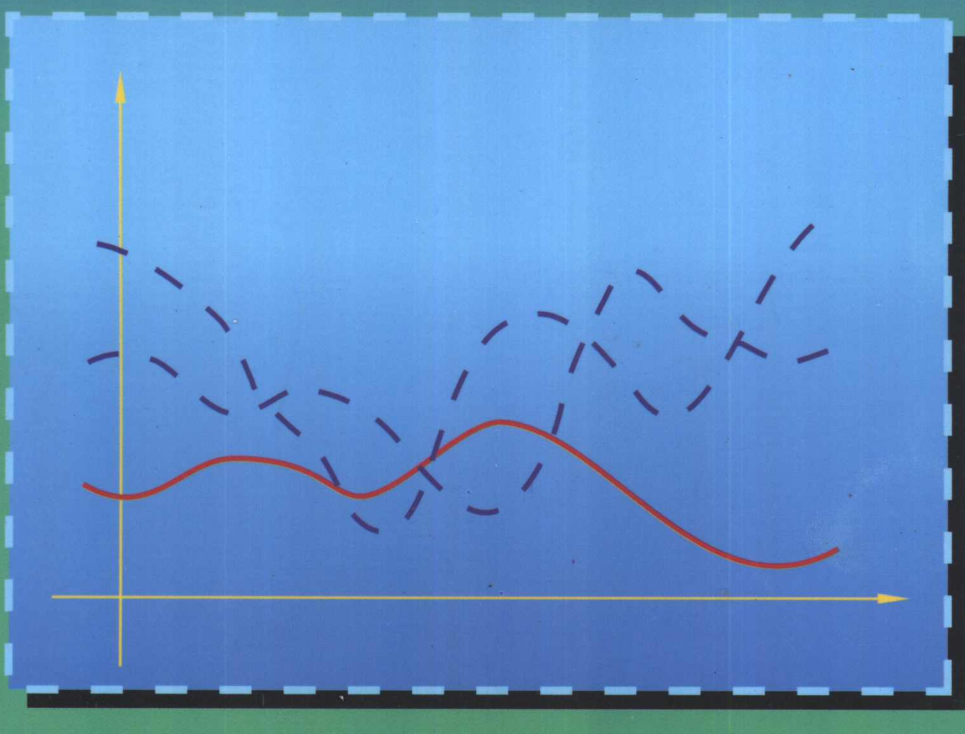
工程随机过程

夏乐天 朱永忠 编著



河海大学出版社

内 容 提 要

本书内容包括预备知识、随机过程的基本概念及分类、几个重要的随机过程简介、马尔可夫过程、平稳过程和时间序列分析共五章。阅读本书只要具备工科高等数学、线性代数和概率论的基本知识即可，是一本适应面宽，少学时的随机过程教材。可供高等院校工科研究生、高年级本科生作为教材使用，也可供工程技术人员参阅。

图书在版编目(CIP)数据

工程随机过程/夏乐天　朱永忠编著. －南京:河海大学出版社,2000.7

ISBN 7-5630-1501-9

Ⅰ.工…　Ⅱ.夏…　Ⅲ.随机过程　Ⅳ.0211.6

中国版本图书馆 CIP 数据核字(2000)第 61471 号

河海大学出版社出版发行

(南京西康路 1 号　邮编:210098)

河海大学印刷厂印刷　　江苏省新华书店发行

2000 年 7 月第 1 版　　2000 年 7 月第 1 次印刷

开本:787 × 1092 毫米　1/16　　印张:13.125

字数:312 千字　　印数:1 ~ 1050 册

定价:20.00 元

前　言

本书是在原讲义《随机过程与时间序列分析》的基础上修订而成的。为了适应相关专业研究生学习随机过程的需要，本着少学时、低起点、便于自学的要求，使之具有一定的通用性，修订后的书稿只要求读者具备高等数学、线性代数和概率论的基本知识，就可以阅读通篇内容。

本书对原讲义所作的修订主要反映在以下三个方面：

(1)辅助知识部分基本上重写并改名为预备知识。

(2)随机过程部分基本上保持了原貌，只在马尔可夫过程这一章作了一些增补，如增写了一节“马尔可夫链的状态分类”，并把“平稳分布与遍历性”扩充重写成一节。

(3)时间序列分析部分将内容大大压缩，使之成为一章，这更符合研究生教学的实际需要。

本书共五章。第一章对随机过程的一些必备知识(概率论与傅里叶变换)作了复习、补充性的简要介绍。第二章是随机过程的基本概念，除了对随机过程的概念、有穷维分布、数字特征作初步介绍外，还列举了常用的几个重要随机过程，第三章是马尔可夫过程。着重介绍马氏链及其状态分类，并对时间连续的马氏链简单地介绍了柯尔莫哥洛夫前进、后退方程。第四章是平稳过程。着重介绍宽平稳过程及其相关函数与功率谱密度函数，并简要介绍了均方随机分析的基本知识，还举例说明了平稳过程在线性系统中的一些应用。第五章介绍实用性很强的时间序列分析的初步知识，在讲授时，可以根据需要和学时数对内容进行取舍。

本书第三章与第四章具有相对的独立性，在教学时，可参考下面的关系图：

$$\begin{tikzcd} \boxed{\text{第一章}} \arrow[r] & \boxed{\text{第二章}} \arrow[r] \arrow[dr] & \boxed{\text{第三章}} \\ & & \boxed{\text{第四章}} \arrow[r] & \boxed{\text{第五章}} \end{tikzcd}$$

如果只讲授第一、二、三、四章，约需 50 学时；如果只讲授第二、三、四、五章，约需 60 学时；如果全书均需讲授而学时又不够，则马氏链的状态分类与平稳过程的谱分解理论可删去不讲。本书每章后面均附有适量的习题，供读者独立完成。

本书第一章至第四章由夏乐天执笔，第五章由朱永忠执笔，全书由夏乐天统稿。河海大学何才法同志审阅了本书的全部原稿，并提出了许多宝贵意见，谨在此表示衷心的感谢。值得说明的是，本书的出版得到了河海大学研究生部、教务处和出版社的大力支持，特别是得到了河海大学研究生部和河海大学出版基金的资助，我们也特在此表示深切的感谢。

由于时间仓促，水平有限，缺点与错误在所难免，恳请读者批评指正。

编　者

2000 年 5 月于河海大学

目　　录

第一章 预备知识

为了使概率论与随机过程更好地衔接起来，本章扼要地复习或加深概率论的某些基本概念，并补充特征函数、正态向量等一般工程数学中不讲授的内容。另外，还简要地介绍了积分变换中的傅里叶(Fourier)变换理论，为读者学习随机过程作好准备。

第一节 概率空间

概率论的基本概念是试验、事件和事件的概率。

所谓随机试验是指满足以下三个条件的试验(也简称为“试验”)：

(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，但能事先明确试验的所有可能的结果。

(3)每次试验前不能确定哪个结果会出现。

我们常用抽象点 ω 来表示试验的一个可能结果，ω 被称为试验的一个基本事件。全体基本事件的集合称为基本事件空间，或样本空间，记为 Ω，即 $\Omega=\{\omega\}$。在任何概率论教程中均已讨论过，一般事件总是由若干个基本事件所组成，因而是 Ω 的一个子集。换言之，每个事件可被看作某些点 $\omega\in\Omega$ 的并，而每个基本事件 $\omega\in\Omega$，不能被表为其他基本事件之并。

表 1.1 给出了概率论与集合论术语之间的关系。由表 1.1 知，Ω 的任一子集都称为事件。然而，无论是从实用观点或从数学观点看，把 Ω 的任一子集都看作是人们感兴趣的事件并没有意义。因此，应该在 Ω 中选出必须讨论的事件类，这个事件类是充分广泛的并包含在解决各种不同的实际问题时会出现的所有事件。另一方面，为了能够有效地利用数学上的技巧，这个事件类也要受到一定的限制。

表 1.1 概率论与集合论术语之间的关系

符 号	集合论术语	概率论术语
Ω	空间或全集	样本空间或必然事件
ω	Ω 中的点	基本事件
$\varnothing$	空集	不可能事件
$A:A\subset\Omega$	Ω 的子集	事件
$A\subset B$	集合 B 包含 A	A 是 B 的子事件
$A\cup B$	集合 A 与 B 之并	A 与 B 的和事件
$A\cap B$	集合 A 与 B 之交	A 与 B 的积事件，可记为 AB
$\bar{A}$	集合 A 之余集	A 的对立事件
$A-B$	集合 A 与 B 之差	A 与 B 的差事件
$A\cap B=\varnothing$	集合 A 与 B 不相交	事件 A 与 B 不相容

现在我们来考虑满足下述三个条件的 Ω 的子集的类 F。

公理 1.1.1 若 $\Omega=\{\omega\}$的一些子集所成的类 F 满足以下条件：

$(F.1)\Omega\in F$；

$(F.2)$若 $A\in F$,则 $\bar{A}\in F$；

$(F.3)$若可列多个 $A_m\in F,m=1,2,\cdots$,则 $\bigcup\limits_{m=1}^{\infty}A_m\in F$。

则称 F 是Ω 中的一个σ 代数或σ 域(若 Ω 是样本空间,则称 F 为Ω 中的一个事件域)。

由这个事件域的定义,易知以下事实：

(1)$\varnothing\in F$。

事实上,只要在$(F.2)$中取 $A=\Omega$,立知$\varnothing=\bar{\Omega}\in F$。

(2)若可列多个 $A_m\in F,m=1,2,\cdots$,则$\bigcap\limits_{m=1}^{\infty}A_m\in F$。

事实上,$\bigcap\limits_{m=1}^{\infty}A_m=\overline{\bigcup\limits_{m=1}^{\infty}\bar{A}_m}\in F$,由$(F.2)$与$(F.3)$及对偶律即得。

(3)若有 $A_m\in F,m=1,2,\cdots,n$,则$\bigcup\limits_{m=1}^{n}A_m\in F$,且$\bigcap\limits_{m=1}^{n}A_m\in F$。

事实上,在$(F.3)$中取 $A_{m+1}=A_{m+2}=\cdots=\varnothing$,即得$\bigcup\limits_{m=1}^{n}A_m\in F$,再与(2)类似可得$\bigcap\limits_{m=1}^{n}A_m\in F$。

(4)若 $A,B\in F$,则 $A-B=A\bar{B}\in F$。

由上述讨论可知,选择 F 作为Ω 中的事件域是足够广泛的,它包含了我们感兴趣的几乎所有事件。另一方面,也可看出,对于不同的问题所选择的 F 一般应该是不同的。但在一般的讨论中,我们总是认为 F 已经选定。

现在我们来引进事件的概率。

公理 1.1.2 对 F 中的每个元素A(事件)定义一个实数 $P(A)$与之对应,若集函数 $P(A)$满足以下条件：

$(P.1)$对每个 $A\in F$,有 $0\leqslant P(A)\leqslant1$(概率的非负性)。

$(P.2)$ $P(\Omega)=1$(概率的规范性)。

$(P.3)$若可列多个 $A_m\in F,m=1,2,\cdots$,且 $A_iA_j=\varnothing,i\neq j$,则 $P(\bigcup\limits_{m=1}^{\infty}A_m)=\sum\limits_{m=1}^{\infty}P(A_m)$(概率的可列可加性或完全可加性)。

则称 $P(A)$为 F 上的概率测度,简称概率。

由$(P.1)$,$(P.2)$与$(P.3)$可以推出概率的如下性质：

(1)$P(\varnothing)=0$； (1.1)

(2)若 $A_m\in F,m=1,2,\cdots,n$,且 $A_iA_j=\varnothing,i\neq j$,则

$$P(\bigcup_{m=1}^{n}A_m)=\sum_{m=1}^{n}P(A_m)\qquad\text{(概率的有限可加性)；}\tag{1.2}$$

(3)若 $A\subset B,A,B\in F$,则

$$P(B-A)=P(B)-P(A),\text{且 }P(A)\leqslant P(B)\text{(概率的单调不减性)；}\tag{1.3}$$

(4)若 $A\in F$,则 $P(\bar{A})=1-P(A)$； (1.4)

(5)若 $A,B\in F$,则 $P(A\cup B)=P(A)+P(B)-P(B)$(概率的加法公式)； (1.5)

(6)若 $A_m\in F$,且 $A_m\subset A_{m+1},m=1,2,\cdots,\lim\limits_{m\to\infty}A_m=\bigcup\limits_{m=1}^{\infty}A_m$,则

$$P(\bigcup_{m=1}^{\infty} A_m) = \lim_{m\to\infty} P(A_m) \qquad \text{(概率的下连续性)}; \tag{1.6}$$

(7)若 $A_m \in F$,且 $A_m \supset A_{m+1}, m = 1,2,\cdots, \lim\limits_{m\to\infty} A_m = \bigcap\limits_{m=1}^{\infty} A_m$,则

$$P(\bigcap_{m=1}^{\infty} A_m) = \lim_{m\to\infty} P(A_m) \qquad \text{(概率的上连续性)}。 \tag{1.7}$$

称三元总体(Ω, F, P)为概率空间,其中 Ω 是针对具体的随机试验所选取的样本空间,F 为Ω 中的事件域,P 为定义在 F 上的概率。概率空间是概率论研究的原始对象。

二元总体(Ω, F)称为可测空间。

设 E 为随机试验,Ω 为其样本空间,$A, B \in F$,若 $P(A) > 0$,则称

$$P(B \mid A) = \frac{P(AB)}{P(A)} \tag{1.8}$$

为事件 A 发生的的条件下事件 B 发生的条件概率,或简称为事件 B 关于事件 A 的条件概率。

条件概率也满足公理$(P.1)$,$(P.2)$与$(P.3)$。事实上,对固定的 $A \in F$,设 $P(A) > 0$,则

(1)对每一个 $B \in F, 0 \leqslant P(B \mid A) \leqslant 1$;

(2)$P(\Omega \mid A) = 1$;

(3)若 $B_m \in F, m = 1,2,\cdots$,且 $B_iB_j = \varnothing, (i \neq j)$,则

$$P(\bigcup_{m=1}^{\infty} B_m \mid A) = \sum_{m=1}^{\infty} P(B_m \mid A)$$

定义 1.1.1 事件组 $A_m \in F, m = 1,2,\cdots,n$ 称为Ω 的一个完备事件组或划分,若它们满足:①$A_iA_j = \varnothing, (i \neq j)$;②$\bigcup\limits_{m=1}^{n} A_m = \Omega$。

当 n 取∞时,也可类似地定义 Ω 的完备事件组。

定理 1.1.1 (乘法公式) 设 $A_1, A_2, \cdots, A_n$ 为任意 n 个事件($n \geqslant 2$),且 $P(A_1A_2\cdots A_{n-1}) > 0$,则

$$P(A_1A_2\cdots A_n) = P(A_1)P(A_2 \mid A_1)\cdots P(A_n \mid A_1A_2\cdots A_{n-1}); \tag{1.9}$$

特别

$$P(A_1A_2) = P(A_1)P(A_2 \mid A_2) \qquad (\text{若 } P(A_1) > 0);$$

$$P(A_1A_2A_3) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1A_2), (\text{若 } P(A_1A_2) > 0)$$

定理 1.1.2 (全概率公式与贝叶斯公式) 设 $A_1, A_2, \cdots, A_n$ 为Ω 的一个完备事件组,且 $P(A_i) > 0, (i = 1,2,\cdots,n)$,则

(1)对任何 $B \in F$,有

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B \mid A_i), \qquad \text{(全概率公式)}; \tag{1.10}$$

(2)对任何 $B \in F$,若 $P(B) > 0$,有

$$P(A_j \mid B) = \frac{P(A_j)P(B \mid A_j)}{\sum\limits_{i=1}^{n} P(A_i)P(B \mid A_i)}, \quad j = 1,2,\cdots,n, \text{(贝叶斯公式)} \tag{1.11}$$

如果事件 A, B 满足

$$P(AB) = P(A)P(B) \tag{1.12}$$

则称事件 A, B 相互独立。

定理 1.1.3 若事件 A,B 相互独立，则下列三对事件：A 与 $\bar{B}$；$\bar{A}$ 与 B；$\bar{A}$ 与 $\bar{B}$ 也分别相互独立。

设 $A_1,A_2,\cdots,A_n\in F$，若对任意 $S(2\leqslant S\leqslant n)$ 和任意 $1\leqslant i_1<i_2<\cdots<i_s\leqslant n$ 有

$$P(A_{i_1}A_{i_2}\cdots A_{i_s}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_s}), \tag{1.13}$$

则称事件 $A_1,A_2,\cdots,A_n$ 相互独立。

定理 1.1.4 设 $A_1,A_2,\cdots,A_n\in F$，相互独立。若其中任意 $m(1\leqslant m\leqslant n)$ 个事件相应地换成它们的对立事件，则所得 n 个事件仍然相互独立。

推论 若事件 $A_1,A_2,\cdots,A_n$ 相互独立，则

$$P(\bigcup_{i=1}^{n}A_i) = 1-\prod_{i=1}^{n}[1-P(A_i)]。\tag{1.14}$$

第二节 随机变量及其分布

设随机试验的样本空间 $\Omega=\{\omega\}$，对应的概率空间为 (Ω,F,P)，如果对于每一个 $\omega\in\Omega$ 都有唯一的一个实数 $X(\omega)$ 与之对应，且对任一实数 x，Ω 中的子集 $\{\omega\mid X(\omega)\leqslant x\}\in F$，则称 $X(\omega)$ 为 (Ω,F,P) 上的随机变量，简记 $X(\omega)$ 为 X。

随机变量 X 取值不超过 x 的概率 $P(X\leqslant x)$ 称为 X 的分布函数（x 为任意实数），记为 $F(x)$，即

$$F(x) = P(X\leqslant x) \qquad (-\infty<x<+\infty)。\tag{1.15}$$

分布函数 $F(x)$ 具有下列性质：

(1) $F(x)$ 是非降函数，即当 $x_1<x_2$ 时，有 $F(x_1)\leqslant F(x_2)$；

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

(3) $F(x)$ 是右连续的，即 $F(x+0)=F(x)$。

具有上述三个性质的函数 $F(x)$ 必是某个随机变量的分布函数。

由分布函数定义及连续性定理可得

$$P(a<X\leqslant b)=F(b)-F(a),$$

$$P(X=x_0)=F(x_0)-F(x_0-0),$$

$$P(X<x_0)=F(x_0-0),$$

$$P(X>x_0)=1-F(x_0),$$

$$P(a\leqslant X\leqslant b)=F(b)-F(a-0),$$

$$P(a<X<b)=F(b-0)-F(a),$$

$$P(a\leqslant X<b)=F(b-0)-F(a-0)。$$

如果随机变量 X 的可能取值仅有有限个或可列无穷多个，则称 X 为离散型随机变量。

设 $x_k(k=1,2,\cdots)$ 是离散型随机变量 X 的所有可能的取值，p_k 是 X 取值 x_k 的概率：

$$P(X=x_k)=p_k, \quad (k=1,2,\cdots); \tag{1.16}$$

则称上式为 X 的概率分布或分布律。可以看出

$$p_k\geqslant 0, \qquad (k=1,2,\cdots);$$

$$\sum_{h=1}^{\infty}p_k = 1$$

如果对于随机变量 X 的分布函数 $F(x)$，存在非负的函数 $f(x)$，使对任意的实数 x 有

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t \tag{1.17}$$

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

(1) $f(x)\geqslant 0 \qquad (-\infty<x<+\infty)$；

(2) $\int_{-\infty}^{+\infty}f(x)dx=1$；

(3)在 $f(x)$ 的连续点 x 处有

$$F'(x)=f(x)$$

对连续型随机变量 X，还成立

$$P(X=x_0)=0$$

$$P(a\leqslant X\leqslant b)=P(a\leqslant X<b)=P(a<X<b)=$$

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

设 E 是一个随机试验，它的样本空间为 Ω，X，Y 是定义在 Ω 上的随机变量；则称它们构成的一个向量 (X,Y) 为二维随机变量或二维随机向量。

二维随机向量 (X,Y) 的分布函数定义为：

$$F(x,y)=P(X\leqslant x,Y\leqslant y), \tag{1.18}$$

其中 x，y 为任意实数。$F(x,y)$ 具有下列性质：

(1) $F(x,y)$ 分别关于 x 和 y 非降，即对于任意固定的 y，当 $x_1<x_2$ 时，$F(x_1,y)\leqslant F(x_2,y)$；对于任意固定的 x，当 $y_1<y_2$ 时，有 $F(x,y_1)\leqslant F(x,y_2)$；

(2) $0\leqslant F(x,y)\leqslant 1$，且

对于任意固定的 y，$F(-\infty,y)=0$，对于任意固定的 x，$F(x,-\infty)=0$，$F(-\infty,-\infty)=0$，$F(+\infty,+\infty)=1$；

(3) $F(x,y)$ 分别关于 x 及 y 右连续，即

$$F(x+0,y)=F(x,y),\quad F(x,y+0)=F(x,y)$$

(4)对于任意的 (x_1,y_1)，(x_2,y_2)，若 $x_1<x_2$，$y_1<y_2$，则有

$$F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)\geqslant 0$$

具有上述四个性质的函数 $F(x,y)$ 必定是某个二维随机向量的分布函数。

如果二维随机向量 (X,Y) 的可能取值仅有有限对或可列无穷多对时，则称 (X,Y) 为二维离散型随机向量。

设 (X,Y) 所有可能取的值为 $(x_i,y_i)(i=1,2,\cdots;j=1,2,\cdots)$，而 p_{ij} 是 (X,Y) 取值 (x_i,y_i) 的概率，即

$$P(X=x_i,Y=y_i)=p_{ij}\qquad(i=1,2,\cdots;\ j=1,2,\cdots) \tag{1.19}$$

则称上式为二维离散型随机向量 (X,Y) 的概率分布或称为随机变量 X 和 Y 的联合分布律，它满足

$$p_{ij}\geqslant 0,\quad \sum_{i=1}^{m}\sum_{j=1}^{m}p_{ij}=1$$

如果存在一个非负的二元函数 $f(x,y)$，使对任意实数 x，y 有

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\mathrm{d}u\mathrm{d}v \tag{1.20}$$

则称(X,Y)为二维连续型随机向量。函数$f(x,y)$称为(X,Y)的概率密度，它具有下列性质：

(1)$f(x,y)\geqslant 0$；

(2) $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x\mathrm{d}y = 1$；

(3)在$f(x,y)$的连续点(x,y)处，有

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

(4)设 G 是 xoy 平面上的一个区域，点(X,Y)落在 G 内的概率

$$P((X,Y)\in G) = \iint_G f(x,y)\mathrm{d}x\mathrm{d}y$$

设(X,Y)的分布函数为 $F(x,y)$；则 X,Y 的分布函数$F_X(x)$，$F_Y(y)$依次称为(X,Y)关于 X 及关于 Y 的边缘分布函数，且有

$$F_X(x) = F(x,+\infty),\quad F_Y(y) = F(+\infty,y) \tag{1.21}$$

对离散型随机变量 X、Y，若分布律为

$$P(X = x_i, Y = y_j) = p_{ij} \qquad (i = 1,2,\cdots)$$

则 X 的分布律

$$p_{i\cdot} = p(X = x_i) = \sum_{j=1}^{\infty} p_{ij},\quad (i = 1,2,\cdots) \tag{1.22}$$

称为(X,Y)关于 X 的边缘分布律，而 Y 的分布律

$$p_{\cdot j} = P(Y = y_j) = \sum_{i=1}^{m} p_{ij},\quad (j = 1,2,\cdots) \tag{1.23}$$

称为(X,Y)关于 Y 的边缘分布律。

若(X,Y)为二维连续型随机向量，概率密度为 $f(x,y)$，X 和 Y 的概率密度依次记为$f_X(x)$，$f_Y(y)$，则

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y, \tag{1.24}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x \tag{1.25}$$

分别称 $f_X(x)$和 $f_Y(y)$为(X,Y)关于 X 和 Y 的边缘概率密度。

设(X,Y)是二维离散型随机向量，对于固定的 j，若 $P(Y=y_j)>0$，则称

$$P(X = x_i \mid Y = y_i) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{\cdot j}},\quad (i = 1,2,\cdots) \tag{1.26}$$

为在 $Y=y_j$ 的条件下，随机变量 X 的条件分布律，若对固定的 i，$P(X=x_i)>0$，则称

$$P(Y = y_j \mid X = x_i) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)} = \frac{p_{ij}}{p_{i.}} \quad (j = 1,2,\cdots) \tag{1.27}$$

为在 $X=x_i$ 的条件下，随机变量 Y 的条件分布律。

若(X,Y)为二维连续型随机向量，概率密度为 $f(x,y)$，X,Y 的概率密度分别为$f_X(x)$，f_Y

(y)。如果$f_X(x)\neq 0$,则在$X=x$条件下,Y的条件概率密度$f(y|x)$为

$$f(y\mid x)=\frac{f(x,y)}{f_X(x)} \tag{1.28}$$

如果$f_Y(y)\neq 0$,则在$Y=y$的条件下,X的条件概率密度$f(x|y)$为

$$f(x\mid y)=\frac{f(x,y)}{f_Y(y)} \tag{1.29}$$

设(X,Y)的分布函数为$F(x,y)$,边缘分布函数分别为$F_x(x),F_y(y)$。如果对于任意实数x,y有

$$F(x,y)=F_X(x)\cdot F_Y(y) \tag{1.30}$$

则称随机变量X和Y是相互独立的。

对于离散型随机变量X与Y,相互独立等价于对所有(X,Y)可能取的值(x_i,y_j)等式

$$P(X=x_i,Y=y_j)=P(X=x_i)\cdot P(Y=y_j),\quad (i=1,2,\cdots;\ j=1,2,\cdots) \tag{1.31}$$

成立。

对于连续型随机变量X与Y,相互独立等价于对所有实数x,y,等式

$$f(x,y)=f_X(x)f_Y(y) \tag{1.32}$$

几乎处处成立,其中$f(x,y),f_X(x),f_Y(y)$分别是(X,Y)的概率密度和关于X,Y的边缘概率密度。

若随机变量X,Y相互独立,则显然有

$$f(y\mid x)=f_Y(y),\quad f(x\mid y)=f_X(x)$$

对于任意n个实数$x_1,x_2,\cdots,x_n$,n元函数。

$$F(x_1,x_2,\cdots,x_n)=(X\leqslant x_1,X\leqslant x_2,\cdots,X\leqslant x_n) \tag{1.33}$$

称为n维随机向量$(X_1,X_2,\cdots,X_n)$的分布函数。

若存在非负n元函数$f(x_1,x_2,\cdots,x_n)$使得

$$F(x_1,x_2,\cdots,x_n)=\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n}f(y_1,y_2,\cdots,y_n)\mathrm{d}y_1\mathrm{d}y_2\cdots\mathrm{d}y_n \tag{1.34}$$

则称$(X_1,X_2,\cdots,X_n)$为连续型n维随机向量,$f(x_1,x_2,\cdots,x_n)$称为概率密度,且有

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

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

若用$+\infty$代替$F(x_1,x_2,\cdots,x_n)$中某些变量,则得到剩下那些变量的分布函数,例如,设四维随机向量(X_1,X_2,X_3,X_4)的分布函数为$F(x_1,x_2,x_3,x_4)$,则(X_1,X_3)的分布函数$F_{X_1X_3}(x_1,x_3)$为

$$F_{X_1X_3}(x_1,x_3)=P(X_1\leqslant x_1,X_3\leqslant x_3)=F(x_1,+\infty,x_3,+\infty) \tag{1.35}$$

若对概率密度$f(x_1,x_2,\cdots,x_n)$的某些变量从$-\infty$到$+\infty$进行积分,则得到剩下的变量的概率密度,例如,设(X_1,X_2,X_3,X_4)的概率密度为$f(x_1,x_2,x_3,x_4)$,则(X_1,X_3)的概率密度$f_{X_1X_3}(x_1,x_3)$为

$$f_{X_1X_3}(x_1,x_3)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f(x_1,x_2,x_3,x_4)\mathrm{d}x_2\mathrm{d}x_4 \tag{1.36}$$

X_1 的概率密度 $f_{X_1}(x_1)$ 为

$$f_{X_1}(x_1) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x_1,x_2,x_3,x_4)\mathrm{d}x_2\mathrm{d}x_3\mathrm{d}x_4 \tag{1.37}$$

我们用 $f(x_1,\cdots,x_h \mid x_n,x_{n-1},\cdots,x_{k+1})$ 表示在 $X_{k+1}=x_{k+1},\cdots,X_n=x_n$ 的条件下，$(X_1,X_2,\cdots,X_k)$ 的条件概率密度，且

$$f(x_1,\cdots,x_k \mid x_n,x_{n-1},\cdots,x_{k+1}) = \frac{f(x_1,\cdots,x_k,\cdots,x_n)}{f_{X_{k+1}\cdots X_n}(X_{k+1},\cdots,x_n)} \tag{1.38}$$

式中 $f(x_1,\cdots,x_k,\cdots,x_n)$ 为 $(X_1,X_2,\cdots,X_n)$ 的概率密度，而 $f_{X_{k+1}\cdots X_n}(x_{k+1},\cdots,x_n)$ 为 $(X_{k+1},\cdots,X_n)$ 的概率密度。

例如：

$$f(x_3 \mid x_2,x_1) = \frac{f(x_1,x_2,x_3)}{f_{X_1X_2}(x_1,x_2)}$$

$$f(x_n \mid x_{n-1},\cdots,x_1) = \frac{f(x_1,x_2,\cdots,x_n)}{f_{X_1X_2\cdots X_{n-1}}(x_1,x_2,\cdots,x_{n-1})}$$

反复利用上式可得

$$f(x_1,x_2,\cdots,x_n) = f(x_1)\cdot f(x_2 \mid x_1)\cdot f(x_3 \mid x_2,x_1)\cdots f(x_n \mid x_{n-1},x_{n-2},\cdots,x_1)$$

式中 $f(x_1)$ 为 X_1 的概率密度；$f(x_2 \mid x_1)$ 为在条件 $X_1=x_1$ 下 X_2 的条件概率密度；…；$f(x_n \mid x_{n-1},x_{n-2},\cdots,x_1)$ 为条件 $X_1=x_1,\cdots,X_{n-1}=x_{n-1}$ 下，X_n 的条件概率密度。

如果 $(X_1,\cdots,X_n)$ 的分布函数为 $F(x_1,\cdots,x_n)$；X_i 的分布函数为 $F_{X_i}(x_i)(i=1,2,\cdots,n)$。若对于任意实数 $x_1,x_2,\cdots,x_n$ 有

$$F(x_1,x_2,\cdots,x_n) = F_{X_1}(x_1)F_{X_2}(x_2)\cdots F_{X_n}(x_n) \tag{1.39}$$

则称 $X_1,\cdots,X_n$ 相互独立。在 $(X_1,\cdots,X_n)$ 为连续型情况下，上式等价于：

$$f(x_1,x_2,\cdots,x_n) = f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n) \tag{1.40}$$

几乎处处成立。其中 $f(x_1,x_2,\cdots,x_n)$ 为 $(X_1,X_2,\cdots,X_n)$ 的概率密度，$f_{X_i}(x_i)(i=1,2,\cdots,m)$ 为 X_i 的概率密度。

当 $X_1,X_2,\cdots,X_n$ 相互独立时，还有

$$f_{X_1\cdots X_{n-1}}(x_1,\cdots,x_{n-1}) = f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_{n-1}}(x_{n-1})$$

$$f_{X_1\cdots X_{n-2}}(x_1,\cdots,x_{n-2}) = f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_{n-2}}(x_{n-2})$$

$$\vdots$$

$$f_{X_1X_2X_3}(x_1,x_2,x_3) = f_{X_1}(x_1)f_{X_2}(x_2)f_{X_3}(x_3)$$

$$f_{X_1X_2}(x_1,x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$$

几乎处处成立。

设有随机变量序列

$$X_1,X_2,\cdots,X_n,\cdots$$

如果对任意正整数 m，$X_1,X_2,\cdots,X_m$ 相互独立，则称随机变量序列是独立的。

设有 n 个二维随机向量

$$(X_1,Y_1),(X_2,Y_2),\cdots,(X_n,Y_n)$$

如果对任何实数 $x_1,y_1,x_2,y_2,\cdots,x_n,y_n$,有

$$P(\{X_1\leqslant x_1,Y_1\leqslant y_1\}\cap\{X_2\leqslant x_2,Y_2\leqslant y_2\}\cap\cdots\cap\{X_n\leqslant x_n,Y_n\leqslant y_n\})$$
$$=P\{X_1\leqslant x_1,Y_1\leqslant y_1\}\cdot P\{X_2\leqslant x,Y_2\leqslant y_2\}\cdots\cdot P\{X_n\leqslant x_n,Y_n\leqslant y_n\};\quad(1.41)$$

则称该 n 个二维随机变量是独立的。

第三节　随机变量的函数的分布

一、一维随机变量的函数

设 X 是一维随机变量,$g(x)$是一元实变函数,则 $Y=g(X)$仍是随机变量,称它为 X 的函数。由 X 的分布和函数关系 $g(x)$可以确定 $Y=g(x)$的分布。

当 X 是离散型随机变量,具有形如式(1.16)的分布律时,$Y=g(X)$的分布律为

$$P(Y=g(x_k))=p_k,\quad(k=1,2,\cdots)\tag{1.42}$$

其中,若 $g(x_k)$有相同的,则对应概率相加。

当 X 是连续型随机变量时,设概率密度为 $f_X(X)$,则 $Y=g(x)$的分布函数为

$$F_Y(y)=P(Y\leqslant y)=P(g(X)\leqslant y)=\int_{S_y}f_X(x)\mathrm{d}x\tag{1.43}$$

其中 $S_y=\{x\mid g(x)\leqslant y\}$。由于式(1.43)的积分范围依赖于 y,积分值一般说来是 y 的函数,将结果对 y 求导就可得出$f_Y(y)$。特别,若 $g(x)$是一个单调可导函数,存在反函数 $h(y)$,则有下列密度变换公式

$$f_Y(y)=\begin{cases}f_X(h(y))\mid h'(y)\mid, & \alpha<y<\beta\\ 0, & \text{其它}\end{cases}\tag{1.44}$$

其中 $\alpha=\min\{g(-\infty),g(+\infty)\}$,$\beta=\max\{g(-\infty),g(+\infty)\}$。

例 1.3.1　设 X 服从正态分布 $N(\mu,\sigma^2)$,求 X 的线性函数 $Y=aX+b(a\neq0)$的概率密度。

解:由 $g(x)=ax+b$ 是单调可导函数,有反函数

$$h(y)=\frac{y-b}{a}$$

现在

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

由式(1.44)即得

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma\mid a\mid}\exp\left\{-\frac{[y-(a\mu+b)]^2}{2(\sigma a)^2}\right\}$$

这说明 $Y=aX+b$ 服从正态分布 $N(a\mu+b,(\sigma a)^2)$,可见,正态随机变量的线性函数仍服从正态分布。特别,若取

$$a=\frac{1}{\sigma},\quad b=-\frac{\mu}{\sigma},\quad\text{即 }Y=\frac{X-\mu}{a}$$

则 Y 服从标准正态分布 $N(0,1)$,所以称这个 Y 是 X 的标准化随机变量。

二、二维随机变量的函数

设(X,Y)是二维随机变量，$g(x,y)$是二元实变函数，则称随机变量 $Z=g(X,Y)$为(X,Y)的函数。

如果(X,Y)是离散型随机变量，具有形如式(1.19)的分布律，那么 $Z=g(X,Y)$的分布律为

$$P(Z = g(x_i,y_j)) = p_{ij}, \quad (i,j = 1,2,\cdots) \tag{1.45}$$

其中，若 $g(x_i,y_i)$有相同的，则对应概率相加。

如果(X,Y)是连续型随机变量，概率密度为$f(x,y)$，那么 $Z=g(X,Y)$的分布函数为

$$F_Z(z) = P(Z \leqslant z) = P(g(X,Y) \leqslant z) = \iint_{D_z} f(x,y)\mathrm{d}x\mathrm{d}y \tag{1.46}$$

其中 $D_z=\{(x,y)\mid g(x,y)\leqslant z\}$与 z 有关，所以积分值一般为 z 的函数，将式(1.46)对 z 求导即可得到$f_Z(z)$。这种求随机变量的函数的概率密度的方法，通常称为“分布函数法”。用此法也可求出多维连续型随机变量的函数的概率密度。

最常用到的函数是求和，即 $Z=X+Y$，Z 的概率密度可按分布函数法求得，也可按下列公式求得：

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x,z-x)\mathrm{d}x = \int_{-\infty}^{+\infty} f(z-y,y)\mathrm{d}y \tag{1.47}$$

若 X,Y 独立，则(1.47)式可写成

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x)\mathrm{d}x = \int_{-\infty}^{+\infty} f_X(z-y)f_Y(y)\mathrm{d}y \tag{1.48}$$

式(1.48)中的积分称为 $f_X(x)$与 $f_Y(y)$的卷积，记作 $f_X * f_Y$。

例 1.3.2 设 X,Y 相互独立，X 服从(0,1)区间上的均匀分布 $U(0,1)$，Y 服从参数为λ 的指数分布 $\mathscr{E}(\lambda)$，其概率密度为

$$f_Y(y) = \begin{cases} \lambda\exp\{-\lambda y\}, & y > 0, \\ 0, & \text{其它}. \end{cases} \quad (\lambda > 0)$$

求 $X+Y$ 的概率密度。

解法Ⅰ （用分布函数法计算）已知(X,Y)的概率密度为

$$f(x,y) = \begin{cases} \lambda\exp\{-\lambda y\}, & 0 < x < 1, y > 0, \\ 0, & \text{其它}。 \end{cases}$$

由式(1.46)知 $X+Y$ 的分布函数为

$$F_Z(z) = \iint_{D_z} f(x,y)\mathrm{d}x\mathrm{d}y$$

其中 $Z=X+Y$，$D_z=\{(x,y)\mid x+y\leqslant z\}$。

当 $z\leqslant 0$ 时，$F_Z(z)=0$；

当 $0 < z \leqslant 1$ 时，$F_Z(X) = \iint\limits_{x+y\leqslant z} \lambda\exp\{-\lambda y\}\mathrm{d}x\mathrm{d}y =$

$$\int_0^z \mathrm{d}x\int_0^{z-x}\lambda\exp\{-\lambda y\}\mathrm{d}y = \int_0^z[1-\exp\{-\lambda(z-x)\}\mathrm{d}x,$$

当 $z > 1$ 时，$F_Z(z) = \int_0^1 dx \int_0^{z-x} \lambda \exp\{-\lambda y\} dy = \int_0^1 [1 - \exp\{-\lambda(z-x)\}] \mathrm{d}x$

故

$$f_Z(z) = \begin{cases} 0, & z \leqslant 0, \\ 1 - \exp\{-\lambda z\}, & 0 < z \leqslant 1, \\ (e^{\lambda} - 1)\exp\{-\lambda z\}, & z > 1。 \end{cases}$$

解法Ⅱ（用公式(1.48)计算）易知

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) \mathrm{d}x$$

其中

$$f_X(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & 其它。 \end{cases}$$

$$f_Y(z-x) = \begin{cases} \lambda \exp\{-\lambda(z-x)\}, & z - x > 0, \\ 0, & 其它。 \end{cases}$$

显然，$f_X(x) f_Y(z-x) > 0$ 等价于

$$\begin{cases} 0 < x < 1, \\ z - x > 0, \end{cases} \quad 等价于 \quad 0 < x < \min\{1, z\}$$

故

$$f_Z(z) = \begin{cases} 0, z \leqslant 0, \\ \int_0^z \lambda \exp\{-\lambda(z-x)\} \mathrm{d}x = 1 - \exp\{-\lambda z\}, o < z \leqslant 1, \\ \int_0^1 \lambda \exp\{-\lambda(z-x)\} \mathrm{d}x = (e^{\lambda} - 1)\exp\{-\lambda z\}, z > 1。 \end{cases}$$

设 X, Y 相互独立，且服从同一分布。如果它们的和 $X + Y$ 仍服从这种分布，就说这种分布具有可加性或再生性。我们熟悉的二项分布 $b(n, p)$，泊松分布 $P(\lambda)$、正态分布 $N(\mu, \sigma^2)$ 都是具有可加性的分布，不过 $b(n, p)$ 仅对参数 n 具有可加性。例如：

如果 X, Y 相互独立，X 服从 $P(\lambda_1)$ 分布，Y 服从 $P(\lambda_2)$ 分布，那么 $X + Y$ 服从 $P(\lambda_1 + \lambda_2)$ 分布。

如果 X, Y 相互独立，X 服从 $N(\mu_1, \sigma_1^2)$ 分布，Y 服从 $N(\mu_2, \sigma_2^2)$ 分布，那么 $X + Y$ 服从 $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ 分布

例 1.3.3　设 $X_1, X_2, \cdots, X_n$ 独立同 $N(\mu, \sigma^2)$ 分布，求 $\frac{1}{n}\sum_{i=1}^{n} X_i$ 的分布。

解　由正态分布的可加性知道 $\sum_{i=1}^{n} X_i$ 服从 $N(n\mu, n\sigma^2)$ 分布，而 $\frac{1}{n}\sum_{i=1}^{n} x_i$ 是 $\sum_{i=1}^{n} x_i$ 的线性函数，由例 1.3.1 可知它服从 $N(\mu, \sigma^2/n)$ 分布。

令 $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$，那么 $\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ 服从标准正态分布 $N(0,1)$。

三、二维随机变量的变换

设 (X, Y) 是二维连续型随机变量，有概率密度 $f(x, y)$，又 $g(x, y)$，$h(x, y)$ 是两个二元实