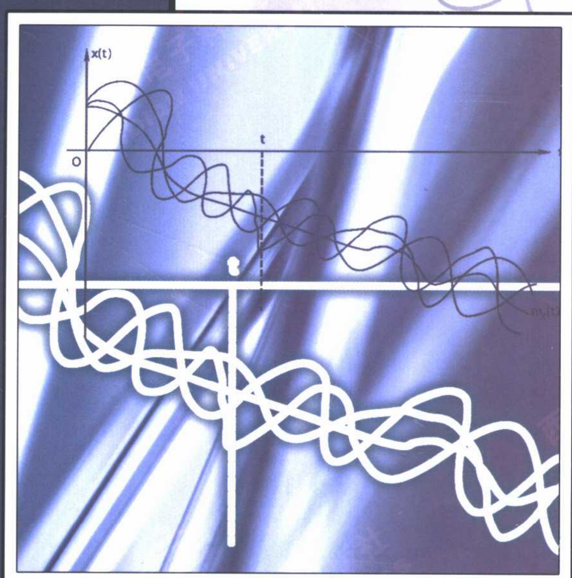
研究生系列教材

随机过程



张卓奎
陈慧婵
编著

西安电子科技大学出版社
http://www.xduph.com

研究生系列教材

随机过程

张卓奎　陈慧婵　编著

西安电子科技大学出版社
2003

内 容 简 介

本书系研究生系列教材之一，是根据工科研究生对学习随机过程的要求而编写的．其内容包括概率论基础、随机过程的基本概念、随机分析、平稳过程、马尔可夫过程、排队和服务系统、更新过程、时间序列分析、鞅过程等．各章后均配有习题．

本书内容简练，通俗易懂，凡具有高等数学基础和工科概率论基础的读者均可阅读．

本书可作为工科研究生和本科高年级学生的教材或教学参考书，也可作为理工科师生和工程技术人员的参考书．

图书在版编目(CIP)数据

随机过程/张卓奎，陈慧婵编著．—西安：西安电子科技大学出版社，2003.9

(研究生系列教材)

ISBN 7-5606-1282-2

Ⅰ．随…　Ⅱ．①张…　②陈…　Ⅲ．随机过程-研究生-教材　Ⅳ．O211.6

中国版本图书馆 CIP 数据核字(2003)第 058671 号

责任编辑　夏大平

出版发行　西安电子科技大学出版社(西安市太白南路 2 号)

电　　话　(029)8242885　8201467　　邮　　编　710071

http://www.xduph.com　　E-mail：xdupfxb@pub.xaonline.com

经　　销　新华书店

印刷单位　西安文化彩印厂

版　　次　2003 年 9 月第 1 版　2003 年 9 月第 1 次印刷

开　　本　787 毫米×1092 毫米　1/16　印张 20.75

字　　数　491 千字

印　　数　1～4000 册

定　　价　25.00 元

ISBN 7-5606-1282-2/O·0066(课)

XDUP 1553001-1

西安电子科技大学
研究生教材建设基金资助

前　言

随机过程是研究随时间演变的随机现象的一门学科．它以概率论为基础，并且是概率论的深入和发展，随着科学技术的发展，它已被广泛地应用到雷达与通信、动态可靠性、自动控制、生物工程、社会科学以及其它工程科学等领域，并且在这些领域显示出十分重要的作用．

目前，在高等院校中，很多工科专业的研究生都要学习随机过程．为了适应不同专业研究生对学习随机过程的需要，编者在多次讲授本课程讲稿的基础上，结合同行专家的优秀成果，经过多次修订和补充写成本书．

在编写过程中，我们有意注重基本理论、基本概念和基本方法的叙述，关注能力的培养．在数学工具的使用上，力求准确、简明和适中，尽量使读者用浅显的数学工具准确、系统地认识和掌握随机过程的基本理论和方法．这样处理既保持了随机过程的数学体系和必要的严密性，又尽可能地结合了工科研究生的知识结构和专业应用．

全书共 9 章．第 1 章为学习随机过程作准备，介绍了概率论基础；第 2 章介绍了随机过程的基本概念和几种常用的随机过程；第 3 章为随机分析，主要讨论二阶矩过程的均方微积分学；第 4 章讨论平稳过程，重点介绍平稳过程的相关函数的性质、各态历经性、功率谱密度、谱分解以及线性系统中的平稳过程；第 5 章为马尔可夫过程，主要介绍马尔可夫链及状态离散参数连续马尔可夫过程的转移概率，状态分类和平稳分布；第 6 章介绍排队和服务系统，讨论了几种简单但又常用的排队系统；第 7 章为更新过程；第 8 章介绍了应用极为广泛的时间序列分析，主要包括线性模型及其性质，参数估计，预报及 Kalman 滤波；鞅过程在第 9 章中介绍．

本书在编写的过程中，得到了西安电子科技大学研究生院副院长曾兴雯教授，研究生院培养办公室顾国其、马丽莎同志的热情支持和帮助；西安电子科技大学出版社的领导和策划编辑夏大平也非常关心此书的出版，夏亲自担任此书的责任编辑，并对该书的出版付出了辛勤劳动，编者在此一并致以诚挚的谢意！

由于编者水平有限，书中难免存在错误和不妥之处，恳请读者批评、指正．

编　者

2003 年 5 月

目　录

第 1 章　概率论基础

概率论中基本概念和基本理论是随机过程的基础．本章扼要地复习概率论中某些基本概念，并补充某些工程概率论中未讲授的内容，为学习随机过程作准备．

1.1　概率空间

概率论中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试验．一个试验所有可能出现的结果的全体称为随机试验的样本空间，记为 Ω，试验的一个结果称为样本点，记为 ω，即 $\Omega=\{\omega\}$．样本空间的某个子集称为随机事件，简称事件．因为事件是集合，所以集合的运算与事件的运算是一致的．根据实际情况，我们并不总是对 Ω 的一切子集有兴趣研究，而是对某些事件类感兴趣，于是便导致了事件域的概念．

定义 1.1.1　设 Ω 是样本空间，$\mathscr{F}$ 是 Ω 的某些子集构成的集合，如果：

(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 $\bigcup\limits_{n=1}^{\infty}A_n\in\mathscr{F}$，

那么称 $\mathscr{F}$ 为一事件域．也称 $\mathscr{F}$ 为 σ 域．

显然，如果 $\mathscr{F}$ 是一事件域，那么

(1) $\varnothing\in\mathscr{F}$；

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

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

定义 1.1.2　设 Ω 是样本空间，$\mathscr{F}$ 是一事件域，定义在 $\mathscr{F}$ 上的实值函数 $P(\cdot)$如果满足：

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

(2) $P(\Omega)=1$，

(3) 若 $A_n\in\mathscr{F}$，$n=1, 2, \cdots$，且 $A_iA_j=\varnothing$，$i\neq j$，i，$j=1, 2, \cdots$，则

$$P\Big(\bigcup_{n=1}^{\infty}A_n\Big)=\sum_{n=1}^{\infty}P(A_n)$$

那么称 P 是二元组$(\Omega, \mathscr{F})$上的概率，称 $P(A)$为事件 A 的概率，称三元组$(\Omega, \mathscr{F}, P)$为

概率空间.

关于事件的概率具有如下性质：

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

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

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

(3) 若 A，$B\in\mathscr{F}$，$A\subset B$，则 $P(B-A)=P(B)-P(A)$;

(4) 若 A，$B\in\mathscr{F}$，$A\subset B$，则 $P(A)\leqslant P(B)$;

(5) 若 $A\in\mathscr{F}$，则 $P(A)\leqslant 1$;

(6) 若 $A\in\mathscr{F}$，则 $P(\overline{A})=1-P(A)$;

(7) 若 $A_n\in\mathscr{F}$，$n=1, 2, \cdots$，则

$$P(\bigcup_{n=1}^{\infty}A_i)\leqslant\sum_{n=1}^{\infty}P(A_n)$$

(8) 若 $A_i\in\mathscr{F}$，$i=1, 2, \cdots, n$，则

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

一列事件 $A_n\in\mathscr{F}$，$n=1, 2, \cdots$ 称为单调递增的事件列，如果 $A_n\subset A_{n+1}$，$n=1,2,\cdots$；一列事件 $A_n\in\mathscr{F}$，$n=1,2,\cdots$ 称为单调递减的事件列，如果 $A_n\supset A_{n+1}$，$n=1,2,\cdots$.

定理 1.1.1 设 $A_n\in\mathscr{F}$，$n=1,2,\cdots$

(1) 若 A_n，$n=1,2,\cdots$ 是单调递增的事件列，则

$$\lim_{n\to\infty}P(A_n)=P(\bigcup_{n=1}^{\infty}A_n)$$

(2) 若 A_n，$n=1,2,\cdots$ 是单调递减的事件列，则

$$\lim_{n\to\infty}P(A_n)=P(\bigcap_{n=1}^{\infty}A_n)$$

证明 (1) 令 $B_1=A_1$，$B_n=A_n\overline{\bigcup_{i=1}^{n-1}A_i}=A_n\overline{A}_{n-1}$，$n=2,3,\cdots$，则 $B_iB_j=\varnothing$，$i\neq j$，$i,j=1,2,\cdots$，且 $\bigcup_{n=1}^{\infty}B_n=\bigcup_{n=1}^{\infty}A_n$，于是

$$P(\bigcup_{n=1}^{\infty}A_n)=P(\bigcup_{n=1}^{\infty}B_n)=\sum_{n=1}^{\infty}P(B_n)=\lim_{n\to\infty}\sum_{k=1}^{n}P(B_k)$$
$$=\lim_{n\to\infty}P(\bigcup_{k=1}^{n}B_k)=\lim_{n\to\infty}P(\bigcup_{k=1}^{n}A_k)=\lim_{n\to\infty}P(A_n)$$

即

$$\lim_{n\to\infty}P(A_n)=P(\bigcup_{n=1}^{\infty}A_n)$$

(2) 若 A_n，$n=1,2,\cdots$ 单调递减，则 $\overline{A}_n$，$n=1,2,\cdots$ 单调递增，由(1)得

$$\lim_{n\to\infty}P(\overline{A}_n)=P(\bigcup_{n=1}^{\infty}\overline{A}_n)$$

而

$$P(\bigcup_{n=1}^{\infty}\overline{A}_n)=P(\overline{\bigcap_{n=1}^{\infty}A_n})=1-P(\bigcap_{n=1}^{\infty}A_n)$$

$$\lim_{n\to\infty} P(\overline{A}_n) = \lim_{n\to\infty}(1 - P(A_n)) = 1 - \lim_{n\to\infty} P(A_n)$$

所以

$$\lim_{n\to\infty} P(A_n) = P(\bigcap_{n=1}^{\infty} A_n)$$

定义 1.1.3 设$(\Omega, \mathscr{F}, P)$为一概率空间，$A,B\in\mathscr{F}$，且$P(A)>0$，则称

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

为在事件 A 发生的条件下事件 B 发生的条件概率.

不难验证，条件概率 $P(\cdot|A)$符合定义 1.1.2 中的三个条件，即

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

(2) $P(\Omega|A)=1$；

(3) 设 $B_n\in\mathscr{F}$，$n=1,2,\cdots$，$B_iB_j=\varnothing$，$i\neq j$，$i,j=1,2,\cdots$，则

$$P(\bigcup_{n=1}^{\infty} B_n|A) = \sum_{n=1}^{\infty} P(B_n|A)$$

既然条件概率符合上述三个条件，故对概率所证明的一些重要结果都适用于条件概率.

定理 1.1.2 设$(\Omega, \mathscr{F}, P)$是一概率空间，有：

(1)（乘法公式） 若 $A_i\in\mathscr{F}$，$i=1,2,\cdots,n$，且 $P(A_1A_2\cdots A_{n-1})>0$，则

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

(2)（全概率公式） 设 $A\in\mathscr{F}$，$B_i\in\mathscr{F}$，$P(B_i)>0$，$i=1,2,\cdots$，且 $B_iB_j=\varnothing$，$i\neq j$，$i,j=1,2,\cdots$，$\bigcup_{i=1}^{\infty}B_i\supset A$，则

$$P(A) = \sum_{i=1}^{\infty} P(B_i)P(A|B_i)$$

(3)（贝叶斯(Bayes)公式） 设 $A\in\mathscr{F}$，$P(A)>0$，$B_i\in\mathscr{F}$，$P(B_i)>0$，$i=1,2,\cdots$，且 $B_iB_j=\varnothing$，$i,j=1,2,\cdots$，$\bigcup_{i=1}^{\infty}B_i\supset A$，则

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{j=1}^{\infty} P(B_j)P(A|B_j)},\ i = 1,2,\cdots$$

定理 1.1.2 的证明请读者自己给出.

定义 1.1.4 设$(\Omega, \mathscr{F}, P)$为一概率空间，$A_i\in\mathscr{F}$，$i=1,2,\cdots,n$，如果对于任意的 $k(1<k\leqslant n)$及任意的 $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$，有

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

则称 $A_1,A_2,\cdots,A_n$ 相互独立.

定理 1.1.3 设 A，$B\in\mathscr{F}$ 相互独立，则 A 与 $\overline{B}$，$\overline{A}$ 与 B，A 与 $\overline{B}$ 也是相互独立的，从而 A 所生成的 σ 域 $\mathscr{F}_A=\{A, \overline{A}, \varnothing, \Omega\}$中的任意一个事件和 B 所生成的 σ 域 $\mathscr{F}_B=\{B, \overline{B}, \varnothing, \Omega\}$中的任意一个事件都相互独立（这时我们称这两个 σ 域 $\mathscr{F}_A$ 和 $\mathscr{F}_B$ 是相互独立的）.

证明 由于 $A=(AB)\cup(A\overline{B})$，且$(AB)(A\overline{B})=\varnothing$，因此 $P(A)=P(AB)+P(A\overline{B})$，于是

$$\begin{aligned}P(A\bar{B}) &= P(A) - P(AB) \\ &= P(A) - P(A)P(B) \\ &= P(A)(1 - P(B)) \\ &= P(A)P(\bar{B})\end{aligned}$$

即 A 与 $\bar{B}$ 相互独立. 类似地可以证明 $\bar{A}$ 与 B, $\bar{A}$ 与 $\bar{B}$ 也相互独立, 从而 $\mathscr{F}_A$ 与 $\mathscr{F}_B$ 相互独立.

定理 1.1.4 设 $A, B, C\in\mathscr{F}$ 相互独立, 则

(1) A 与 BC 相互独立;

(2) A 与 $B\cup C$ 相互独立;

(3) A 与 $B-C$ 相互独立;

(4) A 所生成的 σ 域中的任一事件与 B 和 C 所生成的 σ 域 $\mathscr{F}_{B,C}=\{B, \bar{B}, C, \bar{C}, BC, B\bar{C}, C\bar{B}, \bar{B}\bar{C}, (C\bar{B}\cup B\bar{C}), (BC\cup\bar{B}\bar{C}), B\cup C, \bar{B}\cup C, \bar{B}\cup\bar{C}, B\cup\bar{C}, \Omega, \varnothing\}$中的任意一个事件都相互独立.

证明 (1) 显然.

(2) 由于

$$\begin{aligned}P(A(B\cup C)) &= P(AB\cup AC) \\ &= P(AB) + P(AC) - P(ABC) \\ &= P(A)P(B) + P(A)P(C) - P(A)P(B)P(C) \\ &= P(A)[P(B) + P(C) - P(B)P(C)] \\ &= P(A)[P(B) + P(C) - P(BC)] \\ &= P(A)P(B\cup C)\end{aligned}$$

故 A 与 $B\cup C$ 相互独立.

(3) 与(2)证法类似.

(4) 只需证明 A 与 $B\bar{C}$, $\bar{B}C$, $\bar{B}\bar{C}$ 相互独立. 因为

$$\begin{aligned}P(A(B\bar{C})) &= P(AB\bar{C}) \\ &= P(AB) - P(ABC) \\ &= P(A)P(B) - P(A)P(B)P(C) \\ &= P(A)[P(B) - P(B)P(C)] \\ &= P(A)[P(B) - P(BC)] \\ &= P(A)P(B\bar{C})\end{aligned}$$

所以 A 与 $B\bar{C}$ 相互独立. 其余可以利用加法公式类似地得到.

推论 1.1.1 设 $A, B, C\in\mathscr{F}$ 相互独立, 将 A, B, C 任意分为两组, 则它们各自生成的 σ 域仍然相互独立.

证明 直接由定理 4 推得.

更一般地我们有以下定理.

定理 1.1.5 设 $A_i\in\mathscr{F}$, $i=1, 2, \cdots, n$ 相互独立, 将 A_i, $i=1, 2, \cdots, n$ 任意分成 $m(m\leqslant n)$组, 并对各组中的事件施以积、和、逆运算以后, 所得到的事件 $B_1, B_2, \cdots, B_m$ 也相互独立. 从而这 m 组事件各自所生成的 σ 域也是相互独立的.

证明 与定理 1.1.4 证明类似, 请读者自己给出.

定理 1.1.5 还蕴含了以下有用的具体结论：

(1) 若 A_1，A_2，…，A_n 相互独立，则 $\overline{A}_1$，$\overline{A}_2$，…，$\overline{A}_n$也相互独立，从而有

$$\begin{aligned} P(\bigcup_{i=1}^{n} A_i) &= 1 - P(\overline{\bigcup_{i=1}^{n} A_i}) \\ &= 1 - P(\bigcap_{i=1}^{n} \overline{A}_i) \\ &= 1 - P(\overline{A}_1)P(\overline{A}_2)\cdots P(\overline{A}_n) \end{aligned}$$

(2) 一列独立事件中的任何一部分事件也相互独立.

(3) 若一列事件相互独立，则将其中任一部分改写为对立事件，所得的事件列也相互独立.

1.2　随机变量及其分布

随机变量是概率论的主要研究对象. 随机变量的取值依赖于试验结果或样本点，每次试验之后，其取值是一个实数或直线上的一个点. 描述随机变量的分布用分布函数.

定义 1.2.1　设$(\Omega,\mathscr{F},P)$为一概率空间，定义在 Ω 上的实函数 $X(\cdot)$，如果 $\forall\, x\in\mathbf{R}$，$\{\omega|X(\omega)\leqslant x\}\in\mathscr{F}$，则称 X 是 $\mathscr{F}$ 的随机变量. 称

$$F(x) = P(X \leqslant x),\ -\infty < x < +\infty$$

为随机变量 X 的分布函数.

设 X 是 $\mathscr{F}$ 的随机变量，则不难证明 X 的分布函数 $F(x)$具有如下性质：

(1) $F(x)$是单调不减函数，即若 $x_1<x_2$，则 $F(x_1)\leqslant F(x_2)$；

(2) $F(x)$是右连续函数，即$\forall\, x\in\mathbf{R}$，$F(x+0)=F(x)$；

(3) $F(-\infty)\overset{\text{def}}{=}\lim\limits_{x\to-\infty}F(x)=0$，$F(+\infty)\overset{\text{def}}{=}\lim\limits_{x\to+\infty}F(x)=1$.

同时还可以证明，设 $F(x)$，$x\in\mathbf{R}$ 是单调不减、右连续的函数，并且 $F(-\infty)=0$，$F(+\infty)=1$，则必存在概率空间$(\Omega,\mathscr{F},P)$及其上的一个随机变量 X，使得 X 以 $F(x)$为其分布函数.

在实际应用中，常见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.

若随机变量 X 的可能取值为有限个或可列无限个，则称 X 为离散型随机变量. 离散型随机变量 X 的分布可用分布律来描述，即

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

这时 X 的分布函数为

$$F(x) = \sum_{x_i\leqslant x} p_i,\ x \in \mathbf{R}$$

设随机变量 X 的分布函数为 $F(x)$，如果存在非负可积函数 $f(x)$，$x\in\mathbf{R}$，使得

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

则称 X 为连续型随机变量，$f(x)$称为连续型随机变量 X 的概率密度函数.

定义 1.2.2 设$(\Omega, \mathscr{F}, P)$为一概率空间，定义在Ω上的n元实函数$X(\cdot)=(X_1(\cdot), X_2(\cdot), \cdots, X_n(\cdot))$，如果$\forall\ \boldsymbol{x}=(x_1, x_2, \cdots, x_n)\in\mathbf{R}^n$，$\{\omega \mid X_1(\omega)\leqslant x_1, X_2(\omega)\leqslant x_2, \cdots, X_n(\omega)\leqslant x_n\}\in\mathscr{F}$，则称$\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$为$n$维随机变量或$n$维随机向量．称

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

为$\boldsymbol{X}$的联合分布函数．

设$\boldsymbol{X}$是n维随机变量，则不难证明$\boldsymbol{X}$的联合分布函数具有下列性质：

(1) $F(x_1, x_2, \cdots, x_n)$对任一$x_i(i=1,2,\cdots,n)$是单调不减函数；

(2) $F(x_1, x_2, \cdots, x_n)$对任一$x_i(i=1,2,\cdots,n)$是右连续函数；

(3) $F(x_1, \cdots, x_{i-1}, -\infty, x_{i+1}, \cdots, x_n)=0$，$i=1,2,\cdots,n$，
$F(+\infty, +\infty, \cdots, +\infty)=1$；

(4) 设$x_i\leqslant y_i$，$i=1,2,\cdots,n$，则

$$\begin{aligned}&F(y_1,y_2,\cdots,y_n) - \sum_{i=1}^{n}F(y_1,y_2,\cdots,y_{i-1},x_i,y_{i+1},\cdots,y_n)\\&+\sum_{1\leqslant i<j\leqslant n}F(y_1,y_2,\cdots,y_{i-1},x_i,y_{i+1},\cdots,y_{j-1},x_j,y_{j+1},\cdots,y_n)\\&-\cdots+(-1)^nF(x_1,x_2,\cdots,x_n)\geqslant 0\end{aligned}$$

类似于一维随机变量，可以证明，对于给定的n元函数$F(x_1, x_2, \cdots, x_n)$若满足上面的性质(1)、(2)、(3)、(4)，则必存在概率空间$(\Omega, \mathscr{F}, P)$及其上的n维随机变量$\boldsymbol{X}$，使得$\boldsymbol{X}$以$F(x_1, x_2, \cdots, x_n)$为其联合分布函数．

对于n维随机变量，在实际应用中常见的也有两种类型：离散型n维随机变量和连续型n维随机变量．

若n维随机变量X的可能取值为有限对或可列无限对，则称n维随机变量X为离散型n维随机变量．离散型n维随机变量$\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$的分布可用联合分布律来描述，即

$$P(X_1=x_1,\ X_2=x_2,\ \cdots,\ X_n=x_n)$$

其中$x_i\in I_i$，I_i是离散集，$i=1,2,\cdots,n$．这时X的联合分布函数为

$$F(u_1,u_2,\cdots,u_n) = \sum_{x_1\leqslant u_1}\sum_{x_2\leqslant u_2}\cdots\sum_{x_n\leqslant u_n}P(X_1=x_1,X_2=x_2,\cdots,X_n=x_n)$$

设n维随机变量$\boldsymbol{X}$的联合分布函数为$F(x_1,x_2,\cdots,x_n)$，如果存在非负可积函数$f(x)=f(x_1,x_2,\cdots,x_n)$，$\boldsymbol{x}\in\mathbf{R}^n$，使得

$$F(u_1,u_2,\cdots,u_n) = \int_{-\infty}^{u_1}\int_{-\infty}^{u_2}\cdots\int_{-\infty}^{u_n}f(x_1,x_2,\cdots,x_n)\,\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n$$

则称$\boldsymbol{X}$为连续型n维随机变量，$f(x_1, x_2, \cdots, x_n)$称为连续型n维随机变量$\boldsymbol{X}$的联合概率密度函数．

保留$k(1\leqslant k<n)$个x_i，比如$x_1, x_2, \cdots, x_k$，而令其它x_i都趋于$+\infty$，得到k维边缘分布函数

$$F(x_1,\ x_2,\ \cdots,\ x_k,\ +\infty,\ \cdots,\ +\infty)$$

若 $\boldsymbol{X}$ 是连续型 n 维随机变量，则有

$$
\begin{aligned}
&F(x_1,x_2,\cdots,x_k,+\infty,\cdots,+\infty)\\
&=\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_k}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(y_1,y_2,\cdots,y_n)\,\mathrm{d}y_1\cdots\mathrm{d}y_k\cdots\mathrm{d}y_n
\end{aligned}
$$

可见 $F(x_1,x_2,\cdots,x_k)=F(x_1,x_2,\cdots,x_k,+\infty,\cdots,+\infty)$ 也是连续型 k 维随机变量的联合分布函数，其联合概率密度函数为

$$
g(x_1,x_2,\cdots,x_k)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f(x_1,x_2,\cdots,x_n)\,\mathrm{d}x_{k+1}\cdots\mathrm{d}x_n
$$

特别地，当 $k=1$ 时，n 维随机变量 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$ 的 n 个边缘分布函数和 n 个边缘概率密度函数分别为 $F_{X_1}(x_1)$，$F_{X_2}(x_2)$，$\cdots$，$F_{X_n}(x_n)$ 和 $f_{X_1}(x_1)$，$f_{X_2}(x_2)$，$\cdots$，$f_{X_n}(x_n)$.

定义 1.2.3 设 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$ 是一 n 维随机变量，其联合分布函数和边缘分布函数分别为 $F(x_1, x_2, \cdots, x_n)$，$F_{X_1}(x_1)$，$F_{X_2}(x_2)$，$\cdots$，$F_{X_n}(x_n)$，如果对于任意的 $x_1, x_2, \cdots, x_n\in\mathbf{R}$，有

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

则称随机变量 $X_1, X_2, \cdots, X_n$ 相互独立.

若 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$ 是离散型 n 维随机变量，则 $X_1, X_2, \cdots, X_n$ 相互独立的充要条件是

$$
P(X_1=x_1, X_2=x_2, \cdots, X_n=x_n)=P(X_1=x_1)P(X_2=x_2)\cdots P(X_n=x_n)
$$

其中 x_i 是 X_i，$i=1,2,\cdots,n$ 的所有可能取值.

若 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$ 是连续型 n 维随机变量，则 $X_1, X_2, \cdots, X_n$ 相互独立的充要条件是

$$
f(x_1,x_2,\cdots,x_n)=f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n),\quad x_1,x_2,\cdots,x_n\in\mathbf{R}
$$

如果 X 是随机变量，$g(x)$ 是已知的连续函数，则 $g(X)$ 也是随机变量，关于 $g(X)$（可以是一维也可以是多维）的分布，仅就连续情况不加证明给出.

定理 1.2.1 设连续型 n 维随机变量 $\boldsymbol{X}=(X_1, X_2, \cdots, X_n)$ 的联合概率密度函数为 $f_X(x_1, x_2, \cdots, x_n)$，$n$ 元函数 $y_i=y_i(x_1, x_2, \cdots, x_n)$，$i=1,2,\cdots,n$，满足：

(1) 存在唯一的反函数 $x_i=x_i(y_1, y_2, \cdots, y_n)$，即方程组

$$
\begin{cases}
y_1(x_1, x_2, \cdots, x_n)=y_1\\
y_2(x_1, x_2, \cdots, x_n)=y_2\\
\quad\vdots\\
y_n(x_1, x_2, \cdots, x_n)=y_n
\end{cases}
$$

存在唯一的实数解 $x_i=x_i(y_1, y_2, \cdots, y_n)$，$i=1,2,\cdots,n$；

(2) $y_i=y_i(x_1, x_2, \cdots, x_n)$ 及 $x_i=x_i(y_1, y_2, \cdots, y_n)$，$i=1,2,\cdots,n$ 都是连续的；

(3) $\dfrac{\partial x_i}{\partial y_j}$，$\dfrac{\partial y_i}{\partial x_j}$，$i,j=1,2,\cdots,n$ 存在且连续，令

$$
J=\frac{\partial(x_1, x_2, \cdots, x_n)}{\partial(y_1, y_2, \cdots, y_n)}
$$

则 n 维随机变量 $\boldsymbol{Y}=(Y_1, Y_2, \cdots, Y_n)$，$Y_i=y_i(X_1, X_2, \cdots, X_n)$，$i=1,2,\cdots,n$ 的联合概率密度函数为

$$f_Y(y_1, y_2, \cdots, y_n) = f_X(x_1(y_1, y_2, \cdots, y_n),\ x_2(y_1, y_2, \cdots, y_n), \cdots, x_n(y_1, y_2, \cdots, y_n))\,|J|$$

如果(1)中的方程组有多个解

$$\begin{cases} x_1^{(l)} = x_1^{(l)}(y_1,\ y_2,\ \cdots,\ y_n) \\ x_2^{(l)} = x_2^{(l)}(y_1,\ y_2,\ \cdots,\ y_n) \\ \qquad\vdots \\ x_n^{(l)} = x_n^{(l)}(y_1,\ y_2,\ \cdots,\ y_n) \end{cases}$$

$l=1,2,\cdots$，则 n 维随机变量 $\boldsymbol{Y}=(Y_1,\ Y_2,\ \cdots,\ Y_n)$的联合概率密度函数为

$$\begin{aligned} &f_Y(y_1, y_2, \cdots, y_n) \\ &= \sum_l f_X(x_1^{(l)}(y_1,\ y_2,\ \cdots,\ y_n),\ x_2^{(l)}(y_1,\ y_2,\ \cdots,\ y_n),\ \cdots,\ x_n^{(l)}(y_1,\ y_2,\ \cdots,\ y_n))\,|J| \end{aligned}$$

特别地，若 X 为连续型一维随机变量，其概率密度函数为 $f_X(x)$，则对于 $Y=g(X)$的概率密度函数，有下列结果.

(1) 若 $g(x)$是严格单调可微函数，则 $Y=g(X)$的概率密度函数为

$$f_Y(y) = \begin{cases} f_X(h(y))\,|h'(y)|, & y \in I \\ 0, & y \notin I \end{cases}$$

其中 $h(y)$是 $y=g(x)$的反函数，I 是使 $h(y)$有定义，$h'(y)$有定义及 $f_X(h(y))>0$ 的 y 的取值的公共部分.

(2) 若 $g(x)$不是严格单调的可微函数，则将 $g(x)$在其定义域分成若干个单调分支，在每个单调分支上应用(1)的结果得 $Y=g(X)$的概率密度函数为

$$f_Y(y) = \begin{cases} f_X(h_1(y))\,|h_1'(y)| + f_X(h_2(y))\,|h_2'(y)| + \cdots, & y \in I \\ 0\ , & y \notin I \end{cases}$$

其中 I 是在每个单调分支上按照(1)确定的 y 的取值的公共部分.

例 1.2.1　设 $X \sim U\left[-\dfrac{\pi}{2},\ \dfrac{\pi}{2}\right]$，$Y=\tan X$，试求 Y 的概率密度函数 $f_Y(y)$.

解　由于 $y=\tan x$，故其反函数 $h(y)=\arctan y$，$-\infty<y<+\infty$，并且 $h'(y)=\dfrac{1}{1+y^2}$，$-\infty<y<+\infty$，因此 Y 的概率密度函数

$$f_Y(y) = \frac{1}{\pi}\,\frac{1}{1+y^2},\ -\infty < y < +\infty$$

例 1.2.2　设 $X\sim N(0,\ 1)$，试求 $Y=X^2$ 的概率密度函数 $f_Y(y)$.

解　由于 $y=x^2$ 有两个单调分支，其反函数分别为 $h_1(y)=-\sqrt{y}$，$y\geqslant 0$，$h_2(y)=\sqrt{y}$，$y\geqslant 0$，并且 $h_1'(y)=-\dfrac{1}{2\sqrt{y}}$，$y>0$，$h_2'(y)=\dfrac{1}{2\sqrt{y}}$，$y>0$，因而 $Y=X^2$ 的概率密度函数为

$$f_Y(y) = \begin{cases} f_X(h_1(y))\,|h_1'(y)| + f_X(h_2(y))\,|h_2'(y)| = \dfrac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} \mathrm{e}^{-\frac{y}{2}} & ,\ y > 0 \\ 0 & ,\ y \leqslant 0 \end{cases}$$

例 1.2.3　设(X,Y)为二维随机变量，其中 X，Y 相互独立并且都服从正态分布 $N(0,\ \sigma^2)$，记 Z 为$(X,\ Y)$的模，Θ 为$(X,\ Y)$的辐角，求$(Z,\ \Theta)$的联合概率密度函数及边缘概率密度函数.

解　由于

$$X \sim N(0, \sigma^2),\ f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}},\ -\infty < x < +\infty$$

$$Y \sim N(0, \sigma^2),\ f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}},\ -\infty < y < +\infty$$

X，Y 相互独立，因此

$$f(x, y) = f_X(x) f_Y(y) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}(x^2+y^2)},\ -\infty < x < +\infty,\ -\infty < y < +\infty$$

又因为方程组

$$\begin{cases} z = \sqrt{x^2 + y^2} \\ \theta = \arctan \dfrac{y}{x} \end{cases}$$

有唯一解(反函数)

$$\begin{cases} x = z\cos\theta \\ y = z\sin\theta \end{cases},\quad z > 0,\ -\pi < \theta < \pi$$

$$J = \frac{\partial(x,y)}{\partial(z,\theta)} = \begin{vmatrix} \dfrac{\partial x}{\partial z} & \dfrac{\partial x}{\partial \theta} \\ \dfrac{\partial y}{\partial z} & \dfrac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -z\sin\theta \\ \sin\theta & z\cos\theta \end{vmatrix} = z$$

所以(Z, Θ)的联合概率密度函数为

$$g(z, \theta) = f(z\cos\theta, z\sin\theta)\,|J| = \frac{z}{2\pi\sigma^2} e^{-\frac{z^2}{2\sigma^2}},\ z > 0,\ -\pi < \theta < \pi$$

故

$$g(z, \theta) = \begin{cases} \dfrac{z}{2\pi\sigma^2} e^{-\frac{z^2}{2\sigma^2}}, & z > 0,\ -\pi < \theta < \pi \\ 0, & \text{其它} \end{cases}$$

$$g_Z(z) = \int_{-\infty}^{+\infty} g(z,\theta)\,d\theta = \begin{cases} \displaystyle\int_{-\pi}^{\pi} \frac{z}{2\pi\sigma^2} e^{-\frac{z^2}{2\sigma^2}} d\theta = \frac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}}, & z > 0 \\ 0, & z \leqslant 0 \end{cases}$$

$$g_\Theta(\theta) = \int_{-\infty}^{+\infty} g(z,\theta)\,dz = \begin{cases} \displaystyle\int_{0}^{+\infty} \frac{z}{2\pi\sigma^2} e^{-\frac{z^2}{2\sigma^2}} dz = \frac{1}{2\pi}, & -\pi < \theta < \pi \\ 0, & \text{其它} \end{cases}$$

从 Z，Θ 的概率密度函数可以看出，Z 服从参数为 σ 的 Rayleigh 分布，Θ 服从区间$(-\pi, \pi)$上的均匀分布，并且

$$g(z, \theta) = g_Z(z) g_\Theta(\theta)$$

所以 Z 和 Θ 相互独立.

例 1.2.3 的结果是工程上的一个重要结论：若二维随机变量的两个分量是相互独立且同服从正态分布 $N(0, \sigma^2)$的随机变量，则该二维随机变量的模和辐角也是相互独立的随机变量，并且模服从参数为 σ 的 Rayleigh 分布，辐角服从$(-\pi, \pi)$上的均匀分布.

1.3 随机变量的数字特性

随机变量的分布函数是随机变量概率分布的完整描述，但是要找到随机变量的分布函数是一件不容易的事．另一方面，在实际问题中描述随机变量的概率特征，不一定都要求出它的分布函数，往往需要求出描述随机变量概率特征的几个表征值就够了．这就需要引入随机变量的数字特征，为此我们先介绍 Stieltjes 积分的概念．

定义 1.3.1 设 $f(x)$，$g(x)$是定义在$[a, b]$上的两个有界函数，$a=x_0<x_1<\cdots<x_n=b$ 是区间$[a, b]$的任一划分，$\Delta x_k=x_k-x_{k-1}$，$\Delta=\max\limits_{1\leqslant k\leqslant n}\Delta x_k$，在每一个子区间$[x_{k-1}, x_k]$上任意取一点 ξ_k 作和式

$$S=\sum_{k=1}^{n}f(\xi_k)[g(x_k)-g(x_{k-1})]$$

如果极限

$$\lim_{\Delta\to 0}S=\lim_{\Delta\to 0}\sum_{k=1}^{n}f(\xi_k)[g(x_k)-g(x_{k-1})]$$

存在且与$[a, b]$的分法和 ξ_k 的取法都无关，则称此极限为函数 $f(x)$对函数 $g(x)$在区间$[a, b]$上的 Stieltjes 积分，简称 S 积分，记为 $\int_a^b f(x)\,\mathrm{d}g(x)$. 此时也称 $f(x)$对 $g(x)$在$[a, b]$上 S 可积.

定义 1.3.2 设 $f(x)$，$g(x)$是定义在$(-\infty, +\infty)$上的两个函数，若在任意有限区间$[a, b]$上，$f(x)$对 $g(x)$在$[a, b]$上 S 可积，且极限

$$\lim_{\substack{a\to-\infty\\ b\to+\infty}}\int_a^b f(x)\,\mathrm{d}g(x)$$

存在，则称此极限为 $f(x)$对 $g(x)$在无穷区间$(-\infty, +\infty)$上的 Stieltjes 积分，简称 S 积分，记为 $\int_{-\infty}^{+\infty}f(x)\,\mathrm{d}g(x)$.

在 S 积分中，当 $g(x)$取一些特殊形式时，积分可化为级数或通常积分.

若 $g(x)$在$(-\infty, +\infty)$上是阶梯函数，它的跳跃点为 x_1，x_2，…(有限多个或可列无限多个)，则

$$\int_{-\infty}^{+\infty}f(x)\,\mathrm{d}g(x)=\sum_k f(x_k)[g(x_k+0)-g(x_k-0)]$$

若 $g(x)$在$(-\infty, +\infty)$上是可微函数，它的导函数为 $g'(x)$，则

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

定义 1.3.3 设函数 $g(x)$定义在无限区间$(-\infty, +\infty)$上，若积分

$$\int_{-\infty}^{+\infty}\mathrm{e}^{\mathrm{j}tx}\,\mathrm{d}g(x)=\int_{-\infty}^{+\infty}\cos tx\,\mathrm{d}g(x)+\mathrm{j}\int_{-\infty}^{+\infty}\sin tx\,\mathrm{d}g(x)$$

存在，则称此积分为 $g(x)$的 Fourier - Stieltjes 积分，简称 F - S 积分.

定义 1.3.4 设 X 是一个随机变量，$F(x)$是其分布函数，若 $\int_{-\infty}^{+\infty}|x|\,\mathrm{d}F(x)<+\infty$，