第 1 章　随机信号基础

信号有多种表现形式，主要的形式有电信号、光信号、声信号等；根据表达式的不同，还可以分为连续时间信号和离散时间信号，或者分为确定性信号和随机信号。连续时间信号和离散时间信号的区别在于自变量是连续的还是离散的，而确定性信号和随机信号的区别才是本质的区别，因为确定性信号是以时间为自变量的一般函数，随机信号则是以时间为自变量的随机函数。

在实际应用中，需要处理的信号往往不是确定性信号，而是随机信号与确定性信号的混合信号。由于随机信号与确定性信号有本质上的不同，因此分析方法也不尽相同。

随机信号理论的基础是“概率论”和“信号与系统”，这里假定读者已经掌握了这些知识。本章首先对随机变量的要点做一下系统的回顾；然后介绍用特征函数描述随机变量的方法。本章的后半部分将给出通信与信息处理领域中经常用到的一些随机变量的分布，并重点讨论高斯随机变量。本章还将给出一些随机变量仿真的方法和程序，供读者参考和选用。

1.1　随机变量及其分布

设随机试验的样本空间为 $S=\{e_i\}$，如果对样本空间的每一个元素 $e_i \in S$，都有一个实数 $X\{e_i\}$ 与之对应，对所有的元素 $e \in S$，就得到一个定义在空间 S 上的实单值函数 $X\{e\}$，称 $X\{e\}$ 为随机变量，简写为 X。一般用大写字母 X, Y, Z 来表示随机变量，而用小写字母 x, y, z 表示对应随机变量的可能取值。

引入随机变量可以将随机试验的所有可能结果与对应的概率联系起来。如一段导体中的电子运动引起的电流，接收机的噪声电压，这些都与数值有关。即使像发现目标这样的事件，也可以规定一个数值来表示“发现目标”或“未发现目标”。分布律便表明了随机变量取值与概率的对应关系。

根据随机变量的取值是可列还是不可列的，把随机变量分为离散随机变量和连续随机变量。离散随机变量的样本空间是离散的点，因而取值也是离散的，如图 1.1-1(a) 所示。连续随机变量的样本空间是连续区间，如图 1.1-1(b) 所示，所以取值连续地占据某一区间。接收机的噪声电压是连续随机变量，而探测是否存在目标的试验则是离散随机变量。

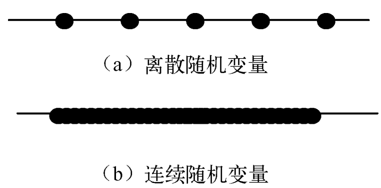


图 1.1-1　随机变量

根据描述随机试验参量的数目，还可以把随机变量分为一维、二维和多维随机变量。例如，随机变量 X 只能用来描述一个随机量，若用它来描述一个随机信号的幅度和相位是不够的，必须用两个随机变量 X 和 Y。对于更复杂的随机试验，可能用更多的随机变量进行描述。

1.1.1　一维随机变量及其分布律

研究确定性函数 $y=f(x)$ 时，如果是单值函数，可以唯一确定 x 和 y 的关系。随机试验

的某一结果是否出现并不能根据函数关系决定，因此无法用函数唯一地表示。例如，在一次掷硬币试验中，在试验前是否“出现正面”是未知的，但是通过大量的试验可以得到“出现正面”的概率为 0.5 这一结论。

通过大量试验得到的结果就是统计规律，那么如何研究随机变量的统计规律呢？分布律就是研究随机变量统计规律的一种方法，它描述了随机变量各可能取值与相应的概率之间的对应关系。

1．一维离散随机变量的概率分布列

设离散随机变量 X 的所有取值为 x_k， $k=1,2,\cdots$ 其概率为

$$P\{X=x_k\}=P_k \qquad k=1,2,\cdots \tag{1.1-1}$$

称式(1.1-1)为离散型随机变量 X 的概率分布列或分布律，并可用表格的形式来表示

X	$x_1 \quad x_2 \cdots x_n \cdots$
P_k	$P_1 \quad P_2 \cdots P_n \cdots$

有
$$P_k \geqslant 0, \quad \sum_k P_k=1 \qquad k=1,2,\cdots$$

2．一维随机变量的概率分布函数

对于连续随机变量 X，由于不能一一列举其可能取值，任意指定的取值的概率趋近于零，不能用概率分布列描述它，所以我们研究取值落在一个区间的概率 $P\{x_1<X\leqslant x_2\}$。

定义随机变量 X 取值不超过 x 的概率为概率分布函数或累积分布函数

$$F(x)=P(X\leqslant x) \tag{1.1-2}$$

如果把一维随机变量看成是数轴上的一个随机点，上式说明了随机变量 X 取值落在区间 $(-\infty,x]$ 的概率，显然它既适用于离散随机变量，也适用于连续随机变量。根据概率分布函数的定义，可得到如下性质。

性质 1　$F(x)$是 x 的单调非减函数。即对于 $x_2>x_1$，有

$$F(x_2)\geqslant F(x_1) \tag{1.1-3}$$

性质 2　$F(x)$非负，且取值满足

$$0\leqslant F(x)\leqslant 1 \tag{1.1-4}$$

性质 3　随机变量在 x_1,x_2 区间内的概率为

$$P(x_1<X\leqslant x_2)=F(x_2)-F(x_1) \tag{1.1-5}$$

性质 4　$F(x)$右连续，即

$$F(x^+)=F(x) \tag{1.1-6}$$

性质 4 对离散随机变量特别有用。对于任意一个函数，看它是否为概率分布函数的正确表达式，只要用性质 1、性质 2 和性质 4 判断即可。离散随机变量的分布函数除满足以上性质外，还具有阶梯形式，阶跃的高度等于随机变量在该点的概率，即

$$F(x)=\sum_{i=1}^{\infty}P(X=x_i)u(x-x_i)=\sum_{i=1}^{\infty}P_i u(x-x_i) \tag{1.1-7}$$

式中，$u(x)$为单位阶跃函数，P_i 为 $X=x_i$ 的概率。

【例 1.1-1】设随机变量 X 只取两个值，其概率分布为 $P\{X=0\}=0.5$，$P\{X=1\}=0.5$，试写出概率分布函数。

解：随机变量 X 的概率分布函数为

$$F(x)=0.5u(x)+0.5u(x-1)$$

或写成

$$F(x)=\begin{cases}0 & x<0\\ 0.5 & 0\leqslant x<1\\ 1 & x\geqslant 1\end{cases}$$

3．一维随机变量的概率密度

分布律的另一种形式是概率密度函数，定义为概率分布函数 $F(x)$对 x 的导数

$$f(x)=\frac{\mathrm{d}F(x)}{\mathrm{d}x} \tag{1.1-8}$$

或写成积分形式

$$F(x)=\int_{-\infty}^{x} f(\lambda)\mathrm{d}\lambda \tag{1.1-9}$$

如果概率分布函数是连续的，其导数一定存在，故概率密度存在。如果概率分布函数存在有限个间断点，则可引入δ函数，因此概率密度总是存在的。由概率密度函数的定义可以看出，$f(x)$是 x 处 $F(x)$的变化率。根据概率分布函数的性质，可得到概率密度的性质。

性质 1　概率密度函数非负，即

$$f(x)\geqslant 0 \tag{1.1-10}$$

性质 2　概率密度函数在整个取值区间积分为 1，即

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x=1 \tag{1.1-11}$$

性质 3　概率密度函数在 $(x_1,x_2]$ 区间积分，给出该区间的取值概率

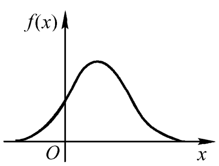
$$P(x_1<X\leqslant x_2)=\int_{x_1}^{x_2} f(x)\mathrm{d}x \tag{1.1-12}$$

这三条性质与概率分布函数的前三条性质是对应的。性质 1 和性质 2 说明概率密度函数是一条在横轴上方且与横轴所围的面积为 1 的曲线，它们也是检验一个函数是否为概率密度的条件。离散随机变量的概率密度为

$$f(x)=\sum_{i=1}^{\infty} P(X=x_i)\delta(x-x_i)=\sum_{i=1}^{\infty} P_i\delta(x-x_i) \tag{1.1-13}$$

式中，$\delta(x)$为单位冲激函数。

概率分布函数 $F(x)$和概率密度 $f(x)$可以充分说明离散随机变量取值落在某点和某个区间的概率，而连续随机变量取值落在某一区间的概率也可由 $F(x)$和 $f(x)$求出。需要注意的是：连续随机变量在某点取值的概率为零。因此，对于连续随机变量，取值区间写成开区间和闭区间是一样的，但对于离散随机变量，开区间和闭区间则是不同的。图 1.1-2 和图 1.1-3 示出了连续随机变量和离散随机变量的分布律。



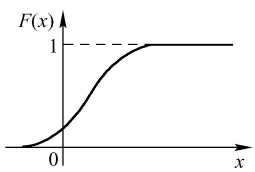


图 1.1-2　连续随机变量概率密度和概率分布函数

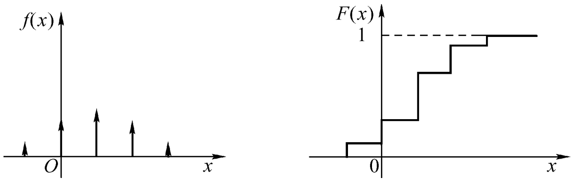


图 1.1−3　离散随机变量的概率密度和概率分布函数

【例 1.1−2】 判断函数 $f(x)=K[u(x)-u(x-a)]$ 满足什么条件才有可能是概率密度函数？当 $a=2$ 和 $a=-2$ 时，K 应该取何值？

解：（1）为保证满足性质 1，概率密度函数非负。当 $a>0$ 时需要 $K>0$；当 $a<0$ 时需要 $K<0$。由性质 2 可知，概率密度函数与横轴包围的面积应该为 1，因此 $K=1/a$。综上，$f(x)=\dfrac{1}{a}[u(x)-u(x-a)]$ 满足概率密度函数的条件。

（2）当 $a=2$ 时 $K=0.5$；当 $a=-2$ 时，$K=-0.5$。

1.1.2　多维随机变量及其分布律

二维随机变量用 (X, Y) 表示，可认为它是二维平面上的一个随机点（图 1.1−4）。n 维随机变量则用 $(X_1, X_2, X_3, \cdots, X_n)$ 表示，可推广为 n 维空间上的一个随机点。

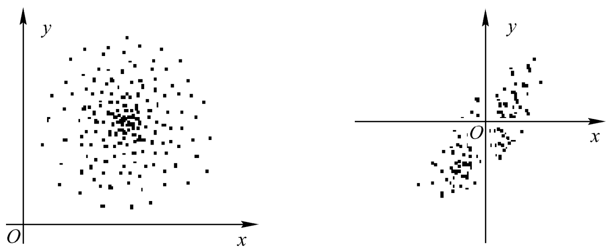


图 1.1−4　二维随机变量——平面上的随机点

多维随机变量不是几个一维随机变量的简单组合，作为一个整体，多维随机变量的统计规律不仅取决于各个随机变量的统计规律，还与几个随机变量之间的关联程度有关。由一维随机变量的分布律不难推广到二维随机变量的分布律。

1．二维离散随机变量的概率分布列

如果二维随机变量 (X, Y) 的可能取值有限个或无限可列，则 (X, Y) 为离散型随机变量。

$$P\{X=x_i, Y=y_j\}=P_{ij}, \qquad i,j=1,2,\cdots \tag{1.1−14}$$

称为 (X, Y) 的概率分布列，或称为 X 和 Y 的联合概率，有

$$P_{ij}\geqslant 0,\ \sum_i\sum_j P_{ij}=1 \tag{1.1−15}$$

2．二维随机变量的概率分布函数

(X, Y) 为二维随机变量，对于任意实数 x, y 的二元函数

$$F_{XY}(x,y)=P(X\leqslant x,\ Y\leqslant y) \tag{1.1−16}$$

称为 (X, Y) 的分布函数或 X 和 Y 的联合分布函数。

若将(X, Y)看成平面上随机点的坐标，则分布函数$F_{XY}(x, y)$在(x, y)处的值即为随机点(X, Y)落在图 1.1−5 所示阴影部分面积内的概率。根据概率分布函数的定义，可得到如下性质。

性质 1　$F_{XY}(x,y)$ 是 x, y 的单调非减函数。

对固定的 y，当 $x_2>x_1$ 时，有

$$F_{XY}(x_2,y) \geqslant F_{XY}(x_1,y) \tag{1.1−17}$$

对固定的 x，当 $y_2>y_1$ 时，有

$$F_{XY}(x,y_2) \geqslant F_{XY}(x,y_1) \tag{1.1−18}$$

性质 2

$$0 \leqslant F_{XY}(x,y) \leqslant 1 \tag{1.1−19}$$

且有

$$F_{XY}(-\infty,y)=0\text{，}\quad F_{XY}(x,-\infty)=0\text{，}\quad F_{XY}(-\infty,-\infty)=0 \tag{1.1−20}$$

$$F_{XY}(+\infty,+\infty)=1 \tag{1.1−21}$$

性质 3　随机变量 X 的边缘分布函数为

$$F_X(x)=F_{XY}(x,+\infty) \tag{1.1−22}$$

随机变量 Y 的边缘分布函数为

$$F_Y(y)=F_{XY}(+\infty,y) \tag{1.1−23}$$

性质 4　对任意(x_1,y_1)和(x_2,y_2)，其中 $x_2>x_1$，$y_2>y_1$，则有

$$P\{x_1<X\leqslant x_2, y_1<Y\leqslant y_2\}=F_{XY}(x_2,y_2)-F_{XY}(x_2,y_1)-F_{XY}(x_1,y_2)+F_{XY}(x_1,y_1) \tag{1.1−24}$$

上式的结果可以通过图 1.1−6 得出。

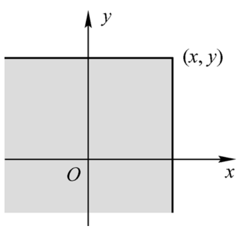


图 1.1−5　二维分布函数图解

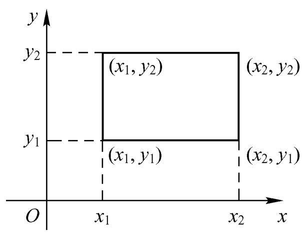


图 1.1−6　二维分布函数性质图解

3. 二维随机变量的概率密度

若二维分布函数$F_{XY}(x, y)$连续并存在二阶混合偏导数，则该二阶偏导数为二维概率密度，即

$$f_{XY}(x,y)=\frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y} \tag{1.1−25}$$

二维概率密度的性质如下：

性质 1　二维概率密度函数非负，即

$$f_{XY}(x,y) \geqslant 0 \tag{1.1−26}$$

性质 2

$$\int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(x,y)\mathrm{d}x\mathrm{d}y = F_{XY}(x,y) \tag{1.1−27}$$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{XY}(x,y)\mathrm{d}x\mathrm{d}y = F_{XY}(+\infty,+\infty)=1 \tag{1.1−28}$$

性质 3　对二维概率密度函数在一个随机变量的所有取值区间上积分，将给出另一个随机变量的概率密度函数

$$f_X(x)=\int_{-\infty}^{\infty} f_{XY}(x,y)\mathrm{d}y \tag{1.1−29}$$

$$f_Y(y)=\int_{-\infty}^{\infty} f_{XY}(x,y)\mathrm{d}x \tag{1.1-30}$$

$f_X(x)$和$f_Y(y)$也称为二维随机变量的边缘概率密度。

性质 4 设 S 是 xOy 平面上的一个区域，则随机点(X, Y)落在该区域内的概率为

$$P\{(X,Y)\in S\}=\oiint_S f_{XY}(x,y)\,\mathrm{d}x\,\mathrm{d}y \tag{1.1-31}$$

上式中的积分是 S 区域的二重面积分。

4. 条件概率分布函数和条件概率密度

在二维分布律中，如果将条件概率的概念引入到分布律中，还可得到条件概率分布函数$F_Y(y|x)$和条件概率密度 $f_Y(y|x)$。在表示概率分布函数和概率密度时，为了区别不同的随机变量，常把随机变量作为下角标。

在已知 $X\leqslant x$ 的条件下，随机变量 Y 的条件概率分布函数和条件概率密度函数分别表示为

$$F_Y(y\,|\,x)=\frac{F_{XY}(x,y)}{F_X(x)} \tag{1.1-32}$$

$$f_Y(y\,|\,x)=\frac{f_{XY}(x,y)}{f_X(x)} \tag{1.1-33}$$

对于所有的 x 和 y，若

$$f_X(x\,|\,y)=f_X(x) \tag{1.1-34}$$

$$f_Y(y\,|\,x)=f_Y(y) \tag{1.1-35}$$

成立，则称 X, Y 是相互统计独立的两个随机变量。将式(1.1−33)～式(1.1−35)联合，便得到两个随机变量 X, Y 相互统计独立的充要条件

$$f_{XY}(x,y)=f_X(x)f_Y(y) \tag{1.1-36}$$

即随机变量 X, Y 的二维联合概率密度等于 X 和 Y 的边缘概率密度的乘积。

5. 多维随机变量的概率分布函数和概率密度

二维分布律是多维分布律最简单的情况，对于 n 维随机变量$(X_1, X_2, X_3, \cdots, X_n)$，仍可仿照式(1.1−16)和式(1.1−25)定义 n 维分布函数和概率密度

$$F_X(x_1,x_2,\cdots,x_n)=P(X_1\leqslant x_1,X_2\leqslant x_2,\cdots,X_n\leqslant x_n) \tag{1.1-37}$$

$$f_X(x_1,x_2,\cdots,x_n)=\frac{\partial^n F_X(x_1,x_2,\cdots,x_n)}{\partial x_1\partial x_2\cdots\partial x_n} \tag{1.1-38}$$

n 维概率密度的性质也类似二维概率密度的性质，对应式(1.1−29)的一条重要性质为

$$f_X(x_1,x_2,\cdots,x_m)=\underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{n-m} f_X(x_1,x_2,\cdots,x_m,\cdots,x_n)\mathrm{d}x_{m+1}\cdots\mathrm{d}x_n \tag{1.1-39}$$

上式说明了高维概率密度可以通过积分降低维数。式(1.1−29)是 $n=2$，$m=1$ 时的情况。

n 维随机变量相互统计独立的充要条件为：对于所有的 $x_1, x_2, x_3,\cdots, x_n$，满足

$$f_X(x_1,x_2,\cdots,x_n)=f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n)=\prod_{i=1}^{n}f_{X_i}(x_i) \tag{1.1-40}$$

若 $n=2$，上式简化为式(1.1-36)。

1.2　随机变量的函数变换

一般来讲，随机变量的分布是由大量的试验获得的。那么，是否所有的随机变量的分布都需要用试验的方法得到呢？试验的高复杂性和高代价促使人们寻求一种间接的方法来确定一个随机变量的分布。

在无线电信号的传输过程中会不可避免地掺杂一些噪声，如果发射的信号为 X，信道中的噪声表示为 Y，那么接收的信号就是 $X+Y$。在已知 X 和 Y 分布的前提下，人们希望能通过一种运算，求得二者之和的分布。如果把随机变量 X_i 作为系统的输入，把随机变量 X_o 作为系统的输出，它们也应满足某种函数关系。直观上看，知道了输入的分布，通过二者的函数关系，一定能得到输出的分布。在进行系统仿真时，常常需要仿真某个分布的信号，当有了均匀分布的随机信号的产生方法后，也可利用函数关系来产生需要的随机信号。这些都是随机变量函数变换的例子。

1．一维变换

设随机变量 X 与 Y 满足下列函数关系

$$Y=\varphi(X) \tag{1.2-1}$$

如果随机变量 X 与 Y 之间的关系是单调的，并且存在反函数

$$X=\varphi^{-1}(Y)=h(Y)$$

若反函数 $h(Y)$的导数也存在，则可利用 X 的概率密度求出 Y 的概率密度。

图 1.2-1 给出了一维随机变量 X 和 Y 的函数关系。根据 X 和 Y 的函数关系，如果 $h(Y)$ 是单调增加的，那么随机变量 Y 的概率分布函数为

$$F_Y(y)=P(Y\leqslant y)=P(\varphi(X)\leqslant y)=P(X\leqslant h(y))=\int_{-\infty}^{h(y)}f(x)\mathrm{d}x \tag{1.2-2}$$

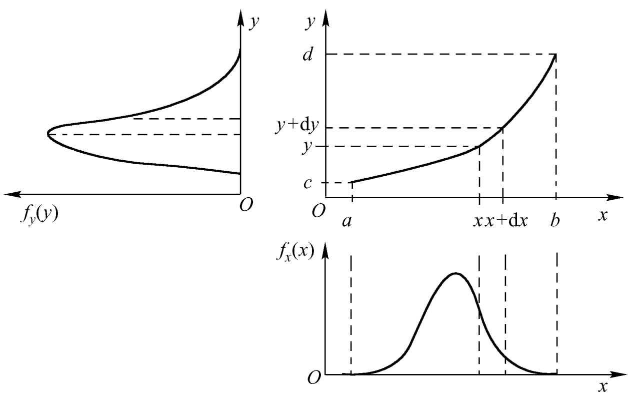


图 1.2-1　一维函数单调变换

将上式对 y 求导，便得到随机变量 Y 的概率密度

$$f_Y(y)=\frac{\mathrm{d}}{\mathrm{d}y}F_Y(y)=f(h(y))\frac{\mathrm{d}}{\mathrm{d}y}h(y)$$

同理，可得到当 $h(Y)$是单调下降时随机变量 Y 的概率密度

$$f_Y(y)=\frac{\mathrm{d}}{\mathrm{d}y}F_Y(y)=-f(h(y))\frac{\mathrm{d}}{\mathrm{d}y}h(y)$$

综合以上两种情况，得到

$$f_Y(y)=f_X(h(y))\left|h'(y)\right| \tag{1.2-3}$$

事实上，也可以从另一个角度来推导上式。设 X 的所有可能值都在区间(a, b)内，对于 Y，所有可能值都在区间(c, d)内，此时应该有

$$P_X(a<X<b)=1,\quad P_Y(c<Y<d)=1$$

由于 X 和 Y 是单调关系，如图 1.2−1 所示，当 $\mathrm{d}x$ 和 $\mathrm{d}y$ 充分小时，随机变量 X 取值落在子区间$(x, x+\mathrm{d}x)$和 Y 的取值落在子区间$(y, y+\mathrm{d}y)$的概率应该相等，即

$$f_X(x)\,\mathrm{d}x=f_Y(y)\,\mathrm{d}y$$

因此

$$f_Y(y)=f_X(x)\frac{\mathrm{d}x}{\mathrm{d}y}=f_X(x)h'(y) \tag{1.2-4}$$

考虑到概率密度非负，无论对单调增函数还是单调减函数，均有

$$f_Y(y)=f_X(h(y))\left|h'(y)\right|$$

【例 1.2−1】 已知随机变量 X 和 Y 满足线性关系 $Y=aX+b$，a 和 b 为常数，X 为高斯变量，其概率密度表示为 $f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_X}\mathrm{e}^{-\frac{(x-m_X)^2}{2\sigma_X^2}}$。求 Y 的概率密度。

解：因为 Y 和 X 是严格单调函数关系，其反函数

$$X=h(Y)=(Y-b)/a$$

的导数存在，即

$$h'(Y)=1/a$$

将上式代入式(1.2−3)，即可得到 Y 的概率密度

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_X}\exp\left[-\frac{\left(\frac{y-b}{a}-m_X\right)^2}{2\sigma_X^2}\right]\left|\frac{1}{a}\right|=\frac{1}{\sqrt{2\pi}\,|a|\,\sigma_X}\exp\left[-\frac{(y-am_X-b)^2}{2a^2\sigma_X^2}\right]$$

$$=\frac{1}{\sqrt{2\pi}\sigma_Y}\exp\left[-\frac{(y-m_Y)^2}{2\sigma_Y^2}\right]$$

上例说明，高斯变量 X 经过线性变换后的随机变量 Y 仍然是高斯分布，其数学期望和方差分别为

$$m_Y=am_X+b,\quad \sigma_Y^2=a^2\sigma_X^2$$

如果 X 和 Y 不是单调关系，那么 Y 的取值 y 就对应 X 的两个或更多的值 $x_1,x_2,x_3,\cdots,x_n$。以双值函数为例，如图 1.2−2 所示，反函数应为

$$\begin{cases}X_1=h_1(Y)\\X_2=h_2(Y)\end{cases} \tag{1.2-5}$$

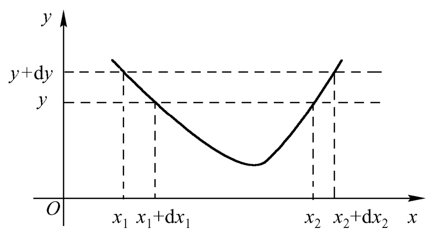


图 1.2−2　一维函数多值变换

这时随机变量 Y 的取值落在子区间 $(y, y+\mathrm{d}\,y)$ 时，对应随机变量 X 的取值应落在两个子区间 $(x_1, x_1+\mathrm{d}\,x_1)$ 和 $(x_2, x_2+\mathrm{d}\,x_2)$ 中，遵循等概率原理，有

$$f_Y(y)\,\mathrm{d}\,y = f_X(x_1)\,\mathrm{d}\,x_1 + f_X(x_2)\,\mathrm{d}\,x_2 \tag{1.2-6}$$

于是

$$f_Y(y) = f_X(h_1(y))\left|h_1'(y)\right| + f_X(h_2(y))\left|h_2'(y)\right| \tag{1.2-7}$$

当 Y 的取值 y 对应多个 x 值时，其概率密度可由上式推广。

2．二维变换

在讨论二维随机变量变换时，仍假定函数的映射关系是单值的。如果已知二维随机变量(X_1, X_2)的联合概率密度 $f_X(x_1, x_2)$，以及二维随机变量(Y_1, Y_2)与(X_1, X_2)之间的函数关系为

$$\begin{cases} Y_1 = \varphi_1(X_1, X_2) \\ Y_2 = \varphi_2(X_1, X_2) \end{cases} \tag{1.2-8}$$

它们的反函数存在

$$\begin{cases} X_1 = h_1(Y_1, Y_2) \\ X_2 = h_2(Y_1, Y_2) \end{cases}$$

即可求出随机变量(Y_1,Y_2)的联合概率密度。仿照图 1.2−1，给出它们之间的映射关系，见图 1.2−3。

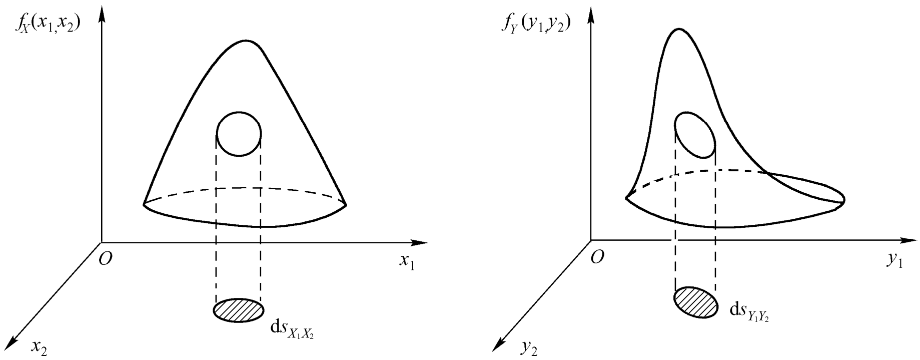


图 1.2−3　二维随机变量函数变换

如果(X_1, X_2)的联合概率密度和(Y_1, Y_2)的联合概率密度之间为单值映射，当 $\mathrm{d}s_{X_1X_2}$ 和 $\mathrm{d}s_{Y_1Y_2}$ 是充分小的区域时，随机变量(X_1, X_2)的取值落在 $\mathrm{d}s_{X_1X_2}$ 区域内的概率应等于随机变量(Y_1, Y_2)取值落在 $\mathrm{d}s_{Y_1Y_2}$ 区域内的概率。一维随机变量在某区间取值的概率等于一维概率密度（曲线）在该区间积分的面积；而二维随机变量(X_1, X_2)或(Y_1, Y_2)在某区域取值的概率应为二维概率密度（曲面）下的体积，于是有

$$f_X(x_1, x_2)\,\mathrm{d}\,s_{X_1X_2} = f_Y(y_1, y_2)\,\mathrm{d}\,s_{Y_1Y_2}$$

注意到联合概率密度非负，应该有

$$f_Y(y_1, y_2) = f_X(x_1, x_2)\left|\frac{\mathrm{d}\,s_{X_1X_2}}{\mathrm{d}\,s_{Y_1Y_2}}\right| \tag{1.2-9}$$

在用二重积分求体积时，若积分区域由 $\mathrm{d}\,s_{X_1X_2}$ 变为 $\mathrm{d}\,s_{Y_1Y_2}$，其变换关系即为雅可比行列式

$$J=\frac{\mathrm{d}s_{X_1X_2}}{\mathrm{d}s_{Y_1Y_2}}=\begin{vmatrix}\dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2}\\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2}\end{vmatrix} \tag{1.2-10}$$

将上式代入式(1.2-9)，并将式中的 x_1 和 x_2 换成 $h_1(y_1,y_2)$和 $h_2(y_1,y_2)$，便得到二维函数变换的最后表达式

$$f_Y(y_1,y_2)=|J|f_X(x_1,x_2)=|J|f_X(h_1(y_1,y_2),h_2(y_1,y_2)) \tag{1.2-11}$$

由于要求概率密度函数非负，无论对 x_1 和 x_2 是否都为单调增函数，都可以利用求雅可比行列式的绝对值来保证概率密度函数非负。下面通过具体的例子说明二维函数变换的应用。

【例 1.2-2】 设 X,Y 是互相独立的高斯变量，它们的联合概率密度为

$$f_{XY}(x,y)=f_X(x)f_Y(y)=\frac{1}{2\pi\sigma^2}\exp\left[-\frac{x^2+y^2}{2\sigma^2}\right]$$

A 和$\varPhi$为随机变量，且$\begin{cases}X=A\cos\varPhi\\ Y=A\sin\varPhi\end{cases}$，$A>0,\ 0\leqslant\varPhi\leqslant 2\pi$。求 $f_{A\varPhi}(a,\varphi)$, $f_A(a)$ 和 $f_\varPhi(\varphi)$ 。

解：由于给出的条件即为反函数，可直接求雅可比行列式

$$J=\begin{vmatrix}\dfrac{\partial x}{\partial a} & \dfrac{\partial x}{\partial\varphi}\\ \dfrac{\partial y}{\partial a} & \dfrac{\partial y}{\partial\varphi}\end{vmatrix}=\begin{vmatrix}\cos\varphi & -a\sin\varphi\\ \sin\varphi & a\cos\varphi\end{vmatrix}=a$$

代入式(1.2-11)，得到 $A,\ \varPhi$ 的联合概率密度

$$f_{A\varPhi}(a,\varphi)=\frac{a}{2\pi\sigma^2}\exp\left[-\frac{x^2+y^2}{2\sigma^2}\right]=\frac{a}{2\pi\sigma^2}\exp\left[-\frac{a^2}{2\sigma^2}\right]$$

式中，$a^2=x^2+y^2$。再利用概率密度降维的性质，对$\varPhi$积分求 A 的概率密度

$$f_A(a)=\int_0^{2\pi}\frac{a}{2\pi\sigma^2}\exp\left[-\frac{a^2}{2\sigma^2}\right]\mathrm{d}\varphi=\frac{a}{\sigma^2}\exp\left[-\frac{a^2}{2\sigma^2}\right]$$

同样可利用概率密度降维的性质，对 A 积分求$\varPhi$的概率密度

$$f_\varPhi(\varphi)=\int_0^{\infty}\frac{a}{2\pi\sigma^2}\exp\left[-\frac{a^2}{2\sigma^2}\right]\mathrm{d}a=\int_0^{\infty}\frac{1}{2\pi}\exp\left[-\frac{a^2}{2\sigma^2}\right]\mathrm{d}\frac{a^2}{2\sigma^2}$$

令 $t=a^2/2\sigma^2$，则 $$f_\varPhi(\varphi)=\frac{1}{2\pi}\int_0^{\infty}\mathrm{e}^{-t}\mathrm{d}t=\frac{1}{2\pi}$$

【例 1.2-3】 已知二维随机变量(X_1, X_2)的联合概率密度 $f_X(x_1,x_2)$，求 $Y=X_1+X_2$ 的概率密度。

解：设 $$\begin{cases}Y_1=X_1\\ Y_2=X_1+X_2\end{cases}$$

做这样的假设是为了保证运算过程的简单，也可做其他形式的假设。先求随机变量 Y_1，Y_2 的反函数及雅可比行列式

$$\begin{cases}X_1=Y_1\\ X_2=Y_2-Y_1\end{cases}$$

$$J=\begin{vmatrix}\dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2}\\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2}\end{vmatrix}=\begin{vmatrix}1 & 0\\ -1 & 1\end{vmatrix}=1$$

代入式(1.2−11)即可得到二维随机变量(Y_1, Y_2)的联合概率密度

$$f_Y(y_1,y_2)=|J|f_X(x_1,x_2)=f_X(x_1,x_2)=f_X(y_1,y_2-y_1)$$

利用概率密度的性质求 Y_2 的边缘概率密度

$$f_{Y_2}(y_2)=\int_{-\infty}^{\infty}f_X(y_1,y_2-y_1)\,\mathrm{d}y_1$$

最后用 Y 和 X_1 代替 Y_2 和 Y_1

$$f_Y(y)=\int_{-\infty}^{\infty}f_X(x_1,y-x_1)\,\mathrm{d}x_1$$

这就是两个随机变量之和的概率密度。进一步地，如果 X_1 和 X_2 互相独立，则

$$f_Y(y)=\int_{-\infty}^{\infty}f_{X_1}(x_1)f_{X_2}(y-x_1)\,\mathrm{d}x_1=f_{X_1}(y)*f_{X_2}(y) \tag{1.2−12}$$

这是常见的卷积公式，也就是说两个互相独立随机变量之和的概率密度等于两个随机变量概率密度的卷积。

这个例子给出了两个随机变量之和的概率密度，用同样的方法也可求出两个随机变量之差、积、商的概率密度。

【例 1.2−4】 任选两个标有阻值 20kΩ的电阻 R_1 和 R_2 串联，两个电阻的误差都在±5%之内，并且在误差之内它们是均匀分布的。求 R_1 和 R_2 串联后误差不超过±2.5%的概率有多大？

解： 由题意已知，电阻 R_1 和 R_2 应在 19～21kΩ内均匀分布。另一方面，虽然电子元件出厂时都存在一定的误差，但由于 R_1 和 R_2 是任选的，两个电阻值应该是互相独立的。因此 R_1 和 R_2 之和的分布应满足式(1.2−12)。

两个矩形脉冲卷积的结果是三角波形，因此可推论两个均匀分布随机变量之和的概率密度应该呈三角形状，如图 1.2−4 所示。假定 $a=19\text{k}\Omega$，$b=21\text{k}\Omega$，R 的概率密度为

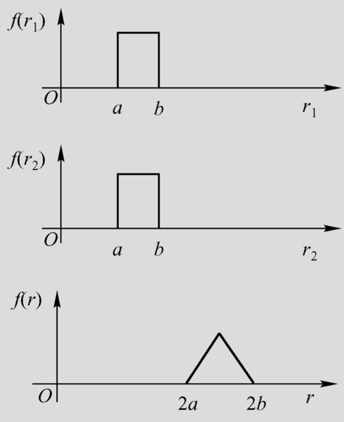


图 1.2−4 例 1.2−4 的图

$$f_R(r)=\begin{cases}\dfrac{r-2a}{(b-a)^2}, & 2a\leqslant r<a+b\\ -\dfrac{r-2b}{(b-a)^2}, & a+b\leqslant r\leqslant 2b\\ 0, & \text{其他}\end{cases}$$

R_1 和 R_2 串联后，$R=R_1+R_2$ 的阻值应是 40kΩ，误差范围也随之增大，这时 R 的取值应该在38～42kΩ之间。求串联后 R 的相对误差在±2.5%之内的概率，也就是求 R 的取值区间在 39～41kΩ之内的概率

$$P(39 \leqslant R \leqslant 41) = \int_{39}^{41} f_R(r)\,\mathrm{d}r = \int_{39}^{40} \frac{r-2a}{(b-a)^2}\,\mathrm{d}r - \int_{40}^{41} \frac{r-2b}{(b-a)^2}\,\mathrm{d}r = \frac{3}{4}$$

即 R_1 和 R_2 串联后误差不超过±2.5%的概率是 0.75。

1.3　随机变量及其函数的数字特征

分布律描述随机变量的统计特征是利用随机变量取值与取值概率的对应关系。在许多实际问题中，概率分布函数和概率密度函数需要大量的试验才能得到。幸运的是有时并不需要对随机变量进行完整的描述，而只要求知道随机变量统计规律的主要特征。另一方面，有时虽然掌握了随机变量的概率分布函数和概率密度函数，但需要更直观地了解它的平均值和偏离平均值的程度，因此引出随机变量的数字特征。

数字特征也称为特征数。数字特征有很多，但主要的数字特征是描述随机变量的集中特性、离散特性和随机变量之间的相关性。

1.3.1　一维随机变量和随机变量函数的数字特征

1．随机变量的数学期望

数学期望又称为统计平均或集合平均，有时更简单地称为均值。数学期望用于描述随机变量的集中特性，用 $E[X]$或 m_X 表示。对于离散随机变量 X，其数学期望

$$E[X] = \sum_{i=1}^{\infty} x_i P(X = x_i) = \sum_{i=1}^{\infty} x_i P_i \tag{1.3-1}$$

如果 X 是连续随机变量，则有

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x \tag{1.3-2}$$

数学期望具有明确的物理意义，如果把概率密度看成具有一定密度的曲线，那么数学期望便是曲线的重心。

在上面的定义中，$E[\cdot]$ 是一个线性算子。在下面随机变量的函数变换中，可以利用 $E[\cdot]$ 的线性性质对随机变量进行运算。

描述随机变量集中特性的统计量还有中位数和众数。使下式成立的 M_e 称为随机变量的中位数

$$P(X < M_\mathrm{e}) = P(X > M_\mathrm{e}) \tag{1.3-3}$$

连续随机变量的中位数将随机变量概率密度下的面积一分为二。离散随机变量的中位数不唯一。概率最大（离散随机变量）或概率密度最大（连续随机变量）的点 x_M 称为众数，记为 M_o。在数字图像处理中，灰度直方图描述了一幅图像的灰度分布。灰度直方图的众数反映了图像的基调，因为在图像上众数这一点的灰度最多。

数学期望、中位数和众数的相对关系如图 1.3-1 所示，若概率密度曲线有单峰且关于峰值点对称，则三者重合。

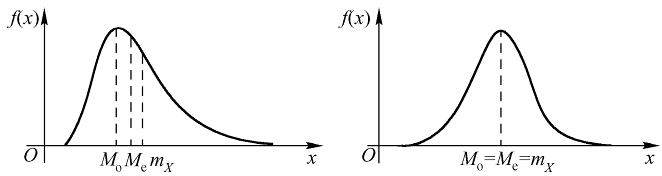


图 1.3-1　表示随机变量集中特性的数字特征

2．随机变量函数的数学期望

利用导出的概率密度变换公式可直接求随机变量函数 Y 的数字特征，而不需要先求出随机变量函数的概率密度 $f_Y(y)$ 。随机变量函数 Y 的数学期望

$$E[Y]=\int_{-\infty}^{\infty} y f_Y(y)\,\mathrm{d}\,y$$

将式(1.2-4)和 $Y=\varphi(X)$ 代入上式

$$E[Y]=\int_{-\infty}^{\infty} y f_X(x)\frac{\mathrm{d}x}{\mathrm{d}y}\mathrm{d}y=\int_{-\infty}^{\infty}\varphi(x) f_X(x)\mathrm{d}x=E[\varphi(X)] \tag{1.3-4}$$

显然，不求 Y 的概率密度 $f_Y(y)$ 就能直接由 X 的概率密度 $f_X(x)$ 和它们的函数关系求得 Y 的数学期望。

3．方差

方差用来度量随机变量偏离其数学期望的程度，或者度量随机变量在数学期望附近的离散程度。因此它描述的是随机变量取值分布的离散特性。方差用 $D[X]$或 σ_X^2 表示。对于离散和连续随机变量，分别有

$$D[X]=E\{(X-E[X])^2\}=\sum_{i=1}^{\infty}(x_i-E[X])^2 P_i \tag{1.3-5}$$

$$D[X]=E\{(X-E[X])^2\}=\int_{-\infty}^{\infty}(x-E[X])^2 f_X(x)\mathrm{d}x \tag{1.3-6}$$

方差开方后称为均方差或标准差

$$\sigma_X=\sqrt{D[X]} \tag{1.3-7}$$

在误差分析中，常用均方差表示误差范围。例如，某一测速雷达的测速精度为 0.1m/s，即均方差 $\sigma=0.1$m/s，是指在多次测量过程中，测速的数值大部分(一般取决于测速误差的概率分布)落在真值的 ± 0.1m/s 范围内。当然也有用 3 倍 σ 表示误差或测量精度的，这时，测量的数值绝大部分落在真值的 ± 3 σ 范围内。

在图像处理中，灰度直方图的方差大致反映了图像的反差。方差较大的图像，层次感较强，而方差较小的图像，图中的景象或物体的轮廓显得不清。

数学期望和方差是随机变量分布的两个重要的特征，图 1.3-2 示出了具有不同数学期望和方差的概率密度。因为概率密度曲线下的面积恒为 1，对于相同分布的随机变量，若数学期望不同但方差相同，表现为概率密度曲线在横轴上平移；若方差不同但数学期望相同，则表现为概率密度曲线在数学期望附近集中的程度，图中 $\sigma_1<\sigma_2<\sigma_3$ 。

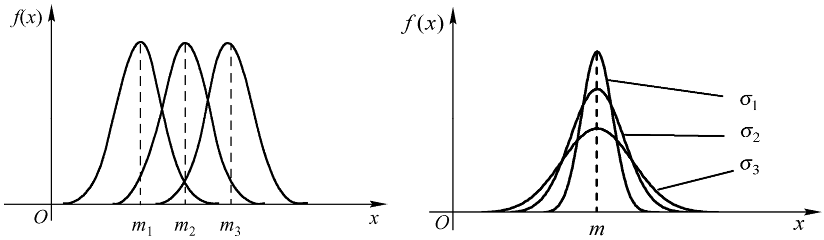


图 1.3−2　具有不同数学期望和方差的随机变量概率密度

仿照随机变量函数的数学期望可得随机变量函数 Y 的方差为

$$D[Y]=\int_{-\infty}^{\infty}[\varphi(x)-m_Y]^2 f_X(x)\mathrm{d}x=D[\varphi(X)] \tag{1.3-8}$$

【例 1.3−1】 已知高斯随机变量 X 的概率密度为

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}}$$

求它的数学期望和方差。

解：根据数学期望和方差的定义，有

$$E[X]=\int_{-\infty}^{\infty}xf_X(x)\mathrm{d}x=\int_{-\infty}^{\infty}\frac{x}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}}\mathrm{d}x$$

令 $t=\dfrac{x-m}{\sigma}$，$\mathrm{d}x=\sigma\mathrm{d}t$，代入上式并整理得

$$E[X]=\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty}t\,\mathrm{e}^{-t^2/2}\mathrm{d}t+\frac{m}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\mathrm{e}^{-t^2/2}\,\mathrm{d}t=0+\frac{m}{\sqrt{2\pi}}\sqrt{2\pi}=m$$

$$D[X]=\int_{-\infty}^{\infty}(x-m)^2 f_X(x)\mathrm{d}x=\int_{-\infty}^{\infty}\frac{(x-m)^2}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}}\mathrm{d}x$$

与前面做同样的变换，即令 $t=\dfrac{x-m}{\sigma}$，整理后得

$$D[X]=\frac{2\sigma^2}{\sqrt{2\pi}}\int_{0}^{\infty}t^2\,\mathrm{e}^{-t^2/2}\mathrm{d}t$$

查数学手册中的积分表 $\displaystyle\int_0^{\infty}x^{2n}\,\mathrm{e}^{-ax^2}\,\mathrm{d}x=\frac{1\times3\times\cdots\times(2n-1)}{2^{n+1}a^n}\sqrt{\frac{\pi}{a}}$

在上式中，令 $n=1$ 及 $a=1/2$，利用积分结果，可得方差

$$D[X]=\frac{2\sigma^2}{\sqrt{2\pi}}\frac{\sqrt{2\pi}}{2}=\sigma^2$$

可见，高斯变量概率密度中的两个量 m 和 σ^2 分别是数学期望和方差，或者说一维高斯变量的概率密度由其数学期望和方差唯一决定。

3．矩函数

矩函数是一种数学定义，根据阶数大小有一阶矩、二阶矩，高于三阶的矩函数称为高阶矩。根据矩函数的计算方式还可以分为原点矩和中心距。下面将会看到一阶原点矩正是曲线

的几何重心。如果曲线是概率密度，那么一阶原点矩就是随机变量的数学期望。

随机变量 X 的 n 阶原点矩定义为

$$m_n = E[X^n], \qquad n = 1, 2, \cdots \tag{1.3-9}$$

根据(1.3-4)式，对于离散和连续随机变量，则分别有

$$m_n = \sum_{i=1}^{\infty} x_i^n P_i, \qquad n = 1, 2, \cdots \tag{1.3-10}$$

$$m_n = \int_{-\infty}^{\infty} x^n f_X(x)\,\mathrm{d}x, \qquad n = 1, 2, \cdots \tag{1.3-11}$$

随机变量 X 的 n 阶中心矩定义为

$$\mu_n = E\{(X - E[X])^n\}, \qquad n = 1, 2, \cdots \tag{1.3-12}$$

类似原点矩的定义式，也可分别写出离散随机变量和连续随机变量中心矩的具体表达式

$$\mu_n = \sum_{i=1}^{\infty} (x_i - E[X])^n P_i, \qquad n = 1, 2, \cdots \tag{1.3-13}$$

$$\mu_n = \int_{-\infty}^{\infty} (x - E[X])^n f_X(x)\,\mathrm{d}x, \qquad n = 1, 2, \cdots \tag{1.3-14}$$

不同阶的矩函数有着不同的意义：

当 $n = 1$ 时，一阶原点矩 m_1 就是数学期望。

当 $n = 2$ 时，二阶中心矩 μ_2 就是方差。

当 $n = 3$ 时，$s = \mu_3 / \sigma^3$ 定义为偏态系数，偏态系数描述概率密度的非对称性，这是因为当概率密度 $f(x)$对称时，奇数阶中心矩为零。在实际应用时，经常用到随机变量三阶中心矩为零的性质。

在图像处理中，灰度直方图的偏态系数是对图像灰度分布偏离对称程度的一种度量。当灰度直方图 $s<0$ 时，图像呈高色调；而当灰度直方图 $s>0$ 时，图像呈低色调。图 1.3−3 示出了具有不同偏态系数的概率密度。

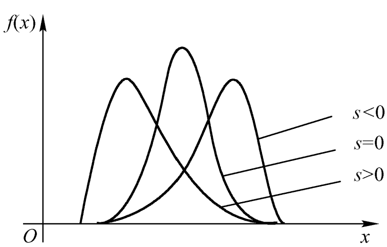


图 1.3−3　不同偏态系数的概率密度

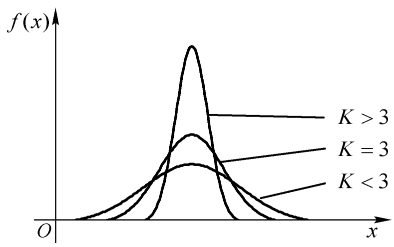


图 1.3−4　不同峰态系数的概率密度

当 $n = 4$ 时，将 $K = \mu_4 / \sigma^4$ 定义为峰态系数，峰态系数描述概率密度的尖锐或平坦程度。高斯概率密度的峰态系数为 3，如图 1.3−4 所示。比较方差相同、具有不同分布的随机变量概率密度，当概率密度的主峰比高斯分布尖锐时，其峰态系数大于 3，反之当概率密度的主峰比高斯分布平坦时，峰态系数小于 3。

图像灰度直方图的峰态系数反映了图像灰度值的分布是聚集在数学期望附近还是分布的比较宽。图像灰度直方图呈现窄峰时，图像的反差小。而当灰度直方图峰态系数较小，灰度分布较宽时，图像具有较多的层次。

1.3.2　二维随机变量的联合矩及统计关系

1．二维随机变量的联合矩

二维随机变量的矩函数包括联合原点矩和联合中心矩。二维随机变量(X, Y)的 $n+k$ 阶联合原点矩定义为

$$m_{nk}=E[X^nY^k]\,,\qquad n=1,2,\cdots,\quad k=1,2,\cdots \tag{1.3-15}$$

仿照一维函数均值的求法，有

$$m_{nk}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x^ny^kf_{XY}(x,y)\,\mathrm{d}x\,\mathrm{d}y \tag{1.3-16}$$

$n+k$ 阶联合中心矩定义为

$$\begin{aligned}\mu_{nk}&=E\{(X-E[X])^n(Y-E[Y])^k\}\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-E[X])^n(y-E[Y])^kf_{XY}(x,y)\mathrm{d}x\mathrm{d}y\end{aligned} \tag{1.3-17}$$

这里只给出了连续随机变量的表达式，参考式(1.3−10)和式(1.3−13)也可得到离散随机变量的矩函数表达式。

当 $n=1$，$k=0$ 和 $n=0$，$k=1$ 时，一阶原点矩分别是 X 和 Y 的数学期望

$$m_{10}=E[X]=m_X \tag{1.3-18}$$

$$m_{01}=E[Y]=m_Y \tag{1.3-19}$$

当 $n=2$，$k=0$ 和 $n=0$，$k=2$ 时，二阶中心矩分别是 X 和 Y 的方差

$$\mu_{20}=E\{(X-E[X])^2\}=\sigma_X^2 \tag{1.3-20}$$

$$\mu_{02}=E\{(Y-E[Y])^2\}=\sigma_Y^2 \tag{1.3-21}$$

当 $n=1$，$k=1$ 时，二阶联合原点矩和二阶联合中心矩分别是 X 和 Y 的相关矩和协方差

$$m_{11}=E[XY]=R_{XY} \tag{1.3-22}$$

$$\mu_{11}=E\{(X-E[X])(Y-E[Y])\}=C_{XY} \tag{1.3-23}$$

这两个统计量反映了两个随机变量之间的关联程度，此外协方差还反映了两个随机变量各自的离散程度。

2．二维随机变量的统计独立、不相关、正交

为了去除两个随机变量离散程度对相关程度的影响，可将协方差对两个随机变量的均方差进行归一化

$$r_{XY}=\frac{C_{XY}}{\sigma_X\sigma_Y},\qquad -1\leqslant r_{XY}\leqslant 1 \tag{1.3-24}$$

归一化协方差也称相关系数，它只反映两个随机变量之间的关联程度，与它们的数学期望和方差无关。当 $r_{XY}=0$ 时，称随机变量 X 和 Y 不相关，否则称为相关；若$|r_{XY}|=1$则称为完全相关。

从原点矩和中心矩的定义看，如果数学期望为零，则原点矩和中心矩相同。当数学期望不为零时，可以根据定义导出原点矩和中心矩之间的关系

$$C_{XY}=E\{(X-E[X])(Y-E[Y])\}$$
$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{xy-yE[X]-xE[Y]+E[X]E[Y]\}f_{XY}(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

整理被积函数第二项，得到

$$E[X]\int_{-\infty}^{\infty}y\left[\int_{-\infty}^{\infty}f_{XY}(x,y)\,\mathrm{d}x\right]\mathrm{d}y=E[X]\int_{-\infty}^{\infty}yf_Y(y)\,\mathrm{d}y=E[X]E[Y]$$

同理可以简化第三项和第四项，得到

$$C_{XY}=E\{(X-E[X])(Y-E[Y])\}=R_{XY}-E[X]E[Y] \tag{1.3-25}$$

当 $X=Y$ 时，协方差退化为方差

$$D[X]=C_{XX}=E\{(X-E[X])^2\}=E[X^2]-(E[X])^2 \tag{1.3-26}$$

【例 1.3-2】 已知 X 与 Y 为互相独立的随机变量，求 X 与 Y 的相关系数。

解： 由于 X 与 Y 互相独立，根据式(1.1-36)

$$f_{XY}(x,y)=f_X(x)f_Y(y)$$

$$C_{XY}=E\{(X-E[X])(Y-E[Y])\}$$
$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-m_X)(y-m_Y)f_{XY}(x,y)\mathrm{d}x\mathrm{d}y$$

由于 X 与 Y 互相独立，可以将二重积分化为两个一重积分

$$C_{XY}=\int_{-\infty}^{\infty}(x-m_X)f_X(x)\mathrm{d}x\int_{-\infty}^{\infty}(y-m_Y)f_Y(y)\mathrm{d}y=0$$

所以，$r_{XY}=0$。

这个例子说明了两个互相独立的随机变量一定是不相关的。统计独立是由概率论中的事件独立推广而来的，对于二维随机变量而言，体现在概率密度满足式(1.1-36)。如果把二维随机变量看成平面上的一个随机点，那么这个随机点的两个坐标就表明了随机点在二维平面上所处的位置。统计独立是指随机点的两个坐标之间是完全随机的，没有任何关系。相关则指随机点两个坐标之间的线性相关程度。如果二维随机变量是完全相关的，那么随机点在平面上的分布是一条直线，每个随机点的两个坐标严格遵循线性方程。如果二维随机变量的相关系数介于 0 和 1(或–1 和 0)之间，则它们可能用一个除直线方程之外的其他方式联系起来。

统计独立与不相关的概念是不同的，相比之下统计独立的条件更严格一些。下面讨论它们满足的条件以及相互之间的关系。

（1）随机变量 X 与 Y 统计独立的充要条件是 $f_{XY}(x,y)=f_X(x)f_Y(y)$。

（2）随机变量 X 与 Y 不相关的充要条件是 $r_{XY}=0$。

由式(1.3–25)，$C_{XY}=R_{XY}-E[X]E[Y]$，若随机变量 X 与 Y 不相关则 $C_{XY}=0$，此时有

$$R_{XY}=E[X]E[Y] \tag{1.3-27}$$

（3）随机变量 X 与 Y 统计独立，它们必然是不相关的。例 1.3–2 已经说明了这个结论。

（4）随机变量 X 与 Y 不相关，它们不一定互相独立。例 1.3–3 可以充分说明这个问题。

（5）若随机变量 X 与 Y 的相关矩为零，即

$$R_{XY}=E[XY]=0 \tag{1.3-28}$$

则称 X 和 Y 互相正交。对于互相正交的两个随机变量，如果其中一个随机变量的数学期望也为零，则二者一定不相关，因为 $C_{XY}=R_{XY}-E[X]E[Y]$，若 $E[X]$和 $E[Y]$之一为零，必有 $C_{XY}=0$。

【例 1.3-3】 已知二维随机变量(X, Y)满足$\begin{cases}X=\cos\varPhi\\Y=\sin\varPhi\end{cases}$，式中，$\varPhi$ 是在$[0, 2\pi]$上均匀分布的随机变量，讨论 X, Y 的独立性和相关性。

解：根据已知条件，$X^2+Y^2=1$，显然它们的取值互相依赖于对方，或者说是通过参变量$\varPhi$互相联系的，因此不可能是互相独立的。另一方面，它们却是不相关的，因为 X 与 Y 的数学期望为

$$E[X]=\int_{-\infty}^{\infty}\cos\varphi f_{\varPhi}(\varphi)\,\mathrm{d}\varphi=\int_{0}^{2\pi}\frac{1}{2\pi}\cos\varphi\,\mathrm{d}\varphi=0$$

$$E[Y]=\int_{-\infty}^{\infty}\sin\varphi f_{\varPhi}(\varphi)\,\mathrm{d}\varphi=\int_{0}^{2\pi}\frac{1}{2\pi}\sin\varphi\,\mathrm{d}\varphi=0$$

X 与 Y 之间的自相关函数和协方差为

$$R_{XY}=E[XY]=E[\sin\varPhi\cos\varPhi]=\frac{1}{2}E[\sin 2\varPhi]=0$$

$$C_{XY}=E[(X-m_X)(Y-m_Y)]=E[XY]=0$$

所以，$r_{XY}=0$，说明 X 与 Y 之间不相关。

不用先求出二维随机变量函数的概率密度，直接用原随机变量的联合概率密度和函数关系即可求一些矩函数。这种直接求随机变量函数数字特征的方法，大大地简化了运算过程。

下面给出一些经常用到且很容易证明的运算法则。

设 X_1 和 X_2 为任意分布的两个随机变量，a 和 b 为常数。

$$E[a]=a,\quad E[aX+b]=aE[X]+b,\quad E[X_1\pm X_2]=E[X_1]\pm E[X_2]$$

$$D[a]=0,\quad D[aX+b]=a^2D[X],\quad D[X_1\pm X_2]=D[X_1]+D[X_2]\pm 2C_{X_1X_2}$$

当 X_1 和 X_2 不相关时 $\quad E[X_1X_2]=E[X_1]E[X_2],\ D[X_1\pm X_2]=D[X_1]+D[X_2]$

【例 1.3-4】 已知高斯随机变量 X 的数学期望和方差分别为 m 和σ^2，求随机变量 $Y=5X+1$ 的 $f_Y(y)$, m_Y, σ_Y^2, R_{XY}, C_{XY}, r_{XY}。

解：本例可以仿照例 1.2-1 通过函数变换求 Y 的概率密度，也可以先根据上面给出的方法求出数学期望和方差，再根据高斯分布的特点写出 Y 的概率密度。

根据 X 的数学期望 m 和方差σ^2，可以得到 Y 的数学期望和方差

$$m_Y=E[Y]=E[5X+1]=5E[X]+1=5m+1$$

$$\sigma_Y^2=D[Y]=D[5X+1]=5^2D[X]=25\sigma^2$$

先写出 Y 的概率密度，再将以上两式代入得

$$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_Y}\exp\left[-\frac{(y-m_Y)^2}{2\sigma_Y^2}\right]=\frac{1}{5\sqrt{2\pi}\sigma}\exp\left[-\frac{(y-5m-1)^2}{50\sigma^2}\right]$$

X 和 Y 的相关矩 $\quad R_{XY}=E[XY]=E[X(5X+1)]=5E[X^2]+E[X]$

由式(1.3-26)，可以由数学期望和方差求二阶原点矩