# 随机信号分析 [第二版]

SUIJI XINHAO FENXI

罗鹏飞　张文明　刘福声　编著

国防科技大学出版社

# 随机信号分析

（第二版）

主编　罗鹏飞
编著　罗鹏飞　张文明
　　　刘福声

国防科技大学出版社
·湖南长沙·

**内容简介**

本书是为高等学校电子科学与技术、信息与通信工程各专业本科生编写的教材，目的是为电子工程与信息工程领域的初学者打好基础，以适应学习统计信号处理的需要。全书共八章，主要内容包括随机变量基础、随机过程的基本概念、随机信号的变换、随机序列分析、窄带随机信号分析、正态随机过程、马尔可夫过程与泊松过程、高阶统计理论等。本书强调对随机信号基本概念的理解和掌握系统的分析方法，注重基本理论与实际系统，特别是电子系统的联系。书中加入了大量的例题和图表，以帮助读者更好地理解课程的内容。每章均附有大量习题，并附有习题答案。

本书可作为高等学校本科生教材，也可供从事相关技术领域的工程技术人员参考。

**图书在版编目(CIP)数据**

随机信号分析/罗鹏飞等编著.—长沙:国防科技大学出版社,2000.8
ISBN 7-81024-661-5

Ⅰ.随... Ⅱ.罗... Ⅲ.随机信号-信号分析-高等学校-教材 Ⅳ.TN911.6

中国版本图书馆 CIP 数据核字(2000)第 35311 号

国防科技大学出版社出版发行
电话:(0731)4555681 邮政编码:410073
E-mail:gfkdcbs@public.cs.hn.cn
责任编辑:卢天贶 责任校对:张 静
新华书店总店北京发行所经销
国防科技大学印刷厂印装
*
开本:787×1092 1/16 印张:14.5 字数:335 千
2003 年 8 月第 2 版第 1 次印刷 印数:5001-9000 册
*
**定价:19.00 元**

# 前　言

随机信号分析是一门研究随机变化过程的特点与规律性的学科，它广泛应用于雷达、通信、自动控制、随机振动、地震信号处理、图像处理、气象预报、生物电子等领域。随着现代科学技术，特别是信息技术的发展，随机信号理论的应用将日益广泛和深入。

为了适应21世纪高等教育教学内容和课程体系改革的需要，作者在多年讲授"随机信号分析"课程讲稿的基础上，根据新的教学大纲、结合教学工作体会和从事的相关科研工作的成果，为电子科学与技术、信息与通信工程各专业本科生编写了本教材，目的是使读者通过本课程的学习，掌握随机信号分析的基本理论和系统的分析方法。

本教材的参考学时数为60学时，带 * 号的章节作为选项，教学中可根据教学需要加以取舍。全书共分八章，第一章随机变量基础，主要是复习概率论中随机变量的分布、数字特征和随机变量的函数基本理论。第二章介绍了随机过程的基本概念，主要包括随机过程的定义和分类，随机过程的统计描述，包括概率分布、数字特征、相关函数与功率谱，以及平稳随机过程的定义及其特点等。第三章系统介绍了随机信号变换的分析方法，包括随机过程变换的基本概念、极限与导数、随机信号通过线性系统和非线性系统的分析方法。第四章介绍了离散随机信号的基本概念、离散随机信号通过系统分析、几种常见的离散信号模型。第五章和第六章介绍了电子系统中常见的窄带随机信号和正态随机信号。第七章介绍了马尔可夫过程和泊松过程。第八章简要介绍了近年发展起来、并得到广泛应用的一种新的随机信号分析理论——高阶矩与高阶累积量。

本书的重点是随机过程的基本概念、随机过程的统计特性描述、随机信号通过系统分析以及电子系统中常见的窄带、正态随机信号。数字技术的发展，使得离散随机信号分析变得越来越重要，因此离散随机信号分析也是本书重点要求掌握的内容。在内容安排上，我们注重从工程应用的角度去讲解复杂的数学关系，而不过多强调数学上的严密性，力求物理概念清晰。书中例题内容广泛，各章附有习题和答案。多年教学实践表明，作适量的习题，对巩固和理解课程内容很有益处。

负责审阅的沈振康教授为本书提出了许多宝贵的意见，本书得到了国防科技大学出版社的大力支持和帮助，在此表示诚挚的谢意。由于作者水平有限，书中难免有缺点和错误，欢迎广大读者对本书提出宝贵的意见和建议，对不妥之处提出批评指正。

编者

2000年6月

# 目　录

## 第一章　随机变量基础



## 第二章　随机过程的基本概念



## 第三章　随机信号的变换

## 第四章 随机序列分析



## 第五章 窄带随机信号分析



## 第六章 正态随机过程



## 第七章 马尔可夫过程与泊松过程



## * 第八章 高阶统计理论

# 第一章 随机变量基础

随机信号分析的基础是概率论与随机变量的理论,本章简要介绍随机变量的基本理论,更为详细的内容请大家参考有关书籍。

## 1.1 概率论的基本术语

**随机试验** 满足下列三个条件的试验称为随机试验:

(1)在相同条件下可重复进行;

(2)试验的结果不止一个,所有可能的结果能事先明确;

(3)每次试验前不能确定会出现哪一个结果。

随机试验通常用 $E$ 表示,比如投掷硬币,就是一个随机试验,它满足以上三个条件,首先,投掷硬币是可以重复进行的,其次,试验的结果可能是正面,也可能是反面,即有两种可能的结果,而且只有这两种结果,事先可以明确,但具体到某次试验,试验前是不能预知会出现哪种结果。

**随机事件** 在随机试验中,对一次试验中可能出现也可能不出现、而在大量重复试验中却具有某种规律性的事件,称为随机事件,简称为事件,如投掷硬币出现正面等。

**基本事件** 随机试验中最简单的随机事件称为基本事件,如投掷骰子出现 1,2,…,6 点是基本事件,出现偶数点是随机事件,但不是基本事件。

**样本空间** 随机试验的所有基本事件组成的集合称为样本空间,记为 $S$,如投掷骰子的样本空间为$\{1,2,3,4,5,6\}$。

## 1.2 随机变量的定义

在随机试验中,试验的结果不止一个,如投掷骰子可能出现的点数,打靶命中的环数及一批产品中的次品数等。另一些随机试验尽管其可能结果与数值间没有直接的联系,如投掷硬币出现正面或反面、雷达探测发现“有目标”或“无目标”等,但我们可以规定一些数值来表示试验的可能结果。比如用“1”表示“正面”,“0”表示“反面”,对雷达探测用“1”表示“有目标”,“0”表示“无目标”。为了表示这些试验的结果,我们定义一个变量,变量的取值反映试验的各种可能结果。由于试验前我们无法确知试验结果,所以变量的值在试验前是无法确知的,即变量的值具有随机性,我们称这个变量为随机变量。下面给出详细的定义。

**定义**:设随机试验 $E$ 的样本空间为$S=\{e\}$,如果对于每一个 $e\in S$,有一个实数 $X(e)$ 与之对应,这样就得到一个定义在 $S$ 上的单值函数$X(e)$,称 $X(e)$为随机变量,简记为 $X$。

从以上的定义可以看出,随机变量是定义在样本空间 $S$ 上的一个单值函数。对应于

不同的样本 $e$,$X(e)$的取值不同,$X(e)$的随机性在样本 $e$ 中体现出来,因为在试验前究竟出现哪个样本事先无法确知,只有试验后才知道。$X$ 的取值可以是连续的,也可以是离散的,所以,根据 $X$ 取值的不同可以分为连续型随机变量和离散型随机变量。

所谓离散型随机变量是指它的全部可能取值为有限个或可列无穷个。离散型随机变量的概率特性通常用概率分布律来描述。

设离散型随机变量 $X$ 的所有可能取值为 $x_k(k=1,\cdots,n)$,其概率为

$$P(X=x_k)=P_k \quad (k=1,2,\cdots,n) \tag{1.2.1}$$

称上式为 $X$ 的概率分布或分布律。通常也用以下表格形式表示:

| $X$ | $x_1$ | $x_2$ | $\cdots$ | $x_n$ |
|---|---|---|---|---|
| $P_k$ | $P_1$ | $P_2$ | $\cdots$ | $P_n$ |

表中

$$\sum_{k=1}^{n} P_k = 1 \tag{1.2.2}$$

以下介绍几种典型的离散随机变量的概率分布。

1.(0,1)分布

设随机变量 $X$ 的可能取值为 0 和 1 两个值,其概率分布为

$$P\{X=1\}=p, \quad P\{X=0\}=1-p \quad (0<p<1) \tag{1.2.3}$$

称 $X$ 服从(0, 1)分布。如投掷硬币的试验,我们假定出现正面用 1 表示,出现反面用 0 表示,用 $X$ 表示试验结果,那么 $X$ 的可能取值为 0、1,$X$ 是一个离散型随机变量,服从(0,1)分布,且

$$P\{X=1\}=P\{X=0\}=0.5$$

2. 二项式分布

设随机试验 $E$ 只有两种可能的结果 $A$ 及 $\bar{A}$,且 $P(A)=p$,$P(\bar{A})=1-p=q$,将 $E$ 独立地重复 $n$ 次,这样的试验称为贝努里(Bernoulli)试验,那么在 $n$ 次试验中事件 $A$ 发生 $m$ 次的概率为

$$P_n(m)=C_n^m p^m q^{n-m} \quad (0<m<n) \tag{1.2.4}$$

(1.2.4)式正是$(p+q)^n$ 展开式的第 $m+1$ 项,故称为二项式分布。

3. 泊松(Poisson)分布

设随机变量 $X$ 的可能取值为 0,1,2,…且概率分布为

$$P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}, k=0,1,\cdots \quad (\lambda>0) \tag{1.2.5}$$

则称 $X$ 服从泊松分布。

## 1.3 随机变量的分布函数与概率密度

对于连续型随机变量,由于不能一一列举其可能的取值,因而不能用概率分布列来描

述，需要用概率分布函数与概率密度来描述。

设 $X$ 为随机变量，$x$ 为实数，定义

$$F(x) = P(X \leq x) \tag{1.3.1}$$

为 $X$ 的概率分布函数或简称为分布函数。

分布函数具有如下性质：

(1) 它是 $x$ 的不减函数，即

$$F(x_2) - F(x_1) \geq 0, x_2 > x_1 \tag{1.3.2}$$

(2) $0 \leq F(x) \leq 1$ (1.3.3)

(3) $F(-\infty) = 0, F(\infty) = 1$ (1.3.4)

(4) 若 $F(x_0) = 0$，则对任何 $x < x_0$，有 $F(x) = 0$ (1.3.5)

(5) $P(X > x) = 1 - F(x)$ (1.3.6)

(6) 函数 $F(x)$ 是右连续的，即

$$F(x^+) = F(x) \tag{1.3.7}$$

(7) 对于任意实数 $x_1, x_2 (x_1 < x_2)$，有

$$\begin{aligned} P(x_1 < X \leq x_2) &= P(X \leq x_2) - P(X \leq x_1) \\ &= F(x_2) - F(x_1) \end{aligned} \tag{1.3.8}$$

(8) $P(X = x) = F(x) - F(x^-)$ (1.3.9)

(9) $P(x_1 \leq X \leq x_2) = F(x_2) - F(x_1^-)$ (1.3.10)

对于连续型随机变量，其分布函数是连续的，在这种情况下，$F(x) = F(x^-)$，所以对于每一个 $x$ 都有

$$P(X = x) = 0 \tag{1.3.11}$$

式(1.3.11)说明连续型随机变量取某一点的概率为零。

对离散型随机变量，分布函数是阶梯型的。设 $x_i$ 表示 $F(x)$ 的不连续点，则

$$F(x_i) - F(x_i^-) = P(X = x_i) = p_i \tag{1.3.12}$$

这时 $X$ 的统计特性由它的取值 $x_i$ 及取值的概率 $p_i$ 确定，也即由概率分布律确定。分布函数可表示为

$$F(x) = \sum_i p_i U(x - x_i) \tag{1.3.13}$$

式中 $p_i = P\{X = x_i\}$，$U(\cdot)$ 为单位阶跃函数。

比如(0,1)分布的随机变量 $X$，其分布函数(见图 1.1)为

$$F(x) = P(X \leq x) = \begin{cases} 0 & , \quad x < 0 \\ 1/2 & , \quad 0 \leq x < 1 \\ 1 & , \quad x \geqslant 1 \end{cases}$$

如果 $X$ 的概率分布既不是连续的，也不是离散的，那么我们称 $X$ 为混合型随机变量。

对于连续型随机变量，如果它的概率分布函数 $F(x)$ 可导，则它的导数定义为随机变量 $X$ 的概率分布密度，简称为概率密度或分布密度，记为 $f(x)$。即

$$f(x) = \frac{\mathrm{d}F(x)}{\mathrm{d}x} \tag{1.3.14}$$

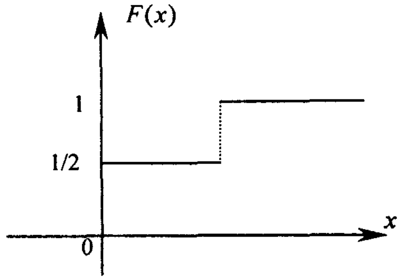


图 1.1 (0,1)分布的分布函数

由概率密度的定义及分布函数的性质,可以得出概率密度的如下性质:

(1) $f(x)\geq 0$,也即概率密度是非负的函数。

(2) $\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1$,即概率密度函数与横轴 $x$ 所围成的面积为 1。

(3) $p\{x_1<X\leq x_2\}=F(x_2)-F(x_1)=\int_{x_1}^{x_2}f(x)\mathrm{d}x$,这说明随机变量 $X$ 落在区间$(x_1,x_2)$上的概率等于图 1.2 中阴影区的面积。从这条性质我们也可以看出,对于连续型随机变量,有 $P\{X=x\}=0$。

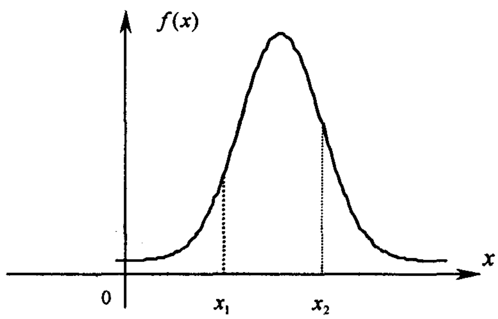


图 1.2 随机变量 $X$ 落在区间$(x_1,x_2)$上的概率

对于离散型随机变量,由于它的概率分布函数是阶梯型,那么它的概率密度函数是一串 $\delta$ 函数之和,$\delta$ 函数出现在随机变量的取值点,强度为取该值的概率。即

$$f(x)=\sum_i p_i\delta(x-x_i) \tag{1.3.15}$$

式中 $x_i$ 为离散型随机变量 $X$ 的取值,且 $p_i=P\{X=x_i\}$。

**离散型随机变量有以下的几种常见分布**

1. 正态分布

若随机变量的概率密度为

$$f(x)=\frac{1}{\sqrt{z\pi}\sigma}\exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\} \tag{1.3.16}$$

式中 $m$、$\sigma$ 为常数,则称 $X$ 服从正态分布,正态分布通常也简记为 $N(m,\sigma^2)$。正态分布随机变量的概率密度是一个高斯曲线,所以又称为高斯分布,概率密度曲线如图 1.3 所示。

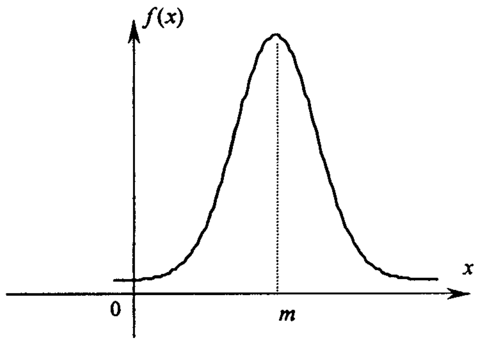


图 1.3 正态分布概率密度曲线

2. 均匀分布

如果随机变量 $X$ 的概率密度函数为

$$f(x)=\begin{cases}\dfrac{1}{b-a} & ,a<x<b\\ 0 & ,\text{其它}\end{cases} \tag{1.3.17}$$

则称 $X$ 在区间$(a,b)$上服从均匀分布。

3. 瑞利分布

如果随机变量 $X$ 的概率密度函数为

$$f(x)=\frac{x}{\sigma^2}\exp\left\{-\frac{x^2}{2\sigma^2}\right\} \quad ,x\geq 0 \tag{1.3.18}$$

式中 $\sigma$ 为常数,则称 $X$ 服从瑞利分布。

4. 指数分布

如果随机变量 $X$ 的概率密度函数为

$$f(x)=ce^{-cx} \quad ,x>0 \tag{1.3.19}$$

式中 $c$ 为常数,则称 $X$ 服从指数分布。

以上是实际应用中常见的几种分布。

## 1.4 多维随机变量及分布

在实际中,实验结果通常需要用多个随机变量才能加以描述,例如回波信号的幅度和相位需要用两个不同的随机变量来描述。我们称由多个随机变量构成的矢量为多维随机变量或随机矢量。

### 1.4.1 二维随机变量

设随机试验 $E$ 的样本空间 $S=\{e\}$,$X=X(e)$和 $Y=Y(e)$是定义在样本空间 $E$ 上的两个随机变量,由 $X$ 和 $Y$ 构成的矢量$(X,Y)$称为二维随机矢量或二维随机变量。

1. 二维分布函数

设 $x$、$y$ 为任意实数,那么二维随机变量$(X,Y)$的分布函数定义为

$$F(x,y)=P\{X\leq x,Y\leq y\} \tag{1.4.1}$$

二维随机变量$(X,Y)$的取值$(x,y)$可以看做为平面上的一个点，那么二维分布函数就是二维随机变量$(X,Y)$的取值落在图 1.4 所示的阴影区域的概率。

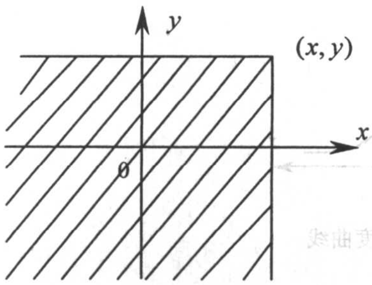


图 1.4　二维分布函数图解

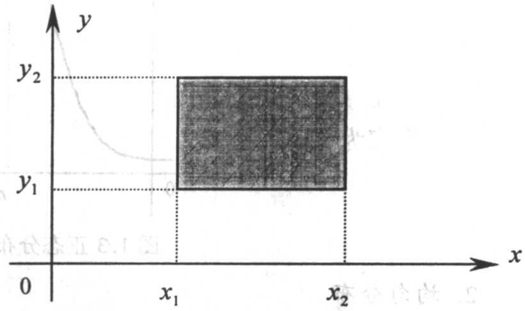


图 1.5　二维随机变量落在某一区域的概率

二维随机变量的分布函数具有下列性质：

(1) $0 \le F(x,y) \le 1$。

(2) 分布函数满足：

$$F(-\infty,y)=0,\ F(x,-\infty)=0,\ F(-\infty,-\infty)=0,\ F(\infty,\infty)=1。$$

(3) $F(x,\infty)=F_X(x)$，$F(\infty,y)=F_Y(y)$。

$F_X(\cdot)$和$F_Y(\cdot)$称为边缘分布，实际上也就是随机变量 $X$ 和 $Y$ 的分布，也就是说由二维分布函数可以求出一维分布函数。

(4) 对于任意的$(x_1,y_1)$和$(x_2,y_2)$，且 $x_2>x_1, y_2>y_1$，则

$$\begin{aligned} &P\{x_1<X\le x_2, y_1<Y\le y_2\} \\ &=F(x_2,y_2)-F(x_2,y_1)-F(x_1,y_2)+F(x_1,y_1)\ge 0 \end{aligned} \tag{1.4.2}$$

式(1.4.2)给出了利用二维分布函数计算二维随机变量落在某一区域的概率的方法，如图 1.5 所示。如果二维随机变量的可能取值为有限个或可列无穷个，则称$(X,Y)$为离散型随机变量。设

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

那么

$$\sum_i \sum_j p_{ij}=1 \tag{1.4.3}$$

$P\{X=x_i,Y=y_j\}=p_{ij}$称为$(X,Y)$的联合概率分布列或简称为分布列。

2. 二维概率密度

若二维分布函数 $F(x,y)$连续并存在二阶偏导数，则定义

$$f(x,y)=\frac{\partial^2 F(x,y)}{\partial x \partial y} \tag{1.4.4}$$

为$(X,Y)$的二维联合概率密度，简称为二维概率密度。

二维概率密度具有以下性质：

(1) $f(x,y)\ge 0$，即概率密度是非负的函数。

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

$$(1.4.5)$$

(3)边缘概率密度可由二维概率密度求得,即

$$f_X(x)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}y,\quad f_Y(y)=\int_{-\infty}^{\infty}f(x,y)\mathrm{d}x \tag{1.4.6}$$

(4)设 $G$ 是 $x-y$ 平面上的一个区域,则二维随机变量$(X,Y)$的取值落在该区域的概率为

$$P\{(X,Y)\in G\}=\iint_G f(x,y)\mathrm{d}x\mathrm{d}y \tag{1.4.7}$$

### 1.4.2 条件分布

设 $X$ 为一随机变量,$A$ 是一随机事件,我们定义

$$F(x\mid A)=P\{X\le x\mid A\} \tag{1.4.8}$$

为随机变量 $X$ 在事件 $A$ 给定时的条件分布函数,相应地,条件概率密度定义为条件分布函数的导数,即

$$f(x\mid A)=\frac{\mathrm{d}F(x\mid A)}{\mathrm{d}x} \tag{1.4.9}$$

由概率的特性,(1.4.8)式可以写成

$$F(x\mid A)=\frac{P\{X\le x,A\}}{P\{A\}} \tag{1.4.10}$$

设有二维随机变量$(X,Y)$,令 $A=\{X=x\}$,定义

$$F_{Y|X}(y\mid x)=P\{Y\le y\mid X=x\} \tag{1.4.11}$$

为随机变量 $Y$ 在 $X=x$ 时的条件分布函数,或称为 $Y$ 对 $X$ 的条件分布函数。相应地我们定义

$$f_{Y|X}(y\mid x)=\frac{\partial F_{Y|X}(y\mid x)}{\partial y} \tag{1.4.12}$$

为随机变量 $Y$ 在 $X=x$ 时的条件概率密度,或称为 $Y$ 对 $X$ 的条件概率密度。

设随机事件 $A=\{x_1<X\le x_2\}$,那么二维条件分布函数为

$$F(x,y\mid x_1<X\le x_2)=\frac{P\{X\le x,Y\le y,x_1<X\le x_2\}}{P\{x_1<X\le x_2\}}$$
$$=\begin{cases}\dfrac{F(x_2,y)-F(x_1,y)}{F_X(x_2)-F_X(x_1)}, & x\ge x_2\\[2ex] \dfrac{F(x,y)-F(x_1,y)}{F_X(x_2)-F_X(x_1)}, & x_1<x\le x_2\\[2ex] 0, & x<x_1\end{cases}$$

二维条件概率密度为

$$f(x,y\mid x_1<X\le x_2)=\frac{\partial^2 F(x,y\mid x_1<X\le x_2\}}{\partial x\partial y}$$

$$
=\begin{cases} \dfrac{f(x,y)}{F_X(x_2)-F_X(x_1)} & ,x_1 < x \le x_2 \\ 0 & , \quad 其他 \end{cases}
$$

在假定 $X=x$ 时,确定 $Y$ 的条件分布密度有着特别的意义,它不能由(1.4.11)式直接推导,因为对于连续型随机变量而言,事件$\{X=x\}$的概率为零。然而我们可以定义一个极限来求。

假定 $A=\{x_1<X\le x_2\}$,则

$$
F(y \mid x_1 < X \le x_2) = \frac{P\{Y \le y, x_1 < X \le x_2\}}{P\{x_1 < X \le x_2\}} = \frac{\int_{-\infty}^{y}\int_{x_1}^{x_2} f(x,y)\mathrm{d}x\mathrm{d}y}{F_X(x_2)-F_X(x_1)}
$$

(1.4.13)

(1.4.13)式对 $y$ 求导,得

$$
f_{(Y \mid x_1 < X \le x_2)}(y \mid x_1 < X \le x_2) = \frac{\int_{x_1}^{x_2} f(x,y)\mathrm{d}x}{F_X(x_2)-F_X(x_1)} \tag{1.4.14}
$$

在(1.4.14)式中,假定 $x_1=x, x_2=x+\Delta x$($\Delta x$ 为无穷小量),则

$$
f_{Y \mid x < X \le x+\Delta x}(y \mid x < X \le x+\Delta x) = \frac{\int_{x}^{x+\Delta x} f(x,y)\mathrm{d}x}{F_X(x+\Delta x)-F_X(x)} \approx \frac{f(x,y)\Delta x}{f_X(x)\Delta x}
$$

因此

$$
f_{Y \mid X}(y \mid x) = \lim_{\Delta x \to 0} f_{Y \mid x < X \le x+\Delta x}(y \mid x < X \le x+\Delta x) = \frac{f(x,y)}{f_X(x)} \tag{1.4.15}
$$

同理可得

$$
f_{X \mid Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)} \tag{1.4.16}
$$

于是有

$$
f(x,y) = f_{X \mid Y}(x \mid y) f_Y(y) = f_{Y \mid X}(y \mid x) f_X(x) \tag{1.4.17}
$$

### 1.4.3 全概率公式和贝叶斯公式

设 $S$ 为随机试验 $E$ 的样本空间,$A_1, A_2, \cdots, A_n$ 为 $S$ 的一个划分,即

$$
A_i \cap A_k = \Phi \quad (i \ne k), \Phi \text{ 为空集},
$$

$$
\bigcup_{i=1}^{n} A_i = S
$$

那么任意随机事件 $B$ 发生的概率为

$$
\begin{aligned} P(B) &= P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) + \cdots + P(B \mid A_n)P(A_n) \\ &= \sum_{i=1}^{n} P(B \mid A_i)P(A_i) \end{aligned}
$$

(1.4.18)

(1.4.18)式称为全概率公式。根据概率特性,

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{P(B)} = \frac{P(B \mid A_i)P(A_i)}{\sum_{i=1}^{n} P(B \mid A_i)P(A_i)} \tag{1.4.19}$$

(1.4.19)式称为贝叶斯公式。

下面我们将全概率公式和贝叶斯公式推广到随机变量的情形。设 $A_1, A_2, \cdots, A_n$ 为样本空间的一个划分，$B = \{X \leq x\}$，则由全概率公式可得

$$P\{X \leq x\} = \sum_{i=1}^{n} P\{X \leq x \mid A_i\} P(A_i)$$

即

$$F(x) = \sum_{i=1}^{n} F(x \mid A_i) P(A_i) \tag{1.4.20}$$

上式两边对 $x$ 求导，得

$$f(x) = \sum_{i=1}^{n} f(x \mid A_i) p(A_i) \tag{1.4.21}$$

(1.4.20)式和(1.4.21)式分别称为分布函数和概率密度的全概率公式。

设随机事件 $B = \{x_1 < X \leq x_2\}$，由等式

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

可得

$$\begin{aligned} P(A \mid x_1 < X \leq x_2) &= \frac{P\{x_1 < X \leq x_2 \mid A\}}{P\{x_1 < X \leq x_2\}} P(A) \\ &= \frac{F(x_2 \mid A) - F(x_1 \mid A)}{F(x_2) - F(x_1)} P(A) \end{aligned} \tag{1.4.22}$$

在上式中令 $x_1 = x, x_2 = x + \Delta x$，那么

$$\begin{aligned} P\{A \mid X = x\} &= \lim_{\Delta x \to 0} P\{A \mid x < X \leq x + \Delta x\} \\ &= \lim_{\Delta x \to 0} \frac{[F(x + \Delta x \mid A) - F(x \mid A)]/\Delta x}{[F(x + \Delta x \mid A) - F(x)]/\Delta x} P(A) \\ &= \frac{f(x \mid A)}{f(x)} P(A) \end{aligned}$$

或

$$P\{A \mid X = x\} \cdot f(x) = f(x \mid A) P(A) \tag{1.4.23}$$

在(1.4.23)式两边对 $x$ 积分，得

$$P(A) = \int_{-\infty}^{\infty} P\{A \mid X = x\} f(x) \mathrm{d}x \tag{1.4.24}$$

(1.4.24)式称为连续形式的全概率公式。

由(1.4.23)式和(1.4.24)式可得

$$f(x \mid A) = \frac{P(A \mid X = x) f(x)}{\int_{-\infty}^{\infty} P\{A \mid X = x\} f(x) \mathrm{d}x} \tag{1.4.25}$$

(1.4.25)式称为连续形式的贝叶斯公式。

### 1.4.4 多维分布

下面我们将二维分布的一些结论直接推广到多维的情况。

1. 多维分布函数

设 $n$ 维随机变量 $(X_1, X_2, \cdots, X_n)$，定义

$$F(x_1, x_2, \cdots, x_n) = P\{(X_1 \leq x_1, X_2 \leq x_2, \cdots, X_n \leq x_n)\tag{1.4.26}$$

为 $n$ 维随机变量的 $n$ 维分布函数。$n$ 维分布函数具有下列性质：

(1) $F(x_1, x_2, \cdots, -\infty, \cdots, x_n) = 0$，式中 $x_i = -\infty (i = 1, 2, \cdots, n)$；

(2) $F(\infty, \infty, \cdots, \infty) = 1$；

(3) $F(\infty, \infty, \cdots, \infty, x_m, x_{m+1}, \cdots, x_n) = F(x_m, x_{m+1}, \cdots, x_n)$。

2. 多维概率密度

若 $n$ 维分布函数的 $n$ 阶混合偏导数存在，那么定义

$$f(x_1, x_2, \cdots, x_n) = \frac{\partial^n F(x_1, x_2, \cdots, x_n)}{\partial x_1 \partial x_2 \cdots \partial x_n}\tag{1.4.27}$$

为 $n$ 维随机变量的 $n$ 维概率密度。显然

$$F(x_1, x_2, \cdots, x_n) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} f(x_1, x_2, \cdots, x_n) \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_n\tag{1.4.28}$$

对于 $n$ 维随机变量，其取值落在区域 $G$ 内的概率可表示为

$$P\{(X_1, X_2, \cdots, X_n) \in G\} = \iint_G f(x_1, x_2, \cdots, x_n) \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_n\tag{1.4.29}$$

3. 多维条件概率密度

在 $n$ 维随机变量 $(X_1, X_2, \cdots, X_n)$ 中，$X_{k+1}, X_{k+2}, \cdots, X_n$ 的取值为 $x_{k+1}, x_{k+2}, \cdots, x_n$ 的条件下，$X_1, X_2, \cdots, X_k$ 的条件概率密度为

$$f(x_1, x_2, \cdots, x_k \mid x_{k+1}, x_{k+2}, \cdots, x_n) = \frac{f(x_1, x_2, \cdots, x_n)}{f(x_{k+1}, x_{k+2}, \cdots, x_n)}\tag{1.4.30}$$

显然 $n$ 维概率密度与条件概率密度之间有如下关系：

$$f(x_1, x_2, \cdots, x_n) = f(x_1) f(x_2 \mid x_1) f(x_3 \mid x_1, x_2) \cdots f(x_n \mid x_1, x_2, \cdots, x_{n-1})\tag{1.4.31}$$

如果 $n$ 维随机变量 $(X_1, X_2, \cdots, X_n)$ 是相互独立的，那么

$$f(x_1, x_2, \cdots, x_n) = f(x_1) f(x_2) \cdots f(x_n)\tag{1.4.32}$$

## 1.5 随机变量的数字特征

随机变量的分布函数或概率密度能完整地描述随机变量的统计特性，但在实际中我们往往很难确定随机变量的分布函数或概率密度，这时我们可以用随机变量的数字特征来加以描述。常用的数字特征有均值、方差、协方差、相关系数等。

### 1.5.1 均值

随机变量 $X$ 均值也称为数学期望，它定义为

$$E(X) = \int_{-\infty}^{+\infty} x f(x) \mathrm{d}x\tag{1.5.1}$$

对于离散型随机变量，假定随机变量 $X$ 有 $n$ 个可能取值，各个取值的概率为 $p_i = P\{X =$