第四节　连续型随机变量及其概率密度

- 连续型随机变量及其概率密度的定义
- 概率密度的性质
- 三种重要的连续型随机变量
- 小结　布置作业

连续型随机变量X所有可能取值充满一个区间，对这种类型的随机变量，不能象离散型随机变量那样，以指定它取每个值概率的方式，去给出其概率分布，而是通过给出所谓“概率密度函数”的方式.

下面我们就来介绍对连续型随机变量的描述方法.

一、 连续型随机变量及其概率密度的定义

对于随机变量 X，如果存在非负可积函数 $f(x)$，$x \in (-\infty, +\infty)$，使得对任意实数 x， 有

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

则称 X为连续型随机变量，称 $f(x)$ 为 X 的概率密度函数，简称为概率密度.

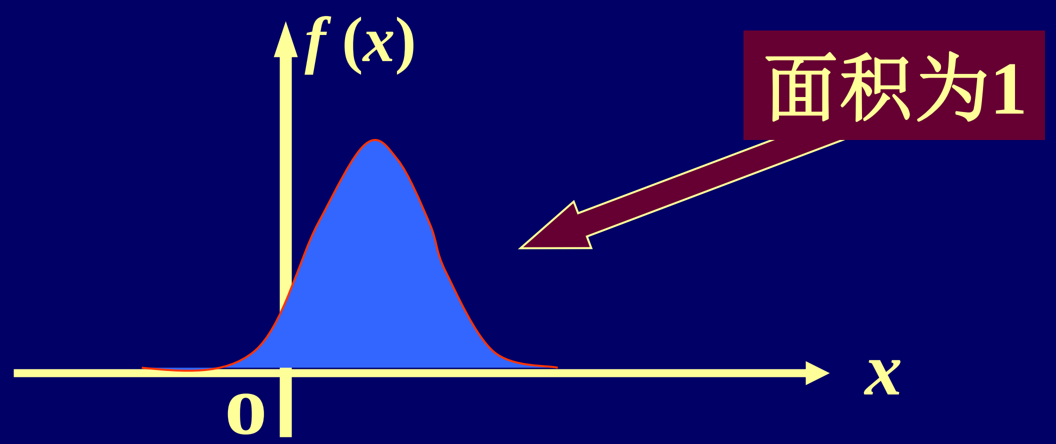
连续型随机变量的分布函数在 R上连续

二、概率密度的性质

1° $f(x) \geq 0$

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

这两条性质是判定一个函数 $f(x)$是否为某$r.v$ X 的概率密度的充要条件



3° 对于任意实数 $x_1, x_2, (x_1 < x_2)$,

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

利用概率密度可确定随机点落在某个范围内的概率

4° 若 $f(x)$ 在点 x 处连续，则有

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

对 $f(x)$的进一步理解:

若 x 是 $f(x)$ 的连续点，则

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

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

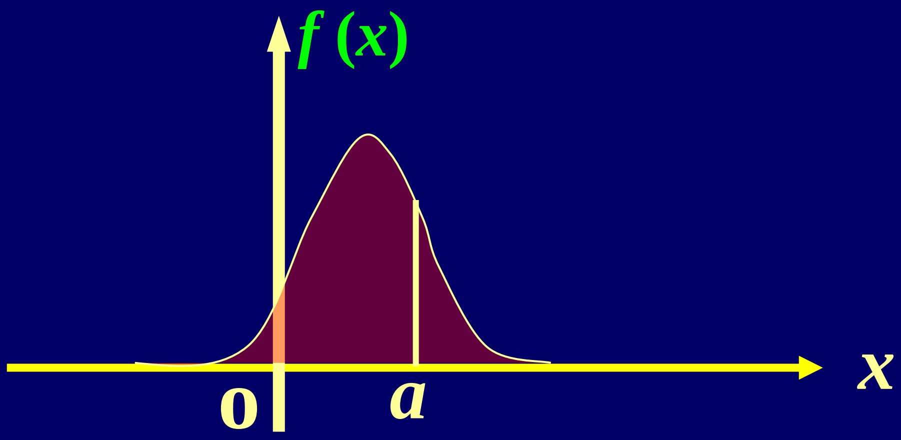
故 X的密度 $f(x)$ 在 x 这一点的值，恰好是X 落在区间$[x, x+\Delta x]$上的概率与区间长度 Δx 之比的极限. 这里，如果把概率理解为质量， $f(x)$ 相当于线密度.

若不计高阶无穷小，有

$$P\{x < X \le x+\Delta x\} = f(x)\Delta x$$

表示随机变量 X 取值于 $(x, x+\Delta x]$ 的概率近似等于 $f(x)\Delta x$.

$f(x)\Delta x$ 在连续型 $r.v$ 理论中所起的作用与 $P(X=x_k)=p_k$ 在离散型 $r.v$ 理论中所起的作用相类似.



要注意的是，密度函数 $f\,(x)$在某点处a的高度，并不反映X取值的概率. 但是，这个高度越大，则X取a附近的值的概率就越大. 也可以说，在某点密度曲线的高度反映了概率集中在该点附近的程度.

请注意:

(1) 连续型*r.v*取任一指定实数值a 的概率均为0. 即

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

这是因为

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

当 $\Delta x\to 0+$时, 得到

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

由$P(A)=0$，不能推出 $A=\varnothing$

由$P(B)=1$，不能推出 $B=S$

(2) 对连续型 r.v X，有

$$P(a\le X\le b)=P(a<X\le b)$$

$$=P(a\le X<b)$$

$$=P(a<X<b)$$

例1　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$F(x)$;

（3）求$P\left\{1<X\le\dfrac{7}{2}\right\}$

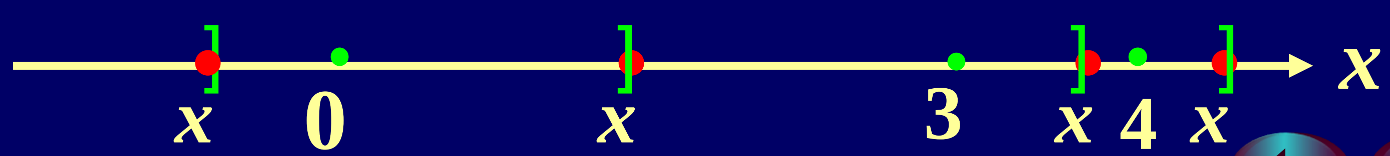
解 $f(x)=\begin{cases} kx, & 0 \le x < 3 \\ 2-\dfrac{x}{2}, & 3 \le x \le 4 \\ 0, & \text{其它} \end{cases}$

(1) 由 $\int_{-\infty}^{+\infty} f(x)dx = 1$ 得 $k=\dfrac{1}{6}$

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

(2) 分布函数

$$F(x)=\begin{cases} 0, & x<0 \\ \displaystyle\int_0^x \frac{x}{6}dx, & 0\le x<3 \\ \displaystyle\int_0^3 \frac{x}{6}dx+\int_3^x\left(2-\frac{x}{2}\right)dx, & 3\le x<4 \\ 1, & x\ge 4 \end{cases}$$



即分布函数

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

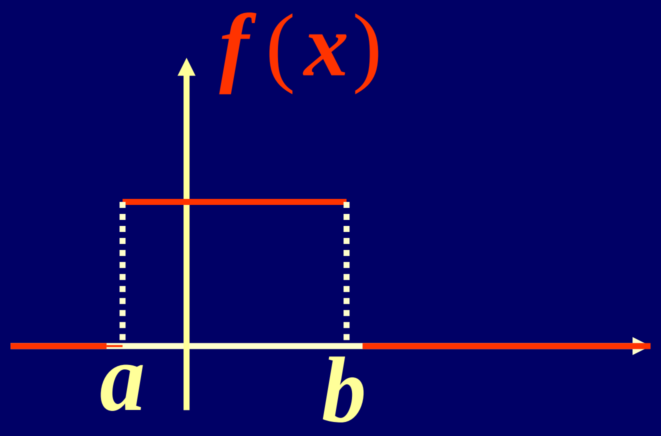
(3) $P\left\{1<X\le\dfrac{7}{2}\right\}=F\left(\dfrac{7}{2}\right)-F(1)=\dfrac{41}{48}$

三、三种重要的连续型随机变量

1. 均匀分布

若 $r.v\ X$的概率密度为:

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



则称X在区间(a, b)上服从均匀分布，记作

$$X \sim U(a, b)$$

若$X \sim U(a,b)$,

1°.对于长度l为的区间$(c,c+l),a \le c < c+l \le b$,有

$$P\{c < X \le c+l\} = \int_{c}^{c+l} \frac{1}{b-a} dx = \frac{l}{b-a}$$

2°.X的分布函数为：

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

均匀分布常见于下列情形：

如在数值计算中，由于四舍五 入，小数点后某一位小数引入的误差；

公交线路上两辆公共汽车前后通过某汽车停车站的时间，即乘客的候车时间等.

例2 某公共汽车站从上午7时起，每15分钟来一班车，即 7:00，7:15，7:30，7:45 等时刻有汽车到达此站，如果乘客到达此站时间 X 是7:00 到 7:30 之间的均匀随机变量，试求他候车时间少于5 分钟的概率.

解 以7:00为起点0，以分为单位

依题意， $X \sim U(0, 30)$

$$f(x)=\begin{cases}\dfrac{1}{30}, & 0<x<30\\ 0, & \text{其它}\end{cases}$$

从上午7时起，每15分钟来一班车，即 7:00，7:15，7:30等时刻有汽车到达汽车站，

为使候车时间X少于 5 分钟，乘客必须在 7:10 到 7:15 之间，或在7:25 到 7:30 之间到达车站.

所求概率为:

$$P\{10 < X < 15\} + P\{25 < X < 30\}$$

$$= \int_{10}^{15} \frac{1}{30} dx + \int_{25}^{30} \frac{1}{30} dx = \frac{1}{3}$$

即乘客候车时间少于5 分钟的概率是1/3.

2 . 指数分布

若 r .v X具有概率密度

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-x/\theta}, & x>0,\\ 0, & \text{其它},\end{cases}$$

其中$\theta>0$为常数, 则称 X 服从参数为θ的指数分布.

指数分布常用于可靠性统计研究中，如元件的寿命.

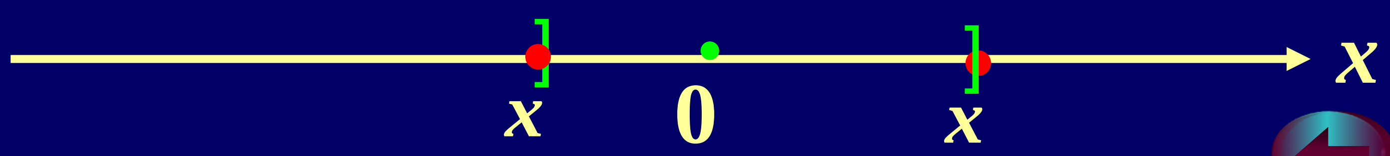
若X 服从参数为 θ 的指数分布, 则其分布函数为

$$F(x)=P\{X\le x\}=\begin{cases}1-e^{-x/\theta}, & x>0\\ 0, & \text{其它}\end{cases}$$

事实上， $F(x)=\int_{-\infty}^{x} f(t)dt$

当 $x\le 0$ 时, $F(x)=\int_{-\infty}^{x} f(t)dt=\int_{-\infty}^{x} 0dt$

当 $x>0$ 时, $F(x)=\int_{-\infty}^{x} f(t)dt=\int_{-\infty}^{0} 0dt+\int_{0}^{x} \frac{1}{\theta}e^{-\frac{t}{\theta}}dt$



3. 正态分布

若连续型 r.v X 的概率密度为

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

其中 μ 和 $\sigma(\sigma>0)$都是常数, 则称X服从参数为μ和σ的正态分布或高斯分布. 记作

$$X \sim N(\mu,\sigma^2)$$

1° $f(x)$具有下述性质：

(1) $f(x) \geq 0$;

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

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

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

$$= \frac{2}{\sqrt{2\pi}\sigma} \int_{0}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

令 $t=\dfrac{x-\mu}{\sqrt{2}\sigma}$，则有

$$\int_0^{+\infty} e^{-t^2}dt=\frac{\sqrt{\pi}}{2}$$

$$\int_{-\infty}^{+\infty} f(x)dx=\frac{2}{\sqrt{\pi}}\int_0^{+\infty} e^{-t^2}dt=\frac{2}{\sqrt{\pi}}\cdot\frac{\sqrt{\pi}}{2}=1$$

(3) 曲线 $f(x)$ 关于 μ 轴对称；

$$\Downarrow$$

$$P(\mu-h<X\le\mu)=P(\mu<X\le\mu+h)\quad(h>0)$$

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

(4) 函数 $f(x)$在 $(-\infty,\mu]$ 上单调增加, 在 $[\mu,+\infty)$ 上单调减少, 在 $x=\mu$ 取得最大值;

$$f'(x)=\frac{\mu-x}{\sqrt{2\pi}\sigma^3}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<+\infty$$

(5) $x=\mu\pm\sigma$为 $f(x)$ 的两个拐点的横坐标;

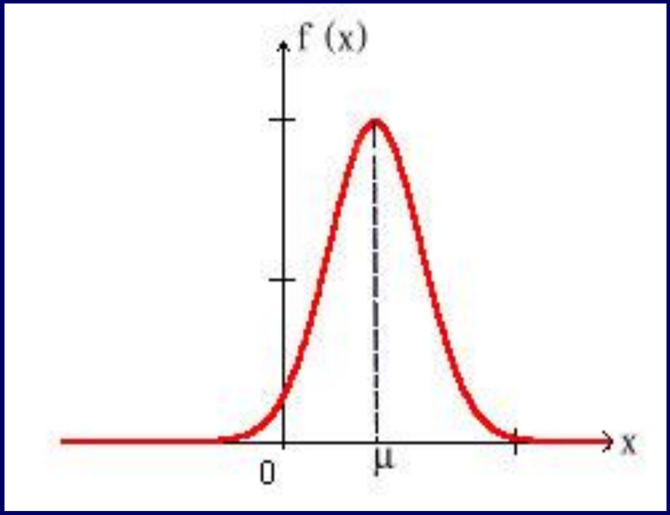
$$f''(x)=\frac{(x-\mu)^2-\sigma^2}{\sqrt{2\pi}\sigma^3}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\quad -\infty<x<+\infty$$

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

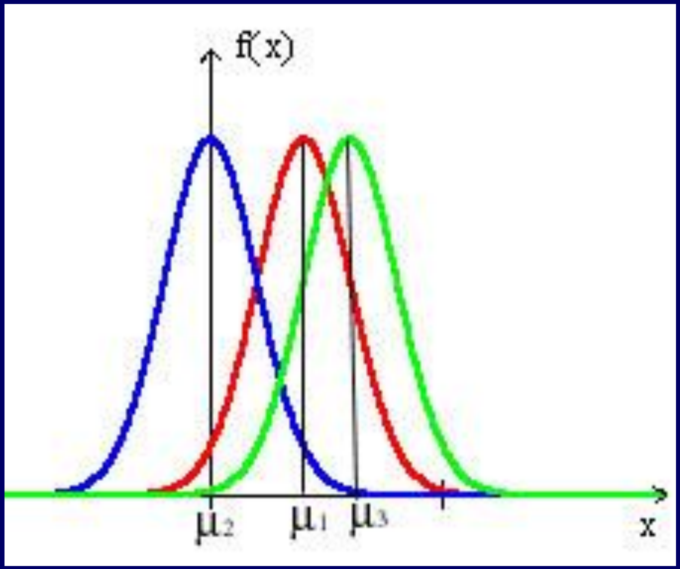
(6) $f(x)$ 以 x 轴为渐近线

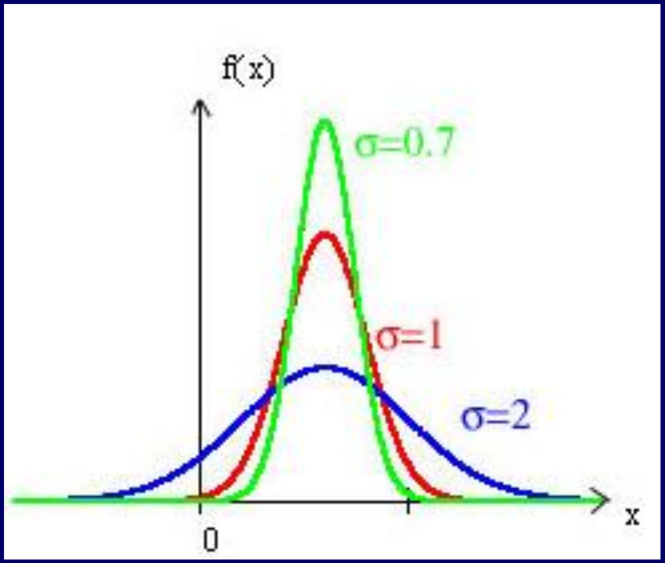
当$x\to\infty$时，$f(x)\to 0$.

根据对密度函数的分析，也可初步画出正态分布的概率密度曲线图.



正态分布 $N(\mu,\sigma^2)$ 的图形特点



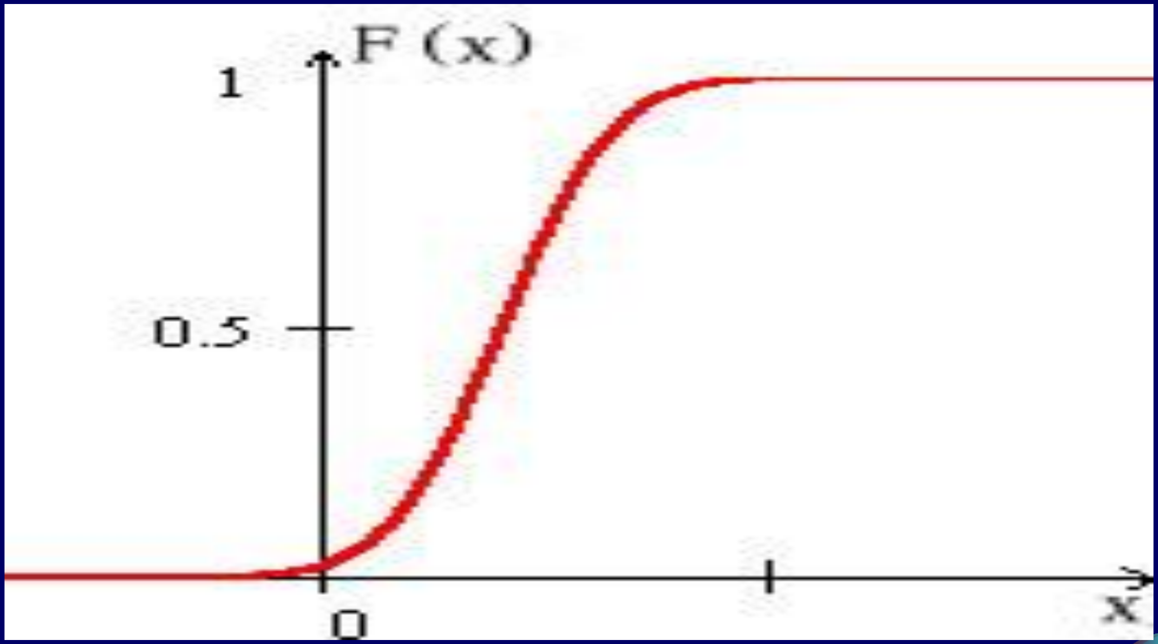


μ 决定了图形的中心位置，σ 决定了图形中峰的陡峭程度.

2° 正态分布 $N(\mu,\sigma^2)$ 的分布函数

设 $X \sim N(\mu,\sigma^2)$， X 的分布函数是

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



正态分布由它的两个参数μ和σ唯一确定， 当μ和σ不同时，是不同的正态分布。

下面我们介绍一种最重要的正态分布

标准正态分布

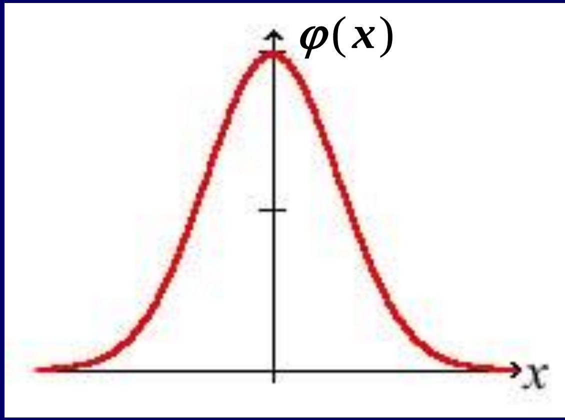
3° 标准正态分布

$\mu=0, \sigma=1$ 的正态分布称为标准正态分布.

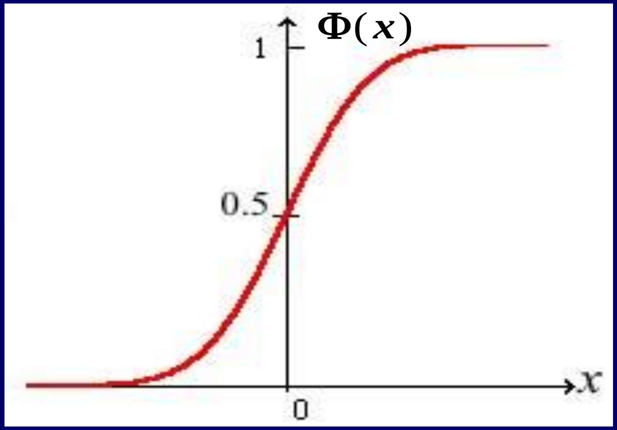
其密度函数和分布函数常用 $\varphi(x)$ 和 $\Phi(x)$ 表示:

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

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



$\varphi(x)$
x



$\Phi(x)$
1
0.5
0
x

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

(1) $\Phi(0)=\dfrac{1}{2}$ ；

$$\left(\ \Phi(0)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0}e^{-\frac{t^2}{2}}dt=\frac{1}{2}\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt=\frac{1}{2}\ \right)$$

(2) $\forall x\in R,\Phi(-x)=1-\Phi(x)$ ；

事实上，$\Phi(-x)=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{-x}e^{-\frac{t^2}{2}}dt$

$$\underline{\underline{u=-t}}\ \frac{1}{\sqrt{2\pi}}\int_{x}^{\infty} e^{-\frac{u^2}{2}}\,du$$

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

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

标准正态分布的重要性在于，任何一个一般的正态分布都可以通过线性变换转化为标准正态分布.

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

证　Z 的分布函数为

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

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

令　$u=\dfrac{t-\mu}{\sigma}$　，　则有

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

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

于是

$$X\sim N(\mu,\sigma^2)$$

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

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

根据定理1, 只要将标准正态分布的分布函数制成表，就可以解决一般正态分布的概率计算问题.

4° 正态分布表

书末附有标准正态分布函数数值表，有了它，可以解决一般正态分布的概率计算查表.

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

表中给的是 $x>0$ 时, $\Phi(x)$的值.

当 $x<0$ 时， $\Phi(-x)=1-\Phi(x)$

若 $X \sim N(0,1)$,

$$P(a < X < b) = \Phi(b) - \Phi(a)$$

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

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

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

5° 3σ 准则

由标准正态分布的查表计算可以求得，

当 $X \sim N(0, 1)$ 时，

$$P(|X| \leq 1) = 2\Phi(1) - 1 = 0.6826$$

$$P(|X| \leq 2) = 2\Phi(2) - 1 = 0.9544$$

$$P(|X| \leq 3) = 2\Phi(3) - 1 = 0.9974$$

这说明，X的取值几乎全部集中在$[-3,3]$区间内，超出这个范围的可能性仅占不到0.3%.

将上述结论推广到一般的正态分布,

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

$$P(|Y-\mu| \le \sigma)=0.6826$$

$$P(|Y-\mu| \le 2\sigma)=0.9544$$

$$P(|Y-\mu| \le 3\sigma)=0.9974$$

可以认为，Y 的取值几乎全部集中在

$[\mu-3\sigma,\mu+3\sigma]$ 区间内.

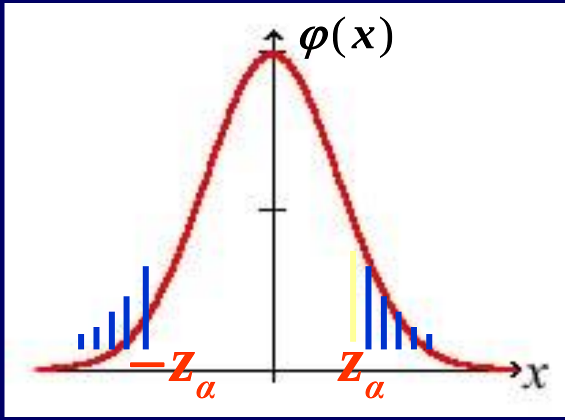
这在统计学上称作“3σ 准则” .

6° 标准正态分布的上α分位点

设 $X \sim N(0,1)$,若数z_α满足条件

$$P\{X > z_\alpha\} = \alpha\,,\, 0 < \alpha < 1 \Rightarrow P\{X < z_{-\alpha}\} = \alpha$$

则称点 z_α 为标准正态分布的上α分位点.



$$P\{X > z_{1-\alpha}\} = 1-\alpha$$

$$\Downarrow$$

$$P\{X < z_{1-\alpha}\} = \alpha$$

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

看一个应用正态分布的例子:

例 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～$N(170, 6^2)$, 问车门高度应如何确定?

解 设车门高度为h cm, 按设计要求

$$P(X\geq h)\leq 0.01$$

或 $$P(X< h)\geq 0.99,$$

下面我们来求满足上式的最小的h .

求满足 $P(X< h) \geq 0.99$ 的最小的 h.

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

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

$$= \Phi\left(\frac{h-170}{6}\right)$$

查表得 $\Phi(2.33)=0.9901>0.99$

因而 $\dfrac{h-170}{6} = 2.33,$

即 $h=170+13.98 \approx 184$

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.

四、小结

这一节，我们介绍了连续型随机变量及三种重要分布.即均匀分布、指数分布、正态分布. 其中正态分布的应用极为广泛，在本课程中我们一直要和它打交道.

后面第五章中，我们还将介绍为什么这么多随机现象都近似服从正态分布 .

练习题

一、设随机变量 X 的分布函数为

$$F_X(x)=\begin{cases}0, x<1,\\ \ln x, 1\le x<e,\\ 1, x\ge e.\end{cases},$$

?

求（1）$P\{X<2\}$, $P\{0<X\leqslant 3\}$；（2）求概率密度 $f_X(x)$.

二、设随机变量 X 的概率密度 $f(x)$为

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

求 X 的分布函数 $F(x)$，并作出 $f(x)$与 $F(x)$的图形。 ?

三、某种型号的电子的寿命 X（以小时计）具有以下的概率密度：

$$f(x)=\begin{cases}\dfrac{1000}{x^2} & x>1000\\ 0 & \text{其它}\end{cases}$$

现有一大批此种管子（设各电子管损坏与否相互独立）。任取 5 只，问其中至少有 2 只寿命大于 1500 小时的概率是多少？

?

四、设 $X\sim N(3,2^2)$

（1）求 $P\{2<X\leqslant 5\}$，$P\{-4<X\leqslant 10\}$，$P\{|X|>2\}$，

（2）决定 C 使得 $P\{X>C\}=P\{X\leqslant C\}$

?

一、设随机变量 X 的分布函数为

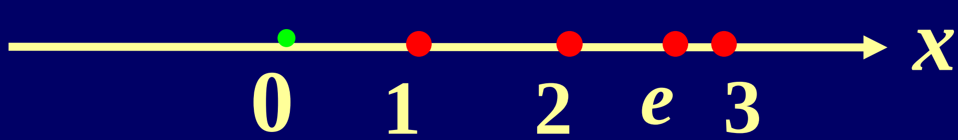
$$F_X(x)=\begin{cases}0, x<1,\\ \ln x, 1\le x<e,\\ 1, x\ge e.\end{cases},$$

求（1）$P\{X<2\}$, $P\{0<X\leqslant 3\}$；（2）求概率密度 $f_X(x)$.

解：$P\{X<2\}=F(2)-P\{X=2\}=F(2)=\ln 2$

$$P\{0<X\le 3\}=F(3)-F(0)=1$$

$$f_X(x)=F_X'(x)=\begin{cases}\dfrac{1}{x}, 1\le x<e,\\ 0, \quad 其它.\end{cases}$$



二、设随机变量 X 的概率密度 $f(x)$ 为

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

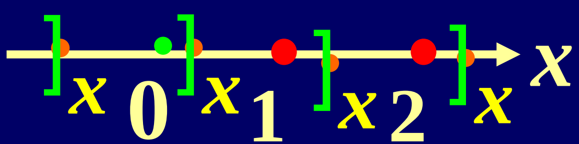
求 X 的分布函数 $F(x)$，并作出 $f(x)$ 与 $F(x)$ 的图形。

解： 当 $x<0$ 时，$F(x)=\int_{-\infty}^{x} f(t)dt=0$

当 $0\le x<1$ 时，$F(x)=\int_{-\infty}^{x} f(t)dt=\int_{0}^{x} tdt=\dfrac{x^2}{2}$

当 $1\le x<2$ 时，$F(x)=\int_{-\infty}^{x} f(t)dt=\int_{0}^{1} tdt+\int_{1}^{x}(2-t)dt$

$$=2x-\frac{x^2}{2}-1$$



当 $x\ge 2$ 时，$F(x)=1$

故

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

三、某种型号的电子的寿命 X（以小时计）具有以下的概率密度：

$$f(x)=\begin{cases}\dfrac{1000}{x^2} & x>1000\\ 0 & \text{其它}\end{cases}$$

现有一大批此种管子（设各电子管损坏与否相互独立）。任取 5 只，问其中至少有 2 只寿命大于 1500 小时的概率是多少？

解：$\{\text{寿命大于1500小时}\}\Leftrightarrow\{X>1500\}$

$$P\{X>1500\}=\int_{1500}^{+\infty}f(x)dx=\int_{1500}^{+\infty}\frac{1000}{x^2}dx=\left[-\frac{1000}{x}\right]_{1500}^{+\infty}=\frac{2}{3}$$

Y 表示5只中寿命大于 1500 小时的管子数

$$Y \sim b\left(5, \frac{2}{3}\right)$$

$$P\{Y \geq 2\} = 1 - P\{Y < 2\}$$

$$= 1 - P\{Y = 0\} - P\{Y = 1\}$$

$$= \frac{232}{243}$$

四、设 $X \sim N(3,2^2)$ $\Rightarrow \dfrac{X-3}{2} \sim N(0,1)$

（1）求 $P\{2<X\leqslant 5\}$，$P\{-4<X\leqslant 10\}$，$P\{|X|>2\}$，

（2）决定 C 使得 $P\{X>C\}=P\{X\leqslant C\}$

解:(1) $$P\{2<X\leq 5\}=P\{\frac{2-3}{2}<\frac{X-3}{2}\leq\frac{5-3}{2}\}$$

$$=\Phi(1)-\Phi(-0.5)=\Phi(1)+\Phi(0.5)-1=0.5328$$

$$P\{-4<X\leq 10\}=P\{\frac{-4-3}{2}<\frac{X-3}{2}\leq\frac{10-3}{2}\}$$

$$=\Phi(3.5)-\Phi(-3.5)=2\Phi(3.5)-1=0.9996$$

$$P\{|X|>2\}=1-P\{\frac{-2-3}{2}<\frac{X-3}{2}\leq\frac{2-3}{2}\}$$

$$=1-\Phi(-0.5)+\Phi(-2.5)=1-\Phi(2.5)+\Phi(0.5)$$

$$=0.6977$$

(2) 由对称性得：$C=3$

五、布置作业

习题2-4（p52）：6、10、14