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。

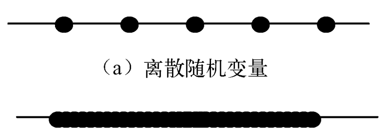
随机信号理论的基础是“概率论”和“信号与系统”，这里假定读者已经掌握了这些知识。本章首先对随机变量的要点做一下系统的回顾；然后介绍用特征函数描述随机变量的方法。本章的后半部分将给出通信与信息处理领域中经常用到的一些随机变量的分布，并重点讨论高斯随机变量。本章还将给出一些随机变量仿真的方法和程序，供读者参考和选用。

1.1　随机变量及其分布

设随机试验的样本空间为 S，如果对样本空间的每一个元素 $e_i \in S$，都有一个实数 $x_i = X(e_i)$ 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。



（a）离散随机变量

（b）连续随机变量

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

概率累积分布函数（以下简称“分布函数”或“累积分布”）是描述随机变量概率特征的最重要函数。分布函数完整地刻画了随机变量的统计规律，并且决定了随机变量的一切概率

特征。分布函数对于任意类型的随机变量均存在。特别地，对于离散型随机变量，分布函数与其分布列具有等价关系。对于连续性随机变量，分布函数与其概率密度函数具有等价关系。

1．一维随机变量的概率分布函数

设 X 是样本空间 S 上的随机变量，x 为任意实数，称函数

$$F(x)=P(X\leqslant x) \tag{1.1-1}$$

为随机变量 X 的概率累积分布函数。注意定义（1.1-1）中 $\{X\leqslant x\}$ 是一集合（事件），而概率总是定义在集合（事件）上的。根据定义，累积分布函数在 x 点的取值表征随机变量 X 取值不超过 x 的概率。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，显然它对任何随机变量都存在。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

$$F(x_2)\geqslant F(x_1) \tag{1.1-2}$$

性质 2　$F(x)$非负，取值满足

$$0\leqslant F(x)\leqslant 1 \tag{1.1-3}$$

并且 $F(-\infty)=0$，$F(+\infty)=1$。

性质 3　随机变量 X 在 x_1,x_2 区间内的概率为

$$P(x_1<X\leqslant x_2)=F(x_2)-F(x_1) \tag{1.1-4}$$

性质 4　$F(x)$右连续，即

$$F(x^+)=F(x) \tag{1.1-5}$$

对于任意一个函数，判断其是否可以作为某随机变量的分布函数，只需要使用性质 1、性质 2 和性质 4 判断即可。

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

解： 根据定义，随机变量 X 的概率分布函数为

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

或写成 $F(x)=0.5u(x)+0.5u(x-1)$，其中 $u(x)$为单位阶跃函数。

从上述例题中可以看出，0-1 分布的随机变量的分布函数在 $x=0$ 和 $x=1$ 两处出现了幅度为 0.5 的跳跃。这个结论具有一般性。特别地，对于所有的离散随机变量，其分布函数均是分段连续的阶梯状函数，阶跃的高度等于随机变量在该点的概率，并且在阶跃处右连续（性质 4）。其分布函数可以写为

$$F(x)=\sum_{i=1}^{\infty}P(X=x_i)u(x-x_i)=\sum_{i=1}^{\infty}P_i u(x-x_i) \tag{1.1-6}$$

式中，$u(x)$为单位阶跃函数，P_i 为$\{X=x_i\}$的概率。

2．一维离散随机变量的概率分布列

如果随机变量 X 只可能取有限多个或者可列多个取值，则称随机变量为离散型随机变

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

$$P(X=x_k)=P_k \qquad k=1,2,\cdots \tag{1.1-7}$$

为离散型随机变量 X 的概率分布列或分布律。

离散随机变量的分布函数 $F(x)$ 与分布列 $\{P_k,k=1,2,\cdots\}$ 具有等价关系。当已知分布函数时，有

$$P_1=P(X=x_1)=P(X\leqslant x_1)=F(x_1)$$

$$P_k=P(X\leqslant x_k)-P(X<x_k)=F(x_k)-F(x_{k-1}),\quad k=2,3,\cdots$$

而当分布列已知时，可以通过式(1.1−6)来获得分布函数。

由分布函数的性质容易得知离散随机变量的分布列满足下面两个性质：

性质 1：$P_k\geqslant 0$， $k=1,2,\cdots$

性质 2：$\sum\limits_{k=1}^{\infty}P_k=1$

离散随机变量的分布列也可用表格的形式来表示（见表 1.1−1）。

表 1.1−1

X	x_1　x_2　$\cdots$　x_n　$\cdots$
P	P_1　P_2　$\cdots$　P_n　$\cdots$

3．一维随机变量的概率密度

对于分布函数为 $F(x)$ 的一维实随机变量 X，如果存在非负函数 $f(x)$，使得对于任意实数 x 均有

$$F(x)=\int_{-\infty}^{x}f(\lambda)\mathrm{d}\lambda \tag{1.1-8}$$

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

从连续型随机变量的定义可知，其概率密度函数可以通过累积分布函数的微分求得

$$f(x)=\frac{\mathrm{d}F(x)}{\mathrm{d}x} \tag{1.1-9}$$

因此，连续型随机变量的累积分布函数与概率密度函数相互等价。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，给出随机变量在该区间的取值概率

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

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充要条件。连续随机变量的分布函数是连续的，因此连续随机变量 X 在任意指定取值的概率为零。性质 3 则刻画了随机变量取值落在一个区间的概率 $P\left(x_1<X\leqslant x_2\right)$。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。

【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 $K=1/a$ 满足概率密度函数的条件。综上，$f(x)=\dfrac{1}{a}[u(x)-u(x-a)]$。

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

连续型随机变量和离散型随机变量并不能囊括所有的随机变量。例如分布函数

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

中既存在间断点（$x=-1$），也存在连续部分（$0\leqslant x<1$）。实际上，这个分布函数所对应的是混合型的随机变量。我们可通过引入 δ 函数的方法，将概率密度函数的定义扩展到分布函数存在有限个间断点的随机变量上。例如，对于离散型随机变量，其概率密度函数可以写为：

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

式中，$\delta(x)$ 为单位冲激函数。

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

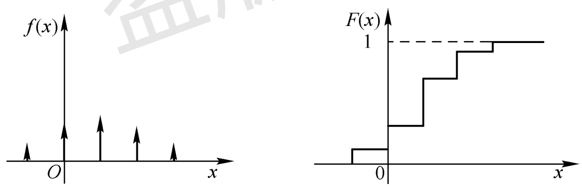


图 1.1-3　离散随机变量的概率密度和概率分布函数

1.1.2　多维随机变量及其分布律

假设 X 和 Y 为同一样本空间 S 上的两个一维随机变量，则称向量 (X, Y) 为 S 上的一个二维随机变量。二维随机变量 (X, Y) 可看成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。

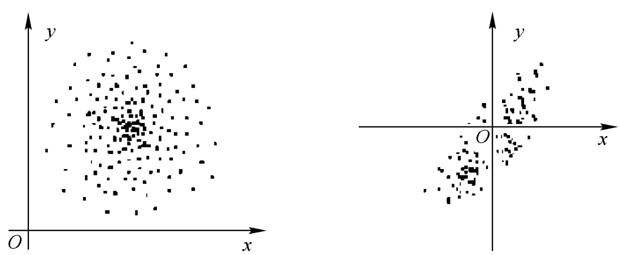


图 1.1-4　二维随机变量——平面上的随机点

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

$$F_{XY}(x,y)=P(X\leqslant x,\ Y\leqslant y) \tag{1.1-14}$$

称为(X, Y)的分布函数或 X 和 Y 的联合概率分布函数。

与一维随机变量的情况相似，联合概率分布函数对于任意类型的二维随机变量都存在，且完整地刻画了这个二维随机变量的统计规律，决定了其一切概率特征。若将(X, Y)看成平面上随机点的坐标，则联合概率分布函数 $F_{XY}(x, y)$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

$$F_{XY}(x_2,y)\geqslant F_{XY}(x_1,y) \tag{1.1-15}$$

对固定的 x，当 $y_2 > y_1$ 时，有

$$F_{XY}(x,y_2)\geqslant F_{XY}(x,y_1) \tag{1.1-16}$$

性质 2　$F_{XY}(x,y)$ 是 x, y 的有界函数。

$$0\leqslant F_{XY}(x,y)\leqslant 1 \tag{1.1-17}$$

且有

$$F_{XY}(-\infty,y)=0\text{，}\ F_{XY}(x,-\infty)=0\text{，}\ F_{XY}(-\infty,-\infty)=0 \tag{1.1-18}$$

$$F_{XY}(+\infty,+\infty)=1 \tag{1.1-19}$$

性质 3　$F_{XY}(x,y)$ 关于 x 或 y 均为右连续函数。

$$F_{XY}(x^+,y)=F(x,y)\text{，}\quad F_{XY}(x,y^+)=F(x,y) \tag{1.1-20}$$

性质 4　随机变量 X 的边缘分布函数为

$$F_X(x)=F_{XY}(x,+\infty) \tag{1.1-21}$$

随机变量 Y 的边缘分布函数为

$$F_Y(y)=F_{XY}(+\infty,y) \tag{1.1-22}$$

性质 5　对任意 (x_1,y_1) 和 (x_2,y_2)，其中 $x_2 > x_1$，$y_2 > y_1$，则有

$$P(x_1<X\leqslant x_2,y_1<Y\leqslant y_2)=F_{XY}(x_2,y_2)-F_{XY}(x_2,y_1)-F_{XY}(x_1,y_2)+F_{XY}(x_1,y_1) \tag{1.1-23}$$

上式的结果可以通过图 1.1−6 得出。

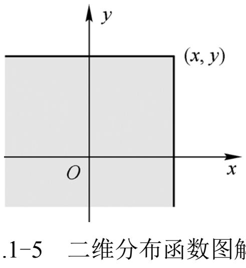


图 1.1−5　二维分布函数图解

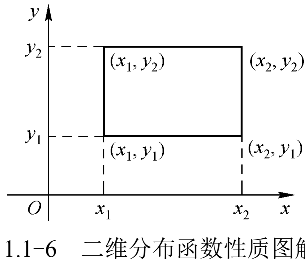


图 1.1−6　二维分布函数性质图解

2．二维离散随机变量的概率分布列

如果二维随机变量(X, Y)的可能取值只有有限对或可列对，则称(X, Y)为二维离散型随机

变量。(X, Y)的联合概率分布列为

$$P(X=x_i, Y=y_j)=P_{ij}, \quad i,j=1,2,\cdots \tag{1.1-24}$$

记

$$P_{i\bullet}=\sum_{j=1}^{\infty}P_{ij}=P(X=x_i), \quad i=1,2,\cdots \tag{1.1-25}$$

$$P_{\bullet j}=\sum_{i=1}^{\infty}P_{ij}=P(Y=y_j), \quad j=1,2,\cdots \tag{1.1-26}$$

分别称 $\{P_{i\bullet}, i=1,2,\cdots\}$ 和 $\{P_{\bullet j}, j=1,2,\cdots\}$ 为(X, Y)关于 X 和 Y 的边缘分布列。

二维离散型随机变量的联合概率分布函数 $F(x,y)$ 与其联合概率分布列 $\{P_{i,j}, i=1,2,\cdots, j=1,2,\cdots\}$ 同样具有等价关系。由分布函数的性质容易得知二维离散随机变量的分布列满足下面两个性质：

性质 1：$P_{ij}\geqslant 0$，$i=1,2,\cdots$，$j=1,2,\cdots$

性质 2：$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_{ij}=1$

二维离散型随机变量的联合概率分布列可以使用表格表示（见表 1.1-2）。

表 1.1-2

Z \ Y	x_1	x_2	$\cdots$	x_i	$\cdots$
y_1	P_{11}	P_{21}	$\cdots$	P_{i1}	$\cdots$
y_2	P_{12}	P_{22}	$\cdots$	P_{i2}	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
y_j	P_{1j}	P_{2j}	$\cdots$	P_{ij}	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

3．二维随机变量的概率密度

对于二维随机变量 (X,Y) 的分布函数 $F_{XY}(x, y)$，如果存在非负可积函数 $f(x,y)$，使得对任意 x,y 均有

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

则称 (X,Y) 为连续型二维随机变量，函数 $f(x,y)$ 称为二维随机变量 (X,Y) 的概率密度函数，或称为 X 和 Y 的联合概率密度函数。记

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

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

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

从上述定义可知，联合概率密度函数可以通过联合概率分布函数的二阶偏导求得，即

$$f_{XY}(x,y)=\frac{\partial^2 F_{XY}(x,y)}{\partial x\partial y} \tag{1.1-30}$$

因此，连续型二维随机变量的累积分布函数与概率密度函数相互等价。从分布函数的性质和二维概率密度的定义可以得到如下性质：

性质 1　二维概率密度函数非负，即

$$f_{XY}(x,y)\geqslant 0 \tag{1.1-31}$$

性质 2　二维概率密度函数在整个平面上的积分为 1，即

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}f_{XY}(x,y)\mathrm{d}x\mathrm{d}y=F_{XY}(+\infty,+\infty)=1 \tag{1.1-32}$$

性质 3　设 R 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

$$P((X,Y)\in R)=\iint_R f_{XY}(x,y)\mathrm{d}x\mathrm{d}y \tag{1.1-33}$$

上式中的积分是 R 区域的二重面积分。

4．条件概率分布函数和条件概率密度

在集合论中，条件概率是指集合（事件）A 在集合（事件）B 发生的条件下的发生概率，表示为

$$P(A|B)=\frac{P(A\cap B)}{P(B)} \tag{1.1-34}$$

在二维分布中，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 $F_Y(y|x)$、条件概率密度 $f_Y(y|x)$或条件分布列 $P_Y(Y|X)$。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。

对于二维离散随机变量，其条件分布列可以写为

$$P_Y(Y=y_j|X=x_i)=\frac{P_Y(Y=y_j,X=x_i)}{P(X=x_i)}=\frac{P_{ij}}{P_{i\bullet}}\quad i=1,2,\cdots;\ j=1,2,\cdots \tag{1.1-35}$$

其中 $P_{i\bullet}=\sum\limits_{j=1}^{\infty}P_{ij}$。

对于二维连续型随机变量，其条件概率分布函数为

$$\begin{aligned}F_Y(y|x)&=\lim_{\Delta x\to 0}\frac{P\left(x<X\leqslant x+\Delta x,Y\leqslant y\right)}{P\left(x<X\leqslant x+\Delta x\right)}=\lim_{\Delta x\to 0}\frac{\int_{-\infty}^{y}\int_{x}^{x+\Delta x}f_{XY}(u,v)\mathrm{d}u\mathrm{d}v}{\int_{x}^{x+\Delta x}f_X(u)\mathrm{d}u}\\&=\lim_{\Delta x\to 0}\frac{\int_{-\infty}^{y}f_{XY}(x,v)\Delta x\mathrm{d}v}{f_X(x)\Delta x}=\int_{-\infty}^{y}\frac{f_{XY}(x,v)}{f_X(x)}\mathrm{d}v\end{aligned}$$

利用概率密度函数为概率分布函数导数的结果，可以得到

$$f_Y(y|x)=\frac{\partial}{\partial y}F_Y\left(y|x\right)=\frac{f_{XY}(x,y)}{f_X(x)} \tag{1.1-36}$$

条件概率还满足全概率公式、贝叶斯公式和链式法则。以连续型随机变量的概率密度函数为例，将这三个重要的公式列出如下。

全概率公式：

$$f(x)=\int_{-\infty}^{+\infty}f(x|y)f(y)\mathrm{d}y \tag{1.1-37}$$

贝叶斯公式：

$$f(x|y)=\frac{f(y|x)f(x)}{\int_{-\infty}^{+\infty}f(y|x)f(x)\mathrm{d}x} \tag{1.1-38}$$

链式公式：

$$f(x_1,x_2,\cdots,x_n)=f(x_1)f(x_2|x_1)\cdots f(x_n|x_1,x_2,\cdots,x_{n-1}) \tag{1.1-39}$$

5．多维随机变量的概率分布和独立随机变量

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（1.1-14）的方式定义 n 维分布函数

$$F_X(x_1,x_2,\cdots,x_n)=P(X_1\leqslant x_1,X_2\leqslant x_2,\cdots,X_n\leqslant x_n) \tag{1.1-40}$$

对于连续型随机变量，可以仿照式（1.1-30）得到其概率密度函数

$$f_X(x_1,x_2,\cdots,x_n)=\frac{\partial^n F_X(x_1,x_2,\cdots,x_n)}{\partial x_1\partial x_2\cdots\partial x_n} \tag{1.1-41}$$

类似于式(1.1−28)或式(1.1−29)，对 n 维联合概率密度函数中的某些随机变量进行积分，可以获得其余随机变量的联合概率密度函数

$$f_X(x_1,x_2,\cdots,x_m)=\underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{n-m} f_X(x_1,x_2,\cdots,x_m,\cdots,x_n)\mathrm{d}x_{m+1}\cdots\mathrm{d}x_n \tag{1.1-42}$$

式(1.1−28)和式(1.1−29)是 $n=2$，$m=1$ 时的情况。

高维随机变量中的一个重要概念是统计独立。对于 n 维随机变量$(X_1,X_2,\cdots,X_n)$，如果对所有的 $x_1,x_2,\cdots,x_n$ 均有

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

则称 $X_1,X_2,\cdots,X_n$ 是相互独立的随机变量。对于连续型随机变量，上述定义可以等价地写为：对于所有的 $x_1,x_2,\cdots,x_n$，满足

$$f_X(x_1,x_2,\cdots,x_n)=f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n)=\prod_{i=1}^{n}f_{X_i}(x_i) \tag{1.1-43}$$

即，当 n 维随机变量的联合概率密度等于各自边缘概率密度的乘积时，各随机变量相互独立。当 $n=2$ 时有

$$f_{XY}(x,y)=f_X(x)f_Y(y) \tag{1.1-44}$$

等式两边同时除以 X 或 Y 的边缘分布，可以得到

$$f_X(x\,|\,y)=f_X(x) \tag{1.1-45}$$

$$f_Y(y\,|\,x)=f_Y(y) \tag{1.1-46}$$

因此，对于二维连续型随机变量，如果边缘概率密度函数与条件概率密度函数相等，则 X 和 Y 相互独立。

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

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

确定随机变量 Y 分布的基本方法是从定义出发：

$$\begin{aligned} F_Y(y) &= P(Y \leqslant y) = P(\varphi(X) \leqslant y) \\ &= P(X \in \{x : \varphi(x) \leqslant y\}) \\ &= \int_{\{x:\varphi(x)\leqslant y\}} f_X(x)\mathrm{d}x \end{aligned} \tag{1.2-2}$$

设随机变量 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，则随机变量 Y 的概率分布函数与 X 的概率分布函数存在如下关系：

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

式(1.2−3)可以由式(1.2−2)化简得到。具体推导过程如下：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

$$F_Y(y) = P(\varphi(X) \leqslant y) = P(X \leqslant h(y)) = \int_{-\infty}^{h(y)} f_X(x)\mathrm{d}x \tag{1.2-4}$$

将上式对 y 求导，便得到随机变量 Y 的概率密度

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

同理，可得到当 $h(Y)$单调下降时随机变量 Y 的概率密度

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

综合以上两种情况，即可以得到式(1.2−3)。

图 1.2−1 给出了当 $h(Y)$是单调增函数情况下，一维随机变量 X 和 Y 的函数关系示意图。图中对式(1.2−3)给予了具有启发性的解释：假设 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

$$f_X(x)\mathrm{d}x = f_Y(y)\mathrm{d}y \tag{1.2-5}$$

因此
$$f_Y(y) = f_X(x)\frac{\mathrm{d}x}{\mathrm{d}y} = f_X(x)h'(y)$$

考虑到概率密度非负，对 $h'(y)$ 项添加绝对值可得式(1.2−3)。这种方法虽然不是严格的数学推导，但在实际中却非常好用。式(1.2−5)被称为等概率原理。

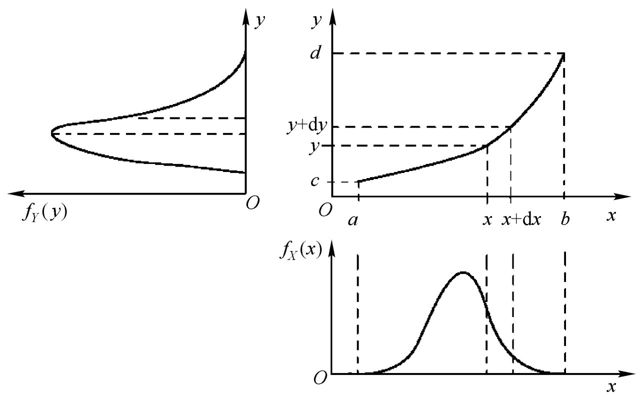


图 1.2−1　一维函数单调变换

【例 1.2−1】 半波整流器的输出 Y 与输入 X 之间的数学模型可以表示为

$$Y=\varphi(X)=\begin{cases}X, & X\geqslant 0\\ 0, & X<0\end{cases}$$

若已知随机变量 X 的概率密度和概率分布函数分别为 $f_X(x)$ 和 $F_X(x)$，试求 Y 的概率密度 $f_Y(y)$ 和概率分布函数 $F_Y(y)$。

解： 由于 $y=\varphi(x)$ 并非严格单调函数，因此通过式(1.2−2)的方法来求解 Y 的分布。具体地，由定义可知 $Y\geqslant 0$，因此当 $y<0$ 时，有

$$F_Y(y)=P(Y\leqslant y)=0$$

当 $y\geqslant 0$ 时，事件 $\{Y\leqslant y\}$ 等价为 $\{X\leqslant y\}$，有

$$F_Y(y)=P(Y\leqslant y)=P(X\leqslant y)=F_X(y)$$

注意到 $F_Y(y)$ 在 $y=0$ 处有跳跃，跳跃幅度为 $F_X(0)$。

综合以上两种情况有

$$F_Y(y)=F_X(y)u(y)$$

$$f_Y(y)=f_X(y)u(y)+F_X(0)\delta(y)$$

【例 1.2−2】 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 $f_X(x)=\dfrac{1}{\sqrt{2\pi}\sigma_X}\mathrm{e}^{-\frac{(x-m_X)^2}{2\sigma_X^2}}$。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

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

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

上例说明，高斯变量 X 经过线性变换后的随机变量 Y 仍然是高斯分布的，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

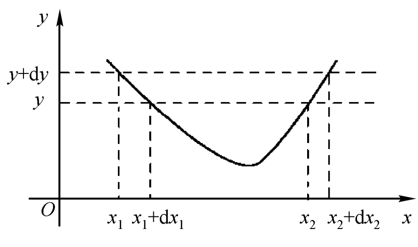


图 1.2−2　一维函数多值变换

$$\begin{cases}X_1=h_1(Y)\\ X_2=h_2(Y)\end{cases}$$

这时随机变量 Y 的取值落在子区间 $(y,y+\mathrm{d}y)$，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。

【例 1.2-3】 平方律检波器的输出 Y 与输入 X 之间的数学模型为

$$Y=\varphi(X)=X^2$$

若已知随机变量 X 的概率密度为 $f_X(x)$，试求 X 的概率密度 $f_Y(y)$。

解：由于 $y=\varphi(x)$ 是分段单调函数，因此通过式(1.2-7)的方法来求解 X 的分布。具体地，令

$$X_1=h_1(Y)=\sqrt{Y} \quad 和 \quad X_2=h_2(Y)=-\sqrt{Y}$$

对函数求导，有 $$h_1'(y)=\frac{1}{2\sqrt{y}} \quad 和 \quad h_2'(y)=-\frac{1}{2\sqrt{y}}$$

因此 $$f_Y(y)=\frac{1}{2\sqrt{y}}\left[f_X\left(\sqrt{y}\right)+f_X\left(-\sqrt{y}\right)\right]$$

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。

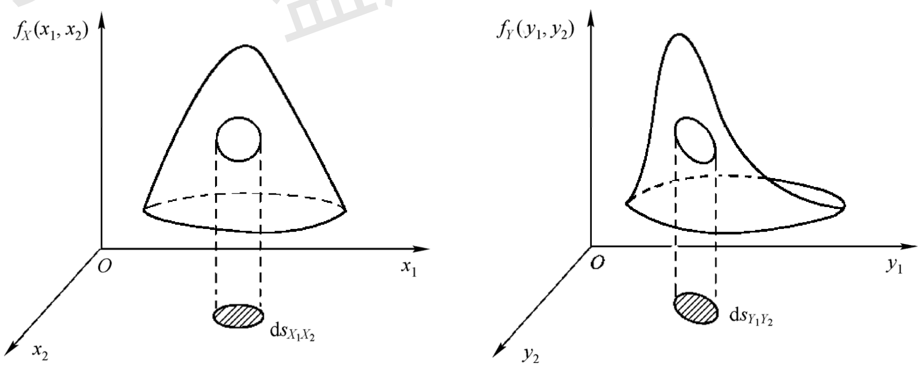


图 1.2-3　二维随机变量函数变换

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

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

注意式(1.2−11)中对雅可比行列式 J 添加了绝对值运算，以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 1.2−4】 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 1.2−5】 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 1.2−6】 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

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

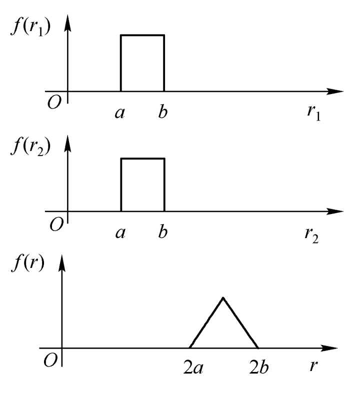


图 1.2−4 例 1.2−4 的图

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，因此引出了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。数学期望刻画了随机变量分布的中心位置，用 $E[X]$或 m_X表示。对于离散随机变量 X，其数学期望为

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

如果 X 是连续随机变量，则有

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

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。

随机变量函数 $Y=g(X)$ 的数学期望可以直接通过随机变量 X 的概率密度求取。可以证明

$$E\left[g(X)\right]=E\left[Y\right]=\int_{-\infty}^{+\infty}yf_Y(y)\mathrm{d}y=\int_{-\infty}^{+\infty}g(x)f_X(x)\mathrm{d}y \tag{1.3-3}$$

上述关系还可以扩展到高维随机变量情况，例如，对 $Y=g(X_1,X_2,\cdots,X_n)$ 求均值，有

$$E\left[Y\right]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}g(x_1,x_2,\cdots,x_n)f(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n \tag{1.3-4}$$

其中 $f(x_1,x_2,\cdots,x_n)$ 是 n 维随机变量 $(X_1,X_2,\cdots,X_n)$ 的概率密度函数。

在上面的定义中，$E[\cdot]$ 是一个线性算子，可以利用 $E[\cdot]$ 的线性性质对复杂的运算进行化简。

中位数同样描述了随机变量分布的中心位置。满足下面关系的常数 M_e 被称为随机变量的中位数

$$P(X\leqslant M_\mathrm{e})\geqslant 1/2 \quad 且 \quad P(X\geqslant M_\mathrm{e})\geqslant 1/2 \tag{1.3-5}$$

上述定义对离散型随机变量和连续型随机变量都成立。对于连续型随机变量，从上述条件可以推导出

$$P(X\geqslant M_\mathrm{e})=P(X\leqslant M_\mathrm{e})=1/2 \tag{1.3-6}$$

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。

描述随机变量分布中心位置的统计量还有众数。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。

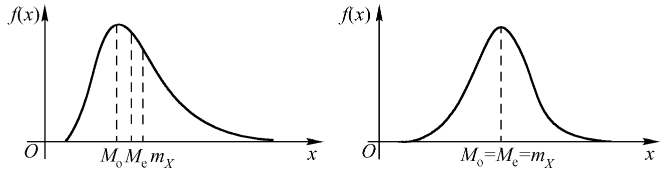


图 1.3−1　数学期望、中位数和众数的相对关系

2．方差

方差用来度量随机变量分布的离散程度，用 $D[X]$或 σ_X^2 表示，其计算公式为

$$D[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2 \tag{1.3-7}$$

对于离散和连续随机变量，分别有

$$D[X] = \sum_{i=1}^{\infty}(x_i - E[X])^2 P_i \tag{1.3-8}$$

$$D[X] = \int_{-\infty}^{\infty}(x - E[X])^2 f_X(x)\mathrm{d}x \tag{1.3-9}$$

方差开方后称为均方差或标准差，即

$$\sigma_X = \sqrt{D[X]} \tag{1.3-10}$$

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随机变量的概率密度。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 $\sigma_1 < \sigma_2 < \sigma_3$ 。

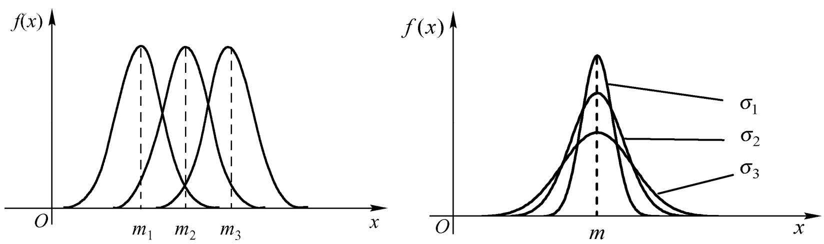


图 1.3−2　具有不同数学期望和方差的随机变量的概率密度

切比雪夫不等式利用均值和方差描述了随机变量分布的位置和聚集程度。设 X 是任一具有有限方差的随机变量，对于任意的 $k > 0$ ，有

$$P(|X - m_X| \geqslant k) \leqslant \sigma_X^2 / k^2 \tag{1.3-11}$$

切比雪夫不等式指出：X 落在区域 $(m_X - k, m_X + k)$ 内的概率要高于 $1 - \sigma_X^2 / k^2$ ，随机变量的

方差越小，随机变量集中在均值附近的概率就越高。特别地，当σ_X=0时，对于任意的$k>0$都有$P(X\in(m_X-k,m_X+k))=1$，因此$P(X=m_X)=1$，此时X几乎确定是一个常数。

通过选取不同k的取值，从切比雪夫不等式中还可以看出，对于任意方差有限的随机变量，其取值落在$m_X\pm\sqrt{2}\sigma_X$范围内的概率不小于50%。其取值落在$m_X\pm 3\sigma_X$范围内的概率不小于 88.89%，取值落在$m_X\pm 5\sigma_X$范围内的概率不小于 96%。因此在误差分析中，用标准差衡量系统的测量精度。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，由此可知，该雷达系统至少有半数的测量值落在真值的±0.14m/s 范围内，而至少 89%的测量值在真值±0.3m/s 范围内。注意切比雪夫不等式只给出了随机变量取值聚集程度的下界，随机变量真正的聚集程度与概率密度函数息息相关。

在图像处理中，灰度直方图的方差大致反映了图像的反差。层次感较强的图像一般对应的方差较大，而轮廓模糊的图像一般对应的方差较小。

对于随机变量函数$Y=\varphi(X)$，利用式（1.3−3）可以得到Y的方差展开式为

$$D[Y]=\int_{-\infty}^{\infty}[\varphi(x)-m_Y]^2 f_X(x)\mathrm{d}x=D[\varphi(X)] \tag{1.3−12}$$

【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

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

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

与前面作同样的变换，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 $$\int_0^{\infty}x^{2n}\mathrm{e}^{-ax^2}\mathrm{d}x=\frac{1\times 3\times\cdots\times(2n-1)}{2^{n+1}a^n}\sqrt{\frac{\pi}{a}}$$

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。

【例 1.3−2】 假设随机变量X满足$E[X^2]<\infty$。求解常数a的最优值，使得函数

$f(a)=E[(X-a)^2]$的取值最小。

解：根据数学期望的线性性质，可得

$$f(a)=a^2-2E[X]a+E[X^2]$$

这是a的一元二次方程，最小值在$a^*=E[X]$处取得，此时$f(a^*)=D[X]$。

由上面例题可知，如果观测者想在试验前对随机变量的实现结果进行猜测，并以猜测误差平方的均值作为代价函数。那么观测者最佳的猜测是选择X的均值，而其对应的最小代价是X的方差。

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，高于二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

$$m_n=E[X^n],\qquad n=1,2,\cdots \tag{1.3-13}$$

根据式(1.3-3)，对于离散和连续随机变量，则分别有

$$m_n=\sum_{i=1}^{\infty}x_i^n P_i,\qquad n=1,2,\cdots \tag{1.3-14}$$

$$m_n=\int_{-\infty}^{\infty}x^n f_X(x)\mathrm{d}x,\qquad n=1,2,\cdots \tag{1.3-15}$$

随机变量X的n阶中心矩定义为

$$\mu_n=E[(X-E[X])^n],\qquad n=1,2,\cdots \tag{1.3-16}$$

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

$$\mu_n=\sum_{i=1}^{\infty}(x_i-E[X])^n P_i,\qquad n=1,2,\cdots \tag{1.3-17}$$

$$\mu_n=\int_{-\infty}^{\infty}(x-E[X])^n f_X(x)\mathrm{d}x,\qquad n=1,2,\cdots \tag{1.3-18}$$

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。

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得比较宽。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。

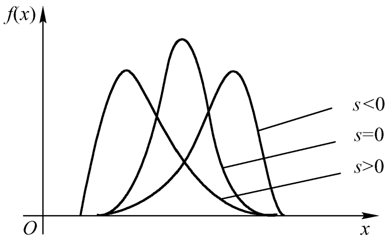


图 1.3−3　不同偏态系数的概率密度

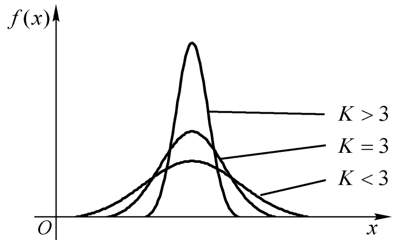


图 1.3−4　不同峰态系数的概率密度

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

$$m_{nk} = E[X^n Y^k], \qquad n = 1,2,\cdots; \ k = 1,2,\cdots \tag{1.3-19}$$

仿照一维函数均值的求法，有

$$m_{nk} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^n y^k f_{XY}(x, y)\mathrm{d}x\mathrm{d}y \tag{1.3-20}$$

$n+k$ 阶联合中心矩定义为

$$\begin{aligned}\mu_{nk} &= E\{(X - E[X])^n (Y - E[Y])^k\} \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - E[X])^n (y - E[Y])^k f_{XY}(x, y)\mathrm{d}x\mathrm{d}y\end{aligned} \tag{1.3-21}$$

这里只给出了连续随机变量的表达式，参考式(1.3−14)和式(1.3−17)也可得到离散随机变量的联合矩函数表达式。

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

$$m_{10} = E[X] = m_X \tag{1.3-22}$$

$$m_{01} = E[Y] = m_Y \tag{1.3-23}$$

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

$$\mu_{20} = E\{(X - E[X])^2\} = \sigma_X^2 \tag{1.3-24}$$

$$\mu_{02} = E\{(Y - E[Y])^2\} = \sigma_Y^2 \tag{1.3-25}$$

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

$$m_{11} = E[XY] = R_{XY} \tag{1.3-26}$$

$$\mu_{11} = E\{(X - E[X])(Y - E[Y])\} = C_{XY} \tag{1.3-27}$$

这两个统计量反映了两个随机变量之间的关联程度，协方差可以看作中心化随机变量(将随机变量的均值置为 0)的相关矩。利用期望的线性性质，容易验证：

$$
\begin{aligned}
C_{XY} &= E\left\{XY - E[X]Y - XE[Y] + E[X]E[Y]\right\} \\
&= E[XY] - E[X]E[Y] = R_{XY} - m_X m_Y
\end{aligned}
\tag{1.3-28}
$$

显然当 $X = Y$ 时，协方差将退化为方差。

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各自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，定义随机变量 X 和 Y 的相关系数为

$$
r_{XY} = \frac{C_{XY}}{\sigma_X \sigma_Y} = \frac{R_{XY} - m_X m_Y}{\sigma_X \sigma_Y}
\tag{1.3-29}
$$

利用柯西施瓦兹不等式，容易证明 $-1 \leqslant r_{XY} \leqslant 1$。相关系数反映两个随机变量之间的线性关联程度，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。

【例 1.3-3】 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(1.1-44)

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

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

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(1.1-44)。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(1.3-28)，$C_{XY} = R_{XY} - E[X]E[Y]$，若随机变量 X 与 Y 不相关则 $C_{XY} = 0$，此时有

$$
R_{XY} = E[X]E[Y]
\tag{1.3-30}
$$

（3）随机变量 X 与 Y 统计独立，它们必然是不相关的。例 1.3-3 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 1.3-4 将说明这个问题。

（5）若随机变量 X 与 Y 的相关矩为零，即

$$R_{XY} = E[XY] = 0 \tag{1.3-31}$$

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 1.3-4】 已知二维随机变量(X, Y)满足 $\begin{cases} X = \cos\Phi \\ Y = \sin\Phi \end{cases}$，式中 Φ 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 Φ 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

$$E[X] = \int_{-\infty}^{\infty} \cos\varphi f_\Phi(\varphi) \mathrm{d}\varphi = \int_0^{2\pi} \frac{1}{2\pi} \cos\varphi \, \mathrm{d}\varphi = 0$$

$$E[Y] = \int_{-\infty}^{\infty} \sin\varphi f_\Phi(\varphi) \, \mathrm{d}\varphi = \int_0^{2\pi} \frac{1}{2\pi} \sin\varphi \, \mathrm{d}\varphi = 0$$

X 与 Y 之间的自相关函数和协方差为

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

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

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。

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

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

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 1.3-5】 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 1.2-2 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

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

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

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

由式(1.3-27)，可以由数学期望和方差求二阶原点矩

$$E[X^2]=D[X]+(E[X])^2$$

则
$$R_{XY}=5D[X]+5(E[X])^2+E[X]=5\sigma^2+5m^2+m$$

X 和 Y 的协方差
$$C_{XY}=R_{XY}-m\,m_Y=5\sigma^2$$

X 和 Y 的相关系数
$$r_{XY}=\frac{C_{XY}}{\sigma\sigma_Y}=\frac{5\sigma^2}{\sigma\sigma_Y}=1$$

【例 1.3-6】 随机变量 $Y=aX+b$，其中 X 为随机变量，a, b 为常数且 $a>0$，求 X 与 Y 的相关系数。

解： Y 的数学期望
$$E[Y]=E[aX+b]=\int_{-\infty}^{\infty}(ax+b)f_X(x)\mathrm{d}x$$
$$=a\int_{-\infty}^{\infty}xf_X(x)\mathrm{d}x+b\int_{-\infty}^{\infty}f_X(x)\mathrm{d}x=aE[X]+b=m_Y$$

Y 的方差
$$D[Y]=\int_{-\infty}^{\infty}(ax+b-m_Y)^2f_X(x)\mathrm{d}x=a^2\int_{-\infty}^{\infty}(x-E[X])^2f_X(x)\mathrm{d}x$$
$$=a^2D[X]=a^2\sigma_X^2$$

X 与 Y 的协方差
$$\begin{aligned}C_{XY}&=E\{(X-E[X])(aX+b-aE[X]-b)\}\\&=E\{(X-E[X])(aX-aE[X])\}\\&=aD[X]\end{aligned}$$

X 与 Y 的相关函数
$$r_{XY}=\frac{C_{XY}}{\sigma_X\sigma_Y}=\frac{aD[X]}{\sqrt{D[X]a^2D[X]}}=1$$

1.4 随机变量的特征函数

随机变量的特征函数不像分布律和数字特征那样具有明显的物理意义，但它的应用价值是不可估量的。一方面，作为一个数学工具，可使很多运算大大简化；另一方面，它又是高阶谱估计的数学基础。

1.4.1 特征函数的定义与性质

特征函数也是一个统计平均量，随机变量 X 的特征函数就是由 X 组成的一个新的随机变量 $\mathrm{e}^{\mathrm{j}\omega X}$ 的数学期望，记为

$$\varPhi(\omega)=E[\mathrm{e}^{\mathrm{j}\omega X}] \tag{1.4-1}$$

离散随机变量和连续随机变量的特征函数分别表示为

$$\varPhi(\omega)=\sum_{i=0}^{\infty}P_i\mathrm{e}^{\mathrm{j}\omega x_i} \tag{1.4-2}$$

$$\varPhi(\omega)=\int_{-\infty}^{\infty}f(x)\mathrm{e}^{\mathrm{j}\omega x}\mathrm{d}x \tag{1.4-3}$$

随机变量 X 的第二特征函数定义为特征函数的对数

$$\varPsi(\omega)=\ln\varPhi(\omega) \tag{1.4-4}$$

下面讨论特征函数的性质。

性质 1
$$\left|\varPhi(\omega)\right|\leqslant\varPhi(0)=1 \tag{1.4-5}$$

由于概率密度非负，且 $\left|\mathrm{e}^{\mathrm{j}\omega X}\right|=1$，所以

$$\left|\int_{-\infty}^{\infty} f(x)\mathrm{e}^{\mathrm{j}\omega x}\mathrm{d}x\right| \leqslant \int_{-\infty}^{\infty} f(x)\mathrm{d}x = \Phi(0) = 1$$

性质 2 若 $Y = aX + b$，a 和 b 为常数，则 Y 的特征函数为

$$\Phi_Y(\omega) = \mathrm{e}^{\mathrm{j}\omega b}\Phi(a\omega) \tag{1.4-6}$$

因为特征函数定义为数学期望，故

$$\Phi_Y(\omega) = E[\mathrm{e}^{\mathrm{j}\omega Y}] = E[\mathrm{e}^{\mathrm{j}\omega(aX+b)}] = \mathrm{e}^{\mathrm{j}\omega b}E[\mathrm{e}^{\mathrm{j}\omega aX}]$$

性质 3 互相独立随机变量之和的特征函数等于各随机变量特征函数之积，即若 $Y = \sum_{n=1}^{N} X_n$，则

$$\Phi_Y(\omega) = E\left[\exp\left(\mathrm{j}\omega\sum_{n=1}^{N} X_n\right)\right] = E\left[\prod_{n=1}^{N}\mathrm{e}^{\mathrm{j}\omega X_n}\right] \tag{1.4-7}$$

如果 $X_n (n = 1,2,\cdots,N)$ 之间互相独立，则

$$\Phi_Y(\omega) = \prod_{n=1}^{N} E[\mathrm{e}^{\mathrm{j}\omega X_n}] = \prod_{n=1}^{N}\Phi_{X_n}(\omega) \tag{1.4-8}$$

【例 1.4-1】 求分布列为 $P(X=1) = p$，$P(X=0) = 1-p$ 的 0-1 分布随机变量 X 的特征函数

解： 根据定义有

$$\Phi_X(\omega) = E[\mathrm{e}^{\mathrm{j}\omega X}] = (1-p)\mathrm{e}^{\mathrm{j}\omega\times 0} + p\mathrm{e}^{\mathrm{j}\omega\times 1} = 1 - p + p\mathrm{e}^{\mathrm{j}\omega}$$

【例 1.4-2】 求二项分布 $X \sim B(n,p)$ 的特征函数。

解： 二项分布可以写为 0-1 分布随机变量的累加和。令 $X = \sum_{i=1}^{n} X_i$，其中 X_i 是独同分布的 0-1 分布随机变量，分布列为 $P(X_i = 1) = p$，$P(X_i = 0) = 1-p$，根据上面例题的结果，有

$$\Phi_{X_i}(\omega) = 1 - p + p\mathrm{e}^{\mathrm{j}\omega}$$

利用性质 3，可知 $\Phi_X(\omega) = \prod_{i=1}^{n}\Phi_{X_i}(\omega) = (1 - p + p\mathrm{e}^{\mathrm{j}\omega})^n$

1.4.2 特征函数与概率密度的关系

根据特征函数的定义，特征函数与概率密度有类似傅里叶变换的关系，即

$$\Phi_X(\omega) = \int_{-\infty}^{\infty} f_X(x)\mathrm{e}^{\mathrm{j}\omega x}\mathrm{d}x \tag{1.4-9}$$

$$f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_X(\omega)\mathrm{e}^{-\mathrm{j}\omega x}\mathrm{d}\omega \tag{1.4-10}$$

因为式(1.4-9)是特征函数的定义，只要证明式(1.4-10)成立即可。将式(1.4-9)代入上式右端，并交换积分顺序

$$\begin{aligned}\frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_X(\omega)\mathrm{e}^{-\mathrm{j}\omega x}\mathrm{d}\omega &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\mathrm{j}\omega x}\left[\int_{-\infty}^{\infty} f_X(\lambda)\mathrm{e}^{\mathrm{j}\omega\lambda}\mathrm{d}\lambda\right]\mathrm{d}\omega \\ &= \int_{-\infty}^{\infty} f_X(\lambda)\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{j}\omega(\lambda - x)}\mathrm{d}\omega\right]\mathrm{d}\lambda \\ &= \int_{-\infty}^{\infty} f_X(\lambda)\delta(\lambda - x)\mathrm{d}\lambda = f_X(x)\end{aligned}$$

即得证。需要注意的是，特征函数与概率密度之间的关系与傅里叶变换略有不同，指数项差一个负号。

【例 1.4−3】 已知随机变量 X_1 和 X_2 为互相独立的高斯变量，数学期望为零，方差为 1。求 $Y=X_1+X_2$ 的概率密度。

解：已知数学期望为零、方差为 1 的高斯变量的概率密度为

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

先根据定义求 X_1 和 X_2 的特征函数，这里需要借助附录 C 中的傅里叶变换表

$$\mathcal{F}[\mathrm{e}^{-\frac{t^2}{2\sigma^2}}]=\sqrt{2\pi}\sigma\mathrm{e}^{-\frac{\sigma^2\omega^2}{2}}$$

令上式 $\sigma=1$，左端 t 换为ω，右端ω换成 x，并做简单的系数变换

$$\mathcal{F}\left[\frac{1}{2\pi}\mathrm{e}^{-\omega^2/2}\right]=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$$

除了系数 1/2π外，求特征函数相当于求傅里叶反变换，由上式得到 X_1 和 X_2 的特征函数

$$\varPhi_{X_1}(\omega)=\int_{-\infty}^{\infty}f_{X_1}(x)\mathrm{e}^{\mathrm{j}\omega x}\mathrm{d}x=\mathrm{e}^{-\omega^2/2},\quad \varPhi_{X_2}(\omega)=\mathrm{e}^{-\omega^2/2}$$

由特征函数的性质 3 $\qquad \varPhi_Y(\omega)=\varPhi_{X_1}(\omega)\varPhi_{X_2}(\omega)=\mathrm{e}^{-\omega^2}$

再由定义式和傅里叶变换表，便可求得 Y 的概率密度

$$f_Y(y)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\varPhi_Y(\omega)\mathrm{e}^{-\mathrm{j}\omega y}\mathrm{d}\omega=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{-\omega^2}\mathrm{e}^{-\mathrm{j}\omega y}\mathrm{d}\omega=\frac{1}{2\sqrt{\pi}}\mathrm{e}^{-y^2/4}$$

由此可见，本例借助傅里叶变换比直接求两个随机变量之和的概率密度要简单得多。

1.4.3 特征函数与矩函数的关系

特征函数与矩函数是一一对应的。一方面，由随机变量的特征函数可以求得其任意阶原点矩；另一方面，如果随机变量的任意阶原点矩都存在，那么可以由原点矩唯一确定随机变量的特征函数。下面先证明原点矩可由特征函数唯一确定，实际上 X 的 n 阶原点矩(如果存在)与特征函数的关系可表达为

$$E[X^n]=\int_{-\infty}^{\infty}x^nf_X(x)\mathrm{d}x=(-\mathrm{j})^n\left.\frac{\mathrm{d}^n\varPhi_X(\omega)}{\mathrm{d}\omega^n}\right|_{\omega=0} \tag{1.4−11}$$

式(1.4−11)的证明可以对特征函数求 n 阶导数，然后令$\omega=0$：

$$\left.\frac{\mathrm{d}^n\varPhi_X(\omega)}{\mathrm{d}\omega^n}\right|_{\omega=0}=\mathrm{j}^n\int_{-\infty}^{\infty}x^n\mathrm{e}^{\mathrm{j}\omega x}f_X(x)\,\mathrm{d}x\Big|_{\omega=0}=\mathrm{j}^n\int_{-\infty}^{\infty}x^nf_X(x)\mathrm{d}x=\mathrm{j}^nE[X^n] \tag{1.4−12}$$

再证逆过程：特征函数由各阶原点矩唯一确定。将特征函数展开成麦克劳林级数

$$\varPhi_X(\omega)=\varPhi_X(0)+\varPhi_X'(0)\omega+\varPhi_X''(0)\frac{\omega^2}{2}+\cdots+\varPhi_X^{(n)}(0)\frac{\omega^n}{n!}+\cdots$$

将特征函数的各阶导数代入上式

$$\Phi_X(\omega)=\sum_{n=0}^{\infty}\frac{\mathrm{d}^n\,\Phi_X(\omega)}{\mathrm{d}\omega^n}\bigg|_{\omega=0}\frac{\omega^n}{n!}=\sum_{n=0}^{\infty}E[X^n]\frac{(\mathrm{j}\omega)^n}{n!} \tag{1.4-13}$$

因此特征函数可以通过各阶原点矩唯一地确定。同样也可把第二特征函数展开成麦克劳林级数

$$\Psi_X(\omega)=\ln\Phi_X(\omega)=\sum_{n=0}^{\infty}c_n\frac{(\mathrm{j}\omega)^n}{n!} \tag{1.4-14}$$

式中

$$c_n=(-\mathrm{j})^n\frac{\mathrm{d}^n}{\mathrm{d}\omega^n}\ln\Phi_X(\omega)\Big|_{\omega=0}=(-\mathrm{j})^n\frac{\mathrm{d}^n}{\mathrm{d}\omega^n}\Psi_X(\omega)\Big|_{\omega=0} \tag{1.4-15}$$

c_n称为随机变量X的n阶累积量，由于c_n是用第二特征函数定义的，因此第二特征函数也称为累积量生成函数。将上式与定义式比较可知，随机变量X的n阶矩和n阶累积量有着密切的联系。有关累积量的性质和更多的知识可参考本书参考文献[10]中的第 1 章。

【例 1.4−4】 求数学期望为零、方差为σ^2的高斯变量X的各阶矩和各阶累积量。

解：已知数学期望为零、方差为σ^2的高斯变量X的概率密度为

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

可求得特征函数

$$\Phi_X(\omega)=\int_{-\infty}^{\infty}f_X(x)\mathrm{e}^{\mathrm{j}\omega x}\mathrm{d}x=\exp\left[-\frac{\sigma^2\omega^2}{2}\right]$$

再利用特征函数的导数来求一、二阶矩

$$E[X]=-\mathrm{j}\left(-\sigma^2\omega\exp\left[-\frac{\sigma^2\omega^2}{2}\right]\right)\bigg|_{\omega=0}=0$$

$$E[X^2]=(-\mathrm{j})^2\left\{(-\sigma^2\omega)^2\exp\left[-\frac{\sigma^2\omega^2}{2}\right]-\sigma^2\exp\left[-\frac{\sigma^2\omega^2}{2}\right]\right\}\bigg|_{\omega=0}=\sigma^2$$

继续求出n阶矩

$$E[X^n]=\begin{cases}1\times3\times5\times\cdots\times(n-1)\sigma^n, & n\text{为偶}\\0, & n\text{为奇}\end{cases}$$

可见，高斯变量的n阶矩除了与阶数有关，还与方差有关。另一方面由第二特征函数

$$\Psi_X(\omega)=\ln\Phi_X(\omega)=-\frac{\sigma^2\omega^2}{2}$$

根据累积量与第二特征函数的关系式，得到各阶累积量

$$c_1=0,\ c_2=\sigma^2,\ c_n=0,(n>2)$$

可见，数学期望为零的高斯变量的前三阶矩与相应阶的累积量相同。

这个例子得到的结论是：对于高斯变量而言，高阶矩只与一、二阶矩有关。从高阶累积量也可得到类似的结果，因高斯变量的n阶累积量在$n>2$时为零。这个结论给出的启示是，当存在加性高斯噪声时，由于高斯噪声的高阶累积量为零，因此可以在高阶累积量上检测非高斯信号。

1.4.4　联合特征函数与联合累积量

二维随机变量(X, Y)的特征函数称为联合特征函数，定义为

$$\Phi_{XY}(\omega_1,\omega_2)=E[\mathrm{e}^{\mathrm{j}(\omega_1 X+\omega_2 Y)}] \tag{1.4-16}$$

与一维随机变量相似，联合特征函数与联合概率密度的关系可表示为

$$\Phi_{XY}(\omega_1,\omega_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\mathrm{e}^{\mathrm{j}(\omega_1 x+\omega_2 y)}\mathrm{d}x\mathrm{d}y \tag{1.4-17}$$

$$f_{XY}(x,y)=\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Phi_{XY}(\omega_1,\omega_2)\mathrm{e}^{-\mathrm{j}(\omega_1 x+\omega_2 y)}\mathrm{d}\omega_1\mathrm{d}\omega_2 \tag{1.4-18}$$

同样，联合特征函数和各阶联合矩有如下关系

$$m_{nk}=(-\mathrm{j})^{n+k}\left.\frac{\partial^{n+k}\Phi_{XY}(\omega_1,\omega_2)}{\partial\omega_1^n\partial\omega_2^k}\right|_{\substack{\omega_1=0\\ \omega_2=0}} \tag{1.4-19}$$

与联合特征函数有关的两个边缘特征函数是

$$\Phi_X(\omega_1)=\Phi_{XY}(\omega_1,0),\ \ \Phi_Y(\omega_2)=\Phi_{XY}(0,\omega_2)$$

第二联合特征函数定义为 $\qquad \Psi_{XY}(\omega_1,\omega_2)=\ln\Phi_{XY}(\omega_1,\omega_2)$ (1.4-20)

联合累积量与第二特征函数的关系式和联合矩与特征函数的关系式相似

$$c_{nk}=(-\mathrm{j})^{n+k}\left.\frac{\partial^{n+k}\Psi_{XY}(\omega_1,\omega_2)}{\partial\omega_1^n\partial\omega_2^k}\right|_{\substack{\omega_1=0\\ \omega_2=0}} \tag{1.4-21}$$

N 维随机变量的联合特征函数可由式(1.4-18)推广得到。N 维联合特征函数的一个重要性质是：当 N 个随机变量互相独立时，其联合特征函数是 N 个随机变量的特征函数的乘积，即

$$\Phi_{X_1X_2\cdots X_N}(\omega_1,\omega_2,\cdots,\omega_N)=\prod_{n=1}^{N}\Phi_{X_n}(\omega_n) \tag{1.4-22}$$

【例 1.4-5】 已知四个数学期望为零、方差为σ^2 且互相独立的高斯变量 X_1,X_2,X_3,X_4，求(X_1,X_2)、(X_1,X_2,X_3)、(X_1,X_2,X_3,X_4)的联合累积量。

解： 将 (X_1,X_2,X_3,X_4) 作为四维随机变量，由于互相独立，第二联合特征函数为

$$\begin{aligned}\Psi_{X_1X_2X_3X_4}(\omega_1,\omega_2,\omega_3,\omega_4)&=\ln\Phi_{X_1X_2X_3X_4}(\omega_1,\omega_2,\omega_3,\omega_4)\\&=\ln[\Phi_{X_1}(\omega_1)\Phi_{X_2}(\omega_2)\Phi_{X_3}(\omega_3)\Phi_{X_4}(\omega_4)]\end{aligned}$$

已知数学期望为零、方差为 σ^2 的高斯变量 X 的第二特征函数为

$$\Psi_X(\omega)=\ln\Phi_X(\omega)=-\frac{\sigma^2\omega^2}{2}$$

可以得到第二联合特征函数

$$\Psi_{X_1X_2X_3X_4}(\omega_1,\omega_2,\omega_3,\omega_4)=-\frac{\sigma^2\omega_1^2}{2}-\frac{\sigma^2\omega_2^2}{2}-\frac{\sigma^2\omega_3^2}{2}-\frac{\sigma^2\omega_4^2}{2}$$

再由式(1.4-21)可以得到(X_1,X_2)的二阶联合累积量

$$c_{X_1X_2}=(-\mathrm{j})^2\left.\frac{\partial^2\psi_{X_1X_2}(\omega_1,\omega_2)}{\partial\omega_1\partial\omega_2}\right|_{\substack{\omega_1=0\\ \omega_2=0}}=(-\mathrm{j})^2\frac{\partial^2}{\partial\omega_1\partial\omega_2}\left[-\frac{\sigma^2\omega_1^2}{2}-\frac{\sigma^2\omega_2^2}{2}\right]\Bigg|_{\substack{\omega_1=0\\ \omega_2=0}}=0$$

由于第二联合特征函数的每一项只有一个变量 ω_i，所以对不同的两个变量求偏导数为零。同理，有三阶联合累积量

$$c_{X_1X_2X_3} = (-\mathrm{j})^3 \left.\frac{\partial^3 \Psi_{X_1X_2X_3}(\omega_1,\omega_2,\omega_3)}{\partial\omega_1\partial\omega_2\partial\omega_3}\right|_{\substack{\omega_1=0\\ \omega_2=0\\ \omega_3=0}} = 0$$

四阶联合累积量
$$c_{X_1X_2X_3X_4} = (-\mathrm{j})^4 \left.\frac{\partial^4 \Psi_{X_1X_2X_3X_4}(\omega_1,\omega_2,\omega_3,\omega_4)}{\partial\omega_1\partial\omega_2\partial\omega_3\partial\omega_4}\right|_{\substack{\omega_1=0\\ \omega_2=0\\ \omega_3=0\\ \omega_4=0}} = 0$$

1.5　随机信号实用分布律

在概率论中已经学过一些分布，本节除了复习一些简单的分布律之外，重点讨论高斯分布及电子信息领域常用的以高斯分布为基础变换的分布律。

1.5.1　一些简单的分布律

1．0-1 分布

0-1 分布又称两点分布或伯努利分布，对应随机变量的取值仅有 0 和 1 两个值。其分布列为

$$P(X=1)=p\,,\quad P(X=0)=1-p$$

分布函数为
$$F(x)=pu(x-1)+(1-p)u(x)$$

通过定义可以直接求得，0-1 分布的均值为 p，方差为 $p(1-p)$。

0-1 分布是最基础的随机分布，其与二项分布、几何分布等都具有密切的联系。而且 0-1 分布具有明确的物理意义，例如在硬币投掷试验、信号有无检测、事件指示函数、随机变量二元量化等方面都有重要应用。

2．二项分布

二项分布 $B(n,p)$ 中的参数 n 表示进行 n 次独立试验，参数 p 表示每次试验成功的概率为 p，试验不成功的概率为 $1-p$。那么二项分布 $B(n,p)$ 表征了在这 n 次独立试验中，成功 m 次的概率分布，即

$$P_n(m)=C_n^m p^m (1-p)^{n-m}\,,\quad m=0,1,\cdots,n \tag{1.5-1}$$

其概率分布函数可表示成

$$F(x)=\sum_{m=0}^{n} C_n^m p^m (1-p)^{n-m} u(x-m) \tag{1.5-2}$$

式中二项式系数
$$C_n^m=\frac{n!}{m!(n-m)!} \tag{1.5-3}$$

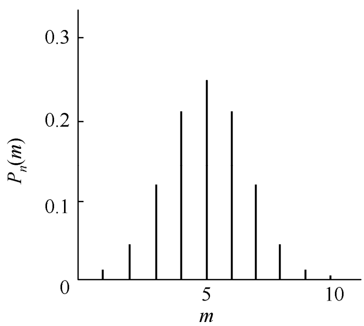


图 1.5-1　二项分布

图 1.5-1 示意了二项分布的取值概率 $P_n(m)$。

二项分布与 0-1 分布有密切的联系。假设 $X_1,X_2,\cdots,X_n$ 是 n 个独立同分布的 0-1 分布随机变量，且 $P(X_i=1)=p$。X 服从二项分布 $B(n,p)$，由二项分布的定义可知 $X=\sum_{i=1}^{n} X_i$。

在信号检测理论中，非参量检测时单次探测的秩为某一值的概率服从二项分布。

【例 1.5−1】 设随机变量 X 为二项式分布 $B(n,p)$，求 X 的数学期望和方差。

解：二项分布可以写为 0-1 分布随机变量的累加和。令 $X=\sum_{i=1}^{n}X_i$，其中 X_i 是独立同分布的 0-1 分布随机变量，利用均值和方差的运算性质，有

$$E[X]=\sum_{m=0}^{n}E[X_i]=np \quad 和 \quad D[X]=\sum_{m=0}^{n}D[X_i]=np(1-p)$$

3．泊松分布

泊松分布常用来描述单位时间内随机事件发生的次数。可以证明，泊松分布可以作为二项分布的极限而得到。若 $X\sim B(n,p)$，其中概率 p 很小，试验次数 n 很大，且 $np=\lambda$ 为常数时，X 的分布接近于泊松分布。泊松分布的分布列为

$$P_n(m)=\frac{\lambda^m}{m!}\mathrm{e}^{-\lambda}，\quad m=0,1,\cdots \qquad (1.5-4)$$

其中$\lambda>0$ 称为速率常数，表示单位时间内随机事件发生的平均次数。若λ 为整数，$P_n(m)$在 $m=\lambda$ 及 $m=\lambda-1$ 时达到最大值。泊松分布是非对称的，但λ 越大，非对称性越不明显。为了比较方便，图 1.5−2 将不同λ 值的 $P_n(m)$画在一起。注意泊松分布是离散分布，图中的曲线是将 m 在整数处的取值连在一起绘制的。

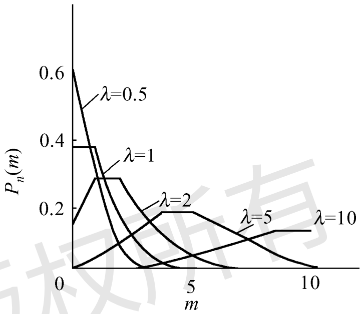


图 1.5−2　泊松分布

【例 1.5−2】 设随机变量 X 是分布列为式(1.5−4)的泊松分布，求 X 的数学期望和方差。

解：$m_X=E[X]=\sum_{m=0}^{n}mP(X=m)=\sum_{m=0}^{n}m\frac{\lambda^m}{m!}\mathrm{e}^{-\lambda}=\lambda\sum_{m=1}^{n}\frac{\lambda^{m-1}}{(m-1)!}\mathrm{e}^{-\lambda}=\lambda\mathrm{e}^{\lambda}\mathrm{e}^{-\lambda}=\lambda$

同样的方法可求得

$$E[X^2]=E[X(X-1)+X]=E[X(X-1)]+E[X]=\sum_{m=0}^{n}m(m-1)\frac{\lambda^m}{m!}\mathrm{e}^{-\lambda}+\lambda$$

$$=\lambda^2\sum_{m=2}^{n}\frac{\lambda^{m-2}}{(m-2)!}\mathrm{e}^{-\lambda}+\lambda=\lambda^2\mathrm{e}^{\lambda}\mathrm{e}^{-\lambda}+\lambda=\lambda^2+\lambda$$

$$D[X]=E[X^2]-E^2[X]=\lambda$$

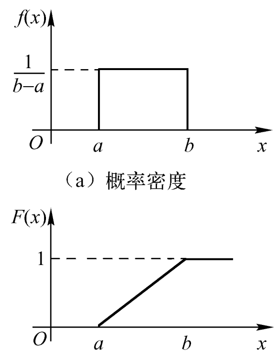
4．均匀分布

如果随机变量 X 的概率密度满足

$$f_X(x)=\begin{cases}\dfrac{1}{b-a}, & a\leqslant x\leqslant b\\ 0, & 其他\end{cases} \qquad (1.5-5)$$

则称 X 为在$[a,b]$区间内均匀分布的随机变量。很容易证明其概率分布函数为

$$F_X(x)=\begin{cases}0, & x<a\\ \dfrac{x-a}{b-a}, & a\leqslant x<b\\ 1, & x\geqslant b\end{cases} \qquad (1.5-6)$$



（a）概率密度

（b）概率分布函数

图 1.5−3　均匀分布随机变量

均匀分布是常用的分布律之一。图 1.5−3 是均匀分布的概率密度和概率分布函数。

均匀分布的数学期望和方差分别为

$$m=\frac{a+b}{2} \qquad 和 \qquad \sigma^2=\frac{(b-a)^2}{12} \tag{1.5-7}$$

在误差分析时经常遇到均匀分布，如数字信号中的量化噪声。由于 A/D 转换器的字长有限，模拟信号通过 A/D 转换时，势必要舍弃部分信息。丢失信息后相当于使信号附加了一部分噪声，称为量化噪声。量化噪声分为截尾噪声和舍入噪声，它们都是均匀分布的，且方差相同，不同的是分布的区间。若量化的最小单位是ε，舍入噪声在$[-\varepsilon/2,\varepsilon/2]$内均匀分布，数学期望为零；而截尾噪声在$[-\varepsilon,0]$内均匀分布，因此数学期望为$-\varepsilon/2$。

1.5.2　高斯分布（正态分布）

高斯分布律不仅在统计数学中占有重要的位置，在通信与信号处理领域中也是应用最广泛的分布。高斯分布广泛应用的主要原因之一在于中心极限定理：大量独立同分布的随机变量之和趋近于高斯分布。除此之外，高斯分布有很多独特的性质，本节将对高斯分布进行集讨论。

1．中心极限定理

中心极限定理的内容如下：设$X_1,X_2,\cdots$是独立同分布的随机变量序列，记$m_X=E[X_1]$，$\sigma_X^2=D[X_1]$。如果$\sigma_X^2<+\infty$，令$S_n=\sum_{i=1}^{n}X_i$，则当$n\to\infty$时

$$\sqrt{n}\left(\frac{1}{n}S_n-m_X\right)\xrightarrow{d}N(0,\sigma_X^2) \tag{1.5-8}$$

其中$\xrightarrow{d}$表示依分布收敛，表示$N(0,\sigma_X^2)$是均值为 0、方差为σ_X^2的正态分布。

出于严谨的数学表达，中心极限定理中对随机变量累加和的均值S_n/n进行了平移和尺度变换，变换后的随机变量依分布收敛于高斯随机变量。在工程应用中，常常对独立同分布的随机变量进行相加后，用高斯分布近似。而且，即使n个独立随机变量不是相同分布的，当n足够大时，如果任意一个随机变量都不占优或对和的影响足够小，那么它们之和的分布也近似于高斯分布。关于中心极限定理及证明，请参考有关书籍。(如参考文献[24]，5.4 节)。

下面以均匀分布为例，演示随机变量累加和分布的演化过程。若 n 个随机变量 X_i $(i=1,2,\cdots,n)$都为[0,1]区间上均匀分布的随机变量，且互相独立，其累加和$S_n=\sum_{i=1}^{n}X_i$的均值和方差分别为

$$m_{S_n}=\sum_{i=1}^{n}m_{X_i} \tag{1.5-9}$$

$$\sigma_{S_n}^2=\sum_{i=1}^{n}\sigma_{X_i}^2 \tag{1.5-10}$$

由于均匀分布的概率密度无峰，在对称分布中，不论是误差分析还是信号检测，都是最不利的分布。$n=1$ 自然是均匀分布，数学期望为 1，方差为 1/12。当$n=2$时，Y的分布为三角形函数，有

$$f(y)=\begin{cases} y, & 0\leqslant y<1 \\ 2-y, & 1\leqslant y<2 \\ 0, & 其他 \end{cases} \tag{1.5-11}$$

它的数学期望为 1，方差为 1/6。$n=3$ 时，Y 的分布为三段抛物线，有

$$f(y)=\begin{cases}\dfrac{1}{2}y^2, & 0\leqslant y<1\\ \dfrac{1}{2}(-2y^2+6y-3), & 1\leqslant y<2\\ \dfrac{1}{2}(y-3)^2, & 2\leqslant y<3\\ 0, & \text{其他}\end{cases} \tag{1.5-12}$$

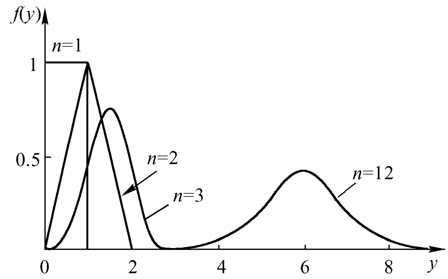


图 1.5-4　n 个独立均匀随机变量之和的分布

其数学期望为 1.5，方差为 1/4。当 $n=12$ 时，Y 就相当接近高斯分布，其数学期望为 6，方差为 1，如图 1.5-4 所示。以上讨论给我们的启示是：在仿真高斯变量时，如果有足够数量的高质量的均匀分布随机数，可以通过相加的方法得到高斯变量(具体方法请见 1.6 节中的例 1.6-4)。

在实际应用中，许多独立噪声之和若满足中心极限定理中某一定理的条件，就认为是高斯分布的。一般情况下，不同分布律的随机变量之和趋向高斯分布的速度是不同的。在工程上，如果不是某个或某些随机变量对和的贡献很大，7～10 个随机变量之和的分布就认为是高斯分布的。

2．一维高斯变量

高斯分布也称正态分布，高斯分布的随机变量 X 的概率密度为

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \tag{1.5-13}$$

式中，常数 m 和 $\sigma(\sigma>0)$ 分别为 X 的均值和标准差。高斯分布记为 $N(m,\sigma^2)$。

由 1.3 节的例 1.3-1，式(1.5-13)中的 m 和 σ^2 恰好是高斯变量的数学期望和方差，因此一维高斯分布律唯一地由它的数学期望和方差决定，见图 1.5-5。当概率密度曲线的一阶导数为零时，极值(最大值) $1/\sqrt{2\pi\sigma^2}$ 发生在 $x=m$ 处。数学期望 m 决定概率密度曲线在横轴所处的位置，方差 σ^2 决定纵向高度。因概率密度曲线下的面积为 1，当 σ 减小时，曲线变得尖锐，取值落在 m 附近固定区间的概率增大，这意味着取值的离散程度减小。当概率密度曲线的二阶导数为零时，拐点发生在 $x=m\pm\sigma$ 处。

对高斯变量进行归一化处理后的随机变量，称为归一化高斯变量。令 $Y=(X-m)/\sigma$，归一化后的概率密度为

$$f_Y(y)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-y^2/2} \tag{1.5-14}$$

归一化高斯变量 Y 是数学期望为 0、方差为 1 的高斯变量，对应的分布称为标准正态分布(见图 1.5-6)。标准正态分布的概率密度曲线对称于纵轴，最大值为 $1/\sqrt{2\pi}$。

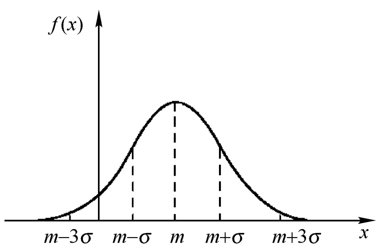


图 1.5-5　高斯变量的概率密度

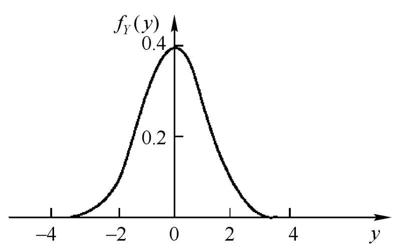


图 1.5-6　归一化高斯变量的概率密度

对概率密度积分可求出概率分布函数

$$F_X(x)=\int_{-\infty}^{x} f_X(x_1)\mathrm{d}x_1=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x_1-m)^2}{2\sigma^2}\right]\mathrm{d}x_1$$

令 $t=\dfrac{x_1-m}{\sigma}$, $\mathrm{d}x_1=\sigma\mathrm{d}t$，则

$$F_X(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{(x-m)/\sigma}\mathrm{e}^{-t^2/2}\mathrm{d}t=\Phi\left(\frac{x-m}{\sigma}\right) \tag{1.5-15}$$

式中

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-t^2/2}\mathrm{d}t \tag{1.5-16}$$

是标准正态分布的分布函数，它的值可通过数学手册中的正态分布表查到。式(1.5-15)说明任意高斯分布的分布函数(因此，在某区间内的概率取值)可通过查阅标准正态分布表得到。

由于分布函数定义为随机变量不超过某值的概率，因此式(1.5-16)表示的分布函数有以下三个主要性质：

性质 1
$$\Phi(-x)=1-\Phi(x) \tag{1.5-17}$$

性质 2
$$F(m)=\Phi(0)=0.5 \tag{1.5-18}$$

性质 3
$$P(\alpha<X\leqslant\beta)=F(\beta)-F(\alpha)=\Phi\left(\frac{\beta-m}{\sigma}\right)-\Phi\left(\frac{\alpha-m}{\sigma}\right) \tag{1.5-19}$$

上式是通过标准正态分布的分布函数给出的随机变量在区间 (α,β) 的取值概率。图 1.5-7 示出了标准正态分布的分布函数性质。

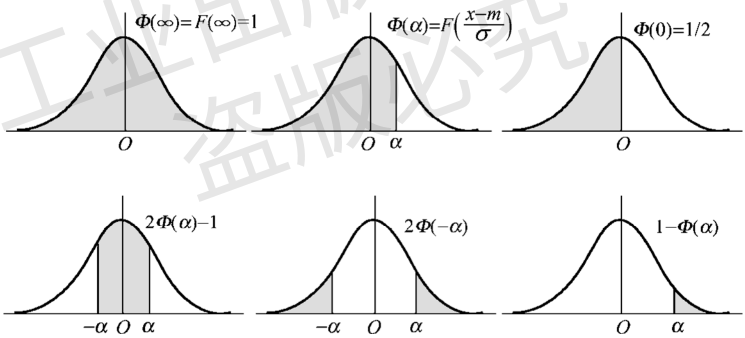


图 1.5-7　标准正态分布的分布函数性质

【例 1.5-3】 求高斯变量在 $m\pm3\sigma$ 区间上的取值概率(参见图 1.5-5)。

解：由性质 3

$$P(m-3\sigma<X<m+3\sigma)=\Phi\left(\frac{m+3\sigma-m}{\sigma}\right)-\Phi\left(\frac{m-3\sigma-m}{\sigma}\right)$$
$$=\Phi(3)-\Phi(-3)=2\Phi(3)-1=0.997$$

式中，$\Phi(3)=0.998\,65$ 是查数学手册中的正态分布表得到的。这说明虽然高斯变量的取值区间为 $(-\infty,\infty)$，但实际取值落在 $x=m\pm3\sigma$ 区间上的概率已高达 99.7%以上。

高斯变量的中心矩为
$$\mu_n=E[(X-m)^n]=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}(x-m)^n\exp\left[-\frac{(x-m)^2}{2\sigma^2}\right]\mathrm{d}x \tag{1.5-20}$$

令 $t=\dfrac{x-m}{\sqrt{2}\sigma}$， $\mathrm{d}t=\dfrac{1}{\sqrt{2}\sigma}\mathrm{d}x$，整理后得到

$$\mu_n = \frac{(\sqrt{2}\sigma)^n}{\sqrt{\pi}}\int_{-\infty}^{\infty} t^n \mathrm{e}^{-t^2}\mathrm{d}t$$

由于 e^{-t^2} 是偶函数，当 n 为奇数时，上式积分为零。当 n 为非零偶数时，令 $n=2m$ $(m=1,2,\cdots)$，根据例 1.3−1 给出的积分

$$\mu_n = \frac{(\sqrt{2}\sigma)^{2m}}{\sqrt{\pi}}\cdot 2\int_0^{\infty} t^{2m}\mathrm{e}^{-t^2}\mathrm{d}t = \frac{(\sqrt{2}\sigma)^{2m}}{\sqrt{\pi}}\cdot\frac{2(2m-1)!!}{2^{m+1}}\sqrt{\pi} = (2m-1)!!\sigma^{2m}$$

再将 $n=2m$ 代入上式，并考虑 n 为奇数的情况

$$\mu_n = \begin{cases}(n-1)!!\sigma^n, & n\text{为偶数}\\ 0, & n\text{为奇数}\end{cases} \tag{1.5−21}$$

在例 1.4−4 中曾经求过数学期望为零的高斯变量的 n 阶原点矩，实际上也就是数学期望为 m 的 n 阶中心矩。由例 1.4−3，已求出归一化高斯变量 Y 的特征函数

$$\varPhi_Y(\omega) = \mathrm{e}^{-\omega^2/2}$$

根据特征函数的性质 2，对于数学期望为 m、方差为 σ^2 的高斯变量 $X=\sigma Y+m$，特征函数为

$$\varPhi_X(\omega) = \mathrm{e}^{\mathrm{j}\omega m-\frac{\omega^2\sigma^2}{2}} \tag{1.5−22}$$

高斯分布律的另一个特点是：高斯变量之和仍为高斯变量。下面以两个高斯变量之和为例来说明这个结论。

【例 1.5−4】 求两个数学期望和方差不同且互相独立的高斯变量 X_1 和 X_2 之和的概率密度。

解： 设 $Y=X_1+X_2$，由式(1.2−12)，两个互相独立的随机变量之和的概率密度为

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(x_1) f_{X_2}(y-x_1)\mathrm{d}x_1$$

将 X_1 和 X_2 的概率密度代入上式

$$f_Y(y) = \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty}\exp\left[-\frac{(x_1-m_1)^2}{2\sigma_1^2}\right]\exp\left[-\frac{(y-x_1-m_2)^2}{2\sigma_2^2}\right]\mathrm{d}x_1 = \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty}\mathrm{e}^{-Ax_1^2+2Bx_1-C}\mathrm{d}x_1$$

利用欧拉公式得 $f_Y(y) = \frac{1}{2\pi\sigma_1\sigma_2}\sqrt{\frac{\pi}{A}}\cdot\exp\left[-\frac{AC-B^2}{A}\right] = \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}}\exp\left[-\frac{[y-(m_1+m_2)]^2}{2(\sigma_1^2+\sigma_2^2)}\right]$

显然，Y 也是高斯变量，且数学期望和方差分别为

$$m_Y = m_1+m_2\text{，}\quad \sigma_Y^2 = \sigma_1^2+\sigma_2^2$$

推广到多个互相独立的随机变量，其和也是高斯分布的。令

$$Y = \sum_{i=1}^{n} X_i \tag{1.5−23}$$

若 $X_i(i=1,2,\cdots,n)$ 的数学期望和方差为 m_i 和 σ_i^2，且 X_i 间相互独立，则 Y 的数学期望和方差分别为

$$m_Y = \sum_{i=1}^{n} m_i \tag{1.5−24}$$

$$\sigma_Y^2 = \sum_{i=1}^{n}\sigma_i^2 \tag{1.5−25}$$

如果求和的高斯变量间不是互相独立的，而是联合高斯的，则可以证明 Y 依然是高斯分布，但式(1.5−25)应修正为

$$\sigma_Y^2=\sum_{i=1}^{n}\sigma_i^2+2\sum_{i<j}r_{ij}\sigma_i\sigma_j \tag{1.5−26}$$

式中，r_{ij} 是 X_i 与 X_j 之间的相关系数。

3．二维高斯分布

（1）联合概率密度

对于互相独立的两个高斯变量 X_1 和 X_2，如果数学期望为零，方差分别为 σ_1^2 和 σ_2^2，则它们的联合概率密度是两个一维概率密度的乘积

$$f_X(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2}\exp\left[-\frac{x_1^2}{2\sigma_1^2}-\frac{x_2^2}{2\sigma_2^2}\right] \tag{1.5−27}$$

下面用互相独立的 X_1 和 X_2 构造两个相关的高斯变量 Y_1 和 Y_2。令

$$\begin{cases}Y_1=X_1\cos\theta-X_2\sin\theta\\ Y_2=X_1\sin\theta+X_2\cos\theta\end{cases} \tag{1.5−28}$$

从上式可看出 Y_1 和 Y_2 的数学期望仍为零，但方差变为

$$\begin{cases}\sigma_{Y_1}^2=\sigma_1^2\cos^2\theta+\sigma_2^2\sin^2\theta\\ \sigma_{Y_2}^2=\sigma_1^2\sin^2\theta+\sigma_2^2\cos^2\theta\end{cases} \tag{1.5−29}$$

Y_1 和 Y_2 的协方差为

$$\mu_{11}=E[Y_1Y_2]=(\sigma_1^2-\sigma_2^2)\sin\theta\cos\theta \tag{1.5−30}$$

下一步利用二维函数变换求 Y_1 和 Y_2 的联合概率密度。先求反函数及雅可比行列式

$$\begin{cases}X_1=Y_1\cos\theta+Y_2\sin\theta\\ X_2=-Y_1\sin\theta+Y_2\cos\theta\end{cases} \tag{1.5−31}$$

$$J=\begin{vmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{vmatrix}=1 \tag{1.5−32}$$

于是

$$\begin{aligned}f_Y(y_1,y_2)&=\frac{1}{2\pi\sigma_1\sigma_2}\exp\left[-\frac{(y_1\cos\theta+y_2\sin\theta)^2}{2\sigma_1^2}-\frac{(-y_1\sin\theta+y_2\cos\theta)^2}{2\sigma_2^2}\right]\\ &=\frac{1}{2\pi\sqrt{\sigma_{Y_1}^2\sigma_{Y_2}^2-\mu_{11}^2}}\exp\left[-\frac{\sigma_{Y_1}^2y_1^2-2\mu_{11}y_1y_2+\sigma_{Y_2}^2y_2^2}{2(\sigma_{Y_1}^2\sigma_{Y_2}^2-\mu_{11}^2)}\right]\end{aligned} \tag{1.5−33}$$

对于归一化随机变量，$\sigma_{Y_1}^2=\sigma_{Y_2}^2=1$，$m_{Y_1}=m_{Y_2}=0$，则

$$f_Y(y_1,y_2)=\frac{1}{2\pi\sqrt{1-r^2}}\exp\left[-\frac{y_1^2-2ry_1y_2+y_2^2}{2(1-r^2)}\right] \tag{1.5−34}$$

式中，r 是 Y_1 和 Y_2 的相关系数。更一般的形式是

$$f_{\boldsymbol{Y}}(y_1,y_2)=\frac{1}{2\pi\sigma_{Y_1}\sigma_{Y_2}\sqrt{1-r^2}}\cdot\exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{(y_1-m_{Y_1})^2}{\sigma_{Y_1}^2}-\frac{2r(y_1-m_{Y_1})(y_2-m_{Y_2})}{\sigma_{Y_1}\sigma_{Y_2}}+\frac{(y_2-m_{Y_2})^2}{\sigma_{Y_2}^2}\right]\right\}$$

（1.5−35）

式中，m_{Y_1} 和 m_{Y_2} 分别为 Y_1 和 Y_2 的数学期望。显而易见，两个非独立的高斯变量的联合概率密度与它们的数学期望、方差和相关系数都有关。

如果令式(1.5−35)中的 r 为零，即假设 Y_1 和 Y_2 是不相关的，于是有

$$\begin{aligned}f_Y(y_1,y_2)&=\frac{1}{2\pi\sigma_{Y_1}\sigma_{Y_2}}\exp\left\{-\frac{1}{2}\left[\frac{(y_1-m_{Y_1})^2}{\sigma_{Y_1}^2}+\frac{(y_2-m_{Y_2})^2}{\sigma_{Y_2}^2}\right]\right\}\\&=\frac{1}{\sqrt{2\pi}\sigma_{Y_1}}\exp\left[-\frac{(y_1-m_{Y_1})^2}{2\sigma_{Y_1}^2}\right]\cdot\frac{1}{\sqrt{2\pi}\sigma_{Y_2}}\exp\left[-\frac{(y_2-m_{Y_2})^2}{2\sigma_{Y_2}^2}\right]\\&=f_{Y_1}(y_1)f_{Y_2}(y_2)\end{aligned}\qquad(1.5\text{−}36)$$

上式说明不相关的高斯变量一定是互相独立的。也就是说，对于高斯变量不相关与统计独立是等价的。

（2）联合特征函数

对于数学期望为零且互相独立的高斯变量 X_1 和 X_2，根据式(1.4−22)和例 1.4−3，其联合特征函数为

$$\Phi_X(\omega_1,\omega_2)=\Phi_{X_1}(\omega_1)\Phi_{X_2}(\omega_2)=\exp\left[-\frac{1}{2}(\sigma_1^2\omega_1^2+\sigma_2^2\omega_2^2)\right]\qquad(1.5\text{−}37)$$

经过式(1.5−28)的变换后，相关系数为 r 的高斯变量 Y_1 和 Y_2 的联合特征函数为

$$\Phi_{\boldsymbol{Y}}(\omega_1,\omega_2)=\exp\left[-\frac{1}{2}(\sigma_{Y_1}^2\omega_1^2+2r\sigma_{Y_1}\sigma_{Y_2}\omega_1\omega_2+\sigma_{Y_2}^2\omega_2^2)\right]\qquad(1.5\text{−}38)$$

如果 Y_1 和 Y_2 是归一化高斯变量，则有

$$\Phi_{\boldsymbol{Y}}(\omega_1,\omega_2)=\exp\left[-\frac{1}{2}(\omega_1^2+2r\omega_1\omega_2+\omega_2^2)\right]\qquad(1.5\text{−}39)$$

而联合特征函数的一般形式为

$$\Phi_{\boldsymbol{Y}}(\omega_1,\omega_2)=\exp\left[\mathrm{j}(m_{Y_1}\omega_1+m_{Y_2}\omega_2)\right]\cdot\exp\left[-\frac{1}{2}(\sigma_{Y_1}^2\omega_1^2+2r\sigma_{Y_1}\sigma_{Y_2}\omega_1\omega_2+\sigma_{Y_2}^2\omega_2^2)\right]\qquad(1.5\text{−}40)$$

对于多维高斯变量，一般写成矩阵形式比较简洁。设 n 维随机变量向量为 $\boldsymbol{Y}$，数学期望和方差向量分别为 $\boldsymbol{m}$ 和 $\boldsymbol{s}$，它们具有如下形式

$$\boldsymbol{Y}=\begin{bmatrix}Y_1\\Y_2\\\vdots\\Y_n\end{bmatrix},\quad\boldsymbol{m}=\begin{bmatrix}m_1\\m_2\\\vdots\\m_n\end{bmatrix},\quad\boldsymbol{s}=\begin{bmatrix}\sigma_1^2\\\sigma_2^2\\\vdots\\\sigma_n^2\end{bmatrix}$$

相关矩阵 $$\boldsymbol{R}=E[\boldsymbol{Y}\boldsymbol{Y}^{\mathrm{T}}]=\begin{bmatrix}E[Y_1Y_1]&E[Y_1Y_2]&\cdots&E[Y_1Y_n]\\E[Y_2Y_1]&E[Y_2Y_2]&\cdots&E[Y_2Y_n]\\\vdots&\vdots&\ddots&\vdots\\E[Y_nY_1]&E[Y_nY_2]&\cdots&E[Y_nY_n]\end{bmatrix}=\begin{bmatrix}R_{11}&R_{12}&\cdots&R_{1n}\\R_{21}&R_{22}&\cdots&R_{2n}\\\vdots&\vdots&\ddots&\vdots\\R_{n1}&R_{n2}&\cdots&R_{nn}\end{bmatrix}$$

式中，R_{ij} 是 Y_i 和 Y_j 的相关矩。

协方差矩阵 $$\boldsymbol{C}=E[(\boldsymbol{Y}-\boldsymbol{m})(\boldsymbol{Y}-\boldsymbol{m})^{\mathrm{T}}]=\begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}=\begin{bmatrix} \sigma_1^2 & C_{12} & \cdots & C_{1n} \\ C_{21} & \sigma_2^2 & \cdots & C_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & \sigma_n^2 \end{bmatrix}$$

式中，C_{ij} 是 Y_i 和 Y_j 的协方差 $E[(Y_i-m_i)(Y_j-m_j)]$。由于 C_{ii} 是第 i 个随机变量的方差，n 维协方差阵的对角线为各随机变量的方差。如果 n 维随机变量是方差均不为零的实随机变量，那么协方差阵是实对称的正定矩阵。方差均不为零的复随机变量的协方差阵是厄米特阵。

对应式(1.5−35)形式的 n 维概率密度函数为

$$f_{\boldsymbol{Y}}(\boldsymbol{y})=\frac{1}{\sqrt{(2\pi)^n|\boldsymbol{C}|}}\exp\left[-\frac{1}{2}(\boldsymbol{y}-\boldsymbol{m})^{\mathrm{T}}\boldsymbol{C}^{-1}(\boldsymbol{y}-\boldsymbol{m})\right] \tag{1.5−41}$$

式中，T 表示矩阵转置，$\boldsymbol{C}^{-1}$ 表示协方差阵的逆矩阵。相应的 n 维特征函数矩阵形式为

$$\varPhi_{\boldsymbol{Y}}(\boldsymbol{\omega})=\exp\left[\mathrm{j}\boldsymbol{m}^{\mathrm{T}}\boldsymbol{\omega}-\frac{1}{2}\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{\omega}\right] \tag{1.5−42}$$

式中，$\boldsymbol{\omega}=[\omega_1,\omega_2,\cdots,\omega_n]^{\mathrm{T}}$。

1.5.3 χ^2 分布

在信号的传输过程中，信号一般是窄带形式的，这样不可避免地要用到包络检波。在小信号检波时，通常采用平方律检波，因此检波器的输出是信号与噪声包络的平方。有时为了使信号检测的错误概率更小，还要对检波器的输出信号进行积累。

如果随机变量 X 是高斯分布，那么平方律检波器的输出 X^2 是什么分布呢？对检波器的输出信号 X^2 进行采样后积累的信号 $Y=\sum\limits_{i=1}^{n}X_i^2$ 又是什么分布呢？

下面将看到 Y 为 χ^2 分布。若 X_i 的数学期望为零，则 Y 为中心 χ^2 分布；若 X_i 的数学期望不为零，则 Y 为非中心 χ^2 分布。积累的次数 n 称为 χ^2 分布的自由度。

1．中心 χ^2 分布

如果 n 个互相独立的高斯变量 $X_1, X_2, \cdots, X_n$ 的数学期望都为零，方差均为 1，则它们的平方和

$$Y=\sum_{i=1}^{n}X_i^2 \tag{1.5−43}$$

的分布是具有 n 个自由度的 χ^2 分布。

由于每个高斯变量 X_i 都是归一化高斯变量，其概率密度

$$f_{X_i}(x_i)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x_i^2/2}$$

如果令 $Y_i=X_i^2$，经函数变换后 Y_i 的分布为

$$f_{Y_i}(y_i)=\frac{1}{\sqrt{2\pi y_i}}\mathrm{e}^{-y_i/2}，\ y_i>0 \tag{1.5−44}$$

利用傅里叶变换求 Y_i 的特征函数

$$\Phi_{Y_i}(\omega)=\int_{-\infty}^{\infty} f_{Y_i}(y_i)\mathrm{e}^{\mathrm{j}\omega y_i}\mathrm{d}y_i=(1-2\mathrm{j}\omega)^{-1/2} \tag{1.5-45}$$

X_i 之间互相独立，Y_i 之间也必然互相独立。根据特征函数的性质，互相独立的随机变量之和的特征函数等于各特征函数之积，所以 Y 的特征函数为

$$\Phi_Y(\omega)=\frac{1}{(1-2\mathrm{j}\omega)^{n/2}} \tag{1.5-46}$$

相应的概率密度可用傅里叶反变换求得

$$f_Y(y)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_Y(\omega)\mathrm{e}^{-\mathrm{j}\omega y}\mathrm{d}\omega=\frac{1}{2^{n/2}\Gamma(n/2)}y^{\frac{n}{2}-1}\mathrm{e}^{-\frac{y}{2}},\quad y\geqslant 0 \tag{1.5-47}$$

上式就是 χ^2 分布。式中的伽马函数由下式计算

$$\Gamma(x)=\int_0^{\infty}t^{x-1}\mathrm{e}^{-t}\mathrm{d}t \tag{1.5-48}$$

当 x 可表示为 n 或 n+1/2 的形式时

$$\Gamma\left(n+\frac{1}{2}\right)=\frac{(2n-1)!!}{2^n}\sqrt{\pi} \tag{1.5-49}$$

$$\Gamma(n+1)=n! \tag{1.5-50}$$

对于不同的自由度 n，其概率密度曲线见图 1.5-8。

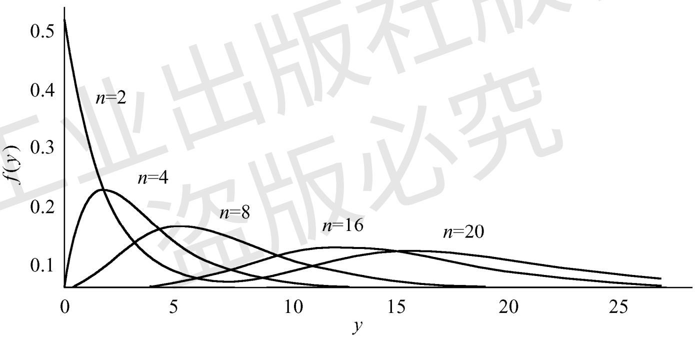


图 1.5-8　不同自由度的 χ^2 分布（σ^2=1）的概率密度曲线

当 $n=1$ 时，1 个自由度的 χ^2 分布为

$$f_Y(y)=\frac{1}{2^{1/2}\Gamma(1/2)}y^{-1/2}\mathrm{e}^{-y/2}=\frac{1}{\sqrt{2\pi y}}\mathrm{e}^{-y/2} \tag{1.5-51}$$

当 $n=2$ 时，2 个自由度的 χ^2 分布简化为指数分布

$$f_Y(y)=\frac{1}{2\Gamma(1)}\mathrm{e}^{-y/2}=\frac{1}{2}\mathrm{e}^{-y/2} \tag{1.5-52}$$

如果互相独立的高斯变量 X_i 的方差不是 1 而是 σ^2，则可做 $\phi(Y)=\sigma^2 Y$ 的变换。变换后的分布为

$$f_Y(y)=\frac{1}{(2\sigma^2)^{n/2}\Gamma(n/2)}y^{\frac{n}{2}-1}\mathrm{e}^{-\frac{y}{2\sigma^2}},\quad y\geqslant 0 \tag{1.5-53}$$

此时 Y 的数学期望和方差为

$$m_Y=n\sigma^2,\quad \sigma_Y^2=2n\sigma^4 \tag{1.5-54}$$

χ^2 分布有一条重要的性质：两个互相独立的具有 χ^2 分布的随机变量之和仍为 χ^2 分布，

若它们的自由度分别为 n_1 和 n_2，其和的自由度为 $n=n_1+n_2$。

在第 4 章中将看到，对窄带信号加窄带高斯噪声进行平方律检波之后，再进行积累就是 χ^2 分布。

2．非中心χ^2分布

如果互相独立的高斯变量 $X_i(i=1,2,\cdots,n)$ 的方差为σ^2，数学期望不是零而是 m_i，则 $Y=\sum_{i=1}^{n}X_i^2$ 为 n 个自由度的非中心χ^2分布。也可把 X_i 看成数学期望仍然为零的高斯变量与确定信号之和。

仍令 $Y_i=X_i^2$，经函数变换后 Y_i 的分布为

$$f_{Y_i}(y_i)=\frac{1}{2\sqrt{2\pi\sigma^2 y_i}}\left\{\exp\left[-\frac{(\sqrt{y_i}-m_i)^2}{2\sigma^2}\right]+\exp\left[-\frac{(-\sqrt{y_i}-m_i)^2}{2\sigma^2}\right]\right\},\qquad y_i\geqslant 0 \tag{1.5-55}$$

经过简化，得到
$$f_{Y_i}(y_i)=\frac{1}{\sqrt{2\pi\sigma^2 y_i}}\exp\left[-\frac{y_i+{m_i}^2}{2\sigma^2}\right]\mathrm{ch}\frac{m_i\sqrt{y_i}}{\sigma^2},\qquad y_i\geqslant 0 \tag{1.5-56}$$

Y_i 的特征函数
$$\varPhi_{Y_i}(\omega)=\frac{1}{\sqrt{1-\mathrm{j}2\sigma^2\omega}}\exp\left[-\frac{m_i^2}{2\sigma^2}\right]\cdot\exp\left[\frac{m_i^2}{2\sigma^2}\cdot\frac{1}{1-\mathrm{j}2\sigma^2\omega}\right] \tag{1.5-57}$$

Y 的特征函数

$$\varPhi_Y(\omega)=\prod_{i=1}^{n}\varPhi_{Y_i}(\omega)=\frac{1}{(1-\mathrm{j}2\sigma^2\omega)^{n/2}}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}m_i^2\right]\cdot\exp\left[\frac{1}{2\sigma^2}\sum_{i=1}^{n}m_i^2\cdot\frac{1}{1-\mathrm{j}2\sigma^2\omega}\right] \tag{1.5-58}$$

通过傅里叶反变换求得 Y 的概率密度

$$f_Y(y)=\frac{1}{2\sigma^2}\left(\frac{y}{\lambda}\right)^{\frac{n-2}{4}}\exp\left[-\frac{y+\lambda}{2\sigma^2}\right]\mathrm{I}_{n/2-1}\left(\frac{\sqrt{\lambda y}}{\sigma^2}\right),\qquad y\geqslant 0 \tag{1.5-59}$$

式中，$\lambda=\sum_{i=1}^{n}m_i^2$ 称为非中心分布参量，$\mathrm{I}_{n/2-1}(x)$为第一类 $n/2-1$ 阶修正贝塞尔函数。

$$\mathrm{I}_n(x)=\sum_{m=0}^{\infty}\frac{(x/2)^{n+2m}}{m!\Gamma(n+m+1)} \tag{1.5-60}$$

非中心χ^2分布的概率密度曲线见图 1.5-9。

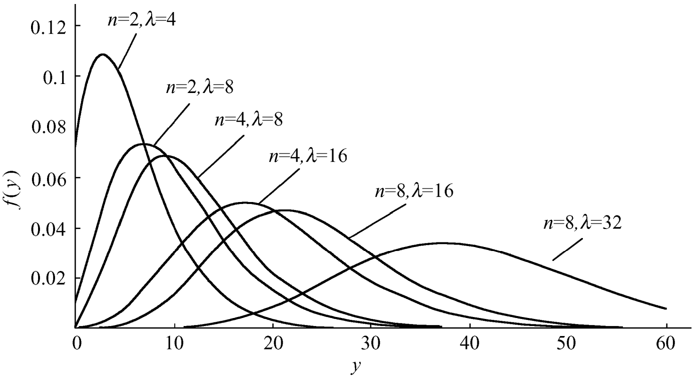


图 1.5-9　不同自由度的非中心χ^2分布（$\sigma^2=1$）的概率密度曲线

非中心χ^2分布 Y 的数学期望和方差分别为

$$\begin{cases} m_Y = n\sigma^2 + \lambda \\ \sigma_Y^2 = 2n\sigma^4 + 4\sigma^2\lambda \end{cases} \tag{1.5-61}$$

非中心χ^2分布也具有与中心χ^2分布类似的特点，两个互相独立的非中心χ^2分布的随机变量之和仍为非中心χ^2分布。若它们的自由度分别为 n_1 和 n_2，非中心分布参量分别为λ_1和λ_2，其和的自由度 $n = n_1 + n_2$，非中心分布参量$\lambda = \lambda_1 + \lambda_2$。

χ^2分布的例子将在第 4 章介绍。

1.5.4　瑞利分布和莱斯分布

在统计数学上，很少用到瑞利分布和莱斯分布，它们主要用于窄带随机信号。瑞利分布和莱斯分布与高斯分布有着一定的联系，确切地说，它们都是高斯分布通过一些变换得到的。另一方面，瑞利分布和莱斯分布又与χ^2分布和非中心χ^2分布联系密切，因为它们分别是由χ^2分布和非中心χ^2分布进行开方变换得来的。

1．瑞利分布

对于两个自由度的χ^2分布，当$Y = X_1^2 + X_2^2$时，$X_i(i=1,2)$是数学期望为零、方差为σ^2且互相独立的高斯变量，Y 服从指数分布

$$f_Y(y) = \frac{1}{2\sigma^2}\mathrm{e}^{-\frac{y}{2\sigma^2}},\ y \geqslant 0 \tag{1.5-62}$$

令

$$R = \sqrt{Y} = \sqrt{X_1^2 + X_2^2}$$

通过函数变换后，得到 R 的概率密度

$$f_R(r) = \frac{r}{\sigma^2}\mathrm{e}^{-\frac{r^2}{2\sigma^2}},\ r \geqslant 0 \tag{1.5-63}$$

R 就是瑞利分布。图 1.5-10 是当$\sigma = 1$时的瑞利分布概率密度曲线。

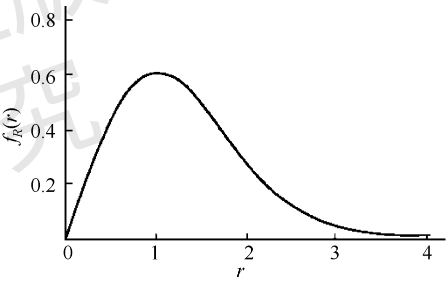


图 1.5-10　瑞利分布的概率密度曲线

在讨论窄带信号时，将看到窄带高斯过程的幅度即为瑞利分布。瑞利分布的各阶原点矩为

$$E[R^k] = (2\sigma^2)^{k/2}\Gamma\left(1 + \frac{k}{2}\right) \tag{1.5-64}$$

式中的伽马函数由式(1.5-49)计算。当 $k = 1$ 时，得数学期望

$$m_R = E[R] = (2\sigma^2)^{1/2}\Gamma\left(1 + \frac{1}{2}\right) = \sqrt{\frac{\pi}{2}}\sigma \tag{1.5-65}$$

可见瑞利分布的数学期望与原高斯变量的均方差成正比。反过来说，当需要估计高斯变量的方差(功率)时，往往通过估计瑞利分布的数学期望来得到，因为估计数学期望一般比估计方差要容易得多。瑞利分布的方差可由二阶原点矩和一阶原点矩获得

$$\sigma_R^2 = E[R^2] - (E[R])^2 = \left(2 - \frac{\pi}{2}\right)\sigma^2 \tag{1.5-66}$$

对 n 个自由度的χ^2分布，若令 $R = \sqrt{Y} = \sqrt{\sum_{i=1}^{n} X_i^2}$，则 R 为广义瑞利分布

$$f_R(r)=\frac{r^{n-1}}{2^{(n-2)/2}\sigma^n\Gamma(n/2)}\mathrm{e}^{-\frac{r^2}{2\sigma^2}},\qquad r\geqslant 0 \tag{1.5-67}$$

当 $n=2$ 时，上式简化为式(1.5−63)。

广义瑞利分布的各阶原点矩为

$$E[R^k]=(2\sigma^2)^{k/2}\frac{\Gamma([n+k]/2)}{\Gamma(n/2)} \tag{1.5-68}$$

当 $n=2$ 时，上式简化为式(1.5−64)。数学期望和方差仍可按上面的方法来求，这里给出数学期望

$$E[R]=(2\sigma^2)^{1/2}\frac{\Gamma(n/2+1/2)}{\Gamma(n/2)} \tag{1.5-69}$$

2．莱斯分布

当高斯变量 $X_i(i=1,2,\cdots,n)$ 的数学期望 m_i 不为零时，$Y=\sum_{i=1}^{n}X_i^2$ 是非中心χ^2 分布，而 $R=\sqrt{Y}$ 则是莱斯分布。当 $n=2$ 时

$$f_R(r)=\frac{r}{\sigma^2}\exp\left[-\frac{r^2+\lambda}{2\sigma^2}\right]\mathrm{I}_0\left(\frac{r\sqrt{\lambda}}{\sigma^2}\right),\qquad r\geqslant 0 \tag{1.5-70}$$

式中，$\mathrm{I}_0(x)$为零阶修正贝塞尔函数，可由下式计算

$$\mathrm{I}_0(x)=1+\sum_{n=1}^{\infty}\left[\frac{(x/2)^n}{n!}\right]^2 \tag{1.5-71}$$

作为式(1.5−70)的推广，对于任意的 n

$$f_R(r)=\frac{r^{n/2}}{\sigma^2\lambda^{n-2}}\exp\left[-\frac{r^2+\lambda}{2\sigma^2}\right]\mathrm{I}_{n/2-1}\left(\frac{r\sqrt{\lambda}}{\sigma^2}\right),\qquad r\geqslant 0 \tag{1.5-72}$$

式中，$\mathrm{I}_{n/2-1}(x)$为 $n/2-1$ 阶修正贝塞尔函数，由式(1.5−60)计算。

特别地，当 $n=2$ 时，上式简化为式(1.5−70)；进一步地，当$\lambda=0$时，式(1.5−72)简化为式(1.5−63)，因此瑞利分布是莱斯分布当$\lambda=0$时的特例。

图 1.5−11 所示为当$\sigma=1$时不同λ 的莱斯分布概率密度曲线。

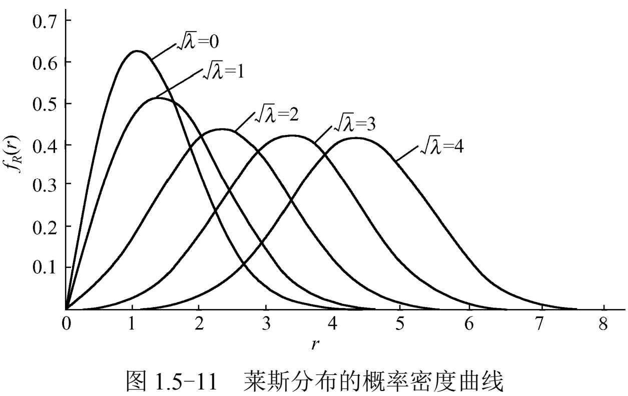


图 1.5−11　莱斯分布的概率密度曲线

一些基于高斯变量变换后的随机变量之间有着密切的关系，在一定的条件下，某个分布可转换为另外一个分布。或者说某个分布是另一个分布在某种条件下的特例。表 1.5−1 给出

了高斯分布和一些基于高斯变量变换后的随机变量之间的关系。

由表 1.5−1 可见，瑞利分布的概率密度由式(1.5−63)给出，它可由 $R=\sqrt{X_1^2+X_2^2}$ 变换而来。其中 X_1 和 X_2 是数学期望为零、方差为σ^2且互相独立的高斯变量。瑞利分布的数学期望为$(\pi/2)^{1/2}\sigma$，方差为$(2-\pi/2)\sigma^2$。此外，瑞利分布还可由 2 个自由度的 χ^2 分布做 $R=Y^{1/2}$ 的变换得到，或对指数分布做 $R=Y^{1/2}$ 的变换得到。

表 1.5−1　基于高斯变量变换的随机变量分布

	非中心 χ^2 分布	χ^2 分布	指数分布	莱斯分布	广义瑞利分布	瑞利分布
概率密度	式(1.5−59)	式(1.5−53)	式(1.5−62)	式(1.5−72)	式(1.5−67)	式(1.5−63)
数学期望	$n\sigma^2+\lambda$	$n\sigma^2$	$2\sigma^2$			$(\pi/2)^{1/2}\sigma$
方差	$2n\sigma^4+4\sigma^2$	$2n\sigma^4$	$4\sigma^4$			$(2-\pi/2)\ \sigma^2$
X_i 为高斯分布，互相独立，方差为σ^2	$Y=\sum_{i=1}^{n}X_i^2$	$Y=\sum_{i=1}^{n}X_i^2$	$Y=\sum_{i=1}^{2}X_i^2$	$R=\sqrt{\sum_{i=1}^{n}X_i^2}$	$R=\sqrt{\sum_{i=1}^{n}X_i^2}$	$R=\sqrt{\sum_{i=1}^{2}X_i^2}$
	X_i 均值不为零	X_i 均值为零	X_i 均值为零	X_i 均值不为零	X_i 均值为零	X_i 均值为零
非中心 χ^2 分布		X_i 均值为零 $\lambda=0$	X_i 均值为零 $\lambda=0, n=2$	$R=Y^{1/2}$	X_i 均值为零 $\lambda=0$	X_i 均值为零 $\lambda=0, n=2$
χ^2 分布			$n=2$		$R=Y^{1/2}$	$R=Y^{1/2}, n=2$
指数分布						$R=Y^{1/2}$
莱斯分布					X_i 均值为零	X_i 均值为零，$n=2$
广义瑞利分布						$n=2$

1.6* 离散随机变量的仿真与计算

计算技术的发展和计算机的普及使计算机仿真的应用越来越广泛。尤其是在实际的系统试验消耗人力物力太多或风险代价太大的情况下，就更能体现出仿真的价值所在。

不论是系统数学模型的建立，还是原始试验数据的产生，最基本的需求是产生一个所需分布的随机变量。比如在通信与信号处理领域中，电子设备的热噪声、通信信道中的加性噪声、图像中的灰度分布、飞行器高度表接收的地面杂波，甚至机械系统的振动噪声等都是遵循某一分布的随机信号。

在很多系统仿真的过程中，需要产生不同分布的随机变量，而随机变量的仿真需要大量的运算。在产生随机变量时，虽然运算量很大，但基本上是简单的重复。利用计算机、单片机或处理器可以很方便地产生不同分布的随机变量，各种分布的随机变量的基础是均匀分布的随机变量。有了均匀分布的随机变量，就可以用函数变换等方法得到其他分布的随机变量。

MATLAB 是由美国 MathWorks 公司开发的面向科学计算、高度集成的计算机语言。MATLAB 最显著的特点是向量化计算，另外功能强大、可视化计算、简捷方便是它流行的主要原因。MATLAB 的工具箱为不同需求提供了有特色的计算工具或函数，既有通常的数值计算，也有符号运算。Simulink 是 MATLAB 中的一个可视化仿真平台，可以完成连续时间系统和离散时间系统的系统级仿真，既适于线性系统，也适于非线性系统，是科学研究中常用仿真计算工具之一。

Python 是另一种跨平台的程序设计语言。作为一种优秀的开源编程语言，Python 提供了

高效的高级数据结构，还能简单有效地面向对象编程。Python 语言的可扩充性极强，本教材中主要使用 Python 的科学计算功能。Python 提供了 Numpy，SciPy 等众多科学计算包，Python 语言在数据统计分析、信号处理、深度学习等领域都有广泛的应用。

本书通过例题介绍了 MATLAB 和 Python 在随机信号分析中的应用，读者可以根据自己的喜好自行选择合适的语言进行学习。

1.6.1 均匀分布随机数的产生

一般来讲，由计算机或处理器产生随机变量时首先要产生均匀分布的随机变量。一个均匀分布的连续随机变量是由若干个样本组成的，而这些样本则是一个个随机的数据。由于 CPU 的算术单元（如累加器）及寄存器是由有限个二进制数组成的，它所完成的计算毕竟只能由有限字长的数字来表示，所以能计算出的样本数也是有限的。例如，当算术单元是 8 位时，它只能表示 256 个不同的数；而当算术单元是 16 位时，它可以表示 65536 个不同的数。与真正均匀分布的连续随机变量相比，这些样本并不是连续地占据某个取值区间的。若把这些样本看成数轴上的随机点，它实质上是图 1.1−1（a）表示的离散随机变量。但当运算器和寄存器字长足够长时，它所能表示或计算的数就能比较密集地充满某一区间，这样可把它当作连续随机变量的一种近似。尽管如此，在一个区间内，计算机能计算出的随机数毕竟是有限的。为了区别真正意义下的随机数，这样计算产生的随机数常称为伪随机数。

产生伪随机数的一种实用方法是同余法，它利用同余运算递推产生伪随机数序列。最简单的方法是加同余法

$$y_{n+1} = y_n + c \quad (\mathrm{mod}\ M) \tag{1.6−1}$$

$$x_{n+1} = y_{n+1} / M \tag{1.6−2}$$

式中，mod M 为取模运算。利用上面的递推公式，可产生[0,1]上均匀分布的随机数。为了保证产生的伪随机数能在[0,1]内均匀分布，需要 M 为正整数，此外常数 c 和初值 y_0 亦为正整数。加同余法虽然简单，但产生的伪随机数效果不好。另一种同余法为乘同余法，它需要两次乘法才能产生一个在[0,1]上均匀分布的随机数

$$y_{n+1} = a y_n \quad (\mathrm{mod}\ M) \tag{1.6−3}$$

$$x_{n+1} = y_{n+1} / M \tag{1.6−4}$$

式中，a 为正整数。用加法和乘法完成递推运算的称为混合同余法，即

$$y_{n+1} = a y_n + c \quad (\mathrm{mod}\ M) \tag{1.6−5}$$

$$x_{n+1} = y_{n+1} / M \tag{1.6−6}$$

用混合同余法产生的伪随机数具有较好的特性，一些程序库中都有成熟的程序供选择。

实际上，由以上各式产生的随机数到了一定数目后，会出现周而复始的现象，就是说产生的随机数存在周期。为了使产生的伪随机数有较大的周期，更接近真正的随机数，M 越大越好。但 M 的取值也不是无限的，一般选择 $M=2^b$，b 是所选计算设备中运算器的字长。M 选择为 2 的幂也有利于减小运算量，因为除以 M 的运算可用移位代替。

周期的大小除了与 M 有关，还与其他几个参数有关。很多人研究过随机数的周期问题，由于要涉及一些数论方面的知识，这里只给出混合同余法达到最大周期的条件。在式（1.6−5）中，若满足：①c 与 M 互素；②对 M 的任意素因子 p，$a \equiv 1(\mathrm{mod}\ \ p)$；③如果 4 是 M 的一个

因子，$a \equiv 1 (\text{mod } 4)$，则产生的随机数可获得的最大周期是 M。而乘同余法和加同余法却不能获得最大周期。能达到最大周期的混合同余法递推公式为

$$y_{n+1} = (4a+1)y_n + (2b+1) \quad (\text{mod } M) \tag{1.6-7}$$

$$x_{n+1} = y_{n+1} / M \tag{1.6-8}$$

式中，a 和 b 为任意正整数。

常用的计算语言如 MATLAB 和 Python 都有产生均匀分布随机数的函数可以调用，只是用各种编程语言对应的函数产生的均匀分布随机数的范围不同，有的函数可能还需要提供种子或初始化。

MATLAB 提供的函数 rand( )可以产生一个在[0,1]区间分布的随机数，rand(2,4)则可以产生一个在[0,1]区间分布的随机数矩阵，矩阵为 2 行 4 列。MATLAB 提供的另一个产生随机数的函数是 random（'unif ',a,b,N,M），unif 表示均匀分布，a 和 b 是均匀分布区间的上下界，N 和 M 分别是矩阵的行和列。

Python 则在 numpy 库中集成了基础运算的众多函数，随机数产生函数也在其中。例如 numpy.random.rand()可以产生[0, 1)之间的随机数，numpy.random.randn()可以产生标准正态分布的随机数，numpy.random.randint()返回随机整数。

【例 1.6−1】 编程产生 1024 个在[2,5]区间上均匀分布随机数的程序。

解：这个程序可以有多种实现方法，如分别用式（1.6−1）～式（1.6−6）递推产生。本例用分别使用 MATLAB 和 Python 函数来产生满足条件的随机数。

- MATLAB 方法

（1）利用 rand( )函数：

```
for n=1:1024                %循环
    y=rand();               %产生 1 个在[0,1]  区间均匀分布的随机数
    x(n)=y*(5−2)+2;         %将在[0,1]内均匀分布的随机数变换到[2,5]区间
end                         %循环结束
plot(x);                    %将 1024 个在[2,5]区间上均匀分布随机数画成一条曲线，便于观察
```

这个程序编写虽然有点麻烦，但是可以很容易转换成其他语言的程序。

（2）利用 rand(M,N)函数：

```
y =rand(1,1024);            %产生 1024 个在[0,1]区间均匀分布的随机数
x=(5−2)*y+2;                %将在[0,1]内均匀分布的随机数变换到[2,5]区间
plot(x);
```

这个程序利用了 MATLAB 向量化计算的优点，节省了程序中的循环

（3）利用 random('unif ', a, b, N, M)函数：

```
x=random('unif',2,5,1,1024);   %产生 1024 个在[2,5]区间均匀分布的随机数
plot(x);
```

这是最简单的方法，充分体现了 MATLAB 的优势，一个语句完成了分布、分布区间、矩阵大小的设定。

- Python 方法

（1）使用 numpy 函数包中的 random.rand，代码如下：

```
import numpy as np
```

```
import matplotlib.pyplot as plt
y = np.random.rand(1024)            #产生 1024 个在[0,1)上均匀分布的随机数
x = (5−2)*y+2   #将在[0,1)内均匀分布的随机数变换到[2,5)区间
plt.plot(x)
plt.show()
```

这个程序与 MATLAB 中的程序（2）相似，通过 random.rand 生成[0,1)上均匀分布的随机数，再通过函数变换方法生成[2,5)区间上的随机数。

（2）使用 numpy 函数包中的 random.uniform，即：

```
import numpy as np
import matplotlib.pyplot as plt
y = np.random.uniform(2, 5, 1024)        #产生 1024 个在[2,5)上均匀分布的随机数
plt.plot(x)
plt.show()
```

这个程序与 MATLAB 中的程序（3）相似，Python 中 numpy 函数包中的 random.uniform 可以直接设置均匀分布的范围和样本数目来产生均匀分布的随机数。

在实际应用中，产生随机数之后，必须对它的统计特性做严格的检验。一般来讲，统计特性的检验包括参数检验、均匀性检验和独立性检验。事实上，如果在二阶矩范围内讨论随机信号，那么参数检验只对产生的随机数一、二阶矩进行检验。此外，参数检验还包括最小值、最大值和周期等。参数检验、均匀性检验和独立性检验的方法请参考有关书籍。

1.6.2　随机变量的仿真

根据随机变量函数变换的原理，如果能将两个分布之间的函数关系用显式表达，那么就可以通过对一种分布的随机变量进行变换得到另一种分布的随机变量。

若 X 是分布函数为 $F_X(x)$的随机变量，且 $F_X(x)$为严格单调升函数，令 $Y=F_X(X)$，则 Y 是在[0,1]上分布的随机变量。若 Y 是在[0,1]上均匀分布的随机变量，那么

$$X = F_X^{-1}(Y) \tag{1.6−9}$$

即是分布函数为 $F_X(x)$ 的随机变量。式中 $F_X^{-1}(\cdot)$ 为 $F_X(\cdot)$ 的反函数。这样，欲求某个分布的随机变量，先产生在[0,1]区间上均匀分布的随机数，再经式(1.6−9)的变换，便可求得所需分布的随机数。

【例 1.6−2】 假定已有均匀分布的随机数，给出产生指数分布随机数的计算公式。

解：指数分布的概率密度为

$$f_X(x) = a\mathrm{e}^{-ax}, \quad x \geqslant 0$$

根据概率密度可求出概率分布函数

$$F_X(x) = \int_0^x a\mathrm{e}^{-a\tau}\mathrm{d}\tau = 1-\mathrm{e}^{-ax}, \quad x \geqslant 0$$

那么

$$Y = F_X(X) = 1-\mathrm{e}^{-aX}$$

就是在[0,1]区间上均匀分布的随机数。而

$$X = F_X^{-1}(Y) = -\frac{\ln(1-Y)}{a}$$

即为指数分布。考虑到 $1-Y$ 也是在[0,1]上均匀分布的随机数，上式可改写为

$$X = F_X^{-1}(Y) = -\frac{\ln Y}{a}$$

当分布函数的反函数比较复杂时，这种函数变换的过程也比较复杂。实际上，产生一个非均匀分布的随机数序列有很多方法，上面讨论的变换法仅是其中的一种。

1.6.3 高斯分布随机数的仿真

由于高斯变量的重要性，有必要重点讨论一下高斯随机数的产生。广泛应用的有两种产生高斯随机数的方法，一种是变换法，一种是近似法。

如果 X_1 和 X_2 是两个互相独立的均匀分布随机数，那么下式给出的 Y_1 和 Y_2

$$\begin{cases} Y_1 = \sigma\sqrt{-2\ln X_1}\cos(2\pi X_2) + m \\ Y_2 = \sigma\sqrt{-2\ln X_1}\sin(2\pi X_2) + m \end{cases} \tag{1.6-10}$$

便是数学期望为 m，方差为 σ^2 的高斯分布随机数，且互相独立，这就是变换法。下面证明一种简单情况，即令数学期望 $m=0$，方差 $\sigma^2=1$。利用上式求反函数

$$\begin{cases} X_1 = \exp\left[-\dfrac{Y_1^2 + Y_2^2}{2}\right] \\ X_2 = \dfrac{1}{2\pi}\arctan\left(\dfrac{Y_2}{Y_1}\right) \end{cases} \tag{1.6-11}$$

雅可比行列式

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} \exp\left[-\dfrac{y_1^2 + y_2^2}{2}\right](-y_1) & \exp\left[-\dfrac{y_1^2 + y_2^2}{2}\right](-y_2) \\ \dfrac{1}{2\pi}\cdot\dfrac{(y_2/y_1^2)}{1+(y_2/y_1)^2} & \dfrac{1}{2\pi}\cdot\dfrac{1/y_1}{1+(y_2/y_1)^2} \end{vmatrix}$$

$$= -\frac{1}{2\pi}\exp\left[-\frac{y_1^2 + y_2^2}{2}\right] \tag{1.6-12}$$

既然 X_1 和 X_2 是两个互相独立的均匀分布随机数，那么 Y_1 和 Y_2 的二维联合概率密度应满足

$$f_Y(y_1, y_2) = \frac{1}{2\pi}\exp\left[-\frac{y_1^2 + y_2^2}{2}\right] = \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{y_1^2}{2}\right]\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{y_2^2}{2}\right] = f_{Y_1}(y_1)f_{Y_2}(y_2) \tag{1.6-13}$$

显然，Y_1 和 Y_2 是高斯分布的，它们的数学期望为零，方差为 1，且互相独立。

【例 1.6-3】 编写一个函数，用式(1.6-10)的方法产生两个互相独立的高斯变量，每个高斯变量有 N 个样本(随机数)，数学期望为 m、方差为 σ^2。

- 使用 MATLAB，实现上述编程要求：

```
function    [xr, xi]= GaussRandomNumbers_1(N, Mu, Sigma)
        a=sqrt(−2.0*log(rand(1, N)));
        b=2*pi*rand(1, N);
        xr= Sigma*a.*cos(b)+Mu;
        xi= Sigma*a.*sin(b)+Mu;
```

函数 GaussRandomNumbers_1( )产生的 xr 和 xi 是两组互相独立的高斯分布随机数，作为函数的返回值。形参 N、Mu 和 Sigma 分别为每组随机数的个数、数学期望和标准差。在这个函数中，pi 为圆周率，是 MATLAB 定义的常数。调用这个函数的主程序为

```
[x1, x2] = GaussRandomNumbers_1(1024, 0,1);
plot(x1,'k');                 % 'k'表示曲线为黑色
hold on;                      % 保持图形，为了将几个曲线画在同一个图中
plot(x2,'r');                 % 'r'表示曲线为红色
```

返回的两组互相独立的高斯分布随机数分别放在数组 x1 和 x2 中，均值为零、方差为 1，并将两个高斯变量对应的随机数画在同一个图中，k 代表曲线为黑色，r 代表曲线为红色。

- 使用 Python，实现上述编程要求：

```
def GaussianRandomNumbers_1(N, Mu, Sigma):
a = np.sqrt(-2.0 * np.log(np.random.rand(N)))
b = 2 * np.pi * np.random.randn(N)
xr = Sigma * np.multiply(a, np.cos(b)) + Mu
xi = Sigma * np.multiply(a, np.sin(b)) + Mu
return xr, xi
```

Python 中定义函数的关键字是 def。这个函数所完成的功能与 MATLAB 函数 GaussRandomNumbers_1( )所完成的功能一致。调用这个函数的 Python 主程序为：

```
x, y = GaussianRandomNumbers_1(1024, 0, 1.0)
fig = plt.figure(5)
plt.plot(x)   # 绘制样本图
plt.plot(y)   # 绘制样本图
plt.show()    # 显示图片
```

另外一种产生高斯随机数的方法是近似法。在学习中心极限定理时，曾提到 n 个独立同分布的随机变量 X_i (i=1,2,⋯,n)，当 n 足够大时，之和的线性变换依分布收敛于高斯分布。由于近似法避免了开方和三角函数运算，计算量大大降低。当精度要求不太高时，近似法还是具有很大应用价值的。

【例 1.6-4】 用近似法编写一个函数，产生服从 $N(m,\sigma^2)$ 的高斯分布随机数。

解：本例中使用独立同分布的均匀分布样本来近似高斯分布的样本。在[0,1]区间上，均匀分布的随机变量 X 的数学期望为 $E[X]=1/2$，方差为 $D[X]=1/12$。

根据中心极限定理，当 n 足够大时，随机变量

$$Y=\frac{\sum_{i=1}^{n}(X_i-1/2)}{\sqrt{n/12}}$$

随机变量 Y 依分布收敛于数学期望为零、方差为 1 的高斯分布。为减小计算量，取 $n=12$ 完成上述近似，此时

$$Y=\sum_{i=1}^{12}(X_i-1/2)=\sum_{i=1}^{12}X_i-6$$

最后做变换 $Z=\sigma Y+m$ 即为所求。使用 Python 实现上述功能，函数如下：

```
import numpy as np
```

```
import matplotlib.pyplot as plt
def GaussRandomNumbers(N, Mu, Sigma):
    x = np.random.rand(12, N)            # 产生 12 组在[0,1)上均匀分布的随机数，每组有 N 个样本
    y = np.sum(x, 0)-6                   # 将每组对应的样本相加后减 6，近似标准正态分布
    return Sigma*y+Mu
z = GaussRandomNumbers(1024, 0.5, 2)     #产生均值为 0.5，方差为 2 的高斯分布随机数
fig = plt.figure(1)
plt.plot(z)                              # 绘制样本图
plt.show()                               # 显示图片
```

高斯随机变量是最常用的随机变量之一。无论是 MATLAB 还是 Python 都集成了高效的高斯随机变量生成方法。例如，MATLAB 中可以直接通过调用命令

x=random('Normal',Mu, Sigma,M,N);

来生成 M（行）×N（列）个独立同分布的高斯随机变量样本，每行样本独立同分布。函数中 Normal 为高斯（正态）分布的标志，Mu 和 Sigma 代表数学期望和均方差。

对于高维高斯分布（其联合概率密度函数见式（1.5−41））随机变量的样本，MATLAB 提供了函数

R = mvnrnd(Mu,Sigma,N)

其中 Mu 是 D 维均值向量，Sigma 是 D×D 的协方差矩阵，N 是样本的个数。函数返回值 R 是 N×D 的矩阵，其中每一行是一个 D 维高斯随机变量的样本。

Python 中可以使用函数 numpy.random.normal 和 numpy.random.multivariate_normal 来产生高斯分布的样本和高维高斯分布的样本。两个函数的功能与 MATLAB 中的 random(‘normal’)和 mvnrnd 接近。

有了高斯随机变量的仿真方法，就可以利用表 1.5−1 构成与高斯变量有关的其他分布随机变量，如瑞利分布、指数分布和χ^2分布随机变量。在仿真随机变量时，一般用直方图近似表示随机变量的分布。画直方图时先将随机变量的取值区间分为 k 个相等的子区间，然后统计随机数落在所有子区间的个数，将 k 个子区间落入随机数的个数画成柱图就是直方图。MATLAB 提供了画直方图的函数 hist（X,Y），其中 X 是随机数，Y 是子区间的坐标。

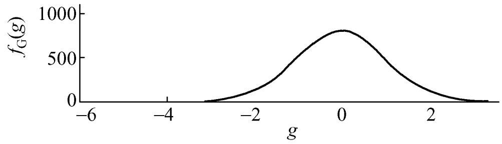
【例 1.6−5】 用 $N(0,1)$ 高斯分布随机数分别仿真瑞利分布随机变量和 4 个自由度的χ^2分布随机变量，并画出直方图。

解： 首先获得数学期望为零、方差为 1 的高斯分布随机数。利用表 1.5−1，根据瑞利分布及χ^2分布与高斯分布的关系就可得到具有所求分布的随机数。假定每个随机变量包括 N 个随机数，分别用 Gi（i = 1,2,3,4）、R 和 X2 表示高斯分布、瑞利分布和χ^2分布的随机数，其程序如下。

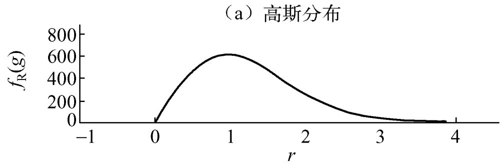
```
N=20000;
g = −5:0.1:5;
G1=random('Normal',0,1,1,N);            %产生 N 个高斯分布随机数，均值为零，方差为 1
G2=random('Normal',0,1,1,N);
G3=random('Normal',0,1,1,N);
G4=random('Normal',0,1,1,N);
R=sqrt(G1.*G1+G2.*G2);                   %由表 1.5−1 高斯与瑞利分布交叉处的公式
X2= G1.*G1+G2.*G2+G3.*G3+G4.*G4;         %由表 1.5−1 高斯与χ2 分布交叉处的公式，n = 4
subplot(311); hist(G1,g);                %在−5～5 范围内画高斯变量的直方图
```

```
subplot(312); hist(R,0:0.05:5);          %在 0～5 范围内画瑞利随机变量的直方图
subplot(313); hist(X2,0:0.2:24);         %在 0～24 范围内画χ² 分布随机变量的直方图
```

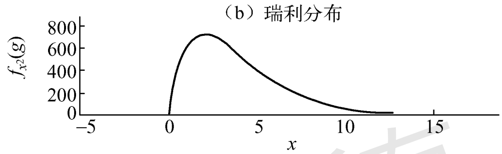
图 1.6-1 示出了由程序得到的直方图，为了清晰地看到每个直方图的细节，各直方图的取值区间不同，如瑞利随机变量的取值区间应该是 0～∞，程序中只求出了 0～5 范围内的直方图，并把 0～5 区间分为 100 个小区间。图中的包络线是理论上的概率密度曲线，当随机数的数目足够大时，由一个优秀设计的程序得到的直方图应该接近理论的概率密度曲线。



（a）高斯分布



（b）瑞利分布



（c）4 个自由度χ^2分布

图 1.6-1 随机变量的直方图

1.6.4 随机变量数字特征的计算

产生的随机数可以作为一个随机变量，也可以视为下一章要讨论的随机过程中的一个样本函数。不论是随机变量还是随机过程的样本函数，都会遇到求其数字特征的情况。在图像处理时，也时常需要计算图像灰度直方图的数学期望、方差、峰态和偏态系数等。

事实上，在很多情况下无法得到或不能利用随机变量的全部样本，只能利用一部分样本(子样)来获得随机变量数字特征的估计值。这时，子样的个数 N 就决定了估计的精度。当 N 增大时，估计值将依概率收敛于被估计的参数。

在实际计算时，根据对计算速度或精度要求的不同，可选择不同的算法。

1．均值的计算

设随机数序列$\{x_1, x_2, \cdots, x_N\}$，一种计算均值的方法是直接计算下式

$$m = \frac{1}{N}\sum_{n=1}^{N} x_n \tag{1.6-14}$$

式中，x_n 为随机数序列中的第 n 个随机数。

另一种方法是利用递推算法，第 n 次迭代的均值即前 n 个随机数的均值为

$$m_n = \frac{n-1}{n} m_{n-1} + \frac{1}{n} x_n = m_{n-1} + \frac{1}{n}(x_n - m_{n-1}) \tag{1.6-15}$$

迭代结束后，便得到随机数序列的均值

$$m = m_N \tag{1.6-16}$$

递推算法的优点是可以实时计算均值，这种方法常用在实时获取数据的场合。

当数据量较大时，为防止计算误差的积累，也可采用

$$m = m_1 + \frac{1}{N}\sum_{n=1}^{N}(x_n - m_1) \tag{1.6-17}$$

式中，m_1 是取一小部分随机数计算的均值。

2．方差的计算

计算方差也分为直接法和递推法。仿照均值的做法

$$\sigma^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n - m)^2 \tag{1.6-18}$$

$$\sigma^2 = \frac{1}{N}\sum_{n=1}^{N}x_n^2 - m^2 \tag{1.6-19}$$

当利用有限字长运算时，式(1.6–18)的运算误差比式(1.6–19)小，但利用式(1.6–19)可节省运算次数。

方差的递推算法需要同时递推均值和方差

$$m_n = m_{n-1} + \frac{1}{n}(x_n - m_{n-1})$$

$$\sigma_n^2 = \frac{n-1}{n}\left[\sigma_{n-1}^2 + \frac{1}{n}(x_n - m_{n-1})^2\right] \tag{1.6-20}$$

迭代结束后，得到随机数序列的方差为

$$\sigma^2 = \sigma_N^2 \tag{1.6-21}$$

其他矩函数也可用类似的方法得到。

图 1.6–2 展示了一组均值为零、方差为 1 的高斯分布随机变量的独立样本。这组随机数共有 300 个样本，用式(1.6–14)和式(1.6–19)计算的均值和方差分别为–0.056 和 1.055。根据式(1.6–15)和式(1.6–20)对这组随机样本所递推计算得到的均值 m_n 和方差 σ_n^2 同样示于图 1.6–2 中，其中 m_n 和 σ_n^2 是使用前 n 个样本得到的结果。当参与递推计算的随机数个数较少时，递推计算的均值 m_n 和方差 σ_n^2 与用式(1.6–14)和式(1.6–19)计算的均值 m 和方差 σ^2 有一定的差距。随着随机数个数的增加，均值和方差分别稳定在均值 m 和方差 σ^2 附近。显然，最终递推计算的均值 m_N 与用式(1.6–14)计算的结果相同，但递推的方差 σ_N^2 =1.052，存在偏差。

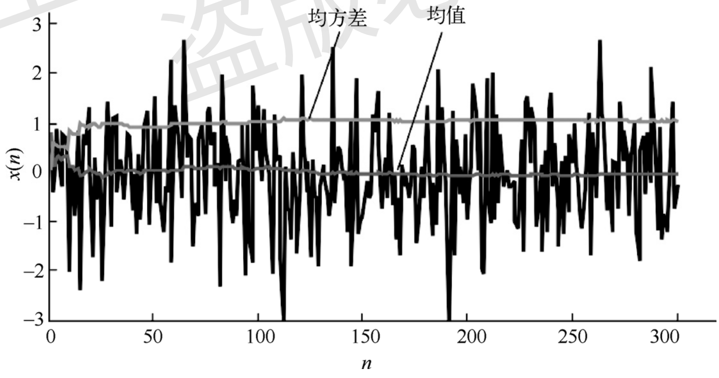


图 1.6–2　高斯分布随机数与递推的均值和均方差

MATLAB 提供了计算随机变量数字特征的函数。如果 y1 为一维数组，即 y1 为 1×*N* 或 *N*×1 矩阵，y2 为 $M \times N$ 矩阵，均值 mean( )和方差 var( )函数的使用方法如下。

```
m=mean(y1)                % y1 的均值
d=var(y1)                 % y1 的方差
m1=mean(y2)               % 按列求均值，m1 是一个 1×N 的数组
m12=mean(y2,2)            % 按行求均值，m12 是一个 M×1 的数组
d1=var(y2)                % 按列求方差，d1 是一个 1×N 的数组
m2=mean(mean(y2))         % 矩阵的均值
d2=var(var(y2))           % 矩阵的方差
```

同样地，Python 中也提供了计算均值和方差的函数。读者可以使用 np.mean( )和 np.var( )来计算向量的均值和方差。

习题一

1.1　设连续随机变量 x 的概率分布函数为

$$F(x)=\begin{cases}0, & x<0\\ 0.5+A\sin\left[\dfrac{\pi}{2}(x-1)\right], & 0\leqslant x<2\\ 1, & x\geqslant 2\end{cases}$$

求：（1）系数 A；（2）X 取值在 $(0.5, 1)$ 内的概率 $p(0.5<x<1)$。

1.2　试确定下列各式是否为连续随机变量的概率分布函数，如果是概率分布函数，求其概率密度。

（1）$F(x)=\begin{cases}1-\mathrm{e}^{-x/2}, & x\geqslant 0\\ 0, & x<0\end{cases}$　　（2）$F(x)=\begin{cases}0, & x<0\\ Ax^2, & 0\leqslant x<1\\ 1, & x\geqslant 1\end{cases}$

（3）$F(x)=\dfrac{x}{a}[u(x)-u(x-a)],\quad a>0$　　（4）$F(x)=\dfrac{x}{a}u(x)-\dfrac{a-x}{a}u(x-a),\quad a>0$

1.3　离散随机变量 X 由 0, 1, 2, 3 四个样本组成，相当于四元通信中的四个电平，四个样本的取值概率顺序为 1/2, 1/4, 1/8, 1/8。求随机变量的数学期望和方差。

1.4　随机变量 X 在 $[\alpha,\beta]$ 上均匀分布，求它的数学期望和方差。

1.5　设随机变量 X 的概率密度 $f_X(x)=\begin{cases}1, & 0\leqslant x\leqslant 1\\ 0, & 其他\end{cases}$，求 $Y=5X+1$ 的概率密度。

1.6　设随机变量 $X_1, X_2,\cdots, X_n$ 在 $[a,b]$ 上均匀分布，且互相独立。若 $Y=\sum\limits_{i=1}^{n}X_i$，求 $n=2$ 时，随机变量 Y 的概率密度。

1.7　设随机变量 X 的数学期望和方差分别为 m 和 σ^2，求随机变量 $Y=-3X-2$ 的数学期望、方差及 X 和 Y 的相关矩。

1.8　已知二维随机变量 (X,Y) 的二阶混合原点矩 m_{11} 及数学期望 m_X 和 m_Y，求随机变量 X,Y 的二阶混合中心矩。

1.9*　随机变量 X 和 Y 分别在 $[0,a]$ 和 $[0,\pi/2]$ 上均匀分布，且互相独立。对于 $b<a$，证明：

$$P(X<b\cos Y)=\frac{2b}{\pi a}$$

1.10*　已知二维随机变量 (X_1, X_2) 的联合概率密度为 $f_{X_1X_2}(x_1,x_2)$，随机变量 (X_1, X_2) 与随机变量 (Y_1, Y_2) 的关系由下式唯一确定

$$\begin{cases}X_1=a_1Y_1+b_1Y_2\\ X_2=c_1Y_1+d_1Y_2\end{cases},\qquad \begin{cases}Y_1=aX_1+bX_2\\ Y_2=cX_1+dX_2\end{cases}$$

证明 (Y_1, Y_2) 的联合概率密度为

$$f_{Y_1Y_2}(y_1,y_2)=\frac{1}{|ad-bc|}f_{X_1X_2}(a_1y_1+b_1y_2,c_1y_1+d_1y_2)$$

式中，$ad-bc\neq 0$。

1.11　随机变量 X, Y 的联合概率密度为

$$f_{XY}(x,y)=A\sin(x+y),\quad 0\leqslant x,y\leqslant \pi/2$$

求：（1）系数 A；（2）数学期望 M_X 和 M_Y；（3）方差 σ_X^2 和 σ_Y^2；（4）相关矩 r_{XY} 及相关系数 r_{XY}。

1.12　若 X 为在[0, 1]区间上均匀分布的随机变量，求 X 的特征函数及 n 阶原点矩。

1.13　已知随机变量 X 的概率密度 $f_X(x)=2\mathrm{e}^{-ax}, x\geqslant 0$，求随机变量 X 的特征函数。

1.14　已知随机变量 X 服从柯西分布 $f(x)=\frac{1}{\pi}\cdot\frac{\alpha}{\alpha^2+x^2}$，求它的特征函数。

1.15　求概率密度 $f(x)=\frac{1}{2}\mathrm{e}^{-|x|}$ 的随机变量 X 的特征函数。

1.16　已知互相独立随机变量 $X_1,X_2,\cdots,X_n$ 的特征函数，求 $X_1,X_2,\cdots,X_n$ 线性组合 $Y=\sum_{i=1}^{n}a_iX_i+c$ 的特征函数。其中 a_i 和 c 是常数。

1.17　已知高斯随机变量 X 的数学期望为零、方差为 1，求 $Y=aX^2\ (a>0)$ 的概率密度。

1.18　已知 X_1,X_2,X_3 是数学期望为零、方差为 1 且互相独立的高斯变量，用特征函数法求 $E[X_1X_2X_3]$。

1.19　如果随机变量 X 服从[0,1]区间的均匀分布，随机变量 Y 的概率密度为

$$f_y(y)=\begin{cases} y, & 0\leqslant y\leqslant 1 \\ 2-y, & 1\leqslant y\leqslant 2 \\ 0, & \text{其他} \end{cases}$$

X 与 Y 互相独立。求 X 与 Y 之和的概率密度。

本章习题解答请扫二维码。