}
P(\limsup_{n\to\infty} E_n) &= P(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} E_k) \\
&= P(\lim_{n\to\infty} \cup_{k=n}^{\infty} E_k) \\
&= \lim_{n\to\infty} P(\cup_{k=n}^{\infty} E_k) \\
&\leq \lim_{n\to\infty} \sum_{k=n}^{\infty} P(E_k) \\
&= 0,
\end{aligned}$$

不等式是由推论 (1.1.4) 得出, 最后一个等号是根据 $\sum_{k=1}^{\infty} P(E_k) < \infty$. ∎

例 1.1.8 设 $P(A_n) = \dfrac{1}{n^2}, P(\overline{A_n}) = 1 - \dfrac{1}{n^2}, n \geq 1$, 故, $\sum_{n=1}^{\infty} P(A_n) < \infty$. 由命题 (1.1.7), $P(\limsup_{n\to\infty} A_n) = 0$. 这是说有无穷多个 A_n 出现是不可能的 (概率为 0). 换言之, 只有有穷多个 A_n 出现是必然的 (概率为 1). 这就是说

$$P(\liminf_{n\to\infty} \overline{A}_n) = 1.$$

命题 1.1.9 Borel-Cantelli 引理的逆

若 $E_1, E_2, \ldots$ 是一列相互独立的事件, 且 $\sum_{n=1}^{\infty} P(E_n) = \infty$, 则

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

证明

$$\begin{aligned} P(\cap_{n=1}^{\infty}\cup_{k=n}^{\infty}E_k) &= P(\lim_{n\to\infty}\cup_{k=n}^{\infty}E_k) = \lim_{n\to\infty}P(\cup_{k=n}^{\infty}E_i) \\ &= \lim_{n\to\infty}[1-P(\cap_{k=n}^{\infty}\overline{E_k})], \end{aligned}$$

但由独立性,

$$\begin{aligned} P(\cap_{k=n}^{\infty}\overline{E_k}) &= \prod_{i=n}^{\infty}P(\overline{E_i}) = \prod_{i=n}^{\infty}(1-P(E_i)) \\ &\le \prod_{i=n}^{\infty}\mathrm{e}^{-P(E_i)} = \exp\left\{-\sum_{i=n}^{\infty}P(E_i)\right\} = 0, \end{aligned}$$

不等号是由于 $1-x\le \mathrm{e}^{-x}$, 最后一个等号是因为对任意 $n, \sum_{i=n}^{\infty}P(E_i)=\infty$. ∎

例 1.1.10 设 $X_1, X_2, \ldots$ 是一列相互独立的随机变量, 对 $n\ge 1$,

$$P(X_n=0)=\frac{1}{n}, P(X_n=1)=1-\frac{1}{n},$$

故 $\sum_{n=1}^{\infty}P(X_n=0)=\infty$, 由命题 1.1.9, 有无穷多个 n 使 $X_n=0$ 是必然的. 同样 $\sum_{n=0}^{\infty}P(X_n=1)=\infty$, 可以推知也有无穷多个 n 使 $X_n=1$ 是必然的. 因此, 以概率 1, $\lim_{n\to\infty}X_n$ 不存在.

例 1.1.11 顾客随机的到达服务台接受服务. 设 A_k 表示在时间 $(0,t)$ 间隔内到达 k 个顾客的事件, 条件 $\sum_{k=0}^{\infty}P(A_k)=1$, 由命题 (1.1.7) 可推出 $P(\limsup_{n\to\infty}A_n)=0$, 亦即无穷多个 A_k 出现的概率是 0. 换言之, 表示在 $(0,t)$ 的时间间隔内以概率 1 只有有穷多个顾客到达.

例 1.1.12 统计物理: 在研究由大量同一性质 "质点" 所构成的具有能量 E 的物理体系的平衡状态时, 我们假定能级 e_1 有 g_1 个微观状态, 能级 e_2 有 g_2 个微观状态, ……, 每一质点处于这些状态之一. 所谓宏观状态, 可由处于各个能级的质点个数来刻画: v_1 个质点处于能级 e_1, v_2 个质点处于能级 $e_2, \ldots$. 设 N 是质点总数. 一个宏观状态就是 $v_1=n_1, v_2=n_2, \ldots$, 这里 $\sum_i n_i=N, \sum_i n_ie_i=E$. 对于很大的 N, 宏观体系平衡状态指的是可能性最大的状态. 我们要计算 n_i 个质点分配到能级 e_i 的 g_i 个状态去的概率, $i=1,2,\ldots$. 特别地求 g_i 个状态中预先指定的 n_i 个状态中各出现一个质点的概率, $n_i\le g_i$.

解 1. (Maxwell-Boltzman) 假定 n_i 个质点可以分辨, 处于每个状态的质点个数是任意的. 这时的样本空间共含有 $g_i^{n_i}$ 个不同的样本点, 因每个质点可以分辨, 所以所求概率为 $p=\dfrac{n_i!}{g_i^{n_i}}$ 这一模型是研究气体的古典理论时提出的. 现代物理表明, 目前已知的基本粒子都不适用于这一模型.

解 2. (Bose-Einstein) 假定 n_i 个质点是不可分辨的, 处于每个状态的质点个数是任意的. g_i 个状态看作 g_i 个小盒, 可以由 g_i-1 个壁 "|" 分开. 用 * 表示质点, 则 $|**|||*|*|$ 表示 $n_i=4$ 个质点分在 $g_i=5$ 个盒中的一种情况. 所有情况的总数是 n_i 个质点与 g_i-1 个 "|" 共 n_i+g_i-1 个位置中选 n_i 个位置放 * 的所有可能数, 即 $\binom{n_i+g_i-1}{n_i}$. 故

$p = \left[\begin{pmatrix} n_i + g_i - 1 \\ n_i \end{pmatrix}\right]^{-1}$. 近代物理已经证明这一模型适用于光子、介子、核子等粒子. 这一模型是 1924 年 Bose 与 Einstein 提出的, 适合称为 Bose 子的基本粒子.

解 3. (Fermi-Dirac) 假定 n_i 个质点不可分辨. 但每个状态只能有一个质点. 此时 $p = \left[\begin{pmatrix} g_i \\ n_i \end{pmatrix}\right]^{-1}$. 这一模型适用于电子、中子和质子, 是 Fermi-Dirac 在 1925 年提出的, 适用于所谓 Fermi 子的粒子.

§1.2 随机变量

给定概率空间 $(\Omega, \mathcal{F}, P)$, 随机变量 $X(\omega)$ 是对 $\omega \in \Omega$ 赋与一个实数值的函数, 它满足条件: 对任一实数 $a, \{\omega : X(\omega) \le a\} \in \mathcal{F}$.

另一个定义是: 随机变量 $X(\omega)$ 是 $(\Omega, \mathcal{F})$ 到可测空间 $(R, \mathcal{B})$ 的可测映射. 这里 $\mathcal{B}$ 是一维 Borel 集类, 或者说, 对一切 $A \in \mathcal{B}$ 有 $\{\omega : X(\omega) \in A\} \in \mathcal{F}$, 我们常常将 ω 略去不写.

对 $A \in \mathcal{B}$, 我们称 $P(X \in A) = P(X^{-1}(A))$ 为随机变量 X 的分布, 它表示随机变量 X 取值于 A 中的概率. 它刻画了该随机变量的全部概率性态, 常记为 $F_X(A) = P(X \in A)$.

随机变量 X 的分布函数 $F_X(x)$ 定义为: 对任意实数 x,

$$F_X(x) = P(X \le x).$$

我们常常略去下脚的 X 直接写 $F(A), F(x)$. 随机变量 X 的分布函数具有性质:
(i) 单调不减; (ii) 右连续; (iii)$F(-\infty) = 0$; (iv)$F(+\infty) = 1$. 反之, 我们有

命题 1.2.1 *设 $F(x), x \in R$ 是单调不减、右连续的函数, 且有 $F(-\infty) = 0, F(+\infty) = 1$, 则必存在概率空间及其上的一个随机变量 ξ, 使 $F_\xi(x) = F(x)$.*

证明 考虑可测空间 $(R, \mathcal{B})$, 令

$$P\{(a, b]\} = F(b) - F(a),$$

其中 $a, b \in R, a \le b$. 由测度扩张定理, 知 P 可以扩张为 $\mathcal{B}$ 上的概率测度. 于是我们得到概率空间 $(R, \mathcal{B}, P)$. 定义随机变量 $\xi(\omega) = \omega, \omega \in R$, 则 $F(x)$ 就是 ξ 的分布函数. ∎

若随机变量 X 的可能值的全体是至多可数的, 就称 X 是离散型随机变量. 对离散型随机变量 X,

$$F(x) = \sum_{y \le x} P(X = y).$$

若对随机变量 X 存在一个函数 $f(x)$, 使得 X 的分布函数可表示为

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

就称 X 是连续型随机变量. $f(x)$ 称为 X 的概率密度函数, 或简称密度. 对连续型随机变量 X,

$$P\{X \in B\} = \int_B f(x) \mathrm{d}x, B \in \mathcal{B},$$

$$\frac{\mathrm{d}}{\mathrm{d}x}F(x) = f(x).$$

概率空间 $(\Omega, \mathcal{F}, P)$ 上有 n 个随机变量 $X_1, X_2, \ldots, X_n$, 则 $X = (X_1, X_2, \ldots, X_n)$ 称为 n 维随机变量, 或称 n 维随机向量.

随机向量 X 是可测空间 $(\Omega, \mathcal{F})$ 到 $(R^n, \mathcal{B}^n)$ 的可测映象. $F_X(A) = P(X \in A), A \in \mathcal{B}^n$ 称为随机向量 X 的分布或称为 $X_1, X_2, \ldots, X_n$ 的联合分布. 特别地取 $A = (-\infty, x_1] \times \ldots \times (-\infty, x_n], x_1, \ldots, x_n \in R$ 就得到随机向量 X 的分布函数, 或 $X_1, X_2, \ldots, X_n$ 的联合分布函数

$$F_X(x_1, x_2, \ldots, x_n) = P(X_1 \le x_1, \ldots, X_n \le x_n),$$

它具有性质:

(i) $F_X(x_1, \ldots, x_n)$ 对任一 x_i 是单调不减的, $i = 1, 2, \ldots, n$;

(ii) $F_X(x_1, \ldots, x_n)$ 对任一 x_i 是右连续的, $i = 1, 2, \ldots, n$;

(iii) $F_X(\infty, \ldots, \infty) = 1$, $F_X(x_1, \ldots, x_{i-1}, -\infty, x_{i+1}, \ldots, x_n) = 0$, $i = 1, 2, \ldots, n$;

(iv) 设 $x_i \le y_i, i = 1, 2, \ldots, n$, 则

$$F(y_1, \ldots, y_n) - \sum_{i=1}^{n} F_i + \sum_{i<j} F_{ij} - \ldots + (-1)^n F(x_1, \ldots, x_n) \ge 0,$$

这里 $F_{ij\ldots k}$ 是 $z_i = x_i, z_j = x_j \ldots, z_k = x_k$ 而其余 $z_l = y_l$ 时 $F(z_1, z_2, \ldots, z_n)$ 的值.

命题 1.2.2 *已给 n 元函数 $F(x_1, \ldots, x_n)$, 满足上面 (i),(ii),(iii),(iv), 则必存在概率空间 $(\Omega, \mathcal{F}, P)$ 及其上的随机向量 ξ, 使 ξ 的分布函数 $F_\xi(x_1 \ldots, x_n) = F(x_1, \ldots, x_n)$.*

证明略.

注意: 性质 (iv) 不能由 (i)(ii)(iii) 推出.

例 1.2.3 定义

$$F(x_1, x_2) = \begin{cases} 1, & x_1 + x_2 \ge 0, \\ 0, & x_1 + x_2 < 0, \end{cases}$$

$F(x_1, x_2)$ 具有性质 (i)(ii)(iii), 但是对 $(x_1, x_2) = (-1, -1), (y_1, y_2) = (1, 1)$,

$$\begin{aligned} & F(1,1) - F(-1,1) - F(1,-1) + F(-1,-1) \\ = \ & 1 - 1 - 1 + 0 \\ = \ & -1. \end{aligned}$$

若随机向量 X 的可能值的全体至多可数, 就称 X 是离散的. 设

$$a_i \in R^n, p_i = P(X = a_i), i = 1, 2, \ldots, p_i \ge 0, \sum_i p_i = 1.$$

若 $a_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \in R^n$, 则

$$F(x_1, \ldots, x_n) = \sum_{\substack{a_{i1} \le x_1 \\ \vdots \\ a_{in} \le x_n}} p_i.$$

若存在函数 $f(x), x \in R^n$ 使 X 的分布函数为

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

则称 X 是连续型的. $f(x)$ 称为 X 的概率密度函数, 或简称密度.

保留 $k(1 \leq k < n)$ 个 x_i, 比如 $x_1, \ldots, x_k$, 而令其它 x_i 都趋于 $+\infty$, 得到 k 维边沿分布函数

$$\begin{aligned} & F(x_1, \ldots, x_k, +\infty, \ldots, +\infty) \\ = & \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(y_1, \ldots, y_n)\mathrm{d}y_1 \ldots \mathrm{d}y_k \ldots \mathrm{d}y_n, \end{aligned}$$

可见 $F(x_1, \ldots, x_k) = F(x_1, \ldots, x_k, +\infty, \ldots, +\infty)$. 也是连续型的. 其密度为

$$g(x_1, \ldots, x_k) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1, \ldots, x_n)\mathrm{d}x_{k+1} \ldots \mathrm{d}x_n.$$

注意边沿分布由分布唯一决定, 但反之不然.

例 1.2.4 两个连续型分布函数 $F(x,y)$ 和 $G(x,y)$ 的密度分别是

$$f(x,y) = \begin{cases} x+y, & 0 \leq x, y \leq 1, \\ 0, & \text{反之}, \end{cases}$$

$$g(x,y) = \begin{cases} (\frac{1}{2}+x)(\frac{1}{2}+y), & 0 \leq x, y \leq 1, \\ 0, & \text{反之}. \end{cases}$$

$f(x,y)$ 与 $g(x,y)$ 不等, 但它们的边沿密度

$$\begin{aligned} \int_{-\infty}^{\infty} f(x,y)\mathrm{d}y &= \int_0^1 (x+y)\mathrm{d}y \\ &= \frac{1}{2}+x, \quad 0 \leq x \leq 1, \\ \int_{-\infty}^{\infty} g(x,y)\mathrm{d}y &= \int_0^1 (\frac{1}{2}+x)(\frac{1}{2}+y)\mathrm{d}y \\ &= \frac{1}{2}+x, \quad 0 \leq x \leq 1 \end{aligned}$$

相同, 另一对边沿密度也相同.

独立性是概率论中重要的特征之一. 称 n 个随机变量 $X_1, X_2, \ldots, X_n$ 是相互独立的, 如果它们的联合分布函数为

$$F(x_1, \ldots, x_n) = F_{X_1}(x_1) \ldots F_{X_n}(x_n),$$

这里 $F_{X_j}(x_j) = \lim_{x_i \to \infty, i \neq j} F(x_1, \ldots, x_n)$.

命题 1.2.5 随机变量 $X_1,\dots,X_n$ 相互独立的充要条件是对任意 $A_i\in\mathcal{B},i=1,2,\dots,n$, $P(X_1\in A_1,\dots,X_n\in A_n)=\prod_{i=1}^n P(X_i\in A_i)$.

推论 1.2.6 若随机变量 $X_1,\dots,X_n$ 相互独立, f_i 为 Borel 可测函数, $i=1,2,\dots,n$. 则 $f_1(X_1),\dots,f_n(X_n)$ 也相互独立.

证明略.

上面我们仅仅就连续型情况进行说明, 对于离散型情况读者可自己写出.

若 X 是随机变量, $g(x)$ 是 Borel 可测函数, 则 $g(X)$ 也是随机变量. 事实上因为 $g(x)$ 为 Borel 可测函数, 对 $A\in\mathcal{B}, g^{-1}(A)\in\mathcal{B}$, 所以有 $\{g(X)\in A\}=\{X\in g^{-1}(A)\}\in\mathcal{F}$. 关于随机向量的变换, 仅就连续型性情形不加证明地给出.

命题 1.2.7 设 n 维随机向量 $X=(X_1,\dots,X_n)$ 的密度为 $f_X(x_1,\dots,x_n)$. 又设 n 元函数 $u_i(x_1,\dots,x_n),i=1,\dots,n$ 满足条件:

(i) 存在唯一反函数 $x_i(y_1,\dots,y_n)$, 即方程

$$\begin{cases} u_1(x_1,\dots,x_n)=y_1, \\ \quad\vdots \\ u_n(x_1,\dots,x_n)=y_n, \end{cases} \tag{2.1}$$

存在唯一实数解 $x_i=x_i(y_1,\dots,y_n),i=1,\dots,n$;

(ii) $u_i(x_1,\dots,x_n)$ 及 $x_i(y_1,\dots,y_n)$ 都连续;

(iii) $\dfrac{\partial x_i}{\partial y_j},\dfrac{\partial y_i}{\partial x_j}$ 存在且连续, $i,j=1,\dots,n$. 令

$$J=\frac{\partial(x_1,\dots,x_n)}{\partial(y_1,\dots,y_n)},$$

则 n 维随机向量 $Y=(Y_1,\dots,Y_n),Y_i=u_i(X_1,\dots,X_n),i=1,2,\dots,n$, 有密度函数为

$$f_Y(y_1,\dots,y_n)=f_X(x_1(y_1,\dots,y_n),\dots,x_n(y_1,\dots,y_n))|J|,$$

只要 $(y_1,\dots,y_n)$ 使 (2.1) 有解.

如果 (2.1) 有多个解

$$\begin{cases} x_1^{(l)}=x_1^{(l)}(y_1\dots,y_n), \\ \quad\vdots \\ x_n^{(l)}=x_n^{(l)}(y_1\dots,y_n), \end{cases}$$

$l=1,2,\dots,$ 则

$$f_Y(y_1,\dots,y_n)=\sum_l f_X(x_1^{(l)}(y_1,\dots,y_n),\dots,x_n^{(l)}(y_1,\dots,y_n))|J^{(l)}|.$$

例 1.2.8 已给 n 阶正定对称矩阵 $B, a=(a_1,\ldots,a_n), x=(x_1,\ldots,x_n)$,

$$f(x)=\frac{1}{(2\pi)^{n/2}}|B|^{-1/2}\exp\{-\frac{1}{2}(x-a)B^{-1}(x-a)'\}$$

是 n 维随机变量的密度. 式中 $|B|$ 表示 B 的行列式的值, C' 表示矩阵 C 的转置矩阵, B^{-1} 表示矩阵 B 的逆矩阵. 下面要证明 $\int_{R^n} f(x)\mathrm{d}x=1$.

因为矩阵 B 对称正定, 故存在正交阵 T, 使

$$TBT'=D=\begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots 0 & \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & d_n \end{pmatrix},$$

其中 d_i 是 B 的特征值并且 $d_i>0$.

作变换 $y=(x-a)T'$, 右乘 T 注意 $T'T=I$, 可得 $x=a+yT$. 因为 $\dfrac{\partial(x)}{\partial(y)}=|T|=\pm1$,

$$\begin{aligned}
|B| &= |T'TBT'T|=|T'DT| \\
&= |T'T||D|=d_1\ldots d_n, \\
(x-a)B^{-1}(x-a)' &= yTB^{-1}T'y'=y(T')^{-1}B^{-1}T^{-1}y'=y(TBT')^{-1}y' \\
&= \sum_{i=1}^{n}\frac{y_i^2}{d_i},
\end{aligned}$$

所以

$$\begin{aligned}
\int_{R^n} f(x)\mathrm{d}x &= \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\frac{1}{(2\pi)^{n/2}\sqrt{d_1\ldots d_n}}\exp\left\{-\sum_{i=1}^{n}\frac{y_i^2}{2d_i}\right\}\mathrm{d}y_1\ldots\mathrm{d}y_n \\
&= \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi d_i}}\int_{-\infty}^{+\infty}\exp\left\{-\frac{y_i^2}{2d_i}\right\}\mathrm{d}y_i \\
&= 1.
\end{aligned}$$

$f(x)$ 是 n 维正态分布的密度函数.

例 1.2.9 对事件 A, 定义一个随机变量

$$I_A(\omega)=\begin{cases} 1, & \omega\in A, \\ 0, & \omega\notin A, \end{cases}$$

称 I_A 为 A 的示性函数. 由 $P(A)=P(I_A=1)$, 知事件的概率可以由该事件的示性函数取值的概率表示, 所以随机变量的概念比事件的概念更一般.

§1.3 随机变量的数字特征

随机变量 X 的数学期望用 EX 表示，定义为

$$\begin{aligned} EX &= \int_{-\infty}^{+\infty} x\mathrm{d}F_X(x) \\ &= \begin{cases} \int_{-\infty}^{+\infty} xf(x)\mathrm{d}x, & \text{如}X\text{是连续型}, \\ \sum_x xP(X=x), & \text{如}X\text{是离散型}, \end{cases} \end{aligned} \tag{3.1}$$

只要上面的积分是存在的. X 的数学期望也可简称为期望或均值. 我们也可以写

$$EX = \int_\Omega XP(\mathrm{d}\omega). \tag{3.2}$$

由积分转化定理，(3.2) 可以转化为 (3.1). (3.2) 表示求 X 的期望，就是求 X 的 (抽象) 积分. 粗略地说 EX 表示大多数 X 的值的集中位置.

若 $y=h(x)$ 是 Borel 函数，则 $Y=h(X)$ 也是随机变量，它的分布函数记为 $F_Y(x)$，则由 (3.1), $EY=\int_{-\infty}^{+\infty} x\mathrm{d}F_Y(x)$. 可以证明

$$EY = Eh(X) = \int_\Omega h(X)P(\mathrm{d}\omega) = \int_{-\infty}^{+\infty} h(x)\mathrm{d}F_X(x), \tag{3.3}$$

(3.3) 式的意义是若一方积分存在，则另一方积分也存在且相等.

将 $h(x)$ 特殊化可以得到其他几个数字特征 ($k \ge 0$):

- 令 $h(x)=x^k, E(X^k)$ 称为 X 的 k 阶矩.
- 令 $h(x)=|x|^k$, 称 $E(|X|^k)$ 为 X 的 k 阶绝对矩.
- 令 $h(x)=(x-EX)^k, E[(X-EX)^k]$ 称为 X 的 k 阶中心矩.

X 的 2 阶中心矩也称为 X 的方差，记为 DX 或 $\mathrm{Var}X$，即

$$DX = E(X-EX)^2 = EX^2-(EX)^2.$$

对于随机向量的情形，$(X_1,\ldots,X_n)$ 是 n 维随机向量，分布函数为 $F(x_1,\ldots,x_n)$, $g(x_1,\ldots,x_n)$ 为 n 维 Borel 函数，则

$$Eg(X_1,\ldots,X_n) = \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} g(x_1,\ldots,x_n)\mathrm{d}F(x_1,\ldots,x_n),$$

只要右方的积分绝对收敛.

n 维随机向量 $X=(X_1,\ldots,X_n)$，如 $EX_i, i=1,\ldots,n$，存在，称向量 $(EX_1,\ldots,EX_n)$ 为 X 的数学期望，记为

$$EX = (EX_1,\ldots,EX_n).$$