第二章 随机变量及其分布

§1 随机变量

在第一章我们看到一些随机试验，它们的结果可以用数来表示．此时样本空间 S 的元素是一个数，如 S_3，S_5；但有些则不然，如 S_1，S_2．当样本空间 S 的元素不是一个数时，人们对于 S 就难以描述和研究．现在来讨论如何引入一个法则，将随机试验的每一个结果，即将 S 的每个元素 e 与实数 x 对应起来．从而引入了随机变量的概念．我们从例题开始讨论．

例 1 在第一章§4 例 1 中，将一枚硬币抛掷三次，观察出现正面和反面的情况，样本空间是

$$S=\{HHH,HHT,HTH,THH,HTT,THT,TTH,TTT\}.$$

以 X 记三次投掷得到正面 H 的总数，那么，对于样本空间 $S=\{e\}$[①]中的每一个样本点 e，X 都有一个数与之对应．X 是定义在样本空间 S 上的一个实值单值函数．它的定义域是样本空间 S，值域是实数集合$\{0,1,2,3\}$．使用函数记号可将 X 写成

$$X=X(e)=\begin{cases}3, & e=HHH,\\ 2, & e=HHT,HTH,THH,\\ 1, & e=HTT,THT,TTH,\\ 0, & e=TTT.\end{cases}$$ □

例 2 在一袋中装有编号分别为 1，2，3 的 3 只球，在袋中任取一只球，放回，再任取一只球，记录它们的号码，试验的样本空间为 $S=\{e\}=\{(i,j)\mid i,j=1,2,3\}$，$i,j$ 分别为第 1，第 2 次取到的球的号码．以 X 记两球号码之和．我们看到，对于试验的每一个结果 $e=(i,j)\in S$，X 都有一个指定的值 $i+j$ 与之对应．（如图 2－1）．X 是定义在样本空间 S 上的单值实值函数．它的定义域是样本空间 S．值域是实数集合$\{2,3,4,5,$

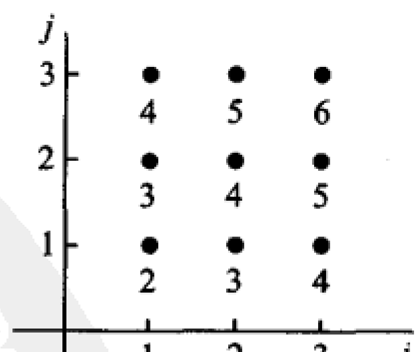

图 2－1

---

① 我们用 e 代表样本空间的元素，而将样本空间记成$\{e\}$．

6}. X 可写成

$$X=X(e)=X((i,j))=i+j,\ i,\ j=1,2,3. \qquad \square$$

一般有以下的定义.

定义 设随机试验的样本空间为 $S=\{e\}$. $X=X(e)$ 是定义在样本空间 S 上的实值单值函数. 称 $X=X(e)$ 为随机变量①.

图 2-2 画出了样本点 e 与实数 $X=X(e)$ 对应的示意图.

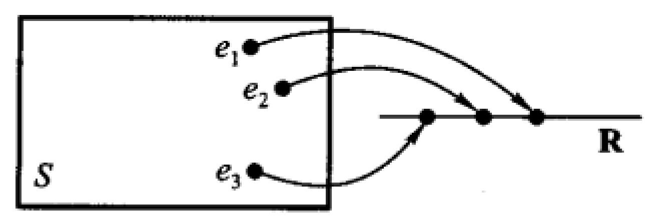

图 2-2

有许多随机试验,它们的结果本身是一个数,即样本点 e 本身是一个数. 我们令 $X=X(e)=e$,那么 X 就是一个随机变量. 例如,用 Y 记某车间一天的缺勤人数,以 W 记某地区第一季度的降雨量,以 Z 记某工厂一天的耗电量,以 N 记某医院某一天的挂号人数. 那么 Y,W,Z,N 都是随机变量.

本书中,我们一般以大写的字母如 $X,Y,Z,W,\cdots$ 表示随机变量,而以小写字母 $x,y,z,w,\cdots$ 表示实数.

随机变量的取值随试验的结果而定,而试验的各个结果出现有一定的概率,因而随机变量的取值有一定的概率. 例如,在例 1 中 X 取值为 2,记成 $\{X=2\}$,对应于样本点的集合 $A=\{HHT,HTH,THH\}$,这是一个事件,当且仅当事件 A 发生时有 $\{X=2\}$. 我们称概率 $P(A)=P\{HHT,HTH,THH\}$ 为 $\{X=2\}$ 的概率,即 $P\{X=2\}=P(A)=3/8$. 以后,还将事件 $A=\{HHT,HTH,THH\}$ 说成是事件 $\{X=2\}$. 类似地有

$$P\{X\leqslant 1\}=P\{HTT,THT,TTH,TTT\}=\frac{1}{2}.$$

一般,若 L 是一个实数集合,将 X 在 L 上取值写成 $\{X\in L\}$. 它表示事件 $B=\{e\mid X(e)\in L\}$,即 B 是由 S 中使得 $X(e)\in L$ 的所有样本点 e 所组成的事件,此时有

$$P\{X\in L\}=P(B)=P\{e\mid X(e)\in L\}.$$

随机变量的取值随试验的结果而定,在试验之前不能预知它取什么值,且它的取值有一定的概率. 这些性质显示了随机变量与普通函数有着本质的差异.

随机变量的引入,使我们能用随机变量来描述各种随机现象,并能利用数学分析的方法对随机试验的结果进行深入广泛的研究和讨论.

---

① 严格地说"对于任意实数 x,集合 $\{e\mid X(e)\leqslant x\}$(即:使得 $X(e)\leqslant x$ 的所有样本点 e 所组成的集合)有确定的概率"这一要求应包括在随机变量的定义之中,一般来说,不满足这一条件的情况,在实际应用中是很少遇到的. 因此,我们在定义中未提及这一要求.

§2　离散型随机变量及其分布律

有些随机变量，它全部可能取到的值是有限个或可列无限多个，这种随机变量称为**离散型随机变量**．例如§1 例 1 中的随机变量 X，它只可能取 0,1,2,3 四个值，它是一个离散型随机变量．又如某城市的 120 急救电话台一昼夜收到的呼唤次数也是离散型随机变量．若以 T 记某元件的寿命，它所可能取的值充满一个区间，是无法按一定次序一一列举出来的，因而它是一个非离散型的随机变量．本节只讨论离散型随机变量．

容易知道，要掌握一个离散型随机变量 X 的统计规律，必须且只需知道 X 的所有可能取值以及取每一个可能值的概率．

设离散型随机变量 X 所有可能取的值为 $x_k(k=1,2,\cdots)$，X 取各个可能值的概率，即事件 $\{X=x_k\}$ 的概率，为

$$P\{X = x_k\} = p_k,\ k = 1,2,\cdots. \tag{2.1}$$

由概率的定义，p_k 满足如下两个条件：

$$1^\circ\quad p_k \geqslant 0,\ k=1,2,\cdots; \tag{2.2}$$

$$2^\circ\quad \sum_{k=1}^{\infty} p_k = 1. \tag{2.3}$$

2° 是由于 $\{X = x_1\} \cup \{X = x_2\} \cup \cdots$ 是必然事件，且 $\{X = x_j\} \cap \{X = x_k\} = \varnothing, k \neq j$，故 $1 = P\left[\bigcup_{k=1}^{\infty} \{X = x_k\}\right] = \sum_{k=1}^{\infty} P\{X = x_k\}$，即 $\sum_{k=1}^{\infty} p_k = 1$.

我们称 (2.1) 式为离散型随机变量 X 的**分布律**．分布律也可以用表格的形式来表示

X	x_1	x_2	$\cdots$	x_n	$\cdots$
p_k	p_1	p_2	$\cdots$	p_n	$\cdots$

(2.4)

(2.4) 直观地表示了随机变量 X 取各个值的概率的规律．X 取各个值各占一些概率，这些概率合起来是 1．可以想像成：概率 1 以一定的规律分布在各个可能值上．这就是 (2.4) 称为分布律的缘故．

例 1　设一汽车在开往目的地的道路上需经过四组信号灯，每组信号灯以 1/2 的概率允许或禁止汽车通过．以 X 表示汽车首次停下时，它已通过的信号灯的组数（设各组信号灯的工作是相互独立的），求 X 的分布律．

解　以 p 表示每组信号灯禁止汽车通过的概率，易知 X 的分布律为

X	0	1	2	3	4
p_k	p	$(1-p)p$	$(1-p)^2p$	$(1-p)^3p$	$(1-p)^4$

或写成

$$P\{X=k\}=(1-p)^k p,k=0,1,2,3,P\{X=4\}=(1-p)^4.$$

以 $p=1/2$ 代入得

X	0	1	2	3	4
p_k	0.5	0.25	0.125	0.0625	0.0625

□

下面介绍三种重要的离散型随机变量.

(一) (0—1)分布

设随机变量 X 只可能取 0 与 1 两个值,它的分布律是

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

则称 X 服从以 p 为参数的 (0—1) 分布或两点分布.

(0—1) 分布的分布律也可写成

X	0	1
p_k	$1-p$	p

对于一个随机试验,如果它的样本空间只包含两个元素,即 $S=\{e_1,e_2\}$,我们总能在 S 上定义一个服从 (0—1) 分布的随机变量

$$X=X(e)=\begin{cases}0,&\text{当 } e=e_1,\\1,&\text{当 } e=e_2\end{cases}$$

来描述这个随机试验的结果. 例如,对新生婴儿的性别进行登记,检查产品的质量是否合格,某车间的电力消耗是否超过负荷以及前面多次讨论过的"抛硬币"试验等都可以用 (0—1) 分布的随机变量来描述. (0—1) 分布是经常遇到的一种分布.

(二) 伯努利试验、二项分布

设试验 E 只有两个可能结果:A 及 $\overline{A}$,则称 E 为**伯努利**(Bernoulli)**试验**. 设 $P(A)=p$ $(0<p<1)$,此时 $P(\overline{A})=1-p$. 将 E 独立重复地进行 n 次,则称这一串重复的独立试验为 **n 重伯努利试验**.

这里"重复"是指在每次试验中 $P(A)=p$ 保持不变;"独立"是指各次试验

的结果互不影响，即若以 C_i 记第 i 次试验的结果，C_i 为 A 或 $\overline{A}$，$i=1,2,\cdots,n$. “独立”是指

$$P(C_1C_2\cdots C_n) = P(C_1)P(C_2)\cdots P(C_n). \tag{2.5}$$

n 重伯努利试验是一种很重要的数学模型，它有广泛的应用，是研究最多的模型之一.

例如，E 是抛一枚硬币观察得到正面或反面. A 表示得正面，这是一个伯努利试验. 如将硬币抛 n 次，就是 n 重伯努利试验. 又如抛一颗骰子，若 A 表示得到“1 点”，$\overline{A}$ 表示得到“非 1 点”. 将骰子抛 n 次，就是 n 重伯努利试验. 再如在袋中装有 a 只白球，b 只黑球. 试验 E 是在袋中任取一只球，观察其颜色. 以 A 表示“取到白球”，$P(A)=a/(a+b)$. 若连续取球 n 次作放回抽样，这就是 n 重伯努利试验. 然而，若作不放回抽样，虽则每次试验都有 $P(A)=a/(a+b)$（见第一章 § 4 例 5），但各次试验不再相互独立[①]，因而不再是 n 重伯努利试验了.

以 X 表示 n 重伯努利试验中事件 A 发生的次数，X 是一个随机变量，我们来求它的分布律. X 所有可能取的值为 $0,1,2,\cdots,n$. 由于各次试验是相互独立的，因此事件 A 在指定的 k $(0\leqslant k\leqslant n)$ 次试验中发生，在其他 $n-k$ 次试验中 A 不发生（例如在前 k 次试验中 A 发生，而后 $n-k$ 次试验中 A 不发生）的概率为

$$\underbrace{p\cdot p\cdot\cdots\cdot p}_{k\text{个}}\cdot\underbrace{(1-p)\cdot(1-p)\cdot\cdots\cdot(1-p)}_{n-k\text{个}} = p^k(1-p)^{n-k}.$$

这种指定的方式共有 $\binom{n}{k}$ 种，它们是两两互不相容的，故在 n 次试验中 A 发生 k 次的概率为 $\binom{n}{k}p^k(1-p)^{n-k}$，记 $q=1-p$，即有

$$P\{X=k\} = \binom{n}{k}p^kq^{n-k},k=0,1,2,\cdots,n. \tag{2.6}$$

显然

$$P\{X=k\}\geqslant 0,\ k=0,1,2,\cdots,n;$$

$$\sum_{k=0}^{n}P\{X=k\} = \sum_{k=0}^{n}\binom{n}{k}p^kq^{n-k} = (p+q)^n = 1.$$

① 对于不放回抽样，以 A_1，A_2 分别记第一次、第二次取到白球，则有 $P(A_2|A_1)=\frac{a-1}{a+b-1}$，而 $P(A_2)=\frac{a}{a+b}$，$P(A_2|A_1)\neq P(A_2)$，故第一次、第二次试验不相互独立. 即知(2.5)不成立.

即 $P\{X=k\}$ 满足条件 (2.2),(2.3). 注意到 $\binom{n}{k}p^kq^{n-k}$ 刚好是二项式 $(p+q)^n$ 的展开式中出现 p^k 的那一项,我们称随机变量 X 服从参数为 n,p 的**二项分布**,并记为 $X\sim b(n,p)$.

特别,当 $n=1$ 时二项分布(2.6)化为

$$P\{X=k\}=p^kq^{1-k},\quad k=0,1.$$

这就是(0—1)分布.

例 2 按规定,某种型号电子元件的使用寿命超过 1 500 小时的为一级品. 已知某一大批产品的一级品率为 0.2,现在从中随机地抽查 20 只. 问 20 只元件中恰有 k 只($k=0,1,\cdots,20$)为一级品的概率是多少?

解 这是不放回抽样. 但由于这批元件的总数很大,且抽查的元件的数量相对于元件的总数来说又很小,因而可以当作放回抽样来处理,这样做会有一些误差,但误差不大. 我们将检查一只元件看它是否为一级品看成是一次试验,检查 20 只元件相当于做 20 重伯努利试验. 以 X 记 20 只元件中一级品的只数,那么,X 是一个随机变量,且有 $X\sim b(20,0.2)$. 由 (2.6) 式即得所求概率为

$$P\{X=k\}=\binom{20}{k}(0.2)^k(0.8)^{20-k},\ k=0,1,\cdots,20.$$

将计算结果列表如下:

$P\{X=0\}=0.012$	$P\{X=4\}=0.218$	$P\{X=8\}=0.022$
$P\{X=1\}=0.058$	$P\{X=5\}=0.175$	$P\{X=9\}=0.007$
$P\{X=2\}=0.137$	$P\{X=6\}=0.109$	$P\{X=10\}=0.002$
$P\{X=3\}=0.205$	$P\{X=7\}=0.055$	
当 $k\geqslant 11$ 时, $P\{X=k\}<0.001$		

为了对本题的结果有一个直观了解,我们作出上表的图形,如图 2—3 所示.

从图 2—3 中看到,当 k 增加时,概率 $P\{X=k\}$ 先是随之增加,直至达到最大值(本例中当 $k=4$ 时取到最大值),随后单调减少. 我们指出,一般,对于固定的 n 及 p,二项分布 $b(n,p)$ 都具有这一性质. □

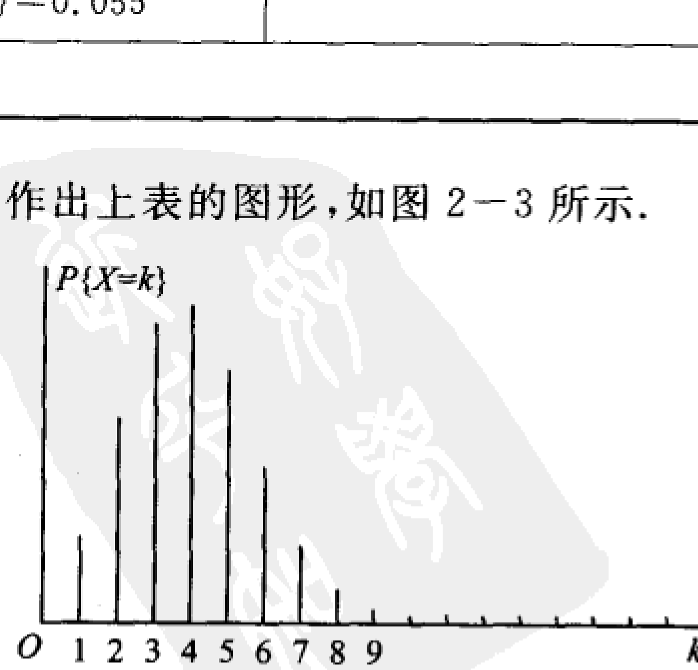

图 2—3

例 3 某人进行射击,设每次射击的命中率为 0.02,独立射击 400 次,试

求至少击中两次的概率.

解　将一次射击看成是一次试验. 设击中的次数为 X,则 $X \sim b(400, 0.02)$. X 的分布律为

$$P\{X=k\}=\binom{400}{k}(0.02)^k(0.98)^{400-k}, k=0,1,\cdots,400.$$

于是所求概率为

$$\begin{aligned}P\{X \geqslant 2\} &= 1-P\{X=0\}-P\{X=1\} \\ &= 1-(0.98)^{400}-400(0.02)(0.98)^{399}=0.9972.\end{aligned}$$

□

这个概率很接近于 1. 我们从两方面来讨论这一结果的实际意义. 其一,虽然每次射击的命中率很小（为 0.02）,但如果射击 400 次,则击中目标至少两次是几乎可以肯定的. 这一事实说明,一个事件尽管在一次试验中发生的概率很小,但只要试验次数很多,而且试验是独立地进行的,那么这一事件的发生几乎是肯定的. 这也告诉人们绝不能轻视小概率事件. 其二,如果射手在 400 次射击中,击中目标的次数竟不到两次,由于概率 $P\{X<2\} \approx 0.003$ 很小,根据实际推断原理,我们将怀疑"每次射击的命中率为 0.02"这一假设,即认为该射手射击的命中率达不到 0.02.

例 4　设有 80 台同类型设备,各台工作是相互独立的,发生故障的概率都是 0.01,且一台设备的故障能由一个人处理. 考虑两种配备维修工人的方法,其一是由 4 人维护,每人负责 20 台;其二是由 3 人共同维护 80 台. 试比较这两种方法在设备发生故障时不能及时维修的概率的大小.

解　按第一种方法. 以 X 记"第 1 人维护的 20 台中同一时刻发生故障的台数",以 $A_i(i=1,2,3,4)$ 表示事件"第 i 人维护的 20 台中发生故障不能及时维修",则知 80 台中发生故障而不能及时维修的概率为

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) \geqslant P(A_1)=P\{X \geqslant 2\}.$$

而 $X \sim b(20, 0.01)$,故有

$$\begin{aligned}P\{X \geqslant 2\} &= 1-\sum_{k=0}^{1} P\{X=k\} \\ &= 1-\sum_{k=0}^{1}\binom{20}{k}(0.01)^k(0.99)^{20-k}=0.0169.\end{aligned}$$

即有

$$P(A_1 \cup A_2 \cup A_3 \cup A_4) \geqslant 0.0169.$$

按第二种方法. 以 Y 记 80 台中同一时刻发生故障的台数. 此时,$Y \sim b(80, 0.01)$,故 80 台中发生故障而不能及时维修的概率为

$$P\{Y \geqslant 4\}=1-\sum_{k=0}^{3}\binom{80}{k}(0.01)^k(0.99)^{80-k}=0.0087.$$

我们发现,在后一种情况尽管任务重了（每人平均维护约 27 台）,但工作效

率不仅没有降低,反而提高了. □

(三)泊松分布

设随机变量 X 所有可能取的值为 $0,1,2,\cdots$,而取各个值的概率为

$$P\{X=k\}=\frac{\lambda^k e^{-\lambda}}{k!},\ k=0,1,2,\cdots,$$

其中 $\lambda>0$ 是常数. 则称 X 服从参数为 λ 的**泊松分布**,记为 $X\sim\pi(\lambda)$.

易知,$P\{X=k\}\geqslant 0,k=0,1,2,\cdots$,且有

$$\sum_{k=0}^{\infty}P\{X=k\}=\sum_{k=0}^{\infty}\frac{\lambda^k e^{-\lambda}}{k!}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=e^{-\lambda}\cdot e^{\lambda}=1.$$

即 $P\{X=k\}$ 满足条件 (2.2),(2.3).

参数 λ 的意义将在第四章说明,有关服从泊松分布的随机变量的数学模型将在第十二章中讨论.

具有泊松分布的随机变量在实际应用中是很多的. 例如,一本书一页中的印刷错误数、某地区在一天内邮递遗失的信件数、某一医院在一天内的急诊病人数、某一地区一个时间间隔内发生交通事故的次数、在一个时间间隔内某种放射性物质发出的、经过计数器的 α 粒子数等都服从泊松分布. 泊松分布也是概率论中的一种重要分布.

下面介绍一个用泊松分布来逼近二项分布的定理.

泊松定理 设 $\lambda>0$ 是一个常数,n 是任意正整数,设 $np_n=\lambda$,则对于任一固定的非负整数 k,有

$$\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=\frac{\lambda^k e^{-\lambda}}{k!}.$$

证 由 $p_n=\dfrac{\lambda}{n}$,有

$$\binom{n}{k}p_n^k(1-p_k)^{n-k}=\frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$=\frac{\lambda^k}{k!}\left[1\cdot\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\right]\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}.$$

对于任意固定的 k,当 $n\to\infty$ 时

$$1\cdot\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\to 1,\left(1-\frac{\lambda}{n}\right)^n\to e^{-\lambda},\left(1-\frac{\lambda}{n}\right)^{-k}\to 1.$$

故有 $$\lim_{n\to\infty}\binom{n}{k}p_n^k(1-p_n)^{n-k}=\frac{\lambda^k e^{-\lambda}}{k!}.$$ □

定理的条件 $np_n=\lambda$(常数)意味着当 n 很大时 p_n 必定很小,因此,上述定理表明当 n 很大,p 很小($np=\lambda$)时有以下近似式

$$\binom{n}{k}p^k(1-p)^{n-k}\approx\frac{\lambda^k e^{-\lambda}}{k!}\quad(\text{其中 }\lambda=np). \tag{2.7}$$

也就是说以 n,p 为参数的二项分布的概率值可以由参数为 $\lambda=np$ 的泊松分布的概率值近似.上式也能用来作二项分布概率的近似计算.

例 5 计算机硬件公司制造某种特殊型号的微型芯片,次品率达 0.1%,各芯片成为次品相互独立.求在 1000 只产品中至少有 2 只次品的概率.以 X 记产品中的次品数,$X\sim b(1000,0.001)$.

解 所求概率为

$$\begin{aligned}P\{X\geqslant 2\}&=1-P\{X=0\}-P\{X=1\}\\&=1-(0.999)^{1000}-\binom{1000}{1}(0.999)^{999}(0.001)\\&\approx 1-0.3676954-0.3680635\approx 0.2642411.\end{aligned}$$

利用(2.7)式来计算得,$\lambda=1000\times 0.001=1$,

$$\begin{aligned}P\{X\geqslant 2\}&=1-P\{X=0\}-P\{X=1\}\\&=1-e^{-1}-e^{-1}\approx 0.2642411.\end{aligned}$$ □

显然利用(2.7)式的计算来得方便.一般,当 $n\geqslant 20,p\leqslant 0.05$ 时用 $\frac{\lambda^k e^{-\lambda}}{k!}$ $(\lambda=np)$ 作为 $\binom{n}{k}p^k(1-p)^{n-k}$ 的近似值效果颇佳.

§3 随机变量的分布函数

对于非离散型随机变量 X,由于其可能取的值不能一一列举出来,因而就不能像离散型随机变量那样可以用分布律来描述它.另外,我们通常所遇到的非离散型随机变量取任一指定的实数值的概率都等于 0(这一点在下一节将会讲到).再者,在实际中,对于这样的随机变量,例如误差 ε,元件的寿命 T 等,我们并不会对误差 $\varepsilon=0.05$ mm,寿命 $T=1\ 251.3$ h 的概率感兴趣,而是考虑误差落在某个区间内的概率,寿命 T 大于某个数的概率.因而我们转而去研究随机变量所取的值落在一个区间$(x_1,x_2]$的概率:$P\{x_1<X\leqslant x_2\}$.但由于

$$P\{x_1<X\leqslant x_2\}=P\{X\leqslant x_2\}-P\{X\leqslant x_1\},$$

所以我们只需知道 $P\{X\leqslant x_2\}$ 和 $P\{X\leqslant x_1\}$ 就可以了.下面引入随机变量的分布函数的概念①.

① 虽然对于离散型随机变量,我们可以用分布律全面地描述它,但为了从数学上能统一地对随机变量进行研究,在这里,我们对离散型随机变量和非离散型随机变量统一地定义了分布函数.

定义 设 X 是一个随机变量，x 是任意实数，函数

$$F(x)=P\{X\leqslant x\},\quad -\infty<x<\infty$$

称为 X 的**分布函数**.

对于任意实数 $x_1,x_2(x_1<x_2)$，有

$$\begin{aligned}P\{x_1<X\leqslant x_2\}&=P\{X\leqslant x_2\}-P\{X\leqslant x_1\}\\&=F(x_2)-F(x_1),\end{aligned}\tag{3.1}$$

因此，若已知 X 的分布函数，我们就知道 X 落在任一区间 $(x_1,x_2]$上的概率，从这个意义上说，分布函数完整地描述了随机变量的统计规律性.

分布函数是一个普通的函数，正是通过它，我们将能用数学分析的方法来研究随机变量.

如果将 X 看成是数轴上的随机点的坐标，那么，分布函数$F(x)$在 x 处的函数值就表示 X 落在区间$(-\infty,x]$上的概率.

分布函数 $F(x)$具有以下的基本性质：

1° $F(x)$ 是一个不减函数.

事实上，由(3.1)式对于任意实数 $x_1,x_2(x_1<x_2)$，有

$$F(x_2)-F(x_1)=P\{x_1<X\leqslant x_2\}\geqslant 0.$$

2° $0\leqslant F(x)\leqslant 1$，且

$$F(-\infty)=\lim_{x\to-\infty}F(x)=0,$$
$$F(\infty)=\lim_{x\to\infty}F(x)=1.$$

上面两个式子，我们只从几何上加以说明. 在图 2－4 中，将区间端点 x 沿数轴无限向左移动（即 $x\to-\infty$），则“随机点 X 落在点 x 左边”这一事件趋于不可能事件，从而其概率趋于 0，即有 $F(-\infty)=0$；又若将点 x 无限右移（即 $x\to\infty$），则“随机点 X 落在点 x 左边”这一事件趋于必然事件，从而其概率趋于 1，即有 $F(\infty)=1$.

图 2—4

3° $F(x+0)=F(x)$，即 $F(x)$是右连续的.(证略)

例 1 设随机变量 X 的分布律为

X	-1	2	3
p_k	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

求 X 的分布函数，并求 $P\left\{X\leqslant\frac{1}{2}\right\}$，$P\left\{\frac{3}{2}<X\leqslant\frac{5}{2}\right\}$，$P\{2\leqslant X\leqslant 3\}$.

解 X 仅在 $x=-1,2,3$ 三点处其概率$\neq 0$，而 $F(x)$的值是 $X\leqslant x$ 的累积概

率值，由概率的有限可加性，知它即为小于或等于 x 的那些 x_k 处的概率 p_k 之和，有

$$F(x)=\begin{cases}0 & x<-1,\\ P\{X=-1\}, & -1\leqslant x<2,\\ P\{X=-1\}+P\{X=2\}, & 2\leqslant x<3,\\ 1, & x\geqslant 3.\end{cases}$$

即
$$F(x)=\begin{cases}0, & x<-1,\\ \dfrac{1}{4}, & -1\leqslant x<2,\\ \dfrac{3}{4}, & 2\leqslant x<3,\\ 1, & x\geqslant 3.\end{cases}$$

$F(x)$的图形如图 2—5 所示，它是一条阶梯形的曲线，在 $x=-1,2,3$ 处有跳跃点，跳跃值分别为$\frac{1}{4}$，$\frac{1}{2}$，$\frac{1}{4}$. 又

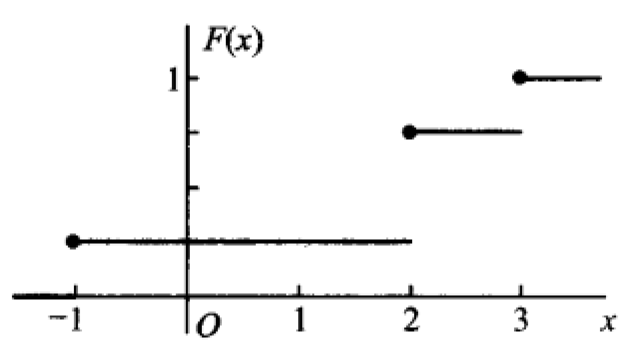

图 2—5

$$P\left\{X\leqslant\frac{1}{2}\right\}=F\left(\frac{1}{2}\right)=\frac{1}{4},$$

$$P\left\{\frac{3}{2}<X\leqslant\frac{5}{2}\right\}=F\left(\frac{5}{2}\right)-F\left(\frac{3}{2}\right)=\frac{3}{4}-\frac{1}{4}=\frac{1}{2}.$$

$$P\{2\leqslant X\leqslant 3\}=F(3)-F(2)+P\{X=2\}$$
$$=1-\frac{3}{4}+\frac{1}{2}=\frac{3}{4}.$$ □

一般，设离散型随机变量 X 的分布律为

$$P\{X=x_k\}=p_k,\quad k=1,2,\cdots.$$

由概率的可列可加性得 X 的分布函数为

$$F(x)=P\{X\leqslant x\}=\sum_{x_k\leqslant x}P\{X=x_k\},$$

即
$$F(x)=\sum_{x_k\leqslant x}p_k, \tag{3.2}$$

这里和式是对于所有满足 $x_k \leqslant x$ 的 k 求和的. 分布函数 $F(x)$ 在 $x=x_k(k=1,2,\cdots)$ 处有跳跃,其跳跃值为 $p_k=P\{X=x_k\}$.

例 2 一个靶子是半径为 2 m 的圆盘,设击中靶上任一同心圆盘上的点的概率与该圆盘的面积成正比,并设射击都能中靶,以 X 表示弹着点与圆心的距离. 试求随机变量 X 的分布函数.

解 若 $x<0$,则 $\{X \leqslant x\}$ 是不可能事件,于是

$$F(x)=P\{X \leqslant x\}=0.$$

若 $0 \leqslant x \leqslant 2$,由题意,$P\{0 \leqslant X \leqslant x\}=kx^2$,$k$ 是某一常数,为了确定 k 的值,取 $x=2$,有 $P\{0 \leqslant X \leqslant 2\}=2^2k$,但已知 $P\{0 \leqslant X \leqslant 2\}=1$,故得 $k=1/4$,即

$$P\{0 \leqslant X \leqslant x\}=\frac{x^2}{4}.$$

于是

$$F(x)=P\{X \leqslant x\}=P\{X<0\}+P\{0 \leqslant X \leqslant x\}=\frac{x^2}{4}.$$

若 $x \geqslant 2$,由题意 $\{X \leqslant x\}$ 是必然事件,于是

$$F(x)=P\{X \leqslant x\}=1.$$

综合上述,即得 X 的分布函数为

$$F(x)=\begin{cases}0, & x<0,\\ \dfrac{x^2}{4}, & 0 \leqslant x<2,\\ 1, & x \geqslant 2.\end{cases}$$

它的图形是一条连续曲线如图 2-6 所示.

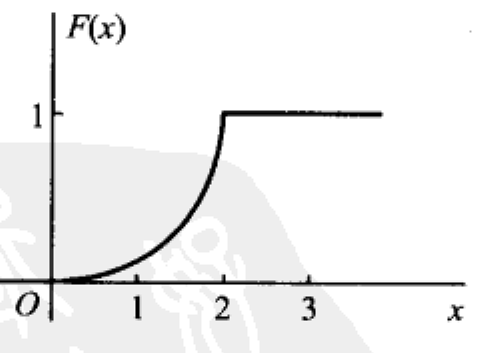

图 2-6

另外,容易看到本例中的分布函数 $F(x)$,对于任意 x 可以写成形式

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

其中

$$f(t)=\begin{cases}\dfrac{t}{2}, & 0<t<2,\\ 0, & \text{其他}.\end{cases}$$

这就是说,$F(x)$ 恰是非负函数 $f(t)$ 在区间 $(-\infty,x]$ 上的积分,在这种情况我们称 X 为连续型随机变量. 下一节我们将给出连续型随机变量的一般定义. □

§4　连续型随机变量及其概率密度

一般，如上节例 2 中的随机变量那样，如果对于随机变量 X 的分布函数 $F(x)$，存在非负函数 $f(x)$，使对于任意实数 x 有

$$F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t, \tag{4.1}$$

则称 X 为**连续型随机变量**，其中函数 $f(x)$ 称为 X 的**概率密度函数**，简称**概率密度**①.

由(4.1)式，据数学分析的知识知连续型随机变量的分布函数是连续函数.

在实际应用中遇到的基本上是离散型或连续型随机变量. 本书只讨论这两种随机变量.

由定义知道，概率密度 $f(x)$ 具有以下性质：

1° $f(x)\geqslant 0$；

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

3° 对于任意实数 $x_1,x_2(x_1\leqslant x_2)$，

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

4° 若 $f(x)$ 在点 x 处连续，则有 $F'(x)=f(x)$.

由性质 2°知道介于曲线 $y=f(x)$ 与 Ox 轴之间的面积等于 1（图 2－7）. 由 3°知道 X 落在区间 $(x_1,x_2]$ 的概率 $P\{x_1<X\leqslant x_2\}$ 等于区间 $(x_1,x_2]$ 上曲线 $y=f(x)$ 之下的曲边梯形的面积（图 2－8）. 由性质 4°在 $f(x)$ 的连续点 x 处有

$$\begin{aligned} f(x) &= \lim_{\Delta x\to 0^+}\frac{F(x+\Delta x)-F(x)}{\Delta x} \\ &= \lim_{\Delta x\to 0^+}\frac{P\{x<X\leqslant x+\Delta x\}}{\Delta x}. \end{aligned} \tag{4.2}$$

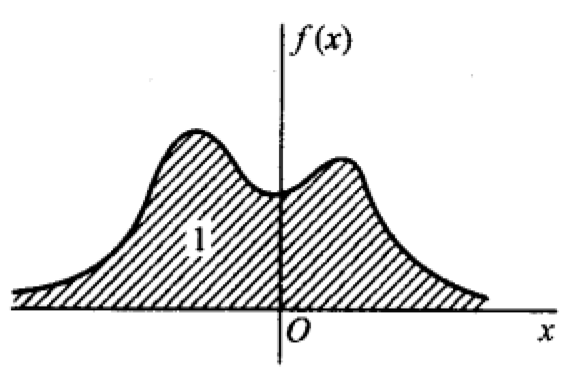

图 2－7

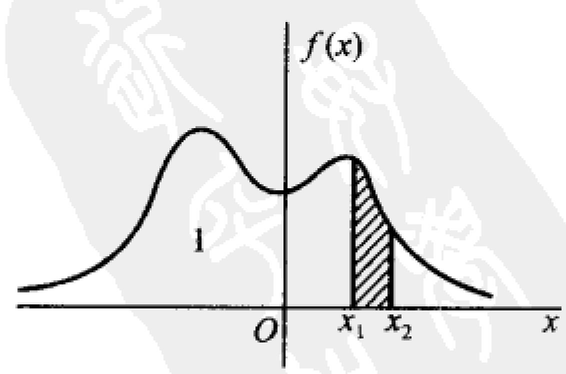

图 2－8

① 由定义知道，改变概率密度 $f(x)$ 在个别点的函数值不影响分布函数 $F(x)$ 的取值. 因此，并不在乎改变概率密度在个别点上的值.

从这里我们看到概率密度的定义与物理学中的线密度的定义相类似,这就是为什么称 $f(x)$ 为概率密度的缘故.

由(4.2)式知道,若不计高阶无穷小,有

$$P\{x<X\leqslant x+\Delta x\}\approx f(x)\Delta x. \tag{4.3}$$

这表示 X 落在小区间 $(x,x+\Delta x]$ 上的概率近似地等于 $f(x)\Delta x$.

例 1 设随机变量 X 具有概率密度

$$f(x)=\begin{cases}kx, & 0\leqslant x<3,\\ 2-\dfrac{x}{2}, & 3\leqslant x\leqslant 4,\\ 0, & \text{其他}.\end{cases}$$

(1) 确定常数 k;(2)求 X 的分布函数 $F(x)$;(3)求 $P\{1<X\leqslant\frac{7}{2}\}$.

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

$$\int_0^3 kx\,\mathrm{d}x+\int_3^4\left(2-\frac{x}{2}\right)\mathrm{d}x=1,$$

解得 $k=\frac{1}{6}$,于是 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{x}{6}, & 0\leqslant x<3,\\ 2-\dfrac{x}{2}, & 3\leqslant x<4,\\ 0, & \text{其他}.\end{cases}$$

(2) X 的分布函数为

$$F(x)=\begin{cases}0, & x<0,\\ \displaystyle\int_0^x\frac{x}{6}\mathrm{d}x, & 0\leqslant x<3,\\ \displaystyle\int_0^3\frac{x}{6}\mathrm{d}x+\int_3^x\left(2-\frac{x}{2}\right)\mathrm{d}x, & 3\leqslant x<4,\\ 1, & x\geqslant 4.\end{cases}$$

即

$$F(x)=\begin{cases}0, & x<0,\\ \dfrac{x^2}{12}, & 0\leqslant x<3,\\ -3+2x-\dfrac{x^2}{4}, & 3\leqslant x<4,\\ 1, & x\geqslant 4.\end{cases}$$

(3) $P\{1<X\leqslant\frac{7}{2}\}=F\left(\frac{7}{2}\right)-F(1)=\frac{41}{48}.$ □

需要指出的是，对于连续型随机变量 X 来说，它取任一指定实数值 a 的概率均为 0，即 $P\{X=a\}=0$. 事实上，设 X 的分布函数为 $F(x)$，$\Delta x>0$，则由 $\{X=a\}\subset\{a-\Delta x<X\leqslant a\}$ 得

$$0\leqslant P\{X=a\}\leqslant P\{a-\Delta x<X\leqslant a\}=F(a)-F(a-\Delta x).$$

在上述不等式中令 $\Delta x\to 0$，并注意到 X 为连续型随机变量，其分布函数 $F(x)$ 是连续的. 即得

$$P\{X=a\}=0. \tag{4.4}$$

据此，在计算连续型随机变量落在某一区间的概率时，可以不必区分该区间是开区间或闭区间或半闭区间. 例如有

$$P\{a<X\leqslant b\}=P\{a\leqslant X\leqslant b\}=P\{a<X<b\}.$$

在这里，事件 $\{X=a\}$ 并非不可能事件，但有 $P\{X=a\}=0$. 这就是说，若 A 是不可能事件，则有 $P(A)=0$；反之，若 $P(A)=0$，并不一定意味着 A 是不可能事件.

以后当我们提到一个随机变量 X 的“概率分布”时，指的是它的分布函数；或者，当 X 是连续型随机变量时，指的是它的概率密度，当 X 是离散型随机变量时，指的是它的分布律.

下面介绍三种重要的连续型随机变量.

（一）均匀分布

若连续型随机变量 X 具有概率密度

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b,\\ 0, & \text{其他},\end{cases} \tag{4.5}$$

则称 X 在区间 (a,b) 上服从**均匀分布**. 记为 $X\sim U(a,b)$.

易知 $f(x)\geqslant 0$，且 $\int_{-\infty}^{\infty}f(x)\mathrm{d}x=1$.

在区间 (a,b) 上服从均匀分布的随机变量 X，具有下述意义的等可能性，即它落在区间 (a,b) 中任意等长度的子区间内的可能性是相同的. 或者说它落在 (a,b) 的子区间内的概率只依赖于子区间的长度而与子区间的位置无关. 事实上，对于任一长度 l 的子区间 $(c,c+l)$，$a\leqslant c<c+l\leqslant b$，有

$$P\{c<X\leqslant c+l\}=\int_c^{c+l}f(x)\mathrm{d}x=\int_c^{c+l}\frac{1}{b-a}\mathrm{d}x=\frac{l}{b-a}.$$

由(4.1)式得 X 的分布函数为

$$F(x)=\begin{cases}0, & x<a,\\ \dfrac{x-a}{b-a}, & a\leqslant x<b,\\ 1, & x\geqslant b.\end{cases} \tag{4.6}$$

$f(x)$及$F(x)$的图形分别如图 2—9,图 2—10 所示.

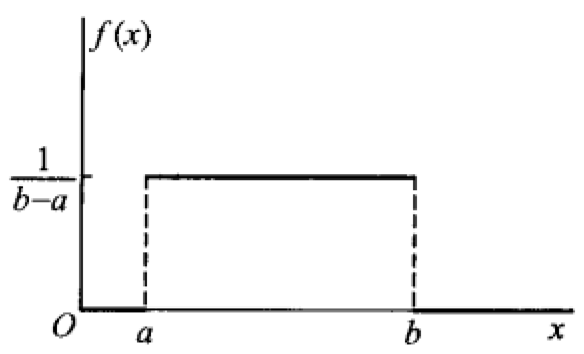

图 2—9

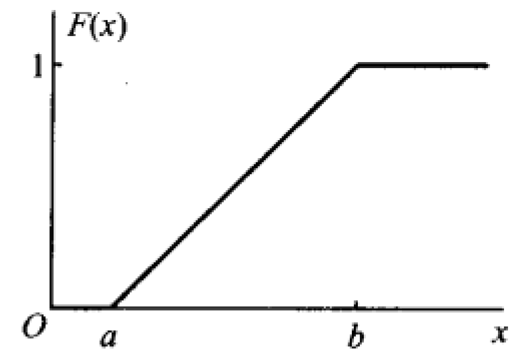

图 2—10

例 2 设电阻值 R 是一个随机变量,均匀分布在 900 Ω~1100 Ω. 求 R 的概率密度及 R 落在 950 Ω~1 050 Ω 的概率.

解 按题意,R 的概率密度为

$$f(r)=\begin{cases}\dfrac{1}{1\,100-900}, & 900<r<1100,\\ 0, & 其他.\end{cases}$$

故有
$$P\{950<R\leqslant 1\,050\}=\int_{950}^{1\,050}\frac{1}{200}\mathrm{d}r=0.5.$$ □

(二) 指数分布

若连续型随机变量 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\theta}\mathrm{e}^{-x/\theta}, x>0,\\ 0, \quad 其他,\end{cases}\tag{4.7}$$

其中 $\theta>0$ 为常数,则称 X 服从参数为 θ 的**指数分布**.

易知 $f(x)\geqslant 0$, 且$\int_{-\infty}^{\infty}f(x)\mathrm{d}x=1$. 图 2—11 中画出了$\theta=1/3,\theta=1,\theta=2$时 $f(x)$的图形.

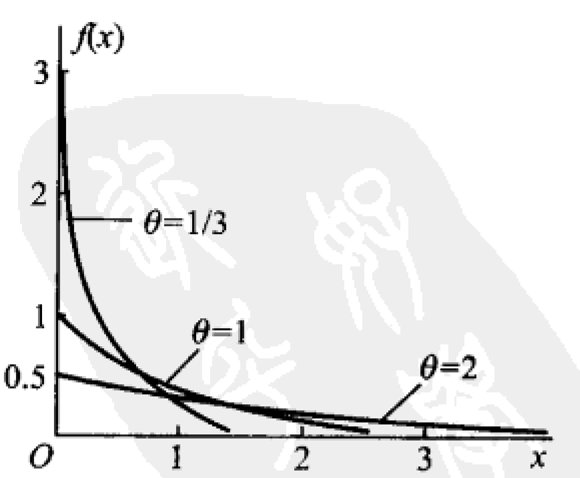

图 2—11

由(4.7)式容易得到随机变量 X 的分布函数为

$$F(x)=\begin{cases}1-\mathrm{e}^{-x/\theta}, & x>0,\\ 0, & 其他.\end{cases}\tag{4.8}$$

服从指数分布的随机变量 X 具有以下有趣的性质:

对于任意 $s,t>0$,有

$$P\{X > s+t \mid X > s\} = P\{X > t\}. \tag{4.9}$$

事实上

$$\begin{aligned} P\{X > s+t \mid X > s\} &= \frac{P\{(X > s+t) \cap (X > s)\}}{P\{X > s\}} \\ &= \frac{P\{X > s+t\}}{P\{X > s\}} = \frac{1-F(s+t)}{1-F(s)} \\ &= \frac{e^{-(s+t)/\theta}}{e^{-s/\theta}} = e^{-t/\theta} \\ &= P\{X > t\}. \end{aligned}$$

性质(4.9)称为**无记忆性**．如果 X 是某一元件的寿命，那么(4.9)式表明：已知元件已使用了 s 小时，它总共能使用至少 $s+t$ 小时的条件概率，与从开始使用时算起它至少能使用 t 小时的概率相等．这就是说，元件对它已使用过 s 小时没有记忆．具有这一性质是指数分布有广泛应用的重要原因．

指数分布在可靠性理论与排队论中有广泛的应用．

(三) 正态分布

若连续型随机变量 X 的概率密度为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty, \tag{4.10}$$

其中 $\mu, \sigma(\sigma > 0)$ 为常数，则称 X 服从参数为 μ, σ 的**正态分布**或**高斯**(Gauss)分布，记为 $X \sim N(\mu, \sigma^2)$.

显然 $f(x) \geqslant 0$，下面来证明 $\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$. 令 $(x-\mu)/\sigma = t$，得到

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2/2} \mathrm{d}t.$$

记　$I = \int_{-\infty}^{\infty} e^{-t^2/2}\mathrm{d}t$，则有 $I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(t^2+u^2)/2}\mathrm{d}t\mathrm{d}u$，利用极坐标将它化成累次积分，得到

$$I^2 = \int_0^{2\pi}\int_0^{\infty} r e^{-r^2/2} \mathrm{d}r\mathrm{d}\theta = 2\pi.$$

而 $I > 0$，故有 $I = \sqrt{2\pi}$，即有

$$\int_{-\infty}^{\infty} e^{-t^2/2}\mathrm{d}t = \sqrt{2\pi}, \tag{4.11}$$

于是

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-t^2/2}\mathrm{d}t = 1.$$

参数 μ, σ 的意义将在第四章中说明. $f(x)$ 的图形如图 2－12 所示，它具有以下的性质.

1° 曲线关于 $x=\mu$ 对称．这表明对于任意 $h>0$ 有（图 2—12）

$$P\{\mu-h<X\leqslant\mu\}=P\{\mu<X\leqslant\mu+h\}.$$

2° 当 $x=\mu$ 时取到最大值

max $$f(\mu)=\frac{1}{\sqrt{2\pi}\sigma}.$$

x 离 μ 越远，$f(x)$ 的值越小．这表明对于同样长度的区间，当区间离 μ 越远，X 落在这个区间上的概率越小．

在 $x=\mu\pm\sigma$ 处曲线有拐点．曲线以 Ox 轴为渐近线．

另外，如果固定 σ，改变 μ 的值，则图形沿着 Ox 轴平移，而不改变其形状（如图 2—12），可见正态分布的概率密度曲线 $y=f(x)$ 的位置完全由参数 μ 所确定．μ 称为位置参数．

如果固定 μ，改变 σ，由于最大值 $f(\mu)=\frac{1}{\sqrt{2\pi}\sigma}$，可知当 σ 越小时图形变得越尖（如图 2—13），因而 X 落在 μ 附近的概率越大．

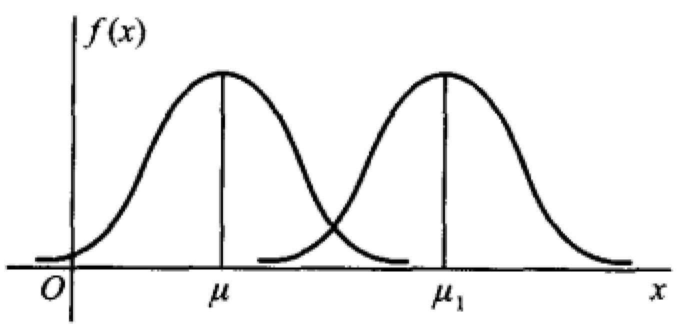

图 2—12

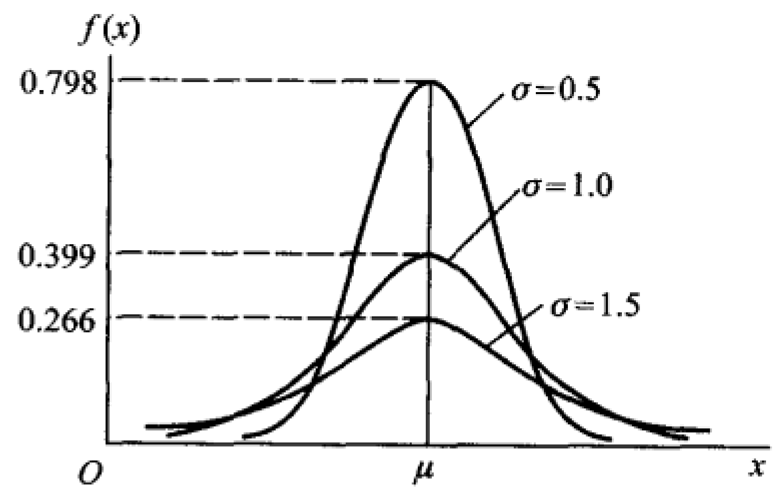

图 2—13

由(4.10)式得 X 的分布函数为（如图 2—14）

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt, \tag{4.12}$$

特别，当 $\mu=0,\sigma=1$ 时称随机变量 X 服从**标准正态分布**．其概率密度和分布函数分别用 $\varphi(x)$，$\Phi(x)$ 表示，即有

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-t^2/2}, \tag{4.13}$$

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-t^2/2}dt. \tag{4.14}$$

易知 $$\Phi(-x)=1-\Phi(x) \tag{4.15}$$

(参见图 2—15).

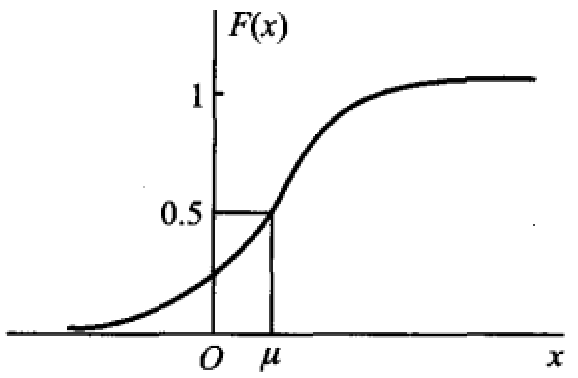

图 2—14

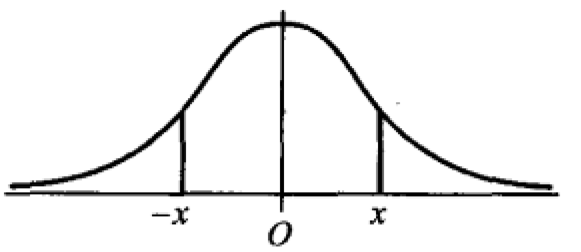

图 2—15

人们已编制了 $\Phi(x)$ 的函数表.可供查用(见附表 2).

一般,若 $X\sim N(\mu,\sigma^2)$,我们只要通过一个线性变换就能将它化成标准正态分布.

引理　若 $X\sim N(\mu,\sigma^2)$,则 $Z=\dfrac{X-\mu}{\sigma}\sim N(0,1)$.

证　$Z=\dfrac{X-\mu}{\sigma}$ 的分布函数为

$$P\{Z\leqslant x\}=P\left\{\frac{X-\mu}{\sigma}\leqslant x\right\}=P\{X\leqslant\mu+\sigma x\}$$

$$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\mu+\sigma x}\mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}\mathrm{d}t,$$

令 $\dfrac{t-\mu}{\sigma}=u$,得

$$P\{Z\leqslant x\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-u^2/2}\mathrm{d}u=\Phi(x),$$

由此知 $Z=\dfrac{X-\mu}{\sigma}\sim N(0,1)$.　□

于是,若 $X\sim N(\mu,\sigma^2)$,则它的分布函数 $F(x)$ 可写成

$$F(x)=P\{X\leqslant x\}=P\left\{\frac{X-\mu}{\sigma}\leqslant\frac{x-\mu}{\sigma}\right\}=\Phi\left(\frac{x-\mu}{\sigma}\right).\tag{4.16}$$

对于任意区间 $(x_1,x_2]$,有

$$P\{x_1<X\leqslant x_2\}=P\left\{\frac{x_1-\mu}{\sigma}<\frac{X-\mu}{\sigma}\leqslant\frac{x_2-\mu}{\sigma}\right\}$$

$$=\Phi\left(\frac{x_2-\mu}{\sigma}\right)-\Phi\left(\frac{x_1-\mu}{\sigma}\right).\tag{4.17}$$

例如,设 $X\sim N(1,4)$,查表得

$$P\{0<X\leqslant 1.6\}=\Phi\left(\frac{1.6-1}{2}\right)-\Phi\left(\frac{0-1}{2}\right)$$

$$=\Phi(0.3)-\Phi(-0.5)$$
$$=0.6179-[1-\Phi(0.5)]$$
$$=0.6179-1+0.6915=0.3094.$$

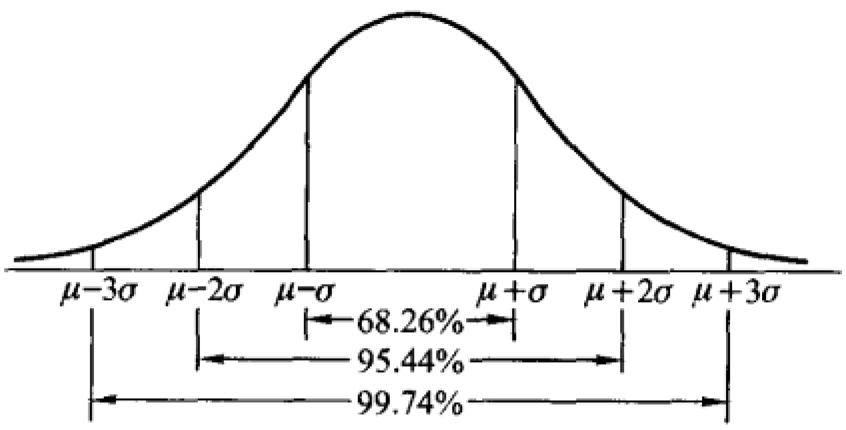

图 2—16

设 $X\sim N(\mu,\sigma^2)$，由 $\Phi(x)$ 的函数表还能得到(图 2—16)：

$$P\{\mu-\sigma<X<\mu+\sigma\}=\Phi(1)-\Phi(-1)$$
$$=2\Phi(1)-1=68.26\%,$$
$$P\{\mu-2\sigma<X<\mu+2\sigma\}=\Phi(2)-\Phi(-2)=95.44\%,$$
$$P\{\mu-3\sigma<X<\mu+3\sigma\}=\Phi(3)-\Phi(-3)=99.74\%.$$

我们看到，尽管正态变量的取值范围是 $(-\infty,\infty)$，但它的值落在 $(\mu-3\sigma,\mu+3\sigma)$ 内几乎是肯定的事．这就是人们所说的"3σ"法则．

例 3 将一温度调节器放置在贮存着某种液体的容器内. 调节器整定在 d ℃，液体的温度 X（以℃计）是一个随机变量，且 $X\sim N(d,0.5^2)$.（1）若 $d=90$ ℃，求 X 小于 89 ℃的概率.（2）若要求保持液体的温度至少为 80 ℃的概率不低于 0.99，问 d 至少为多少？

解 （1）所求概率为

$$P\{X<89\}=P\left\{\frac{X-90}{0.5}<\frac{89-90}{0.5}\right\}$$
$$=\Phi\left(\frac{89-90}{0.5}\right)=\Phi(-2)$$
$$=1-\Phi(2)=1-0.9772=0.0228.$$

（2）按题意需求 d 满足

$$0.99\leqslant P\{X\geqslant 80\}=P\left\{\frac{X-d}{0.5}\geqslant\frac{80-d}{0.5}\right\}$$
$$=1-P\left\{\frac{X-d}{0.5}<\frac{80-d}{0.5}\right\}=1-\Phi\left(\frac{80-d}{0.5}\right).$$

即
$$\Phi\left(\frac{d-80}{0.5}\right)\geqslant 0.99=\Phi(2.327),$$

亦即
$$\frac{d-80}{0.5}\geqslant 2.327.$$
故需
$$d>81.1635.$$
□

为了便于今后在数理统计中的应用，对于标准正态随机变量，我们引入上 α 分位点的定义.

设 $X\sim N(0,1)$，若 z_α 满足条件
$$P\{X>z_\alpha\}=\alpha,\ 0<\alpha<1,\qquad (4.18)$$
则称点 z_α 为标准正态分布的**上 α 分位点**（如图 2－17）. 下面列出了几个常用的 z_α 的值.

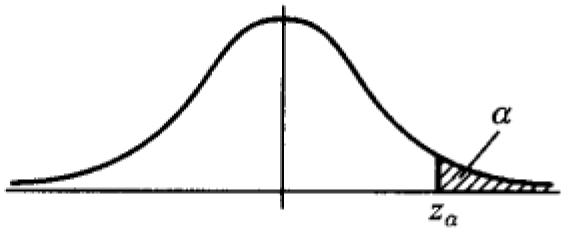

图 2－17

α	0.001	0.005	0.01	0.025	0.05	0.10
z_α	3.090	2.576	2.326	1.960	1.645	1.282

另外，由 $\varphi(x)$ 图形的对称性知道 $z_{1-\alpha}=-z_\alpha$.

在自然现象和社会现象中，大量随机变量都服从或近似服从正态分布. 例如，一个地区的男性成年人的身高，测量某零件长度的误差，海洋波浪的高度，半导体器件中的热噪声电流或电压等，都服从正态分布. 在概率论与数理统计的理论研究和实际应用中正态随机变量起着特别重要的作用. 在第五章我们将进一步说明正态随机变量的重要性.

§5　随机变量的函数的分布

在实际中，我们常对某些随机变量的函数更感兴趣. 例如，在一些试验中，所关心的随机变量往往不能由直接测量得到，而它却是某个能直接测量的随机变量的函数. 比如我们能测量圆轴截面的直径 d，而关心的却是截面面积 $A=\frac{1}{4}\pi d^2$. 这里，随机变量 A 是随机变量 d 的函数. 在这一节中，我们将讨论如何由已知的随机变量 X 的概率分布去求得它的函数 $Y=g(X)$（$g(\cdot)$是已知的连续函数）的概率分布. 这里 Y 是这样的随机变量，当 X 取值 x 时，Y 取值 $g(x)$.

例 1　设随机变量 X 具有以下的分布律，试求 $Y=(X-1)^2$ 的分布律.

X	-1	0	1	2
p_k	0.2	0.3	0.1	0.4

解　Y 所有可能取的值为 0，1，4. 由
$$P\{Y=0\}=P\{(X-1)^2=0\}=P\{X=1\}=0.1,$$
$$P\{Y=1\}=P\{X=0\}+P\{X=2\}=0.7,$$
$$P\{Y=4\}=P\{X=-1\}=0.2,$$

即得 Y 的分布律为

Y	0	1	4
p_k	0.1	0.7	0.2

□

例 2 设随机变量 X 具有概率密度

$$f_X(x)=\begin{cases}\dfrac{x}{8}, & 0<x<4,\\ 0, & \text{其他}.\end{cases}$$

求随机变量 $Y=2X+8$ 的概率密度.

解 分别记 X,Y 的分布函数为 $F_X(x),F_Y(y)$. 下面先来求 $F_Y(y)$.

$$\begin{aligned}F_Y(y)&=P\{Y\leqslant y\}=P\{2X+8\leqslant y\}\\&=P\left\{X\leqslant\frac{y-8}{2}\right\}=F_X\left(\frac{y-8}{2}\right).\end{aligned}$$

将 $F_Y(y)$ 关于 y 求导数,得 $Y=2X+8$ 的概率密度为

$$\begin{aligned}f_Y(y)&=f_X\left(\frac{y-8}{2}\right)\left(\frac{y-8}{2}\right)'\\&=\begin{cases}\dfrac{1}{8}\left(\dfrac{y-8}{2}\right)\cdot\dfrac{1}{2}, & 0<\dfrac{y-8}{2}<4,\\ 0, & \text{其他}\end{cases}\\&=\begin{cases}\dfrac{y-8}{32}, & 8<y<16,\\ 0, & \text{其他}.\end{cases}\end{aligned}$$

□

例 3 设随机变量 X 具有概率密度 $f_X(x),-\infty<x<\infty$,求 $Y=X^2$ 的概率密度.

解 分别记 X,Y 的分布函数为 $F_X(x),F_Y(y)$. 先来求 Y 的分布函数 $F_Y(y)$. 由于 $Y=X^2\geqslant0$,故当 $y\leqslant0$ 时 $F_Y(y)=0$. 当 $y>0$ 时有

$$\begin{aligned}F_Y(y)&=P\{Y\leqslant y\}=P\{X^2\leqslant y\}\\&=P\{-\sqrt{y}\leqslant X\leqslant\sqrt{y}\}\\&=F_X(\sqrt{y})-F_X(-\sqrt{y}).\end{aligned}$$

将 $F_Y(y)$ 关于 y 求导数,即得 Y 的概率密度为

$$f_Y(y)=\begin{cases}\dfrac{1}{2\sqrt{y}}[f_X(\sqrt{y})+f_X(-\sqrt{y})], & y>0,\\ 0, & y\leqslant0.\end{cases}\tag{5.1}$$

□

例如,设 $X\sim N(0,1)$,其概率密度为

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2},-\infty<x<\infty.$$

由(5.1)得 $Y=X^2$ 的概率密度为

$$f_Y(y)=\begin{cases}\dfrac{1}{\sqrt{2\pi}}y^{-1/2}e^{-y/2}, & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

此时称 Y 服从自由度为 1 的 χ^2 分布.

上述两个例子解法的关键一步是在“$Y\leqslant y$”中，即在“$g(X)\leqslant y$”中解出 X，从而得到一个与“$g(X)\leqslant y$”等价的 X 的不等式，并以后者代替“$g(X)\leqslant y$”. 例如，在例 2 中以“$X\leqslant\dfrac{y-8}{2}$”代替“$2X+8\leqslant y$”；在例 3 中，当 $y>0$ 时以“$-\sqrt{y}\leqslant X\leqslant\sqrt{y}$”代替“$X^2\leqslant y$”. 一般来说，可以用这样的方法①求连续型随机变量的函数的分布函数或概率密度. 下面我们仅对 $Y=g(X)$，其中 $g(\cdot)$ 是严格单调函数的情况，写出一般的结果.

定理　设随机变量 X 具有概率密度 $f_X(x)$，$-\infty<x<\infty$，又设函数 $g(x)$ 处处可导且恒有 $g'(x)>0$（或恒有 $g'(x)<0$），则 $Y=g(X)$ 是连续型随机变量，其概率密度为

$$f_Y(y)=\begin{cases}f_X[h(y)]|h'(y)|, & \alpha<y<\beta,\\ 0, & \text{其他},\end{cases}\tag{5.2}$$

其中 $\alpha=\min\{g(-\infty),g(\infty)\}$，$\beta=\max\{g(-\infty),g(\infty)\}$，$h(y)$ 是 $g(x)$ 的反函数.

我们只证 $g'(x)>0$ 的情况. 此时 $g(x)$ 在 $(-\infty,\infty)$ 严格单调增加，它的反函数 $h(y)$ 存在，且在 (α,β) 严格单调增加，可导. 分别记 X,Y 的分布函数为 $F_X(x)$，$F_Y(y)$. 现在先来求 Y 的分布函数 $F_Y(y)$.

因为 $Y=g(X)$ 在 (α,β) 取值，故当 $y\leqslant\alpha$ 时 $F_Y(y)=P\{Y\leqslant y\}=0$；当 $y\geqslant\beta$ 时，$F_Y(y)=P\{Y\leqslant y\}=1$.

当 $\alpha<y<\beta$ 时，

$$\begin{aligned}F_Y(y)&=P\{Y\leqslant y\}=P\{g(X)\leqslant y\}\\&=P\{X\leqslant h(y)\}=F_X[h(y)].\end{aligned}$$

将 $F_Y(Y)$ 关于 y 求导数，即得 Y 的概率密度

$$f_Y(y)=\begin{cases}f_X[h(y)]h'(y), & \alpha<y<\beta,\\ 0, & \text{其他}.\end{cases}\tag{5.3}$$

对于 $g'(x)<0$ 的情况可以同样地证明，此时有

$$f_Y(y)=\begin{cases}f_X[h(y)][-h'(y)], & \alpha<y<\beta,\\ 0, & \text{其他}.\end{cases}\tag{5.4}$$

① 连续型随机变量 X 的函数 $Y=g(X)$ 不一定是连续型的随机变量.

合并(5.3)与(5.4)两式,(5.2)得证. □

若 $f(x)$ 在有限区间 $[a,b]$ 以外等于零,则只需假设在 $[a,b]$ 上恒有 $g'(x)>0$(或恒有 $g'(x)<0$),此时

$$\alpha=\min\{g(a),g(b)\},\beta=\max\{g(a),g(b)\}.$$

例 4 设随机变量 $X\sim N(\mu,\sigma^2)$. 试证明 X 的线性函数 $Y=aX+b\ (a\neq 0)$ 也服从正态分布.

证 X 的概率密度为

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<\infty.$$

现在 $y=g(x)=ax+b$,由这一式子解得

$$x=h(y)=\frac{y-b}{a},\text{且有 } h'(y)=\frac{1}{a}.$$

由(5.2)式得 $Y=aX+b$ 的概率密度为

$$f_Y(y)=\frac{1}{|a|}f_X\left(\frac{y-b}{a}\right),\quad -\infty<y<\infty.$$

即

$$f_Y(y)=\frac{1}{|a|}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}}$$

$$=\frac{1}{|a|\sigma\sqrt{2\pi}}e^{-\frac{[y-(b+a\mu)]^2}{2(a\sigma)^2}},\quad -\infty<y<\infty.$$

即有 $\quad Y=aX+b\sim N(a\mu+b,(a\sigma)^2)$.

特别,在上例中取 $a=\frac{1}{\sigma},b=-\frac{\mu}{\sigma}$ 得

$$Y=\frac{X-\mu}{\sigma}\sim N(0,1).$$

这就是上一节引理的结果. □

例 5 设电压 $V=A\sin\Theta$,其中 A 是一个已知的正常数,相角 Θ 是一个随机变量,且有 $\Theta\sim U\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$,试求电压 V 的概率密度.

解 现在 $v=g(\theta)=A\sin\theta$ 在 $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ 上恒有 $g'(\theta)=A\cos\theta>0$,且有反函数

$$\theta=h(v)=\arcsin\frac{v}{A},\quad h'(v)=\frac{1}{\sqrt{A^2-v^2}},$$

又,Θ 的概率密度为

$$f(\theta)=\begin{cases}\frac{1}{\pi}, & -\frac{\pi}{2}<\theta<\frac{\pi}{2},\\ 0, & \text{其他}.\end{cases}$$

由(5.2)式得 $V=A\sin\Theta$ 的概率密度为

$$\psi(v)=\begin{cases}\dfrac{1}{\pi}\cdot\dfrac{1}{\sqrt{A^2-v^2}}, & -A<v<A,\\ 0, & \text{其他}.\end{cases}$$ □

若在上题中 $\Theta\sim U(0,\pi)$,因为此时 $v=g(\theta)=A\sin\theta$ 在 $(0,\pi)$ 上不是单调函数,上述定理失效,应仍按例 3 的方法来做. 请读者自行求出其结果.

小结

随机变量 $X=X(e)$ 是定义在样本空间 $S=\{e\}$ 上的实值单值函数. 也就是说,它是随机试验结果的函数. 它的取值随试验的结果而定,是不能预先确定的,它的取值有一定的概率. 随机变量的引入,使概率论的研究由个别随机事件扩大为随机变量所表征的随机现象的研究. 今后,我们主要研究随机变量和它的分布.

一个随机变量,如果它所有可能的值是有限个或可列无限个,这种随机变量称为离散型随机变量,不是这种情况则称为非离散型的. 不论是离散型的或非离散型的随机变量 X,都可以借助分布函数

$$F(x)=P\{X\leqslant x\},\quad -\infty<x<\infty$$

来描述. 若已知随机变量 X 的分布函数,就能知道 X 落在任一区间 $(x_1,x_2]$ 上的概率

$$P\{x_1<X\leqslant x_2\}=F(x_2)-F(x_1),\quad x_1<x_2.$$

这样,分布函数就能完整地描述随机变量取值的统计规律性.

对于离散型随机变量,我们需要掌握的是它可能取哪些值,以及它以怎样的概率取这些值,这就是离散型随机变量取值的统计规律性. 因而,对于离散型随机变量,用分布律

$$P\{X=x_k\}=p_k,\quad k=1,2,\cdots$$

或写成

X	x_1	x_2	$\cdots$	x_k	$\cdots$
p_k	p_1	p_2	$\cdots$	p_k	$\cdots$

(这里,$\sum\limits_{k=1}^{\infty}p_k=1$)来描述它的取值的统计规律性较为直观和简洁. 分布律与分布函数有以下的关系

$$F(x)=P\{X\leqslant x\}=\sum_{x_k\leqslant x}P\{X=x_k\},$$

它们是一一对应的.

设随机变量 X 的分布函数为 $F(x)$,如果存在非负可积函数 $f(x)$,使得对于任意 x,有

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

则称 X 是连续型随机变量,其中 $f(x)\geqslant 0$ 称为 X 的概率密度.

给定 X 的概率密度 $f(x)$ 就能确定 $F(x)$,由于 $f(x)$ 位于积分号之内,故改变 $f(x)$ 在个别点上的函数值并不改变 $F(x)$ 的值. 因此,改变 $f(x)$ 在个别点上的值,是无关紧要的.

连续型随机变量 X 的分布函数是连续的,连续型随机变量取任一指定实数值 a 的概率

为 0，即 $P\{X=a\}=0$. 这两点性质离散型随机变量是不具备的.

我们将随机变量分成为

$$\text{随机变量}\begin{cases}\text{离散型}\\ \text{非离散型}\begin{cases}\text{连续型}\\ \text{其他}\end{cases}\end{cases}$$

读者不要误以为，一个随机变量，如果它不是离散型的那一定是连续型的. 但本书只讨论两类重要的随机变量：离散型和连续型随机变量.

读者应掌握分布函数、分布律、概率密度的性质. 本章引入了几种重要的随机变量的分布：(0－1)分布，二项分布，泊松分布，指数分布，均匀分布和正态分布. 读者必须熟知这几种随机变量的分布律或概率密度.

随机变量 X 的函数 $Y=g(X)$ 也是一个随机变量，要掌握如何由已知的 X 的分布(X 的分布律或概率密度)去求得 $Y=g(X)$ 的分布(Y 的分布律或概率密度).

■ 重要术语及主题

随机变量　分布函数　离散型随机变量及其分布律　连续型随机变量及其概率密度　伯努利试验　(0－1)分布　n 重伯努利试验　二项分布　泊松分布　指数分布　均匀分布　正态分布　随机变量函数的分布

习题

1. 考虑为期一年的一张保险单，若投保人在投保后一年内因意外死亡，则公司赔付 20 万元，若投保人因其他原因死亡，则公司赔付 5 万元，若投保人在投保期末生存，则公司无需付给任何费用. 若投保人在一年内因意外死亡的概率为 0.0002，因其他原因死亡的概率为 0.0010，求公司赔付金额的分布律.

2. (1) 一袋中装有 5 只球，编号为 1,2,3,4,5. 在袋中同时取 3 只，以 X 表示取出的 3 只球中的最大号码，写出随机变量 X 的分布律.

(2) 将一颗骰子抛掷两次，以 X 表示两次中得到的小的点数，试求 X 的分布律.

3. 设在 15 只同类型的零件中有 2 只是次品，在其中取 3 次，每次任取 1 只，作不放回抽样. 以 X 表示取出的次品的只数.

(1) 求 X 的分布律.

(2) 画出分布律的图形.

4. 进行重复独立试验，设每次试验的成功概率为 p，失败概率为 $q=1-p$ $(0<p<1)$.

(1) 将试验进行到出现一次成功为止，以 X 表示所需的试验次数，求 X 的分布律.(此时称 X 服从以 p 为参数的**几何分布**.)

(2) 将试验进行到出现 r 次成功为止，以 Y 表示所需的试验次数，求 Y 的分布律.(此时称 Y 服从以 r,p 为参数的**巴斯卡分布或负二项分布**.)

(3) 一篮球运动员的投篮命中率为 45%. 以 X 表示他首次投中时累计已投篮的次数，写出 X 的分布律，并计算 X 取偶数的概率.

5. 一房间有 3 扇同样大小的窗子，其中只有一扇是打开的. 有一只鸟自开着的窗子飞入了房间，它只能从开着的窗子飞出去. 鸟在房子里飞来飞去，试图飞出房间. 假定鸟是没有记

忆的，它飞向各扇窗子是随机的.

(1) 以 X 表示鸟为了飞出房间试飞的次数，求 X 的分布律.

(2) 户主声称，他养的一只鸟是有记忆的，它飞向任一窗子的尝试不多于一次. 以 Y 表示这只聪明的鸟为了飞出房间试飞的次数. 如户主所说是确实的，试求 Y 的分布律.

(3) 求试飞次数 X 小于 Y 的概率和试飞次数 Y 小于 X 的概率.

6. 一大楼装有 5 台同类型的供水设备. 设各台设备是否被使用相互独立. 调查表明在任一时刻 t 每台设备被使用的概率为 0.1，问在同一时刻，

(1) 恰有 2 台设备被使用的概率是多少？

(2) 至少有 3 台设备被使用的概率是多少？

(3) 至多有 3 台设备被使用的概率是多少？

(4) 至少有 1 台设备被使用的概率是多少？

7. 设事件 A 在每次试验发生的概率为 0.3. A 发生不少于 3 次时，指示灯发出信号.

(1) 进行了 5 次重复独立试验，求指示灯发出信号的概率.

(2) 进行了 7 次重复独立试验，求指示灯发出信号的概率.

8. 甲、乙两人投篮，投中的概率分别为 0.6，0.7. 今各投 3 次. 求

(1) 两人投中次数相等的概率；

(2) 甲比乙投中次数多的概率.

9. 有一大批产品，其验收方案如下，先作第一次检验：从中任取 10 件，经检验无次品接受这批产品，次品数大于 2 拒收；否则作第二次检验，其做法是从中再任取 5 件，仅当 5 件中无次品时接受这批产品. 若产品的次品率为 10%，求

(1) 这批产品经第一次检验就能接受的概率.

(2) 需作第二次检验的概率.

(3) 这批产品按第二次检验的标准被接受的概率.

(4) 这批产品在第一次检验未能作决定且第二次检验时被通过的概率.

(5) 这批产品被接受的概率.

10. 有甲、乙两种味道和颜色都极为相似的名酒各 4 杯. 如果从中挑 4 杯，能将甲种酒全部挑出来，算是试验成功一次.

(1) 某人随机地去猜，问他试验成功一次的概率是多少？

(2) 某人声称他通过品尝能区分两种酒. 他连续试验 10 次，成功 3 次. 试推断他是猜对的，还是他确有区分的能力(设各次试验是相互独立的).

11. 尽管在几何教科书中已经讲过仅用圆规和直尺三等分一个任意角是不可能的，但每一年总是有一些"发明者"撰写关于仅用圆规和直尺将角三等分的文章. 设某地区每年撰写此类文章的篇数 X 服从参数为 6 的泊松分布. 求明年没有此类文章的概率.

12. 一电话总机每分钟收到呼唤的次数服从参数为 4 的泊松分布. 求

(1) 某一分钟恰有 8 次呼唤的概率；

(2) 某一分钟的呼唤次数大于 3 的概率.

13. 某一公安局在长度为 t 的时间间隔内收到的紧急呼救的次数 X 服从参数为 $(1/2)t$ 的泊松分布，而与时间间隔的起点无关(时间以小时计).

(1) 求某一天中午 12 时至下午 3 时未收到紧急呼救的概率.

(2) 求某一天中午 12 时至下午 5 时至少收到 1 次紧急呼救的概率.

14. 某人家中在时间间隔 t(小时)内接到电话的次数 X 服从参数为 $2t$ 的泊松分布.

(1) 若他外出计划用时 10 分钟,问其间有电话铃响一次的概率是多少?

(2) 若他希望外出时没有电话的概率至少为 0.5,问他外出应控制最长时间是多少?

15. 保险公司在一天内承保了 5000 张相同年龄,为期一年的寿险保单,每人一份.在合同有效期内若投保人死亡,则公司需赔付 3 万元.设在一年内,该年龄段的死亡率为 0.0015,且各投保人是否死亡相互独立.求该公司对于这批投保人的赔付总额不超过 30 万元的概率(利用泊松定理计算).

16. 有一繁忙的汽车站,每天有大量汽车通过,设一辆汽车在一天的某段时间内出事故的概率为 0.0001.在某天的该时间段内有 1000 辆汽车通过.问出事故的车辆数不小于 2 的概率是多少?(利用泊松定理计算)

17. (1) 设 X 服从(0-1)分布,其分布律为 $P\{X=k\}=p^k(1-p)^{1-k}, k=0,1$,求 X 的分布函数,并作出其图形.

(2) 求第 2 题(1)中的随机变量的分布函数.

18. 在区间$[0,a]$上任意投掷一个质点,以 X 表示这个质点的坐标.设这个质点落在$[0,a]$中任意小区间内的概率与这个小区间的长度成正比例.试求 X 的分布函数.

19. 以 X 表示某商店从早晨开始营业起直到第一个顾客到达的等待时间(以分计),X 的分布函数是

$$F_X(x)=\begin{cases}1-\mathrm{e}^{-0.4x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

求下述概率:

(1) $P\{$至多 3 分钟$\}$.

(2) $P\{$至少 4 分钟$\}$.

(3) $P\{$3 分钟至 4 分钟之间$\}$.

(4) $P\{$至多 3 分钟或至少 4 分钟$\}$.

(5) $P\{$恰好 2.5 分钟$\}$.

20. 设随机变量 X 的分布函数为

$$F_X(x)=\begin{cases}0, & x<1,\\ \ln x, & 1\leqslant x<\mathrm{e},\\ 1, & x\geqslant \mathrm{e}.\end{cases}$$

(1) 求 $P\{X<2\}, P\{0<X\leqslant 3\}, P\{2<X<5/2\}$.

(2) 求概率密度 $f_X(x)$.

21. 设随机变量 X 的概率密度为

(1) $f(x)=\begin{cases}2(1-1/x^2), & 1\leqslant x\leqslant 2,\\ 0, & \text{其他}.\end{cases}$

(2) $f(x)=\begin{cases}x, & 0\leqslant x<1,\\ 2-x, & 1\leqslant x<2,\\ 0, & \text{其他}.\end{cases}$

求 X 的分布函数 $F(x)$，并画出(2)中的 $f(x)$ 及 $F(x)$ 的图形.

22. (1) 分子运动速度的绝对值 X 服从麦克斯韦(Maxwell)分布，其概率密度为

$$f(x)=\begin{cases}Ax^2e^{-x^2/b}, & x>0,\\ 0, & \text{其他},\end{cases}$$

其中 $b=m/(2kT)$，k 为玻耳兹曼(Boltzmann)常数，T 为绝对温度，m 是分子的质量，试确定常数 A.

(2) 研究了英格兰在 1875 年～1951 年期间，在矿山发生导致不少于 10 人死亡的事故的频繁程度，得知相继两次事故之间的时间 T(日)服从指数分布，其概率密度为

$$f_T(t)=\begin{cases}\dfrac{1}{241}e^{-t/241}, & t>0,\\ 0, & \text{其他}.\end{cases}$$

求分布函数 $F_T(t)$，并求概率 $P\{50<T<100\}$.

23. 某种型号器件的寿命 X(以小时计)具有概率密度

$$f(x)=\begin{cases}\dfrac{1000}{x^2}, & x>1000,\\ 0, & \text{其他}.\end{cases}$$

现有一大批此种器件(设各器件损坏与否相互独立)，任取 5 只，问其中至少有 2 只寿命大于 1500 小时的概率是多少?

24. 设顾客在某银行的窗口等待服务的时间 X(min)服从指数分布，其概率密度为

$$f_X(x)=\begin{cases}\dfrac{1}{5}e^{-x/5}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

某顾客在窗口等待服务，若超过 10 min，他就离开. 他一个月要到银行 5 次. 以 Y 表示一个月内他未等到服务而离开窗口的次数. 写出 Y 的分布律，并求 $P\{Y\geqslant 1\}$.

25. 设 K 在(0,5)服从均匀分布，求 x 的方程

$$4x^2+4Kx+K+2=0$$

有实根的概率.

26. 设 $X\sim N(3,2^2)$.

(1) 求 $P\{2<X\leqslant 5\}$，$P\{-4<X\leqslant 10\}$，$P\{|X|>2\}$，$P\{X>3\}$.

(2) 确定 c，使得 $P\{X>c\}=P\{X\leqslant c\}$.

(3) 设 d 满足 $P\{X>d\}\geqslant 0.9$，问 d 至多为多少?

27. 某地区 18 岁的女青年的血压(收缩压，以 mmHg 计)服从 $N(110,12^2)$ 分布. 在该地区任选一 18 岁的女青年，测量她的血压 X. 求

(1) $P\{X\leqslant 105\}$，$P\{100<X\leqslant 120\}$;

(2) 确定最小的 x，使 $P\{X>x\}\leqslant 0.05$.

28. 由某机器生产的螺栓的长度(cm)服从参数 $\mu=10.05$，$\sigma=0.06$ 的正态分布. 规定长度在范围 10.05 ± 0.12 内为合格品，求一螺栓为不合格品的概率.

29. 一工厂生产的某种元件的寿命 X(以小时计)服从参数为 $\mu=160$，$\sigma(\sigma>0)$ 的正态分布. 若要求 $P\{120<X\leqslant 200\}\geqslant 0.80$，允许 σ 最大为多少?

30. 设在一电路中，电阻两端的电压(V)服从 $N(120,2^2)$，今独立测量了 5 次，试确定有 2 次测定值落在区间[118,122]之外的概率.

31. 某人上班，自家里去办公楼要经过一交通指示灯，这一指示灯有 80%时间亮红灯，此时他在指示灯旁等待直至绿灯亮. 等待时间在区间[0,30](以秒计)服从均匀分布. 以 X 表示他的等待时间，求 X 的分布函数$F(x)$. 画出 $F(x)$的图形，并问 X 是否为连续型随机变量，是否为离散型的?(要说明理由)

32. 设 $f(x)$，$g(x)$都是概率密度函数，求证

$$h(x)=\alpha f(x)+(1-\alpha)g(x),\quad 0\leqslant\alpha\leqslant1$$

也是一个概率密度函数.

33. 设随机变量 X 的分布律为

X	-2	-1	0	1	3
p_k	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{11}{30}$

求 $Y=X^2$ 的分布律.

34. 设随机变量 X 在区间(0,1)服从均匀分布.

(1) 求 $Y=e^X$ 的概率密度.

(2) 求 $Y=-2\ln X$ 的概率密度.

35. 设 $X\sim N(0,1)$.

(1) 求 $Y=e^X$ 的概率密度.

(2) 求$Y=2X^2+1$的概率密度.

(3) 求 $Y=|X|$的概率密度.

36. (1) 设随机变量 X 的概率密度为 $f(x)$，$-\infty<x<\infty$. 求 $Y=X^3$ 的概率密度.

(2) 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}e^{-x}, & x>0,\\ 0, & \text{其他}.\end{cases}$$

求 $Y=X^2$ 的概率密度.

37. 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{2x}{\pi^2}, & 0<x<\pi,\\ 0, & \text{其他}.\end{cases}$$

求 $Y=\sin X$ 的概率密度.

38. 设电流 I 是一个随机变量，它均匀分布在 9 A～11 A 之间. 若此电流通过 2 Ω 的电阻，在其上消耗的功率 $W=2I^2$. 求 W 的概率密度.

39. 某物体的温度 T(°F)是随机变量，且有 $T\sim N(98.6,2)$，已知 $\Theta=\frac{5}{9}(T-32)$，试求 Θ(℃)的概率密度.