§ 2.1　随机变量及其分布函数

- **定义**：设X(e)是定义于样本空间Ω上的单值实函数，则称X(e)为**随机变量**。 随机变量（r.v.）有时也简记为X。

  随机变量一般用罗马字母X、Y、Z等表示，也可用希腊字母 ξ、η、ζ等来表示。

- **例**：投掷一枚硬币； X:正面->1,反面->0;

  投掷一粒骰子； X:出现第i点->i;

  一夫妇生了三个小孩； X:女孩的个数->i;

  乘客的候车时间; 学生的成绩等等

- **例：**从含有4件次 品的9件产品中任意取3件，若以X表示取出的3件产品中的次品数，则X是一个随机变量。
- 事件{取出的3件产品中有2件次品}就可以简单地表示为{X=2},
- 事件{取出的3件产品中的次品数小于2件}就可以表示为{X<2}.
- **随机变量=变量+随机**

$$(\Omega, P)\xrightarrow{X}(R, P_X)$$

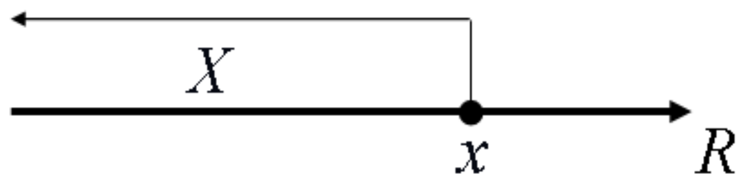
$$\forall A \subset R, P_X(A) = P(X^{-1}(A))$$

函数 概率（分布函数）

- **定义**：设X是一随机变量，x是任意实数，称函数

$$F(x)=P(X\le x)$$

为随机变量X的**分布函数**。



$P(X\le x)= F(x)$ $P(X>x)= 1-F(x)$

$P(a<X\le b)= F(b)- F(a)$ $P(X<x)=?$

- 分布函数完整地描述了随机变量的概率规律性。显然，分布函数是一个普通函数，因此我们可以通过它，利用高等数学的方法来研究随机变量（随机现象）

分布函数$F(x)$的性质

- （1） F(x)是个单调不减函数（单调递增的函数）
- （2）$0 \le F(x) \le 1$

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

$$F(-\infty) = \lim_{x \to -\infty} F(x) = 0$$

- （3） F(x) 是右连续的，即F(x) =F(x+0)
- 证明：

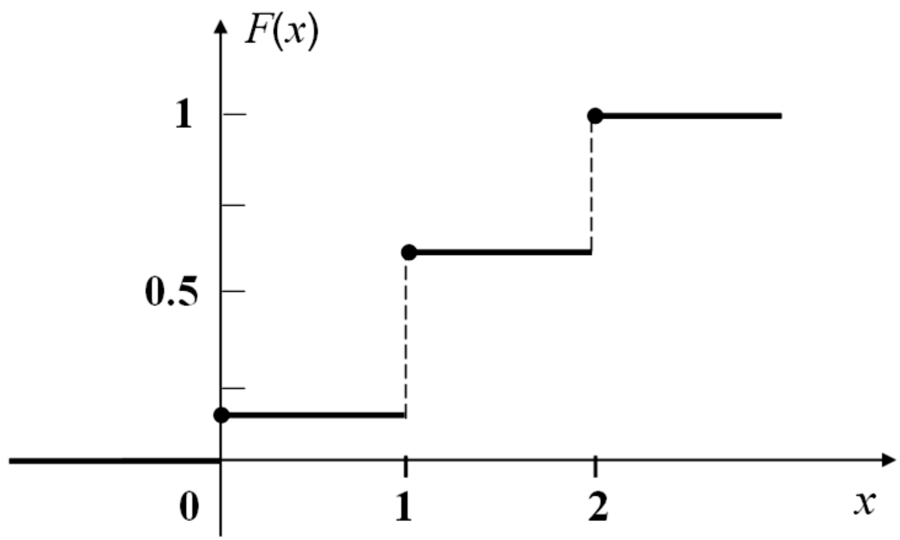
$$\lim_{x \to +\infty} F(x) = \lim_{x \to +\infty} P(X \le x) = P(\lim_{x \to +\infty} \{X \le x\}) = P(\Omega) = 1$$

$$\lim_{x \to x_0+0} F(x) = \lim_{x \to x_0+0} P(X \le x) = P(\lim_{x \to x_0+0} \{X \le x\}) = P(\{X \le x_0) = F(x_0)$$

- **例**：设X的分布函数为：

$$F(x)=\begin{cases}0 & x<0\\ 0.16 & 0\le x<1\\ 0.64 & 1\le x<2\\ 1 & x\ge 2\end{cases}$$

- **解**：图形如下：



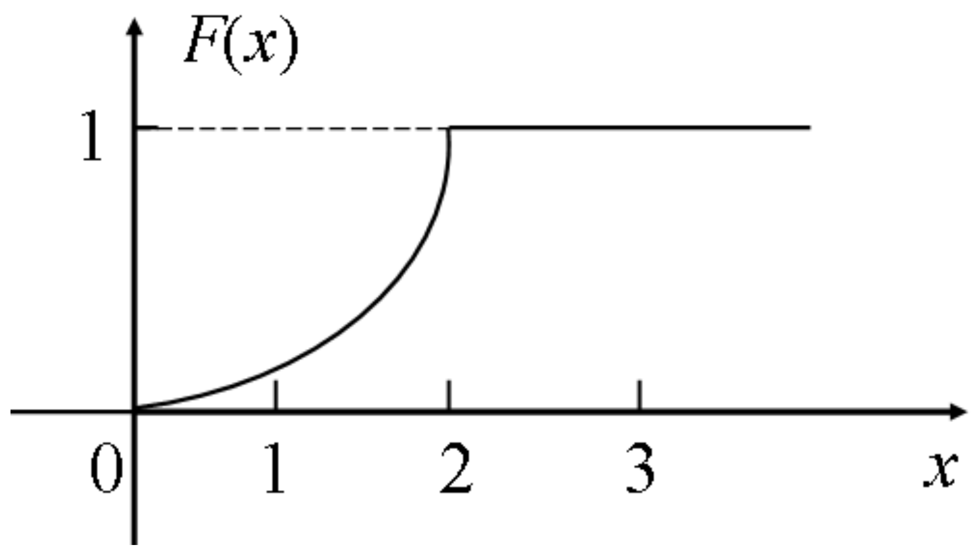
- 例：设随机变量X的分布函数为：

1. F(x)=A+Barctanx;
2. $$F(x)=\begin{cases} 0 & x \le -a \\ A+B\arcsin(x/a) & -a < x \le a \\ 1 & x > a \end{cases}$$

求A，B。

- **例**：个靶子是半径为 2 米的圆盘，设击中靶上任一同心圆盘上的点的概率与该圆盘的面积成正比，若射击都能中靶，以X表示弹着点与圆心的距离。求随机变量X的分布函数。

- **解**：



§ 2.2 离散型随机变量

- **定义：** 设$x_1, x_2, x_3\ldots$为离散型随机变量X的所有可能取值， 记

$$p_k=P(X=x_k)，k=1,2,\ldots$$

称为**离散型随机变量*X*的分布律（或概率分布）**。

- 分布律也可以用表格的形式来表示:

X	x_1	x_2	x_3	$\cdots$
P	p_1	p_2	p_3	$\cdots$

或 $$X \sim \begin{pmatrix} x_1 & x_2 & x_3 & \cdots \\ p_1 & p_2 & p_3 & \cdots \end{pmatrix}$$

离散型随机变量的分布律具有下列性质：

- 1. 非负性， $p_k \geq 0$，$k=1,2,\ldots$
- 2. 正规性，

$$\sum_{k=1}^{\infty} p_k = 1$$

分布函数与概率的计算：

$$P(a \leq X \leq b) = \sum_{k: a \leq x_k \leq b} p_k \quad P(X \in I) = \sum_{k: x_k \in I} p_k$$

$$F(x) = P(X \leq x) = \sum_{k: x_k \leq x} p_k$$

几种常用离散型分布

1.（0—1）分布

- 如果r.v. X 的所有可能取值为0，1，且有分布律

$$P(X=k)=p^k(1-p)^{1-k},\ k=0,1\ (0<p<1)$$

则称X服从参数为p的**（0—1）分布**或两点分布。

- （0—1）分布的分布律也可表示成：

X	0	1
P	$1-p$	p

2. 二项分布

- 在n重贝努里试验中，若设X为n次试验中A出现的次数，设P(A)=p (0<p<1)，则X是一个离散型随机变量，其分布律为：

$$p_k = P\{X = k\} = C_n^k p^k q^{n-k}, k = 0,1,\cdots,n.$$

我们称这样的随机变量X服从参数为n，p的**二项分布**，记为X~B(n,p)。

- 特别地，当n =1时的二项分布即为（0—1）分布.

- 下面研究二项分布的**最大值点**：

$$\frac{p_k}{p_{k-1}} = \frac{C_n^k p^k q^{n-k}}{C_n^{k-1} p^{k-1} q^{n-k+1}} = \frac{(n-k+1)p}{kq} = 1 + \frac{(n+1)p-k}{kq}$$

- 1. 当 (n+1)p 为整数时，在k=(n+1)p和k=(n+1)p-1达到最大，
- 2. 当 (n+1)p 不是整数时，在 k=[(n+1)p] 达到最大。其中，[*x*] 表示不超过*x*的最大整数。

- **例**：假设一厂家生产的每台仪器，以概率0.7可以直接出厂；以概率0.3需进一步调试，经调试后以概率0.8可以出厂；以概率0.2定为不合格不能出厂。现该厂新生产了n(n >3)台仪器（假设各台仪器的生产过程相互独立），求(1)全部能出厂的概率；(2)其中恰好有两件不能出厂的概率；(3)其中至少有两件不能出厂的概率.
- **解**： 设A={仪器需进一步调试}，B={仪器可出厂}.**(1)** $P(B)=P(A)P(B\mid A)+P(\bar{A})P(B\mid \bar{A})=0.3\times 0.8+0.7\times 1=0.94$

令X={n台仪器可出厂的台数} ,$X\sim B$(n,0.94)

(2) $P(X=n-2)=C_n^{n-2}0.94^{n-2}0.06^2$

(3) $P(X\le n-2)=1-C_n^{n-1}0.94^{n-1}0.06\ -0.94^n$

3. 泊松分布

- 设r.v. X的所有可能取值为0，1，2，…，且

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda},k=0,1,\cdots$$

其中 λ>0是常数，则称X服从**参数为λ的泊松（Poisson）分布**，记为X～ P(λ)。

- 泊松分布的应用：
稀有事件发生的次数；
质点流的流量；
作为二项分布的近似。

$\lambda=4$. 平均卖4件.

- **例**：设某商店某种商品每月销售数服从参数为4的泊松分布，问在月初进货时应进多少件此商品，才能保证当月不脱销的概率至少是90%？
- **解**：设该商店每月销售此种商品X件，月初进货数为a件，那么当X ≤ a时就不会脱销．已知X～P(4)，上述要求即为$P(X \le a) \ge 90\%$,，即求满足下式的最小的a：

$$\sum_{k=0}^{a} \frac{4^k}{k!} e^{-4} \ge 0.9$$

- 查表知：当a=6时，上式左边=0.8893；当a=7时，上式左边=0.9489。因此，该商店只要在月初进7件该商品就能保证当月不脱销的概率至少是90%.

4. 几何分布

- 设r.v. X的所有可能取值为1，2，…，且

$$P(X=k)=pq^{k-1}, k=1,2,\cdots$$

其中 0<p<1是常数，q=1-p。则称X服从**参数为p的几何分布**，记为X～ g(p).

- 1. 几何分布的概率背景：贝努里概型中事件A首次发生的试验次数；

- 2.几何分布具有无记忆性。即P(X=s+t|X>t)=P(X=s)

- **例**：一个人有n把钥匙，其中只有一把能打开门。一天此人醉酒回家，随机抽取一把开门（若该钥匙不能开门，则放回重新抽取），问此人在第k次才打开门的概率是多少？
- **解**：每一次抽取钥匙都是一次贝努里试验，$p=1/n$，以X表示将门打开所需试验的次数，则

$$P(X=k)=(1-\frac{1}{n})^{k-1}\frac{1}{n},k=1,2,\cdots$$

§ 2.3 连续型随机变量

- **定义**： 若X是随机变量，其分布函数为F(x)，如果存在非负函数p(x)，使得对于任意实数x，有

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

则称X是**连续型随机变量**，而称p(x)为X的**概率密度函数**（简称**密度函数**）。

密度函数$p(x)$具有以下结论：

- 1. $p(x) \geq 0; \int_{-\infty}^{+\infty} p(x)dx = 1.$
- 2. $P(x_1 < X \leq x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} p(x)dx$
- 3.对任意x，P(X=x)=0。即F(x)为连续函数，且

$$P(x_1 < X < x_2) = P(x_1 < X \leq x_2) = P(x_1 \leq X \leq x_2)$$

$$= P(x_1 \leq X < x_2) = \int_{x_1}^{x_2} p(t)dt$$

- 4. 若$p(x)$在x处连续，则$F'(x) = p(x)$
- 5.密度函数p(x)反映了随机变量X在x点处附近概率的大小。

几种常用的连续型分布

1. 均匀分布

- 若随机变量X的密度函数为

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

则称X 服从**区间[a,b]上的均匀分布**，记为X～U[a,b]。

- X的分布函数为

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

- 特别的，在区间[0，1]上的均匀分布X ～$U[0,1]$
- 密度函数为

$$p(x)=\begin{cases}1, & 0\le x\le 1\\ 0, & \text{其它}\end{cases}$$

- 分布函数为

$$F(x)=\begin{cases}0, & x\le 0\\ x, & 0<x\le 1\\ 1, & x>1\end{cases}$$

2. 指数分布

- 若随机变量X的密度函数为

$$p(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & \text{其它}\end{cases}$$

其中λ >0, 则称X服从参数为λ 的指数分布，记为X~E(λ)。分布函数为：

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

- 指数分布具有无记忆性。

3.正态分布

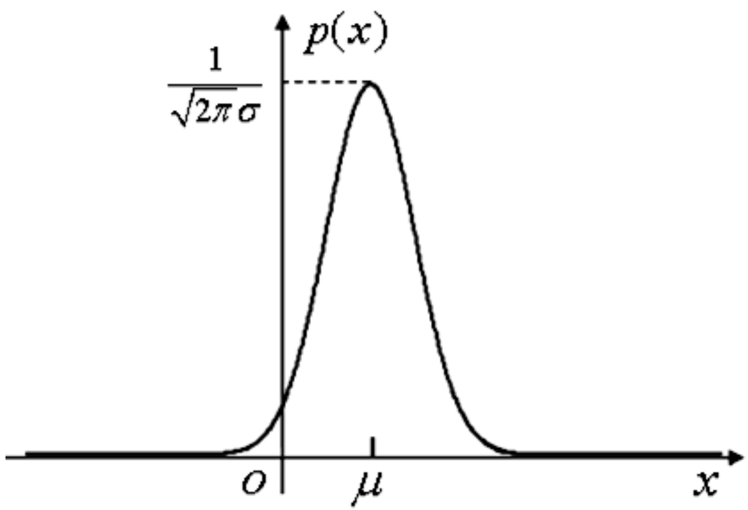
- 若随机变量X的密度函数为

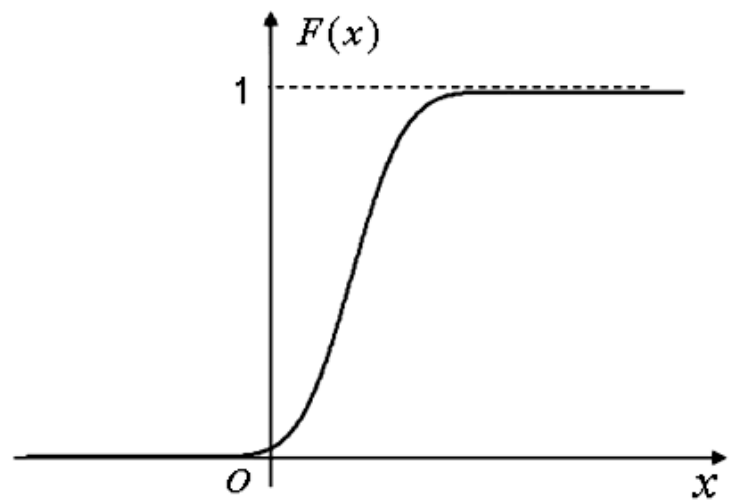
$$p(x)=\frac{1}{\sqrt{2\pi}\cdot\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad x\in R$$

其中μ,σ(σ>0)是常数，则称X服从**参数为μ,σ²的正态分布(或高斯分布)**，记为X~N(μ,σ²).分布函数为：

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

正态分布的密度函数与分布函数曲线如下：





- 1. $p(x)$是关于直线$x=\mu$对称的，在$x=\mu$处取得极大值。

- 2. 固定μ，σ的值越大，$p(x)$的图形就越平坦；σ的值越小，$p(x)$的图形就越陡峭。σ决定图形的形状。

- 3. 固定σ，μ的值增大，$p(x)$的图形就向右移动；μ的值减小，$p(x)$的图形就向左移动。μ决定图形的位置。

标准正态分布

- 特别地，当μ=0,，σ=1时，正态分布称为标准正态分布，记为X~N(0,1)，相应的密度函数以及分布函数分别记为φ(x)和Φ(x)。

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

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

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

- 定理：设$X \sim N(\mu,\sigma^2)$，则$(X-\mu)/\sigma \sim N(0,1)$

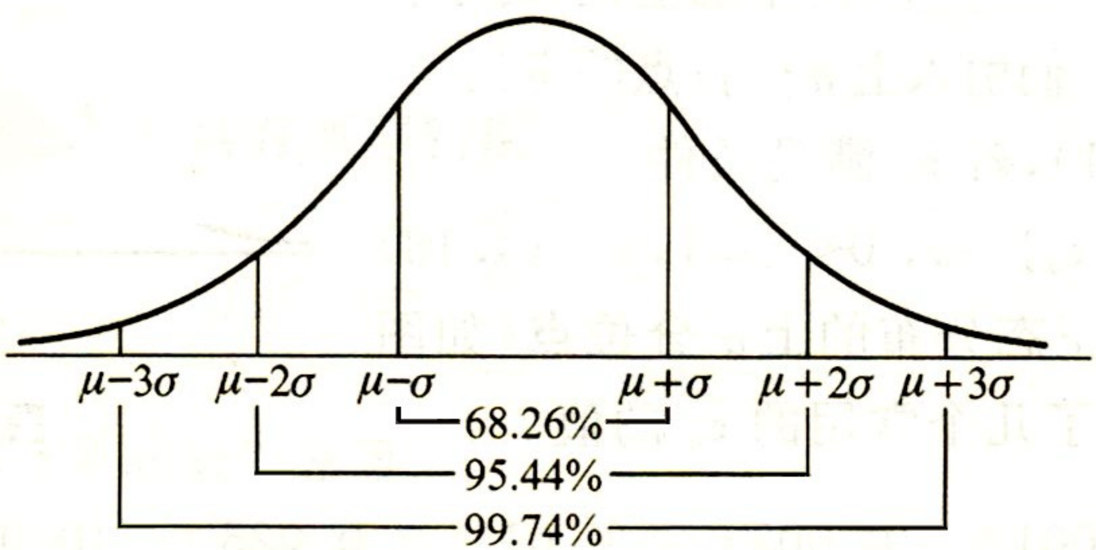
- 一般的，设$X \sim N(\mu,\sigma^2)$，则$aX+b \sim N(a\mu+b,(a\sigma)^2)$

正态分布的 3 σ 原则

$$P(|X-\mu|<\sigma)=2\Phi(1)-1=2\times 0.8413-1=0.6826$$

$$P(|X-\mu|<2\sigma)=2\Phi(2)-1=0.9545$$

$$P(|X-\mu|<3\sigma)=2\Phi(3)-1=0.9973$$



- **例**：设某次外语统考的成绩服从正态分布，平均成绩为78分，92分以上的占学生总数的2.28%，求学生成绩在71分至85分之间的概率.
- **解**：

设$X \sim N(78,\sigma^2)$,则$\dfrac{X-78}{\sigma} \sim N(0,1)$.

$$P(X>92)=P\left(\frac{X-78}{\sigma}>\frac{92-78}{\sigma}\right)=1-\Phi\left(\frac{92-78}{\sigma}\right)=0.0228$$

查表σ=7.

$$P(71<X<85)=P\left(\frac{71-78}{7}<\frac{X-78}{7}<\frac{85-75}{7}\right)=2\Phi(1)-1=0.6826$$

一个既不是离散型，也不是连续型的随机变量

- **例**：设随机变量X~E(λ)，求Y=max{X,2}的分布函数。
- **解**：

$$F_Y(y) = P(\max\{X,2\} \le y)$$

$$\text{当} y < 2, F_Y(y) = 0;$$

$$\text{当} y \ge 2, F_Y(y) = P(\max\{X,2\} \le y)$$

$$= P(X \le y, 2 \le y) = P(X \le y) = 1 - e^{-\lambda y}$$

$$\text{故} F_Y(y) = \begin{cases} 0 & y < 2 \\ 1 - e^{-\lambda y} & y \ge 2 \end{cases}.$$

$F_Y(y)$在$y = 2$处不连续，故不是连续型随机变量。

§ 2.4 随机变量函数的分布

在许多实际问题中，不仅要研究随机变量的分布问题，还要研究随机变量函数的分布问题.

即已知随机变量X的分布，要求随机变量X的函数Y=g(X)的分布，实际上也就是求随机变量Y的分布.

例如，设X是物体的速率，而Y是物体的动能，则$Y=\frac{1}{2}mX^2$（其中m是物体的质量）是X的函数.

一、离散型随机变量函数的分布

- 设X的分布列为：

X	x_1	x_2	x_3	$\cdots$
P	p_1	p_2	p_3	$\cdots$

- Y=g(X)为X的函数，则Y的分布列为：

Y	$g(x_1)$	$g(x_2)$	$g(x_3)$	$\cdots$
P	p_1	p_2	p_3	$\cdots$

二、连续型随机变量函数的分布

- 设X是连续型随机变量,密度函数为 p(x). Y=g(X)是X的连续函数，则Y也是一个连续型随机变量。计算它的分布函数的**一般方法**是：
- 先求Y的分布函数$F_Y(y)$，

$$F_Y(y) = P(g(X) \le y) = \int_{\{x|g(x)\le y\}} p_X(x)dx$$

然后再通过求导得出Y的密度函数$p_Y(y)$。

- **定理:** 设X是一个连续型随机变量，其密度函数为p(x)，取值于（a，b），又函数y=g(x)在（a，b）上是严格单调的可导函数，则Y=g(X)也是一个连续型随机变量，且其密度函数为

$$p_Y(y)=\begin{cases} p_X[h(y)]\cdot\left|h'(y)\right|, & \alpha<y<\beta \\ 0, & \text{其它} \end{cases}$$

其中α=min{g(a),g(b)},β=max{g(a),g(b)}，h(y)是y=g(x)的反函数。

- 例：设X~N(μ,σ²)，则aX+b~N(aμ+b,(aσ)²)
- 解：y=ax+b是严格单调函数，于是

$$p_Y(y) = p_X[h(y)]\cdot\left|h'(y)\right| = p_X\left(\frac{y-b}{a}\right)\cdot\left|\frac{1}{a}\right|$$

$$= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}} \left|\frac{1}{a}\right| = \frac{1}{\sqrt{2\pi}\sigma\,|a|} e^{-\frac{(y-a\mu-b)^2}{2(a\sigma)^2}}$$

- **例**：设随机变量X的密度函数为

$$f_X(x)=\begin{cases}\frac{1}{2} & -1<x<0,\\ \frac{1}{4} & 0<x<2,\\ 0 & \text{其它}.\end{cases}$$

求$Y=X^2$的密度函数。

$(0,4]$,

解:

$$F_Y(y)=\begin{cases}0 & y\le 0,\\ \frac{3}{4}\sqrt{y} & 0<y\le 1,\\ \frac{1}{2}+\frac{1}{4}\sqrt{y} & 1<y\le 4,\\ 1 & y>4.\end{cases}$$

因此

$$f_Y(y)=\begin{cases}\frac{3}{8\sqrt{y}} & 0<y\le 1,\\ \frac{1}{8\sqrt{y}} & 1<y\le 4,\\ 0 & \text{其他}.\end{cases}$$

- 例：设X服从N(0,1),求Y=X²的密度函数。
- 解：

$$\text{当}y>0,\ F_Y(y)=P(Y\le y)=P(X^2\le y)=P(-\sqrt{y}\le X\le\sqrt{y})$$

$$=\int_{-\sqrt{y}}^{\sqrt{y}}\phi(x)dx=2\int_{0}^{\sqrt{y}}\phi(x)dx,$$

$$p_Y(y)=F_Y{}'(y)=\begin{cases}\dfrac{1}{\sqrt{2\pi}}y^{-1/2}e^{-y/2} & \text{当}y>0\\ 0 & \text{当}y\le 0\end{cases}\text{。}$$