哈爾濱工業大學

第3章　随机变量及其分布

第16讲　正态分布

正态分布(Normal)

- 定义 若随机变量X的概率密度为

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

μ,σ为常数，且$\sigma>0$, 称X 服从参数为μ,σ的正态分布或高斯(Gauss)分布，也称X为正态变量,

记作$X \sim N(\mu,\sigma^2)$.

可以验证：$f(x)\geq 0$,

概率密度的性质

下证：$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1.$　　令$t = \dfrac{x-\mu}{\sigma}$

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

$$\boxed{\text{记 } I = \int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{t^2}{2}} \mathrm{d}t} \qquad I^2 = \left(\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{t^2}{2}} \mathrm{d}t\right)\left(\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{u^2}{2}} \mathrm{d}u\right)$$

$$I^2 = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{t^2+u^2}{2}} \mathrm{d}t\mathrm{d}u \qquad \text{令}\begin{cases} t = r\cos\theta, \\ u = r\sin\theta. \end{cases}$$

$$= \int_0^{2\pi} \mathrm{d}\theta \int_0^{+\infty} r\mathrm{e}^{-\frac{r^2}{2}} \mathrm{d}r = 2\pi. \Rightarrow I = \sqrt{2\pi} \Rightarrow \int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1.$$

正态分布密度曲线特征

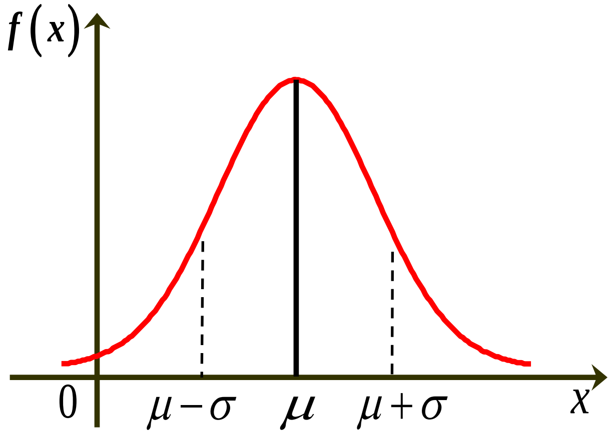
(1) 关于 $x=\mu$ 对称;

(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, $f(x) \to 0$;

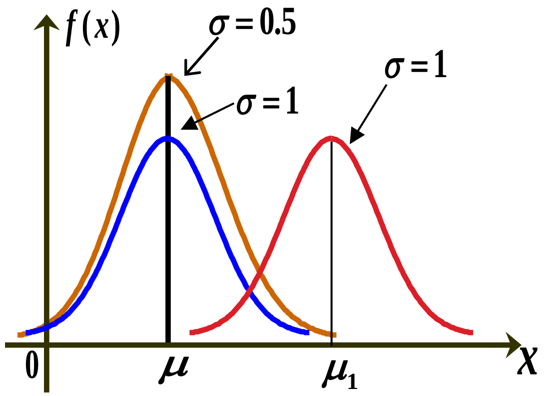
(4) 曲线在 $x=\mu\pm\sigma$ 处有拐点;



两个参数的含义

(5) μ决定对称轴的位置.

当固定σ 值，改变 μ 值时，$f(x)$的形状不变，只是沿着x轴平移；



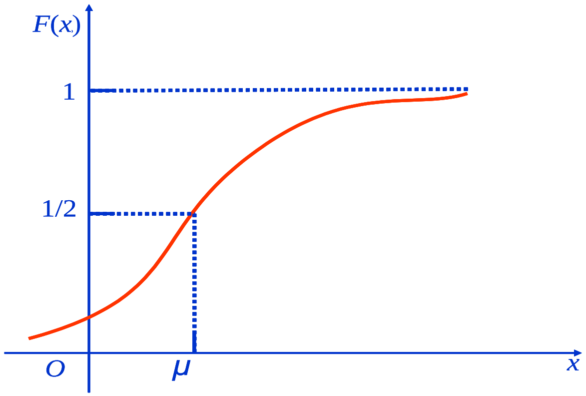
(6) σ决定离散程度.

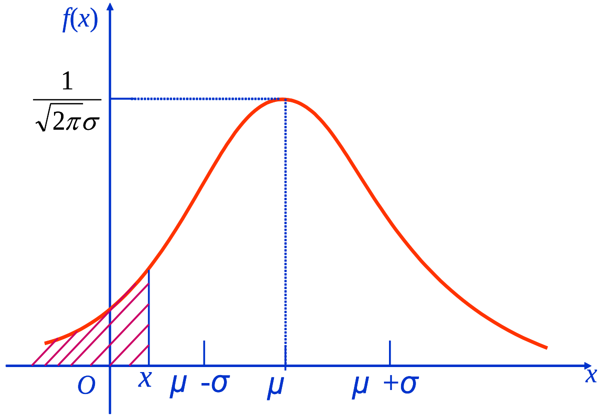
当固定μ值，改变σ 值时，$f(x)$的对称轴不变，但形状改变. σ 越大，图形越矮越胖，σ 越小，图形越高越瘦.

正态变量X的分布函数

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

$$-\infty < x < +\infty$$



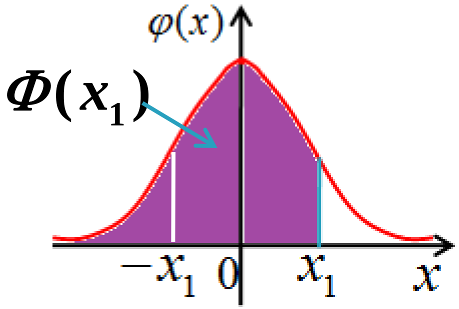


$$F(\mu-x)+F(\mu+x)=1$$

标准正态分布定义

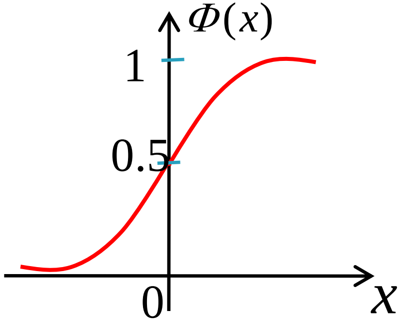
- 若$X \sim N(0,1)$，称X服从标准正态分布.
- 概率密度为

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



- 分布函数为

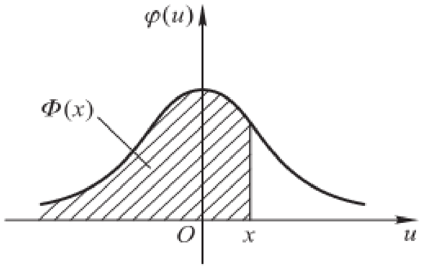
$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{t^2}{2}}dt\quad(-\infty<x<+\infty)$$



- 性质

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

附表 2　标准正态分布函数值表



$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{u^2}{2}du} = \int_{-\infty}^{x}\varphi(u)\,du$$

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.500 0	0.504 0	0.508 0	0.512 0	0.516 0	0.519 9	0.523 9	0.527 9	0.531 9	0.535 9
0.1	0.539 8	0.543 8	0.547 8	0.551 7	0.555 7	0.559 6	0.563 6	0.567 5	0.571 4	0.575 3
0.2	0.579 3	0.583 2	0.587 1	0.591 0	0.594 8	0.598 7	0.602 6	0.606 4	0.610 3	0.614 1
0.3	0.617 9	0.621 7	0.625 5	0.629 3	0.633 1	0.636 8	0.640 6	0.644 3	0.648 0	0.651 7
0.4	0.655 4	0.659 1	0.662 8	0.666 4	0.670 0	0.673 6	0.677 2	0.680 8	0.684 4	0.687 9
0.5	0.691 5	0.695 0	0.698 5	0.701 9	0.705 4	0.708 8	0.712 3	0.715 7	0.719 0	0.722 4
0.6	0.725 7	0.729 1	0.732 4	0.735 7	0.738 9	0.742 2	0.745 4	0.748 6	0.751 7	0.754 9
0.7	0.758 0	0.761 1	0.764 2	0.767 3	0.770 3	0.773 4	0.776 4	0.779 4	0.782 3	0.785 2

$$\Phi(0.76)=0.7764$$

一般的正态分布$N(\mu, \sigma^2)$的分布函数$F(x)$与标准正态分布的分布函数$\Phi(x)$间的关系为

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

证明

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}}\,dt \qquad \text{令} v = \frac{t-\mu}{\sigma}$$

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

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

若$X \sim N(\mu,\sigma^2)$,$\forall a<b$有

$$P(a<X<b)=F(b)-F(a)=\Phi(\frac{b-\mu}{\sigma})-\Phi(\frac{a-\mu}{\sigma}).$$

例1 某种电池的寿命(h) $X \sim N(160,20^2)$,分别求寿命小于120h和寿命大于200h的概率.

解 $$P(X<120)=F(120)=\Phi(\frac{120-160}{20})=\Phi(-2).$$

$$=1-\Phi(2)\overset{查表}{==}1-0.9772=0.0228.$$

$$P(X>200)=1-F(200)=1-\Phi(\frac{200-160}{20})$$

$$=1-\Phi(2)\overset{查表}{==}1-0.9772=0.0228.$$

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

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

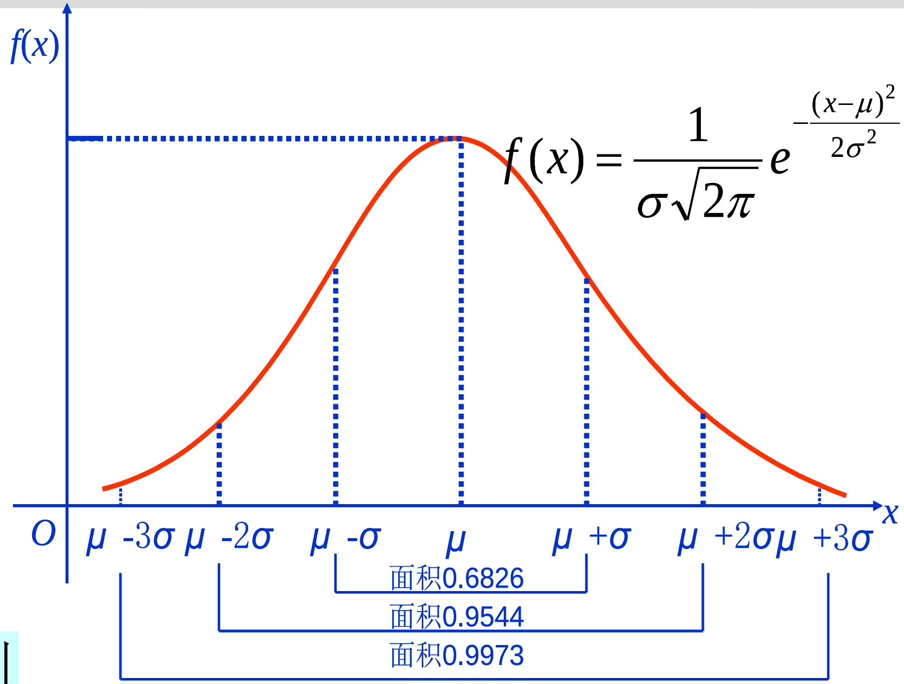
$$= \Phi\left(\frac{\mu+\sigma-\mu}{\sigma}\right) - \Phi\left(\frac{\mu-\sigma-\mu}{\sigma}\right) = \Phi(1) - \Phi(-1)$$

$$= \Phi(1) - [1 - \Phi(1)] = 2\,\Phi(1) - 1 = 0.6826.$$

同理有 $P\{|X-\mu|<2\sigma\} = 0.9544,$

$$P\{|X-\mu|<3\sigma\} = 0.9973.$$

X几乎总是落在 $(\mu-3\sigma, \mu+3\sigma)$ 之中.



f(x)
$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
O
x
μ -3σ
μ -2σ
μ -σ
μ
μ +σ
μ +2σ
μ +3σ
面积0.6826
面积0.9544
面积0.9973

3σ准则

正态分布的应用

在实际中，许多随机变量都服从或近似地服从这种“中间大，两头小”的正态分布.

例如，

✓测量一个零件长度的测量误差，

✓向一中心点射击的横向偏差或纵向偏差，

✓电子管中的噪声电流或电压，

✓飞机材料的疲劳应力等等.

练习

1. 公共汽车车门的高度是按男子与车门顶头碰头机会在0.01以下来设计的. 设男子身高$X\sim N(170,6^2)$,问(1)车门高度应如何确定?(2)若车门高为182cm, 求100个成年男子与车门顶头碰头人数不多于2的概率.

解 (1)设车门高度为h cm, 按设计要求 $P(X\geq h)\leq 0.01$

或 $P(X< h)\geq 0.99$,

下面我们来求满足$P(X< h)\geq 0.99$

的最小的h.

因为$X \sim N(170, 6^2)$，$\dfrac{X-170}{6} \sim N(0,1)$

故 $P(X < h) = \Phi\left(\dfrac{h-170}{6}\right) \geq 0.99$

查表得 $\Phi(2.33) = 0.9901 > 0.99$

$$\frac{h-170}{6} = 2.33$$

即 h=170+13. 98≈184

设计车门高度为184厘米时，可使男子与车门碰头机会不超过0.01.

练习

1. 公共汽车车门的高度是按男子与车门顶头碰头机会在0.01以下来设计的. 设男子身高$X\sim N(170,6^2)$,问(2) 若车门高为182cm, 求100个成年男子与车门顶头碰头人数不多于2的概率.

解 (2) 设任一男子身高超过182cm的概率为 p.

则$p=P\{X>182\}=1-F(182)=1-\Phi\left(\dfrac{182-170}{6}\right)=1-\Phi(2)=0.0228.$

设 Y 为100个男子中身高超过182cm的人数,

则$Y\sim B(100,0.0228)$, $\quad P\{Y=k\}=C_{100}^{k}\times 0.0228^{k}\times 0.9772^{100-k}$,

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

练习

1. 公共汽车车门的高度是按男子与车门顶头碰头机会在0.01以下来设计的. 设男子身高$X\sim N(170,6^2)$,问(2)若车门高为182cm, 求100个成年男子与车门顶头碰头人数不多于2的概率.

解 所求概率为 $P\{Y\le 2\}=1-P\{Y>2\}$

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

$$P\{Y\le 2\}=1-P\{Y>2\}\approx 1-\sum_{k=3}^{100}\frac{2.28^k}{k!}e^{-2.28}\approx 1-\sum_{k=3}^{+\infty}\frac{2.28^k}{k!}e^{-2.28}$$

$$\approx 0.6013$$

谢　　谢！