第二章

随机变量及其分布

在第一章研究随机试验时，只是孤立地考虑每个随机事件的概率，研究方法缺乏一般性且无法用高等数学的工具加以研究. 从本章开始，我们引入随机变量及其分布的概念. 随机变量概念的建立是概率论发展史上的重大突破。对于随机变量的分布函数，我们能够用微积分为工具进行研究，强有力的高等数学的工具大大增强了我们研究随机现象的手段，从而使概率论的发展进入了一个新的阶段.

第一节　随机变量及其分布函数

一、随机变量

在第一章中，随机事件通过样本点所构成的集合来表示，并运用初等数学的方法求出一些事件的概率，但这种表示方式大大限制了计算概率的方法. 我们希望将所有随机事件的集合统一成数集，这就需要将随机试验所产生的样本点与实数对应. 有些试验的样本点本身就与数值有关. 例如，取出的产品中的废品数；掷一颗骰子，观察其出现的点数；观察一个元件的使用寿命；等等. 但也有一些试验的结果与数值无关，我们可以将其量化. 例如，抛掷一枚硬币，可以将“出现正面”记为 1，“出现反面”记为 0；取一件产品，当取到次品时记为 1，当取到正品时记为 0. 这样一来，对任意的随机试验，即一种对应关系，可以给出如下定义：

定义 2.1　设随机试验的样本空间为 $\Omega=\{\omega\}$，若对每一个样本点 ω 有唯一的实数 $X(\omega)$ 与之对应，则称这个定义在 Ω 上的单值实函数 $X=X(\omega)$ 为随机变量，简记为 X.

随机变量一般用大写的字母 X、Y、Z 或希腊字母 ξ、η、ζ 来表示.

引入随机变量以后，随机试验所产生的任意事件都可以通过随机变量的取值表达出来.

例如，设 X 表示电话总机在单位时间内接到呼唤的次数，则样本空间 $\Omega=\{X=0,1,2,\cdots\}$，事件“接到至少 1 次呼唤”可用$\{X\geqslant 1\}$来表示，事件“恰好接到 3 次呼唤”可用$\{X=3\}$来表示. 反之，若 X 表示一个随机变量，则 X 的任意取值所形成的集合也都表示一个随机事件，如$\{X\geqslant a\}$、$\{X=a\}$、$\{a<X\leqslant b\}$(其中 a、b 为任意实数)等.

这样，对于随机事件概率的研究就可以转化为对随机变量取值的概率的研究，这使我们有可能用微积分的方法对各种有关的问题进行深入的研究，使我们的研究更为方便.

二、分布函数

对于随机变量 X，一方面要明确其各种可能的取值及取值范围，另一方面要掌握其取值的概率分布. 那如何来描写其取值的概率分布呢？在后面的讨论中经常需要考虑随机变量的取值落在某个区间上的概率. 而常数 $a<b$ 有

$$\{a<X\leqslant b\}=\{X\leqslant b\}-\{X\leqslant a\},$$

且$\{X\leqslant a\}\subset\{X\leqslant b\}$，由概率的差事件的运算性质可得

$$P(a<X\leqslant b)=P(X\leqslant b)-P(X\leqslant a).$$

因此，对于随机变量 X，对任意实数 x 只要知道 $P(X\leqslant x)$，就可以求出 X 的取值落在任意区间$(a,b]$上的概率. 为此，我们给出随机变量分布函数的概念.

定义 2.2 设 X 是一个随机变量，对任意的 $x\in(-\infty,+\infty)$，称函数

$$F(x)=P(X\leqslant x)$$

为随机变量 X 的分布函数.

由定义可知，$F(x)$是一个定义在$(-\infty,+\infty)$上、值域为$[0,1]$的函数，且对任意 $a<b$，有

$$P(a<X\leqslant b)=F(b)-F(a),P(X\leqslant a)=F(a),P(X>a)=1-F(a).$$

分布函数 $F(x)$是事件$\{X\leqslant x\}$的概率，也是 x 的一个普通函数，因而通过分布函数我们就可以用高等数学的方法来研究随机变量. 如果将 X 看成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 $F(x)$具有如下性质：

定理 2.1 设 $F(x)$是随机变量 X 的分布函数，则

(1)有界性：$0\leqslant F(x)\leqslant 1,x\in(-\infty,+\infty)$；

(2)单调非降性：若 $x_1<x_2$，则 $F(x_1)\leqslant F(x_2)$；

(3)右连续性：对任意实数 x_0，$F(x)$在 x_0 点处右连续，即 $\lim\limits_{x\to x_0^+}F(x)=F(x_0)$；

(4)规范性：$\lim\limits_{x\to+\infty}F(x)=1$，$\lim\limits_{x\to-\infty}F(x)=0$.

证明 (1)由概率的有界性可知.

(2)若 $x_1<x_2$,则$\{X\leqslant x_1\}\subset\{X\leqslant x_2\}$,由概率的单调性可知 $P(X\leqslant x_1)\leqslant P(X\leqslant x_2)$,即 $F(x_1)\leqslant F(x_2)$.

(3)该项证明已超出本书范围,从略.

(4)该项的严格证明已超出本书的范围,但可以做一个直观的说明.当 $x\to+\infty$时,事件$\{X\leqslant x\}$趋于必然事件,从而其概率 $F(x)$趋于必然事件的概率 1.同理,当 $x\to-\infty$时,事件$\{X\leqslant x\}$趋于不可能事件,从而 $F(x)$趋于 0.

反之,任意一个满足以上这四条性质的函数,一定可以作为某个随机变量的分布函数.

例 2-1 设随机变量 X 的分布函数 $F(x)=A+B\arctan x, x\in(-\infty,+\infty)$,求:

(1)常数 A、B;

(2)$P(-1<X\leqslant1)$.

解 (1)由分布函数的规范性可得

$$\lim_{x\to+\infty}F(x)=\lim_{x\to+\infty}(A+B\arctan x)=A+\frac{\pi}{2}B=1;$$

$$\lim_{x\to-\infty}F(x)=\lim_{x\to-\infty}(A+B\arctan x)=A-\frac{\pi}{2}B=0,$$

解得

$$A=\frac{1}{2}, B=\frac{1}{\pi}.$$

(2)$P(-1<X\leqslant1)=F(1)-F(-1)=\left(\frac{1}{2}+\frac{1}{\pi}\arctan 1\right)-\left[\frac{1}{2}+\frac{1}{\pi}\arctan(-1)\right]=\frac{1}{2}.$

第二节　离散型随机变量及其分布

按照随机变量的取值情况,可以将随机变量分为离散型随机变量和非离散型随机变量.非离散型随机变量包括的范围比较广,情况比较复杂,其中最重要也最常用到的是连续型随机变量.因此,这里只讨论离散型随机变量和连续型随机变量及它们的分布.这一节中,我们先来研究离散型随机变量及其分布.

一、离散型随机变量及其分布列

定义 2.3 若随机变量 X 只能取到有限个可能值或至多可列个可能值,则称 X 为**离散型随机变量**.设 X 所有可能的取值为 $x_1, x_2, \cdots, x_i, \cdots$,则称

$$P(X=x_i)=p_i, i=1,2,3\cdots$$

为随机变量 X 的**概率分布列**，简称为**分布列**.

离散型随机变量分布列也可用如下表的形式表示：

X	x_1	x_2	…	x_i	…
P	p_1	p_2	…	p_i	…

其中第一行是 X 所有可能的取值，第二行是 X 取相应值时的概率.

离散型随机变量分布列中的数列 $\{p_i\}$ 应满足以下两条**基本性质**：

(1)非负性：$p_i \geqslant 0, i=1,2,\cdots$；

(2)规范性：$\sum\limits_{i=1}^{\infty} p_i = 1$.

反之，如果一个数列满足以上两条基本性质，则该数列就可以作为某个离散型随机变量的分布列.

如果已知离散型随机变量 X 的分布列为 $P(X=x_i)=p_i, i=1,2,\cdots$，则事件 $\{X=x_i\}$，$i=1,2,\cdots$ 两两互不相容，所以 X 的取值落入任意一个集合 A 的概率，根据可加性为

$$P(X\in A)=\sum_{x_i\in A} P(X=x_i).$$

特别地，可给出由分布列求分布函数的计算公式：

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

显然，分布列不仅给出了离散型随机变量的所有取值，同时也给出了取每一个值时的概率.

例 2-2 设离散型随机变量 X 的分布列为 $P(X=k)=\dfrac{c}{3^k}, k=1,2,\cdots$，试求：

(1)常数 c；

(2)$P(X$ 为偶数$)$.

解 (1)由分布列的规范性，有

$$\sum_{k=1}^{\infty} P(X=k)=\sum_{k=1}^{\infty}\frac{c}{3^k}=c\cdot\frac{\frac{1}{3}}{1-\frac{1}{3}}=\frac{c}{2}=1,$$

所以常数 $c=2$.

(2)$P(X$ 为偶数$)=P(X=2)+P(X=4)+\cdots+P(X=2k)+\cdots$

$$=\frac{2}{3^2}+\frac{2}{3^4}+\cdots+\frac{2}{3^{2k}}+\cdots=2\times\frac{\frac{1}{9}}{1-\frac{1}{9}}=\frac{1}{4}.$$

例 2-3 袋中有 2 个白球和 3 个黑球，每次从中任取一球，直到取出白球为止. 设随机变

量 X 表示取球次数，试求：

(1)X 的分布列；

(2)X 的分布函数.

解 (1)由题意知，X 可取的值为 1、2、3、4.

$$P(X=1)=\frac{2}{5}=0.4,$$

$$P(X=2)=\frac{3\times 2}{5\times 4}=\frac{3}{10}=0.3,$$

$$P(X=3)=\frac{3\times 2\times 2}{5\times 4\times 3}=\frac{1}{5}=0.2,$$

$$P(X=4)=\frac{3\times 2\times 1\times 2}{5\times 4\times 3\times 2}=\frac{1}{10}=0.1.$$

(2)由 X 的分布函数 $F(X)=P(X\leqslant x)=\sum\limits_{k\leqslant x}P(X=k)$ 得：

当 $x<1$ 时，事件 $\{X\leqslant x\}$ 中不含 X 的任何可能取值，所以 $F(x)=0$；

当 $1\leqslant x<2$ 时，事件 $\{X\leqslant x\}$ 中只含有 $\{X=1\}$ 这个事件，所以 $F(x)=P(X=1)=0.4$.
同理可得：

当 $2\leqslant x<3$ 时，$F(x)=P(X=1)+P(X=2)=0.7$；

当 $3\leqslant x<4$ 时，$F(x)=P(X=1)+P(X=2)+P(X=3)=0.9$；

当 $x\geqslant 4$ 时，$F(x)=P(X=1)+P(X=2)+P(X=3)+P(X=4)=1$.

综上，得 X 的分布函数为

$$F(x)=\begin{cases}0, & x<1,\\ 0.4, & 1\leqslant x<2,\\ 0.7, & 2\leqslant x<3,\\ 0.9, & 3\leqslant x<4,\\ 1, & x\geqslant 4.\end{cases}$$

其图形如图 2-1 所示，从图形中可以看出，离散型随机变量的分布函数为阶梯形分段函数，并且在每个分段点处都是右连续的.

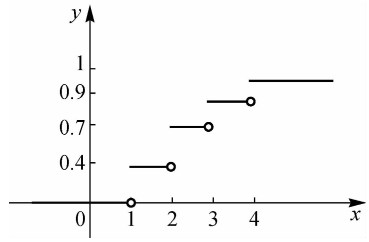

图 2-1

二、常见的离散型分布

1. (0—1)分布

定义 2.4 若随机变量 X 的分布列为

X	0	1
P	$1-p$	p

即

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

则称 X 服从**(0—1)分布或两点分布**.

显然,若一个随机试验只有两种可能的结果 A 和 $\bar{A}$,且 $P(A)=p(0<p<1)$,则一定可以定义一个服从(0—1)分布的随机变量来描述这个随机试验的结果.

在概率论中,像这样只考虑两个可能结果的随机试验,称为**伯努利试验**;将同一个伯努利试验独立的重复 n 次的随机试验称为 ***n* 重伯努利试验**. n 重伯努利试验是一种很重要的概率模型,它有着广泛的应用. 下面就来介绍一个在 n 重伯努利试验中产生的离散型分布.

2. 二项分布

假设一个 n 重伯努利试验,每次试验中事件 A 出现的概率 $P(A)=p(0<p<1)$,以 X 表示 n 重伯努利试验中事件 A 出现的次数. 下面来求它的分布列.

显然,X 所有可能取值为 $0,1,2,\cdots,n$. 事件$\{X=0\}$表示 n 次试验中都是 $\bar{A}$ 发生,由于 n 次试验相互独立,故

$$P\{X=0\}=P\{\underbrace{\bar{A}\bar{A}\cdots\bar{A}}_{n}\}=[P(\bar{A})]^n=(1-p)^n,$$

事件$\{X=1\}$表示 n 次试验中事件 A 出现 1 次,$\bar{A}$ 出现 $n-1$ 次,故

$$\begin{aligned}P(X=1)&=\underbrace{P(A\underbrace{\bar{A}\cdots\bar{A}}_{n-1})+P(\bar{A}A\bar{A}\cdots\bar{A})+\cdots+P(\underbrace{\bar{A}\bar{A}\cdots\bar{A}}_{n-1}A)}_{n}\\&=\underbrace{p(1-p)^{n-1}+p(1-p)^{n-1}+\cdots+p(1-p)^{n-1}}_{n}\\&=np(1-p)^{n-1}=\mathrm{C}_n^1p(1-p)^{n-1}.\end{aligned}$$

同理,$\{X=k\}$表示 n 次试验中事件 A 出现 k 次,$\bar{A}$ 出现 $n-k$ 次. 例如,前 k 次事件 A 发生,而后 $n-k$ 次 $\bar{A}$ 发生,此时

$$P(\underbrace{A\cdots A}_{k}\underbrace{\bar{A}\cdots\bar{A}}_{n-k})=p^k(1-p)^{n-k},$$

显然,事件$\{X=k\}$中包含 C_n^k 种这种形式,故

$$P(X=k)=C_n^k p^k(1-p)^{n-k},$$

所以,随机变量 X 的分布列为

$$P(X=k)=C_n^k p^k(1-p)^{n-k},k=0,1,2,\cdots,n.$$

一般地,有如下定义:

定义 2.5 若随机变量 X 的分布列为

$$P(X=k)=C_n^k p^k(1-p)^{n-k},k=0,1,2,\cdots,n,$$

其中 $0<p<1$ 为常数,则称 X 服从参数为 n、p 的二项分布,记为 $X\sim B(n,p)$.

特别地,当 $n=1$ 时,二项分布化为

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

即为$(0-1)$分布.

例 2-4 已知 100 件产品中有 5 件次品,现从中有放回地取 3 次,每次取 1 件,以 X 记所取 3 件产品中的次品数.

(1)写出 X 的分布列;

(2)求在 3 件产品中恰好有 1 件次品的概率.

解 将抽取一次产品看作一次试验,这是一个三重伯努利试验.由题意,每次试验中取出次品的概率 $p=\frac{5}{100}=0.05$,所以 $X\sim B(3,0.05)$,于是

(1)X 的分布列为

$$P(X=k)=C_n^k 0.05^k(1-0.05)^{n-k}=C_3^k\,0.05^k\,0.95^{3-k},k=0,1,2,3.$$

(2)$P(X=1)=C_3^1 0.05\times0.95^2\approx0.135\ 4$.

若将本例中的“有放回抽样”改为“无放回抽样”,这就不是 n 重伯努利试验了,X 也就不服从二项分布,这时 X 服从后面将要讲的超几何分布.

但在实际问题中,真正能在完全相同条件下进行的试验并不多见.对于产品抽样问题,当产品的数量很大,而抽查的产品数量相对于产品总数来说很小时,“无放回”可以当作“有放回”来处理.

例 2-5 已知一大批产品的次品率为 0.1,从中任取 20 件产品,求 20 件产品中次品不超过 2 件的概率.

解 设 X 表示 20 件产品中次品的个数,则 $X\sim B(20,0.1)$,所以

$$
\begin{aligned}
P(X\leqslant 2)&=\sum_{k=0}^{2}\mathrm{C}_{20}^{k}\times 0.1^{k}\times 0.9^{20-k}\\
&=0.9^{20}+20\times 0.1\times 0.9^{19}+190\times 0.1^{2}\times 0.9^{18}\\
&\approx 0.667.
\end{aligned}
$$

在有些问题中，当伯努利试验的次数 n 很大，而每次试验中事件 A 出现的概率 p 很小时，计算是相当麻烦的. 为了简化计算，人们得到如下定理：

定理 2.2(泊松定理) n 重伯努利试验中，设事件 A 出现的概率为 p_n（与试验总数 n 有关），若 $\lim\limits_{n\to\infty}np_n=\lambda(\lambda>0)$，则有

$$\lim_{n\to\infty}\mathrm{C}_n^k p_n^k(1-p_n)^{n-k}=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},k=0,1,2,\cdots$$

证明 记 $np_n=\lambda_n$，则 $\lambda_n\to\lambda$，且 $p_n=\dfrac{\lambda_n}{n}$，所以

$$
\begin{aligned}
\mathrm{C}_n^k p_n^k(1-p_n)^{n-k}&=\frac{n!}{k!(n-k)!}\left(\frac{\lambda_n}{n}\right)^k\left(1-\frac{\lambda_n}{n}\right)^{n-k}\\
&=\frac{\lambda_n}{k!}\cdot\frac{n(n-1)\cdots(n-k+1)}{n^k}\left(1-\frac{\lambda_n}{n}\right)^n\left(1-\frac{\lambda_n}{n}\right)^{-k}.
\end{aligned}
$$

对于固定的 k，当 $n\to\infty$ 时，

$$\frac{n(n-1)\cdots(n-k+1)}{n^k}=\frac{n}{n}\cdot\frac{n-1}{n}\cdot\cdots\cdot\frac{n-k+1}{n}\to 1,$$

$$\left(1-\frac{\lambda_n}{n}\right)^{-k}\to 1,\left(1-\frac{\lambda_n}{n}\right)^{n}=\left[\left(1-\frac{\lambda_n}{n}\right)^{-\frac{n}{\lambda_n}}\right]^{-\lambda_n}\to\mathrm{e}^{-\lambda},$$

所以

$$\lim_{n\to\infty}\mathrm{C}_n^k p_n^k(1-p_n)^{n-k}=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},k=0,1,2,\cdots$$

利用级数 $\sum\limits_{k=0}^{\infty}\dfrac{x^k}{k!}=\mathrm{e}^x$，易知 $\sum\limits_{k=0}^{\infty}\dfrac{\lambda^k}{k!}\mathrm{e}^{-\lambda}=1$. 从而引入一种新的离散型分布，即泊松分布.

3. 泊松分布

定义 2.6 若随机变量 X 的分布列为

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

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

由定理 2.2 可知，泊松分布可以看作大量试验中小概率事件发生次数的概率分布的一个近似数学模型. 故如果 $X\sim B(n,p)$，当 n 很大而 p 很小时，二项分布可用泊松分布近似，即

$$P(X=k)=\mathrm{C}_n^k p_n^k(1-p_n)^{n-k}\approx\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda},\lambda=np,k=0,1,2,\cdots$$

泊松分布是一种重要的离散型分布，有着广泛的应用. 泊松分布常见于社会生活与具有

物理性质的问题中，如一段时间内电话局收到用户的呼叫次数、放射性物质放射的粒子数、车站内到达的乘客数等，都服从泊松分布.

为方便泊松分布的计算，有现成的分布表(见附表1)可供查阅.

例 2-6 某商店根据过去的销售记录知道某种商品每月的销售量 X 服从参数为 5 的泊松分布，为了以 98%以上的概率保证不脱销，问商店在月初应进货多少(假设上月无剩余)?

解 由题意知销售量 $X\sim P(5)$，设月初的进货为 N 件，故

$$P(X\leqslant N)\geqslant 0.98,$$

即

$$P(X>N)=P(X\geqslant N+1)\leqslant 0.02,$$

查附表 1 可得

$$P(X\geqslant 11)=0.013\ 695.$$

所以 $N+1\geqslant 11$，即 $N\geqslant 10$ 时，才能以 98%以上的概率保证不脱销.

例 2-7 设共有 2 000 个条件相同的人参加了人寿保险，每个参加保险的人一年交付保费 12 元，如果一年内死亡，则保险公司支付赔偿金 2 000 元，设每个参保人一年内死亡的概率为 0.002，求保险公司亏本的概率.

解 设一年内死亡的人数为 X，则 $X\sim B(2\ 000,0.002)$，故 X 近似地服从泊松分布，其参数 $\lambda=np=4$. 保险公司亏本的概率为

$$p=P(12\times 2\ 000-2\ 000X<0)=P(X>12)=P(X\geqslant 13).$$

查附表 1 可得 $p=0.000\ 274$.

例 2-8 设有 40 台同类型设备，各台工作是相互独立的，每台发生故障的概率为 0.01，且一台发生故障只需一人维修，考虑两种维修方法：方法一是由 2 人共同维护 40 台设备，方法二是 2 人分别负责 20 台. 问哪种维修方式使得设备发生故障时不能及时维修的概率较小.

解 按第一种方法，以 X 表示发生故障的设备数，显然 $X\sim B(40,0.01)$，故 X 近似地服从 $P(0.4)$. 当发生故障的台数超过维修工人数时，就不能及时维修，故

$$P(X>2)=P(X\geqslant 3)=0.007\ 9.$$

按第二种方法，以 X_1、X_2 分别表示两组发生故障的设备数，显然 $X_1\sim B(20,0.01)$和$X_2\sim B(20,0.01)$都近似地服从 $P(2)$. 记事件 A_i 表示“第 i 组发生故障不能及时维修”，$i=1,2$，则

$$P(A_1)=P(X_1>1)=P(X_1\geqslant 2)=0.017\ 5,$$

同理，$P(A_2)=0.017\ 5$.

不难看出，事件 A_1、A_2 相互独立，故按第二种方法发生故障不能及时维修的概率为

$$P(A_1\cup A_2)=1-P(\bar{A}_1\bar{A}_2)=1-P(\bar{A}_1)P(\bar{A}_2)\approx 0.034\ 7.$$

由此可看出，第一种方法使得发生故障而不能及时维修的概率要比第二种方法发生的概率要小，故第一种方法的维修效率更高.

4. 几何分布

设 X 是一个无穷次伯努利试验序列中事件 A 首次发生时所需的试验次数，显然 X 为一个离散型随机变量，如果每次试验中事件 A 出现的概率为 $p(0<p<1)$，则 X 的分布列为

$$P(X=k)=P(\underbrace{\bar{A}\cdots\bar{A}}_{k-1}A)=(1-p)^{k-1}p, k=1,2,\cdots.$$

定义 2.7 设随机变量 X 的分布列为

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

其中 $0<p<1$，则称 X 服从参数为 p 的**几何分布**，记为 $X\sim G(p)$.

例 2-9 设某人在求职过程中，每次求职成功的概率为 0.4，问该人至少要求职多少次，才能以 0.9 以上的把握获得一个就业机会.

解 设 X 表示首次求职成功所需的求职次数，显然 $X\sim G(0.4)$. 设该人共求职 N 次，则由题意得

$$P(X\leqslant N)\geqslant 0.9,$$

即

$$\begin{aligned}P(X=1)+P(X=2)+\cdots+P(X=N)&=0.4+0.4\times 0.6+\cdots+0.4\times 0.6^{N-1}\\&=0.4\times\frac{1-0.6^N}{0.4}=1-0.6^N,\end{aligned}$$

得 $N\geqslant 5$，故至少求职 5 次才能以 0.9 以上的把握获得一个就业机会.

5. 超几何分布

设共有 N 件产品，其中次品有 M 件，正品有 $N-M$ 件. 从中不放回地抽取 $n(n\leqslant N)$ 件产品，其中的次品数 X 是一个离散型随机变量，其分布列为

$$P(X=k)=\frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, k=0,1,2,\cdots,\min\{M,n\}.$$

具有此分布列的随机变量，称为服从参数 n、M、N 的超几何分布，记为 $H(n,M,N)$.

第三节　连续型随机变量及其分布

在第二节中研究的离散型随机变量取值只限于有限个或可列无穷多个. 但在很多随机试验中，如身高、候车时间、产品的使用寿命、测量误差等可以取某一区间或整个实数轴上所

有的值，这类随机变量称为连续型随机变量.

一、连续型随机变量及其概率密度函数

对于连续型随机变量，不能采用离散型随机变量的概率函数来描述，于是我们引入概率密度的概念.

考虑连续型随机变量 X，设分布函数为 $F(x)$，则其落在区间 $(x, x+\Delta x]$ 内的概率 $P(x<X<x+\Delta x)=F(x+\Delta x)-F(x)$，其中 x 为任意实数，$\Delta x>0$ 为区间长度，若极限

$$\lim_{\Delta x\to 0}\frac{P(x<X<x+\Delta x)}{\Delta x}=\lim_{\Delta x\to 0}\frac{F(x+\Delta x)-F(x)}{\Delta x}$$

存在，则这极限叫作 X 在 x 处的概率密度，记作 $f(x)$. 由上述极限可知概率密度 $f(x)$ 与分布函数 $F(x)$ 之间的关系为

(1) $F'(x)=f(x)$，即连续型随机变量的概率密度 $f(x)$ 是分布函数 $F(x)$ 的导函数；

(2) $F(x)=P(X\leqslant x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$.

由此我们给出连续型随机变量的定义.

定义 2.8 若对于随机变量 X 的分布函数 $F(x)$ 存在可积的非负函数 $f(x)$，则对于任意实数 x 有

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

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

由上述定义可知，每个连续型随机变量都与一个概率密度函数相对应. 由定义可知，概率密度函数 $f(x)$ 具有下列性质：

(1) 非负性：$f(x)\geqslant 0$；

(2) 规范性：$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$.

若一个函数满足上述两个性质，一定可以作为某个连续型随机变量的概率密度.

由定义及以上性质可知，概率密度函数曲线总是位于 x 轴上方，并且与 x 轴所围成图形的面积恒为 1（图 2-2）；对任意 $a\in(-\infty, +\infty)$，分布函数 $F(a)$ 的值等于密度函数曲线与 x 轴及直线 $x=a$ 所围左侧无限曲边梯形的面积（图 2-3 中阴影部分的面积）.

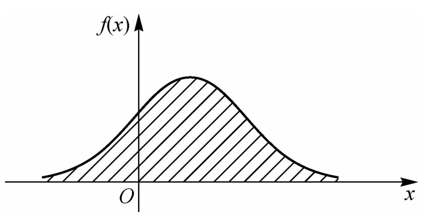

图 2-2

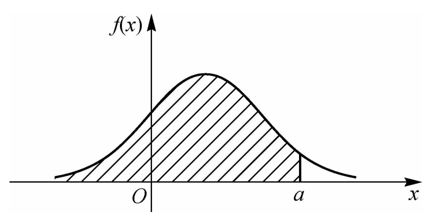

图 2-3

根据连续型随机变量的定义还可以给出以下性质：

(1)对任意的实数 a、$b(a<b)$，有

$$P(a<X\leqslant b)=F(b)-F(a)=\int_a^b f(x)\mathrm{d}x.$$

该性质由定义 2.8 及积分区间的可加性很容易得到，并且根据定积分的几何意义，该事件的概率等于区间$(a,b]$上概率密度函数曲线之下的曲面梯形的面积(图 2-4).

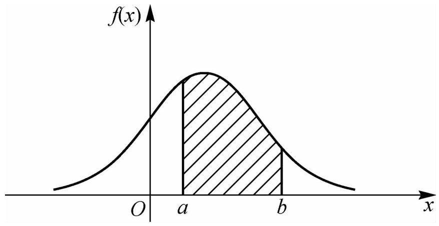

图 2-4

(2)连续型随机变量取任意固定值 a 的概率都为零，即 $P(X=a)=0$.

实际上，对任意 $a\in(-\infty,+\infty)$，$\varepsilon>0$，有

$$0\leqslant P(X=a)\leqslant P(a-\varepsilon<X<a+\varepsilon)=\int_{a-\varepsilon}^{a+\varepsilon} f(x)\mathrm{d}x,$$

而

$$\lim_{\varepsilon\to 0}\int_{a-\varepsilon}^{a+\varepsilon} f(x)\mathrm{d}x=0,$$

所以

$$P(X=a)=0.$$

这个性质说明在计算连续型随机变量落在某区间上的概率时，区间是否包含端点对事件的概率是没有影响的，即

$$\begin{aligned}P(a<X<b)&=P(a<X\leqslant b)=P(a\leqslant X<b)=P(a\leqslant X\leqslant b)\\&=F(b)-F(a)=\int_a^b f(x)\mathrm{d}x.\end{aligned}$$

由这个性质也可以看出，在概率论中，概率为 0 的事件不一定是不可能事件，因为连续型随机变量取任何一点都是有可能发生的. 同样，概率为 1 的事件也就不一定是必然事件.

(3)设 $F(x)$是连续型随机变量 X 的分布函数，则 $F(x)$是连续函数.

实际上，对任意的实数 x，有

$$\begin{aligned}\lim_{\Delta x\to 0}\Delta F&=\lim_{\Delta x\to 0}[F(x+\Delta x)-F(x)]\\&=\lim_{\Delta x\to 0}\int_x^{x+\Delta t} f(t)\mathrm{d}t=0.\end{aligned}$$

所以，$F(x)$是连续型随机变量.

这是连续型随机变量的一个重要性质.

(4)设 $F(x)$ 和 $f(x)$ 分别是连续型随机变量 X 的分布函数和密度函数,则由定义 2.8 可知,$F(x)$ 是 $f(x)$ 的积分上限函数.所以,在 $f(x)$ 的连续点 x 处,有

$$F'(x)=f(x).$$

这个性质表明,连续型随机变量的分布函数和密度函数是相互确定的,当已知分布函数时,求导可得其密度函数;当已知密度函数时,积分可得分布函数.

例 2-10 设连续型随机变量 X 的密度函数为

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

求:

(1)常数 a;

(2)X 的分布函数;

(3)概率 $P\left(-1<X<\frac{1}{2}\right)$.

解 (1)根据密度函数的规范性 $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$,可得

$$\int_0^1 ax(1-x)\mathrm{d}x=\frac{1}{2}a-\frac{1}{3}a=\frac{1}{6}a=1,$$

解得

$$a=6.$$

(2)由定义 2.8 的 $F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$,可得

当 $x\leqslant 0$ 时,$F(x)=0$;

当 $0<x<1$ 时,$F(x)=\int_0^x 6t(1-t)\mathrm{d}t=3x^2-2x^3$;

当 $x\geqslant 1$ 时,$F(x)=1$.

即

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

(3)可以利用密度函数来求,即

$$P\left(-1<X<\frac{1}{2}\right)=\int_{-1}^{\frac{1}{2}} f(x)\mathrm{d}x=\int_0^{\frac{1}{2}} 6x(1-x)\mathrm{d}x=\frac{1}{2}.$$

也可以利用分布函数来求,即

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

例 2-11 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases} A+Be^{-\lambda x}, & x\geqslant 0, \\ 0, & x<0. \end{cases}$$

其中 $\lambda>0$,求:

(1)常数 A、B;

(2)$P(X\geqslant 3)$;

(3)X 的密度函数.

解 (1)利用分布函数的规范性 $\lim\limits_{x\to+\infty}F(x)=1$,得

$$\lim_{x\to+\infty}(A+Be^{-\lambda x})=A=1;$$

利用连续型随机变量分布函数的连续性,可知

$$\lim_{x\to 0^-}F(x)=0=F(0)=A+B,$$

所以 $A=1$,$B=-1$.

(2)$P(X\geqslant 3)=1-F(3)=1-(1-e^{-3\lambda})=e^{-3\lambda}$.

(3)随机变量 X 的密度函数为

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

二、常见的连续型分布

1. 均匀分布

定义 2.9 若连续型随机变量 X 的密度函数为

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

则称 X 服从区间(a,b)上的均匀分布,记为 $X\sim U(a,b)$.

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

容易求得均匀分布的分布函数为

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

直观地讲,均匀分布反映了随机变量 X 在(a,b)上各点取值的等可能性.严格地说,X 落在(a,b)中任意子区间上的概率只与区间长度有关,而与区间位置无关,即与该子区间的

长度成正比. 事实上，对于任意的$[c,c+l]\subset(a,b)$，都有

$$P(c\leqslant X\leqslant c+l)=\frac{l}{b-a}.$$

这个值与 c 无关.

在实际问题中，服从均匀分布的例子很多. 例如，测量一个圆的直径所产生的误差服从$(-0.5,0.5)$上的均匀分布；乘客在 1 h 内任意时刻到达车站是等可能的，则其到达的时刻服从$(0,60)$上的均匀分布；等等.

2. 指数分布

定义 2.10 若随机变量 X 的密度函数为

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0,\end{cases}$$

其中 $\lambda>0$，则称随机变量 X 服从参数为 λ 的指数分布，记作 $X\sim E(\lambda)$.

容易求得指数分布的分布函数为

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

指数分布常用来作为各种“寿命”分布的近似，如某种电器的使用寿命、随机服务时间等.

例 2-12 某电子元件的使用寿命 $X\sim E(3)$.

(1)求该电子元件使用寿命至少是 5 年的概率；

(2)已知该元件已经使用了 3 年，求其再使用 5 年以上的概率.

解 由题意知 X 的分布函数为

$$F(x)=\begin{cases}1-e^{-3x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

(1)$P(X\geqslant 5)=1-F(5)=1-(1-e^{-15})=e^{-15}$；

(2)$P(X\geqslant 8\mid X\geqslant 3)=\dfrac{P(X\geqslant 8)}{P(X\geqslant 3)}=\dfrac{1-F(8)}{1-F(3)}=\dfrac{e^{-24}}{e^{-9}}=e^{-15}$.

通过这个例子可以看出，使用寿命服从指数分布的电子元件在使用了 3 年后再使用 5 年以上的概率与直接使用 5 年以上的概率相同，即与已使用了 3 年无关. 这是指数分布的一个重要性质，称为无记忆性. 即对任意的 $s>0$、$t>0$，有

$$P(X>s+t\mid X>t)=P(X>s)=e^{-\lambda s}.$$

指数分布是唯一具有这个性质的连续型随机变量.

3. 正态分布

定义 2.11 若随机变量 X 的密度函数为

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

其中$-\infty<\mu<+\infty$，$\sigma>0$为常数，则称X服从参数为μ、σ^2的正态分布，记作$X\sim N(\mu,\sigma^2)$.

若$X\sim N(\mu,\sigma^2)$，则其分布函数为

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

正态分布的密度函数的图形如图2-5所示，它具有如下性质：

(1)图像关于$x=\mu$对称，故$P(X>\mu)=P(X\leqslant\mu)=\frac{1}{2}$；并在$x=\mu$处取得最大值$\frac{1}{\sqrt{2\pi}\sigma}$，在$x=\mu\pm\sigma$处有拐点，以$x$轴为水平渐近线.

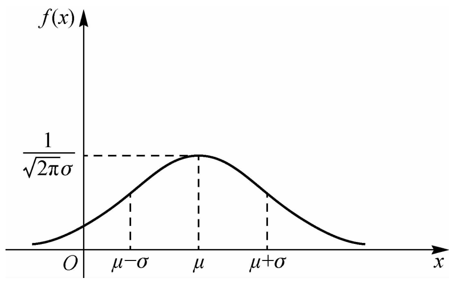

图 2-5

(2)参数μ决定了图形的中心位置，σ决定了图形的最高位置. 当固定σ、改变μ时，曲线沿x轴平行移动，形状不变(图2-6). 当固定μ、改变σ时，曲线的对称轴不变，形状发生改变，σ越大，曲线越平坦，σ越小，曲线越陡峭(图2-7).

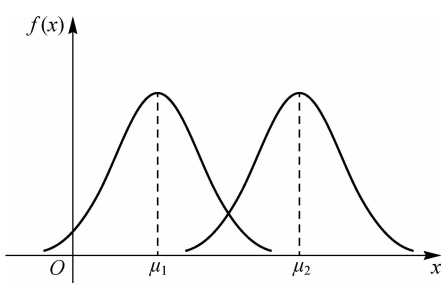

图 2-6

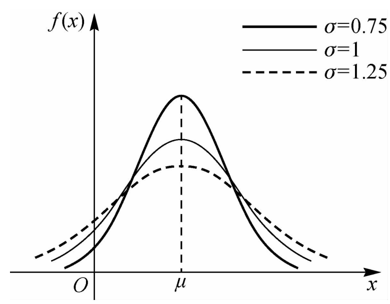

图 2-7

正态分布在概率论中占有非常重要的地位. 实际问题中大量的随机变量都服从或近似服从正态分布. 例如，人的身高体重、农作物的收获量、零件的长度、材料的强度等，都服从或近似服从正态分布. 一般说来，若影响某个数量指标的随机因素很多，而每个因素所起的作用又不太大，则这个指标就服从正态分布. 在概率论和数理统计的理论研究与实际应用中，服从正态分布的随机变量起着非常重要的作用.

特别地，在正态分布中，当$\mu=0$，$\sigma=1$时，则称随机变量X服从**标准正态分布**，记作

$X \sim N(0,1)$. 此时概率密度为

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

分布函数为

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

显然,标准正态分布的密度函数 $\varphi(x)$ 的图形关于 y 轴对称,$\varphi(0)=\frac{1}{\sqrt{2\pi}}$ 为 $\varphi(x)$ 的最大值,在 $x=\pm 1$ 处有拐点(图 2-8).

根据 $\varphi(x)$ 的对称性,由图 2-9 易知

$$\Phi(-x)=1-\Phi(x).$$

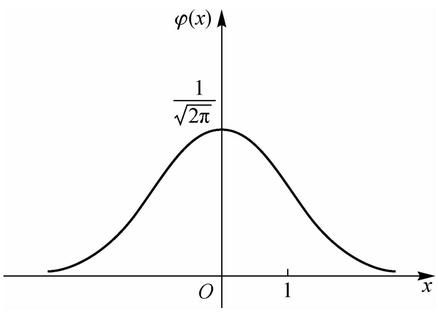

图 2-8

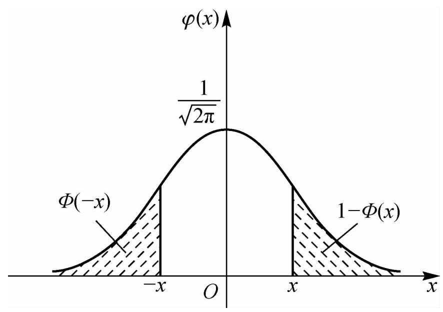

图 2-9

因此,对于标准正态分布的概率计算,只要解决 $x \geqslant 0$ 的计算就可以了. 在附表 2 中给出了 $x \geqslant 0$ 的 $\Phi(x)$ 的数值表. 而对于一般的正态分布,显然不能直接通过密度函数的积分得到事件的概率,我们需要来研究一下它与标准正态分布之间的关系.

定理 2.3 若 $X \sim N(\mu,\sigma^2)$ 的分布函数为 $F(x)$,则对任意 $x \in (-\infty,+\infty)$ 都有

$$F(x)=\Phi\left(\frac{x-\mu}{\sigma}\right).$$

证明 由分布函数的定义知

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

令 $v=\frac{t-\mu}{\sigma}$,可得

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\frac{x-\mu}{\sigma}}\sigma\mathrm{e}^{-\frac{v^2}{2}}\mathrm{d}v=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{x-\mu}{\sigma}}\mathrm{e}^{-\frac{v^2}{2}}\mathrm{d}v=\Phi\left(\frac{x-\mu}{\sigma}\right).$$

由上述定理可知,若 $X \sim N(\mu,\sigma^2)$,则对任意的实数 x_1、x_2 $(x_1<x_2)$,有

$$P(x_1<X<x_2)=F(x_1)-F(x_2)=\Phi\left(\frac{x_2-\mu}{\sigma}\right)-\Phi\left(\frac{x_1-\mu}{\sigma}\right).$$

定理 2.4 若 $X\sim N(\mu,\sigma^2)$，记 $Y=\dfrac{X-\mu}{\sigma}$，则 $Y\sim N(0,1)$.

证明 要证明 $Y\sim N(0,1)$，只要证明 Y 的分布函数为 $\Phi(x)$ 即可.

$$F_Y(x)=P(Y\leqslant x)=P\left(\frac{X-\mu}{\sigma}\leqslant x\right)=P(X\leqslant \mu+\sigma x)=F_X(\mu+\sigma x)=\Phi(x).$$

由上述两个定理知，一般正态分布的概率计算需转化为标准正态分布的概率计算.

例 2-13 若 $X\sim N(\mu,\sigma^2)$，计算 $P(|X-\mu|<k\sigma,k=1,2,3$.

解
$$\begin{aligned}P(|X-\mu|<k\sigma)&=P(\mu-k\sigma<X<\mu+k\sigma)=F(\mu+k\sigma)-F(\mu-k\sigma)\\&=\Phi(k)-\Phi(-k)=2\Phi(k)-1\end{aligned}$$

查附表 2 得，$\Phi(1)=0.841\ 3$，$\Phi(2)=0.977\ 2$，$\Phi(3)=0.998\ 7$，将其分别代入得

$$P(|X-\mu|<\sigma)=P(\mu-\sigma<X<\mu+\sigma)=2\times0.841\ 3-1=0.682\ 6,$$
$$P(|X-\mu|<2\sigma)=P(\mu-2\sigma<X<\mu+2\sigma)=0.954\ 4,$$
$$P(|X-\mu|<3\sigma)=P(\mu-3\sigma<X<\mu+3\sigma)=0.997\ 4.$$

由例 2-13 可以看到，尽管正态分布的随机变量的取值范围是$(-\infty,+\infty)$，但它的值落在$(\mu-3\sigma,\mu+3\sigma)$上几乎是肯定的，这就是通常所说的 3σ 原则(图 2-10).

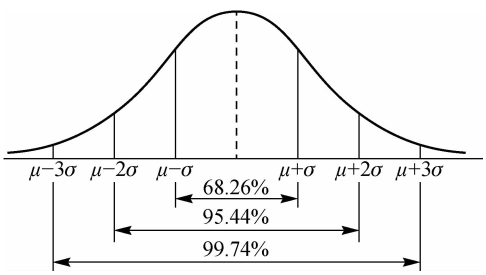

图 2-10

例 2-14 设 $X\sim N(3,2^2)$，求：

(1) $P(2<X<5)$；

(2)确定常数 c，使得 $P(X>c)=P(X<c)$；

(3)设常数 d 满足 $P(X\geqslant d)\geqslant 0.901\ 5$，试给出 d 的取值范围.

解 (1)
$$\begin{aligned}P(2<X<5)&=F(5)-F(2)=\Phi\left(\frac{5-3}{2}\right)-\Phi\left(\frac{2-3}{2}\right)=\Phi(1)-\Phi(-0.5)\\&=\Phi(1)+\Phi(0.5)-1=0.841\ 3+0.691\ 5-1=0.532\ 8.\end{aligned}$$

(2)由题意知

$$P(X>c)=P(X<c)=\frac{1}{2},$$

根据正态分布密度函数的对称性，可知 $c=\mu=3$.

(3)由于 $P(X\geqslant d)\geqslant 0.9>0.5$，所以 $d<3$. 故

$$P(X\geqslant d)=1-P(X<d)=1-\Phi\left(\frac{d-3}{2}\right)=\Phi\left(\frac{3-d}{2}\right)\geqslant 0.9015,$$

查附表 2 得 $\Phi(1.29)=0.9014$，根据分布函数的单调不减性，可知

$$\frac{3-d}{2}\geqslant 1.29,$$

解得 $d\leqslant 0.42$.

例 2-15 设测量中产生的误差 $X\sim N(0,10^2)$，现进行 100 次独立测量，求误差的绝对值超过 19.6 的次数小于 3 的概率.

解 设随机变量 Y 为 100 次测量中事件 $|X|>19.6$ 出现的次数，则 $X\sim B(100,p)$，其中 p 是每次测量中事件 $|X|>19.6$ 出现的概率，即

$$\begin{aligned}p&=P(|X|>19.6)=1-P(|X|\leqslant 19.6)\\&=1-\left[\Phi\left(\frac{19.6}{10}\right)-\Phi\left(-\frac{19.6}{10}\right)\right]=2-2\Phi(1.96),\end{aligned}$$

查附表 2 得 $\Phi(1.96)=0.975$，代入上式得

$$p=2-2\times 0.975=0.05,$$

所以 $X\sim B(100,0.05)$，这里 n 较大，p 较小，可用泊松逼近，即

$$\begin{aligned}P(Y<3)&=P(X=0)+P(X=1)+P(X=2)\\&\approx e^{-5}-\frac{5}{1!}e^{-5}-\frac{5^2}{2!}e^{-5}=1-18.5e^{-5}\approx 0.88.\end{aligned}$$

第四节　随机变量函数的分布

在很多实际问题中，我们所关心的随机变量的分布往往难以直接得到，但是与它们有关的另一些随机变量的分布却容易得到. 例如，对于某圆形的面积不好直接测量，但是可以很容易地测量出它的直径. 因此，我们需要讨论如何根据直径的分布求出面积的分布. 这一节我们将讨论如何由随机变量 X 的概率分布求得它的函数 $Y=g(X)$ 的概率分布.

一、离散型随机变量函数的分布

当 X 为离散型随机变量时，$Y=g(X)$ 的分布可用列举法直接由 X 的分布求得.

例 2-16 设随机变量 X 的概率分布为

X	-2	-1	0	1	2
p	0.2	0.3	0.1	0.3	0.1

求 $Y=2X+1$ 及 $Z=(X-1)^2$ 的分布列.

解 显然函数的取值由 X 的取值及函数关系所确定,故可列表

X	-2	-1	0	1	2
$Y=2X+1$	-3	-1	1	3	5
$Z=(X-1)^2$	9	4	1	0	1
p	0.2	0.3	0.1	0.3	0.1

其中

$$P(Z=1)=P(X=0)+P(X=2)=0.2.$$

所以,可得 Y、Z 的分布列分别为

Y	-3	-1	1	3	5
p	0.2	0.3	0.1	0.3	0.1

Z	9	4	1	0
p	0.2	0.3	0.2	0.3

一般地,设离散型随机变量 X 的分布列为

$$P(X=x_k)=p_k,k=1,2,\cdots,$$

则对 X 的函数 $Y=g(X)$,记 $y_k=g(x_k)(k=1,2,\cdots)$,则

(1)如果 y_k 的值互不相等,由 $P(Y=y_k)=P(X=x_k)$知 Y 的分布列为

$$P(Y=y_k)=p_k,k=1,2,\cdots;$$

(2)如果 y_k 中有相等的值,则将这些相等的值分别合并,并把相应的概率相加,便可得到 Y 的分布列.

例 2-17 已知离散型随机变量 X 的分布列为

$$P(X=k)=\frac{1}{2^k},k=1,2,\cdots,$$

试求函数 $Y=\cos\left(\frac{\pi}{2}X\right)$的分布列.

解 由于

$$\cos\left(\frac{\pi}{2}X\right)=\begin{cases}-1, & k=4n-2,\\ 0, & k=2n-1\\ 1, & k=4n,\end{cases},n=1,2,\cdots,$$

所以 $Y=\cos\left(\frac{\pi}{2}X\right)$ 的所有可能取值为-1、0、1，且取每个值的概率为

$$P(Y=-1)=P(X=2)+P(X=6)+P(X=10)+\cdots$$

$$=\frac{1}{2^2}+\frac{1}{2^6}+\frac{1}{2^{10}}+\cdots=\frac{\frac{1}{4}}{1-\frac{1}{16}}=\frac{4}{15};$$

$$P(Y=0)=P(X=1)+P(X=3)+P(X=5)+\cdots$$

$$=\frac{1}{2}+\frac{1}{2^3}+\frac{1}{2^5}+\cdots=\frac{\frac{1}{2}}{1-\frac{1}{4}}=\frac{2}{3};$$

$$P(Y=1)=P(X=4)+P(X=8)+P(X=12)+\cdots$$

$$=\frac{1}{2^4}+\frac{1}{2^8}+\frac{1}{2^{12}}+\cdots=\frac{\frac{1}{16}}{1-\frac{1}{16}}=\frac{1}{15}.$$

二、连续型随机变量函数的分布

设 X 为连续型随机变量，其密度函数 $f_X(x)$已知，求其函数 $Y=g(X)$的密度函数$f_Y(y)$一般有两种方法：分布函数法（一般方法）和公式法.

1. 分布函数法（一般方法）

为了求 Y 的密度函数 $f_Y(y)$，可先求 Y 的分布函数 $F_Y(y)$，即

$$F_Y(y)=P(Y\leqslant y)=P(g(X)\leqslant y),$$

利用 X 的分布将上述事件的概率求出或表示出，然后再关于 y 求导，即得 Y 的密度函数.

例 2-18 设随机变量 $X\sim U(-1,1)$，求 $Y=X^2$ 的密度函数.

解 由题意知

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

由分布函数的定义知

$$F_Y(y)=P(Y\leqslant y)=P(X^2\leqslant y),$$

当 $y\leqslant 0$ 时，显然上述事件为不可能事件，所以 $F_Y(y)=0$；

当 $0<y<1$ 时，$F_Y(y)=P(X^2\leqslant y)=P(-\sqrt{y}\leqslant X\leqslant \sqrt{y})=\sqrt{y}$；

当 $y\geqslant 1$ 时，$F_Y(y)=P(X^2\leqslant y)=P(-\sqrt{y}\leqslant X\leqslant \sqrt{y})=P(-1<X<1)=1$.

求导得 Y 的密度函数为

$$f_Y(y)=\begin{cases}\dfrac{1}{2\sqrt{y}}, & 0<y<1,\\ 0, & \text{其他}.\end{cases}$$

在这个例子中，我们完整地求出了 Y 的分布函数；但对于某些题目，当 Y 的分布函数不容易求出时，也可通过 X 的分布函数来表示.

例 2-19 设随机变量 $X\sim N(0,1)$，求 $Y=X^2$ 的密度函数.

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

由分布函数的定义知

$$F_Y(y)=P(Y\leqslant y)=P(X^2\leqslant y),$$

当 $y\leqslant 0$ 时，显然上述事件为不可能事件，所以 $F_Y(y)=0$；

当 $y>0$ 时，

$$F_Y(y)=P(X^2\leqslant y)=P(-\sqrt{y}\leqslant X\leqslant \sqrt{y})=2\Phi(\sqrt{y})-1;$$

当 $y>0$ 时，

$$f_Y(y)=\frac{dF_Y(y)}{dy}=2\phi(\sqrt{y})\left(-\frac{1}{2\sqrt{y}}\right)=\frac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}}.$$

所以，$Y=X^2$ 的密度函数为

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

2. 公式法

利用分布函数法可以得到如下公式：

定理 2.5 已知连续型随机变量 X 的密度函数为 $f_X(x)$，若函数 $y=g(x)$ 满足：

(1) $P(a<X<b)=1,-\infty\leqslant a<b\leqslant +\infty$；

(2) 函数 $g(x)$ 是 (a,b) 上严格单调的连续函数；

(3) 反函数 $x=g^{-1}(y)$ 有连续导数.

则 $Y=g(X)$ 的密度函数为

$$f_Y(y)=\begin{cases}f_X[g^{-1}(y)]\,|[g^{-1}(y)]'|, & \alpha<y<\beta,\\ 0, & \text{其他}.\end{cases}$$

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

证明 对于 $y=g(x)$ 严格增加的情形，其反函数 $x=g^{-1}(y)$ 也是严格增加的，此时 $Y=g(X)$ 的可能取值落在区间 $(\alpha=g(a),b=g(b))$ 上，所以，

当 $y\leqslant\alpha$ 时，$F_Y(y)=P(Y\leqslant y)=0$；

当 $\alpha<y<\beta$ 时，$F_Y(y)=P(Y\leqslant y)=P(g(X)\leqslant y)=P(X\leqslant g^{-1}(y))=F_X[g^{-1}(y)]$；

当 $y\geqslant\beta$ 时，$F_Y(y)=P(Y\leqslant y)=1$.

于是，$Y=g(X)$ 的概率密度为

$$f_Y(y)=F'_Y(y)=\begin{cases} f_X[g^{-1}(y)][g^{-1}(y)]', & \alpha<y<\beta, \\ 0, & \text{其他}. \end{cases}$$

对于 $y=g(x)$ 严格单调减少的情形，其反函数 $x=g^{-1}(y)$ 也是严格减少的，所以 $[g^{-1}(y)]'<0$，此时 $Y=g(X)$ 的可能取值落在区间 $(\alpha=g(b),b=g(a))$ 上，所以，

当 $y\leqslant\alpha$ 时，$F_Y(y)=P(Y\leqslant y)=0$；

当 $\alpha<y<\beta$ 时，$F_Y(y)=P(Y\leqslant y)=P(g(X)\leqslant y)=P(X\geqslant g^{-1}(y))=1-F_X[g^{-1}(y)]$；

当 $y\geqslant\beta$ 时，$F_Y(y)=P(Y\leqslant y)=1$.

于是，$Y=g(X)$ 的概率密度为

$$f_Y(y)=F'{}_Y(y)=\begin{cases} -f_X[g^{-1}(y)][g^{-1}(y)]', & \alpha<y<\beta, \\ 0, & \text{其他}. \end{cases}$$

综上，得到 $Y=g(X)$ 的概率密度为

$$f_Y(y)=\begin{cases} f_X[g^{-1}(y)]\left|[g^{-1}(y)]'\right|, & \alpha<y<\beta, \\ 0, & \text{其他}. \end{cases}$$

例 2-20 设随机变量 X 的密度函数为

$$f_X(x)=\begin{cases} \dfrac{1}{\pi(1+x^2)}, & x>0, \\ 0, & \text{其他}, \end{cases}$$

试求随机变量 $Y=\ln X$ 的概率密度.

解 显然，函数 $y=\ln x$ 是严格单调增加的，它的反函数为 $x=g^{-1}(y)=\mathrm{e}^y(-\infty<y<+\infty)$，反函数的导数 $[g^{-1}(y)]'=\mathrm{e}^y$，利用定理 2.5 的结论可得 $Y=\ln X$ 的概率密度为

$$f_Y(y)=\frac{1}{\pi(1+\mathrm{e}^{2y})}\mathrm{e}^y=\frac{\mathrm{e}^y}{\pi(1+\mathrm{e}^{2y})},-\infty<y<+\infty.$$

例 2-21 设随机变量 $X\sim N(\mu,\sigma^2)$，求 $Y=aX+b(a\neq 0)$ 的概率密度.

解 由题意知 X 的密度函数为

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

因 $y=g(x)=ax+b(a\neq 0)$ 严格单调，故它的反函数为

$$x=g^{-1}(y)=\frac{y-b}{a},-\infty<y<+\infty,$$

且其导数$[g^{-1}(y)]'=\frac{1}{a}$，利用定理2.5的结论可得$Y=aX+b(a\neq 0)$的概率密度为

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

由此可知，当$X\sim N(\mu,\sigma^2)$且$a\neq 0$时，有

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

即服从正态分布的随机变量的线性函数仍然服从正态分布.特别地，令$a=\frac{1}{\sigma}$，$b=-\frac{\mu}{\sigma}$，则$aX+b=\frac{X-\mu}{\sigma}\sim N(0,1)$，即定理2.4的结论.

习题

(1)一口袋中装有6个形状相同的球，分别标号为1、2、3、4、5、6，现从袋中任取3个，设X是取出球的最小号码，写出随机变量X的分布列，并求$P(X\geqslant 3)$.

(2)设随机变量X的分布列为$P(X=k)=\frac{c}{k+1}$，$k=0,1,3,5$，求：

①常数c；

②$P(X<2)$；

③$P(X<3|X\neq 1)$.

(3)两名射手轮流射击某一目标.第一位射手命中目标的概率为1/2，第二位射手命中目标的概率为1/3.为击中目标共射击了X次，求X的分布列.

(4)试确定a的值，使下列形式为某个离散型随机变量的分布列：

①$P(X=k)=ae^{-k+2}(k=0,1,2,\cdots)$；

②$P(X=k)=\frac{a}{3^k k!}(k=0,1,2,\cdots)$.

(5)自动生产线在调整之后出现废品的概率为p，在生产过程中出现废品时立即重新进行调整，求在两次调整之间生产合格品X的分布列.

(6)甲、乙两人投篮，投中的概率分别为0.6、0.7，今各投3次，求：

①两人投中次数相同的概率；

②甲比乙投中次数多的概率.

(7)某公安局在长度为t(单位为h)的事件间隔内收到的报警次数X服从参数为$\frac{t}{2}$的泊松分布，而与时间间隔起点无关.求：

①某天 12:00～15:00 没有收到报警的概率；

②某天 12:00～17:00 至少收到 1 次报警的概率.

(8)已知 $X\sim B(5,p)$，并且 $P(X=1)=P(X=2)$，求 $P(X=4)$.

(9)已知随机变量 X 的概率密度为 $f(x)=Ae^{-|x|}$，$-\infty<x<+\infty$，求：

①A 的值；

②$P(0<X<1)$；

③分布函数 $F(x)$.

(10)设随机变量 $X\sim U(2,5)$，现对 X 进行 3 次独立观测，求至少有两次观测值大于 3 的概率.

(11)设随机变量 $\xi\sim U(1,6)$，则方程 $x^2+\xi x+1=0$ 有实根的概率是多少？

(12)某电子元件的使用寿命 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{100}{x^2}, & x\geqslant 100,\\ 0, & \text{其他}.\end{cases}$$

若 1 台电视上装有 3 个这种元件，求使用最初的 150 h 内 3 个元件都没有发生故障的概率.

(13)设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases}A+Be^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0.\end{cases}$$

其中 $\lambda>0$. 求：

①常数 A、B；

②$P(-1<X<1)$，$P(X\geqslant 3)$；

③概率密度 $f(x)$.

(14)设成年男子的身高(以 cm 计)$X\sim N(170,6^2)$，某种公共汽车车门的高度是按成年男子碰头的概率在 1%以下来设计的，问车门的高度最少应为多少厘米？

(15)设 $X\sim N(\mu,\sigma^2)$ 且 $P(X\leqslant -5)=0.045$，$P(X>3)=0.382$，求 μ 和 σ.

(16)设车床加工金属圆杆，已知圆杆的直径(以 cm 计)$X\sim N(12.4,\sigma^2)$，规定直径为 12.0～12.8 cm 的概率至少为 0.95，试确定 σ 最大为多少.

(17)设随机变量 X 的分布列为

X	-2	-1	0	1	2
p	0.15	0.32	0.24	0.11	0.18

求：

①$Y=2X+1$；

②$Y=|X|-1$ 的分布列.

(18)一食品厂一天的产量 X(以 t 计)的概率密度为

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

一天的产值是 $Y=5X+1$(以千元计),求 Y 的概率密度.

(19)设随机变量 X 服从(0,9)上的均匀分布,随机变量 Y 是 X 的函数:

$$Y=\begin{cases}-1, & X<1,\\ 1, & X=1,\\ 2, & 1<X\leqslant 6,\\ 3, & 6<X\leqslant 9,\end{cases}$$

求 Y 的分布列.

(20)一台电话总机共有 225 部分机,每部分机呼叫外线的概率为 0.02.

①若设置 12 条外线,则每部分机呼叫外线失败的概率是多少?

②这台总机需要设置多少外线,才能保证每部分机呼叫外线失败的概率低于 0.01.

(21)设随机变量 $X\sim U(-1,1)$. 求下列函数的概率密度:

①$Y=(X+1)^2$;

②$Z=\ln(X+1)$;

③$T=2X^2+1$.

(22)设随机变量 $X\sim E(\lambda)$,其分布函数为 $F(x)$. 求随机变量 $Y=F(x)$ 的概率密度函数.

(23)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.