第 2 章 随机变量及其分布

2.1 内容提要

2.1.1 随机变量的概念

1. 随机变量

设随机试验的样本空间是S, 如果$X = X(\omega)$是定义在样本空间S上的实值函数, 即对于每一个$\omega \in S$, 总有一个确定的实数$X(\omega)$与其对应, 则称$X = X(\omega)$为随机变量. 一般用大写英文字母X, Y, Z或希腊字母ξ, η, ζ等表示随机变量, 其可能的取值用小写字母x, y, z等表示.

随机事件A可以用随机变量X的取值表示出来, 即$A = (X \in S)$, 其中$S \subset R$(实数集).

随机变量按取值情况可分为离散型和非离散型两个类型, 其中非离散型随机变量中最重要的, 也是应用最广的是连续型随机变量.

2. 随机变量的分布函数

设X为随机变量, 对任意实数x, 则称函数$F(x) = P\{X \le x\}$为随机变量X的分布函数. 分布函数具有以下性质:

(1) **有界性** $0 \le F(x) \le 1$;

(2) **单调非降性** 对任意$x_1 < x_2$, 有$F(x_1) \le F(x_2)$;

(3) $F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0$, $F(+\infty) = \lim\limits_{x \to +\infty} F(x) = 1$;

(4) **右连续性** 对任意实数x_0, 都有$F(x_0) = \lim\limits_{x \to x_0 + 0} F(x)$;

(5) $P\{a < X \le b\} = P\{X \le b\} - P\{X \le a\} = F(b) - F(a)$.

实际上, 满足条件(1), (2), (3), (4)的实值函数$F(x)$一定可以做为某个随机变量的分布函数. 分布函数是随机变量的一般特征, 无论是离散型随机变量还是连续型随机变量都有分布函数.

2.1.2 随机变量的概率分布

随机变量的概率分布就是随机变量取值的概率规律, 简称分布.

1. 离散型随机变量及其分布律

如果随机变量X的所有可能的取值为有限个或可列个, 则称X为离散型随机变量. 设X

的所有可能取值为 $x_1, x_2, x_3, \cdots x_n \cdots$，且 X 取以上各值的概率分别为 $p_1, p_2, \cdots p_n, \cdots$ 即

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

这一系列的式子称为离散型随机变量 X 的分布律, 通常也写成表格的形式:

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

离散型随机变量的概率分布具有以下性质:

(1) $p_k \ge 0$，$k = 1, 2, \cdots, n, \cdots$；

(2) $\sum\limits_{k=1}^{\infty} p_k = 1$；

(3) $P(a < \xi \le b) = \sum\limits_{a < x_k \le b} p_k$.

实际上, 满足上述(1), (2)两个条件的数列 $p_1, p_2, \cdots, p_n, \cdots$, 一定可以作为某个离散型随机变量的概率分布.

离散型随机变量的分布函数为

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

可以看出, 离散型随机变量的分布函数 $F(x)$ 在 X 可能的取值 x_k 处发生跳跃, 其跳跃的高度为 X 取该值的概率, 它是单调, 非降的阶梯函数.

2. 连续型随机变量及其概率密度

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

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

则称 X 为连续型随机变量, $f(x)$ 称为 X 的概率密度函数, 简称概率密度或密度函数, 简记 $X \sim f(x)$.

概率密度 $f(x)$ 具有以下性质:

(1) $f(x) \ge 0 \quad (-\infty < x < +\infty)$；

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

(3) $P(a < X \le b) = \int_a^b f(x)\mathrm{d}x$.

对于连续型随机变量 X, 它还有下面重要性质:

(1) 连续型随机变量取任意给定数值的概率都是零, 即 $P\{X = a\} = 0$, 其中 a 为任意实数. 因而也有

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

必须注意, 上式对于离散型随机变量一般不成立.

(2) 若密度函数 $f(x)$ 在 x 点处连续, 则 $F'(x) = f(x)$.

2.1.3 几种常见分布

1. 常见离散型随机变量的分布

(1) $(0-1)$分布

当随机试验只有两种可能结果时, 我们常常把这两个值取为0和1, 这时称随机变量X服从参数为p的$0-1$分布, 其概率分布为

X	0	1
P	$1-p$	p

(2) 二项分布

若随机变量X的分布律为

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

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

一般地, 在n重伯努利试验中, 如果每次试验中事件A发生的概率为p, 用X表示A发生的次数, 这时X服从二项分布.

(3) 泊松分布

若随机变量X的分布律为

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

其中$\lambda>0$, 则称X服从参数为λ的泊松分布, 记作$X \sim P(\lambda)$.

二项分布与泊松分布的关系:

(**泊松定理**) 假设在n重伯努利试验中, 随着试验次数n无限增大, 而事件出现的概率p_n无限缩小, 且当$n\to+\infty$时有$np_n\to\lambda$, 则

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

(4) 几何分布

若离散型随机变量X的分布律为:

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

则称随机变量X服从参数为p的几何分布, 记为$X\sim g(p)$.

一般地, 在伯努利试验中, 设每次试验中事件A发生的概率为p, 记X为首次发生事件A的试验次数, 则X服从参数为p的几何分布.

(5) 超几何分布

若离散型随机变量X的分布律为:

$$P\{X=m\}=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n} \qquad (m=0,1,\cdots,n),$$

则称随机变量 X 服从超几何分布.

一般地,超几何分布的典型例子是:假设有 N 个产品,其中 M 个是正品, $N-M$ 个次品,从中无放回取出 n 个产品,则其中含有的正品数 X 的分布律为超几何分布.

(6) 幂律分布

若离散型随机变量 X 的分布律为:

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

则称随机变量 X 服从参数为 γ 的幂律分布,其中幂次 $\gamma>1$, C 为归一化常数.

注 幂律分布的类型有好几种,上面提到的只是其中的一种.幂律分布被称为复杂系统的"指纹".关于幂律分布的普适性研究目前仍然为科学前沿的热点之一,本书特别给以介绍,只是为了表明其重要性.

2. 常见连续型随机变量的分布

(1) 均匀分布

若连续型随机变量 X 的密度函数为:

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

则称 X 在区间 $[a,b]$ 上服从均匀分布,记作 $X\sim U[a,b]$. 其分布函数为

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

(2) 指数分布

若连续型随机变量 X 的密度函数为:

$$f(x)=\begin{cases}\lambda \mathrm{e}^{-\lambda x}, & x\ge 0,\\ 0, & x<0,\end{cases}$$

则称随机变量 X 服从参数为 λ 指数分布,记作 $X\sim E(\lambda)$, 其分布函数为

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

(3) 柯西分布

若连续型随机变量 X 的密度函数为:

$$f(x)=\frac{1}{\pi(1+x^2)} \quad (-\infty<x<+\infty),$$

则称随机变量 X 服从柯西分布, 其分布函数为

$$F(x)=\frac{1}{\pi}\arctan x+\frac{1}{2}.$$

（4） 正态分布

若连续型随机变量 X 的密度函数为:

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

则称随机变量 X 服从参数为 μ,σ^2 的正态分布, 记作 $X\sim N(\mu,\sigma^2)$.

特别地, 称 $\mu=0,\sigma=1$ 时的正态分布为标准正态分布, 其密度函数、分布函数分别用 $\varphi(x)$ 和 $\Phi(x)$ 表示, 即

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

一般概率统计教材后附有 $\Phi(x)$ 的数值表, 供查用.

正态分布的重要性质:

① 若 $X\sim N(\mu,\sigma^2)$, 则 $aX+b\sim N(a\mu+b,a^2\sigma^2)\quad(a\neq 0)$. 特别地, X 的标准化随机变量 $\dfrac{X-\mu}{\sigma}\sim N(0,1)$;

② “3σ 规则”: 设随机变量 $X\sim N(\mu,\sigma^2)$, 则

$P(|X-\mu|\leq\sigma)=0.683$,

$P(|X-\mu|\leq 2\sigma)=0.954$,

$P(|X-\mu|\leq 3\sigma)=0.997$;

③ 记 $F(x)$ 与 $\Phi(x)$ 分别为一般正态分布 $N(\mu,\sigma^2)$ 与标准正态分布的分布函数, 则对任意实数 x, 都有 $F(x)=\Phi(\dfrac{x-\mu}{\sigma})$, 且

$$P\{a\leq X\leq b\}=F(b)-F(a)=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right) \quad ;$$

④ 对于任意的 x, 均有 $\Phi(-x)=1-\Phi(x)$.

若 $x>0$, 可直接查表得到 $\Phi(x)$ 的值；若 $x<0$, 则利用上述公式, 再查表, 即可得到 $\Phi(x)$ 的值.

2.1.4 随机变量函数的分布

1. 离散型随机变量函数的分布

随机变量 X 的分布律为

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

又 $y=h(x)$ 是连续函数, 则 $Y=h(X)$ 也是一个随机变量, Y 的分布律可由下表求得

Y	$y_1=h(x_1)$	$y_2=h(x_2)$	$\cdots$	$y_n=h(x_n)$	$\cdots$
$P\{Y=y_k\}$	p_1	p_2	$\cdots$	p_n	$\cdots$

若 $h(x_k)\quad(k=1,2,\cdots)$ 的值互不相同, 则上表就是 Y 的分布律; 若 $h(x_k)\quad(k=1,2,\cdots)$ 的值中有相等的, 则应把那些相等的取值合并, 同时把对应的概率相加, 从而得到 Y 的分布律.

2. 连续型随机变量函数的分布

设连续型随机变量 X 的分布函数为 $F_X(x)$, 概率密度为 $f_X(x)$, $y=h(x)$ 是连续函数, 求 $Y=h(X)$ 的分布函数 $F_Y(y)$ 或概率密度 $f_Y(y)$ 的方法主要有:

(1) 定义法

定义法也称分布函数法, 关键是设法找出 Y 的分布函数 $F_Y(y)$ 与 X 的分布函数 $F_X(x)$ 之间的关系.

首先按定义写出 Y 的分布函数 $F_Y(y)=P\{Y\le y\}$; 然后利用关系式 $Y=h(X)$, 把事件 $\{Y\le y\}$ 转化为等价事件 $\{X\in S\}$, 其中 $S\subset R$, 并将其概率用 X 的分布函数表示出来, 记作 $F_X[u(y)]$, 即得 $F_Y(y)=F_X[u(y)]$; 最后两边关于 y 求导, 即可得 Y 的密度函数 $f_Y(y)$.

(2) 公式法

如果 $y=h(x)$ 为单调函数, 最小值为 α, 最大值为 β, $h(x)$ 处处可导, 且导数不为零, 那么随机变量 $Y=h(X)$ 的密度函数

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

其中 $x=h^{-1}(y)$ 为 $y=h(x)$ 的反函数, $f_X(x)$ 为 X 的密度函数.

2.2 习题详解

2.2 练习题

1. 填空题

(1) 已知某自动生产线加工出的产品次品率为 0.01, 检验人员每天检验 8 次, 每次从已生

产出的产品中随意取 10 件进行检验,如果发现其中有次品就去调整设备,那么一天至少要调整设备一次的概率为__________.

解 发现的次品数 $X \sim B(80, 0.01)$,若设备不需调整,即加工出的产品全是正品.那么一天至少要调整设备一次的概率为

$$1-P\{X=0\}=1-C_{80}^{0}\times 0.01^{0}\times 0.99^{80}\approx 0.55.$$

(2) 袋中有 8 个球,其中 3 个白球,5 个黑球.现从中随意取出 4 个球,如果 4 个球中有 2 个白球 2 个黑球,试验停止,否则将 4 个球放回袋中重新抽去 4 个球,直至取到 2 个白球 2 个黑球为止.用 X 表示抽取次数,则 $P\{X=k\}=$__________.

解 取出的 4 个球中有 2 个白球 2 个黑球的概率为 $\dfrac{C_3^2C_5^2}{C_8^4}=\dfrac{3}{7}$,此时 X 服从几何分布,分布律为:$P\{X=k\}=\dfrac{3}{7}(\dfrac{4}{7})^{k-1}$.

2. 选择题

设随机变量 X 服从参数为 λ 的泊松分布,已知 $2P\{X=1\}=P\{X=2\}$,则参数 λ 等于( ).

A. 1 B. 2 C. 3 D. 4

解 因为 $2P\{X=1\}=P\{X=2\}$,即 $2\dfrac{\lambda^1}{1!}\mathrm{e}^{-\lambda}=\dfrac{\lambda^2}{2!}\mathrm{e}^{-\lambda}$,得 $\lambda=4$,故本题应选 D.

2.3 练习题

1. 填空题

设随机变量 X 的分布函数为 $F(x)=A+B\arctan x$,则 $A=$_______, $B=$________.

解 由分布函数的性质 $F(+\infty)=1$, $F(-\infty)=0$,得

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

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

解上述两式,得 $A=\dfrac{1}{2}, B=\dfrac{1}{\pi}$.

2. 选择题

设 X 与 Y 是任意两个随机变量,它们的分布函数分别为 $F_1(x)$和$F_2(x)$,则( ).

A. $F_1(x)+F_2(x)$ 必为某一随机变量的分布函数

B. $F_1(x)-F_2(x)$ 必为某一随机变量的分布函数

C. $\dfrac{1}{2}(F_1(x)+2F_2(x))$ 必为某一随机变量的分布函数

D. $F_1(x)F_2(x)$ 必为某一随机变量的分布函数

分析 要判断 $F(x)$ 是否为分布函数,需要验证 $F(x)$ 是否同时满足: $0 \le F(x) \le 1$,单调非降,右连续,及 $F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$, $F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$.

解 首先否定 A, B, C.这是因为 $F_1(+\infty)+F_2(+\infty)=1+1=2$, $F_1(+\infty)-F_2(+\infty)=0$, $\dfrac{1}{2}(F_1(+\infty)+2F_2(+\infty))=\dfrac{3}{2}$,因此,它们均不是随机变量的分布函数.而在 D 中易知 $F(x)$ 同时满足分布函数的几条性质,故本题应选 D.

2.4 练习题

1. 填空题

(1) 设连续型随机变量 X 的概率密度为 $f(x)$,若 $\lim\limits_{x\to\infty}f(x)$ 存在,则 $\lim\limits_{x\to\infty}f(x)$ =________.

解 因为概率密度 $f(x)$ 是要同时满足: $f(x)\ge 0$, $\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1$. 若 $\lim\limits_{x\to\infty}f(x)$ 存在且大于零,则 $\int_{-\infty}^{+\infty}f(x)\mathrm{d}x\ne 1$, 故 $\lim\limits_{x\to\infty}f(x)=0$.

(2) 设随机变量 X 服从正态分布 $N(\mu,\sigma^2)(\sigma>0)$,且二次方程 $y^2+4y+4X=0$ 无实根的概率为 $\dfrac{1}{2}$,则 μ =__________.

解 由 $\Delta=16-16X<0,$ 可得当 $X>1$ 时方程无实根,所以 $P\{X>1\}=\dfrac{1}{2}$, 由于正态分布的概率密度关于 $X=\mu$ 对称,故 $\mu=1$.

2. 选择题

(1) 设 X 与 Y 是任意两个连续型随机变量,它们的概率密度分别为 $f_1(x)$ 和 $f_2(x)$,则(　　).

A. $f_1(x)+f_2(x)$ 必为某一随机变量的概率密度

B. $\dfrac{1}{2}(f_1(x)+f_2(x))$ 必为某一随机变量的概率密度

C. $f_1(x)-f_2(x)$ 必为某一随机变量的概率密度

D. $f_1(x)f_2(x)$ 必为某一随机变量的概率密度

分析 要判断某个函数 $f(x)$ 是否为某连续型随机变量的概率密度,需要验证 $f(x)$ 是否同时满足: (1) $f(x)\ge 0 \quad (-\infty<x<+\infty)$; (2) $\int_{-\infty}^{+\infty}f(x)dx=1$.

解 首先否定 A 与 C. 这是因为

$$\int_{-\infty}^{+\infty}[f_1(x)+f_2(x)]dx=1+1=2,\ \int_{-\infty}^{+\infty}[f_1(x)-f_2(x)]dx=1-1=0.$$

因此, 它们均不是随机变量的概率密度.

对于选项 D, 若

$$f_1(x)=\begin{cases}1, & -2<x<-1,\\ 0, & \text{其他},\end{cases}\quad f_2(x)=\begin{cases}1, & 0<x<1,\\ 0, & \text{其他},\end{cases}$$

则对任意一 $x\in(-\infty,+\infty)$, $f_1(x)f_2(x)=0$, 则 $\int_{-\infty}^{+\infty}f_1(x)f_2(x)dx=0\neq 1$, 因而否定 D. 综上分析, 用排除法选 B.

进一步分析可知

$$\frac{1}{2}[f_1(x)+f_2(x)]\geq 0,\ \int_{-\infty}^{+\infty}\frac{1}{2}[f_1(x)+f_2(x)]dx=\frac{1}{2}(1=1)=1.$$

故本题应选 B.

(2) 假设随机变量 X 的概率密度 $f(x)$ 是偶函数, 分布函数为 $F(x)$, 则(　　).

A. $F(x)$ 是偶函数　　B. $F(x)$ 是奇函数

C. $F(x)+F(-x)=1$　　D. $2F(x)-F(-x)=1$

解 由于 $F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$, $F(+\infty)=\lim\limits_{x\to+\infty}F(x)=1$, 因而否定 A 与 B.

因为概率密度 $f(x)$ 是偶函数, 所以 $P\{X\leq -x\}=P\{X\geq x\}$, 从而有

$$F(x)+F(-x)=P\{X\leq x\}+P\{X\leq -x\}=P\{X\leq x\}+P\{X\geq x\}=1,$$

$$2F(x)-F(-x)=2P\{X\leq x\}-P\{X\leq -x\}=2P\{X\leq x\}-P\{X\geq x\}\neq 1,$$

故本题应选 C.

2.5 练习题

1. 填空题

(1) 设随机变量 X 服从 $N(0,\sigma^2)$, 则 $Y=-X$ 服从__________.

解 设 $F_X(x),F_Y(y)$ 分别为随机变量 X 与 Y 的分布函数.

$$F_Y(y)=P\{Y\leq y\}=P\{-X\leq y\}=P\{X\geq -y\}=1-F_X(-y),$$

由于 X 服从 $N(0,\sigma^2)$, X 的概率密度函数关于 $x=0$ 即 y 轴对称, 是偶函数, 故 $1-F_X(-y)=F_X(y)$, 则 $F_Y(y)=F_X(y)$, 从而 $Y=-X$ 与 X 具有相同的分布服从 $N(0,\sigma^2)$.

(2) 设随机变量 X 的概率密度 $f_X(x)$ 是偶函数, 则 $Y=|X|$ 的概率密度 $f_Y(y)=$______.

解 设 $F_X(x),F_Y(y)$ 分别为随机变量 X 与 Y 的分布函数.

当 $y<0$ 时， $F_Y(y)=P\{Y\le y\}=P\{|X|\le y\}=0$,

当 $y\ge 0$ 时， $F_Y(y)=P\{Y\le y\}=P\{|X|\le y\}$

$$=P\{-y\le X\le y\}=F_X(y)-F_X(-y),$$

从而 $Y=|X|$ 的密度函数为

$$f_Y(y)=F_Y'(y)=\begin{cases}f_X(y)+f_X(-y), & y>0,\\ 0, & y\le 0.\end{cases}$$

由于概率密度 $f_X(x)$ 是偶函数, 故

$$f_Y(y)=\begin{cases}2f_X(y), & y>0,\\ 0, & y\le 0.\end{cases}$$

2. 选择题

设随机变量 X 的概率密度 $f(x)$ 是偶函数, 则下列结论错误的是(　　　).

A. X 的分布函数 $F(x)$ 是偶函数　　　　B. $-X$ 与 X 有相同的概率密度

C. $-X$ 与 X 具有相同的分布函数　　　　D. $F(-x)+F(x)=1$

解　本题应选 A. 参见 2.4 练习题中的选择题(2)及本节填空题(1).

习题二

1. 掷一颗均匀骰子两次, 以 X 表示前后两次出现的点数之和, Y 表示两次中所得的最小点数, 求:(1) X 的分布律;(2) Y 的分布律.

分析　求离散型随机变量分布律的步骤为: (1) 找出随机变量 X 的所有可能的取值, (2) 求出 X 取各可能值的事件的概率.

解　记 X_1, X_2 分别为掷一颗均匀骰子两次出现的点数, 则

$$P\{X_1=k\}=\frac{1}{6},\ P\{X_2=k\}=\frac{1}{6},\ k=1,2,\cdots,6.$$

由于事件 $\{X_1=i\}$ 与事件 $\{X_2=j\}$ 相互独立, 因此

$$P\{X_1=i, X_2=j\}=P\{X_1=i\}\cdot P\{X_2=j\}=\frac{1}{6}\cdot\frac{1}{6}=\frac{1}{36},\ i,j=1,\cdots,6.$$

(1) 随机变量 $X=X_1+X_2$ 的可能值为 $2,3,\cdots,12$,

$$P\{X=2\}=P\{X_1=1, X_2=1\}=P\{X_1=1\}\cdot P\{X_2=1\}=\frac{1}{36},$$

$$P\{X=3\}=P\{X_1=1, X_2=2\}+P\{X_1=2, X_2=1\}=\frac{1}{36}+\frac{1}{36}=\frac{2}{36},$$

$$P\{X=4\}=P\{X_1=1, X_2=3\}+P\{X_1=2, X_2=2\}+P\{X_1=3, X_2=1\}=\frac{3}{36},$$

依次可得 X 的分布律为

X	2	3	4	5	6	7	8	9	10	11	12
P	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

(2) 随机变量 $Y=\min(X_1,X_2)$ 的可能值为 $1,2,3,4,5,6$, 当且仅当以下三种情况之一发生时事件 $\{Y=k\}, k=1,2,3,4,5,6$ 发生:

1) $X_1=k$ 且 $X_2=k+1,k+2,\cdots,6$(共有 $6-k$ 个点);

2) $X_2=k$ 且 $X_1=k+1,k+2,\cdots,6$(共有 $6-k$ 个点);

3) $X_1=k$ 且 $X_2=k$(仅有1个点).

因此事件 $\{Y=k\}$ 共包含 $(6-k)+(6-k)+1=13-2k$ 个样本点, 于是 Y 的分布律为

$$P\{Y=k\}=\frac{13-2k}{36},\quad k=1,2,3,4,5,6.$$

因此 Y 的分布律为

Y	1	2	3	4	5	6
P	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$

2. 口袋中有 7 只白球、3 只黑球, 每次从中任取一个, 如果取出黑球则不放回, 而另外放入一只白球, 求首次取出白球时的取球次数 X 的分布律.

解 X 的可能取值为 $1,2,3,4$, 且"$X=1$"表示第一次就取到白球, 其概率为 $\frac{7}{10}$, "$X=2$"表示第一次取到黑球, 第二次取到白球, 其概率为 $\frac{3}{10}\cdot\frac{8}{10}$, 类似可得

$$P\{X=1\}=\frac{7}{10}=0.7,\qquad P\{X=2\}=\frac{3}{10}\cdot\frac{8}{10}=0.24,$$

$$P\{X=3\}=\frac{3}{10}\cdot\frac{2}{10}\cdot\frac{9}{10}=0.054,\ P\{X=4\}=\frac{3}{10}\cdot\frac{2}{10}\cdot\frac{1}{10}\cdot\frac{10}{10}=0.006.$$

因此 X 的分布律为

X	1	2	3	4
P	0.7	0.24	0.054	0.006

3. 一台设备由三个部件构成, 在设备运转中各部件需要调整的概率相应为 0.10, 0.20 和 0.30, 假设各部件的状态相互独立, 以 X 表示同时需要调整的部件数, 试求 X 的分布律.

解 设事件 A_i = "第 i 个部件需要调整", $i=1,2,3$. 依题意 A_1,A_2,A_3 相互独立, 且

$$P(A_1)=0.10,\ P(A_2)=0.20,\ P(A_3)=0.30.$$

显然 X 的可能取值为 $0,1,2,3$.

$$P\{X=0\}=P(\overline{A}_1\overline{A}_2\overline{A}_3)=P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3)=0.9\times0.8\times0.7=0.504,$$

$$P\{X=1\}=P(A_1\overline{A}_2\overline{A}_3)+P(\overline{A}_1A_2\overline{A}_3)+P(\overline{A}_1\overline{A}_2A_3)$$

$$=0.1\times0.8\times0.7+0.9\times0.2\times0.7+0.9\times0.8\times0.3=0.398,$$

$$P\{X=3\}=P(A_1A_2A_3)=0.1\times0.2\times0.3=0.006,$$

$$P\{X=2\}=1-P\{X=0\}-P\{X=1\}-P\{X=3\}$$

$$=1-0.504-0.398-0.006=0.092.$$

即 X 的分布律为

X	0	1	2	3
P	0. 504	0. 398	0. 092	0. 006

4. 一批产品共有 100 件, 其中 10 件是次品, 从中任取 5 件产品进行检验, 如果 5 件都是正品, 则这批产品被接收, 否则不接收这批产品, 求:(1) 5 件产品中次品数 X 的分布律;(2) 不接收这批产品的概率.

解 (1) X 的所有可能取值为 $0,1,2,3,4,5$, 其分布律为

$$P\{X=k\}=\frac{C_{10}^{k}C_{90}^{5-k}}{C_{100}^{5}},\ k=0,1,2,3,4,5;$$

(2) 不接收这批产品的概率为 $1-\dfrac{C_{90}^{5}}{C_{100}^{5}}$.

5. 设离散型随机变量 X 的分布律为

$$P\{X=k\}=\frac{C}{15},\quad k=1,2,3,4,5$$

(1) 试确定常数 C; (2) 求 $P\{1\le X\le3\}$;(3) $P\{0.5<X<2.5\}$.

解 (1) 由分布律的性质, 有

$$1=\sum_{k=1}^{5}P\{X=k\}=\sum_{k=1}^{5}\frac{C}{15},$$

由此得, $C=3$;

(2) $P\{1\le X\le3\}=P\{X=1\}+P\{X=2\}+P\{X=3\}=\dfrac{9}{15}=0.6$;

(3) $P\{0.5<X<2.5\}=P\{X=1\}+P\{X=2\}=\dfrac{6}{15}=0.4$.

6. 设一个试验只有两种结果:成功或失败, 且每次试验成功的概率为 p ($0<p<1$), 现反复试验, 直到获得 k 次成功为止. 以 X 表示试验停止时一共进行的试验次数, 求 X 的分布律.

分析 考虑的是独立重复试验序列, "直至事件 A 发生 k 次为止" 表示至少需要进行 k 次试验, 如果需要进行 $k+r$ 次试验, 则表明前 $k+r-1$ 次试验中事件 A 恰好发生了 $k-1$ 次, 而

第 $k+r$ 次试验中事件 A 发生.

解 设事件 A 发生 k 次时所需要进行的试验次数为 X，则 X 的取值范围为 $k,k+1,k+2,\cdots$. 事件 $\{X=n\}$ 表示前 $n-1$ 次试验中事件 A 恰好发生 $k-1$ 次, 并且第 n 次试验中事件 A 发生, 所以

$$P\{X=n\}=C_{n-1}^{k-1}p^{k-1}(1-p)^{n-k}\cdot p=C_{n-1}^{k-1}p^{k}(1-p)^{n-k},\ n=k,k+1,\cdots.$$

注 此分布称为帕斯卡分布或负二项分布, 当 $k=1$ 时, 即为几何分布. 因此可以说几何分布是负二项分布的特殊情形.

7. 设某射手每次射击命中目标的概率为 0.8, 现射击了 20 次, 求射中目标次数的分布律.

分析 在很多实际问题中, 随机变量的分布往往是一些常见的分布(如二项分布, 几何分布, 泊松分布, 指数分布, 正态分布等), 因此我们应该熟悉这些常见分布所描述的一些典型的概率模型. 本题中, 射手每次射击可以看成是一次试验, 命中与否是其两个结果, 射击 20 次相当于 20 次重复试验, 且各次试验相互独立, 因此是 n 重伯努利试验.

解 用 A 表示事件“射手射击命中目标”, 则命中目标的次数 X，就是 n 重伯努利试验中事件 A 出现的次数, 因此 X 服从二项分布, 其参数 $n=20$，$p=P(A)=0.8$，即 $X\sim B(20,0.8)$., 故 X 的分布律为

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

8. 一个工人同时看管 5 部机器, 在一小时内每部机器需要照看的概率是 $\frac{1}{3}$, 求: (1) 在一小时内没有 1 部机器需要照看的概率; (2) 在一小时内至少有 4 部机器需要照看的概率.

解 设在一小时内需要照看的机器数为 X， 则 $X\sim B(5,\frac{1}{3})$.

(1) 在一小时内没有 1 部机器需要照看的概率为

$$P\{X=0\}=C_5^0\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^5=\frac{32}{243};$$

(2) 在一小时内至少有 4 部机器需要照看的概率为

$$P\{X\geq 4\}=P\{X=4\}+\{X=5\}=C_5^4\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^1+C_5^5\left(\frac{1}{3}\right)^5\left(\frac{2}{3}\right)^0=\frac{11}{243}.$$

9. 甲、乙二人投篮, 投中的概率分布为 0.6、0.7. 二人各投 3 次, 求: (1) 二人投中次数相等的概率; (2) 甲比乙投中次数多的概率.

解 以 X, Y 分别表示甲、乙投中的次数, 则 $X\sim B(3,0.6)$, $Y\sim B(3,0.7)$.

(1) 按题意需求事件 $\{X=Y\}$ 的概率, 而事件 $\{X=Y\}$ 是下列 4 个两两互不相容的事件之和, 即

$$\{X=0\}\cap\{Y=0\},\ \{X=1\}\cap\{Y=1\},\ \{X=2\}\cap\{Y=2\},\ \{X=3\}\cap\{Y=3\}.$$

自然, 甲、乙投中与否被认为是相互独立的, 从而

$$P\{X=Y\}=\sum_{i=0}^{3}P[\{X=i\}\cap\{Y=i\}]=\sum_{i=0}^{3}P\{X=i\}P\{Y=i\}$$
$$=(1-0.6)^3(1-0.7)^3+C_3^1 0.6(1-0.6)^2 C_3^1 0.7(1-0.7)^2$$
$$+C_3^2 0.6^2(1-0.6)C_3^2 0.7^2(1-0.7)+0.6^3\times 0.7^3=0.3208.$$

(2) 按题意需求事件$\{X>Y\}$的概率, 而事件$\{X>Y\}$可表示为下列两两互不相容的事件之和, 即

$$\{X>Y\}=[\{X=1\}\cap\{Y=0\}]\cup[\{X=2\}\cap\{Y\le 1\}]\cup[\{X=3\}\cap\{Y\le 2\}].$$

由于甲、乙投中与否相互独立, 所以

$$P\{X>Y\}=P[\{X=1\}\cap\{Y=0\}]+P[\{X=2\}\cap\{Y\le 1\}]+P[\{X=3\}\cap\{Y\le 2\}]$$
$$=P\{X=1\}P\{Y=0\}+P\{X=2\}\cdot[P\{Y=0\}+P\{Y=1\}]$$
$$+P\{X=3\}\cdot[1-P\{Y=3\}]$$
$$=C_3^1 0.6(1-0.6)^2(1-0.7)^3+C_3^2 0.6^2(1-0.6)\times[(1-0.7)^3$$
$$+C_3^1 0.7(1-0.7)^2]+0.6^3(1-0.7^3)$$
$$=0.2430.$$

10. 某产品的不合格率为 0.1, 每次随机抽取 10 件进行检验, 若发现有不合格品, 就去调整设备. 若检验员每天检验 4 次, 试求每天调整次数的分布律.

分析 把每抽取一件产品看作一次试验, 每一次试验的结果只有两个: 抽取合格品或不合格品. 现在所给的试验(序列)模型是: 若抽取了一个不合格品, 则去调整设备; 若调整设备后抽取了一个是合格品, 则继续抽取下去, 直至抽取了一个不合格品, 又得去调整设备. 因此, 在 10 件产品中至少有一个不合格品就需调整设备.

解 记A表示事件“抽取的一个产品是不合格品”, 由题意

$$P(A)=0.1, P(\overline{A})=1-P(A)=0.9,$$

设抽取的不合格品数为X, 则$X\sim B(10,0.1)$.则设备需要调整的概率为

$$P\{X\ge 1\}=1-P\{X=0\}=1-C_{10}^0 0.1^0\times 0.9^{10}=0.6513,$$

由于检验员每天检验 4 次, 因此每天调整次数Y的分布律为$Y\sim B(4,0.6513)$.

11. 保险公司在一天内承保了 5000 份相同年龄为期一年的寿险保单, 每人一份. 在合同的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).

分析 这是二项分布的概率计算问题, 由于n较大, p很小, 所以用泊松定理作近似计算.

解 用X表示 5000 个投保人最终死亡的人数, 则$X\sim B(5000,0.0015)$, $n=5000$, $p=0.0015$, $np=7.5$, 由泊松定理知, X近似服从参数为 7.5 的泊松分布.

$$P\{3X\le 30\}=P\{X\le 10\}=\sum_{k=0}^{10}C_{5000}^k\cdot 0.0015^k\cdot 0.9985^{5000-k}$$

$$\approx \sum_{k=0}^{10} \frac{7.5^k}{k!} e^{-7.5} = 0.8622 \cdot$$

注 二项分布的泊松近似(即泊松定理),常常应用于如下问题:在一次试验中事件A发生的概率很小,但独立重复试验的次数n很大,求事件A恰好或至少发生一次或几次的概率.如求某段高速公路上至少发生一起交通事故的概率,或求保险业务中恰有、多于或少于几起理赔发生的概率等.

12. 设随机变量X服从泊松分布,且已知$P\{X=1\}=P\{X=2\}$,求$P\{X=4\}$.

解 X的分布律为

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

由$P\{X=1\}=P\{X=2\}$,即

$$\frac{\lambda^1}{1}e^{-\lambda}=\frac{\lambda^2}{2}e^{-\lambda},$$

解得$\lambda=2$,($\lambda=0$舍去).从而得

$$P\{X=4\}=\frac{2^4}{4!}e^{-2}=\frac{2}{3}e^{-2}.$$

13. 假设某电话总机每分钟接到的呼唤次数服从参数为5的泊松分布,求:(1) 某分钟内恰好接到6次呼唤的概率;(2) 某分钟内接到的呼唤次数多于4次的概率.

解 设X是电话总机每分钟接到的呼唤次数,则$X\sim P(5)$,从而

(1) 某分钟内恰好接到6次呼唤的概率为

$$P\{X=6\}=\frac{5^6}{6!}e^{-5}\approx 0.1462;$$

(2) 某分钟内接到的呼唤次数多于4次的概率为

$$P\{X>4\}=1-P\{X\le 4\} = 1-\sum_{k=0}^{4}\frac{5^k}{k!}e^{-5}\approx 0.5595.$$

注 泊松分布不仅仅是二项分布的一种近似分布,在实际生活有大量的随机变量都服从泊松分布.例如,在一定时间内传呼台收到的呼叫次数;一定时间内,在超级市场排队等候付款的顾客人数;一匹布上的瑕点个数;一定区域内在显微镜下观察到的细菌个数;一定页数的书上出现印刷错误的页数等.

14. 某110接警台在长度为t(单位:h)的时间间隔内收到的报警电话次数服从参数为$2t$的泊松分布,而且与时间间隔的起点无关,求:(1)某天 8 点到 11 点没有接到报警电话的概率;(2)某天8点到12点至少接到1个报警电话的概率.

解 已知报警电话次数$X\sim P(2t)$.

(1) $t=3$,所求概率为

$$P\{X=0\}=\frac{6^0}{0!}e^{-6}=e^{-6};$$

(2) $t=4$, 所求概率为

$$P\{X\geq 1\}=1-P\{X=0\}=1-\frac{8^0}{0!}e^{-8}=1-e^{-8}.$$

15. 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<-1,\\ 0.4, & -1\leq x<1,\\ 0.8, & 1\leq x<3,\\ 1, & x\geq 3,\end{cases}$$

试求 X 的分布律.

分析 这是已知离散型随机变量 X 的分布函数求其分布律的问题, 只需对 $F(x)$ 的所有分段点 x_0, (也就是 X 的所有可能取值点) 利用公式 $P\{X=x_0\}=F(x_0)-F(x_0-0)$ 即可.

解 X 的所有可能取值为 $-1,1,3$,

$$P\{X=-1\}=F(-1)-F(-1-0)=0.4-0=0.4,$$

$$P\{X=1\}=F(1)-F(1-0)=0.8-0.4=0.4,$$

$$P\{X=3\}=F(3)-F(3-0)=1-0.8=0.2,$$

从而 X 的分布律为

X	-1	1	3
P	0.4	0.4	0.2

16. 设随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<0,\\ x^2, & 0\leq x<1,\\ 1, & x\geq 1,\end{cases}$$

试求 $P\{X\leq 0.5\}$, $P\{-1<X\leq 0.25\}$.

解 由连续型随机变量分布函数的定义和性质, 有

$$P\{X\leq 0.5\}=F(0.5)=0.5^2=\frac{1}{4},$$

$$P\{-1<X\leq 0.25\}=F(0.25)-F(-1)=0.25^2-0=\frac{1}{16}.$$

17. 设连续型随机变量 X 的分布函数为

$$F(x)=\begin{cases}a+be^{-\frac{x^2}{2}}, & x\geq 0,\\ 0, & x<0,\end{cases}$$

求:(1) 常数 a 和 b;(2) 随机变量 X 的密度函数.

分析 求分布函数中的待定常数, 一般做法是: 根据分布函数的性质,(主要是 $\lim\limits_{x\to+\infty} F(x)=1$, $\lim\limits_{x\to-\infty} F(x)=0$, 连续性), 列出含有待定常数的等式(方程)解之即可.

解 (1) 因为连续型随机变量的分布函数为连续函数, 则

$$\lim_{x\to+\infty} F(x)=\lim_{x\to+\infty}(a+b\mathrm{e}^{-\frac{x^2}{2}})=a=1,$$

$$\lim_{x\to0^+} F(x)=\lim_{x\to0^+}(1+b\mathrm{e}^{-\frac{x^2}{2}})=1+b=0=F(0),$$

所以 $b=-1$;

(2) 随机变量 X 的密度函数

$$f(x)=F'(x)=\begin{cases} x\mathrm{e}^{-\frac{x^2}{2}}, & x>0, \\ 0, & x\le 0. \end{cases}$$

18. 设随机变量 X 的密度函数为

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

求:(1) 常数 k;(2) $P\{-0.5<X\le 0.5\}$;(3) 分布函数 $F(x)$.

分析 (1) 连续型随机变量 X 的密度函数 $f(x)$ 中的一个待定常数, 可由性质 $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$ 确定. (2) 求连续型随机变量 X 落在某一区间 $[a,b]$ 上的概率, 有两种方法: 一种是计算分布函数在区间 $[a,b]$ 上的增量, 即 $F(b)-F(\mathrm{a})$; 另一种是计算密度函数在区间 $[a,b]$ 上的积分, 即 $\int_a^b f(x)\mathrm{d}x$. (3) 由于密度函数 $f(x)$ 是分段表示的, 所以求分布函数 $F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$ 时, 必须对 x 分区间进行讨论.

解 (1) $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=\int_{-1}^{1}(k-|x|)\mathrm{d}x=2\int_0^1(k-x)\mathrm{d}x=2k-1=1$, 从而解得 $k=1$;

(2) $P\{-0.5<X\le 0.5\}=\int_{-0.5}^{0.5}(1-|x|)\,\mathrm{d}x=2\int_0^{0.5}(1-x)\mathrm{d}x=1-x^2\Big|_0^{0.5}=0.75$;

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

当 $x<-1$ 时,

$$F(x)=P\{X\le x\}=\int_{-\infty}^{x} f(t)\,\mathrm{d}t=0;$$

当 $-1\le x<0$ 时,

$$F(x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t=\int_{-1}^{x} f(t)\,\mathrm{d}t=\int_{-1}^{x}(1+t)\mathrm{d}t=0.5x^2+x+0.5;$$

当 $0 \le x < 1$ 时,

$$F(x) = \int_{-1}^{0}(1+t)\mathrm{dt} + \int_{0}^{x}(1-t)\mathrm{d}t = -0.5x^2 + x + 0.5\ ;$$

当 $x \ge 1$ 时,

$$F(x) = \int_{-1}^{0}(1+t)\mathrm{dt} + \int_{0}^{1}(1-t)\mathrm{d}t = 1\ ,$$

因此, 分布函数为

$$F(x) = \begin{cases} 0, & x < -1, \\ 0.5x^2 + x + 0.5, & -1 \le x < 0, \\ -0.5x^2 + x + 0.5, & 0 \le x < 1, \\ 1, & x \ge 1. \end{cases}$$

19. 设随机变量 X 的密度函数为

$$f(x) = \begin{cases} A(9 - x^2), & -3 \le x \le 3, \\ 0, & 其他, \end{cases}$$

求:(1) 常数 A ;(2) $P\{X < 0\}$, $P\{X > 2\}$, $P\{-1 < X < 1\}$;(3) 分布函数 $F(x)$.

解 (1) $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = \int_{-3}^{3} A(9 - x^2)\mathrm{d}x = 2A\int_{0}^{3}(9 - x^2)\mathrm{d}x = 36A = 1$,

从而解得 $A = \dfrac{1}{36}$;

(2) $P\{X < 0\} = \int_{-3}^{0} \dfrac{1}{36}(9 - x^2)\,\mathrm{d}x = 0.5$,

$$P\{X > 2\} = \int_{2}^{3} \frac{1}{36}(9 - x^2)\,\mathrm{d}x = \frac{2}{27},$$

$$P\{-1 < X < 1\} = 2\int_{0}^{1} \frac{1}{36}(9 - x^2)\,\mathrm{d}x = \frac{13}{27}\ ;$$

(3) 当 $x < -3$ 时, $F(x) = 0$;

当 $-3 \le x < 3$ 时,

$$F(x) = \int_{-3}^{x} f(t)\,\mathrm{d}t = \int_{-3}^{x} \frac{1}{36}(9 - t^2)\,\mathrm{d}t == \frac{1}{2} + \frac{1}{4}x - \frac{1}{108}x^3\ ;$$

当 $x \ge 3$ 时, $F(x) = \int_{-3}^{3} \dfrac{1}{36}(9 - t^2)\,\mathrm{d}t = 1$,

因此, 分布函数为

$$F(x) = \begin{cases} 0, & x < -3, \\ \frac{1}{2} + \frac{1}{4}x - \frac{1}{108}x^3, & -3 \le x < 3, \\ 1, & x \ge 3. \end{cases}$$

20. 已知随机变量 X 的密度函数为 $f(x)=Ae^{-|x|}$，试求:(1) 常数 A;(2) X 的分布函数.

解(1) 因为 $\int_{-\infty}^{+\infty}Ae^{-|x|}dx=2A\int_{0}^{+\infty}e^{-x}dx=-2Ae^{-x}\Big|_{0}^{+\infty}=2A=1$，从而 $A=0.5$;

(2) 当 $x<0$ 时，$F(x)=P\{X\le x\}=\int_{-\infty}^{x}0.5e^{t}dt=0.5e^{x}$;

当 $x\ge 0$ 时，$F(x)=\int_{-\infty}^{0}0.5e^{t}dt+\int_{0}^{x}0.5e^{-t}dt=1-0.5e^{-x}$,

因此分布函数为

$$F(x)=\begin{cases}0.5e^{x}, & x<0,\\ 1-0.5e^{-x}, & x\ge 0.\end{cases}$$

21. 某城市每天用电量不超过一百万 kWh，以 X 表示每天的耗电率(即用电量除以百万 kWh)，它具有密度函数:

$$f(x)=\begin{cases}12x(1-x)^2, & 0\le x\le 1,\\ 0, & \text{其他}.\end{cases}$$

若该城市每天供电量仅 80 万 kWh，求供电量不够需要的概率.若每天的供电量上升到 90 万 kWh，每天供电量不足的概率是多少?

解 若该城市每天供电量仅 80 万 kWh，则供电量不够需要的概率即为

$$P\{X>\frac{80}{100}\}=\int_{0.8}^{1}12x(1-x)^2dx=12\int_{0.8}^{1}(x-2x^2+x^3)dx=\frac{17}{625},$$

若每天的供电量上升到 90 万 kWh. 时，则每天供电量不足的概率为

$$P\{X>\frac{90}{100}\}=\int_{0.9}^{1}12x(1-x)^2dx=0.0037.$$

22. 假设某种设备的使用寿命 X (年)服从参数为 0.25 的指数分布. 制造这种设备的厂家规定，若设备在一年内损坏，则可以调换. 如果厂家每售出一台设备可赢利 100 元，而设备一台设备厂家要花费 300 元，求每台设备所获利润的分布律.

解 因为该设备的使用寿命 $X\sim E(0.25)$，故 X 的密度函数为

$$f(x)=\begin{cases}0.25e^{-0.25x}, & x>0,\\ 0, & x\le 0.\end{cases}$$

设事件 A 表示设备需调换，则

$$P(A)=P\{X<1\}=\int_{0}^{1}0.25e^{-0.25t}dt=-e^{-0.25t}\Big|_{0}^{1}=1-e^{-0.25},$$

令随机变量 Y 表示每台设备所获利润，则 Y 的分布律为

Y	100	-200
P	$e^{-0.25}$	$1-e^{-0.25}$

23. 设随机变量 X 的密度函数为

$$f(x)=\begin{cases}A\sin x, & 0\le x\le \pi,\\ 0, & 其他,\end{cases}$$

对 X 独立观察 4 次, 随机变量 Y 表示观察值大于 $\dfrac{\pi}{3}$ 的次数, 求: (1) 常数 A; (2) Y 的分布律.

解 (1) 由 $\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=\int_{-\infty}^{0}0\mathrm{d}t+\int_{0}^{\pi}A\sin t\mathrm{d}t+\int_{\pi}^{+\infty}0\mathrm{d}x=-A\cos t\Big|_{0}^{\pi}=2A=1$,

解得 $A=0.5$;

(2) 设事件 A 表示观察值大于 $\dfrac{\pi}{3}$, 则

$$P(A)=P\{X>\frac{\pi}{3}\}=\int_{\frac{\pi}{3}}^{\pi}0.5\mathrm{sint}\mathrm{d}t=\text{-}0.5\mathrm{cost}\Big|_{\frac{\pi}{3}}^{\pi}=0.75,$$

依题意, Y 服从二项分布 $B(4,0.75)$.

24. 某仪器装有 3 个独立工作的同型号电子元件, 其寿命 X (单位: h) 的密度函数为

$$f(x)=\begin{cases}\dfrac{100}{x^2}, & x>100,\\ 0, & x\le 100,\end{cases}$$

试求: (1) X 的分布函数; (2) 在最初的 150 小时内没有一个电子元件损坏的概率.

解 (1) 当 $x<100$ 时, $F(x)=0$,

当 $x\ge 100$ 时,

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=\int_{100}^{x}\frac{100}{t^2}\mathrm{d}t=1-\frac{100}{x},$$

因此分布函数为

$$F(x)=\begin{cases}1-\dfrac{100}{x}, & x\ge 100,\\ 0, & x<100.\end{cases}$$

(2) 没有一个电子元件损坏的充要条件是每个元件都能正常工作, 而这里三个元件的工作是相互独立的, 因此, 若用 A 表示“在最初的 150 小时内没有一个电子元件损坏”, 则

$$P(A)=[P\{X>150\}]^3=\left(\int_{150}^{+\infty}\frac{100}{x^2}\mathrm{d}x\right)^3=\left(\frac{2}{3}\right)^3=\frac{8}{27},$$

25. 公共汽车站每隔 10 分钟有一辆汽车通过, 乘客到达汽车站的是等可能的, 求乘客候车时间不超过 3 分钟的概率.

解 设乘客的候车时间为 X, 则 $X\sim U[0,10]$, 其密度函数为

$$f(x)=\begin{cases}\dfrac{1}{10}, & 0\le x\le 10,\\ 0, & \text{其他},\end{cases}$$

因此乘客候车时间不超过 3 分钟的概率为

$$P\{X\le 3\}=\int_0^3 f(x)\mathrm{d}x=\int_0^3\frac{1}{10}\mathrm{d}x=0.3.$$

26. 设 $X\sim N(1,4)$,(1) 求 $P\{0<X<5\}$;(2) 求 $P\{|X|>2\}$;(3) 设 c 满足 $P\{X>c\}\ge 0.95$,问 c 至多为多少?

分析 服从正态分布的随机变量 X 落在某区间内的概率的计算,通常是先将 X 标准化,然后查标准正态分布表即可.

解 因为 $X\sim N(\mu,\sigma^2)$,所以 $\dfrac{X-\mu}{\sigma}\sim N(0,1)$,这里 $\mu=1,\ \sigma=2$, 故

(1) $P\{0<X<5\}=P\{\dfrac{0-1}{2}<\dfrac{X-1}{2}<\dfrac{5-1}{2}\}=\Phi(2)-\Phi(-0.5)$

$$=\Phi(2)+\Phi(0.5)-1\overset{\text{查表得}}{=}0.9772+0.6915-1=0.6687;$$

(2) $P(|X|>2)=1-P\{|X|\le 2\}=1-P\{-2\le X\le 2\}$

$$=1-[\Phi(0.5)-\Phi(-1.5)]=2-\Phi(0.5)-\Phi(1.5)=0.3753.$$

(3) 因为 $P\{X>c\}=1-P\{X\le c\}=1-\Phi\left(\dfrac{c-1}{2}\right)\ge 0.95$,

即 $\Phi\left(\dfrac{1-c}{2}\right)=0.95$,而 $\Phi(1.64)=0.95$,则 $\dfrac{1-c}{2}=1.64$,因此 $c=-2.2897$.

27. 从南郊某地乘车前往北区火车站乘火车有二条线路可走,第一条路线穿过市区,路程较短,但交通拥挤,所需时间(单位:min)服从正态分布 $N(50,100)$,第二条路线沿环城公路走,路程较长,但意外阻塞少,所需时间服从正态分布 $N(60,16)$,(1) 假如有 70 min 可用,应走哪一条路线?(2) 若只有 65 min 可用,应走哪一条路线?

解 记行走时间为 t,则 $t\sim N(50,100)$,$t\sim N(60,16)$.

(1) 走第一条路线能及时赶到的概率为

$$P\{t\le 70\}=\Phi(\frac{70-50}{10})=\Phi(2)=0.9772,$$

走第二条路线能及时赶到的概率为

$$P\{t\le 70\}=\Phi(\frac{70-60}{4})=\Phi(2.5)=0.9938,$$

因此,若有 70 分钟可用,应选第二条路线;

（2）第一条：$P\{t \le 65\} = \Phi(\frac{65-50}{10}) = \Phi(1.5) = 0.9332$，

第二条：$P\{t \le 65\} = \Phi(\frac{65-60}{4}) = \Phi(1.25) = 0.8944$，

因此，若只有 65 分钟可用，应选第一条路线.

28. 某地抽样调查结果表明，考生的外语成绩(百分制)服从正态分布 $N(72,\sigma^2)$，已知 96 分以上的占考生总数的 2.3%，试求考生的外语成绩在 60 分至 84 分之间的概率.

解 设考生的英语成绩为 X，则 $X \sim N(\mu,\sigma^2)$，其中 $\mu = 72$，由题意，

$$P\{X \ge 96\} = 1 - P\{X < 96\} = 1 - P\{\frac{X-72}{\sigma} < \frac{96-72}{\sigma}\} = 1 - \Phi(\frac{24}{\sigma}) = 0.023,$$

查表得 $\frac{24}{\sigma} = 2$，从而 $\sigma = 12$，故 $X \sim N(72,12^2)$，则

$$P\{60 \le X \le 84\} = P\{\frac{60-72}{12} \le \frac{X-72}{12} \le \frac{84-72}{12}\} = 2\Phi(1) - 1 \approx 0.6826.$$

29．设某地区成人的身高(单位：cm)服从正态分布 $N(172,64)$，问公共汽车的车门的高度为多少时才能以 95%的概率保证该地区的成人在乘车时不会碰到车门？

解 设该地区成人的身高为 X，车门高度为 h，本题为已知 $P\{X < h\} \ge 0.95$，求 h.

因为 $X \sim N(172, 8^2)$，由题设：

$$P\{X < h\} = P\{\frac{X-172}{8} < \frac{h-172}{8}\} = \Phi(\frac{h-172}{8}) \ge 0.95,$$

查表可知，

$$\Phi(1.65) = 0.9505 > 0.95,$$

于是，$\frac{h-172}{8} = 1.65$，解得 $h = 185.1588$. 故取 $h = 186$，即车门高度应定为 186 厘米，男子与车门碰头的机会不超过 0.05.

30. 在电源电压(单位：V)不超过 200、在 200—240 和超过 240 三种情形下，某种电子元件损坏的概率分别为 0.1，0.001 和 0.2. 假设电源电压 X 服从正态分布 $N(220,25^2)$，试求：(1) 该电子元件损坏的概率；(2) 该电子元件损坏时，电源电压在 200—240 的概率.

解 设事件 $A_1 =$ {电压不超过 200 伏}，$A_2 =$ {电压不超过 200-240 伏}，$A_3 =$ {电压超过 240 伏}；$B =$ {电子元件损坏}.

由条件知 $X \sim N(220,25^2)$，因此

$$P(A_1) = P\{X \le 200\} = P\left\{\frac{X-220}{25} \le \frac{200-220}{25}\right\} = 1 - \Phi(0.8) = 0.2119,$$

$$P(A_2) = P\{200 \le X \le 240\} = P\left\{-0.8 \le \frac{X-220}{25} \le 0.8\right\} = 2\Phi(0.8) - 1 = 0.5762,$$

$P(A_3) = P\{X > 240\} = 1 - P\{X \le 240\} = 1 - \Phi(0.8) = 0.2119$.

(1) 由题设条件知

$$P(B \mid A_1) = 0.1, \quad P(B \mid A_2) = 0.001, \quad P(B \mid A_3) = 0.2,$$

于是由全概率公式得

$$P(B) = \sum_{i=1}^{3} P(A_i) P(B \mid A_i) = 0.0641;$$

(2) 由贝叶斯公式得

$$P(A_2 \mid B) = \frac{P(A_2)P(B \mid A_2)}{P(B)} \approx 0.0090.$$

31. 设随机变量 X 的分布律为

X	-1	0	1	4
P	0.1	0.4	0.3	0.2

试求 $Y = X^2$ 的分布律.

分析 求离散型随机变量函数 $Y = g(X)$ 的分布律的步骤: 第一步, 求 Y 的所有可能值: $y_i = g(x_i)$, $i = 1,2,3,\cdots$; 第二步, 求 Y 取每一个可能值 y_i 的概率 $P\{Y = y_i\} = P\{X = x_i\}$ ($i = 1,2,3,\cdots$). 注意应将相同的 y_i 的值所对应的概率相加.

解 Y 的所有可能取值 $0, 1, 16$, 由

$$P\{Y = 0\} = P\{X^2 = 0\} = P\{X = 0\} = 0.4,$$
$$P\{Y = 1\} = P\{X = 1\} + P\{X = -1\} = 0.4,$$
$$P\{Y = 16\} = P\{X = 4\} = 0.2,$$

得 Y 的分布律为

Y	0	1	16
P	0.4	0.4	0.2

32. 设随机变量 X 服从 $U(-1,2)$, 定义

$$Y = \begin{cases} 1, & X \ge 0, \\ -1, & X < 0, \end{cases}$$

试求随机变量 Y 的分布律.

解 因为 $X \sim U(-1,2)$, 其密度函数为

$$f(x) = \begin{cases} \dfrac{1}{3}, & -1 \le x \le 2, \\ 0, & \text{其他}, \end{cases}$$

因此

$$P\{Y=1\}=P\{X\geq 0\}=\int_0^2 f(x)\mathrm{d}x=\int_0^2\frac{1}{3}\mathrm{d}x=\frac{2}{3},$$

$$P\{Y=-1\}=P\{X<0\}=\int_{-1}^0 f(x)\mathrm{d}x=\int_{-1}^0\frac{1}{3}\mathrm{d}x=\frac{1}{3},$$

所以Y的分布律为

Y	-1	1
P	$\frac{1}{3}$	$\frac{2}{3}$

33. 设随机变量X服从$U(0,2)$，试求随机变量$Y=X^2$的密度函数.

分析 求连续型随机变量函数$Y=X^2$的密度函数，由于$y=x^2$不是单调函数，因此用分布函数法(一般步骤见内容提要).

解 设X,Y的分布函数分别为$F_X(x)$，$F_Y(y)$，由$Y=X^2$可知，

当$y<0$时，

$$F_Y(y)=P\{Y\leq y\}=P\{X^2\leq y\}=0,$$

于是，Y的密度函数$f_Y(y)=F_Y'(y)=0$；

当$0\leq y<4$时，

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

上式两边关于y求导，注意到复合函数求导法则，有

$$\begin{aligned}f_Y(y)=F_Y'(y)&=F_X'(\sqrt{y})\cdot(\sqrt{y})'-F_X'(-\sqrt{y})\cdot(-\sqrt{y})'\\&=\frac{1}{2\sqrt{y}}f_X(\sqrt{y})+\frac{1}{2\sqrt{y}}f_X(-\sqrt{y}).\end{aligned}$$

因此，Y的密度函数为

$$f_Y(y)=\begin{cases}\dfrac{1}{2\sqrt{y}}[f_X(\sqrt{y})+f_X(-\sqrt{y})], & 0\leq y<4,\\ 0, & y<0.\end{cases}$$

由于$X\sim U(0,2)$，其密度函数为

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

从而 $$f_Y(y)=\begin{cases}\dfrac{1}{4\sqrt{y}}, & 0<y<4,\\ 0, & \text{其他}.\end{cases}$$

34. 设随机变量 X 的密度函数为

$$f(x)=\begin{cases}\mathrm{e}^{-x}, & x\ge 0,\\ 0, & x<0.\end{cases}$$

求随机变量 $Y=\mathrm{e}^{X}$ 的密度函数.

分析 求连续型随机变量函数 $Y=\mathrm{e}^{X}$ 的密度函数, 由于 $y=\mathrm{e}^{x}$ 是单调函数, 因此用连续型随机变量的单调函数的分布密度的公式法(一般步骤见内容提要), 使用时必须注意条件.

解 因为 $y=\mathrm{e}^{x}$ 为严格单调增函数, 其反函数为 $x=h(y)=\ln y$, 又 $y=\mathrm{e}^{x}$ 的值域为 $[1,+\infty)$, 由连续型随机变量函数的密度函数的公式,

当 $y>1$ 时, 有

$$f_Y(y)=f_X(h(y))\cdot|h'(y)|=f_X(\ln y)\cdot\frac{1}{y}=\mathrm{e}^{-\ln y}\cdot\frac{1}{y}=\frac{1}{y^2},$$

因而 Y 的密度函数为

$$f_Y(y)=\begin{cases}\dfrac{1}{y^2}, & y>1,\\ 0, & 0<y\le 1.\end{cases}$$

35. 假设随机变量 $X\sim N(0,1)$, 求下列随机变量 Y 的密度函数.

(1) $Y=\mathrm{e}^{X}$; (2) $Y=2X^2+1$; (3) $Y=|X|$.

解 设 $F_Y(y), f_Y(y)$ 分别为随机变量 Y 的分布函数和概率密度函数.

(1) 当 $y<0$ 时, $F_Y(y)=P\{Y\le y\}=P\{\mathrm{e}^{X}\le y\}=0$,

当 $y\ge 0$ 时,

$$F_Y(y)=P\{Y\le y\}=P\{\mathrm{e}^{X}\le y\}=P\{X\le \ln y\}$$

$$=F_X(\ln y)=\Phi(\ln y)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\ln y}\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x,$$

再由 $f_Y(y)=F_Y'(y)$, 利用变限积分求导, 得

$$f_Y(y)=F_Y'(y)=\begin{cases}\dfrac{1}{y}\varphi(\ln y), & y>0,\\ 0, & y\le 0\end{cases}=\begin{cases}\dfrac{1}{y\sqrt{2\pi}}\mathrm{e}^{-\frac{(\ln y)^2}{2}}, & y>0,\\ 0, & \text{其他}.\end{cases}$$

注 通常称上式中的 Y 服从对数正态分布, 它也是一种常用寿命分布.

(2) 当 $y<1$ 时, $F_Y(y)=P\{Y\le y\}=P\{2X^2+1\le y\}=0$,

当 $y\ge 1$ 时,

$$F_Y(y)=P\{Y\le y\}=P\{2X^2+1\le y\}$$

$$=P\{-\sqrt{\frac{y-1}{2}}\le X\le\sqrt{\frac{y-1}{2}}\}=2\Phi(\sqrt{\frac{y-1}{2}})-1,$$

从而$Y=2X^2+1$的密度函数为

$$f_Y(y)=F_Y'(y)=\begin{cases}\dfrac{1}{\sqrt{2(y-1)}}\varphi(\sqrt{\dfrac{y-1}{2}}), & y>1,\\ 0, & y\le 1\end{cases}=\begin{cases}\dfrac{1}{2\sqrt{\pi(y-1)}}\mathrm{e}^{-\frac{y-1}{4}}, & y>1,\\ 0, & y\le 1.\end{cases}$$

（3）当$y<0$时，$F_Y(y)=P\{Y\le y\}=P\{|X|\le y\}=0$，

当$y\ge 0$时,

$$F_Y(y)=P\{Y\le y\}=P\{|X|\le y\}=P\{-y\le X\le y\}=2\Phi(y)-1,$$

从而$Y=|X|$的密度函数为

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

36. 设随机变量X的密度函数为

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

求随机变量$Y=\ln X$的密度函数.

解　因为$y=\ln x$为严格单调增函数, 其反函数为$x=h(y)=\mathrm{e}^y$, 又$y=\ln x$在区间（0, 1）上的值域为$(-\infty,0)$, 由连续型随机变量函数的密度函数的公式,

当$y<0$时,

$$f_Y(y)=f_X(h(y))\cdot|h'(y)|=f_X(\mathrm{e}^y)\cdot\mathrm{e}^y=2\mathrm{e}^y\cdot\mathrm{e}^y=2\mathrm{e}^{2y},$$

因而Y的密度函数为

$$f_Y(y)=\begin{cases}2\mathrm{e}^{2y}, & y<0,\\ 0, & y\ge 0.\end{cases}$$

37. 设随机变量X服从参数为 2 的指数分布, 证明$Y=\mathrm{e}^{-2X}$服从$U(0,1)$.

证法 1　因为$X\sim E(2)$, 所以X的密度函数为

$$f_X(x)=\begin{cases}2\mathrm{e}^{-2x}, & x\ge 0,\\ 0, & x<0.\end{cases}$$

当$y<0$时, $F_Y(y)=0$；

当$0\le y<1$时,

$$\begin{aligned}F_Y(y)&=P\{\mathrm{e}^{-2X}\le y\}=P\{X\ge-\frac{1}{2}\ln y\}\\&=1-P\{X<-\frac{1}{2}\ln y\}=1-\int_0^{-\frac{1}{2}\ln y}2\mathrm{e}^{-2x}\mathrm{d}x,\end{aligned}$$

利用变限积分求导, 得

$$F_Y'(y) = 2e^{\ln y} \cdot \frac{1}{2} \cdot \frac{1}{y} = 1,$$

当 $y \geq 1$ 时,

$$F_Y(y) = P\{e^{-2X} \leq y\} = 1,$$

于是

$$f_Y(y) = F_Y'(y) = \begin{cases} 1, & 0 < y < 1, \\ 0, & 其他, \end{cases}$$

即 Y 服从 $(0,1)$ 上的均匀分布.

证法 2 因为 $Y = e^{-2X}$ 为严格单调减函数, 其反函数为 $x = h(y) = -\frac{1}{2}\ln y$, 又 $y = e^{-2x}$ 的值域为 $(0,1)$, 由连续型随机变量函数的密度函数的公式,

当 $0 < y < 1$ 时,

$$\begin{aligned} f_Y(y) &= f_X(h(y)) \cdot |h'(y)| \\ &= f_X(-\frac{1}{2}\ln y) \cdot \left|(-\frac{1}{2}) \cdot \frac{1}{y}\right| = 2e^{\ln y} \cdot \frac{1}{2} \cdot \frac{1}{y} = 1, \end{aligned}$$

因而

$$f_Y(y) = \begin{cases} 1, & 0 < y < 1, \\ 0, & 其他, \end{cases}$$

即 Y 服从 $(0,1)$ 上的均匀分布.