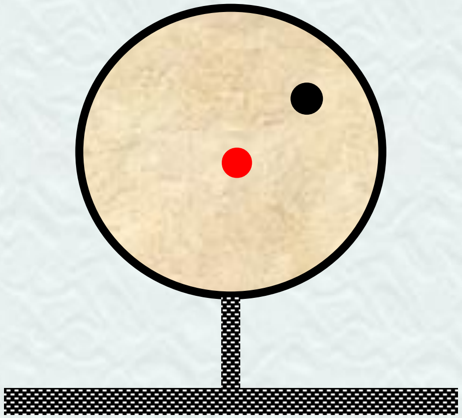
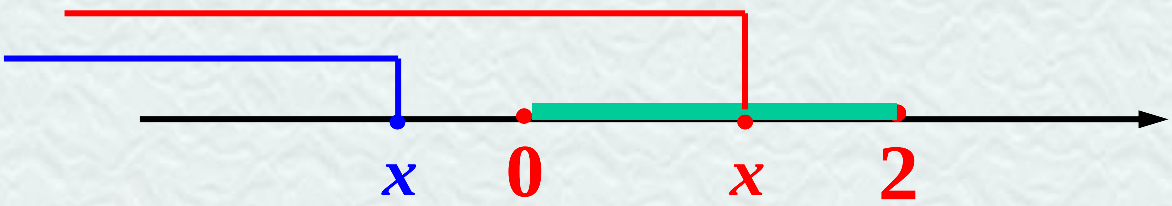
引例 一个靶子是半径为2m的圆盘,设击中靶上任一同心圆盘上的点的概率与该圆盘的面积成正比,并设射击都能中靶,以X表示弹着点与圆心的距离.试求随机变量 X 的分布函数.



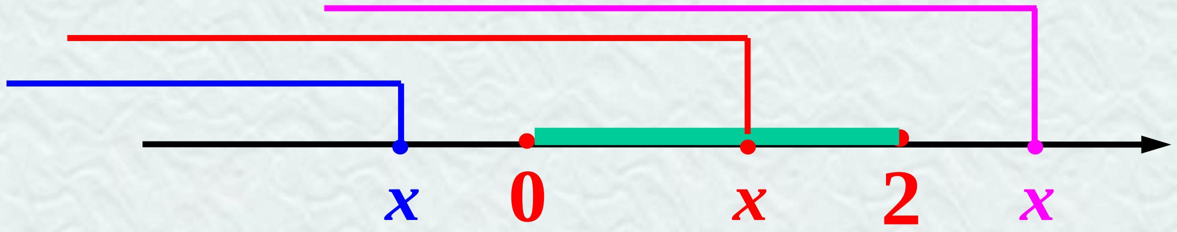
解 r.v. X 的取值范围为 $[0,\ 2]$



当 $x<0$时, $F(x)=P\{X\le x\}=0;$

当$0\le x<2$ 时, $F(x)=P\{X\le x\}$

$$=P\{X<0\}+P\{0\le X\le x\}=0+\frac{\pi x^2}{4\pi}=\frac{x^2}{4}.$$



当 $x < 0$ 时, $F(x) = P\{X \le x\} = 0$;

当 $0 \le x < 2$ 时,

$$F(x) = P\{X \le x\} = P\{0 \le X \le x\} = \frac{\pi x^2}{4\pi} = \frac{x^2}{4}.$$

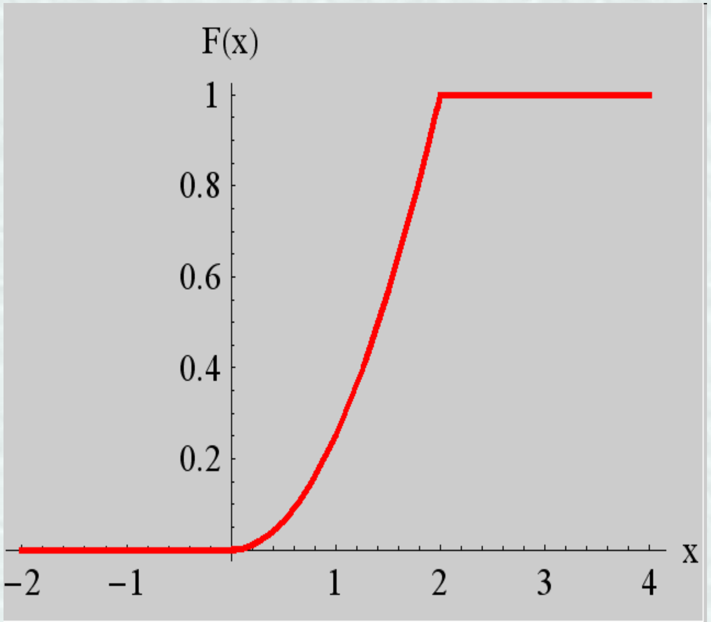
当 $x \ge 2$ 时, $F(x) = P\{X \le x\}$

$$= P\{X < 0\} + P\{0 \le X \le 2\} + P\{2 < X \le x\}$$

$$= P\{0 \le X \le 2\} = 1.$$

故连续型r.v.X 的分布函数为

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



其图形为一连续曲线

注意 两类随机变量的分布函数图形的特点不一样。离散型r.v.的分布函数图形为阶梯状曲线；连续型r.v. 的分布函数图形是连续曲线。

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

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

$F(x)$ 恰是非负函数 $f(t)$ 在区间 $(-\infty,x]$ 上的积分,

此时称 X 为连续型随机变量.

§2.3 一维连续型随机变量

定义 2.3.1 对于随机变量X的分布函数$F(x)$,如果存在非负函数$f(x)$，使对于任意实数有

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

则X称为连续型随机变量，其中函数$f(x)$称为X的概率密度函数，简称密度函数.

由定义知道，概率密度$f(x)$具有以下性质

(1) $f(x) \ge 0$　　(2) $\int_{-\infty}^{+\infty} f(x)dx = 1$

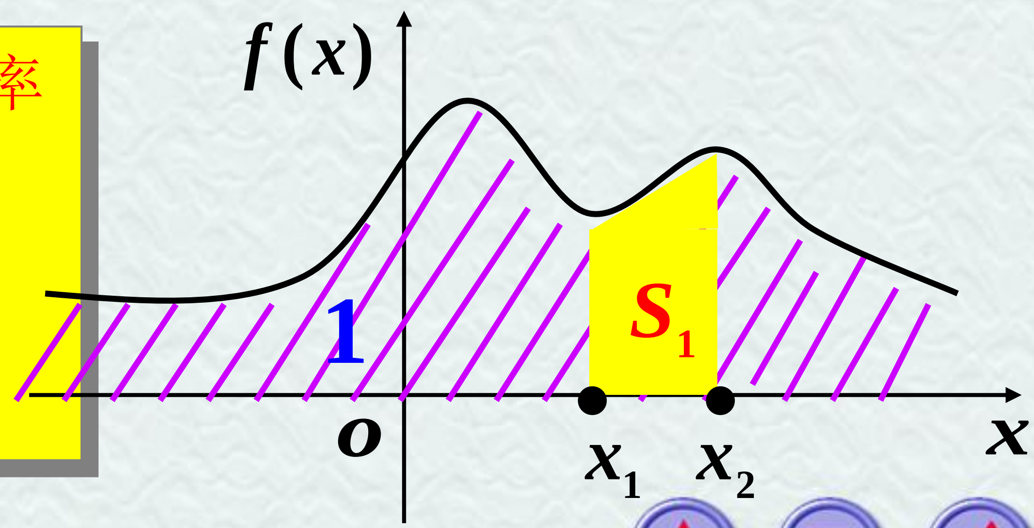
性质(1),(2)是两个最基本的性质

(3) 对于任意实数$x_1, x_2, x_1 \le x_2$,有

$$P\{x_1 < X \le x_2\} = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)dx$$

(4) 若$f(x)$在点x连续，则有$F'(x) = f(x)$

X落在区间$(x_1, x_2]$上的概率$P\{x_1 \le X \le x_2\}$等于区间$(x_1, x_2]$上曲线$f(x)$之下的曲边梯形的面积(如图)



由性质(4)知,对于$f(x)$的连续点x有

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

上式表明密度函数$f(x)$不是随机变量X取值x的概率,而是X在点x的概率分布的密集程度,$f(x)$的大小能反映出X取x附近的值的概率大小．因此对于连续型随机变量，用密度函数描述它的分布比分布函数直观.

注意 对于任意可能值 a ,连续型随机变量取 a 的概率等于零.即

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

证明 $P\{X=a\}=\lim_{\Delta x\to 0}\int_a^{a+\Delta x}f(x)\,\mathrm{d}x=0.$

由此可得

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

连续型随机变量取值落在某一区间的概率与区间的开闭无关

注意

若 X 为离散型随机变量,

$\{X=a\}$是不可能事件 $\Leftrightarrow P\{X=a\}=0.$

若X是连续型随机变量,

$\{X=a\}$是不可能事件 $\underset{\not\Longleftarrow}{\Longrightarrow} P\{X=a\}=0.$

概率为0的事件不一定是不可能事件

概率为1的事件不一定是必然事件

例2.3.2　设连续型随机变量X的分布函数

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

(1)求常数a;

(2)求X的密度函数$f(x)$;

(3)求$P\{|X|\le 0.5\}$.

解 (1) 因为 X 是连续型随机变量, 所以 $F(x)$ 连续,

故有 $\lim\limits_{x\to 1} F(x) = F(1)$, 从而 $\lim\limits_{x\to 1^-} ax^2 = a = 1.$

故 X 的分布函数 $F(x) = \begin{cases} 0, & x < 0, \\ x^2, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$

(2) X 的密度函数 $f(x) = F'(x) = \begin{cases} 2x, & 0 \le x \le 1, \\ 0, & \text{其他}. \end{cases}$

X 的分布函数

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

X 的密度函数

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

(3) $P\{|X|\le 0.5\}$

$$=P\{-0.5\le X\le 0.5\}=F(0.5)-F(-0.5)=0.5^2-0=0.25$$

或 $=\int_{-0.5}^{0.5} f(x)dx=\int_{-0.5}^{0} 0dx+\int_{0}^{0.5} 2xdx=x^2\Big|_0^{0.5}=0.25$

例2.3.3 设连续型随机变量X具有概率密度

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

(1)确定常数k

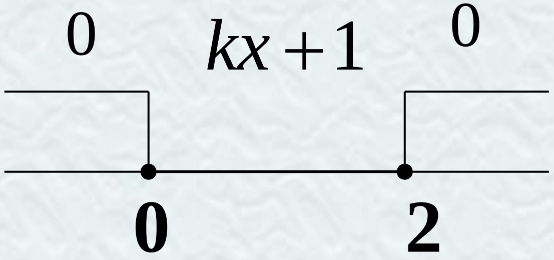
(2)求X的分布函数$F(x)$

(3)求$P\{-1 < X \le 1\}$

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

$$=\int_{-\infty}^{0} 0dx+\int_{0}^{2}(kx+1)\mathrm{d}x+\int_{2}^{+\infty} 0dx \quad \text{解得} k=-1/2$$

(2) X的分布函数为



$$F(x)=P\{X\le x\}=\int_{-\infty}^{x} f(t)dt$$

$$=\begin{cases} \int_{-\infty}^{x} 0dt=0, & x<0 \\ \int_{-\infty}^{0} 0dt+\int_{0}^{x}(-\frac{1}{2}t+1)dt=-\frac{1}{4}x^2+x, & 0\le x<2 \\ \int_{-\infty}^{0} 0dt+\int_{0}^{2}(-\frac{1}{2}t+1)dt+\int_{2}^{x} 0dt=1, & x\ge 2 \end{cases}$$

$$F(x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t=\begin{cases}0, & x<0\\ -\dfrac{1}{4}x^{2}+x, & 0\le x<2\\ 1, & x\ge 2\end{cases}$$

(3) $P\{-1<X\le 1\}=F(1)-F(-1)=\left(-\dfrac{1}{4}+1\right)-0=\dfrac{3}{4}$

或 $=\int_{-1}^{1} f(x)dx=\int_{-1}^{0}0dx+\int_{0}^{1}(-\dfrac{1}{2}x+1)dx=\dfrac{3}{4}$

题型1：求概率密度函数中未知常数

题型2：求连续型r.v.落在区间里的概率

题型3：根据概率密度函数求连续型r.v.的分布函数

题型4：根据分布函数求连续型r.v.的概率密度函数

练习 设随机变量 X 具有概率密度

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

(1) 确定常数 k; (2) 求 X 的分布函数;

(3) 求 $P\{1<X\le \frac{7}{2}\}$.

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

得 $\int_0^3 kx\,\mathrm{d}x + \int_3^4 (2 - \frac{x}{2})\,\mathrm{d}x = 1,$ 解之得 $k = \frac{1}{6}.$

(2) 由 $k = \frac{1}{6}$ 知 X 的概率密度为

$$f(x) = \begin{cases} \frac{x}{6}, & 0 \le x < 3, \\ 2 - \frac{x}{2}, & 3 \le x \le 4, \\ 0, & \text{其它}. \end{cases}$$

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

$$=\begin{cases}\int_{-\infty}^{x} 0dt, & x<0,\\ \int_{-\infty}^{0} 0dt+\int_{0}^{x}\frac{t}{6}\mathrm{d}t, & 0\le x<3,\\ \int_{-\infty}^{0} 0dt+\int_{0}^{3}\frac{t}{6}\mathrm{d}t+\int_{3}^{x}(2-\frac{t}{2})\mathrm{d}t, & 3\le x<4,\\ \int_{-\infty}^{0} 0dt+\int_{0}^{3}\frac{t}{6}\mathrm{d}t+\int_{3}^{4}(2-\frac{t}{2})\mathrm{d}t+\int_{4}^{x}0dt, & x\ge 4.\end{cases}$$

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

(3) $P\{1<X\le \frac{7}{2}\}=F(\frac{7}{2})-F(1)=\dfrac{41}{48}.$

或 $=\int_1^{7/2} f(x)dx=\int_1^3 \dfrac{x}{6}dx+\int_3^{7/2}(2-\dfrac{x}{2})dx=\dfrac{41}{48}$

几种常见的连续型随机变量的分布

（一）均匀分布

设连续型随机变量X具有概率密度函数

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

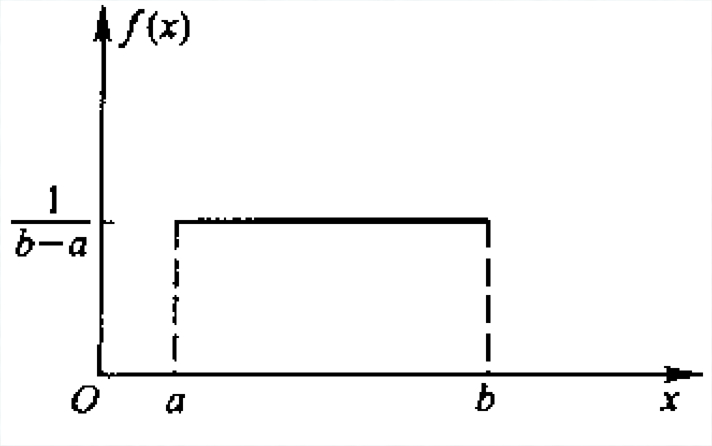
则称X在区间(a,b)上服从均匀分布,记为$X\sim U(a,b)$.

易知$\boxed{f(x)\geq 0,\text{且}\int_{-\infty}^{+\infty}f(x)\mathrm{d}x=1}$

满足连续型随机变量的两个最基本性质

$$=\int_{-\infty}^{a}0dx+\int_{a}^{b}\frac{1}{b-a}dx+\int_{b}^{+\infty}0dx$$

$f(x)$的图形

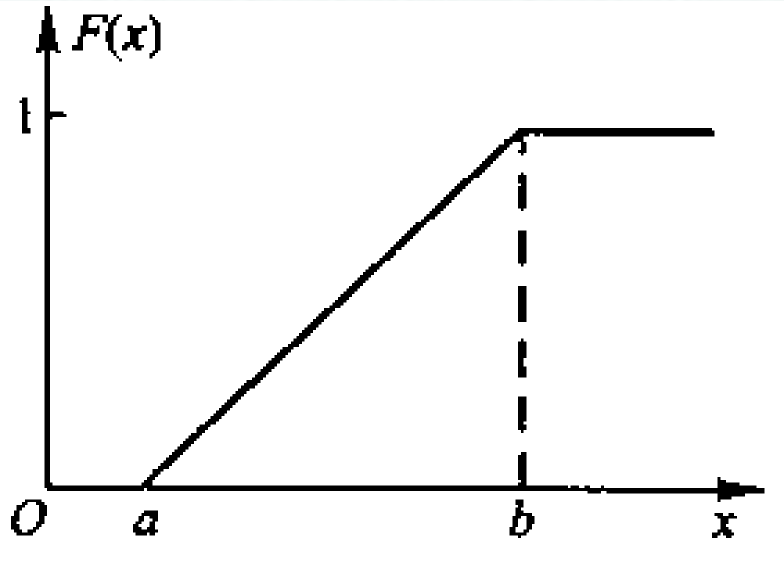


在(a,b)上服从均匀分布的随机变量X落在(a,b)中任一等长度的子区间内的可能性是相同的,或者说X落在(a,b)子区间的概率只依赖于子区间的长度,而与子区间的位置无关.

X的分布函数为

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

$F(x)$相应的图形为



例1 设电阻值 R 是一个随机变量，均匀分布在 $900\,\Omega \sim 1100\,\Omega$. 求 R 的概率密度及 R 落在 $950\,\Omega \sim 1050\,\Omega$ 的概率.

解 由题意,R 的概率密度为

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

故有 $$P\{950 < R \le 1050\} = \int_{950}^{1050} \frac{1}{200}\,\mathrm{d}r = 0.5.$$

例2 设随机变量X在[2，5]上服从均匀分布，现对X进行三次独立观测。求至少有两次观测值大于3的概率。

解： X 的概率密度函数为

$$f(x)=\begin{cases}\frac{1}{3}, & 2\le x\le 5,\\ 0, & \text{其他}.\end{cases}$$

则$P(X>3)=\int_3^5 \frac{1}{3}\mathrm{d}x=\frac{2}{3}$

设Y表示三次独立观测其观测值大于3的次数，则

$$Y\sim b\left(3,\frac{2}{3}\right).$$

$$P(Y\ge 2)=C_3^2\left(\frac{2}{3}\right)^2\cdot\frac{1}{3}+C_3^3\left(\frac{2}{3}\right)^3=\frac{20}{27}$$

（二） 指数分布 记为$X \sim E(\lambda)$.

若连续型随机变量X的概率密度

满足连续型随机变量的两个最基本性质

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

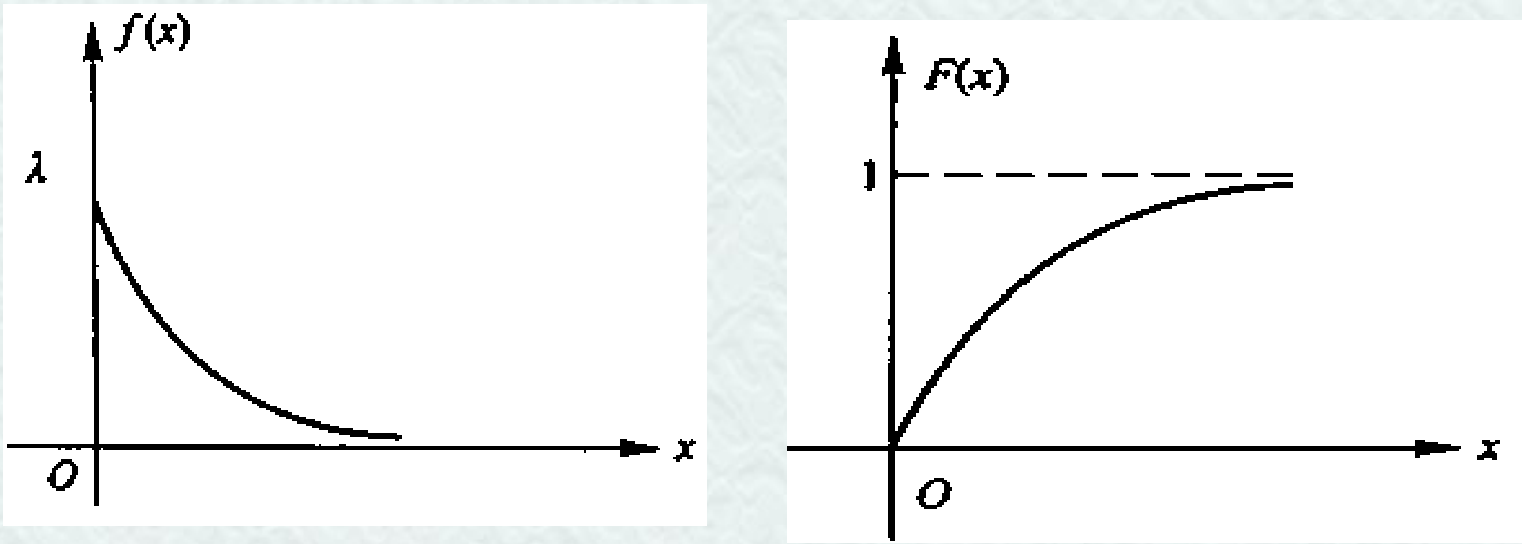
其中 $\lambda>0$ 为常数，则称X服从参数为λ 的指数分布。

易知 $\boxed{f(x)\geq 0,\int_{-\infty}^{+\infty}f(x)dx=\int_{0}^{+\infty}\lambda e^{-\lambda x}dx=1}$

X的分布函数为

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

指数分布的概率密度及分布函数分别如图所示



应用 某些元件或设备的寿命服从指数分布. 例如无线电元件的寿命、电力设备的寿命、动物的寿命等都服从指数分布.

指数分布的重要性质 :“无记忆性”.

$$P\{X>s+t \mid X>s\}=\frac{P\{X>s+t, X>s\}}{P\{X>s\}}$$

$$=\frac{P\{X>s+t\}}{P\{X>s\}}=\frac{1-P\{X\leq s+t\}}{1-P\{X\leq s\}}=\frac{1-F(s+t)}{1-F(s)}$$

$$=\frac{1-(1-e^{-\lambda(s+t)})}{1-(1-e^{-\lambda s})}=\frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}=e^{-\lambda t}=P(X>t)$$

（与s无关）

（三）正态分布(或高斯分

满足连续型随机变量的两个最基本性质

定义 设连续型随机变量

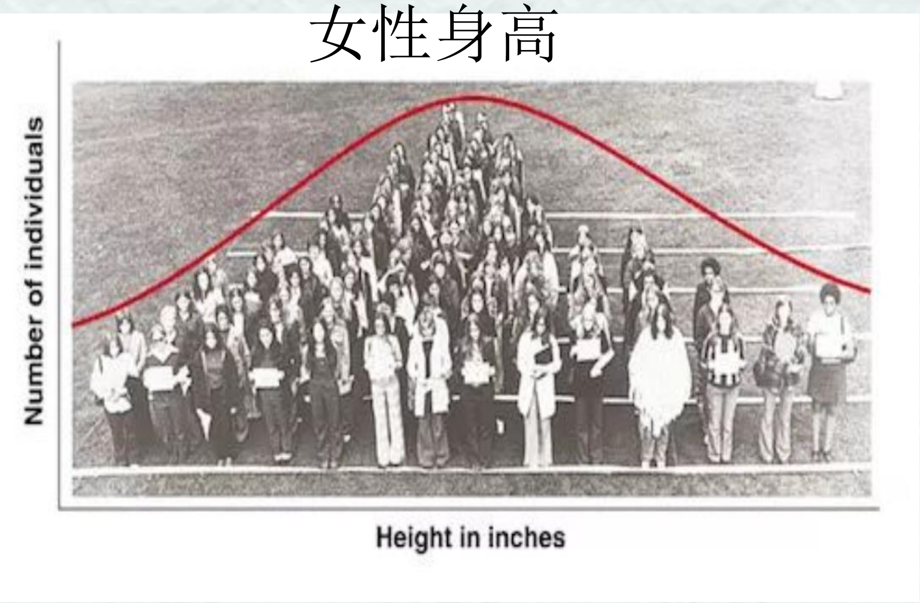
$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x$$

其中 $\mu,\sigma(\sigma>0)$ 为常数,则称 服从参数为 μ,σ 的正态分布或高斯分布,记为 $X\sim N(\mu,\sigma^2)$.

易知 $f(x)\geq 0$, 可以证明$\int_{-\infty}^{+\infty}f(x)dx=1$

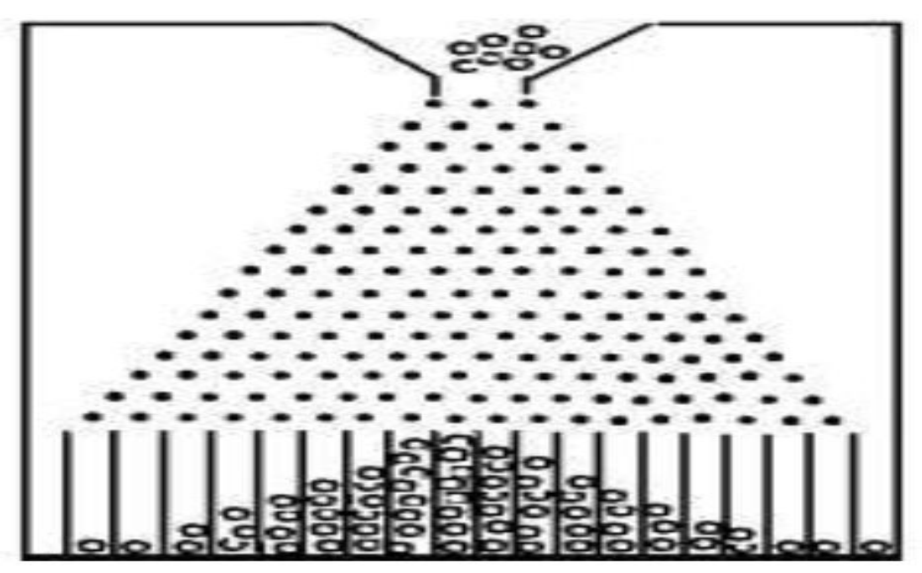
正态分布的应用与背景

正态分布是最常见最重要的一种分布,例如测量误差, 人的生理特征尺寸如身高、体重等；正常情况下生产的产品尺寸:直径、长度、重量高度等都近似服从正态分布.



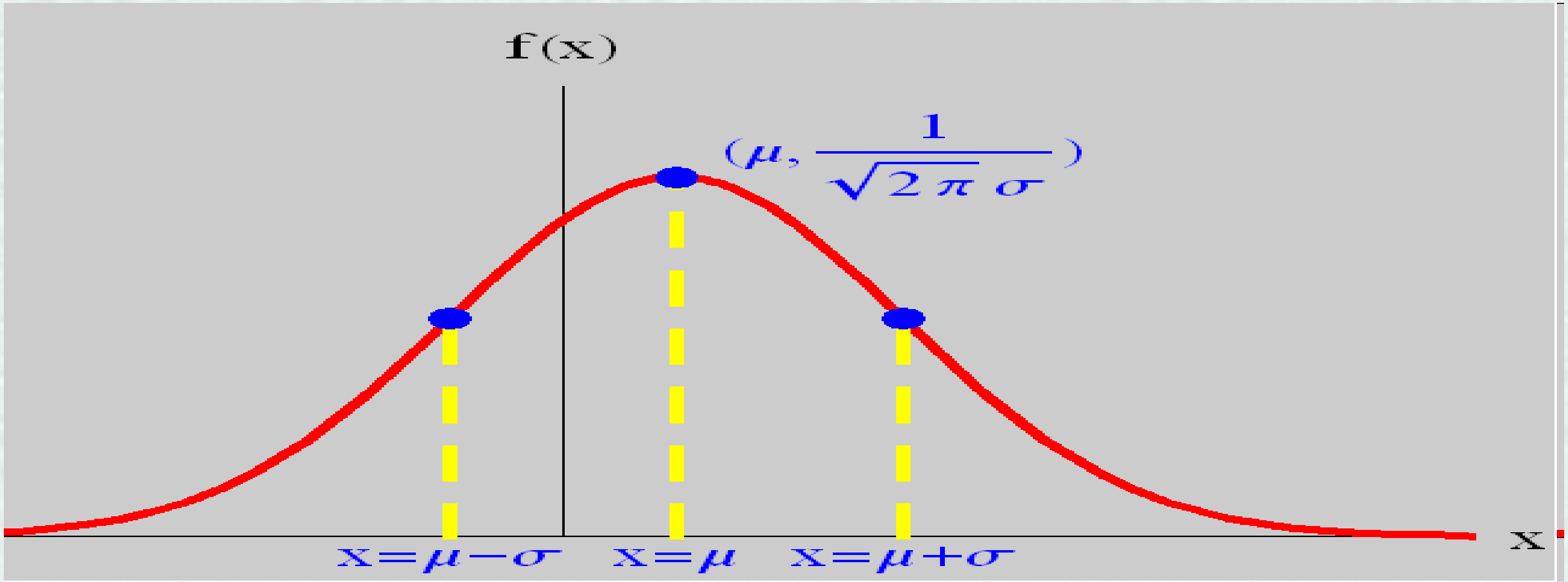
高尔顿钉板

如下图，每一点表示钉在板上的一颗钉子，它们彼此的距离均相等。当小圆球向下降落过程中，碰到钉子后皆以1/2 的概率向左或向右滚下。



如果有 n 排钉子，则小圆球落入各槽内的概率满足二项分布 b(n,1/2), 当n 较大时，接近正态分布。

正态概率密度函数的几何特征



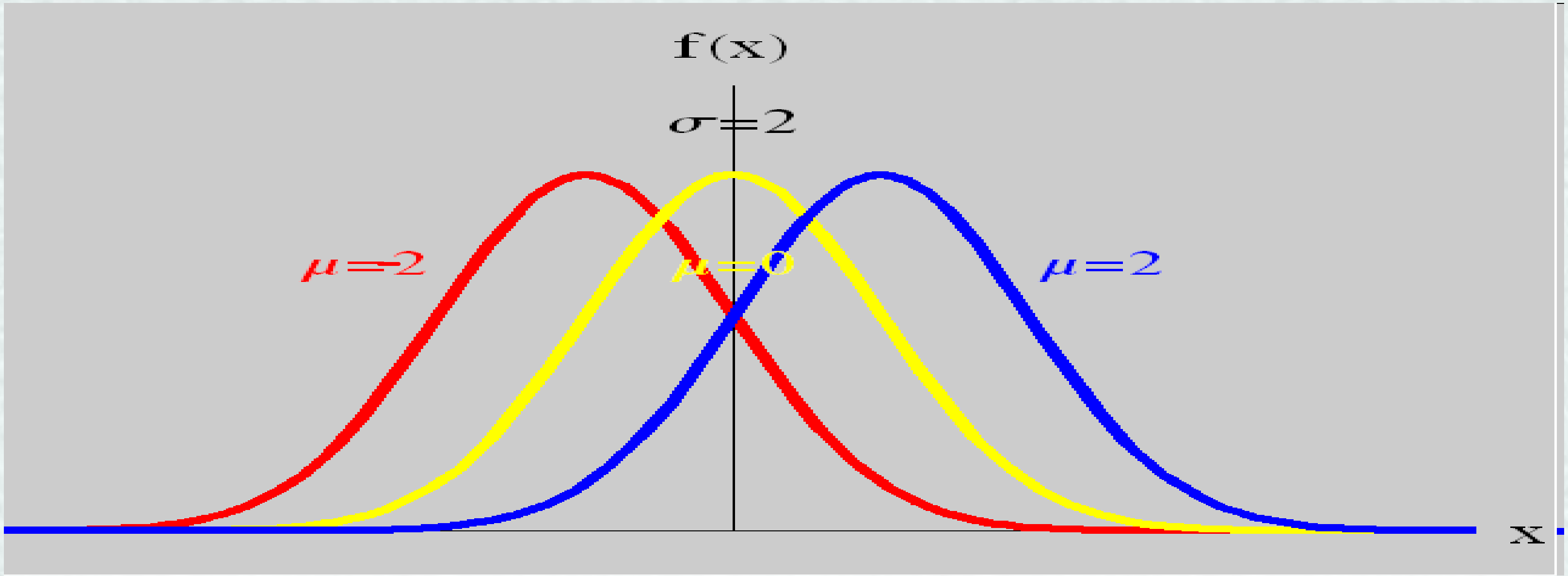
(1) 曲线关于 $x=\mu$ 对称;

(2) 当 $x=\mu$ 时, $f(x)$ 取得最大值 $\dfrac{1}{\sqrt{2\pi}\sigma}$;

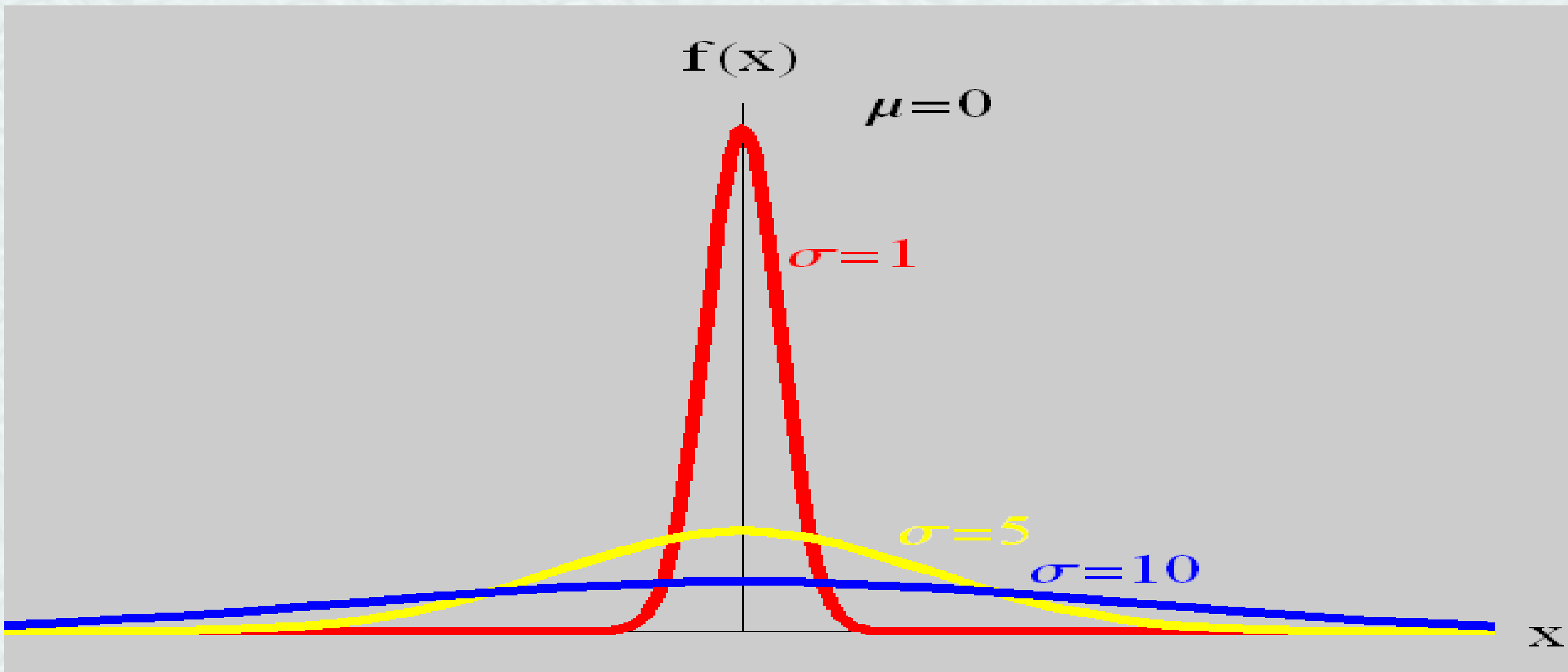
(3) 当 $x \to \pm\infty$ 时, $f(x) \to 0$; 即曲线以x轴为渐近线

(4) 曲线在 $x=\mu\pm\sigma$ 处有拐点;

(5) 当固定 σ, 改变 μ 的大小时,
$f(x)$ 图形的形状不变, 只是沿
着 x 轴作平移变换; 故称μ为位置参数

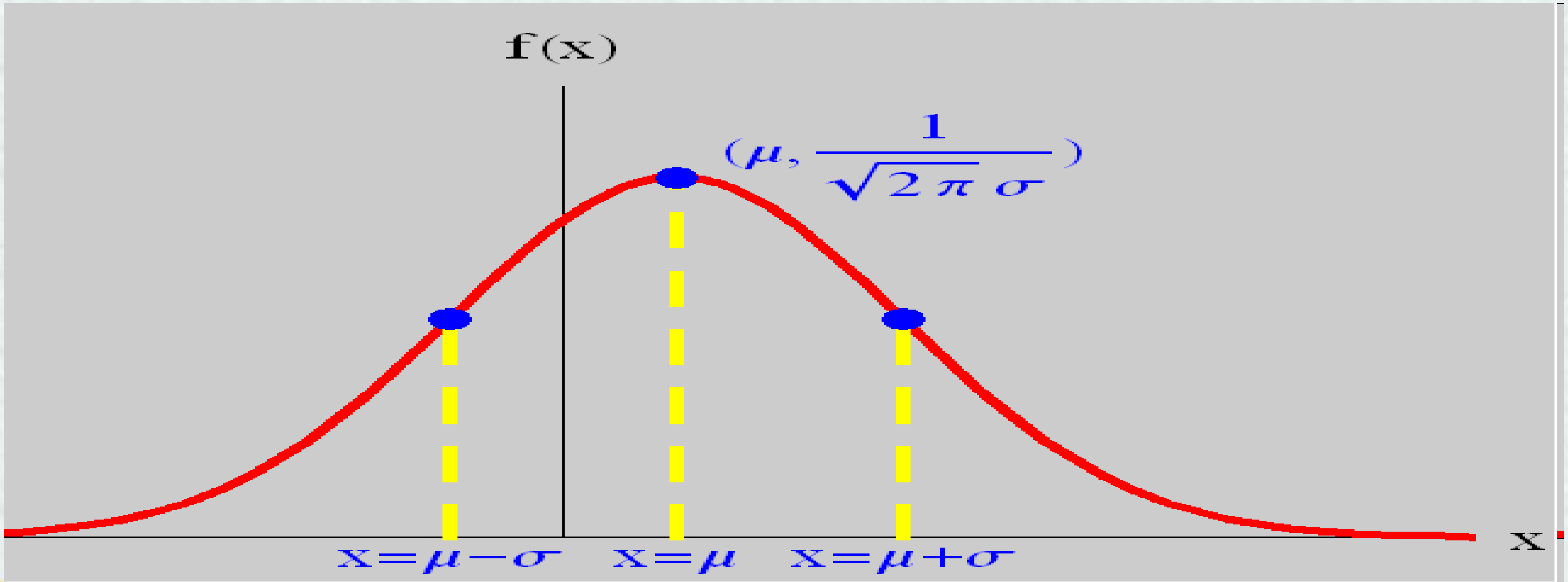


(6) 当固定 μ, 改变 σ 的大小时, $f(x)$ 图形的对称轴不变, 而形状在改变, σ 越小, 图形越高越瘦, σ 越大, 图形越矮越胖.　故称σ为形状参数



正态分布的分布函数 $F(x)=\dfrac{1}{\sqrt{2\pi}\sigma}\displaystyle\int_{-\infty}^{x}\mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}\,\mathrm{d}t$ **＝？**

但 $P(X\le\mu)=F(\mu)=\dfrac{1}{2}$



标准正态分布

当正态分布 $N(\mu,\sigma^2)$ 中的 $\mu=0,\sigma=1$ 时,这样的正态分布称为标准正 态分布,记为 $N(0,1)$.

标准正态分布的概率密度表示为

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

标准正态分布的分布函数表示为

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

正态分布下的概率计算

例3 已知 $X \sim N(0,1)$,求 $P\{1.25 \le X < 2\}$.

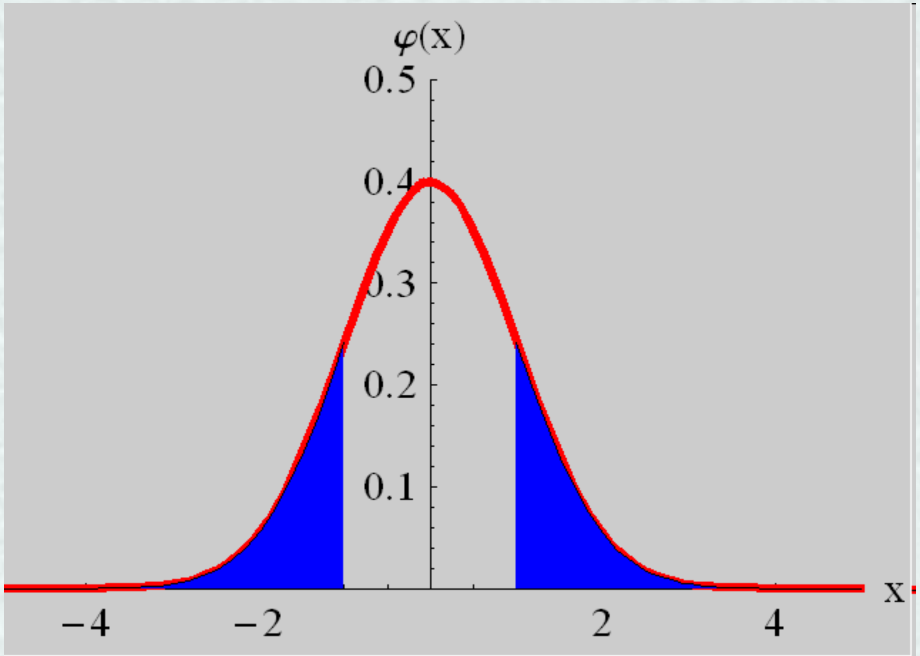
解 $P\{1.25 \le X < 2\}$

$= \Phi(2) - \Phi(1.25)$ 查p215标准正态分布表

$= 0.9773 - 0.8944$

$= 0.0829$.

易知 $\boxed{\Phi(-x)=1-\Phi(x)}$



例4 设$X \sim N(0,1)$，求$P(|X|<1.96)$

解 $P(|X|<1.96) = P(-1.96<X<1.96)$

$$= \Phi(1.96) - \Phi(-1.96)$$

$$= \Phi(1.96) - [1 - \Phi(1.96)]$$

$$= 2\Phi(1.96) - 1$$

$$= 2 \times 0.975 - 1 = 0.95$$

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

$$P\{a < X \le b\} = P\left\{\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right\} = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

例2. 3. 4

已知 $X \sim N(1.5, 4)$，求 $P\{X \le 3.5\}, P\{X > 2.5\}$ 及 $P\{|X| \le 3\}$

解

$$P\{X \le 3.5\} = \Phi\left(\frac{3.5-1.5}{2}\right) = \Phi(1) = 0.8413$$

$$P\{X > 2.5\} = 1 - \Phi\left(\frac{2.5-1.5}{2}\right) = 1 - \Phi(0.5) = 0.3085$$

$$P\{|X| < 3\} = \Phi\left(\frac{3-1.5}{2}\right) - \Phi\left(\frac{-3-1.5}{2}\right) = \Phi(0.75) - \Phi(-2.25)$$

$$= 0.7734 - (1 - 0.9878) = 0.7612$$

例5 设 $X \sim N(\mu, \sigma^2)$ 求X落在区间$(\mu - k\sigma, \mu + k\sigma)$内的概率 $k = (1,2,3,\cdots)$

解 $P\{|X-\mu| < k\sigma\} = P\{\mu - k\sigma < X < \mu + k\sigma\}$

$$= P\{-k < \frac{X-\mu}{\sigma} < k\} = \Phi(k) - \Phi(-k) = 2\Phi(k) - 1$$

于是 $P\{|X-\mu| < \sigma\} = 2\Phi(1) - 1 = 0.6826$

$P\{|X-\mu| < 2\sigma\} = 2\Phi(2) - 1 = 0.9546$

$P\{|X-\mu| < 3\sigma\} = 2\Phi(3) - 1 = 0.9973$

则有$P\{|X-\mu| \geq 3\sigma\} = 1 - P\{|X-\mu| < 3\sigma\} = 0.0027 < 0.003$

$\boldsymbol{X}$落在$(\boldsymbol{\mu} - 3\boldsymbol{\sigma}, \boldsymbol{\mu} + 3\boldsymbol{\sigma})$以外的概率小于0.003,

在实际问题中常认为它不会发生.

例2.3.5 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 $d^{\circ}C$, 液体的温度 X (以$^{\circ}C$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$80^{\circ}C$ 的概率不低于 0.99, 问d至少为多少?

解 (1) 所求概率为

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

$$= 1 - \Phi(2) = 1 - 0.9773 = 0.0227.$$

(2) $P\{X \geq 80\} \geq 0.99$

$$\Rightarrow 1 - P\{X < 80\} \geq 0.99$$

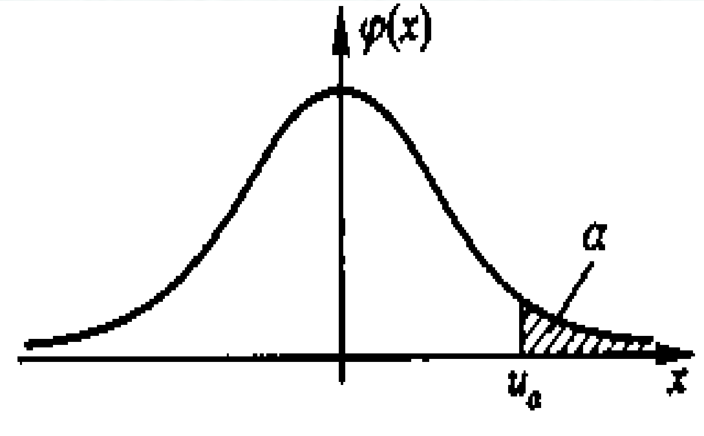
$$\Rightarrow 1 - \Phi\left(\frac{80 - d}{0.5}\right) \geq 0.99$$

$$\Rightarrow \Phi\left(\frac{d - 80}{0.5}\right) \geq 0.99,$$

$$\Rightarrow \frac{d - 80}{0.5} \geq 2.33 \qquad \Rightarrow d \geq 81.165.$$

定义2.3.5

设$X \sim N(0,1)$,若u_α满足条件$P\{X > u_\alpha\} = \alpha, 0 < \alpha < 1$,则称点$u_\alpha$为标准正态分布的上α分位点(如下图所示)



由$\varphi(x)$的图形的对称性可知：$u_{1-\alpha} = -u_\alpha$

§ 2.4　随机变量的函数的分布

设 $g(x)$ 是定义在随机变量 X 的一切可能值 x 的集合上的函数，若随机变量 Y 随着 X 取值 x 而取 $y=g(x)$ 的值，则称随机变量 Y 为随机变量 X 的函数，记作 $Y=g(X)$.

问题

如何根据已知的随机变量 X 的分布求得随机变量 $Y=g(X)$ 的分布？

一、一维离散型随机变量的函数的分布

例1 设随机变量X具有以下分布(如下表)，

试求(1)$Y=2X$;　(2)$Z=2X^2+1$的分布律。

X	-1	0	1	2
P	0.2	0.1	0.3	0.4

解　(1)Y的所有可能取值为$-2,0,2,4$。

由$P\{Y=2k\}=P\{X=k\}=p_k$得Y的分布律为

Y	-2	0	2	4
P	0.2	0.3	0.1	0.4

(2)Z的所有可能取值1, 3, 9

$$P\{Z=1\}=P\{2X^2+1=1\}=P\{X=0\}=0.1$$

$$P\{Z=3\}=P\{2X^2+1=3\}=P\{X=1\}+P\{X=-1\}=0.5$$

$$P\{Z=9\}=P\{2X^2+1=9\}=P\{X=2\}=0.4$$

故z的分布律为

Z	1	3	9
P	0.1	0.5	0.4

离散型随机变量的函数的分布

如果 X 是离散型随机变量,其函数 $Y=g(X)$ 也是离散型随机变量.若 X 的分布律为

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

则 $Y=g(X)$的分布律为

$Y=g(X)$	$g(x_1)$	$g(x_2)$	$\cdots$	$g(x_k)$	$\cdots$
P	p_1	p_2	$\cdots$	p_k	$\cdots$

若 $g(x_k)$ 中有值相同的,应将相应的 p_k 合并.

练习 设

X	-1	1	2
P	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$

求 $Y = X^2 - 5$ 的分布律.

解 Y 的分布律为

Y	-4	-1
P	$\frac{1}{2}$	$\frac{1}{2}$

二、一维连续型随机变量的函数的分布

例2.4.2 设随机变量 X 的概率密度为

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

求随机变量 $Y=2X+1$ 的概率密度.

解 第一步 先求Y=2X+1 的分布函数 $F_Y(y)$.

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

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

$$= P\{X \le \frac{y-1}{2}\} = F_X\left(\frac{y-1}{2}\right)$$

第二步　由分布函数求概率密度.

$$f_Y(y) = F_Y'(y) = \left(F_X\left(\frac{y-1}{2}\right)\right)' = f_X\left(\frac{y-1}{2}\right)\left(\frac{y-1}{2}\right)'$$

$$= f_X\left(\frac{y-1}{2}\right)\cdot\frac{1}{2} = \begin{cases} \frac{1}{2}\left(\frac{y-1}{2}\right)\cdot\frac{1}{2}, & 0 < \frac{y-1}{2} < 2, \\ 0, & \text{其他}. \end{cases}$$

所以 $$f_Y(y)=\begin{cases}\dfrac{1}{2}\left(\dfrac{y-1}{2}\right)\cdot\dfrac{1}{2}, & 0<\dfrac{y-1}{2}<2,\\ 0, & \text{其他}.\end{cases}$$

$$=\begin{cases}\dfrac{y-1}{8}, & 1<y<5,\\ 0, & \text{其他}.\end{cases}$$

例2.4.3 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$F_X(x), F_Y(y)$

先求Y的分布函数$F_Y(y) = P(Y \le y) = P(X^2 \le y)$

$$= \begin{cases} P(\varnothing) = 0, & y < 0, \\ P\left(X = 0\right) = 0, & y = 0 \\ P\left(-\sqrt{y} \le X \le \sqrt{y}\right) = F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right), & y > 0. \end{cases}$$

将$F_Y(y)$关于y求导,即得y的概率密度为

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

书P59例2.4.3, 设$X\sim N(0,1)$,其概率密度为

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

由（1）得$Y=X^2$的概率密度为

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

$Y\sim\chi^2(1).$

此时称Y服从自由度为1的χ^2分布.

例2.4.4 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.

证明 设$a > 0$, 下面先求$F_Y(y)$

$$F_Y(y) = P\{Y \leq y\} = P\{aX + b \leq y\}$$

$$= P\{X \leq \frac{y-b}{a}\} = F_X(\frac{y-b}{a})$$

将$F_Y(y)$关于y求导，得$Y=aX+b$的概率密度为

$$f_Y(y) = f_X(\frac{y-b}{a})(\frac{y-b}{a})' = \frac{1}{a} f_X(\frac{y-b}{a})$$

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

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

若$a<0$,以同样的方法可以求得

$$F_Y(y)=P\{Y\le y\}=P\{aX+b\le y\}$$

$$=P\{X\ge\frac{y-b}{a}\}=1-F_X(\frac{y-b}{a})$$

$$f_Y(y)=-f_X(\frac{y-b}{a})(\frac{y-b}{a})'=-\frac{1}{a}f_X(\frac{y-b}{a})$$

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

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

特别地,在上例中取 $a=\dfrac{1}{\sigma},b=-\dfrac{\mu}{\sigma}$ 得 $Y=\dfrac{X-\mu}{\sigma}\sim N(0,1)$

这就是上一节引理的结果

定理2.4.1 设随机变量 X 的具有概率密度 $f_X(x)$, 其中 $-\infty < x < +\infty$, 又设函数 $g(x)$ 处处可导, 且恒有 $g'(x) > 0$(或恒有 $g'(x) < 0$), 则称 $Y = g(X)$ 是连续型随机变量, 其概率密度为

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

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

例2.4.4 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.

证明 X 的概率密度为

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

设 $y = g(x) = ax + b$,

得 $x = h(y) = \dfrac{y-b}{a}$, 知 $h'(y) = \dfrac{1}{a} \neq 0.$

由公式 $f_Y(y)=\begin{cases} f_X[h(y)]\left|h'(y)\right|, \alpha<y<\beta, \\ 0, \qquad\qquad 其它. \end{cases}$

得 $Y=aX+b$ 的概率密度为

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

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

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

得 $Y=aX+b$

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

连续型随机变量的函数的分布

方法1 $F_Y(y)=P\{Y\le y\}=P\{g(X)\le y\}$

$=P\{X$的取值范围$\}=$用F_X(或f_X)表示

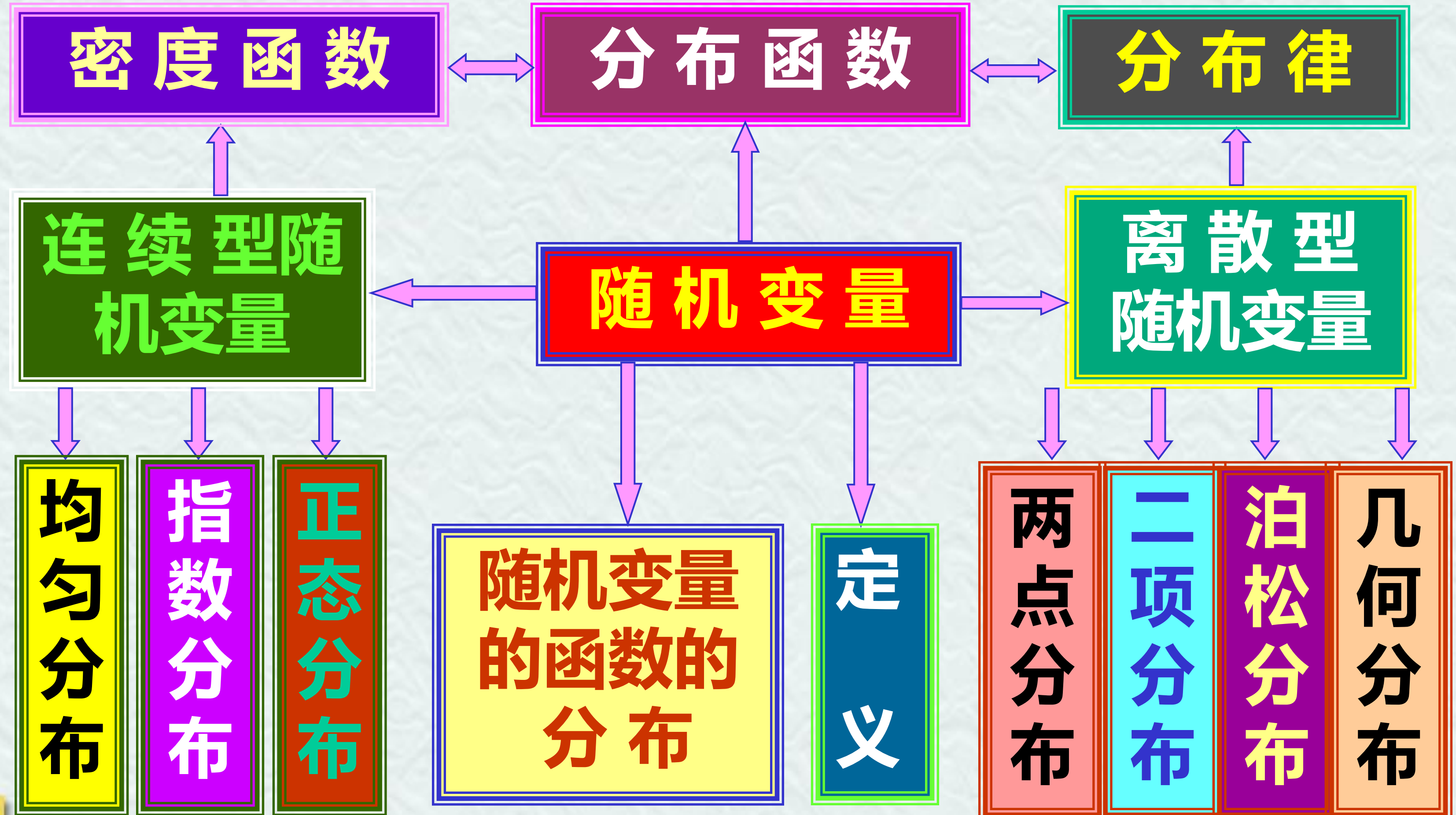
再对 $F_Y(y)$ 求导得到Y 的密度函数 $f_Y(y)$.

注意y的取值范围的确定。

方法2 $$f_Y(y)=\begin{cases} f_X[h(y)]\,|h'(y)|, & \alpha<y<\beta, \\ 0, & \text{其他}. \end{cases}$$

注意“单调”这一条件.

复习：一维r.v.及其分布的主要内容



重点与难点

1.重点

离散型r.v的分布律 ⟷ 分布函数

连续性r.v.的概率密度函数

根据分布求r.v.落在某个区间的概率

常用的离散型和连续型分布

随机变量函数的分布

2.难点

连续型随机变量（的函数）的分布计算

典型例题

例1 已知离散型随机变量 X 的可能取值为 $-2, 0, 2, \sqrt{5}$, 相应的概率依次为 $\dfrac{1}{a}, \dfrac{3}{2a}, \dfrac{5}{4a}, \dfrac{7}{8a}$, 试求概率 $P\{|X| \le 2 \mid X \ge 0\}$.

[思路] 首先根据概率分布的性质求出常数 a 的值, 然后确定概率分布律的具体形式,最后再计算条件概率.

解 利用概率分布律的性质 $\sum_i p_i = 1$,

有 $1=\sum_i p_i=\frac{1}{a}+\frac{3}{2a}+\frac{5}{4a}+\frac{7}{8a}=\frac{37}{8a},$

故 $a=\frac{37}{8},$

因此 X 的分布律为

X	-2	0	2	$\sqrt{5}$
P	$\frac{8}{37}$	$\frac{12}{37}$	$\frac{10}{37}$	$\frac{7}{37}$

从而

$$P\{|X| \le 2 | X \ge 0\} = \frac{P\{|X| \le 2, X \ge 0\}}{P\{X \ge 0\}}$$

$$= \frac{P\{X = 0\} + P\{X = 2\}}{P\{X = 0\} + P\{X = 2\} + P\{X = \sqrt{5}\}}$$

$$= \frac{22}{29}.$$

例2 袋中有球10个，7红3黑，逐个随机抽取，若取出黑球，则放入红球取代，依次进行，直到取到红球为止。求（1）抽取次数N的分布律；

（2）N的分布函数。

解 N的分布律为

N	1	2	3	4
P	0.7	0.24	0.054	0.006

例3 设离散型随机变量 X 的分布函数为

$$F(x)=\begin{cases}0, & x<-1,\\ a, & -1\le x<1,\\ \dfrac{2}{3}-a, & 1\le x<2,\\ a+b, & x\ge 2.\end{cases}$$

且 $P\{X=2\}=\dfrac{1}{2}$，试确定常数 a,b，并求 X 的分布律.

[思路] 首先利用分布函数的性质求出常数 a, b，再用已确定的分布函数来求分布律.

解 利用分布函数 $F(x)$ 的性质：

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

$$F(+\infty) = 1,$$

知 $\dfrac{1}{2} = P\{X = 2\}$

$$= (a+b) - (\frac{2}{3} - a)$$

$$= 2a + b - \frac{2}{3},$$

且 $a + b = 1.$

由此解得 $a = \dfrac{1}{6}, b = \dfrac{5}{6}.$

因此有 $$F(x)=\begin{cases}0, & x<-1,\\ \dfrac{1}{6}, & -1\le x<1,\\ \dfrac{1}{2}, & 1\le x<2,\\ 1, & x\ge 2.\end{cases}$$

从而 X 的分布律为

X	-1	1	2
P	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$

例4 已知随机变量 X 的概率密度为

$$f(x)=Ae^{-|x|},\quad -\infty<x<+\infty.$$

(1) 求系数 A;

(2) 求 X 的分布函数 $F(x)$;

(3) 求 $Y=X^2$ 的概率密度.

(4) 求$P(-1<X<1)$.

解 (1)由概率密度的性质,有

$$1=\int_{-\infty}^{+\infty}f(x)\,\mathrm{d}x=\int_{-\infty}^{+\infty}Ae^{-|x|}\,\mathrm{d}x=2\int_{0}^{+\infty}Ae^{-x}\,\mathrm{d}x$$

$$=2A,\qquad \text{故}\quad A=\frac{1}{2}.$$

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

当 $x<0$ 时,有 $F(x)=\frac{1}{2}\int_{-\infty}^{x}\mathrm{e}^{t}\,\mathrm{d}t=\frac{1}{2}\mathrm{e}^{x};$

当 $x\geq 0$ 时,有 $F(x)=\frac{1}{2}[\int_{-\infty}^{0}\mathrm{e}^{t}\,\mathrm{d}t+\int_{0}^{x}\mathrm{e}^{-t}\,\mathrm{d}t]=1-\frac{1}{2}\mathrm{e}^{-x};$

所以 X 的分布函数为

$$F(x)=\begin{cases}\frac{1}{2}\mathrm{e}^{x}, & x<0,\\ 1-\frac{1}{2}\mathrm{e}^{-x}, & x\geq 0.\end{cases}$$

(3) 由于 $Y = X^2 \geq 0$,

故当 $y \leq 0$ 时,有 $F_Y(y) = P\{Y \leq y\} = 0;$ $f_Y(y) = 0$

当 $y > 0$ 时,有

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

$$= P\{-\sqrt{y} \leq X \leq \sqrt{y}\} = F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right)$$

$$f_Y(y) = F_Y'(y) = f_X\left(\sqrt{y}\right) \cdot \left(\sqrt{y}\right)' - f_X\left(-\sqrt{y}\right) \cdot \left(-\sqrt{y}\right)'$$

$$= \mathrm{e}^{-\sqrt{y}} \cdot \frac{1}{2\sqrt{y}},$$

从而, Y 的概率密度为

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

(4)

$$P(-1<X<1)=F(1)-F(-1)=(1-\frac{1}{2}e^{-1})-\frac{1}{2}e^{-1}=1-e^{-1}$$

或

$$P(-1<X<1)=\int_{-1}^{0}\frac{1}{2}e^{x}dx+\int_{0}^{1}\frac{1}{2}e^{-x}dx=1-e^{-1}$$

例5 设连续型随机变量 X 的分布函数为

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

求：(1) 系数 A, B 的值;

(2) $P\{-a<X<\dfrac{a}{2}\}$;

(3) 随机变量 X 的概率密度.

解 (1) 因为 X 是连续型随机变量, 所以 $F(x)$ 连续,

故有 $F(-a)=\lim\limits_{x\to -a}F(x)$,

$$F(a)=\lim_{x\to a}F(x),$$

即 $$A+B\arcsin\left(\frac{-a}{a}\right)=A-\frac{\pi}{2}B=0,$$

$$A+B\arcsin\left(\frac{a}{a}\right)=A+\frac{\pi}{2}B=1,$$

解之得 $A = \dfrac{1}{2}, \quad B = \dfrac{1}{\pi}.$

所以 $$F(x) = \begin{cases} 0, & x \le -a, \\ \dfrac{1}{2} + \dfrac{1}{\pi}\arcsin\dfrac{x}{a}, & -a < x \le a, \\ 1, & x > a. \end{cases}$$

(2) $P\{-a < X < \frac{a}{2}\} = F(\frac{a}{2}) - F(-a)$

$$= \frac{1}{2} + \frac{1}{\pi}\arcsin(\frac{a}{2a}) - 0$$

$$= \frac{1}{2} + \frac{1}{\pi} \times \frac{\pi}{6} = \frac{2}{3}.$$

(3) 随机变量 X 的概率密度为

$$f(x) = F'(x) = \begin{cases} 1/\pi\sqrt{a^2 - x^2}, & -a < x < a, \\ 0, & \text{其它}. \end{cases}$$

例6 设某城市成年男子的身高 $X \sim N(170, 6^2)$ (单位 : cm)

(1) 问应如何设计公共汽车车门的高度,使男子与车门顶碰头的几率小于 0.01?

(2) 若车门高为 182 cm,求 100 个成年男子与车门顶碰头的人数不多于 2 的概率.

[思路] 设车门高度为 l cm, 那么按设计要求应有 $P\{X > l\} < 0.01$,确定 l. 第二问首先要求出 100名男子中身高超过 182cm 的人数的分布律,然后用此分布律,求其不超过2的概率.

解　(1) 由题设知 $X \sim N(170, 6^2)$,

$$P\{X > l\} = 1 - P\{X \le l\}$$

$$= 1 - P\left\{\frac{X - 170}{6} \le \frac{l - 170}{6}\right\}$$

$$= 1 - \Phi\left(\frac{l - 170}{6}\right) < 0.01,$$

即 $\Phi\left(\frac{l - 170}{6}\right) > 0.99.$　查表得 $\frac{l - 170}{6} > 2.33,$

故　$l > 183.98(\text{cm}).$

(2) 设任一男子身高超过 182cm 的概率为 p.

则 $$p = P\{X > 182\} = P\left\{\frac{X-170}{6} > \frac{182-170}{6}\right\}$$

$$= 1 - \Phi(2) = 0.0228.$$

设 Y 为 100 个男子中身高超过 182cm 的人数,

则 $Y \sim B(100, 0.0228)$, 其中

$$P\{Y = k\} = \binom{100}{k} \times 0.0228^{k} \times 0.9772^{100-k},$$

$$k = 0,1,\cdots,100.$$

所求概率为

$$P\{Y \le 2\} = P\{Y = 0\} + P\{Y = 1\} + P\{Y = 2\},$$

由于 $n = 100$ 较大, $p = 0.0228$ 较小,故可用泊松分布来计算,其中 $\lambda = np = 2.28$,

从而

$$P\{Y \le 2\} \approx \frac{2.28^0 e^{-2.28}}{0!} + \frac{2.28 e^{-2.28}}{1!} + \frac{2.28^2 e^{-2.28}}{2!}$$

$$= 0.6013.$$