复习-----描述随机变量统计规律的工具.

一、分布函数　适用范围—连续型、离散型R.V

$$F(x)=P(X\le x)\quad(-\infty<x<+\infty)$$

性质

(1) 即 $\forall\, x_1<x_2$，有 $F(x_1)\le F(x_2)$;

(2) $0\le F(x)\le 1$　$F(-\infty)=\lim\limits_{x\to-\infty}F(x)=0$

$$F(+\infty)=\lim_{x\to+\infty}F(x)=1$$

(3) $F(x)$ 右连续，即 $\lim\limits_{x\to x_0^+}F(x)=F(x_0)$

(4) $P(x_1<X\le x_2)=F(x_2)-F(x_1)$

2.分布律　　适用范围----离散型随机变量

$$P\{X = x_k\} = p_k, k = 1,2,\cdots$$

X	x_1	x_2	$\cdots$	x_k	$\cdots$
p_k	p_1	p_2	$\cdots$	p_k	$\cdots$

性质

（1）$p_k \geq 0$,　$k=1,2,\ldots$

（2）$\sum_k p_k = 1$

（3）$P(x \leq X \leq y) = \sum_{x \leq x_k \leq y} P(X = x_k) = \sum_{x \leq x_k \leq y} P_k$

（4）分布函数是一条跳跃型的阶梯形曲线

3.密度函数　适用范围----连续型随机变量

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

已知 f(x) 求分布函数F(x)

性质

（1）　$f(x)\geq 0$

求 f(x)表达式中的未知参数

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

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

计算X落入一个区间内的概率

（4）若 $f(x)$ 在点 x 处连续，则

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

已知 F(x) 求密度函数f(x)

分布律 ----离散型随机变量

$$P\{X = x_k\} = p_k, k = 1,2,\cdots$$

性质

（1）$p_k \geq 0,\quad k=1,2,\ldots$

（2）$\sum_k p_k = 1$

(3) $P(x \leq X \leq y) = \sum_{x \leq x_k \leq y} P_k$

(4) $F(x) = \sum_{x_k \leq x} P_k$

密度函数 ----连续型R.V

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

（1）$f(x) \geq 0$

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

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

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

（5）若 $f(x)$ 在点 x 处连续，则

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

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$P\left\{1 < X \le \dfrac{7}{2}\right\}$

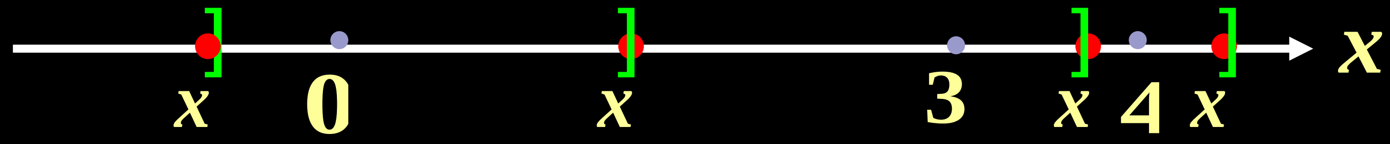
解 $$f(x)=\begin{cases} kx, & 0 \le x < 3 \\ 2-\dfrac{x}{2}, & 3 \le x \le 4 \\ 0, & \text{其它} \end{cases}$$

(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

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



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

例2：设连续型随机变量X的分布函数为

$$F(x)=\begin{cases}0 & x<-a\\ A+B\arcsin\dfrac{x}{a} & -a\le x<a\\ 1 & x\ge a\end{cases}$$

其中 a>0，试求：（1）常数A，B；（2）$P(|X|<\frac{a}{2})$;

（3）概率密度 f(x)。

解：（1）因为 $F(x)$在 $(-\infty,+\infty)$ 上连续，故有

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

$$A-\frac{\pi}{2}B=0,\ \ 1=A+\frac{\pi}{2}B$$ 解得：$A=\frac{1}{2}$, $B=\frac{1}{\pi}$

例2：设连续型随机变量X的分布函数为

$$F(x)=\begin{cases}0 & x<-a\\ A+B\arcsin\dfrac{x}{a} & -a\le x<a\\ 1 & x\ge a\end{cases}$$

其中 a>0，试求:（1）常数A，B;（2）$P(|X|<\frac{a}{2})$;（3）分布密度 f(x)。

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

$$=\frac{1}{2}+\frac{1}{\pi}\arcsin\frac{1}{2}-(\frac{1}{2}-\frac{1}{\pi}\arcsin\frac{1}{2})=\frac{2}{\pi}\arcsin\frac{1}{2}=\frac{1}{3}$$

（3）$f(x)=F'(x)=\begin{cases}\dfrac{1}{\pi\sqrt{a^2-x^2}} & |x|<a\\ 0 & |x|\ge a\end{cases}$

例3 已知某型号电子管的使用寿命 X 为连续随机变量, 其密度函数为

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

(1) 求常数 c. $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=\int_{1000}^{+\infty}\dfrac{c}{x^2}\mathrm{d}x=1$

(2)求 任取一只寿命大于1500小时的概率 $P(X>1500)$

(3) 计算 $P(X\le 1700 \mid 1500<X<2000)$

(4) 已知一设备装有3个这样的电子管, 每个电子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. $Y\sim b(3,p)$

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

$\longrightarrow$ $c = 1000$

(2) $P(X > 1500)$

$$= \int_{1500}^{+\infty} \frac{1000}{x^2}\,\mathrm{d}x = \left(-\frac{1000}{x}\right)\Big|_{1500}^{+\infty} = \frac{2}{3}$$

密度函数 $f(x) = \begin{cases} \dfrac{1000}{x^2}, & x > 1000 \\ 0, & \text{其他} \end{cases}$

(3) $P(X \le 1700 \mid 1500 < X < 2000)$

$$= P(X \le 1700, 1500 < X < 2000) \Big/ P(1500 < X < 2000)$$

$$= P(1500 < X \le 1700) \Big/ P(1500 < X < 2000)$$

$$= \int_{1500}^{1700} \frac{1000}{x^2} d\,x \Big/ \int_{1500}^{2000} \frac{1000}{x^2} d\,x$$

$$= \frac{4}{51} \Big/ \frac{1}{6} = \frac{24}{51} \approx 0.4706.$$

(4)

设A 表示一个电子管的寿命小于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小时三个电子管中损坏的个数为 Y

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

§ 2.4-2几种常见连续型随机变量

均匀分布

指数分布

正态分布

1. 均匀分布 *Uniform distribution*

(1) 均匀分布　　若 r.v X 的密度函数为

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\[2ex] 0, & \text{其他}\end{cases}$$

则称 X 服从区间(a,b)上的均匀分布或称

X 服从参数为 a,b的均匀分布. 记作

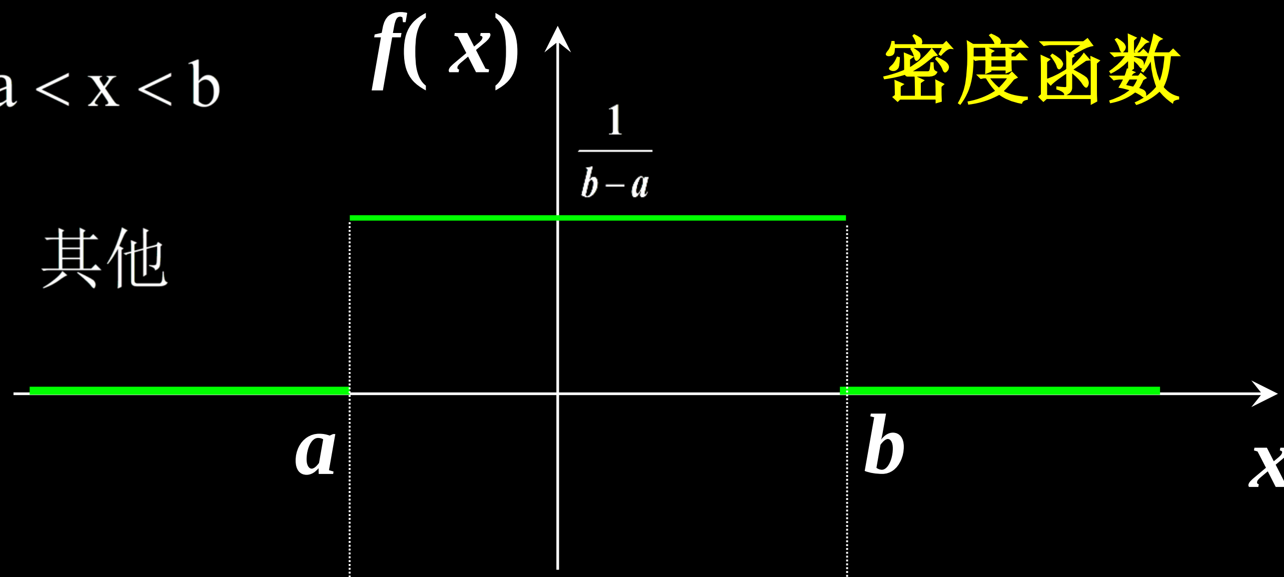
$$X\sim U(a,b)$$

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

X 的分布函数为

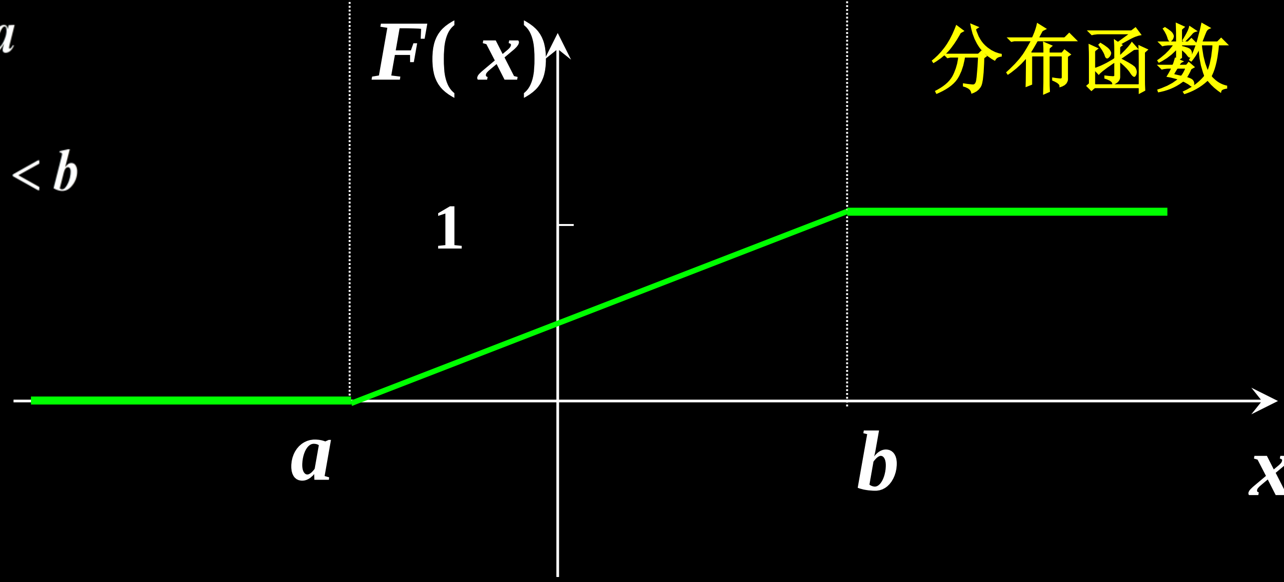
$$F(x)=\int_{-\infty}^{x} f(t)\,dt \quad =\begin{cases}0, & x<a,\\ \dfrac{x-a}{b-a}, & a\le x<b,\\ 1 & x\ge b\end{cases}$$

$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\[2ex] 0, & \text{其他}\end{cases}$$



密度函数

$$F(x)=\begin{cases}0, & x<a\\[1ex] \dfrac{x-a}{b-a}, & a\le x<b\\[2ex] 1 & x\ge b\end{cases}$$



分布函数

例1. 长途汽车起点站于每时的10分、25分、55分发车，设乘客不知发车时间，于每小时的任意时刻随机地到达车站，(1)求乘客候车时间超过10分钟的概率. (2)现有4名乘客，求至少有三人候车时间超过10分钟的概率，（其中4名乘客到达车站时间相互独立）

解：设A —乘客候车时间超过10分钟

X—乘客于某时X 分钟到达，则X ~U(0,60)

Y—4人中候车时间超过10分钟的人数，则Y~b(4,P(A))

$$P(A) = P\{10 < X \le 15\} + P\{25 < X \le 45\} + P\{55 < X \le 60\}$$

$$= \frac{5+20+5}{60} = \frac{1}{2}$$

$$P(Y \ge 3) = C_4^3 0.5^3 \cdot 0.5 + C_4^4 0.5^4 = 5 \times 0.5^4$$

2 指数分布

1）指数分布的密度函数与分布函数

若 X 的密度函数为

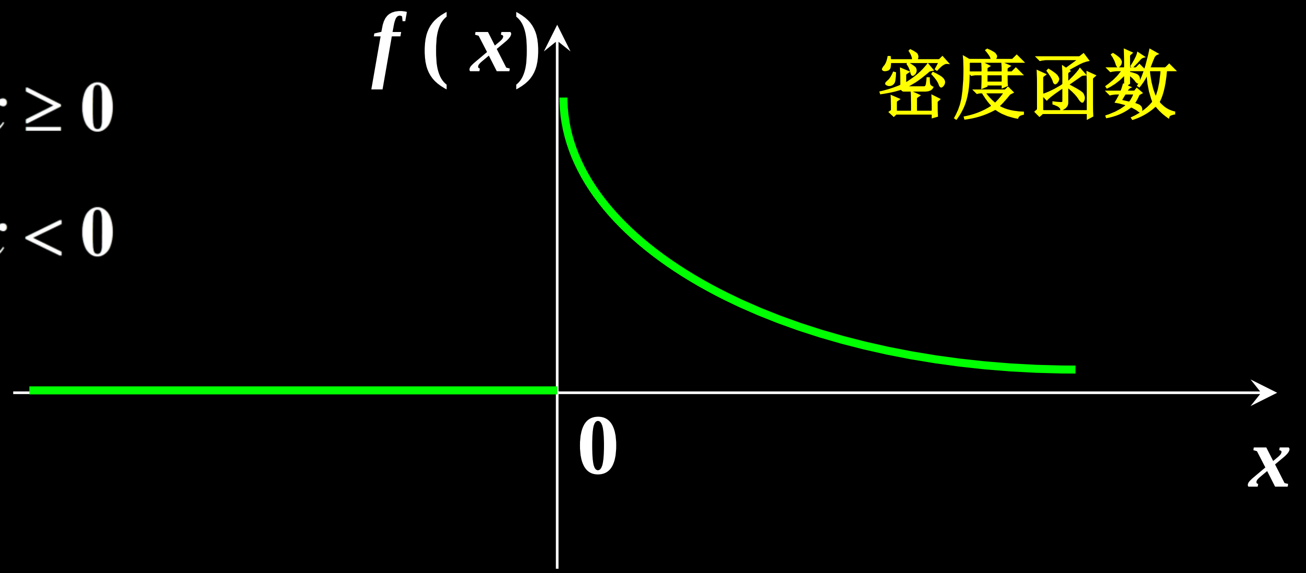
$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x\ge 0\\ 0, & x<0\end{cases}\qquad \lambda>0\text{ 为常数}$$

则称 X 服从 参数为λ的指数分布

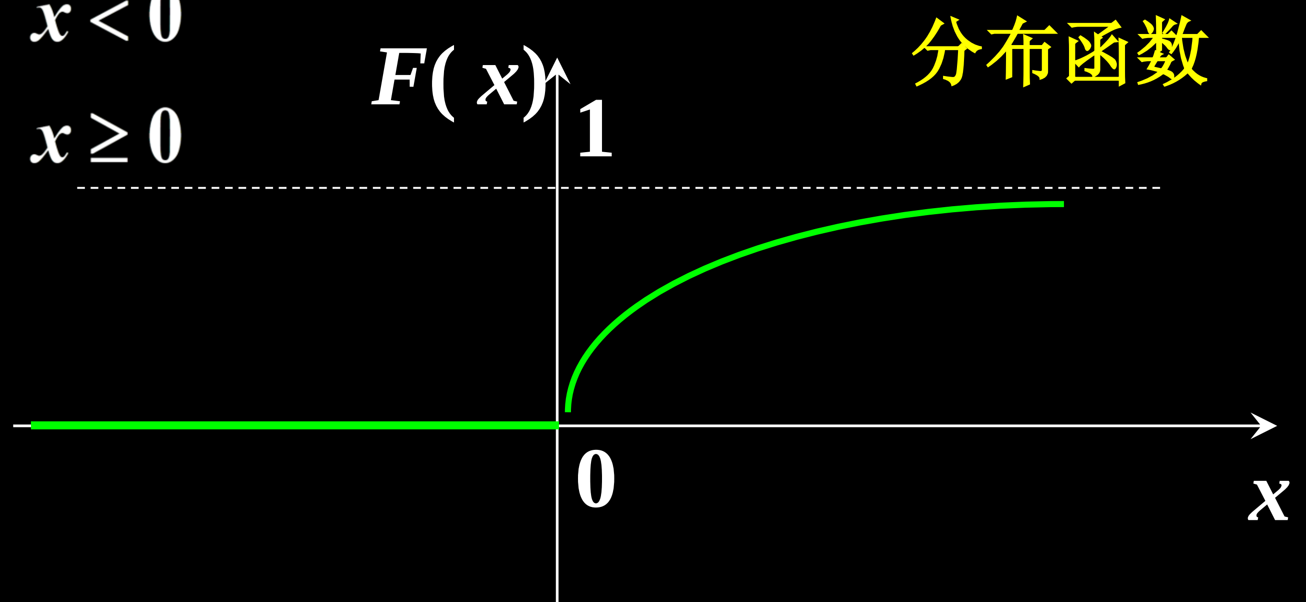
记作$X\sim Exp(\lambda)$

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

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



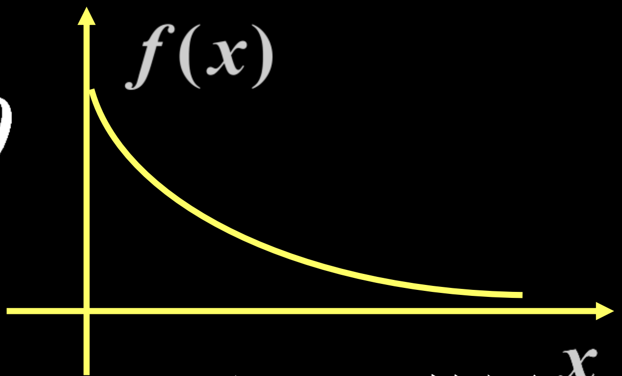
$$F(x)=\begin{cases}0, & x<0\\ 1-e^{-\lambda x}, & x\ge 0\end{cases}$$



指数分布的密度函数有时也写成

$$f(x)=\begin{cases}\frac{1}{\theta}e^{-x/\theta}, & x\ge 0\\ 0, & x<0\end{cases}$$

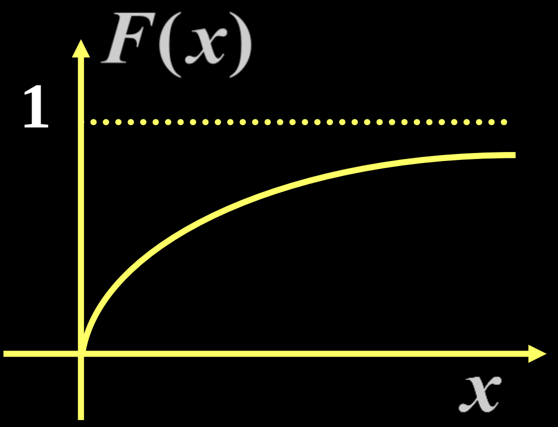
其中 θ 是常数



称此随机变量X 服从参数为 θ 的指数分布，f(x) 和 F(x) 的图形

由此可得出指数分布的分布函数：

$$F(x)=\begin{cases}1-e^{-x/\theta} & x\ge 0\\ 0 & x<0\end{cases}$$



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

指数分布常作为各种“寿命”分布的近似

2）指数分布的“无记忆性”

若 $X \sim Exp(\lambda)$, 则对任意 $s>0,t>0$ 有

$$P(X > s + t | X > s) = P(X > t)$$

证明：

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

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

故又把指数分布称为“永远年轻”的分布

例2: 假定自动取款机对每位顾客的服务时间(单位:分钟)服从 $\theta = 3$ 的指数分布. 如果有一顾客恰好在你前头走到空闲的取款机. 求你

(1) 至少等候3分钟;

(2) 等待时间在3分钟至6分钟之间的概率.

如果你到达取款机时, 正有一名顾客使用着取款机, 上述概率又是多少?

解: 以X表示你前面这位顾客所用服务时间. $F(x)$ 为X的分布函数. 则所求的概率为:

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

(2) $$P\{3 < X < 6\} = F(6) - F(3) = e^{-1} - e^{-2}$$

例3: 设 X 服从参数 $\theta = 1$ 的指数分布，求方程 $4x^2 + 4xX + X + 2 = 0$ 无实根的概率（关于x 的二次方程）。

解: 要方程无实根，必须满足

$$\Delta = (4X)^2 - 4 \times 4 \times (X + 2) < 0 \qquad (X+1)(X-2) < 0$$

解得 $-1 < X < 2$

由于X的分布密度为 $f(x) = \begin{cases} e^{-x} & x \ge 0 \\ 0 & x < 0 \end{cases}$

所以

$$P(-1 < X < 2) = \int_{-1}^{0} f(x)dx + \int_{0}^{2} f(x)dx$$

$$= \int_{0}^{2} e^{-x}dx = 1 - e^{-2} = 0.8647$$

3. 正态分布

德莫佛

1）正态分布的密度函数和分布函数

若r.v X的概率密度为

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

其中 μ 和 σ^2 都是常数，μ 任意，$\sigma>0$，
则称X 服从参数为μ 和 σ^2 的正态分布.

称X 为正态变量

记作 $X\sim N(\mu,\sigma^2)$

$f(x)$所确定的曲线叫作正态曲线.

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

证明思路：作变量代换 $t=\frac{x-\mu}{\sigma}$

即证 $I=\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt=\sqrt{2\pi}$

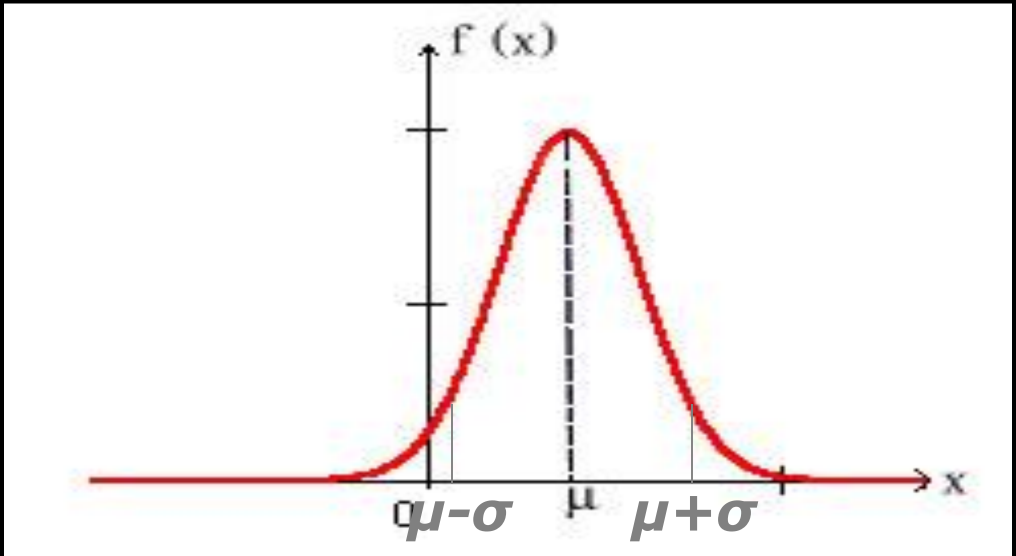
先证：

$$I^2=\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt\int_{-\infty}^{\infty}e^{-\frac{u^2}{2}}du=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{e}^{-(t^2+u^2)/2}dtdu=2\pi$$

转化为极坐标：

$$\begin{cases}t=r\cos\theta\\ u=r\sin\theta\end{cases}\qquad I^2=\int_0^{2\pi}d\theta\int_0^{\infty}e^{-\frac{r^2}{2}}rdr=2\pi$$

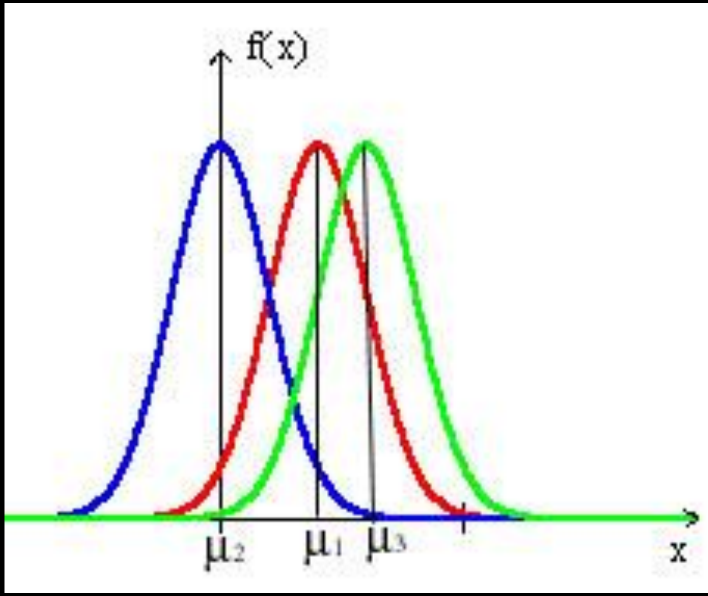
正态分布$N(\mu,\sigma^2)$ 的图形特点

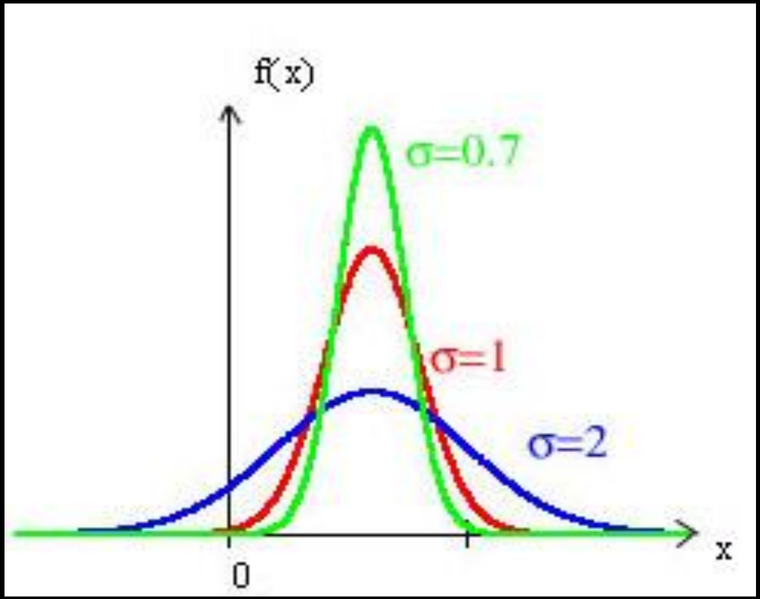


正态分布的密度曲线是一条关于 μ 对称的钟形曲线.

特点是“两头低，中间高，左右关于μ对称”.

正态分布$N(\mu,\sigma^2)$的图形特点



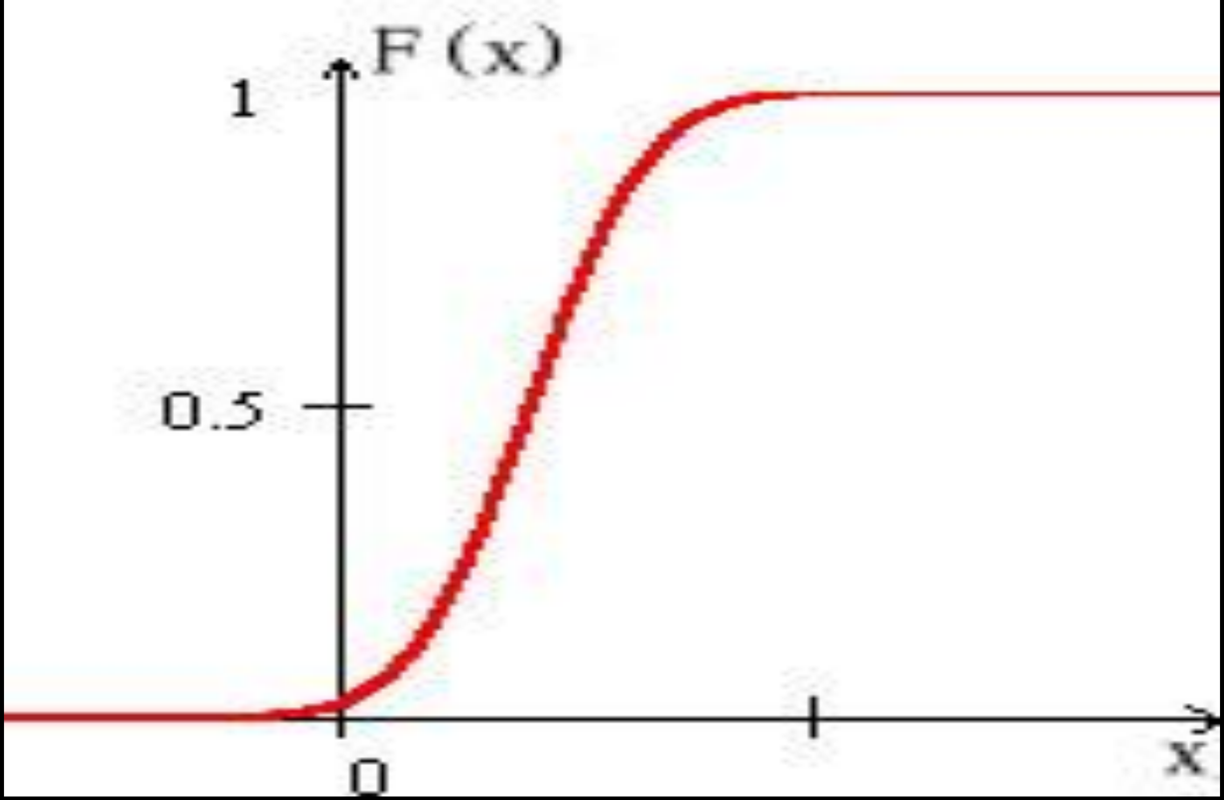


μ决定了图形的中心位置，σ决定了图形中峰的陡峭程度.

μ称为位置参数，σ称为尺度参数

设$X \sim N(\mu,\sigma^2)$， X的分布函数是

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



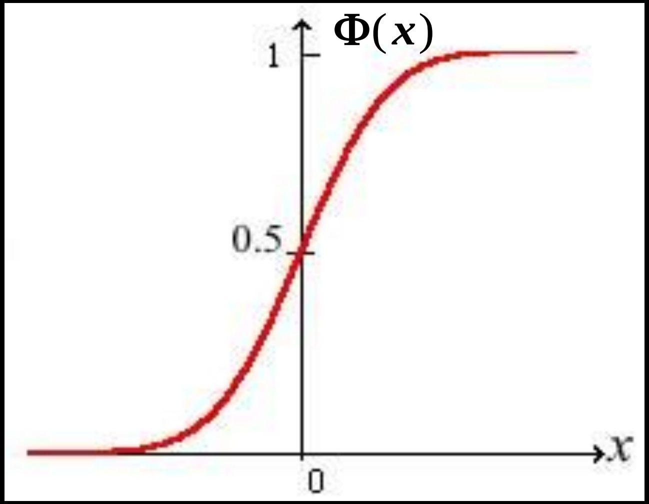
2）标准正态分布

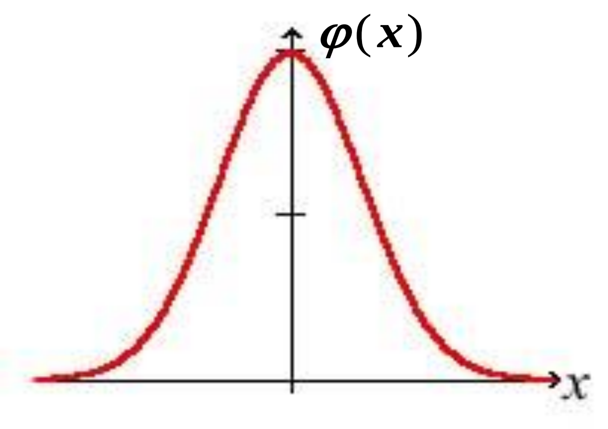
$\mu=0,\sigma=1$的正态分布称为标准正态分布.

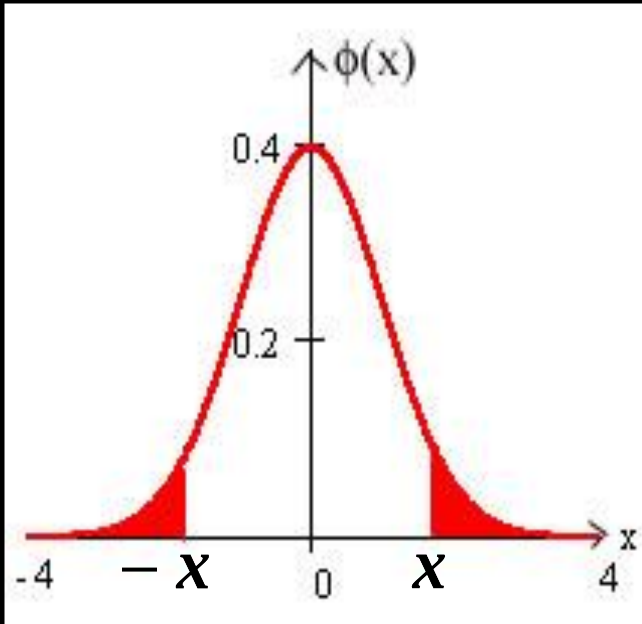
其密度函数和分布函数常用 $\varphi(x)$和 $\Phi(x)$表示：

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

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







设$X \sim N(0,1)$

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

表中给的是$x>0$时, $\Phi(x)$的值.

当$-x<0$时

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

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

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

标准正态分布表

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

x	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5	0.504	0.508	0.512	0.516	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.591	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.648	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.67	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.695	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.719	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.758	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.791	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.834	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.877	0.879	0.881	0.883
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.898	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.937	0.9382	0.9394	0.9406	0.9418	0.943	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9648	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.97	0.9706

例 4　设$X \sim N(0,1)$,求下列事件的概率:

(1) $P(X < 0.84) = \Phi(0.84) = 0.7995$

(2) $P(X < -0.84) = \Phi(-0.84)$

$$= 1 - \Phi(0.84) = 0.2005$$

(3) $P(-0.41 \le X < 0.84) = \Phi(0.84) - \Phi(-0.41)$

$$= 0.7995 - (1 - 0.6591)$$

$$= 0.4586$$

(4) $P(|X| < 0.84) = 2\Phi(0.84) - 1$

$$= 2 \times 0.7995 - 1 = 0.5990$$

3）一般正态分布的标准化

若 $X \sim N(\mu, \sigma^2)$，则

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

事实上

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

令 $\frac{t-\mu}{\sigma} = u$

$$\overline{\overline{\qquad}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

若 $X \sim N(0,1)$,

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

若 $X \sim N(\mu, \sigma^2)$,

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

例5　设$X \sim N(108,9)$,求

(1) $P(102 < X < 117)$;

(2) 使得$P(X < a) = 0.9505$,则$a = ?$;

解:(1) $P(102 < X < 117)$

$$= \Phi(\frac{117-108}{3}) - \Phi(\frac{102-108}{3})$$

$$= \Phi(3) - \Phi(-2) = \Phi(3) + \Phi(2) - 1$$

$$= 0.9987 + 0.9772 - 1 = 0.9759$$

(2) $P(X < a) = \Phi(\frac{a-108}{3}) = 0.9505$

$$\Phi(1.65) = 0.9505$$

$$\frac{a-108}{3} = 1.65 \longrightarrow a = 112.95$$

例6: 设 $X \sim N(\mu,\sigma^2)$, $\lambda > 0$, 证明X落在$(\mu-\lambda\sigma,\mu+\lambda\sigma)$内的概率只与 λ 有关而与 μ,σ 无关。

证: $P\{\mu-\lambda\sigma < X < \mu+\lambda\sigma\}$

$$= F(\mu+\lambda\sigma) - F(\mu-\lambda\sigma) = \Phi(\lambda) - \Phi(-\lambda)$$

$$= 2\Phi(\lambda) - 1$$

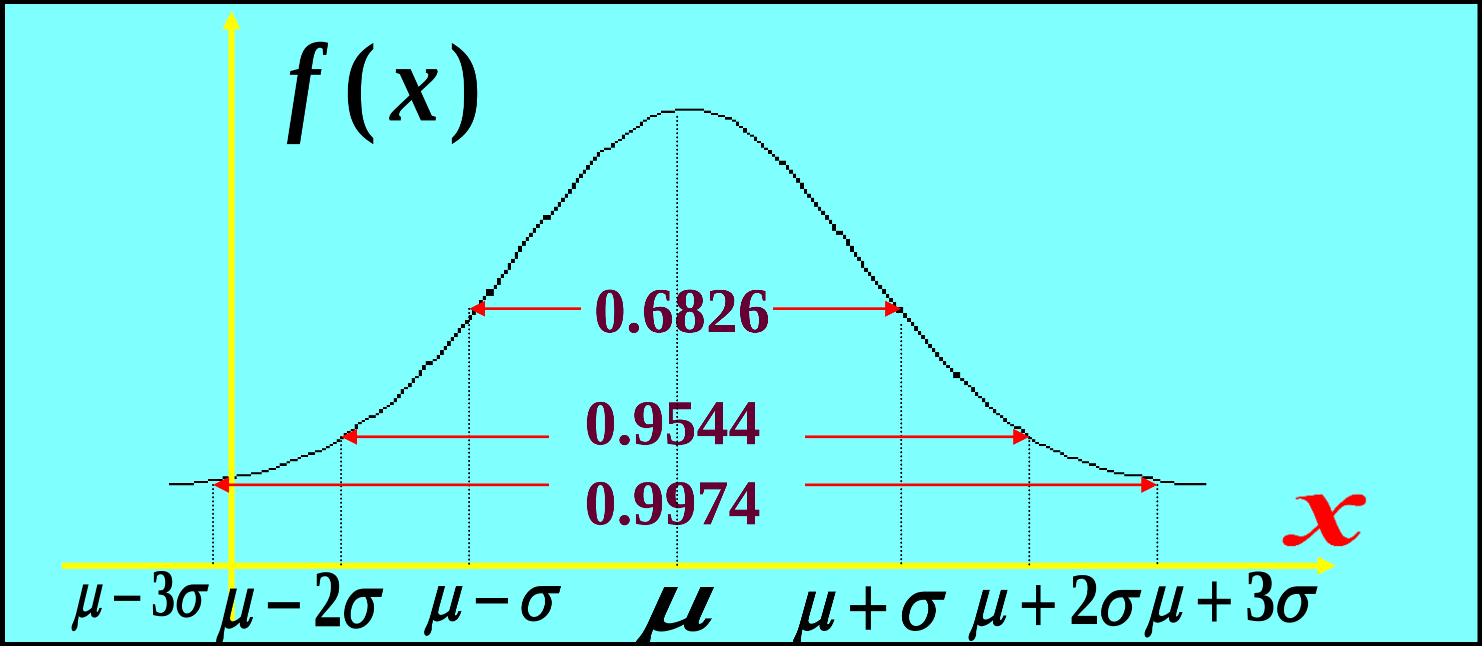
$\therefore$ X落在区间$(\mu-\lambda\sigma,\mu+\lambda\sigma)$内的概率只与$\lambda$有关而与 μ,σ 无关。

特别当 λ =1，2，3时，可查表求得

$$P\{\mu-\sigma < X < \mu+\sigma\} = 2\Phi(1) - 1 = 0.6826$$

$$P\{\mu-2\sigma < X < \mu+2\sigma\} = 2\Phi(2) - 1 = 0.9544$$

$$P\{\mu-3\sigma < X < \mu+3\sigma\} = 2\Phi(3) - 1 = 0.9974$$



可见服从正态分布 $N(\mu,\sigma^2)$ 的随机变量X 之值基本上落在区间 $(\mu-2\sigma,\mu+2\sigma)$ 内，而几乎不落在 $(\mu-3\sigma,\mu+3\sigma)$ 之外，在实际应用中称为 **3σ原则**。

练习．设随机变量X的密度函数为

$$f(x)=\begin{cases} A\cos x & -\dfrac{\pi}{4}\le x\le\dfrac{\pi}{4} \\ 0 & \text{其它} \end{cases}$$

（1）求系数 A；（2）求X的分布函数F(x)；

（3）求概率 $P(0<X<\dfrac{\pi}{6})$

.