2.4 连续型随机变量及其概率密度

连续型随机变量的定义

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

容易看到，本例的随机变量 X 的取值充满区间$[0, a]$，其分布函数

$F(x)$对于任意的x可以表示成 $F(x)=\int_{-\infty}^{x} f(t)\,\mathrm{d}t,$

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

综上，随机变量X的分布函数为

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

这就是说，$F(x)$恰好是非负可积函数 $f(x)$在区间$(-\infty，x]$上的积分，对于这类随机变量有以下的一般定义。

定义 设 X 是随机变量，其分布函数为 $F(x)$，如果存在**非负可积函数** $f(x)$，使得

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

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

连续型随机变量的性质

性质1 $f(x)\geq 0,\quad -\infty<x<+\infty$

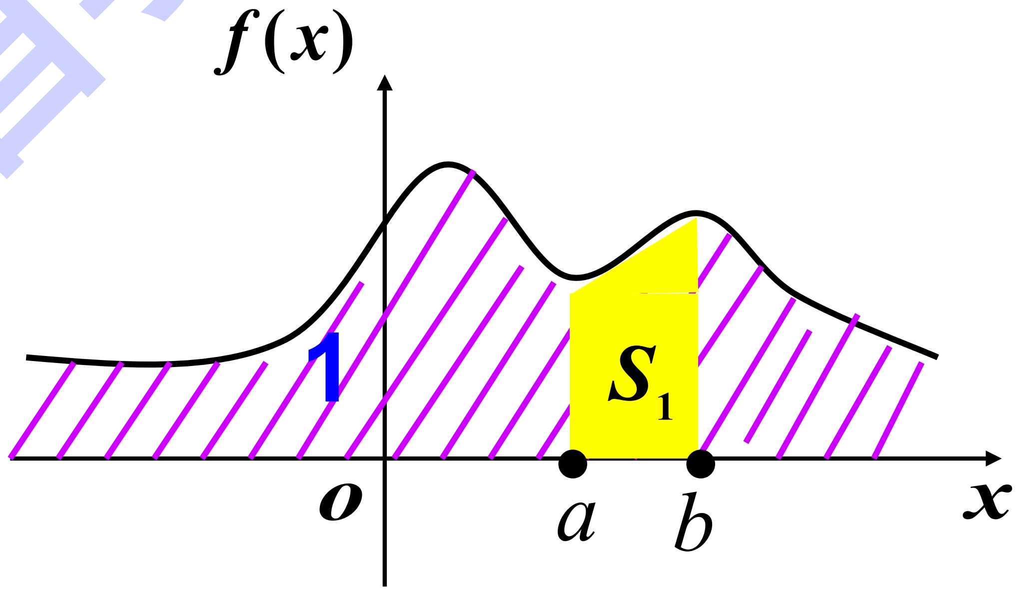
由定义1直接得证

性质2 $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x=1.$

证明 $1=F(\infty)=\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x.$

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

$$S_1=\int_{a}^{b} f(x)\,\mathrm{d}x$$



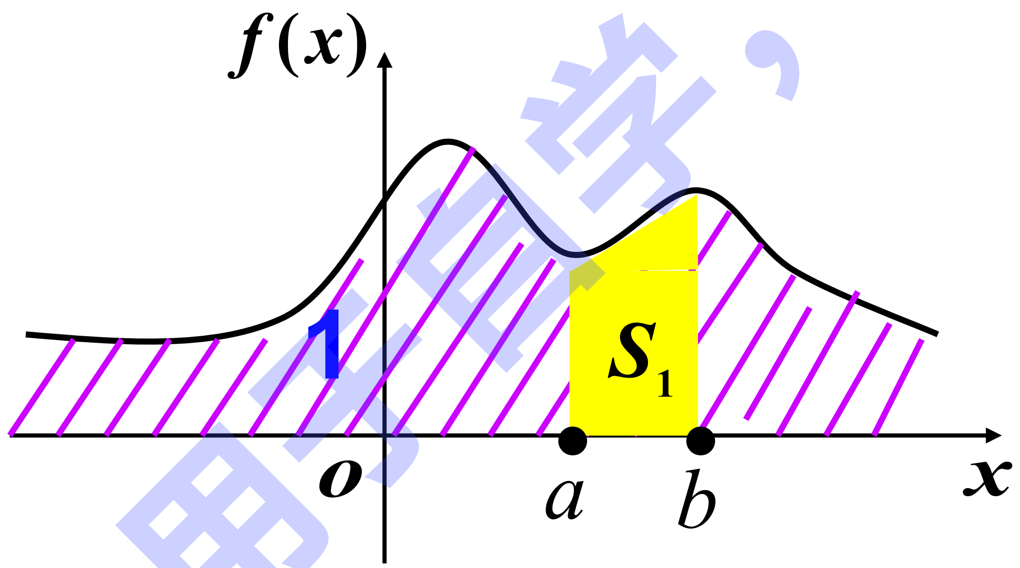
性质3 $P\{a<X\le b\}==\int_a^b f(x)\,\mathrm{d}x$

证明 $P\{a<X\le b\}=P(x\le b)-P(x\le a)$

$$=\int_{-\infty}^{b} f(x)\,\mathrm{d}x-\int_{-\infty}^{a} f(x)\,\mathrm{d}x=\int_a^b f(x)\,\mathrm{d}x.$$

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

$$S_1=\int_a^b f(x)\,\mathrm{d}x$$



同时得以下计算公式

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

$$P\{X>a\}=1-P\{X\le a\}=1-F(a)$$

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

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

$$=\int_{a}^{\infty} f(x)\,\mathrm{d}x.$$

性质4 若 $f(x)$ 在点 x 处连续,则有 $F'(x)=f(x)$.

证明 在 $f(x)$ 的连续点 x 处,由于 $F(x)=\int_{-\infty}^{x} f(t)dt$,

根据高等数学中所学习的变上限积分求导法则可得 $F'(x)=f(x)$.

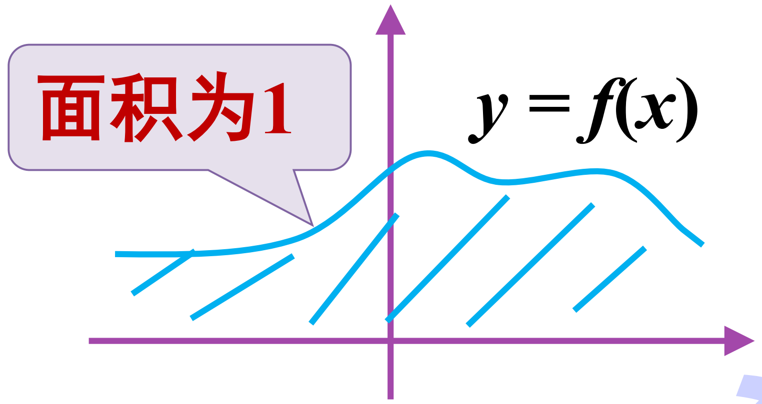
需要指出的是,

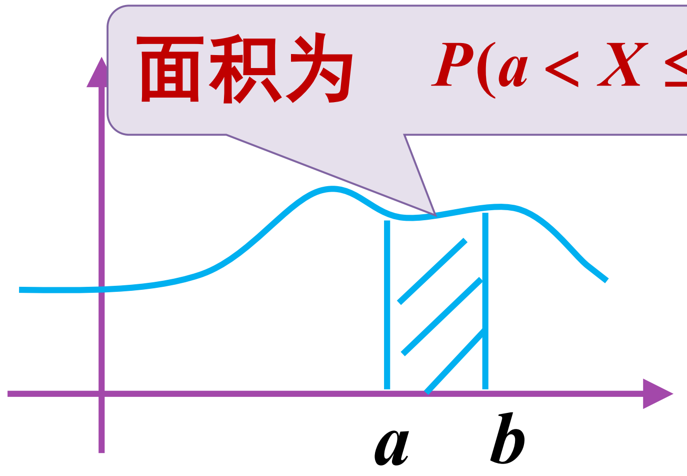
如果函数 $f(x)$满足上述性质 1 与性质 2，那么 $f(x)$一定是某个连续型随机变量的概率密度

定理 1 设连续型随机变量 X 的概率密度为 $f(x)$，分布函数为 $F(x)$，则 $F(x)$在$(-\infty, +\infty)$上连续。

证 由微积分学基本定理即可得证。

概率意义





概率密度与质量、电量的线密度定义类似 $P(a<X\le a+\Delta x)\approx f(x)\,\Delta x$

X落在小区间$(a, a+\Delta x]$的概率近似等于$f(x)\,\Delta x$

X落在一段区间的概率可知，那么取某一实数值的概率是？即求$P(X=a)$

定理 2 设 X 为连续型随机变量，则随机变量 X 取任一实数值 a 的概率均为零，即$P(X=a)=0$

证 设 X 的分布函数为 $F(x)$， $\Delta x>0$， 则由$\{X=a\}\subset\{a-\Delta x<X\le a\}$得

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

在上述不等式中令 $\Delta x\to 0$， 由定理 1 得

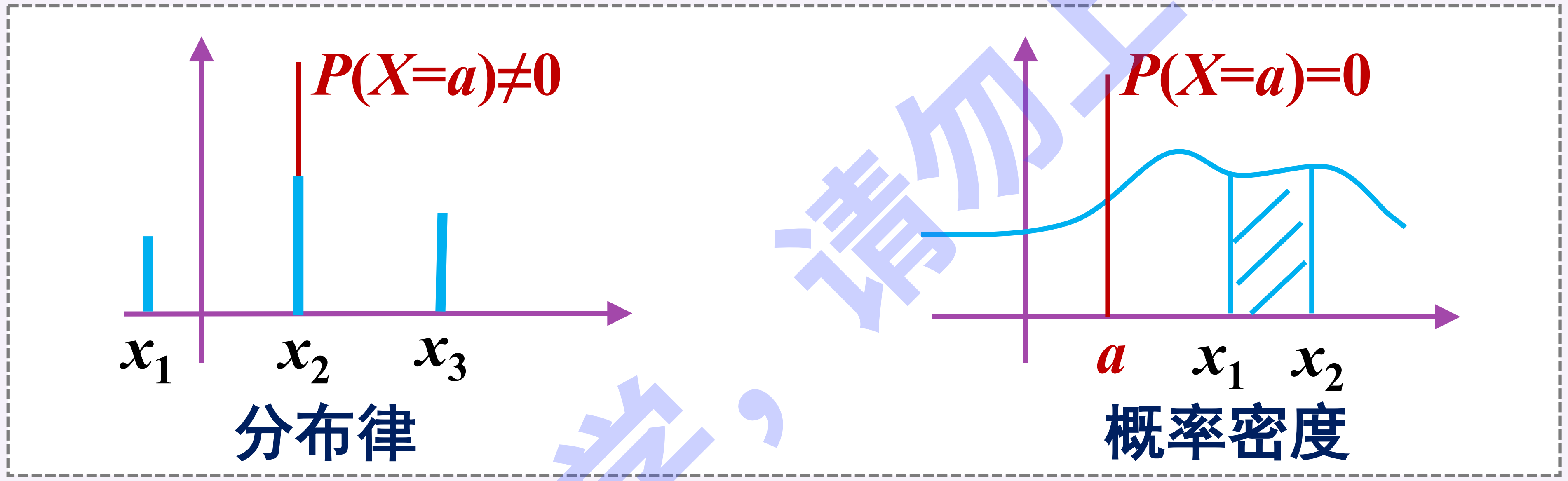
$$P\{X=a\}=0.$$

由定理2知，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，即连续积分可忽略端点处.从而

$$P\{a<X\le b\}=P\{a\le X<b\}=P\{a<X<b\}=P\{a\le X\le b\}=\int_a^b f(x)dx$$

注意

1° 连续型随机变量必有处处$P(X=a)=0$；
离散型随机变量可以有$P(X=a)\neq 0$



若 A 为离散型随机变量，

A是不可能事件$\leftrightarrow P(A)=0$

2° 概率为0不一定是不可能事件，例如连续型随机变量$P(X=a)=0$；但是，
若A为不可能事件，必有$P(A)=0$

3° 由2°知，概率为1不一定是必然事件，例如连续型随机变量$P(X\neq a)=1$；
但是，若A为必然事件，必有$P(A)=1$

例 设连续型随机变量 X 的概率密度为 $f(x)=\dfrac{C}{1+x^2},-\infty<x<+\infty$

试求：(1)常数 C；(2) X 的分布函数 $F(x)$；(3) $P(X>1)$。

解 (1)由 $\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x=1$ 得，$1=\int_{-\infty}^{+\infty}\dfrac{C}{1+x^2}\mathrm{d}x=\pi C$，故 $C=\dfrac{1}{\pi}$

$$(2)F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t=\int_{-\infty}^{x}\frac{1}{\pi(1+t^2)}\mathrm{d}t=\left[\frac{1}{\pi}\arctan t\right]_{-\infty}^{x}=\frac{1}{\pi}\arctan x+\frac{1}{2}$$

$$(3)P(X>1)=\int_{1}^{+\infty}\frac{1}{\pi(1+x^2)}\mathrm{d}x=\left[\frac{1}{\pi}\arctan x\right]_{1}^{+\infty}=\frac{1}{\pi}\left(\frac{\pi}{2}-\frac{\pi}{4}\right)=\frac{1}{4}$$

例 设X的概率密度为 $f(x)=\begin{cases} c & 0<x<1 \\ 2/9 & 3<x<6 \\ 0 & \text{其他} \end{cases}$

(i) 求常数 c 的值； (ii) 写出X的分布函数；(iii) 要使 $P(X<k)=2/3$，求 k 的值。

解 (i) $1=\int_{-\infty}^{+\infty} f(t)dt=c\int_0^1 dt+\frac{2}{9}\int_3^6 dt=\frac{2}{3}+c\Rightarrow c=\frac{1}{3}$

(ii) $$F(x)=P(X\le x)=\begin{cases} 0 & x\le 0 \\ \int_0^x \frac{1}{3}dt & 0<x<1 \\ \int_0^1 \frac{1}{3}dt & 1\le x\le 3 \\ \int_0^1 \frac{1}{3}dt+\int_3^x \frac{2}{9}dt & 3<x<6 \\ 1 & x\ge 6 \end{cases}=\begin{cases} 0 & x\le 0 \\ x/3 & 0<x<1 \\ 1/3 & 1\le x\le 3 \\ (2x-3)/9 & 3<x<6 \\ 1 & x\ge 6 \end{cases}$$

(iii) $k=4.5$

例 **设连续型随机变量 X 的概率密度为** $f(x)=\begin{cases} cx(1-x), & 0<x<1, \\ 0, & 其它. \end{cases}$

求：(1)常数 c; (2) $P\left(|X|\le\frac{1}{2}\right)$、$P\left(X=\frac{1}{2}\right)$ 和 $P\left(X\ge\frac{1}{3}\right)$； 3) X的分布函数$F(x)$。

解 (1) 由$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$ 得 $1=\int_0^1 cx(1-x)\mathrm{d}x=c\left(\frac{1}{2}-\frac{1}{3}\right)=\frac{c}{6}$ 解得$c=6$

(2) 由于$c=6$，因此X的概率密度为 $f(x)=\begin{cases} 6x(1-x), & 0<x<1, \\ 0, & 其它. \end{cases}$

又由于X是连续型随机变量

$$P\left(|X|\le\frac{1}{2}\right)=P\left(-\frac{1}{2}\le X\le\frac{1}{2}\right)=P\left(0\le X\le\frac{1}{2}\right)=\int_0^{\frac{1}{2}}6x(1-x)\mathrm{d}x=\frac{1}{2}$$

$$P\left(X=\frac{1}{2}\right)=0 \qquad P\left(X\ge\frac{1}{3}\right)=\int_{\frac{1}{3}}^{1}6x(1-x)\mathrm{d}x=\frac{20}{27}$$

例 **设连续型随机变量 X 的概率密度为** $f(x)=\begin{cases} cx(1-x), & 0<x<1, \\ 0, & \text{其它}. \end{cases}$

求：(1)常数 c; (2) $P\left(|X|\le\frac{1}{2}\right)$、$P\left(X=\frac{1}{2}\right)$ 和 $P\left(X\ge\frac{1}{3}\right)$； 3) X的分布函数$F(x)$。

解 (3) 因为$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t, x\in\mathbb{R}$,则当$x<0$时，$F(x)=0$

当$0\le x<1$时，$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t=\int_{0}^{x}6t(1-t)\mathrm{d}t=x^2(3-2x)$

当$x\ge1$时，$F(x)=1$

即X的分布函数为 $$F(x)=\begin{cases} 0, & x<0 \\ x^2(3-2x), & 0\le x<1 \\ 1, & x\ge1 \end{cases}$$

例 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

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

即 $$A+B\arcsin\left(\frac{-a}{a}\right)=A-\frac{\pi}{2}B=0,\qquad A+B\arcsin\left(\frac{a}{a}\right)=A+\frac{\pi}{2}B=1,$$

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

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

$$=\frac{1}{2}+\frac{1}{\pi}\arcsin(\frac{a}{2a})-0 \quad =\frac{1}{2}+\frac{1}{\pi}\times\frac{\pi}{6} \quad =\frac{2}{3}.$$

(3) 随机变量 X 的概率密度为

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

三种重要的连续型随机变量

1. 均匀分布

X具有概率密度 $f(x)=\begin{cases}\dfrac{1}{b-a} & x\in(a,b)\\ 0 & \text{其他}\end{cases}$

称X在区间(a, b)上服从均匀分布，

记为$X\sim U(a, b)$

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

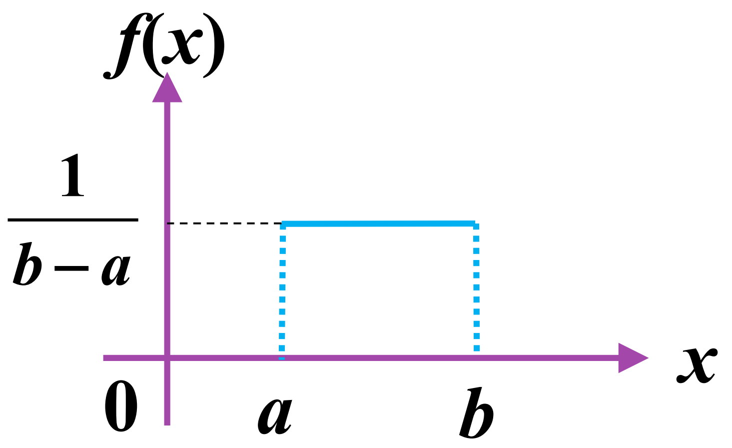
设 $a\leq c<c+l\leq b$

$$P(c<X\leq c+l)=\int_c^{c+l}\frac{1}{b-a}dt=\frac{l}{b-a}$$ 与c无关

X落在(a, b) 区间中任意等长度子区间的概率相同

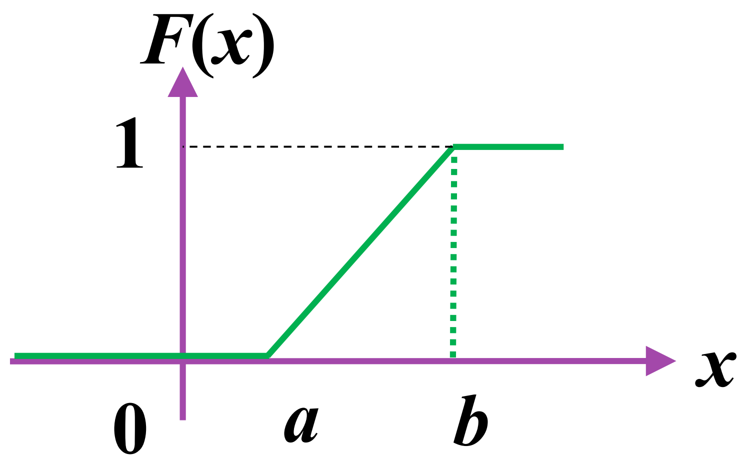
概率密度

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



分布函数

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



例 在区间(-1, 2)上随机取一数X，试写出X 的概率密度。
并求$P(X>0)$的值；

若在该区间上随机取10个数，求10个数中恰有两个数大于0的概率。

解 X 在区间(-1, 2) 上均匀分布

$$f(x)=\begin{cases}\frac{1}{3}, & -1<x<2\\ 0, & \text{其他}\end{cases} \qquad P(X>0)=\frac{2}{3}$$

设10个数中有Y个数大于0，

则 $Y\sim B(10,\frac{2}{3})$ $\qquad P(Y=2)=C_{10}^{2}\left(\frac{2}{3}\right)^{2}\left(\frac{1}{3}\right)^{8}$

例 若随机变量 ξ 在区间$(0，5)$上服从均匀分布，求方程 $4x^2+4\xi x+\xi+2=0$有实根的概率。

解 由于随机变量 ξ 在区间$(0，5)$上服从均匀分布，因此 ξ 的概率密度为

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

$$P(\text{“方程 } 4x^2+4\xi x+\xi+2=0 \text{ 有实根”})=P\left((4\xi)^2-4\cdot4\cdot(\xi+2)\geq0\right)$$

$$=P\left(\xi^2-\xi-2\geq0\right)=P\left(\xi\geq2\right)+P\left(\xi\leq-1\right)$$

$$=\int_2^5\frac{1}{5}dx=\frac{3}{5}$$

例 某人上班，自家里去办公室要经过一交通信号灯，这一信号灯有 80% 时间亮红灯，这时此人就在信号灯旁等待直至绿灯亮，等待时间(单位:秒)在区间[0， 30]上服从均匀分布。以 X 表示此人的等待时间，求 X 的分布函数。

解 设 A 表示事件"信号灯亮绿灯"， X 的分布函数为 $F(x)$，由全概率公式及题设得

当 $x<0$ 时，事件$\{X\le x\}$是不可能事件，则 $F(x)=P\left(X\le x\right)=0$

当$0\le x<30$时，$F(x)=P\left(X\le x\right)=P(A)P(X\le x\mid A)+P(\overline{A})P(X\le x\mid \overline{A})$

$$=0.2\times 1+0.8\times\frac{x}{30}=0.2+\frac{2}{75}x$$

当$x\ge 30$时，事件$\{X\le x\}$是必然事件，则$F(x)=1$

即 X 的分布函数为 $F(x)=\begin{cases}0, & x<0,\\ 0.2+\dfrac{2}{75}x, & 0\le x<30,\\ 1, & x\ge 30.\end{cases}$

结合几何概率，容易得到：

定理 3 设随机变量 $X \sim U[a, b]$，则 X 在 $[a, b]$ 的任一子区间上取值的概率等价于以 a、b 为端点的直线线段上的几何概率。

例 设随机变量 $X \sim U[-2, 3]$，试求 $P(-1 \le X \le 0)$、$P\left(-2 < X \le -\frac{1}{2}\right)$、$P\left(-\frac{1}{2} < X < \frac{1}{2}\right)$ 和 $P(|X| > 1)$

解 由于以 -2、3 为端点的直线线段长为 5，以 -1、0 为端点的直线线段是以 -2、3 为端点的直线线段的一部分，且长为 1，因此

$$P(-1 \le X \le 0) = \frac{1}{5}$$

同理可得，$$P\left(-2 < X \le -\frac{1}{2}\right) = \frac{3}{10}, P\left(-\frac{1}{2} < X < \frac{1}{2}\right) = \frac{1}{5}, P(|X| > 1) = \frac{3}{5}$$

2. 正态分布

若连续型随机变量 X 的概率密度为 $f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x<+\infty,$

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

显然，$f(x)\geq 0$。下面验证 $\int_{-\infty}^{+\infty}f(x)dx=1$

证

$$\int_{-\infty}^{+\infty}f(x)dx \overset{令t=\frac{x-\mu}{\sigma}}{=\!=} \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-t^2/2}\sigma dt=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt$$

$$记\ I=\int_{-\infty}^{+\infty}e^{-t^2/2}dt \Rightarrow I^2=\iint e^{-\frac{(x^2+y^2)}{2}}dxdy=\int_0^{2\pi}d\theta\int_0^{+\infty}re^{-r^2/2}dr$$

$$\Rightarrow I=\sqrt{2\pi} \Rightarrow \int_{-\infty}^{+\infty}f(x)dx=1$$

高斯

Carl Friedrich Gauss

Born: 30 Apr. 1777 in Brunswick, Duchy of Brunswick (now Germany)

Died: 23 Feb. 1855 in Göttingen, Hanover (now Germany)

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

$f(x)$的图像如图 2—5 所示，它关于 $x=\mu$ 对称，在 $x=\mu$ 处取得最大值 $f(u)=\frac{1}{\sqrt{2\pi}\sigma}$

若改变 μ (固定 σ)值，它将沿 x 轴平移，但其形状不变，如图 2 － 6 所示，称 μ 为位置参数;

若改变 σ (固定 μ)值，它的扁尖程度将改变，当 σ 越大时图形变得越扁，当 σ 越小时图形变得越尖，如图 2－7 所示，当 σ 变小时 X 落在 μ 附近的概率越大。

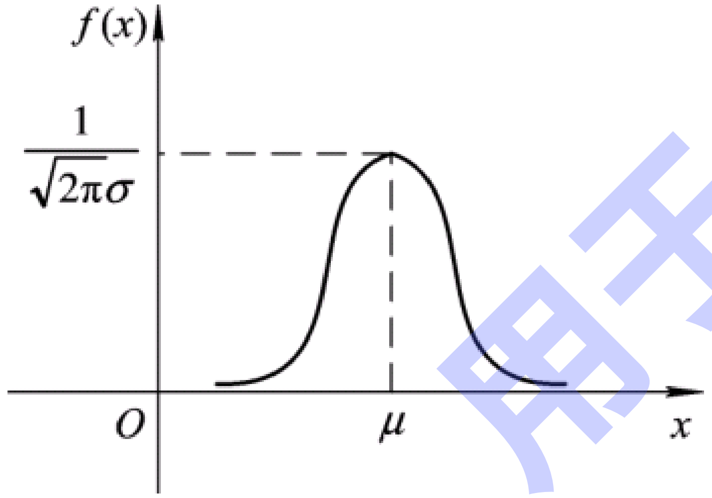


图 2–5

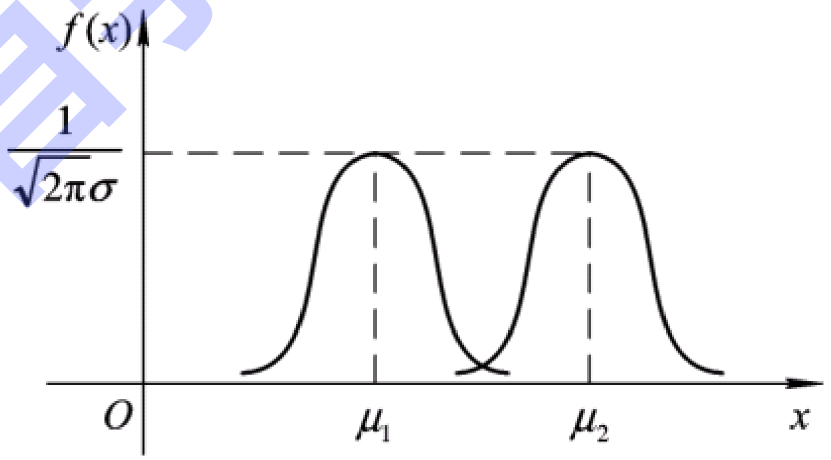


图 2–6

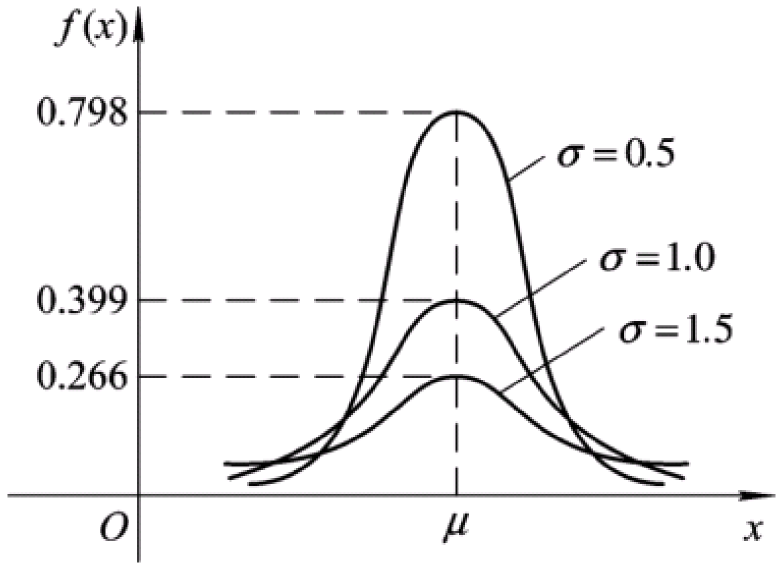


图2–7

正态随机变量 X 的概率密度为： $f(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}},-\infty<x<+\infty,$

正态随机变量 X 的分布函数为： $F(x)=\int_{-\infty}^{x}\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t-\mu)^2}{2\sigma^2}}\mathrm{d}t,\qquad -\infty<x<+\infty$

$F(x)$的图像如图2-8所示。显然，$F(\mu)=\frac{1}{2}$。

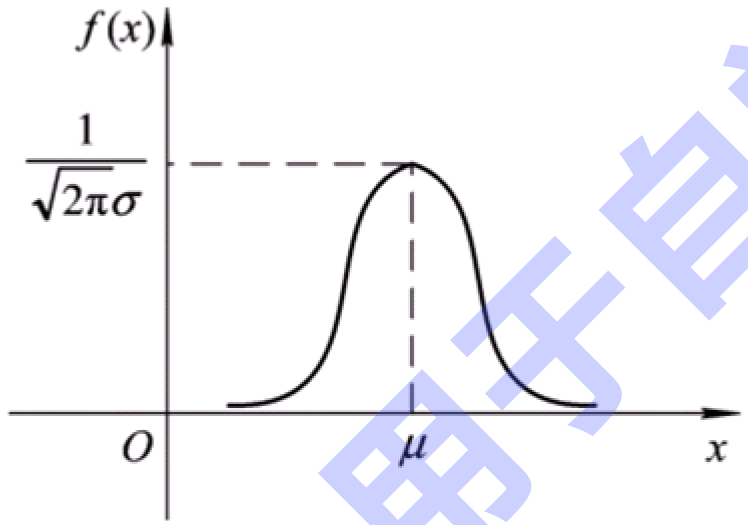


图 2-5

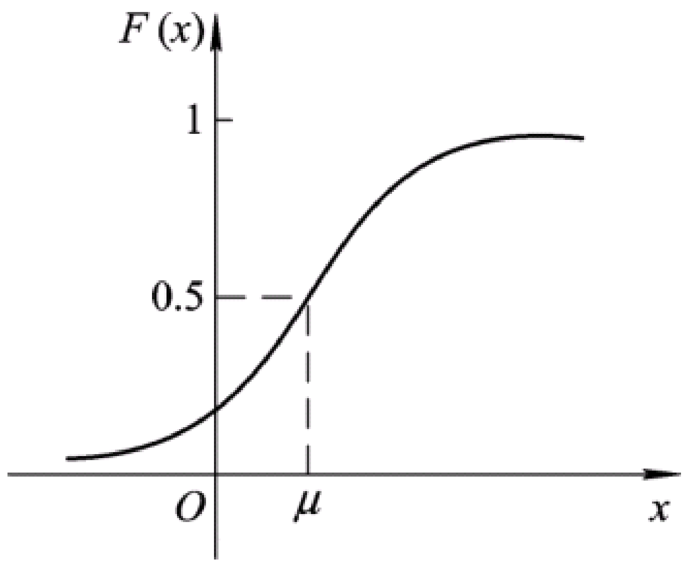
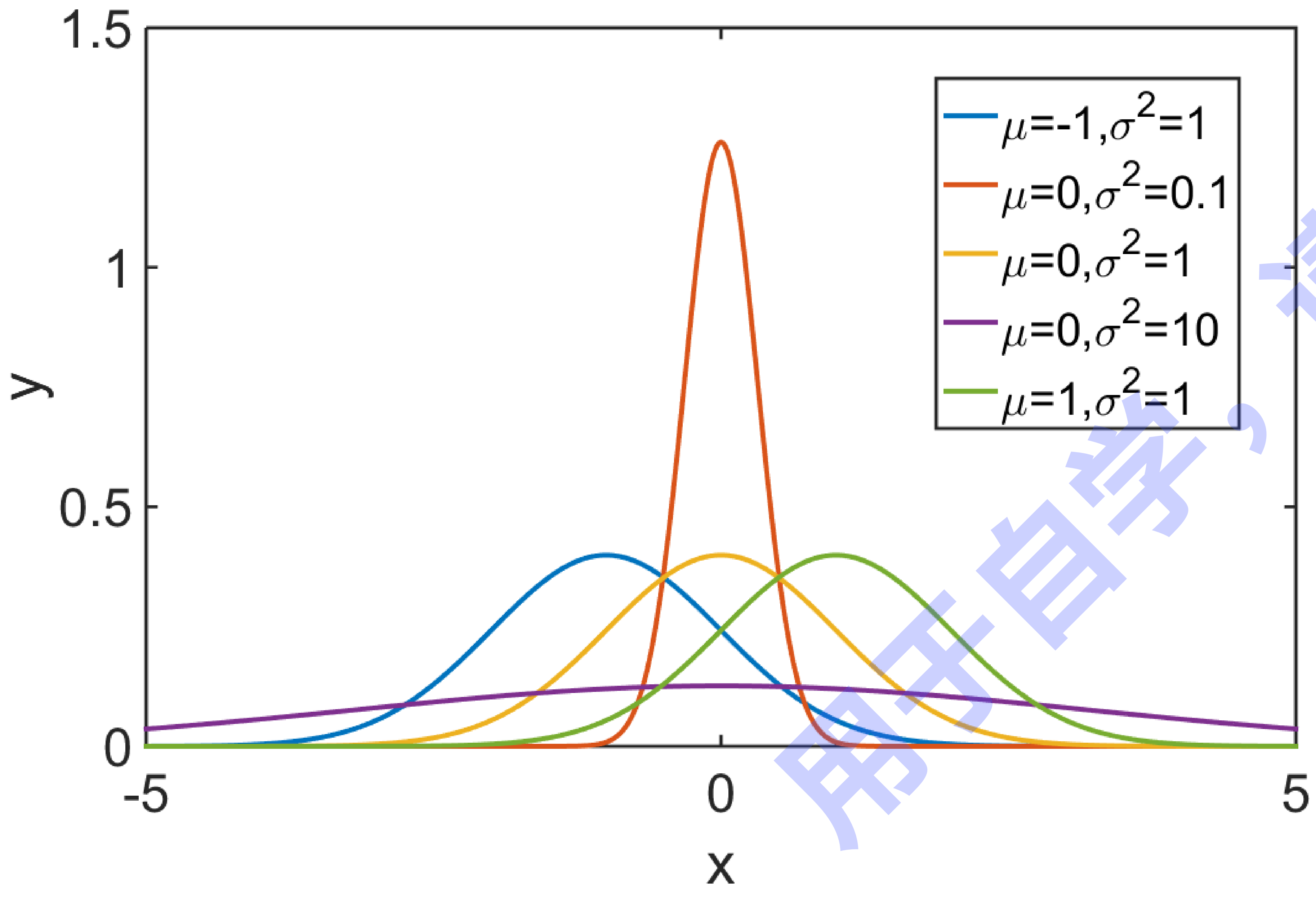
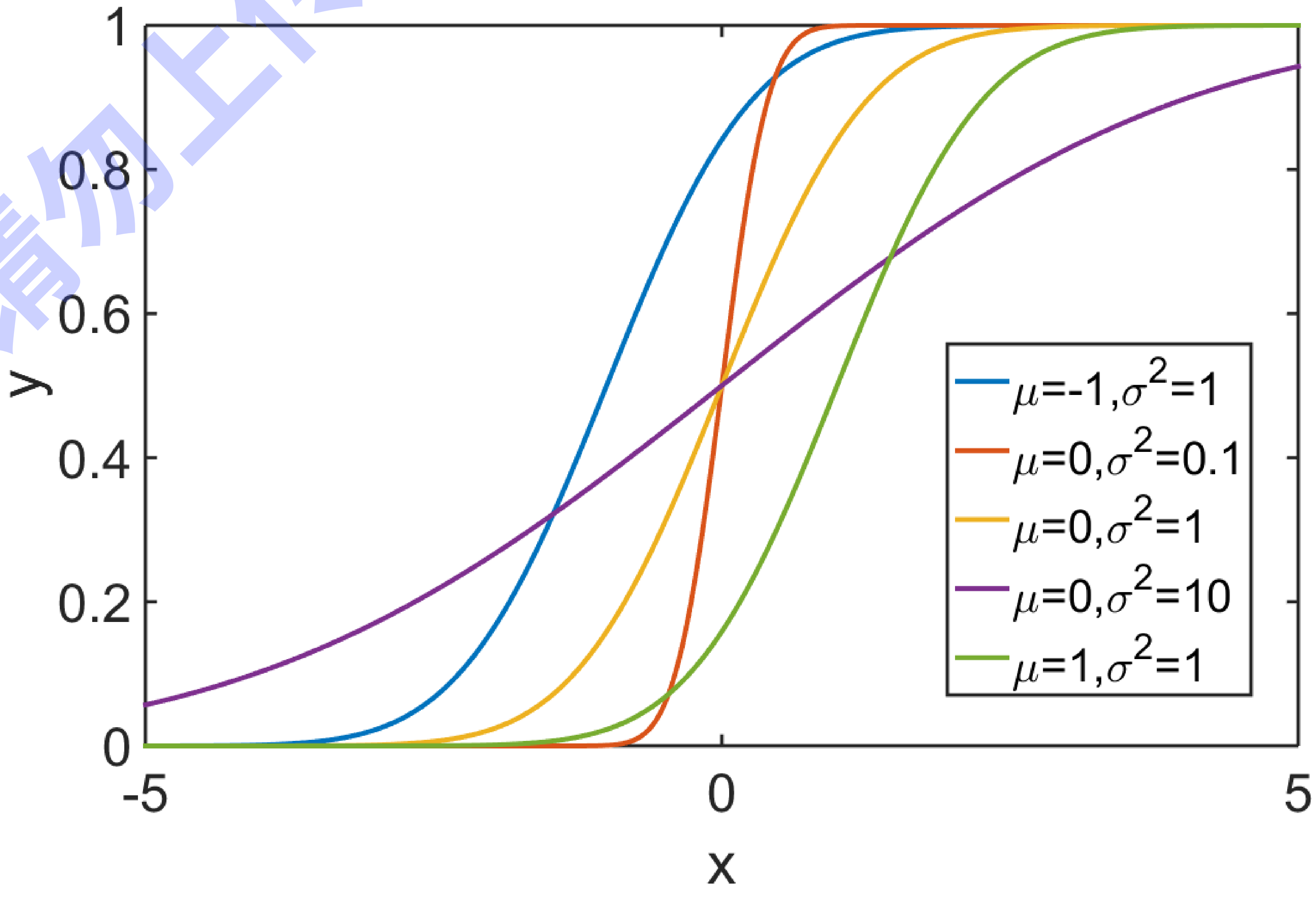


图2-8

不同参数设置正态分布的概率密度



不同参数设置正态分布的分布函数



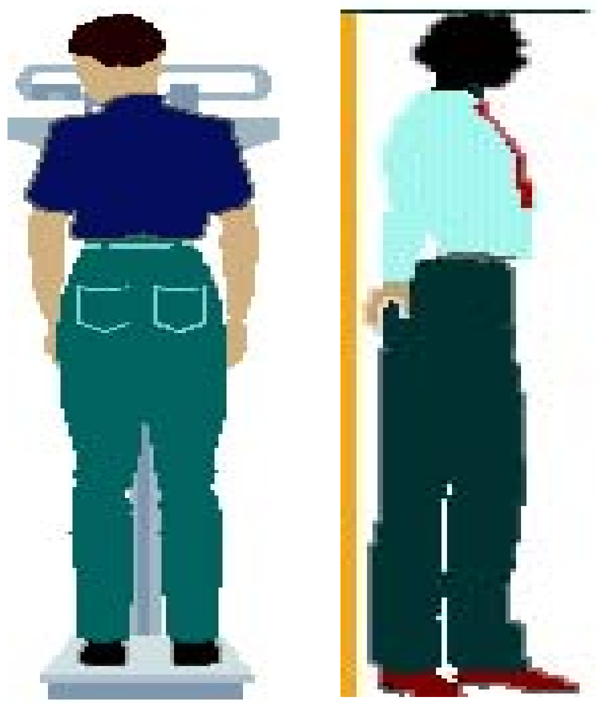
正态分布的应用与背景

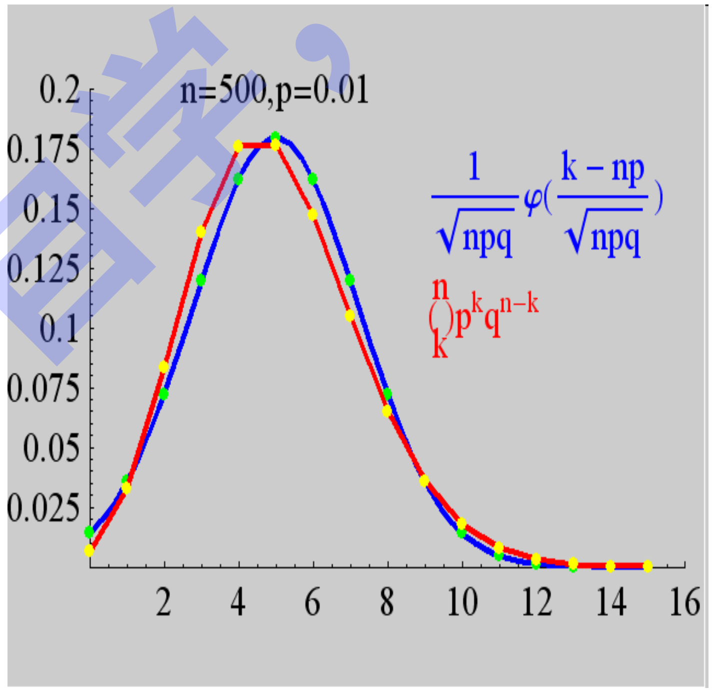
正态分布是最常见最重要的一种分布,例如

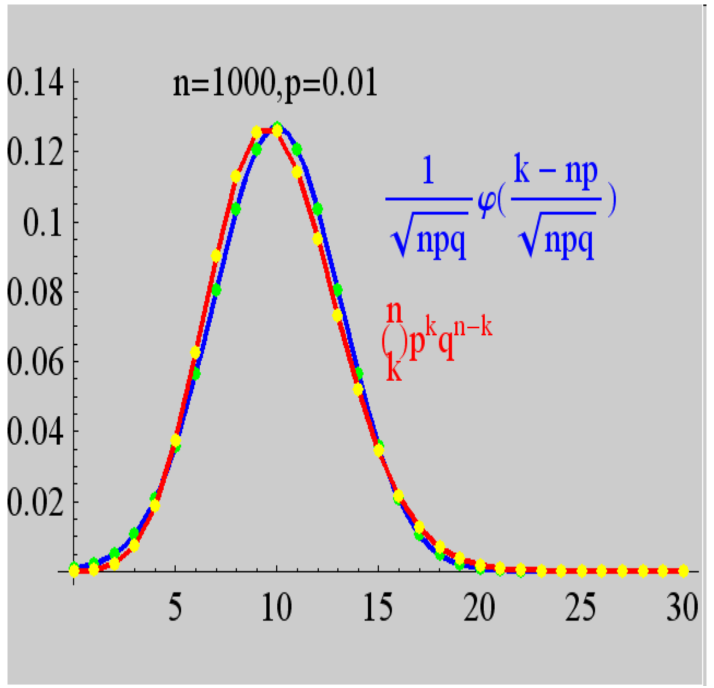
测量误差, 人的生理特征尺寸如身高、体重等 ;正常情况下生产的产品尺寸:直径、长度、重量、高度等都近 似服从正态分布.

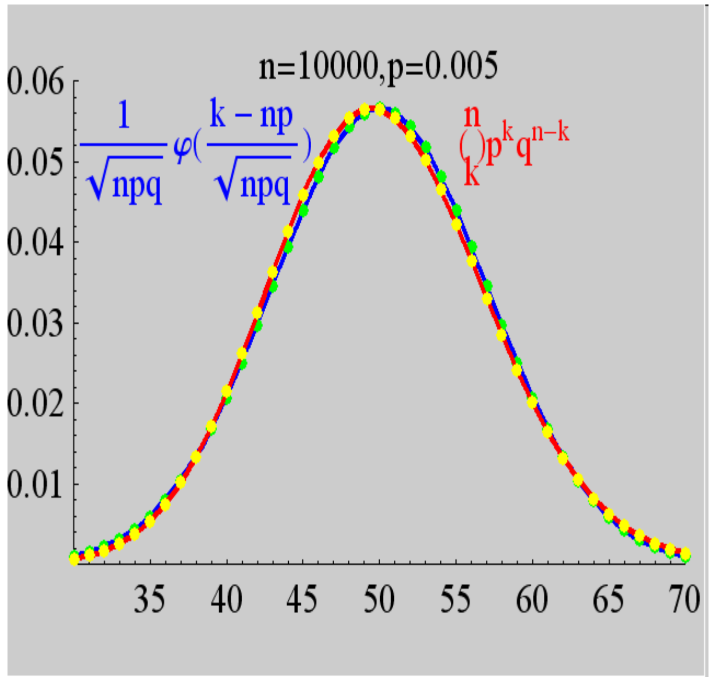
另一方面,有些分布(如二项分布、泊松分布)的极限分布是正态分布.所以,无论在实践中,还是在理论上,正态分布是概率论中最重要的一种分布.

二项分布向正态分布的转换









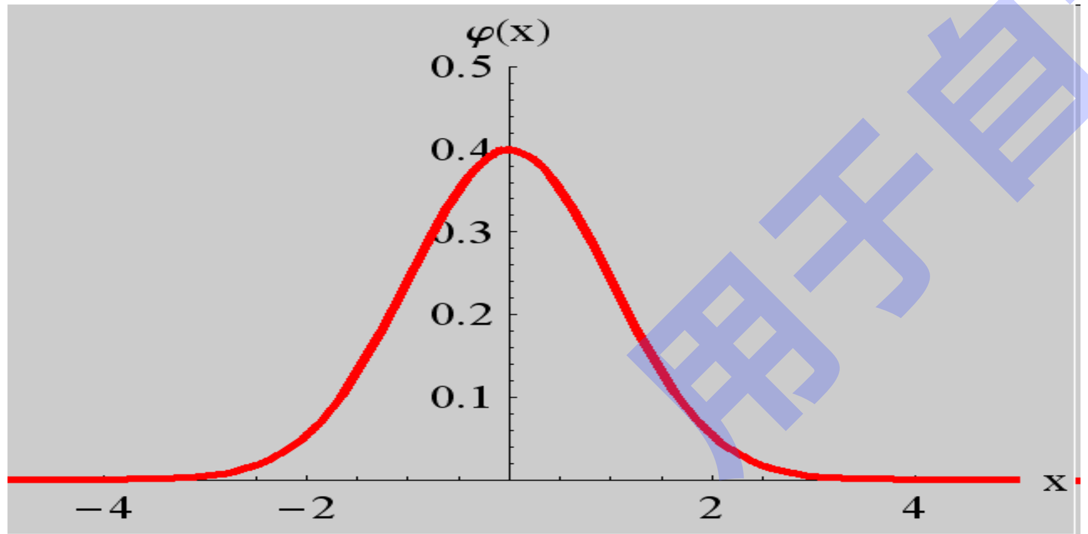
定义 设$X\sim N(\mu,\ \sigma^2)$，若$\mu=0$，$\sigma^2=1$，则称X服从**标准正态分布**，记为$X\sim N(0,\ 1)$。

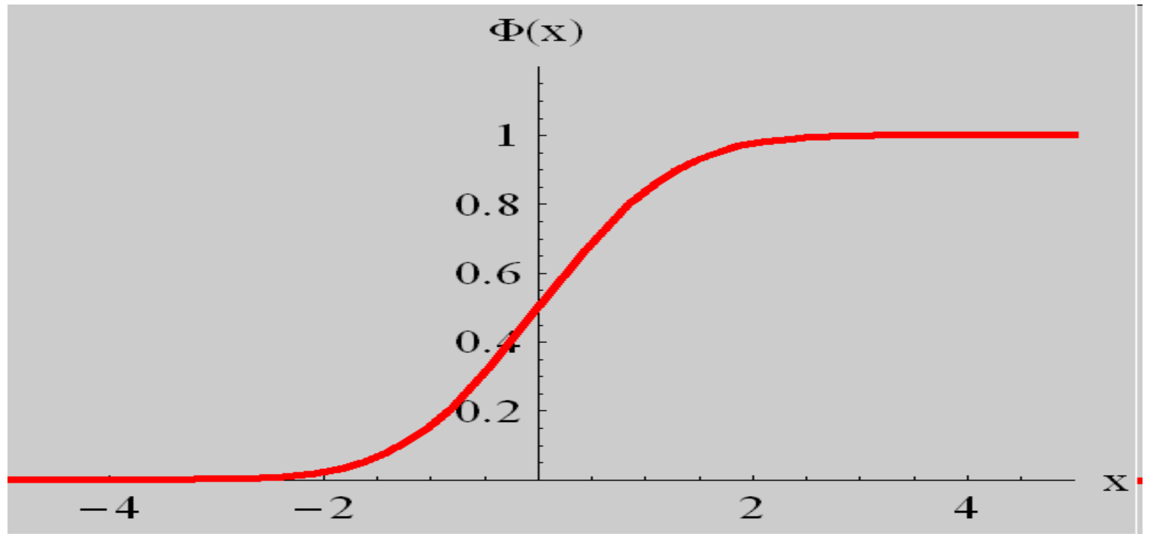
标准正态分布的概率密度$\varphi(x)$和分布函数$\Phi(x)$分别为：

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

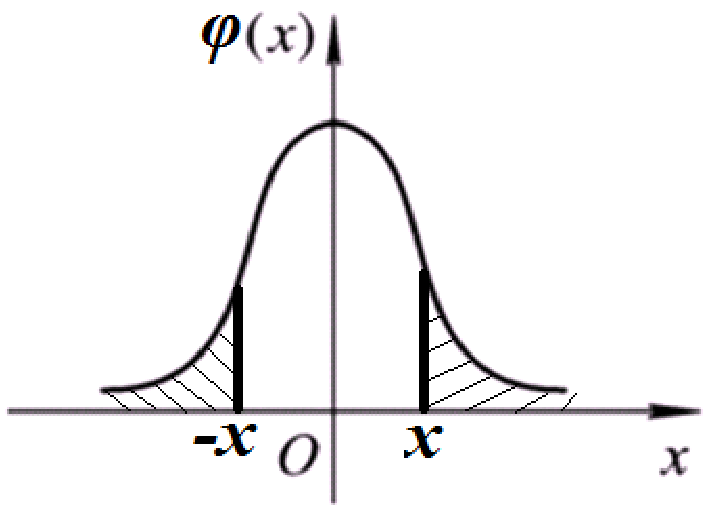
$$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\,\mathrm{d}t,\quad -\infty<x<+\infty.$$

$\varphi(x)$和$\Phi(x)$的图像分别如图2−9和2−10所示。





定理 4 设$X \sim N(0, 1)$，则$\Phi(-x)=1-\Phi(x)$.



证

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

$$=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}\,\mathrm{d}u-\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}\,\mathrm{d}u$$

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

由于**标准正态分布**在工程技术中有着重要的应用，因此人们为了使用方便，通过计算$\Phi(x)$的值，编制了**$\Phi(x)$的函数表(见附表 1)**，供实践中查用。这样，就解决了标准正态分布的问题，但在实际工作中，人们也会经常遇到一般的正态分布$X \sim N(\mu, \sigma^2)$，要解决一般正态分布的问题，自然需要知道其分布函数，那么如何得到$F(x)$的值呢？这时可以运用如下定理。

标准正态分布$N(0, 1)$的 x与$\Phi(x)$ 可以查表可知，那么其他正态分布$N(\mu, \sigma^2)$的$F(x)$呢？

定理 5 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

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

$$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\mu+\sigma x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt \overset{令v=\frac{t-\mu}{\sigma}}{=\!=} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-v^2/2}dv \quad =\Phi(x)$$

$$P\left(\frac{X-\mu}{\sigma}\le x\right)=\Phi(x) \qquad Z=\frac{X-\mu}{\sigma}\sim N(0,1)$$

定理 6 若$X \sim N(\mu, \sigma^2)$，则 $F(x) = P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$

对任意区间 $(x_1, x_2]$有 $P(x_1 < X \le x_2) = P\left(\frac{x_1-\mu}{\sigma} < \frac{X-\mu}{\sigma} \le \frac{x_2-\mu}{\sigma}\right) = \Phi\left(\frac{x_2-\mu}{\sigma}\right) - \Phi\left(\frac{x_1-\mu}{\sigma}\right)$

证 由定理 5，得

$$F(x) = P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = P\left(Z \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

于是由分布函数的性质可得：

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

例 设随机变量 $X \sim N(10,\ 0.022)$，已知 $\Phi(2.5)=0.9938$，求 X 在区间 $(9.95,\ 10.05)$ 内的概率。

解

$$P(9.95 < X < 10.05) = \Phi\left(\frac{10.05-10}{0.02}\right) - \Phi\left(\frac{9.95-10}{0.02}\right) = \Phi(2.5) - \Phi(-2.5)$$

$$= 2\Phi(2.5) - 1 = 2 \times 0.9938 - 1 = 0.9876$$

例 设随机变量 $X \sim N(1.5,\ 4)$，求 $P(X<3.5)$、$P(X>2)$、$P(X<-2)$ 和 $P(2<X<4)$。

解

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

$$P(X > 2) = 1 - P(X \le 2) = 1 - \Phi\left(\frac{2-1.5}{2}\right) = 1 - \Phi(0.25) = 1 - 0.5987 = 0.4013$$

$$P(X < -2) = \Phi\left(\frac{-2-1.5}{2}\right) = \Phi(-1.75) = 1 - \Phi(1.75) = 1 - 0.9599 = 0.0401$$

$$P(2 < X < 4) = \Phi\left(\frac{4-1.5}{2}\right) - \Phi\left(\frac{2-1.5}{2}\right) = \Phi(1.25) - \Phi(0.25) = 0.8944 - 0.5987 = 0.2957$$

例 设$X\sim N(2,\ \sigma^2)$，且$P(2<X<4)=0.3$，求$P(X<0)$。

解 由 $P(2<X<4)=\Phi\left(\frac{4-2}{\sigma}\right)-\Phi\left(\frac{2-2}{\sigma}\right)=\Phi\left(\frac{2}{\sigma}\right)-\Phi(0)=0.3$ 得

$$\Phi\left(\frac{2}{\sigma}\right)=0.5+0.3=0.8$$

从而，$P(X<0)=\Phi\left(\frac{0-2}{\sigma}\right)=1-\Phi\left(\frac{2}{\sigma}\right)=1-0.8=0.2$

例 若$X\sim N(\mu,\sigma^2)$，则

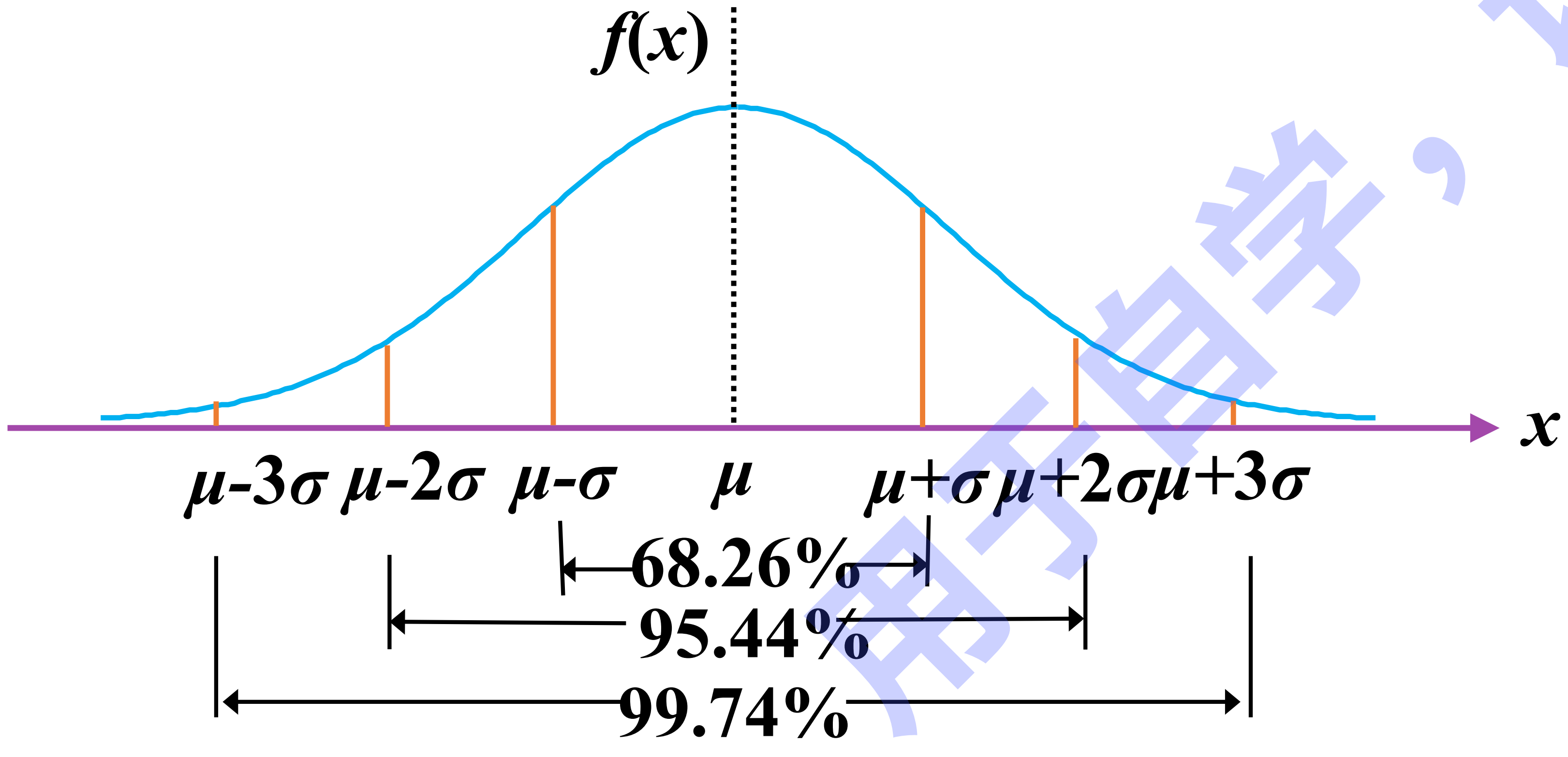
$$P(\mu-\sigma<X<\mu+\sigma)=\Phi(1)-\Phi(-1)=2\Phi(1)-1=68.26\%$$

$$P(\mu-2\sigma<X<\mu+2\sigma)=\Phi(2)-\Phi(-2)=95.44\%$$

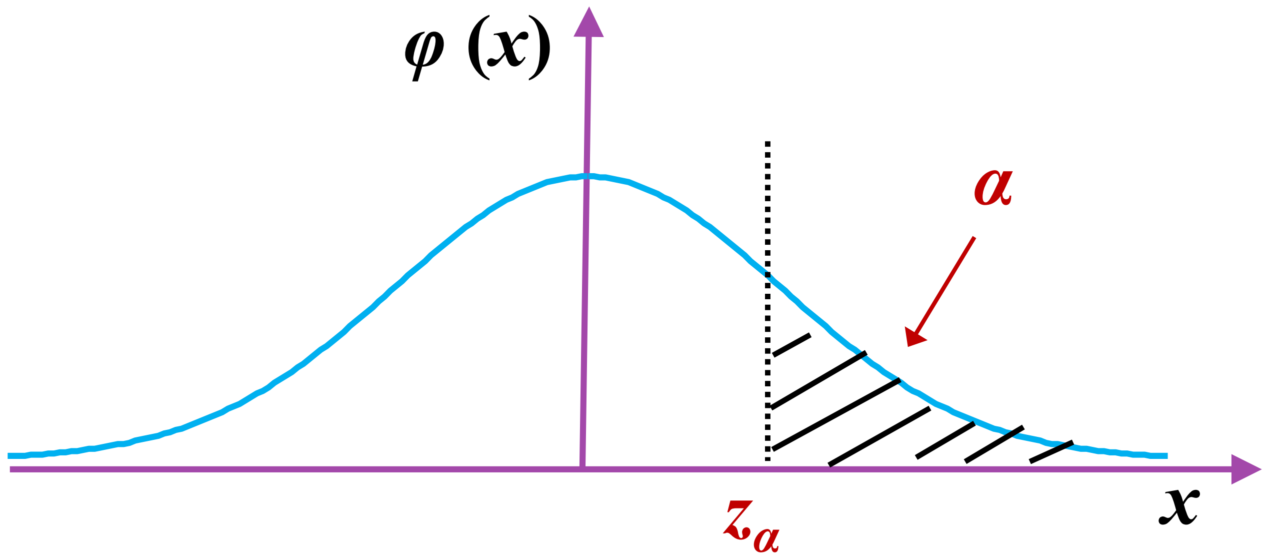
$$P(\mu-3\sigma<X<\mu+3\sigma)=\Phi(3)-\Phi(-3)=99.74\%$$

在$(-\infty,\infty)$落在$(\mu-3\sigma,\mu+3\sigma)$区间几乎是必然的

“3σ规则”



上α分位点 $X\sim N(0,1)$



$P(X>z_\alpha)=\alpha$ 称z_α是标准正态分布的**上α分位点**

$z_{1-\alpha}=-z_\alpha$

例 设某地区男子身高$X\,(\mathrm{cm}) \sim N(169.4, 4.1^2)$

(i) 从该地区随机找一男子测身高，求他的身高大于175 cm的概率；

(ii) 若从中随机找5个男子测身高，问至少有一人身高大于175 cm的概率是多少？

恰有一人身高大于175 cm的概率为多少？

解 (i) $P(X>175)=1-\Phi\left(\dfrac{175-169.7}{4.1}\right)=1-\Phi(1.293)\overset{\text{查附表}}{=\!=}1-0.9015=0.0985$

(ii) 设5人中有Y人身高大于175 cm ，则 $Y\sim B(5,p)$，其中 $p=0.0985$

$$P(Y\ge 1)=1-P(Y=0)=1-(1-p)^5=0.4045$$

$$P(Y=1)=C_5^1p^1(1-p)^4=0.3253$$

例 一批钢材长度（线材，单位cm）$X \sim N(\mu, \sigma^2)$

(i) 若 $\mu=100$，$\sigma=2$，求这批钢材长度小于97.8 cm 的概率；

(ii) 若 $\mu=100$，要使这批钢材的长度至少有90%落在区间 (97, 103)内，问 σ 至多取何值？

解 (i) $$P(X<97.8)=\Phi\left(\frac{97.8-100}{2}\right)=1-\Phi(1.1)\overset{\text{查附表}}{=}1-0.8643=0.1357$$

(ii) 令 $P(97<X<103)\geq 90\%$

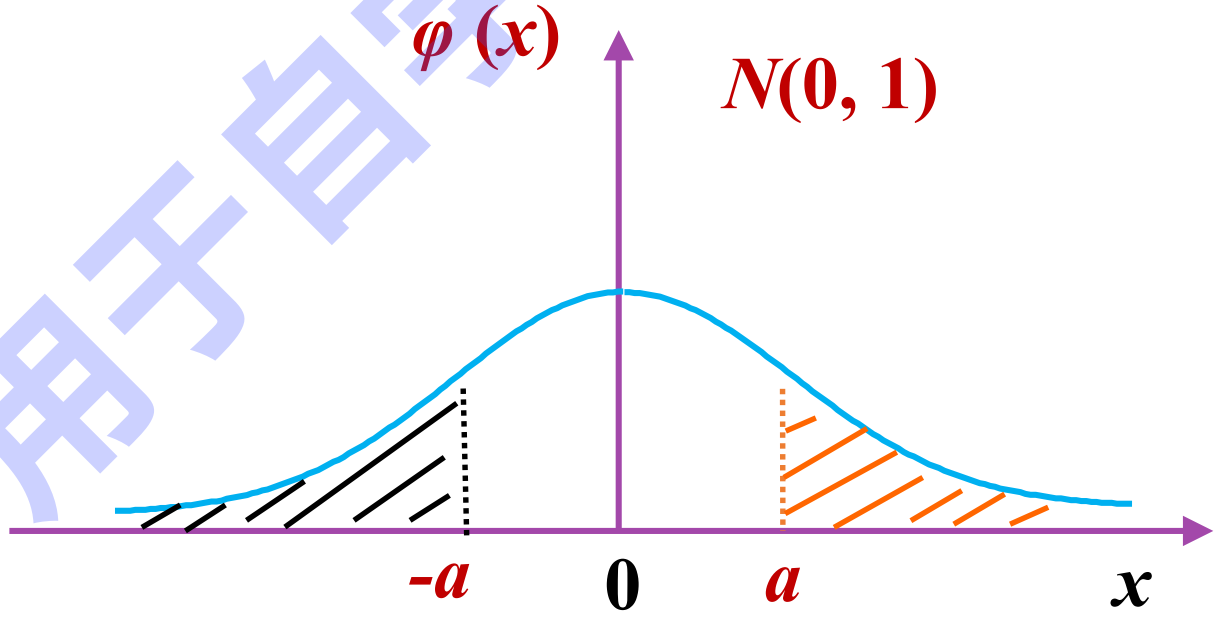
即 $$\Phi\left(\frac{103-100}{\sigma}\right)-\Phi\left(\frac{97-100}{\sigma}\right)=2\Phi\left(\frac{3}{\sigma}\right)-1\geq 90\%$$

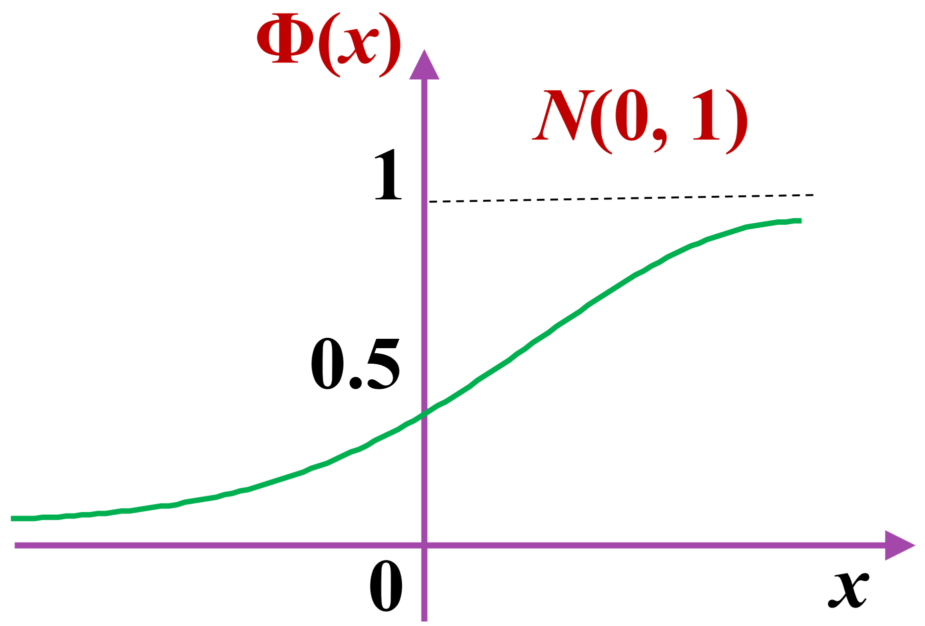
$$\Phi\left(\frac{3}{\sigma}\right)\geq 0.95 \Rightarrow \frac{3}{\sigma}\geq 1.645 \Rightarrow \sigma\leq 1.8237$$

	概率密度	分布函数
正态分布$X \sim N(\mu, \sigma^2)$	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt$
标准正态分布$X \sim N(0, 1)$	$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$	$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-t^2/2}dt$

$N(0, 1)$的分布函数可以查$\Phi(x)$表获得

$\Phi(-x)=1-\Phi(x) \quad x \geq 0$





3. 指数分布

若连续型随机变量 X 的概率密度为

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

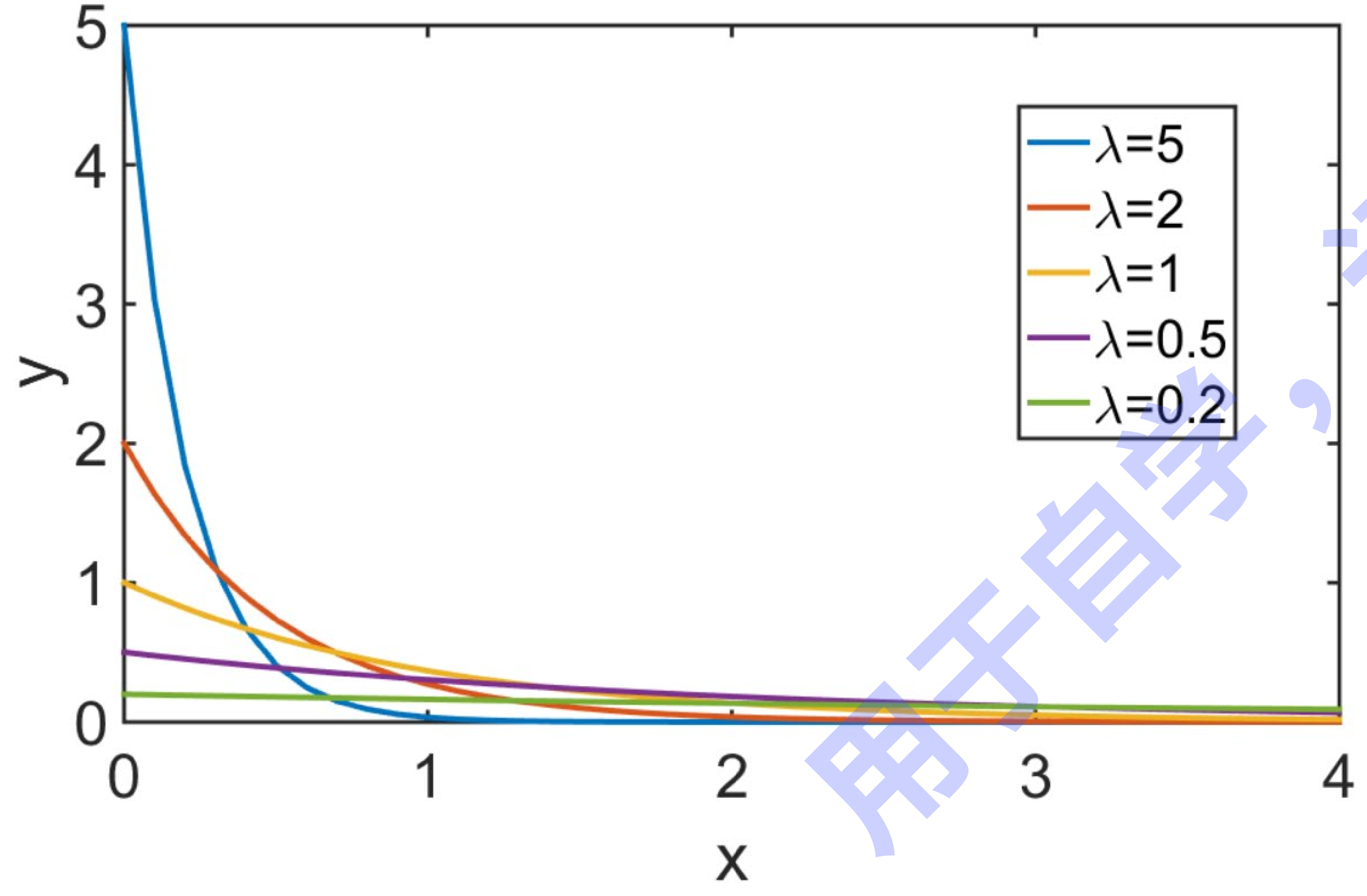
其中$\lambda>0$ 为常数，则称 X 服从参数为 λ的指数分布，记为$X\sim E(\lambda)$。

显然，$f(x)\ge 0$，且 $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=\int_{0}^{+\infty}\lambda e^{-\lambda x}\mathrm{d}x=\left[-e^{-\lambda x}\right]_0^{+\infty}=1$

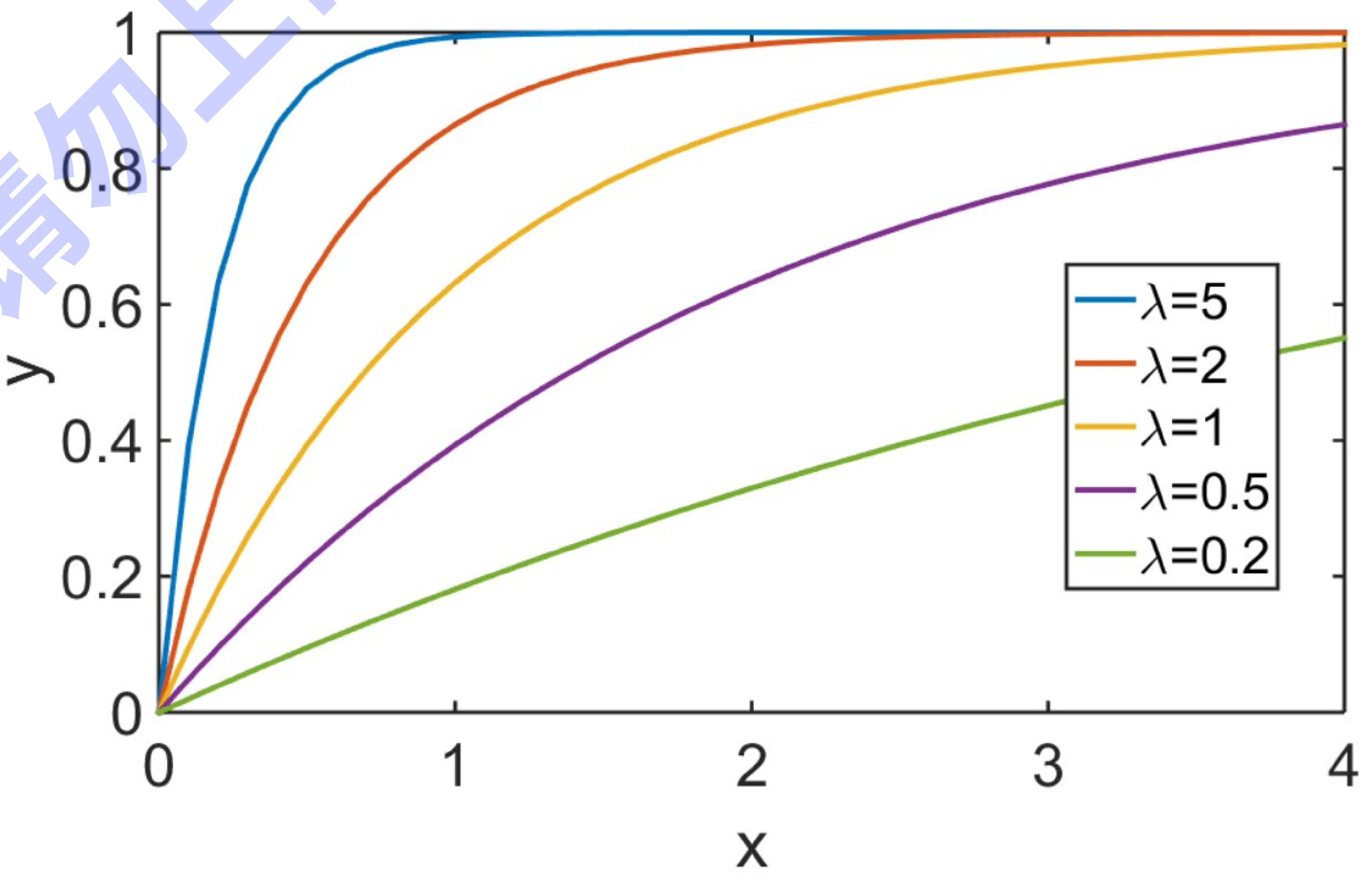
X的分布函数为 $F(x)=\begin{cases}1-e^{-\lambda x}, & x>0\\ 0, & x\le 0\end{cases}$

指数分布有着重要的应用，常被用来描述寿命类随机变量的分布。如电子元件的寿命、生物的寿命、电话的通话时间、随机服务系统的服务时间等都可以认为服从指数分布。

不同参数设置指数分布的概率密度



不同参数设置指数分布的分布函数



定理 7 指数分布具有**无记忆性**，即设 $X \sim E(\lambda)$，则 $\forall s, t > 0$，

$$P(X > s + t \mid X > s) = P(X > t)$$

证

$$P(X > s + t | X > s) = \frac{P(X > s + t, X > s)}{P(X > s)}$$

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

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

例 设 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 的分布函数为

$$F(x)=\begin{cases}1-e^{-0.4x}, & x>0\\ 0, & x\le 0\end{cases}$$

求：（1）至多等待3分钟的概率；
（2）至少等待4分钟的概率；
（3）等待3分钟至4分钟的概率；
（4）至多等待3分钟或至少等待4分钟的概率；
（5）恰好等待2.5分钟的概率

解 X服从指数分布，是连续型随机变量

$$(1)\ P(X\le 3)=F(3)=1-e^{-1.2}$$

$$(2)\ P(X\ge 4)=1-P(X<4)=1-P(X\le 4)=e^{-1.6}$$

$$(3)\ P(3\le X\le 4)=P(3<X\le 4)=F(4)-F(3)=e^{-1.2}-e^{-1.6}$$

$$(4)\ P(\{X\le 3\}\cup\{X\ge 4\})=P(X\le 3)+P(X\ge 4)=1-e^{-1.2}+e^{-1.6}$$

$$(5)\ P(X=2.5)=0$$

例 设随机变量 X 服从参数为 $\lambda\,(\lambda>0)$ 的指数分布，且 $P(X\le 1)=1/2$。

求: (1)参数 λ; (2)$P(X>2\mid X>1)$。

解 随机变量 X 的分布函数为

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

(1)由于$1/2=P(X\le 1)=F(1)=1-e^{-\lambda}$，因此$\lambda=\ln 2$

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

例 **假设一设备开机后无故障工作的时间 X (单位:小时)服从参数 $\lambda=1/5$ 的指数分布。设备定时开机，出现故障时自动关闭，而在无故障的情况下工作 2 小时便关机。**

试求设备每次开动无故障工作的时间 Y 的分布函数 $F(y)$。

解 **由于 X 服从参数 $\lambda=1/5$ 的指数分布，因此 X 的概率密度为** $f(x)=\begin{cases}\frac{1}{5}e^{-\frac{x}{5}}, & x>0\\ 0, & x\le 0\end{cases}$

显然$Y=\min\{X,2\}$

当$y<0$时，事件$\{Y\le y\}$是不可能事件，则$F(y)=P(Y\le y)=0$

当$0\le y<2$时，$F(y)=P(Y\le y)=P(\min\{X,2\}\le y)=P(X\le y)=\int_0^y \frac{1}{5}e^{-\frac{x}{5}}\mathrm{d}x=1-e^{-\frac{y}{5}}$

当$y\ge 2$时，事件$\{Y\le y\}$是必然事件，则$F(y)=1$

即Y的分布函数为 $F(y)=\begin{cases}0, & y<0\\ 1-e^{-\frac{y}{5}}, & 0\le y<2\\ 1, & y\ge 2\end{cases}$

例 假设一大型设备在任何长为 t 的时间(单位:小时)内发生故障的次数 $N(t)$ 服从参数为 λ 的 Poisson 分布。

（1）求相继两次故障之间时间间隔 T 的概率分布;

（2）求在设备无故障工作 8 小时的情形下，再无故障工作 8 小时的概率。

解 (1)先求 T 的分布函数为 $F_T(t)$。

当 $t<0$ 时，事件$\{T\le t\}$是不可能事件，则 $F_T(t)=P(T\le t)=0$

当 $t\ge 0$ 时， $F_T(t)=P(T\le t)=1-P(T>t)=1-P(N(t)=0)=1-e^{-\lambda t}$

即T的分布函数为 $F_T(t)=\begin{cases}1-e^{-\lambda t}, & t\ge 0\\ 0, & t<0\end{cases}$, 从而$T$的概率密度为 $f_T(t)=\begin{cases}\lambda e^{-\lambda t}, & t\ge 0\\ 0, & t<0\end{cases}$

即 T 服从参数为 λ 的指数分布。

(2) $P(T\ge 16|T\ge 8)=P(T\ge 8+8|T\ge 8)=P(T>8)=1-P(T\le 8)=e^{-8\lambda}$

（指数分布的无记忆性）

本节回顾

- 连续型随机变量

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

分布函数　概率密度

- 常见连续型随机变量的分布

$$\begin{cases} \text{均匀分布} \\ \text{正态分布（或高斯分布）} \\ \text{指数分布} \end{cases}$$