第五章 正态分布与自然指数分布族

正态分布的重要性

正态分布是概率论中最重要的分布，这可以由以下情形加以说明：

(1)正态分布是自然界及工程技术中最常见的分布之一，大量的随机现象都是服从或近似服从正态分布的.

(2)正态分布有许多良好的性质，这些性质是其它许多分布所不具备的.

(3)正态分布可以作为许多分布的近似分布.

5.1 正态分布及其密度函数和分布函数

标准正态分布N(0,1)

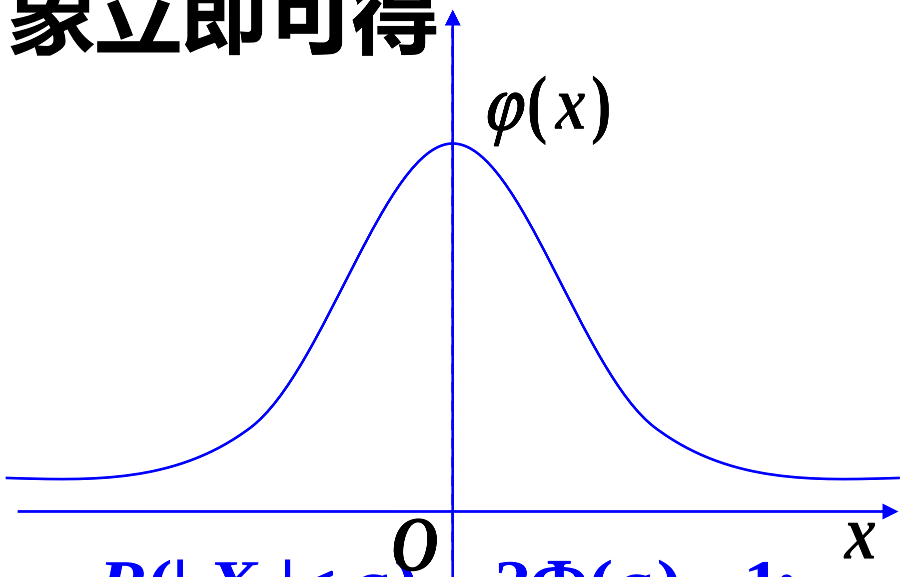
定义5.1 若随机变量**X**的密度函数为：

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

则称**X**服从标准正态分布，记为 $X\sim N(0,1)$

其分布函数为:$\Phi(x)=P(X\le x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$

标准正态分布的密度函数为**偶函数**,由其图象立即可得



$i)\ \ \Phi(0)=\dfrac{1}{2};$

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

$or\ \ \Phi(-x)+\Phi(x)=1;$

$iii)\ \ if\ \ X \sim N(0,1),\ \ then$

$P(|X|<a)=2\Phi(a)-1;\qquad P(|X|\geq a)=2\big(1-\Phi(a)\big).$

关于查表

书后附表给出了在 $x\geq 0$ 时标准正态分布函数的数值，当 $x<0$ 时应用公式 $\Phi(x)=1-\Phi(-x)$ 来计算 $\Phi(x)=P(X\leq x)$ 的值。另外还需注意的是

$\Phi(x)\approx 0(x\leq -4),\ \ \Phi(x)\approx 1(x\geq 4).$

一般正态分布

定义 若随机变量X 的密度函数为

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

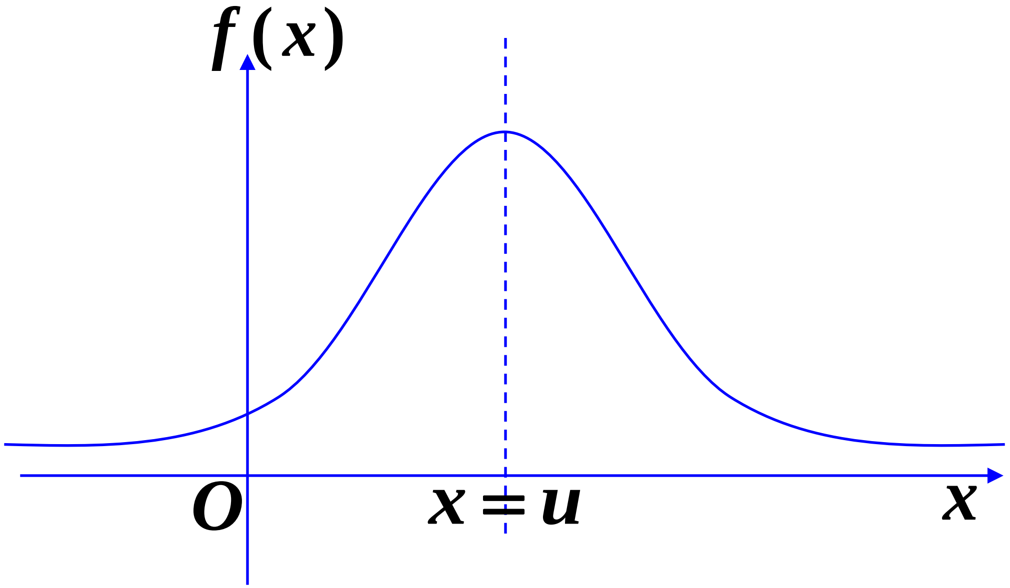
则称X 服从参数为 u,σ^2 的正态分布，

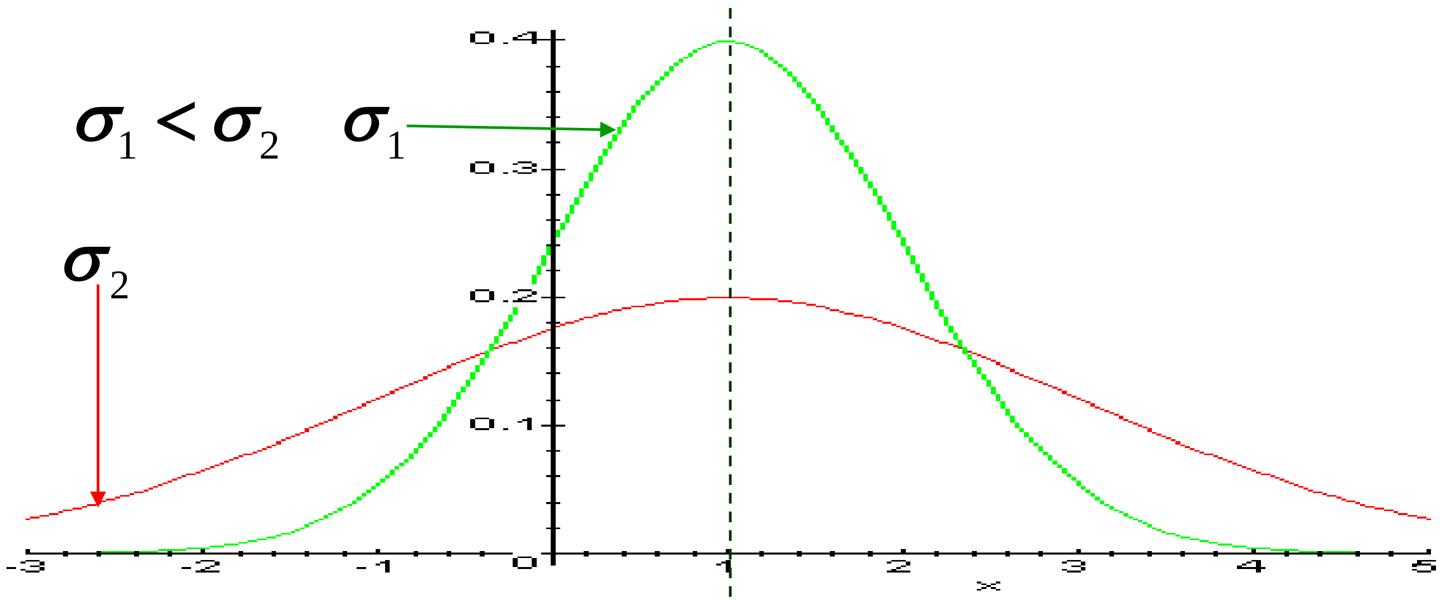
记为 $X \sim N(u,\sigma^2)$,其中的 $u\in R,\sigma>0$ 为常数。

正态分布

$\boxed{X \sim N(u,\sigma^2)}$

的密度函数的图象为





显然可见，正态分布$N(u,\sigma^2)$的密度函数关于x=u对称。

当u给定之后,σ越大，图象越矮越胖；σ越小，图象越高越瘦。

标准正态分布与一般正态分布有如下关系：

定理5.1：若 $X \sim N(\mu,\sigma^2)$，分布函数为 $F(x)$ 则

$$(1)\ F(x)=\Phi\left(\frac{x-\mu}{\sigma}\right),-\infty<x<+\infty,$$

（2）$\forall a,b\in R, a<b,$

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

$$P(X\le b)=\Phi\left(\frac{b-\mu}{\sigma}\right),\quad P(X>a)=1-\Phi\left(\frac{a-\mu}{\sigma}\right).$$

$$(3)\ Y=\frac{X-\mu}{\sigma}\sim N(0,1).$$

证：（1）$F(x)=\int_{-\infty}^{x} f(t)dt=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt$

$$=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}d\left(\frac{t-u}{\sigma}\right)$$

$$=\int_{-\infty}^{\frac{x-u}{\sigma}}\frac{1}{\sqrt{2\pi}}e^{-\frac{s^2}{2}}ds \quad （令\frac{t-\mu}{\sigma}=s）$$

$$=\Phi(\frac{x-\mu}{\sigma}).$$

(3)利用公式：若y=g(x)严格单调，其反函数$x=g^{-1}(y)$具有连续导数，则Y=g(X)的密度函数

$$f_Y(y)=\begin{cases} f_X(g^{-1}(y))\bullet|[g^{-1}(y)]'|, & y\in R(Y), \\ 0 & y\notin R(Y). \end{cases}$$

特别地，$Y=aX+b$的密度为

$$f_Y(y)=\begin{cases} \frac{1}{|a|}f_X(\frac{y-b}{a}), & y\in R(Y), \\ 0 & y\notin R(Y). \end{cases}$$

例：若 $X \sim N(1,16)$，求如下概率：$P(X \le 0.6)$, $P(|X-1| \le 2), P(|X-18| > 2), P(-17 < X-1 < 6.58)$.

解：由于 $X \sim N(\mu,\sigma^2)$, $\mu=1,\sigma=4$, 所以

$$P(X \le 0.6)\left(=\Phi(\frac{x-\mu}{\sigma})\right)=\Phi(\frac{0.6-1}{4})=\Phi(-0.1)$$

$$=1-\Phi(0.1)=1-0.5398=0.4602;$$

$$P(|X-1| \le 2)=P(-1 \le X \le 3)=\Phi(\frac{3-1}{4})-\Phi(\frac{-1-1}{4})$$

$$=\Phi(0.5)-\Phi(-0.5)=\Phi(0.5)-\left(1-\Phi(0.5)\right)$$

$$=2\Phi(0.5)-1=2\times 0.6915-1=0.3830;$$

$$P(|X-18|>2)=P(X>20 \quad or \quad X<16)=P(X>20)+P(X<16)$$

$$=1-P(X\le 20)+P(X\le 16)$$

$$=1-\Phi(\frac{20-1}{4})+\Phi(\frac{16-1}{4})$$

$$=1-\Phi(4.75)+\Phi(3.75)=1-1+0.9999=0.9999;$$

$$P(-17<X-1<6.58)=P(-16<X<7.58)$$

$$=\Phi(\frac{7.58-1}{4})-\Phi(\frac{-16-1}{4})=\Phi(1.645)-\Phi(-4.25)$$

$$=0.95-0=0.95.$$

例 5.3　设 $X: N(u,\sigma^2)$, 求　$P(|X-u|\le 3\sigma)$。

解：$P(|X-u|\le 3\sigma)=P(-3\sigma\le X-u\le 3\sigma)$

$$=\Phi(\frac{u+3\sigma-u}{\sigma})-\Phi(\frac{u-3\sigma-u}{\sigma})=2\Phi(3)-1=0.9974.$$

注：例5.3说，X落在$(u-3\sigma,u+3\sigma)$内的概率为0.9974，即X几乎总是落在（$u-3\sigma,u+3\sigma$）内。所以在应用中，常视（$u-3\sigma,u+3\sigma$）为X的实际取值区间。此称“　原则”3σ。

定义：设$X\sim N(0,1)$，则称随机变量$Y=X^2$服从自由度为1的χ^2分布，

记为$Y\sim\chi^2(1)$.书上例5.4表明$\chi^2(1)=\Gamma(\frac{1}{2},\frac{1}{2})$.