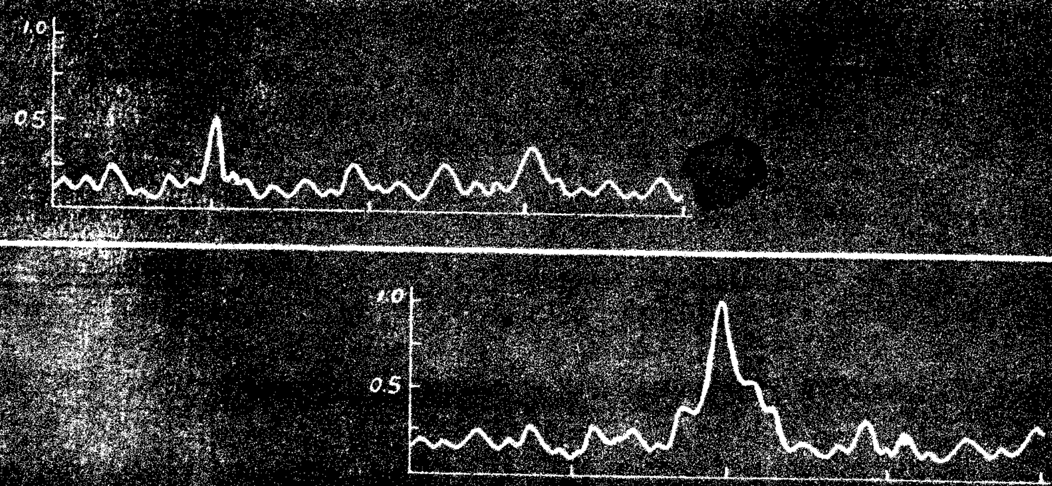
SJXHCHLDL

随机信号处理导论

吕扬生 边莫英 编



天津大学出版社

内容提要

本书较系统地介绍了几种随机信号处理的基本方法。前三章从复习概率论的基本概念入手，介绍了随机信号和它的统计特性的非参数估计方法；第四章讨论了信号检测和估计的基本原理；后三章则从线性系统理论入手介绍了AR谱估计、自适应滤波、卡尔曼滤波等信号处理方法。

本书着重于基本概念的讨论和分析，避免艰深的数学推导；适合于非电类专业高年级大学生、研究生作为教科书和参考书，也可供高等院校教师和工程技术人员参考。

随机信号处理导论

吕扬生　边莫英　编

天津大学出版社出版

（天津大学内）

河北省邮电印刷厂印刷

新华书店天津发行所发行

开本：787×1092毫米1/16　印张：$15^1/_4$字数：380千字

1988年5月第一版　　1988年5月第一次印刷

印数 1—2500

ISBN 7-5618-0051-7

TB·1

定价：2.60元

前　言

随着计算机应用的普及和计算技术的进步，信号处理技术的应用范围日益扩大；它不仅用于处理声音、图象、雷达等信号，而且更加广泛地用于机械、气象、医学甚至经济学等社会科学学科中去。

实际上几乎所有信号，包括来自自然的和社会的，都伴随着噪声；或者本身就以随机形式存在。因此，掌握处理随机信号的方法就会加深人们对自然，社会的了解，增强改造自然、改造社会的能力。

信号处理是一个比较复杂的学科，它用到包括概率统计、积分变换、线性代数等方面的数学知识，还牵涉到线性系统分析方面的内容。因此，关于这一学科的全面的严格的论述对于广大读者来说就显得过于深奥。实际的信号处理计算程序常常比较简单；有时甚至不懂原理也能进行计算，可是不懂原理就无法正确理解和应用分析的结果。鉴于这种情况就需要一种比较简明而又能系统解释信号处理技术的基本概念的实用读物。它可以在不损害科学性的条件下，尽量利用直观的讨论来代替严格的证明。掌握了这些内容，读者就可以在实践中应用，并且还可以此为起点，去掌握更深入的内容。向读者提供这样一本书，正是我们编写这本书的出发点。

描述随机信号的数学工具是随机过程，本书第二章就讨论这个内容。为使读者易于接受这一概念，在第一章复习了概率论的基本知识。第三章介绍根据观测数据估计平稳信号统计特性的非参数方法。第四章在贝叶斯风险准则的基础上，介绍了几种信号检测方法；从多元判定引出了信号估计理论和方法。第五章和第七章分别讨论了信号差分模型和状态变量模型以及应用它们进行信号参数估计的方法。第六章介绍了自适应滤波原理，这是一项具有很大应用潜力的技术。这些内容除了它们本身的实用价值以外，还为读者进一步学习模式识别理论打下了基础。为了便于读者理解本书用到的数学知识，我们又写了个附录补充了若干工科大学工程数学未学的内容。

本书是在几年教学活动和课题研究工作基础上写成的。全书由吴泳诗、王予宏和李修恕三位教授审阅，他们对本书的编写给予了极大的支持和关注，并提出许多宝贵的意见。鉴于时间仓促及作者水平所限，书中不足之处乃至错误敬请读者指正。

目　录

第一章　随机变量

§1.1　确定性信号与随机信号

人类在认识世界、改造世界的过程中需要研究各种客观事物的运动变化，收集有关的信息。为了便于观察和分析，人们通常利用某种装置把研究对象的运动从它原来的形式（如温度、压力等）转化为某种特定形式的量，比如电量，即为信号。然后再通过信号进行深入的研究。

客观事物的运动变化都是在一定条件下发生的。为了寻找它的规律，人们总要设法控制或掌握一部份影响较大的、较重要的条件；对其余的条件则往往不了解，一般把它们当作随机因素处理。如果这些条件对研究对象运动的综合影响很小，就可以观察到较有规律的信号。这种信号容易再现，也容易用数学关系式来表达，它们就是所谓的确定性信号。

与此不同，当那些未加控制的条件的影响相当大，以至不能忽略时，由于对这些条件的变化规律不了解，观察到的信号就成为随机的。这类信号不能用人为控制条件的方法准确再现，也不能准确预言这类信号在未来时刻的数值。因此，这类信号不能用确定的数学关系式来表达，只能用概率和统计的方法描述。

自然界中存在的物理现象大多具有随机性质。例如，用地震法勘察地层时，由于地质条件的千变万化，接受到的信号是随机信号，在分析噪声研究机器缺陷时，要分析的也是随机信号。即使是确定性信号，通常也伴随有噪声，造成了一定的随机性。因此，认识随机信号的特性，掌握处理它们的方法是十分重要的。分析和处理随机信号的能力越强，就越能从复杂的自然现象中获得更多的有用信息，增强人类认识自然、改造自然的能力。

为了建立随机信号的基本概念，图1.1给出了某一随机信号的记录。这些信号是这样产生的：同样的试验在相同的条件下重复N次，得到N个分立的记录。每个记录都以$t=0$为时间参考点，这些记录构成了一个记录的集合，每个记录都是这个随机信号的一个样本。从样本记录可以看到，尽管试验条件相同，但任何一个样本都不是其他样本的再现，根据$N-1$个样本记录不能准确地预测第N个样本的特性是典型的随机信号。

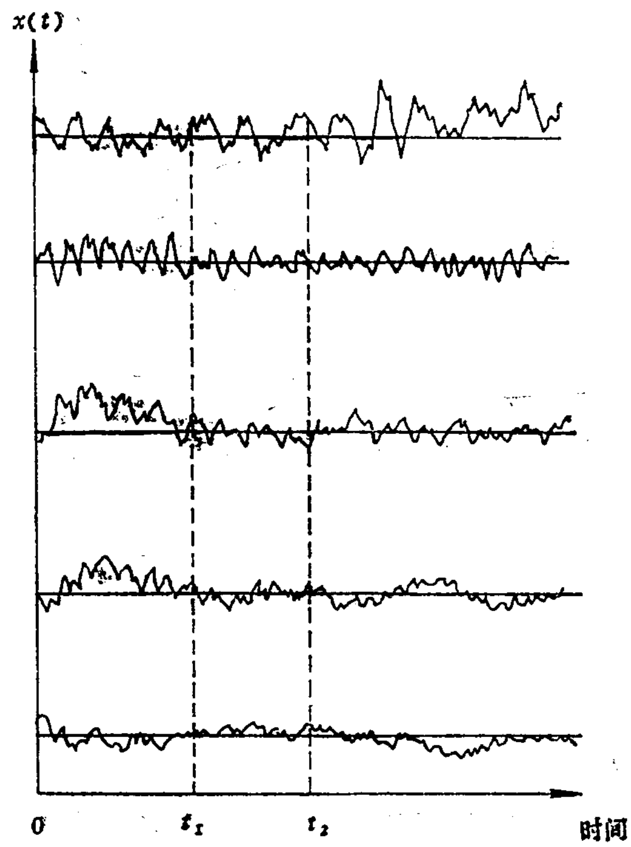


图 1.1　随机信号的N个记录

这一样本集合的统计量可以这样来获得，从这些记录中读出某一特定时刻（例如t_1）信号的值 $x_k(t_1)$，$k=1$，2，…，N，用这些值可以计算出一系列的统计量来。例如在时刻

t_1的均值为

$$m_X(t_1)=\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}x_k(t_1) \tag{1.1.1}$$

再如，在时刻t_1样本值的方差为

$$VarX(t_1)=\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}[x_k(t_1)-m_X(t_1)]^2 \tag{1.1.2}$$

上述取得统计量的过程称为集合平均，这些统计量就是集合统计量。上述分析指出，对应于每一时刻随机信号的值都是一个随机变量。而随机信号则是由许多存在某种关系的随机变量构成的随机函数，函数的变量一般用 t 表示，故又称为随机过程。

§1.2 随机变量的分布和数字特征

1-2-1 概率分布

如前所述，随机信号是一个随机过程，对应于每一个特定时刻的随机信号为一随机变量。随机变量有两种不同类型。一种随机现象，其试验结果X所可能取的值是可列的，例如电话呼叫次数，废品数等。这类随机变量的取值空间是离散的，称为离散型随机变量。

设$\{x_i\}$，$i=1，2，\cdots$，为离散型随机变量X的所有可能取值。记$P(x_i)$为X取x_i的概率，即

$$P\{X=x_i\}=P(x_i),\quad i=1，2，\cdots \tag{1.2.1}$$

那么离散型随机变量的概率分布可表示成下述阵列

$$X:\begin{pmatrix} x_1, & x_2, & \cdots, & x_n \\ P(x_1), & P(x_2), & \cdots, & P(x_n) \end{pmatrix} \tag{1.2.2}$$

阵列（1.2.2）称为分布列。由分布列能一目了然地看出随机变量X的概率分布。显然，这种概率分布应满足下述关系

$$P(x_i)\geqslant 0,\qquad i=1，2，\cdots \tag{1.2.3}$$

$$\sum_{i=1}^{\infty}P(x_i)=1 \tag{1.2.4}$$

有了概率分布，可通过下式求得分布函数

$$F(x)=P\{X\leqslant x\}=\sum_{x\geqslant x_K}P(x_K) \tag{1.2.5}$$

这时$F(x)$是一个跳跃函数，它在每个x_K处跳跃了$P(x_K)$。由式（1.2.5）知，$F(x)$是由$P(x_K)$，$K=1，2，\cdots$，唯一决定的。反之，有了$F(x)$也可唯一决定x_K和$P(x_K)$。因此，用分布列或分布函数均能描述离散型随机变量。

例1.1　二项分布

设一随机试验的结果只能取A和$\overline{A}$两种状态，A出现的概率为P，（$0<P<1$），$\overline{A}$出现的概率为$1-P$。同样的试验独立地进行N次，随机变量X为N次试验中A出现的次数。显然X是在0到N之间的一个整数。在N次试验中，结果A出现x次的不同方式有$N!/[x!(N-x)!]$种。每种方式出现的概率为$P^x(1-P)^{N-x}$。因此，随机变量X取值x的概率为

$$P(x)=\frac{N!}{x!(N-x)!}P^x(1-P)^{N-x}, \qquad x=0,\ 1,\ 2,\ \cdots,\ N \tag{1.2.6}$$

由二项式定理给出

$$(p+q)^N=\sum_{x=0}^{N}\frac{N!}{x!(N-x)!}P^x q^{N-x} \tag{1.2.7}$$

所以

$$\sum P(x)=[P+(1-P)]^N=1 \tag{1.2.8}$$

同时，因 $0<P<1$，故 $0<1-P<1$。

因此证明了$P(x)$满足（1.2.3）和（1.2.4）的关系，具备了离散型概率分布的特性。

与离散型随机变量不同，另一种随机现象的试验结果，其取值连续分布在取值空间的某一区域中，不可能一一列举出来。例如，机械零件的加工误差、电路中的噪声、风速等。这类随机变量称为连续型随机变量。这类随机变量所取的值不能一一列出，且它取某个特定值的概率常常为零，因此，用描写离散型随机变量分布的方法来描写连续型随机变量是不可能的。

对于连续型随机变量，如零件加工误差X，人们所关心的往往不是它取某个特定值的概率，而是它的取值小于某个特定值x的概率$P\{X\leqslant x\}$，或是它的取值x落在某个区间（x，$x+\triangle x$）内的概率$P\{x<X\leqslant x+\triangle x\}$。于是，连续型随机变量$X$的概率分布函数定义为

$$F(x)=P\{X\leqslant x\} \tag{1.2.9}$$

由定义可知，概率分布函数具有下列特性：

1） 当x趋于$-\infty$时，$F(x)$趋于0；

2） 当x趋于$+\infty$时，$F(x)$趋于1；

3） $F(x)$是x的连续的非下降函数。

包含在分布函数中的信息也可以用导数的形式来表示。定义概率密度函数$f(x)$为

$$f(x)=\frac{\mathrm{d}}{\mathrm{d}x}F(x) \tag{1.2.10}$$

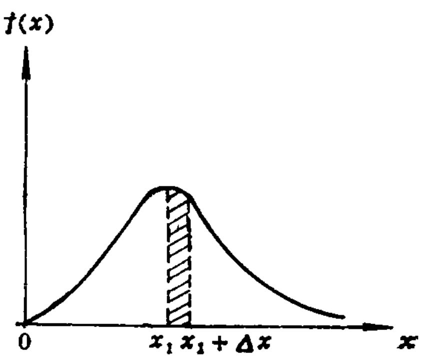


图 1.2 概率密度函数

因此，概率密度函数有下列特性：

1） $f(x)$是一个非负函数；

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

根据式（1.2.10）可知

$$\int_{x_1}^{x_1+\triangle x} f(t)\mathrm{d}t=F(x_1+\Delta x)-F(x_1)$$
$$=P\{x_1<X\leqslant x_1+\Delta x\}$$

图1.2中阴影部分的面积，正好是所表示的随机变量X的取值落在x_1和$x_1+\Delta x$之间的概率。如果Δx为无穷小，那么阴影部分的面积近似为$f(x)\ \Delta x$，因此称$f(x)$为概率密度，有时$f(x)$又称为似然值。

例1.2 一元正态密度函数

正态分布是经常遇到的最重要的一种分布，其概率密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left[-\frac{(x-m)^2}{2\sigma^2}\right], \qquad -\infty<x<\infty \tag{1.2.11}$$

式中$\sigma>0$，m与σ均为常数。

由上式不难看出，$f(x)$是非负的。可以证明，密度函数$f(x)$满足

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

因此，$f(x)$具有概率密度函数的特性。

1-2-2 数字特征与特征函数

由上节可见，概率分布已完整地描述了随机变量。但这种表示方法不够简练。因此，需找出一些量来更简单明确地描述随机变量的统计特性，这些量就是随机变量的数字特征，它包括随机变量的各阶矩。

首先定义随机变量的均值或数学期望，记作$E\{X\}$。对于离散型随机变量X，$E\{X\}$定义为

$$E\{X\}=\sum x_i\cdot P(x_i) \tag{1.2.12}$$

式中x_i为随机变量X的全部可能取值，$P(x_i)$是X取值为x_i的概率。对于连续型随机变量，$E\{X\}$可定义为

$$E\{X\}=\int_{-\infty}^{\infty} xf(x)\mathrm{d}x \tag{1.2.13}$$

式中$f(x)$为X的概率密度函数。

从图1.2看出，数学期望的几何意义为曲线$f(x)$与x轴所包围的面积对于坐标原点的一次矩，在此基础上产生了一类常用的数字特征——矩。最常用的有两种矩，即原点矩和中心矩。

K阶原点矩定义为随机变量X的K次方的数学期望，即

$$M_K=E\{X^K\} \tag{1.2.14}$$

对离散型随机变量X，为

$$M_K=\sum x_i{}^K P(x_i) \tag{1.2.15}$$

对连续型随机变量X为

$$M_K=\int_{-\infty}^{\infty} x^K f(x)\mathrm{d}x \tag{1.2.16}$$

因此，均值就是随机变量X的一阶原点矩M_1，记为m。

K阶中心矩定义为

$$C_K=E\{(X-m)^K\} \tag{1.2.17}$$

对离散型随机变量有

$$C_K=\sum (x_i-m)^K P(x_i) \tag{1.2.18}$$

对连续型随机变量有

$$C_K=\int_{-\infty}^{\infty} (x-m)^K f(x)\mathrm{d}x \tag{1.2.19}$$

由此可知，随机变量的二阶中心矩为

$$C_2=E\{(x-m)^2\} \tag{1.2.20}$$

它又称为随机变量的方差，可用符号表示为$\mathrm{Var}\{x\}$。方差的平方根是一个很重要的参数，称为标准差，即

$$X\text{的标准差}=\sqrt{\mathrm{Var}\{X\}}$$

例1.3　一元正态分布的均值和方差

正态概率密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left[-\frac{1}{2\sigma^2}(x-m)^2\right]$$

不难证明，其均值为

$$E\{X\}=\int_{-\infty}^{\infty}xf(x)dx=m$$

其方差为

$$\mathrm{Var}\{X\}=\int_{-\infty}^{\infty}(x-m)^2f(x)\mathrm{d}x=\sigma^2$$

因此，一元正态分布只决定于均值m和方差σ^2，故可简记为$N(m,\ \sigma^2)$。

数字特征反映了概率分布的某些侧面，但一般不能通过它完全确定分布函数。下面引进的特征函数，既能完全确定分布函数又具有良好的分析性质。

随机变量X的特征函数定义为复随机变量$e^{j\omega X}$的数学期望，用符号$\Psi(\omega)$表示。即

$$\Psi(\omega)=E\{e^{j\omega X}\} \tag{1.2.21}$$

离散型随机变量的特征函数为

$$\Psi(\omega)=\sum P_i e^{j\omega x_i} \tag{1.2.22}$$

连续型随机变量的特征函数为

$$\Psi(\omega)=\int_{-\infty}^{\infty}f(x)e^{j\omega x}\,\mathrm{d}x \tag{1.2.23}$$

这时，特征函数就是密度函数$f(x)$的傅里叶变换。因此，在计算特征函数时，利用傅里叶变换的一些定理（和表）是十分方便的。

在计算X的各阶矩时，特征函数特别有用。若X为零均值随机变量，则X的各阶原点矩与中心矩完全一致，可以写成

$$\begin{aligned}E\{X\}&=\int_{-\infty}^{\infty}xf(x)\mathrm{d}x\\E\{X^2\}&=\int_{-\infty}^{\infty}x^2f(x)\mathrm{d}x\\&\cdots\cdots\end{aligned} \tag{1.2.24}$$

同时，$\Psi(\omega)$的各次导数在$\omega=0$的值为

$$\begin{aligned}\left[\frac{\mathrm{d}\Psi}{\mathrm{d}\omega}\right]_{\omega=0}&=\left[\int_{-\infty}^{\infty}jxf(x)e^{j\omega x}\,\mathrm{d}x\right]_{\omega=0}\\&=\int_{-\infty}^{\infty}jxf(x)\mathrm{d}x\\\left[\frac{\mathrm{d}^2\Psi}{\mathrm{d}\omega^2}\right]_{\omega=0}&=\left[\int_{-\infty}^{\infty}(jx)^2f(x)e^{j\omega x}\,\mathrm{d}x\right]_{\omega=0}\\&=\int_{-\infty}^{\infty}j^2x^2f(x)\mathrm{d}x\\&\cdots\cdots\end{aligned} \tag{1.2.25}$$

对照式（1.2.24）与式（1.2.25）可见

$$E\{X\}=\frac{1}{j}\left[\frac{\mathrm{d}\Psi}{\mathrm{d}\omega}\right]_{\omega=0} \tag{1.2.26}$$

$$E\{X^2\}=\frac{1}{j^2}\left[\frac{\mathrm{d}^2\Psi}{\mathrm{d}\omega^2}\right]_{\omega=0} \tag{1.2.27}$$

其高阶矩可依次类推。

因此，借助于傅里叶变换表，不进行按其定义规定的积分，就能求出X的各阶矩。

例1.4　正态分布$N(m,\ \sigma^2)$的特征函数

$$\begin{aligned}\Psi(\omega)&=\int_{-\infty}^{\infty}e^{j\omega x}f(x)\,\mathrm{d}x\\&=\frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{\infty}e^{j\omega x-(x-m)^2/2\sigma^2}\mathrm{d}x\\&=e^{jm\omega-\frac{1}{2}\sigma^2\omega^2}\frac{1}{\sqrt{2\pi}}\int_{-\infty-j\omega\sigma}^{\infty-j\omega\sigma}e^{-z^2/2}\mathrm{d}z=e^{jm\omega-\frac{1}{2}\sigma^2\omega^2}\end{aligned}$$

§1.3　随机向量和随机变量的相关性

1-3-1　随机向量及其分布

在某些随机现象中，每次试验的结果不能只用一个随机数来描述，而要用几个随机数来表示。例如钢的成份，需要同时指出它的含碳量、含硫量、含磷量等等，这几种成份各为一个随机变量，分别表示成X_1，X_2，…，X_n，它们一起构成了一个多元随机变量或随 机向量$\mathbf{X}$

$$\mathbf{X}=(X_1X_2\cdots X_n)^T \tag{1.3.1}$$

称为n元函数。

$$F(x_1x_2\cdots x_n)=P\{X_1\leqslant x_1,\ X_2\leqslant x_2,\ \cdots,\ X_n\leqslant x_n\} \tag{1.3.2}$$

为随机向量$\mathbf{X}$的联合分布函数。

类似于一元的情况，随机向量的联合分布函数具有下列性质：

1)　$F(x_1,\ x_2,\ \cdots,\ -\infty,\ \cdots,\ x_n)=0$　(1.3.3)

2)　$F(\infty,\ \infty,\ \cdots,\ \infty)=1$　(1.3.4)

3)　$F(x_1,\ x_2,\ \cdots,\ x_n)$对每个变元均有左连续、单调、非降的性质。

随机向量的分布函数，也分离散型和连续型。对于离散型，概率分布集中在可列点上，可以把离散型一元随机变量的概率分布的概念推广到此。对于连续型随机向量，可由式（1.3.2）表示的联合分布函数定义它的联合密度函数

$$f(x_1,\ x_2,\ \cdots,\ x_n)=\frac{\partial^n F(x_1,x_2,\cdots,x_n)}{\partial x_1\partial x_2\partial x_3\cdots\partial x_n} \tag{1.3.5}$$

密度函数满足下述两个条件：

1)　$f(x_1,\ x_2,\ \cdots,\ x_n)\geqslant 0$　(1.3.6)

2)　$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f(x_1,\ x_2\cdots,\ x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n=1$　(1.3.7)

对多维随机向量概率分布性质的叙述和讨论在数学上较为复杂，因此首先讨论它的特例——二维随机向量，再把讨论的结果引用到多维随机向量。

1-3-2　二维随机向量和随机变量的独立性

设离散型二维随机向量$(X,\ Y)$，X的取值为x_1，x_2，…，Y的取值为y_1，y_2，…。

记二维随机向量联合概率分布为

$$p(x_i,\ y_i)=P\{X=x_i,\ Y=y_i\} \tag{1.3.8}$$

并记

$$p_1(x_i)=P\{x=x_i\}=\sum_j p(x_i,\ y_j) \tag{1.3.9}$$

和

$$p_2(y_j)=P\{y=y_j\}=\sum_i p(x_i,\ y_j) \tag{1.3.10}$$

上面的$p_1(x_i)$和$p_2(y_j)$称为边缘概率或非条件概率。可以把X和Y的联合分布看作一个二维概率阵列，阵列中的每个元素表示X和Y的一个具体值组合出现的概率，阵列中全部数的总和显然应该是1，即 $\sum_{i,j} p(x_i,\ y_j)=1$。

此外，把阵列水平方向的各个值或垂直方向的各个值加起来得到的就是边缘概率。下面举例说明这些概念。

例1.5 一口袋里装有2只白球和3只黑球，现连续进行两次摸球，球取出后不放回。定义随机变量X为第一次摸球的结果，随机变量Y为第二次摸球的结果。它们的取值为

$$x_i=\begin{cases}1, & \text{第一次模出白球}\\ 0, & \text{第一次模出黑球}\end{cases}$$

$$y_j=\begin{cases}1, & \text{第二次模出白球}\\ 0, & \text{第二次模出黑球}\end{cases}$$

〔X，Y〕的联合概率分布与非条件概率分布由表1.1给出。中央四方格中为〔X，Y〕的联合概率分布的值。最右一列表示Y的非条件分布；最下一行则是X的非条件分布的值。

表 1.1 不放回摸球的概率分布

Y \\ X	0	1	$p_2(y_j)$
0	$\frac{3}{5}\cdot\frac{2}{4}$	$\frac{2}{5}\cdot\frac{3}{4}$	$\frac{3}{5}$
1	$\frac{3}{5}\cdot\frac{2}{4}$	$\frac{2}{5}\cdot\frac{1}{4}$	$\frac{2}{5}$
$p_1(x_i)$	$\frac{3}{5}$	$\frac{2}{5}$	

从上例还可以研究某次试验结果已知条件下另一次试验结果的概率分布，即条件概率分布

$$p(x_i/y_j)=P\{X=x_i/Y=y_j\}$$

和

$$p(y_j/x_i)=P\{Y=y_j/X=x_i\}$$

它们与联合概率分布和非条件概率分布之间存在下述关系

$$p(x_i/y_j)=p(x_i,\ y_j)/p_2(y_j) \tag{1.3.11}$$

$$p(y_j/x_i)=p(x_i,\ y_j)/p_1(x_i) \tag{1.3.12}$$

以及贝叶斯规则

$$p(x_i/y_j)=p(y_j/x_i)p_1(x_i)/p_2(y_j) \tag{1.3.13}$$

如果对所有的x_i和y_j，存在下述关系式

$$p(x_i,\ y_j)=p_1(x_i)\cdot p_2(y_j) \tag{1.3.14}$$

那么，把它代入式（1.3.11）和式（1.3.12），分别可得

$$p(x_i/y_j)=p_1(x_i) \tag{1.3.15}$$

和

$$p(y_j/x_i)=p_2(y_j) \tag{1.3.16}$$

以上两式说明，随机变量X的概率分布不受Y取值的影响，随机变量Y的概率分布也不受X

取值的影响。因此，定义使（1.3.14）式成立的随机变量X和Y是统计独立的。

与一元随机变量的情况一样，二元连续随机变量（X，Y）的概率分布可以用分布函数或者分布密度函数来描述。这里X和Y均为连续型随机变量。根据式（1.3.2），它们的联合分布函数定义为

$$F(x,\ y)=P\{X\leqslant x,\ Y\leqslant y\} \tag{1.3.17}$$

连续型随机变量X和Y的联合密度函数由下式给出，即

$$f(x,y)=\frac{\partial^2 F(x,y)}{\partial x\partial y} \tag{1.3.18}$$

X和Y的取值落在xy平面上某个R区域内的联合概率，可以用二重积分公式表示

$$P\{X和Y落在R区域内\}=\iint_R f(x,\ y)\mathrm{d}x\,\mathrm{d}y \tag{1.3.19}$$

图1.3是它的几何表示。如果域R是一个矩形微分域（如图1.4所示），则X和Y落在该矩形内的概率为

$$P\{x_0<X\leqslant x_0+\mathrm{d}x,\ y_0<Y\leqslant y_0+\mathrm{d}y\}=f(x_0,\ y_0)\mathrm{d}x\,\mathrm{d}y \tag{1.3.20}$$

可以看出，在二元随机变量的情况下，$f(x,\ y)$仍具有密度的意义。

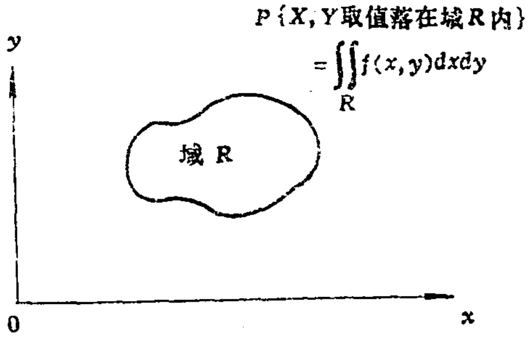


图 1.3　在XY平面上的域R

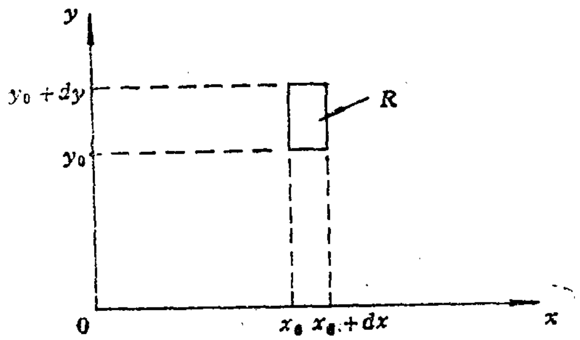


图 1.4　在XY平面上的矩形微分域R

用与离散型情况类似的方法，可以得到边缘密度或非条件密度，所不同的是用积分而不是用求和来表示。因此有

$$f_1(x)=\int_{-\infty}^{\infty}f(x,\ y)\mathrm{d}y \tag{1.3.21}$$

和

$$f_2(y)=\int_{-\infty}^{\infty}f(x,\ y)\mathrm{d}x \tag{1.3.22}$$

离散条件概率的方程（1.3.11）和（1.3.12）也可以应用到微分区域的密度函数，结果导出类似的条件密度方程。图1.4所示的微分域可以用于下面的推导。从条件概率的基本定义出发，有

$$P\{X落在区间\mathrm{d}x内/Y落在区间\mathrm{d}y内\}=\frac{f(x_0,y_0)\mathrm{d}x\mathrm{d}y}{f_2(y_0)\mathrm{d}y} \tag{1.3.23}$$

和

$$P\{x_0<X\leqslant x_0+\mathrm{d}x/Y=y_0\}=\left[\frac{f(x_0,y_0)}{f_2(y_0)}\right]\mathrm{d}x \tag{1.3.24}$$

式(1.3.24)右边括号中的量具有密度函数的全部特征，式左边表示的是概率。因此，可以定义

条件密度为

$$f(x/y)=f(x, y)/f_2(y) \tag{1.3.25}$$

与上述过程类似，$f(y/x)$定义为

$$f(y/x)=f(x, y)/f_1(x) \tag{1.3.26}$$

由方式（1.3.25）和式（1.3.26）可得

$$f(x/y)=f(y/x)f_1(x)/f_2(y) \tag{1.3.27}$$

定义连续型随机变量X和Y的统计独立性，可以用类似于离散型随机变量定义统计独立性的方法，即如

$$f(x, y)=f_1(x)f_2(y) \tag{1.3.28}$$

成立，则X和Y是统计独立的。

1-3-3 相关、协方差和正交性

两随机变量X和Y乘积的数学期望是特别重要的。在一般情况下，可以由下式给出

$$E\{XY\}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x,y)\mathrm{d}x\,\mathrm{d}y \tag{1.3.29}$$

当X和Y彼此独立时，式（1.3.29）就显得特别简单。这时，$f(X, Y)$可以分解（见式（1.3.28））。因而式（1.3.29）可简化为

$$\begin{aligned}E\{XY\}&=\int_{-\infty}^{\infty}xf_1(x)\,\mathrm{d}x\int_{-\infty}^{\infty}yf_2(y)\mathrm{d}y\\&=E\{X\}E\{Y\}\end{aligned} \tag{1.3.30}$$

如果X和Y具有式（1.3.30）所表示的特性，即如果随机变量乘积的数学期望是单个随机变量数学期望的乘积，则它们是不相关的。显然，如果X和Y是独立的，那么它们也是不相关的。但是反过来不一定成立。即X和Y虽不相关，但并不一定是统计独立的。

如果$E\{XY\}=0$，就说X和Y正交。

X和Y的协方差用$C(X, Y)$来表示，并由下式定义

$$C(X, Y)=E\{(X-m_X)(Y-m_Y)\} \tag{1.3.31}$$

式中m_X、m_Y分别为随机变量X和Y的均值。

由式（1.3.31），可定义两个随机变量的相关系数为

$$\begin{aligned}\rho(\text{相关系数})&=\frac{C(X,Y)}{\sqrt{\mathrm{Var}X}\sqrt{\mathrm{Var}Y}}\\&=\frac{E\{(X-m_X)(Y-m_Y)\}}{\sqrt{\mathrm{Var}X}\sqrt{\mathrm{Var}Y}}\end{aligned} \tag{1.3.32}$$

相关系数是两个随机变量相关程度的归一化量度。规一化的ρ值总是在-1和$+1$之间，即$-1\leqslant\rho\leqslant1$。考察下面三种特殊情况可以说明$\rho$的含义（这里不是证明）。

1) $Y=X$（最大值正相关）。

当$Y=X$时，式（1.3.32）为

$$\rho=\frac{E\{(X-m_X)(X-m_X)\}}{\sqrt{E\{(X-m_X)^2\}}\sqrt{E\{(X-m_X)^2\}}}=1$$

2) $Y=-X$(最大值负相关)。

当$Y=-X$时，式(1.3.32)为

$$\rho=\frac{E\{(X-m_X)(-X+m_X)\}}{\sqrt{E\{(X-m_X)^2\}}\sqrt{E\{(-X+m_X)^2\}}}=-1$$

3) X和Y不相关，即$E\{XY\}=E\{X\}E\{Y\}$。这时，把式（1.3.32）展开，得到

$$\begin{aligned}E\{(X-m_X)(Y-m_Y)\}&\\&=E\{XY\}-m_XE\{Y\}-m_YE\{X\}+m_Xm_Y\\&=m_Xm_Y-m_Xm_Y-m_Ym_X+m_Xm_Y=0\end{aligned}$$

因此，$\rho=0$

上面验证了正相关、负相关和零相关三种极端情况，其他情况处于它们之间。

1-3-4 二元正态分布

设随机变量X和Y的均值分别为

$$m_X=E\{X\}和m_Y=E\{Y\} \tag{1.3.33}$$

方差分别为

$$\sigma_X{}^2=E\{(X-m_X)^2\}$$

和

$$\sigma_Y{}^2=E\{(Y-m_Y\}^2\} \tag{1.3.34}$$

二者之间的相关系数为

$$\rho=\frac{E\{X-m_X)(Y-m_Y)\}}{\sigma_X\sigma_Y} \tag{1.3.35}$$

这样，二维正态随机向量〔X，Y〕的联合分布密度函数为

$$f(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp\left[-\frac{\dfrac{(x-m_X)^2}{2\sigma_X{}^2}-\rho\dfrac{(x-m_X)(y-m_Y)}{\sigma_X\sigma_Y}+\dfrac{(y-m_Y)^2}{2\sigma_Y{}^2}}{1-\rho^2}\right] \tag{1.3.36}$$

因为$\rho^2\leqslant1$，且 $\left(\dfrac{x-m_X}{\sigma_X}\pm\dfrac{y-m_Y}{\rho_Y}\right)^2\geqslant0$，所以式（1.3.36）中的指数永远为负， 在（$m_X$，$m_Y$）处取极大值零。分布密度函数的极大值为

$$f(m_X,m_Y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

在图1.5中用xy平面上的等概率密度线表示这个分布。这些等概率密度线是同轴椭圆，其主轴的倾斜角是$\alpha=\frac{1}{2}\tan^{-1}2\rho\sigma_X\sigma_Y/(\sigma_X{}^2-\sigma_Y{}^2)$。如$\rho=0$，即$X$和$Y$不相关。这时，椭圆的长半轴和短半轴分别平行于$x$轴和$y$轴。如果进一步假设$\sigma_X=\sigma_Y$（且$\rho=0$），则椭圆退化为圆。反之，$|\rho|$越接近于1，椭圆就越扁。

$\rho=0$的非相关情况是特别重要的。这时，$f(x，y)$简化形式由式（1.3.37）给出，即

$$f(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y}\exp\left\{-\frac{1}{2}\left[\frac{(x-m_X)^2}{\sigma_X{}^2}+\frac{(y-m_Y)^2}{\sigma_Y{}^2}\right]\right\}$$

$$=\frac{1}{\sqrt{2\pi}\,\sigma_X}e^{-(x-m_X)^2/2\sigma_X}\ \frac{1}{\sqrt{2\pi}\,\sigma_Y}e^{-(y-m_Y)^2/2\sigma_Y{}^2} \tag{1.3.37}$$

联合分布密度函数分解为$f_1(x)$和$f_2(y)$的乘积。由此得出正态分布的一个重要性质：两个