§2.2 连续型随机变量

一、连续型随机变量的概念

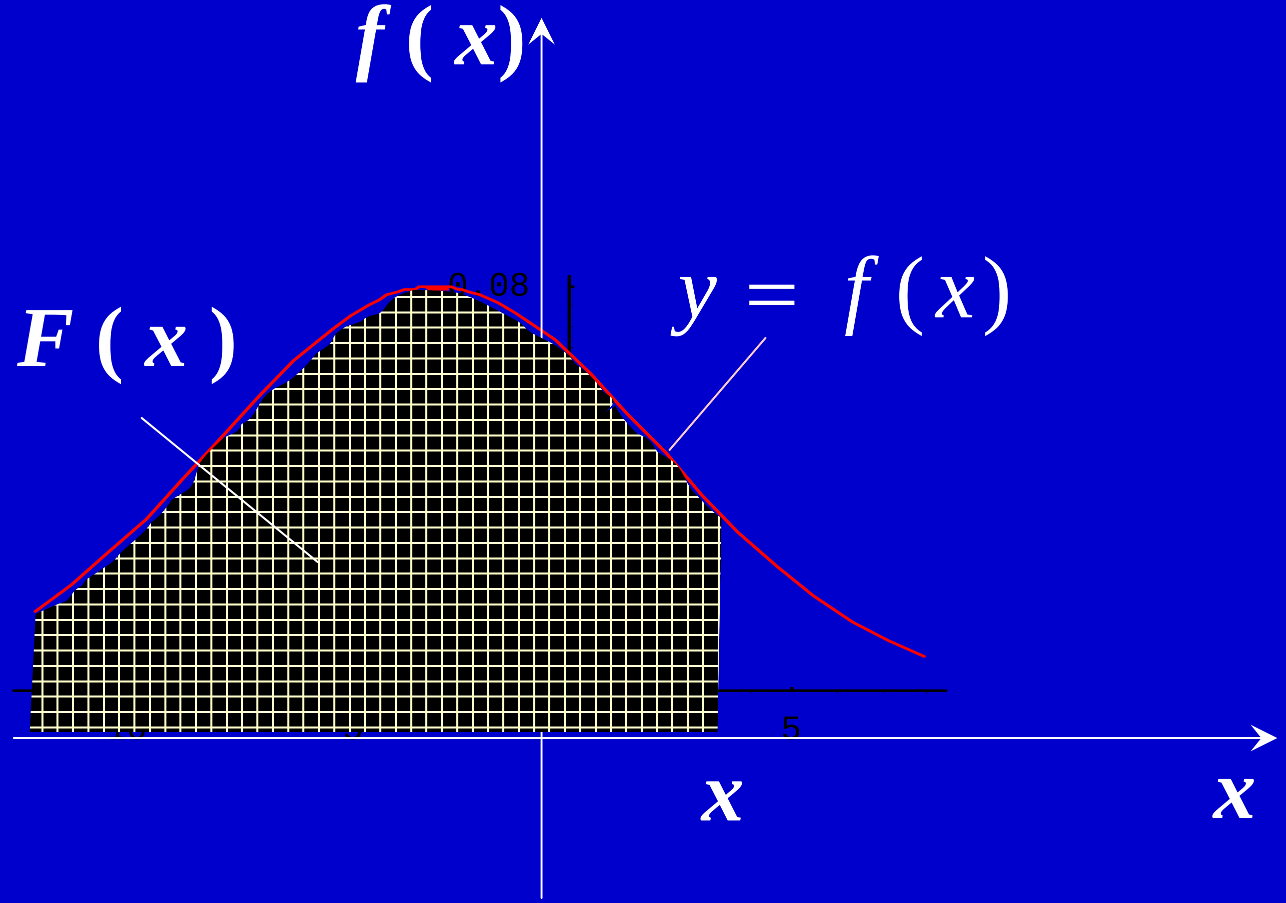
1、定义 设 X 是一随机变量，若存在一个非负可积函数 $f(x)$, 使得

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

其中$F(x)$是它的分布函数

则称 X 是连续型随机变量，$f(x)$是它的概率密度函数(p.d.f.)，简称为概率密度或密度函数

分布函数$F(x)$与概率密度$f(x)$的几何意义



2、概率密度 $f(x)$的性质

1) $f(x) \geq 0$

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

常利用这两个性质检验一个函数能否作为连续型随机变量的概率密度，或求其中的未知参数

3) 在 $f(x)$ 的连续点处，

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

$f(x)$ 描述了X 在 x 附近单位长度的区间内取值的概率

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

不是Cauchy 积分，而是Lesbesgue 意义下的积分，所得的变上限的函数是绝对连续的，因此几乎处处可导

$$F'(x_0)=\lim_{\Delta x\to +0}\frac{F(x_0+\Delta x)-F(x_0)}{\Delta x}$$

$$=\lim_{\Delta x\to +0}\frac{P\{x_0<X\le x_0+\Delta x\}}{\Delta x} \quad = f(x_0)$$

$$f(x_0)\,\Delta x \approx \boxed{P\{x_0<X\le x_0+\Delta x\}}$$

密度 长度 线段质量

注意: 对于连续型随机变量X, $P\{X=a\}=0$
这里 a 可以是随机变量 X 的任意一个可能的取值

事实上 $\{X=a\} \subset \{a-\Delta x < X \le a\}$ $\Delta x > 0$

$$0 \le P\{X=a\} \le P\{a-\Delta x < X \le a\} = \int_{a-\Delta x}^{a} f(x)\mathrm{d}x$$

$$0 \le P\{X=a\} \le \lim_{\Delta x \to +0} \int_{a-\Delta x}^{a} f(x)\mathrm{d}x = 0$$

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

命题 连续型随机变量取任一常数的概率为零

强调 概率为1 (零) 的事件未必发生 (不发生)

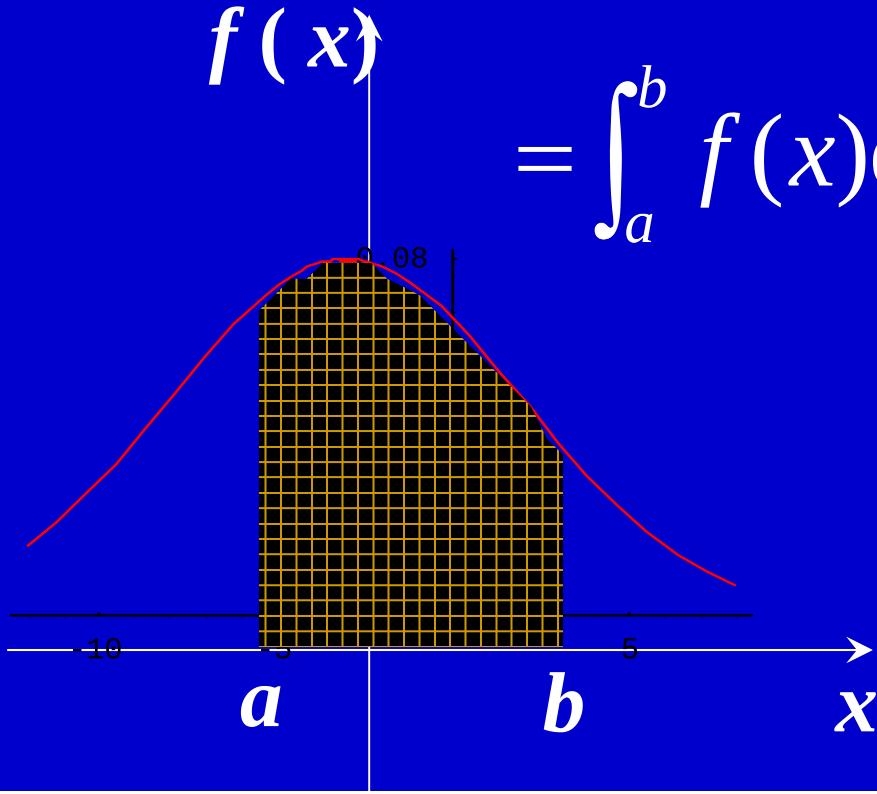
对于连续型随机变量X

$$P\{a < X \leq b\} = P\{a \leq X \leq b\}$$

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

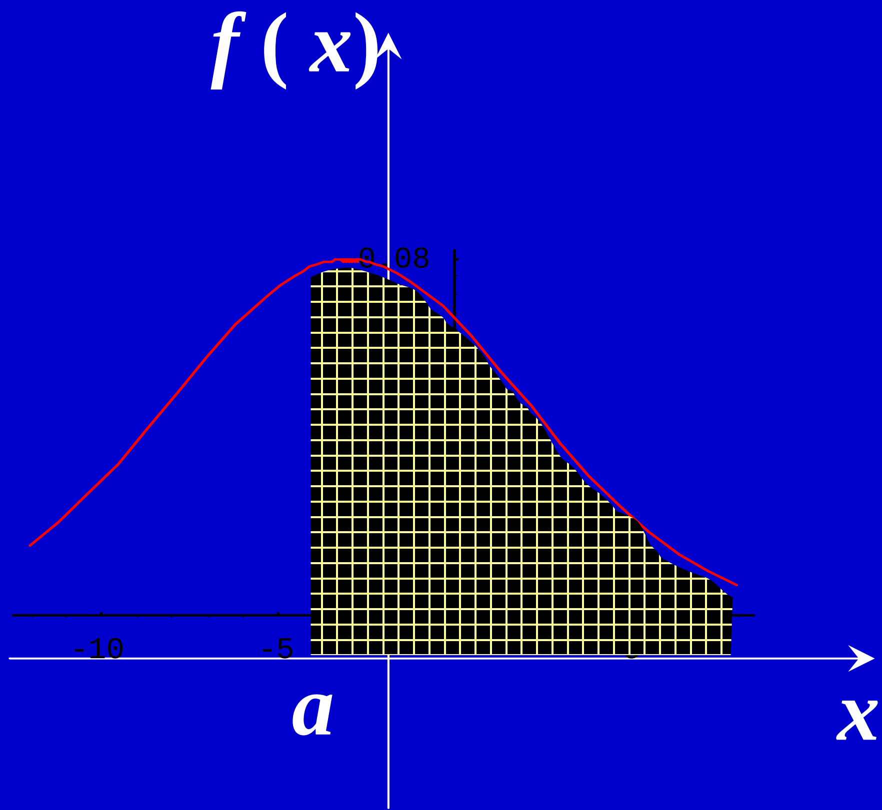
$$= P\{a \leq X < b\}$$

$$= \int_a^b f(x)\mathrm{d}x = F(b) - F(a)$$



$$P\{X \le b\} = P\{X < b\} = F(b)$$

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



例1 有一批晶体管，已知每只的使用寿命 X 为连续型随机变量，其概率密度函数为

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

(1) 求常数 c

(2) 已知一只收音机上装有3只这样的晶体管，每只晶体管能否正常工作相互独立，求在使用的最初1500小时只有一个损坏的概率.

解 (1) $$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = \int_{1000}^{+\infty} \frac{c}{x^2}\mathrm{d}x \overset{令}{=} 1$$

$$\longrightarrow \quad c = 1000$$

(2) 设事件 A 表示一只晶体管的寿命小于1500小时

$$P(A) = P\{0 \le X < 1500\}$$

$$= \int_{1000}^{1500} \frac{1000}{x^2}\mathrm{d}x = \frac{1}{3}$$

设在使用的最初1500小时三只晶体管中损坏的只数为 $Y \sim B\left(3, \frac{1}{3}\right)$

$$P\{Y = 1\} = C_3^1\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^2 = \frac{4}{9}$$

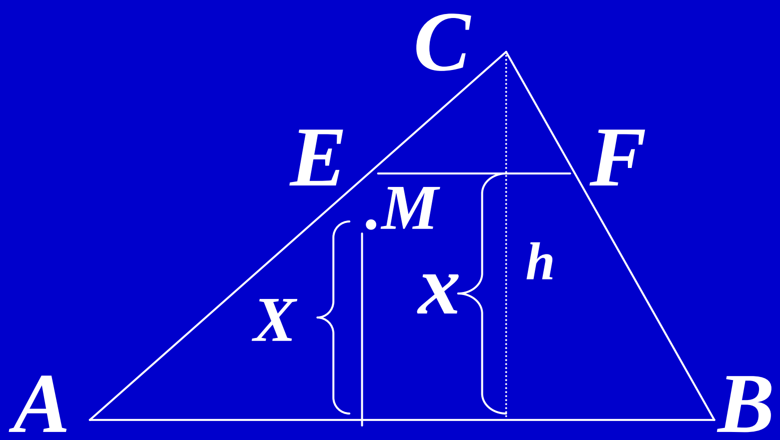
例2 在高为h 的$\triangle ABC$ 中任取一点M ,点M到AB 的距离为X , 求X 的概率密度函数 $f(x)$.

解 作$EF \parallel AB$, 使EF 与AB间的距离为x

当 $0 \le x \le h$ 时

$$F(x) = P\{X \le x\} = \frac{S_{\square EFBA}}{S_{\triangle ABC}}$$

$$= 1 - \frac{S_{\triangle CEF}}{S_{\triangle ABC}} = 1 - \left(\frac{h-x}{h}\right)^2$$



于是

$$F(x)=\begin{cases} 0 & x<0 \\ 1-(\dfrac{h-x}{h})^2 & 0\le x\le h \\ 1 & x>h \end{cases}$$

当$0<x<h$时，$f(x)=F'(x)=\dfrac{2(h-x)}{h^2}$

$$\therefore f(x)=\begin{cases} \dfrac{2(h-x)}{h^2} & 0<x<h \\ 0 & \text{其他} \end{cases}$$

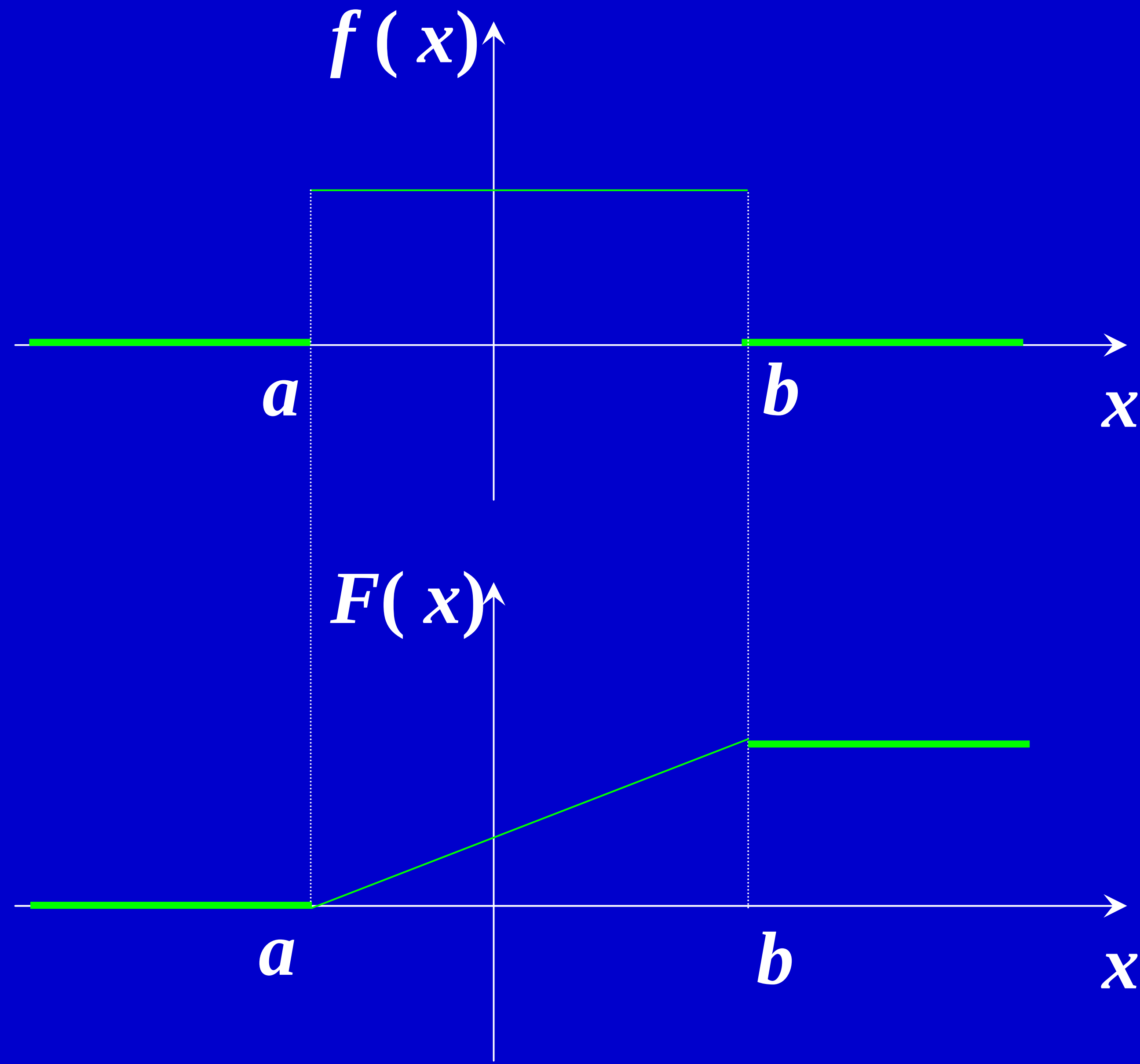
二、常见的连续型随机变量的分布

1、 均匀分布

若 X 的概率密度为 $f(x)$，则称 X 服从**区间 (a,b) 上的均匀分布** 记作 $X \sim U(a,b)$

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

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)
a
b
x
F(x)
a
b
x

$\forall (c,d) \subset (a,b)$

$$P\{c < X < d\} = \int_c^d \frac{1}{b-a} \mathrm{d}x = \frac{d-c}{b-a}$$

即 X 的取值在(a,b)内任何长为 $d-c$ 的小区间的概率与小区间的位置无关，只与其长度成正比．这正是几何概型的情形.

应用场合

在进行大量数值计算时，如果在小数点后第 k 位进行四舍五入，则产生的误差可以看作

服从 $U\left(-\frac{1}{2}10^{-k}, \frac{1}{2}10^{-k}\right)$

例3 秒表的最小刻度差为0.01秒. 若计时精度是取最近的刻度值, 求使用该秒表计时产生的随机误差X的概率密度, 并计算误差的绝对值不超过0.004秒的概率.

解 由题设知随机误差 X 等可能地取得区间 $[-0.005,0.005]$上的任一值，则

$$X \sim U[-0.005, 0.005]$$

$$f(x)=\begin{cases}100, & |x| \le 0.005 \\ 0, & \text{其他}\end{cases}$$

所以 $P\{|X| \le 0.004\} = \int_{-0.004}^{0.004} 100\, dx = 0.8$

2、 指数分布

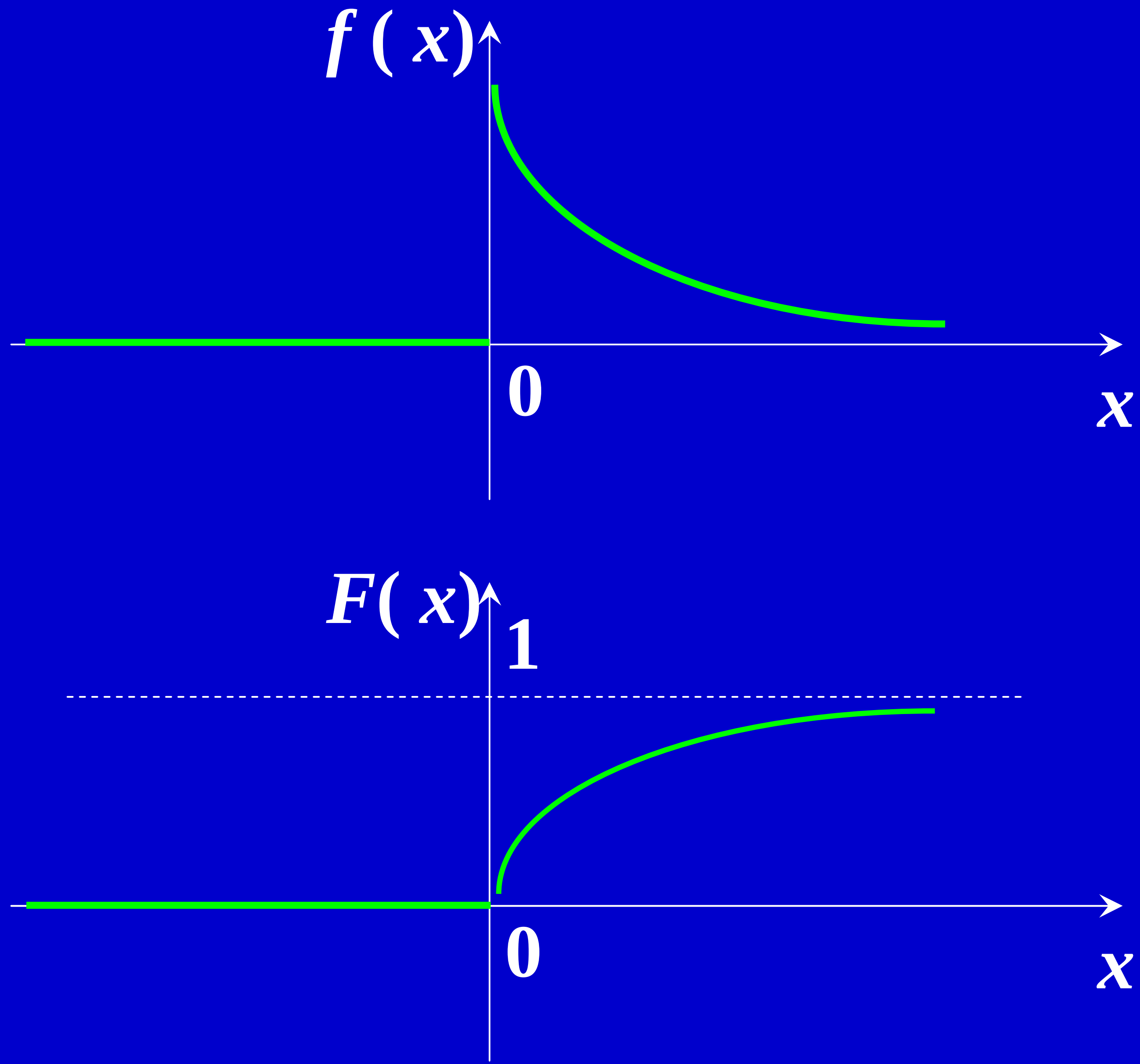
若 X 的概率密度为

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

则称 X 服从 参数为λ的指数分布

记作 $X \sim E(\lambda)$

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



f (x)
0
x
F(x)
1
0
x

对于任意的 $0 < a < b$,

$$P\{a < X < b\} = \int_a^b \lambda e^{-\lambda x}\,\mathrm{d}x$$
$$= F(b) - F(a)$$
$$= e^{-\lambda a} - e^{-\lambda b}$$

应用场合 用指数分布描述的实例有：

随机服务系统中的服务时间

电话问题中的通话时间

无线电元件的寿命

动物的寿命

指数分布常作为各种"寿命"分布的近似

指数分布的“无记忆性”

若 $X \sim E(\lambda)$,则

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

事实上

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

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

所以，又把指数分布称为“永远年轻”的分布

例4 假定一大型设备在任何长为 t 的时间内发生故障的次数 $N(t)$ 服从参数为 λt 的Poisson分布，

(1) 求相继两次故障的时间间隔 T 的概率分布

(2) 求设备已经无故障运行８小时的情况下，再无故障运行 10 小时的概率.

解 (1) $F_T(t) = P\{T \le t\}$

$$= \begin{cases} 0, & t \le 0 \\ 1 - P\{T > t\}, & t > 0 \end{cases}$$

$$P\{T > t\} = P\{N(t) = 0\}$$

$$= \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t}$$

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

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

即 $T\sim E(\lambda)$

(2) 由指数分布的“无记忆性”

$$P\{T>18\,|\,T>8\}=P\{T>8+10\,|\,T>8\}$$

$$=P\{T>10\}=e^{-10\lambda}$$

3、 正态分布

若X 的概率密度为

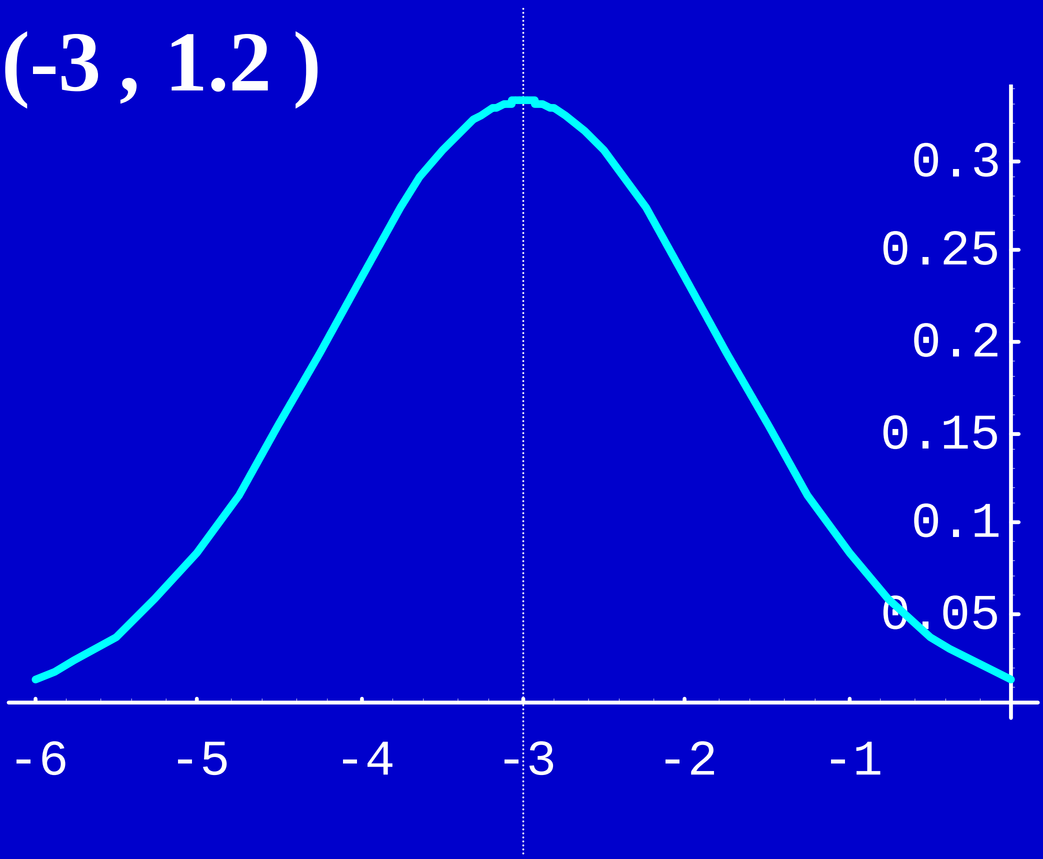
$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < +\infty$$

μ,σ 为常数，$\sigma > 0$

则称 X 服从参数为 $\mu,\ \sigma^2$ 的正态分布

记作 $X \sim N(\mu,\ \sigma^2)$

$N(-3, 1.2)$



$$\mu = -3$$

$f(x)$ 的性质：

- 图形关于直线 $x=\mu$ 对称：$f(\mu+x)=f(\mu-x)$

  在 $x=\mu$ 时, $f(x)$ 取得最大值

  $$\frac{1}{\sqrt{2\pi}\sigma}$$

  在 $x=\mu\pm\sigma$ 时, 曲线 $y=f(x)$ 在对应的点处有拐点

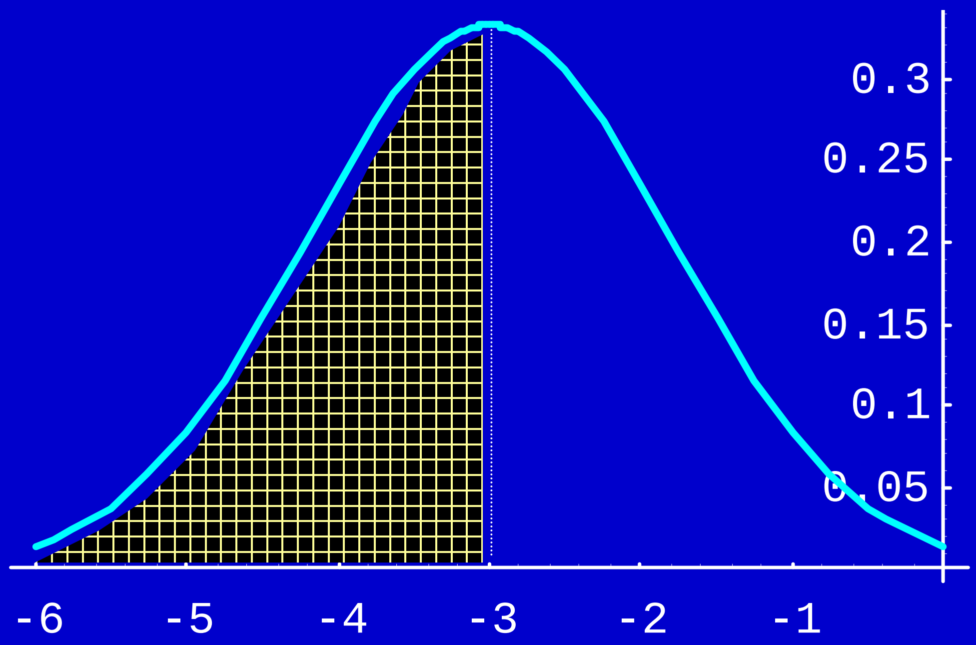
  曲线 $y=f(x)$ 以x轴为渐近线

  曲线 $y=f(x)$ 的图形呈单峰状

$$P\{X \le \mu\} = F(\mu)$$

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

$$= \frac{1}{2}$$



- $f(x)$ 的两个参数：

μ— 位置参数

即固定 σ, 对于不同的 μ, 对应的 $f(x)$
的形状不变化，只是位置不同

σ— 形状参数

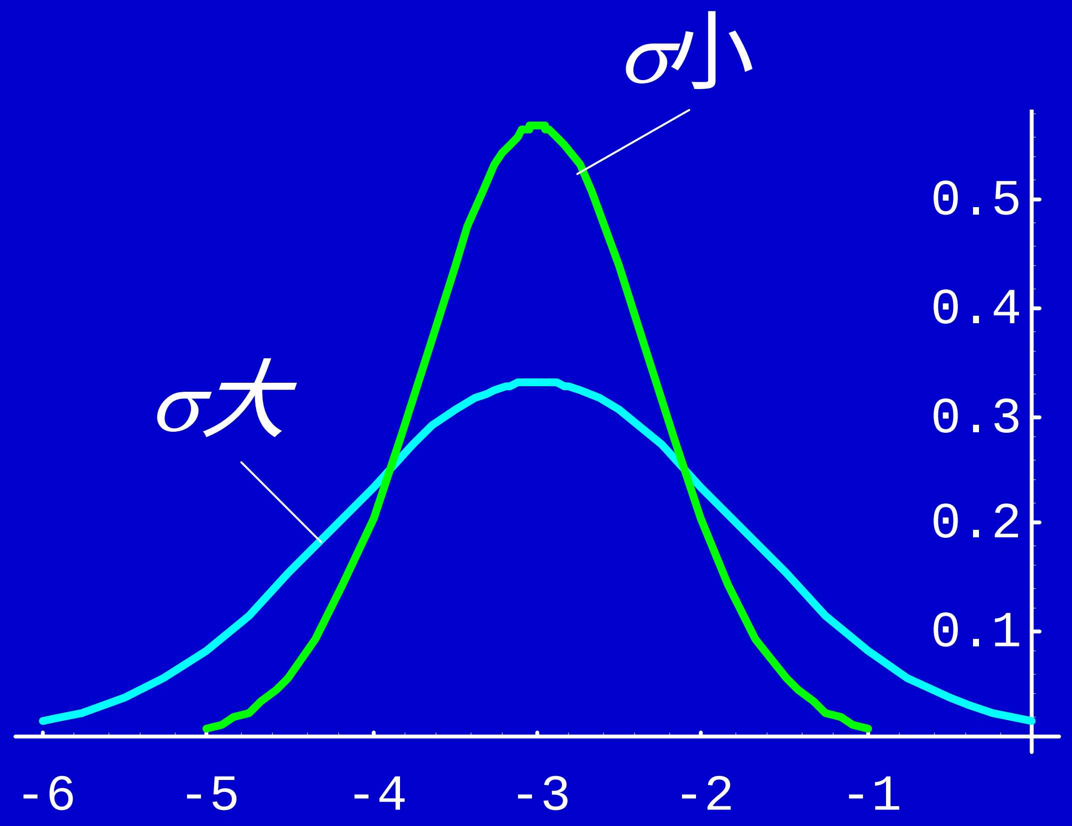
固定 μ，对于不同的σ，$f(x)$ 的形状不同.

若 $\sigma_1 < \sigma_2$ 则 $\frac{1}{\sqrt{2\pi}\sigma_1} > \frac{1}{\sqrt{2\pi}\sigma_2}$ 前者取 μ

附近值的概率更大. $x = \mu \pm \sigma_1$ 所对应的拐点

比 $x = \mu \pm \sigma_2$ 所对应的拐点更靠近直线 $x = \mu$

```
Show[fn1,fn3]
```



应用场合

若随机变量 X 受到众多相互独立的随机因素的影响，而每一个别因素的影响都是微小的，且这些影响可以叠加，则 X 服从正态分布.

可用正态变量描述的实例非常之多：

各种测量的误差；　人的生理特征；

工厂产品的尺寸；　农作物的收获量；

海洋波浪的高度；　金属线的抗拉强度；

热噪声电流强度；　学生们的考试成绩；

……　……

一种重要的正态分布：$N(0,1)$ — 标准正态分布

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

$\varphi(x)$ 是偶函数，其图形关于纵轴对称

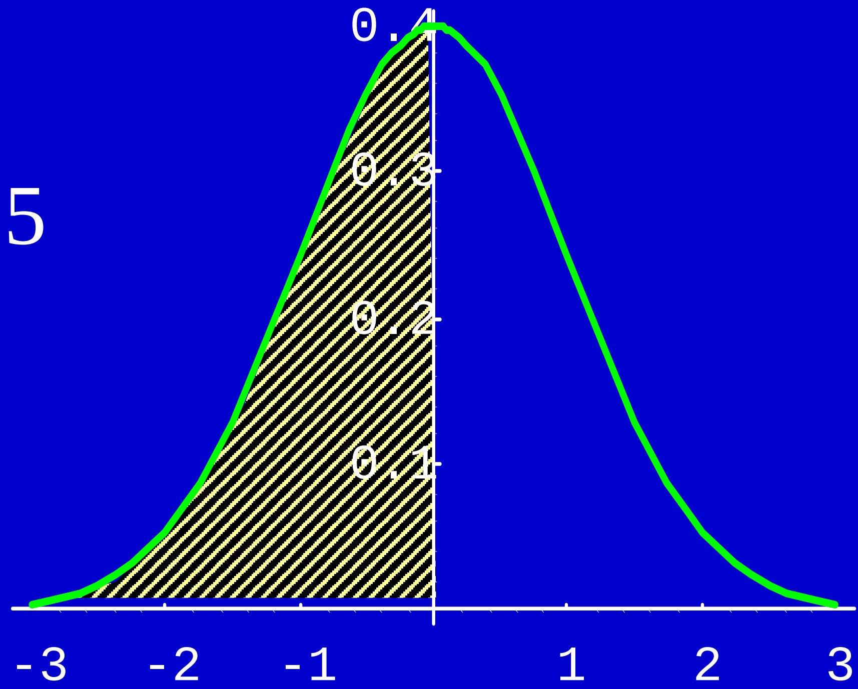
它的分布函数记为 $\Phi(x)$，其值有专门的表可查

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

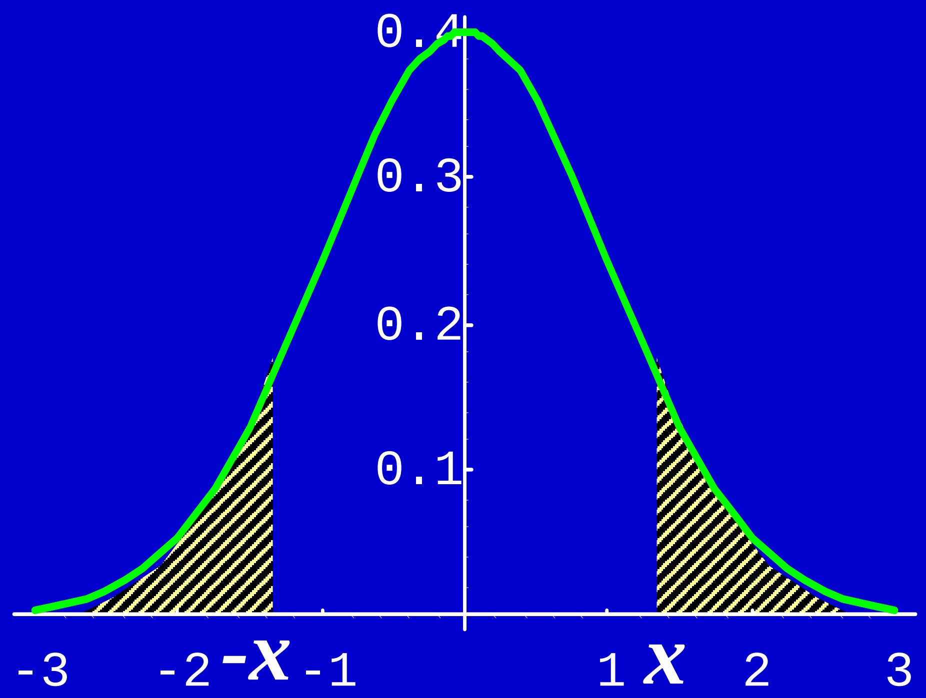
$$\Phi(0)=0.5 \qquad \Phi(-x)=1-\Phi(x)$$

$$P\{|X|<a\}=2\Phi(a)-1$$

$$\Phi(0) = 0.5$$



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



$$P(|X| < a) = 2\Phi(a) - 1$$

对一般的正态分布：$X \sim N(\mu, \sigma^2)$

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

作变量代换 $s = \dfrac{t-\mu}{\sigma}$ $\longrightarrow$ $F(x) = \Phi\left(\dfrac{x-\mu}{\sigma}\right)$

$$\longrightarrow \quad P\{a < X < b\} = F(b) - F(a)$$

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

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

$$= 1 - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

例5 设 $X \sim N(1,4)$，求 $P\{0 \le X \le 1.6\}$

解

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

$$= \Phi(0.3) - \Phi(-0.5)$$

$$= \Phi(0.3) - [1 - \Phi(0.5)]$$

P334 附表2

$$= 0.6179 - [1 - 0.6915]$$

$$= 0.3094$$

例6 已知 $X \sim N(2, \sigma^2)$ 且 $P\{\, 2 < X < 4 \,\} = 0.3$,

求 $P\{\, X < 0 \,\}$.

解一 $$P\{X < 0\} = \Phi\left(\frac{0-2}{\sigma}\right) = 1 - \Phi\left(\frac{2}{\sigma}\right)$$

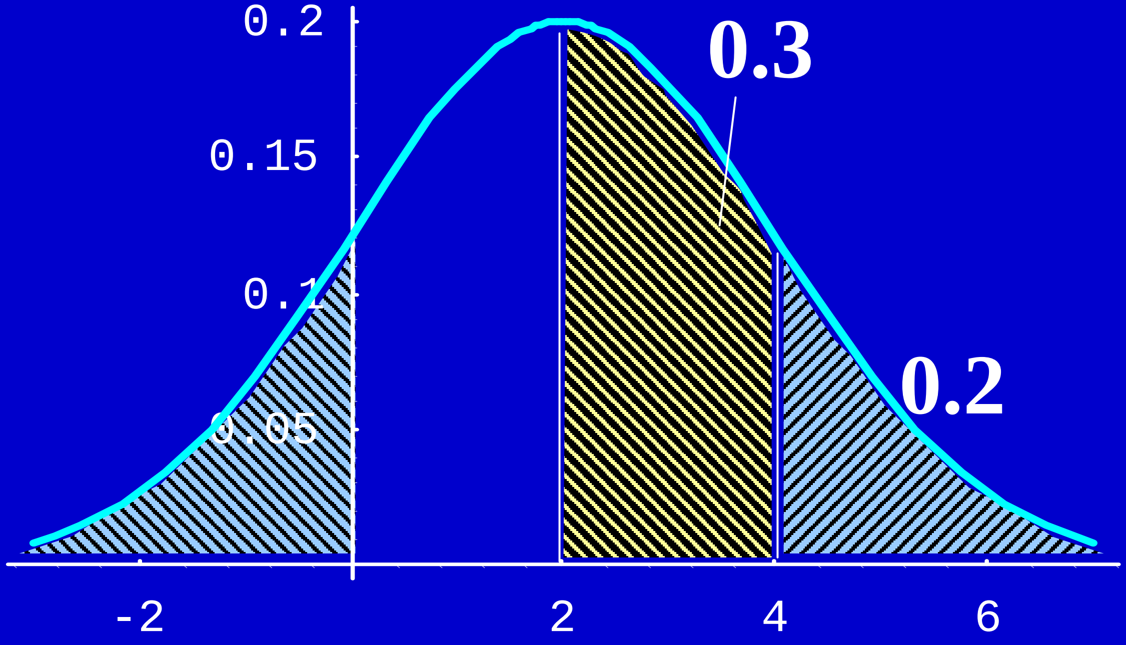
$$P\{2 < X < 4\} = \Phi\left(\frac{4-2}{\sigma}\right) - \Phi\left(\frac{2-2}{\sigma}\right)$$

$$= \Phi\left(\frac{2}{\sigma}\right) - \Phi(0) = 0.3$$

$$\longrightarrow \quad \Phi\left(\frac{2}{\sigma}\right) = 0.8$$

$$\longrightarrow \quad P\{X < 0\} = 0.2$$

解二 图解法



由图

$$P\{X < 0\} = 0.2$$

例 3σ 原理

设 $X \sim N(\mu, \sigma^2)$, 求 $P(|X-\mu|<3\sigma)$

解 $P\{|X-\mu|<3\sigma\} = P\{\mu-3\sigma < X < \mu+3\sigma\}$

$$= \Phi\left(\frac{\mu+3\sigma-\mu}{\sigma}\right) - \Phi\left(\frac{\mu-3\sigma-\mu}{\sigma}\right)$$

$$= \Phi(3) - \Phi(-3)$$

$$= 2\Phi(3) - 1 \quad = 2\times 0.9987 - 1 = 0.9974$$

在一次试验中, X 落入区间(μ - 3σ, μ +3σ) 的概率为 0.9974, 而超出此区间的可能性很小

由3σ 原理知,

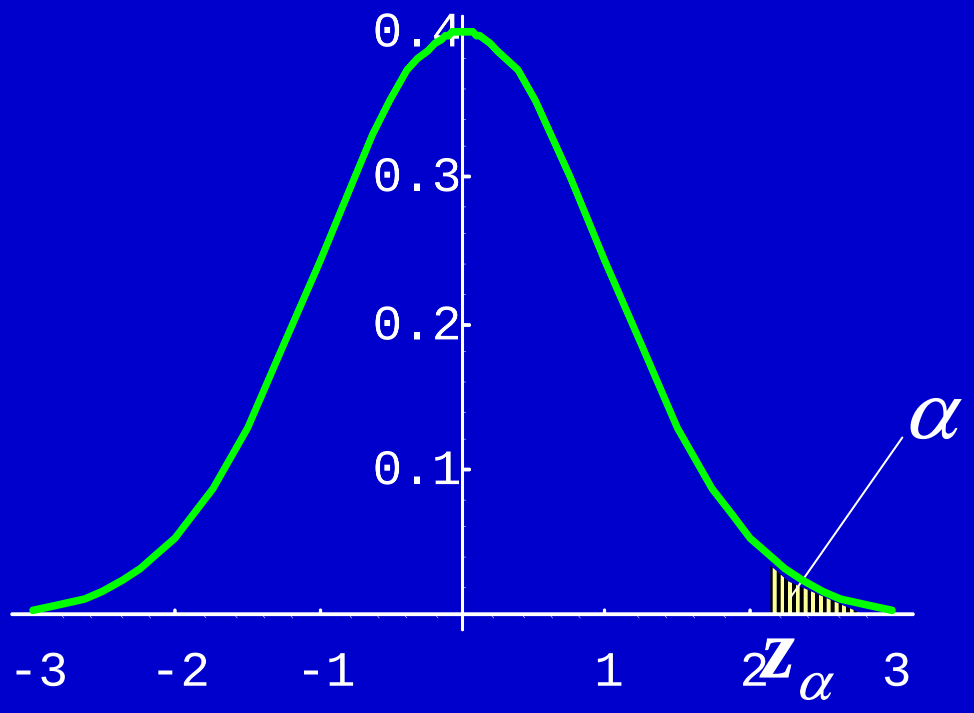
当 $a < -3$ 时 $\Phi(a) \approx 0$, $b > 3$ 时 $\Phi(b) \approx 1$

标准正态分布的上 α 分位数 z_α

设 $X \sim N(0,1)$, $0 < \alpha < 1$, 称满足

$$P\{X > z_\alpha\} = \alpha$$

的点 z_α 为X 的上α 分位数



常用的几个数据

$$z_{0.05} = 1.645$$

$$z_{0.025} = 1.96$$

例7 设测量的误差 $X \sim N(7.5,100)$(单位：米),问要进行多少次独立测量，才能使至少有一次误差的绝对值不超过10米的概率大于0.9？

解
$$P\{|X| \le 10\} = \Phi\left(\frac{10-7.5}{10}\right) - \Phi\left(\frac{-10-7.5}{10}\right)$$
$$= \Phi(0.25) - \Phi(-1.75)$$
$$= \Phi(0.25) - [1 - \Phi(1.75)]$$
$$= 0.5586$$

设 A 表示进行 n 次独立测量至少有一次误差的绝对值不超过10米

$P(A) = 1-(1-0.5586)^n > 0.9 \longrightarrow n > 3$

所以至少要进行 4 次独立测量才能满足要求.