回 顾

- 定义：几何概型、条件概率、随机事件独立性、
  样本空间的划分；
- 模型：乘法定理、全概率公式、贝叶斯公式；
- 定理：独立性定理1、独立性定理2；
- 伯努利试验，n重伯努利试验；

第二章 随机变量及其分布

- 2.1 随机变量的概念
- 2.2 离散型随机变量及其分布律
- 2.3 随机变量的分布函数
- 2.4 连续型随机变量的概率密度
- 2.5 随机变量的函数分布

2.1 随机变量的概念

- 例 1. 从一批产品中任意抽取 k 件，观察出现的“废品数” X_1，依试验结果不同，X_1 的所有可能为: 0, 1, 2, …, k 等 $k+1$ 个结果，可用 $\{X_1 = j,\quad j = 0, 1, 2, \ldots, k\}$ 表示.
- 例 2. 记录某接待站一天中来访的人数 X_2，依试验结果不同，X_2 的所有可能取值为: 0, 1, 2, …. “接待 k 个人” 可用 $\{X_2 = k\}$ 表示.

例3 **测试灯泡寿命的试验中,随不同的试验,"灯泡寿命"** X_3 **可以取所有非负的实数值,"灯泡寿命为** t **小时"可以用** $\{X_3=t\}$ **来表示.**

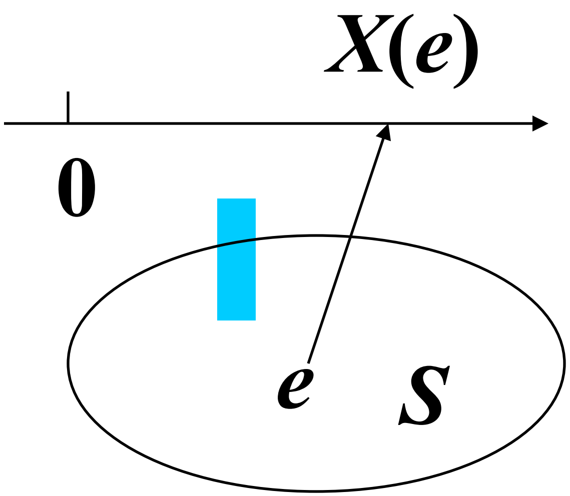
例4 掷一枚硬币观察正反面.试验结果为:e_1={正面}, e_2={反面}.试验的结果可以用变量 X_4 表示.

$$X_4 = X_4(e) = \begin{cases} 1, \text{当} e = e_1 \\ 0, \text{当} e = e_2 \end{cases}$$

对于试验 E 引进变量 $X(e)$，它定义在 S 上，依试验结果 e 不同取不同实值，$X(e)$ 取的不同实值也与不同试验结果对应.由于试验结果 e 发生是随机的，故称 $X(e)$ 为随机变量.

➢ 定义2.1 如果对于样本空间中每个样本点 e，都有唯一的一个实数 $X(e)$ 与之对应，则称 $X(e)$ 为随机变量.
简记 $X(e)$为 X.

- “废品数大于10”可用$\{X_1>10\}$表示;
- “来访人数为5” 用$\{X_2=5\}$表示.
- “灯泡寿命在3000到5000小时之间用$\{3000<X_3<5000\}$表示;
- “硬币出现正面” 用 $\{X_4=1\}$表示.



➢ 若 L 是一个实数集合，将 X 在 L上取值写成 $\{X \in L\}$；

它表示事件$B = \{e \mid X(e) \in L\}$，即$B$ 是由S 中使得$X(e) \in L$ 的所有样本点e 所组成的事件，此时有 $P\{X \in L\} = P(B) = P\{e \mid X(e) \in L\}$.

2. 分类：(1) 离散型随机变量;
(2) 连续型随机变量；
(3) 其他.

➢ 目的：用随机变量表示试验发生的结果以及事件，比较方便，并且可以进行各种数学运算；通过随机变量来研究随机试验，全面揭示随机现象的统计规律.

2.2 离散型随机变量及其分布律

- **定义：** 若随机变量全部可能取到的值是有限多个或可列无限多个，则称为**离散型随机变量**。

离散型 random variable (r.v.)的分布律：

设离散型 r.v.X 所有可能取值为$x_k\ (k=1,2,3,\ldots)$

$$P\{X=x_k\}=p_k, \quad k=1,2,\ldots \qquad (1)$$

则称式(1)为离散型r.v.X的**分布律**或 **概率分布**.

2.2 离散型随机变量及其分布律

设离散型 r.v.X 所有可能取值为$x_k\,(k=1,2,3,\ldots)$

$$P\{X=x_k\}=p_k,\quad k=1,2,\ldots \qquad (1)$$

式(1)也可用表格形式表示:

X	x_1	x_2	$\cdots$	x_n	$\cdots$
p_k	p_1	p_2	$\cdots$	p_n	$\cdots$

分布律的性质：

(1) $p_k \geq 0,\ k = 1,2,\ldots$

(2) $\sum_{k=1}^{\infty} p_k = 1.$

例1. 设一汽车在开往目的地的道路上需经过四盏信号灯，每盏信号灯以概率 p 禁止汽车通过， 以 X 表示汽车首次停下时已通过信号灯的盏数， 求 X 的分布律. (设各信号灯的工作是相互独立的).

解:

X	0	1	2	3	4
p_k	p	$(1-p)p$	$(1-p)^2p$	$(1-p)^3p$	$(1-p)^4$

即 $P\{X=k\}=(1-p)^k p,\quad k=0, 1, 2, 3.$

$$P\{X=4\}=(1-p)^4$$

例2. 袋中装有 3 只白球和 2 只红球，从袋中任取两球，用 X 表示取到的白球数，则 X 是一取值为0, 1, 2的离散型随机变量，其分布律为

解：

$$P\{X=0\}=\frac{C_2^2}{C_5^2}=\frac{1}{10},\qquad P\{X=1\}=\frac{C_3^1\cdot C_2^1}{C_5^2}=\frac{3}{5}$$

$$P\{X=2\}=\frac{C_3^2}{C_5^2}=\frac{3}{10}$$

或

X	0	1	2
p_k	1/10	3/5	3/10

例3 设X的分布律为$P\{X=k\}=\dfrac{1}{3^n}\cdot C_n^k\cdot a^k \quad (k=0,1,2,\cdots,n)$,
试确定常数a.

解：由分布律的性质可得

$$1=\sum_{k=0}^{n}P\{X=k\}=\sum_{k=0}^{n}\frac{1}{3^n}C_n^k a^k$$

$$=\sum_{k=0}^{n}C_n^k\left(\frac{a}{3}\right)^k\left(\frac{1}{3}\right)^{n-k} \quad =\left(\frac{a}{3}+\frac{1}{3}\right)^n$$

故 $a=2.$

几种重要的离散型 r.v. 的分布律

（一） 0-1分布

设随机变量 X 只可能取 0 和 1 两个数值，其分布为

$P\{X=1\}=p$， $P\{X=0\}=1-p$；

或表为

X	0	1
p_k	$1-p$	p

其中$0<p<1$, 则称 X 服从 (0-1) 分布或两点分布。

设 A 是随机事件，$P(A)=p$ $(0<p<1)$，记: $X=\begin{cases}1 & A\text{发生}\\ 0 & \overline{A}\text{发生}\end{cases}$，

则 X 服从 (0-1)分布.

- 一般地，若某随机试验 E 只有两个可能的结果，这种随机试验就可用 0-1 分布来描述.

  - 如：产品是否合格，射击是否命中，婴儿的性别等；

(二) 二项分布

定义： 设试验E只有两个可能结果A 与$\overline{A}$, 且$P(A)=p \quad (0<p<1)$, 将试验E 独立重复地进行n 次, 这样的试验称为n 重伯努利(Bernoulli)试验.

➢ 若以 X 表示 n 重伯努利试验中事件 A 发生的次数，则 X 为取值 0, 1, …, n 的离散随机变量，且分布律为：

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

称 r.v. X 服从参数为 n, p 的二项分布， 记为: $X \sim b(n,p)$.

(当 $n=1$ 时, $b(1,p)$就是$(0-1)$分布)

- n 重伯努利试验中事件 A 出现次数 X 的分布律，以 A_i 表示第 i 次试验中 A 发生，以 $\overline{A_i}$ 表示第 i 次试验中 A 不发生，令 $B_k=$有k次A发生

上式右边共有C_n^k个，且两两互斥. 由试验独立性

$$P(A_1\cdots A_k\overline{A}_{k+1}\cdots\overline{A}_n)=P(A_1)\cdots P(A_k)P(\overline{A}_{k+1})\cdots P(\overline{A}_n)=p^kq^{n-k}$$

再由概率可加性得：

$$P(B_k)=C_n^k p^k q^{n-k} \qquad (k=0,1,2,\cdots,n)$$

例4. 某种电子元件的使用寿命超过1500小时为一级品， 已知一大批该产品的一级品率为 0.2， 从中随机抽查 20 只， 求这20只元件中一级品只数 X 的分布律.

解: 检查一只元件看是否为一级品可以看作是 一次试验，抽查20只元件可以看作20次试验；

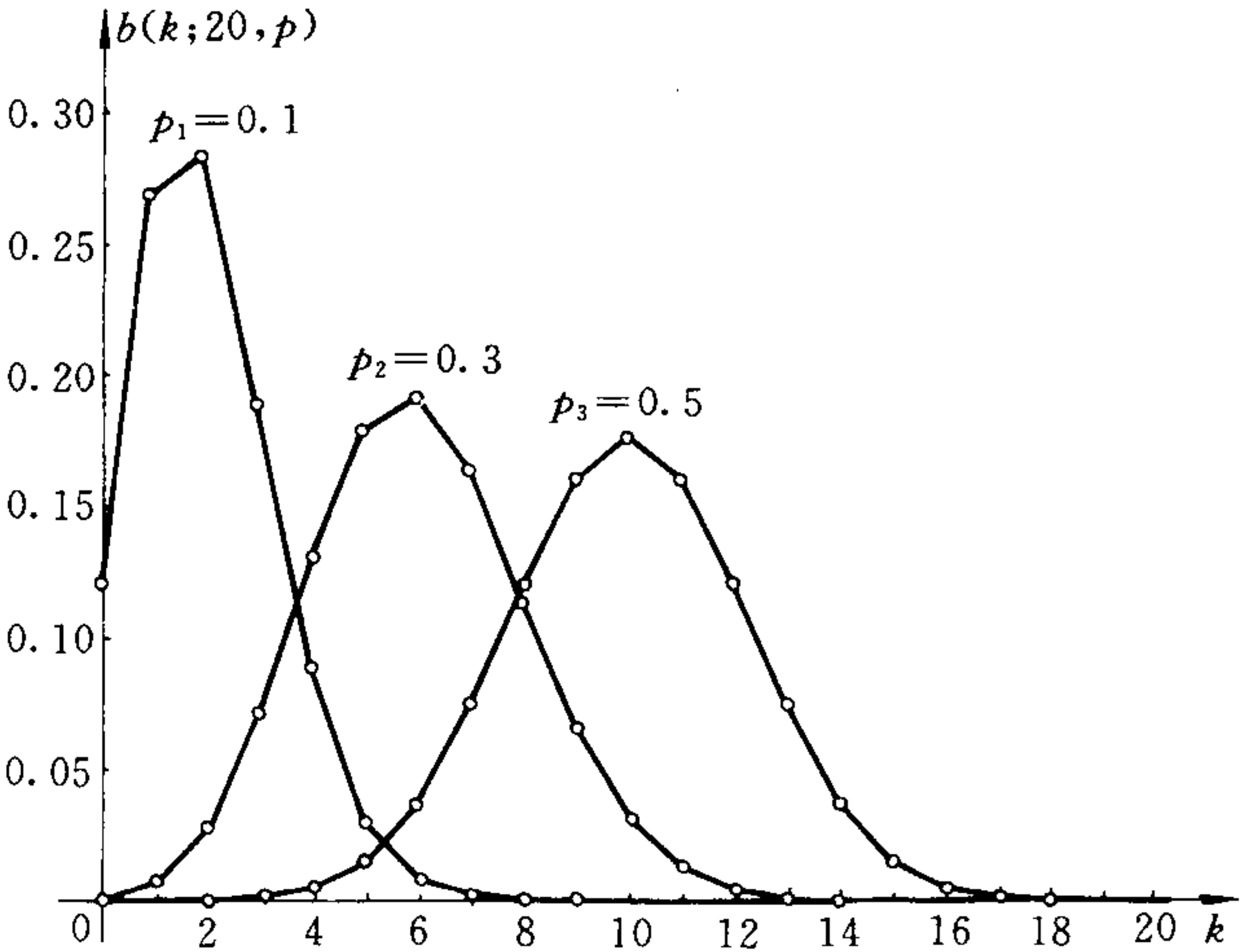
又因这批元件总数很大，不放回抽样可以近似看作放回抽样处理，所以这是20重伯努利试验，则$X \sim b(20, 0.2)$.

$$P\{X=k\} = C_{20}^{k}(0.2)^{k}(0.8)^{20-k},\ k=0,\ 1,\ 2,\ \ldots,\ 20.$$

k	0	1	2	3	4
p	0.0115	0.0576	0.1369	0.2054	0.2182
k	5	6	7	….	20
p	0.1745	0.1091	0.0546	….	1.048E-14

- 从上表可以看出，当 k 从 0 到 20 变化时，对应的概率值先变大，后变小；
- 其实，这个概率规律是所有二项分布共有的性质，这就需要求出具有最大概率的项；

二项分布图



b(k;20,p)
0.30
0.25
0.20
0.15
0.10
0.05
p₁=0.1
p₂=0.3
p₃=0.5
0
2
4
6
8
10
12
14
16
18
20
k

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

$$=1+\frac{(n+1)p-k}{kq}$$

于是： 当$k<(n+1)p$时，$P\{X=k\}>P\{X=k-1\}$

当$k=(n+1)p$时，$P\{X=k\}=P\{X=k-1\}$

当$k>(n+1)p$时，$P\{X=k\}<P\{X=k-1\}$

- 所以有:

(1) 当 $(n+1)p$ 为整时，$P\{X=k\}$在 $k=(n+1)p$ 和 $k=(n+1)p-1$ 处同时达到最大。

(2) $(n+1)p$ 非整时，$P\{X=k\}$在 $k=[(n+1)p]$ 处达到最大值。

$$P\{X=k\} = C_{20}^{k}(0.2)^{k}(0.8)^{20-k},\ k=0,\ 1,2,\ldots,20.$$

如上例中, $(n+1)p=(20+1)\times 0.2=4.2$, 所以 $P\{X=k\}$ 有最大值 $P\{X=4\}=0.2182$，即最有可能抽到4个一级品.

使得 P$\{X=k\}$ 达到最大值的数 k 称为最可能成功的次数。

例5. 某人进行射击, 每次命中率为0.02, 独立射击400次, 试求至少击中两次的概率.

解： 设400次射击中击中的次数为X, 则$X \sim b(400, 0.02)$.

$$P\{X=k\}=C_{400}^{k}(0.02)^{k}(0.98)^{400-k}, \quad k=0, 1, \ldots, 400.$$

则 $P\{X \geq 2\} = 1 - P\{X=0\} - P\{X=1\}$

$$=1-(0.98)^{400}-400\times(0.02)\times(0.98)^{399}.$$

- 当 n 较大, p又较小时, 二项分布的计算比较困难, 例如 0.98^{400}， 0.02^{400}, …, 可以用后面的 Poisson 分布近似计算.

(三) 泊松分布(Poisson)

若离散随机变量 X 的分布为

$$P\{X=k\}=\frac{\lambda^k \mathrm{e}^{-\lambda}}{k!},\quad k=0,\ 1,\ 2,\ \dots \qquad ,$$

其中$\lambda>0$是常数, 则称 X 服从参数为 λ 的泊松分布,

记为$X\sim\pi(\lambda)$.

可以验证: $$\sum_{k=0}^{\infty}P\{X=k\}=\sum_{k=0}^{\infty}\frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}=\mathrm{e}^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=1$$

$$e^{\lambda}=\lim_{n\to\infty}\left(1+\frac{\lambda}{n}\right)^n=\lim_{n\to\infty}\sum_{k=0}^{n}C_n^k\left(\frac{\lambda}{n}\right)^k=\lim_{n\to\infty}\sum_{k=0}^{n}\frac{A_n^k}{n^k}\cdot\frac{\lambda^k}{k!}$$

例如, 一定时间间隔内电话交换台收到的呼唤次数; 一本书的印刷错误数; 某一地区一个时间间隔内发生的交通事故数等都服从泊松分布.

二项分布可用泊松分布来近似, 有如下定理:

泊松 (Poisson) 定理: 设随机变量序列$\{X_n\}$, $X_n \sim b(n, p_n)$,

且 $np_n = \lambda > 0$为常数, k 为任一固定的非负整数，则

$$\lim_{n\to\infty} P\{X_n = k\} = \lim_{n\to\infty} C_n^k p_n^k (1-p_n)^{n-k} = \frac{\lambda^k e^{-\lambda}}{k!},$$

证明:　由 $np_n=\lambda, p_n=\lambda/n$. 于是

$$P\{X=k\}=C_n^k p_n^k(1-p_n)^{n-k}$$

$$=\frac{n(n-1)\ldots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$=\frac{\lambda^k}{k!}\left[1\cdot\left(1-\frac{1}{n}\right)\cdot\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\right]\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}$$

$$\to\frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}.(n\to\infty)$$

泊松定理的意义：

1. 在定理的条件下，二项分布的极限分布是 泊松分布.

2. 当 n 很大且 p 又较小时，这就是二项分布的概率近似计算公式 .

$$C_n^k p^k (1-p)^{n-k} \approx \frac{\lambda^k e^{-\lambda}}{k!}, \quad \text{其中}\lambda = np,$$

例6. 某人进行射击, 每次命中率为0.02, 独立射击400次, 试求至少击中两次的概率.

解： 设400次射击中击中的次数为X, 则$X \sim b(400, 0.02)$.

$$P\{X = k\} = C_{400}^{k}(0.02)^{k}(0.98)^{400-k}, \quad k = 0, 1, \dots, 400.$$

则 $P\{X \geq 2\} = 1 - P\{X = 0\} - P\{X = 1\}$

在例3中, $X \sim b(400, 0.02)$, $\lambda = np = 400 \times 0.02 = 8$,

$$P\{X \geq 2\} = 1 - P\{X = 0\} - P\{X = 1\}$$
$$= 1 - (0.98)^{400} - 400 \times (0.02) \times (0.98)^{399}$$
$$\approx 1 - e^{-8} - 8e^{-8} \approx 0.997.$$

例7. 设有同类型设备300台, 各台工作是相互独立的, 发生故障的概率都是0.01, 设一台设备的故障由一个人处理, 问：至少需配备多少工人，才能保证当设备发生故障但不能及时维修的概率小于 0.01?

解: 设需配备 N 个工人，记同一时刻发生故障的设备台数为 X;

"设备发生故障不能及时维修" 的数学表达式为:

"$X>N$", 即求最小的N, 使得 $P\{X>N\}<0.01$.

其中 n=300，p=0.01，λ=np=3，由泊松近似公式

$$P\{X > N\} = \sum_{k=N+1}^{300} C_{300}^{k} (0.01)^{k} (0.99)^{300-k}$$

$$\approx \sum_{k=N+1}^{300} \frac{3^{k} \mathrm{e}^{-3}}{k!} < 0.01.$$

查$Poisson$ 分布表可知, 当 $N \geq 8$ 时，有$P\{X \leq N\} \geq 0.99$

故最小的$N = 8$

(四) 几何分布

进行重复独立试验, 设每次试验成功的概率为p, 失败的概率为 $1-p=q$ $(0<p<1)$, 将试验进行到出现一次成功为止, 以 X 表示所需的试验次数, 则X 的分布律为: $P\{X=k\}=q^{k-1}p$, $k=1, 2, \ldots$

称为 X 服从参数为 p 的几何分布.

例. 设某种社会定期发行的奖券，每券 1 元，中奖率为 $p=0.0001$, 某人每次购买1张奖券, 如果没有中奖下次继续买，直到中奖止, 求购买次数 X 的分布律.

解： $P\{X=k\}=p(1-p)^{k-1}$, $k=1, 2, 3, \ldots$

$$P\{X>1000\}=p\sum_{k=1001}^{\infty}q^{k-1}=q^{1000}=0.9047$$

- 对某批 N 件产品进行有放回抽样调查，若产品中有 M 件次品，现从整批产品中随机抽取 n (n<=N-M) 件产品，则在这 n 件产品中出现的次品数 X 是一个所有可能取的值为0 , 1, 2, …, l（其中 l=min{M, n}) 的离散型随机变量，其分布律为：

$$P(X=k)=C_n^k\left(\frac{M}{N}\right)^k\left(1-\frac{M}{N}\right)^{n-k}$$

（五）超几何分布

对某批 N 件产品进行无放回抽样调查，若产品中有 M 件次品，现从整批产品中随机抽取 n ($n<=N-M$) 件产品，则在这 n 件产品中出现的次品数 X 是一个所有可能取的值为0 , 1, 2, …, l（其中 l=min{M, n}）的离散型随机变量，其分布律为：

$$P\{X=k\}=\frac{C_M^k \cdot C_{N-M}^{n-k}}{C_N^n}(k=0,1,2,\cdots,l)$$

这个分布称为 超几何分布；

（六）负二项分布

若随机变量 X 的分布律为 $P\{X=k\}=C_{k-1}^{r-1}p^r(1-p)^{k-r}\quad k=r,r+1,\cdots$

其中，$0<p<1$，则称 X 服从负二项分布.

- 若令 X 表示贝努里试验序列中事件 A 第 r 次出现时所需要的试验次数，则 X 服从负二项分布.
  - 如第 r 次击中目标时所射击的次数是 k 的概率；
  - 投掷硬币试验中第 r 次掷出正面时投掷的次数是 k 的概率；

3.3 随机变量的分布函数

➢ 对于非离散型 r. v. 已不能用分布律来描述它， 需要考虑 r.v. 的值落入一个区间的概率， 如 $P\{x_1<X\leq x_2\}$, $P\{X\leq x\}$等，为此引入随机变量的分布函数.

1. 定义：设 r.v. X, $x\in R^1$, 则 $F(x)=P\{X\leq x\}$称为 X 的分布函数.

(1) $P\{x_1<X\leq x_2\}=P\{X\leq x_2\}-P\{X\leq x_1\}$
$=F(x_2)-F(x_1)$.

(2) 无论是离散型 r.v. 还是非离散型 r.v.， 分布函数都可以描述其统计规律性.

2. 性质:

(1) $F(x)$是单调不减函数. (单调性)

$\forall x_2>x_1$, $F(x_2)-F(x_1)=P\{x_1<X\leq x_2\}\geq 0$.

(2) $0\leq F(x)\leq 1$ 且 (规范性)

$$F(-\infty)=\lim_{x\to-\infty}F(x)=0,\quad F(+\infty)=\lim_{x\to+\infty}F(x)=1$$

2. 性质:

(3) $F(x)$至多有可列个间断点，而在其间断点上也是右连续的，$F(x+0)=F(x)$， 即在间断点 x_0 处，有

$$\lim_{x\to x_0^+} F(x) = F(x_0+0) = F(x_0) \quad \text{(右连续性)}$$

说明: 性质 (1), (2), (3) 是鉴别一个函数 $F(x)$ 是否为某随机变量 X 的分布函数的充要条件.

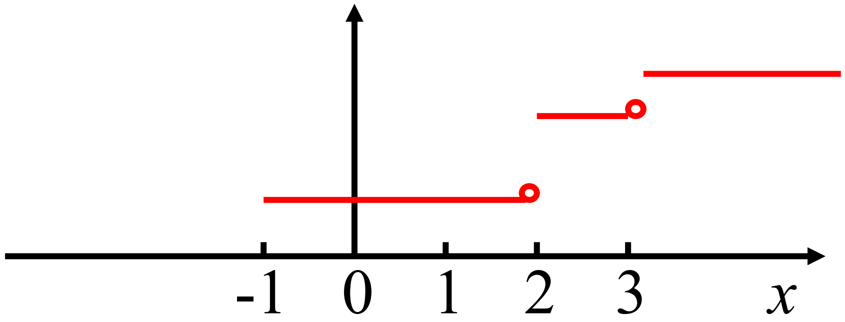
例1. 离散型 r.v., 已知分布律求其分布函数.

X	-1	2	3
p_k	1/4	1/2	1/4

求： X 的分布函数, 并求$P\{X\le 1/2\}$, $P\{3/2 < X \le 5/2\}$.

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

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



$P\{X \le 1/2\} = F(1/2) = 1/4$ 或由分布律直接得：

$P\{X \le 1/2\} = P\{X = -1\} = 1/4$,

$P\{3/2 < X \le 5/2\} = F(5/2) - F(3/2) = 1/2$.

若离散型随机变量 X 的分布律为：$P\{X=x_k\}=p_k \quad (k=1,2,\cdots)$

则 X 的分布函数为 $F(x)=P\{X\le x\}=P\{\bigcup_{x_k\le x}\{X=x_k\}\}=\sum_{x_k\le x}P\{X=x_k\}$

即 $F(x)=\sum_{x_k\le x}p_k$

- 离散型 r.v. X 的分布函数 $F(x)$ 的图形呈阶梯状(如上例), x_1=-1，x_2=2，x_3=3 都是**第一类间断点**（跳跃式的)；
- $F(x)$的图形在这些点处都有一个跃度，在 x_k 处的跃度就是 X 取值 x_k 的概率 p_k.

3.4 连续型随机变量的概率密度

1. 定义：对于$r.v.X$ 的分布函数$F(x)$，存在非负函数$f(x)$，使对于任意的实数 x，有 $F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$

则称 X 为连续型 r.v.，其中函数 $f(x)$ 称为 X 的概率密度函数，简称概率密度.

> ➢ 连续型 r.v. 的分布函数是连续函数，这种 r.v. 的取值是充满某个区间的.

2. 概率密度$f(x)$的性质: (1) $f(x)\geq 0$. (2) $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$.

(3) $P\{x_1 < X \le x_2\} = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)\mathrm{d}x,$

$(x_1 \le x_2)$

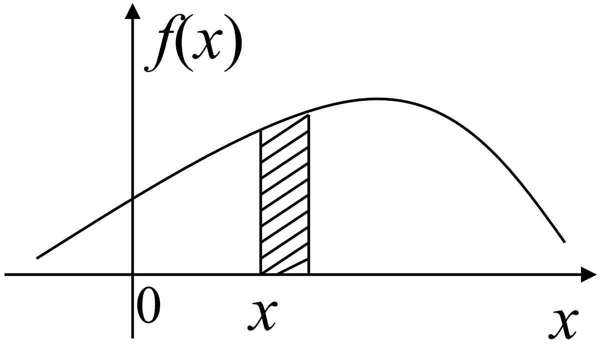
(4) 若 $f(x)$ 在点 x 处连续，则有$F'(x) = f(x)$

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

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

上式可知当 Δx 很小时，有

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



当 $\Delta x \to 0$，？

3. 关于连续型 r.v. 的一个重要结论

定理: 设 X 为连续型 r.v., 它取任一指定的实数值 a 的概率均为0.
即 $P\{X=a\}=0.$

证明: 设X 的分布函数为$F(x)$, $\Delta x>0$

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

令 $\Delta x \to 0$, 由$F(x)$ 的连续性可得 $P\{X=a\}=0.$

所以 $$P\{x_1 < X < x_2\} = P\{x_1 \le X \le x_2\}$$
$$= P\{x_1 \le X < x_2\} = P\{x_1 < X \le x_2\}$$
$$= F(x_2)-F(x_1)$$

例1. 设随机变量X 具有概率密度 $f(x)$，试确定常数k，并求分布函数$F(x)$ 和 $P\{X>0.1\}$.

$$f(x)=\begin{cases} k\mathrm{e}^{-3x}, & x>0, \\ 0, & x\le 0, \end{cases}$$

例1. 设随机变量X 具有概率密度 $f(x)$，试确定常数k，并求分布函数$F(x)$ 和 $P\{X>0.1\}$.

$$f(x)=\begin{cases} k\mathrm{e}^{-3x}, & x>0, \\ 0, & x\le 0, \end{cases}$$

解：由 $1=\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x=\int_0^{+\infty} k\mathrm{e}^{-3x}\,\mathrm{d}x=k/3$ 得 $k=3.$

当$x>0$时，$F(x)=\int_{-\infty}^{x} f(x)\mathrm{d}x=\int_0^x 3\mathrm{e}^{-3x}\mathrm{d}x=1-\mathrm{e}^{-3x}$

于是分布函数 $F(x)=\begin{cases} 1-\mathrm{e}^{-3x}, & x>0 \\ 0, & x\le 0 \end{cases}$

$$P\{X>0.1\}=\int_{0.1}^{+\infty} 3\mathrm{e}^{-3x}\mathrm{d}x=\mathrm{e}^{-0.3}=0.7408$$

例2. 连续型随机变量 X 的分布函数为 $F(x)$，

$$F(x)=\begin{cases} A+Be^{-\frac{x^2}{2}}, x>0 \\ 0, x\le 0 \end{cases},$$

求：(1) A, B ；(2) 概率密度函数；(3) 落入 $(1,2)$ 的概率；

解: (1) 因为 $F(+\infty)=1$, 故有 $\lim\limits_{x\to\infty}(A+Be^{-\frac{x^2}{2}})=1$, 因此 $A=1$.

又因 $\lim\limits_{x\to 0^-}F(x)=0=\lim\limits_{x\to 0^+}F(x)=A+B$,

所以有 $B=-A=-1$ 于是 $F(x)=\begin{cases} 1-e^{-\frac{x^2}{2}}, x>0 \\ 0, x\le 0 \end{cases}$

$$F(x)=\begin{cases}1-\mathrm{e}^{-\frac{x^2}{2}}, x>0\\ 0, x\le 0\end{cases}$$

(2) 对 $F(x)$ 求导，得 x 的概率密度为

$$f(x)=F'(x)=\begin{cases}x\mathrm{e}^{-\frac{x^2}{2}}, x>0\\ 0, x\le 0\end{cases}$$

(3) X 落入 (1, 2) 内的概率为

$$P\{1<X<2\}=\int_1^2 x\mathrm{e}^{-\frac{x^2}{2}}\mathrm{d}x(=F(2)-F(1))$$

$$=\mathrm{e}^{-\frac{1}{2}}-\mathrm{e}^{-2}=0.4721.$$

连续型随机变量的概率密度

1. **定义**：对于$r.v.X$ 的分布函数$F(x)$，存在非负函数$f(x)$，使对于任意的实数 x, 有 $F(x)=\int_{-\infty}^{x} f(t)\mathrm{d}t$

则称 X 为连续型 r.v.，其中函数 $f(x)$ 称为 X 的概率密度函数, 简称概率密度.

4. 几个常用的连续型 r.v. 分布

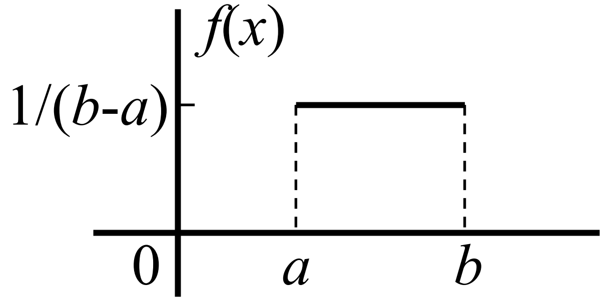
常见的连续型分布有 均匀分布，指数分布，正态分布，伽玛分布等.

(一) 均匀分布:

设随机变量X在区间$[a,b]$上取值，且概率密度为:

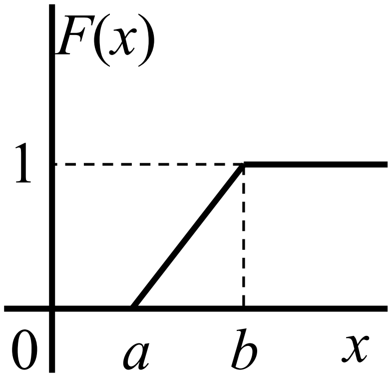
$$f(x)=\begin{cases}\dfrac{1}{b-a}, & a<x<b\\ 0, & \text{其它}\end{cases}$$

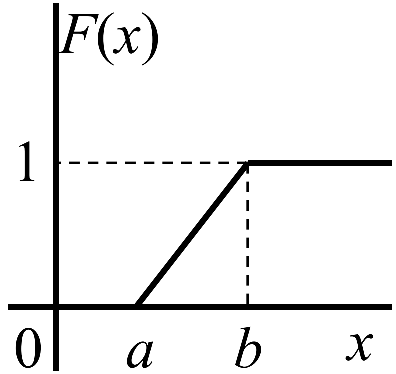
则称随机变量 X 在 (a, b) 上服从均匀分布, 记作 $X\sim U(a, b)$.



其分布函数为:

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





若$X \sim U(a,b)$，且$(c,c+d) \subset (a,b)$，则

$$P\{c < X \leq c+d\} = \int_c^{c+d} \frac{1}{b-a} \mathrm{d}x = \frac{d}{b-a},$$

➢ 此概率与子区间长度成正比，而与子区间的起点无关，这也是均匀分布的由来.

(二) 指数分布:

1. 定义: 如果连续型随机变量 X 的概率密度为:

$$f(x)=\begin{cases}\dfrac{1}{\theta}\mathrm{e}^{-x/\theta}, x>0, \\ 0, \qquad x\le 0,\end{cases}, \theta>0.$$

则称X 服从参数为θ 的指数分布，记为$X\sim e(\theta)$

有些书上称 $\lambda=1/\theta$ 时上述的概率分布是参数为λ 的指数分布。

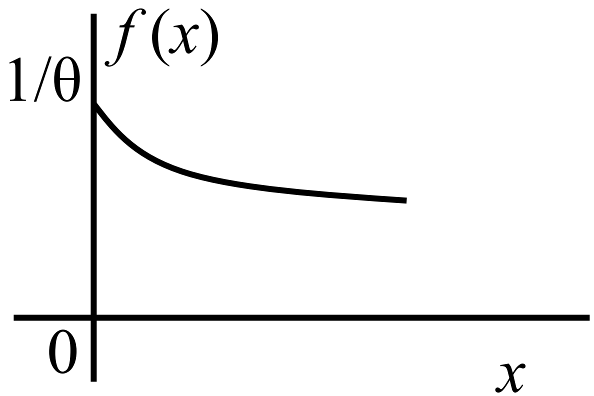
(二) 指数分布:

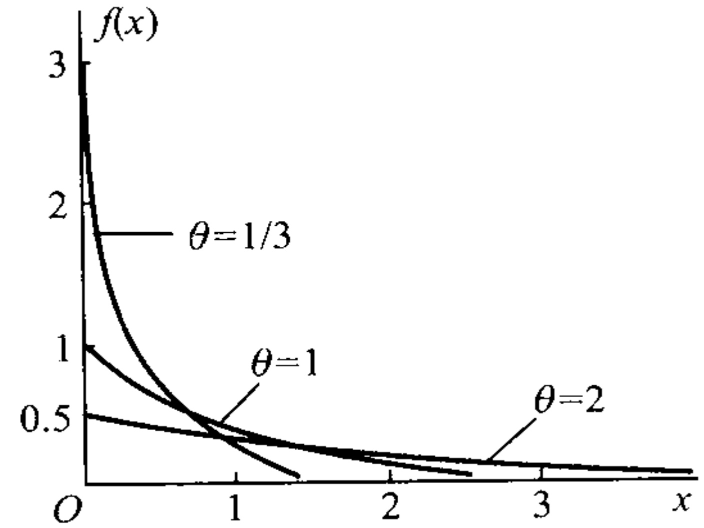
X 的分布函数为：$F(x)=\begin{cases}1-\mathrm{e}^{-x/\theta}, x>0\\ 0, \quad x\le 0\end{cases}$

- 指数分布常用来做各种寿命分布的近似分布，如电子元件，动物寿命等，通话时间，随机服务时间也近似服从指数分布.

指数分布的概率密度：

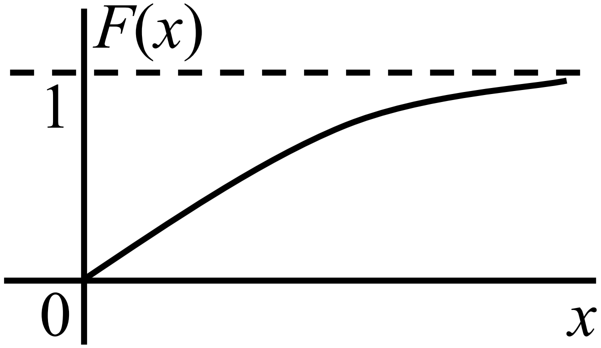
$$f(x)=\begin{cases}\dfrac{1}{\theta}\mathrm{e}^{-x/\theta}, & x>0\\ 0, & x\le 0\end{cases}$$





指数分布的分布函数:

$$F(x)=\begin{cases}1-\mathrm{e}^{-x/\theta}, x>0\\ 0, \quad x\le 0\end{cases}$$



➢ 指数分布的无记忆性:

定理：若 $X \sim e(\theta)$, 则对任意的正数 s, t 有

$$P\{X>s+t \mid X>s\}=P\{X>t\}$$

证：$$P\{X>s+t \mid X>s\}=\frac{P\{X>s+t\}}{P\{X>s\}}=\frac{1-F(s+t)}{1-F(s)}$$

$$=\frac{\mathrm{e}^{-(s+t)/\theta}}{\mathrm{e}^{-s/\theta}}=\mathrm{e}^{-t/\theta}=P\{X>t\}$$

> ➢ 无记忆性表明寿命 X 大于 s 时，再活 t 年的概率与年龄 s 无关，即寿命“无老化”,“永远年轻”.

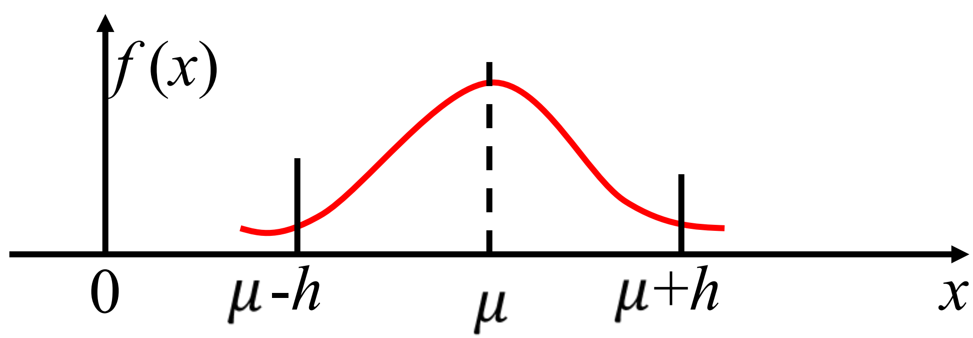
(三) 正态分布:

(1) 设随机变量X的概率密度为

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

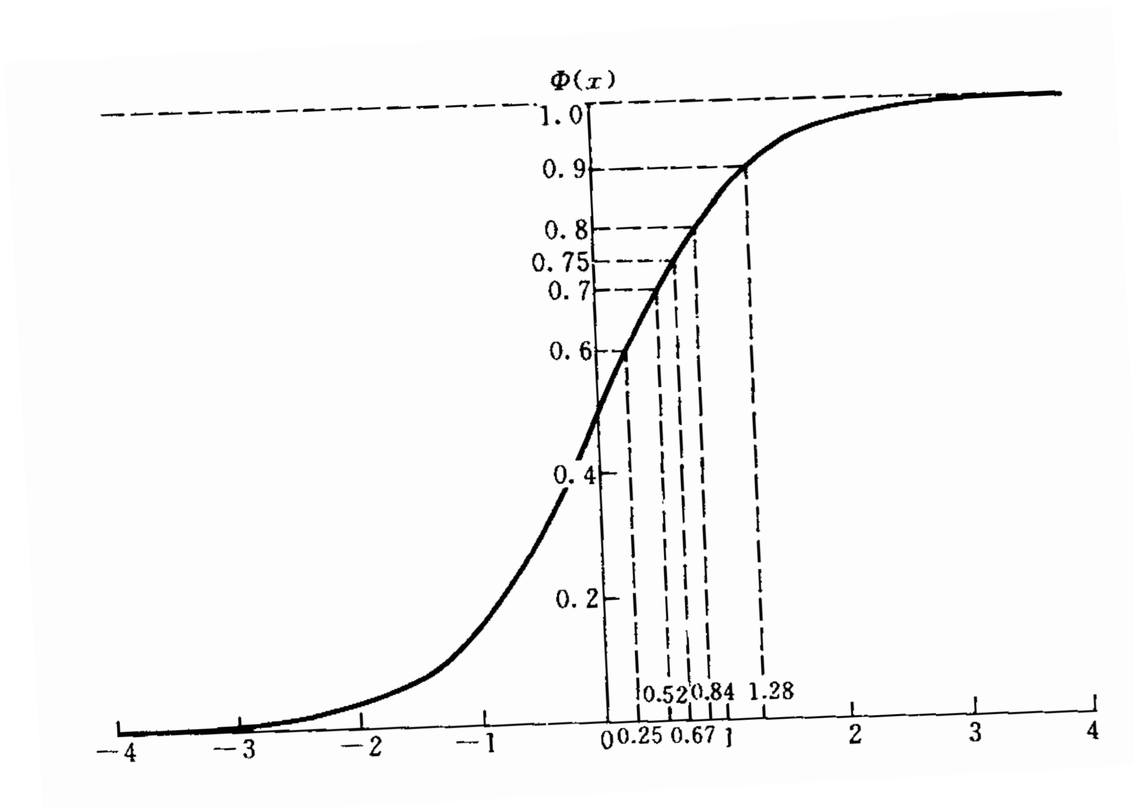
其中 $\mu,\sigma(\sigma>0)$为常数, 则称 X 服从参数为 μ,σ 的正态分布或高斯($Gauss$)分布, 记作 $X\sim N(\mu,\sigma^2)$.

其图像为:



正态分布函数为：

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



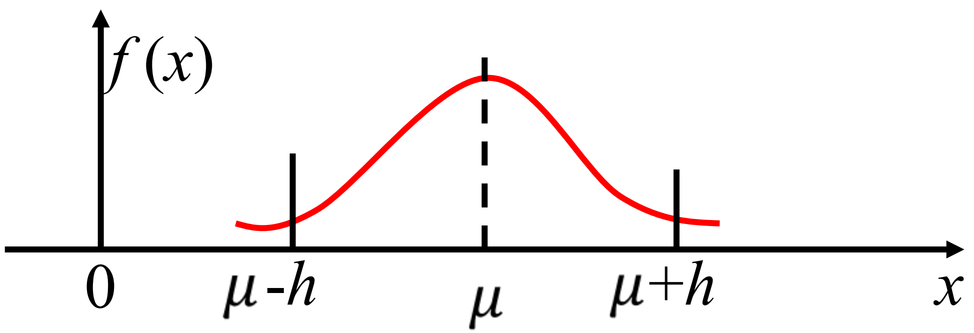
正态分布密度函数：

性质：(1) 曲线关于 $x=\mu$ 对称, 这表明对 $\forall h>0$ 有

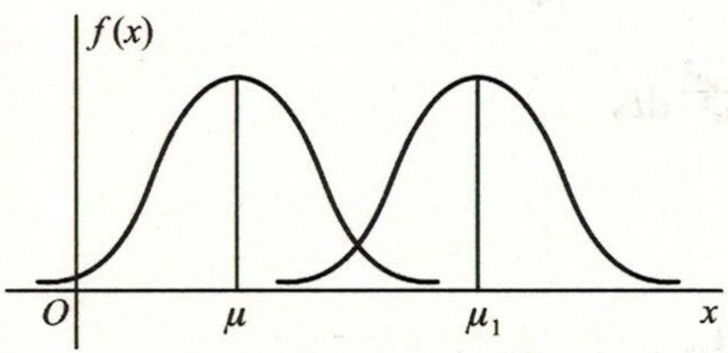
$$P\{\mu-h<X\le\mu\}=P\{\mu<X\le\mu+h\}.$$

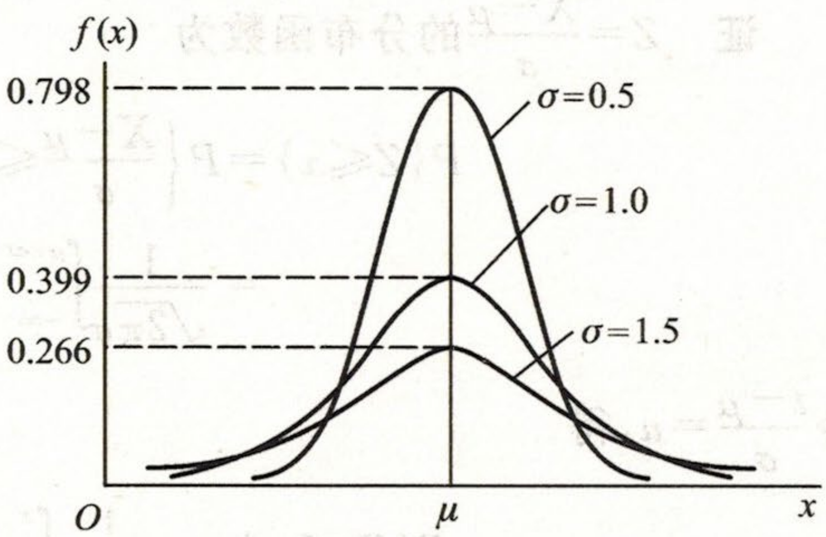
(2) 当 $x=\mu$ 时取最大值 $f(\mu)=\dfrac{1}{\sqrt{2\pi}\sigma}$.表明 X 取值在 $x=\mu$ 附近较集中.

(3) $f(x)$ 以 x 轴为渐近线.



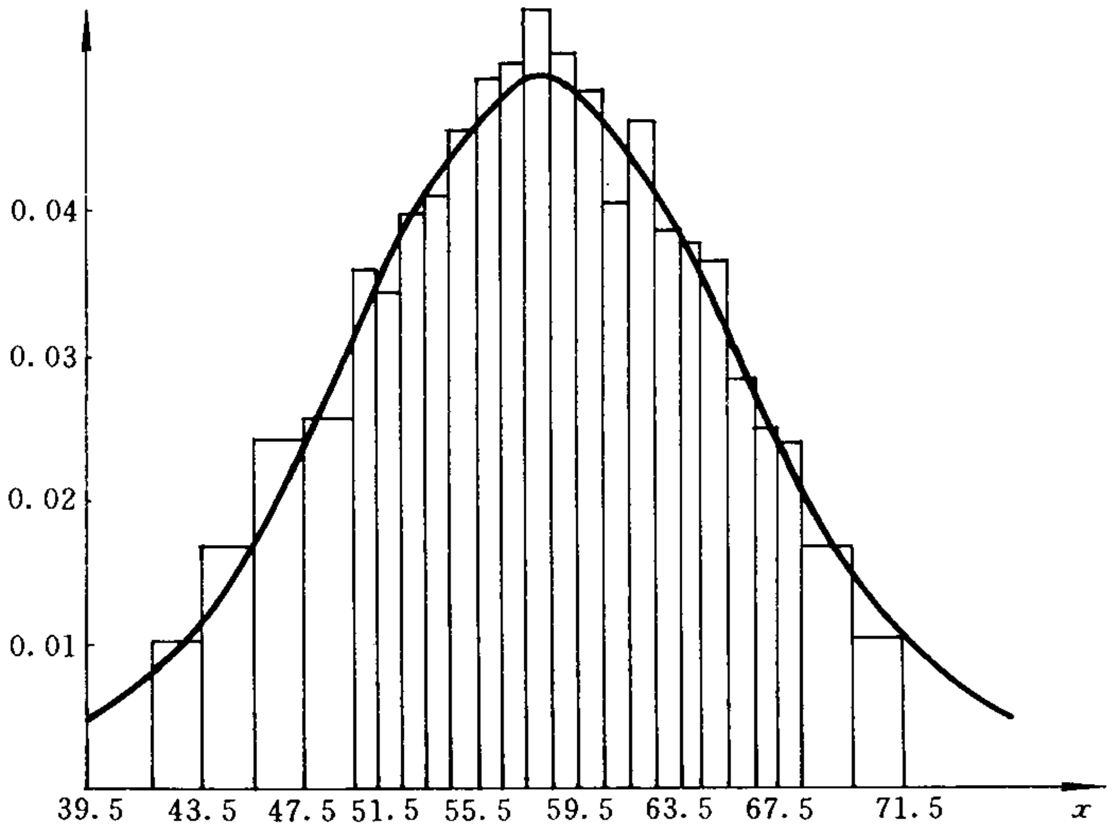
(4) $f(x)$ 的图形依赖于两个参数 μ,σ. 若固定 σ, 改变 μ, 则 $f(x)$ 的图形沿 x 轴平移而形状不变. 可见 $f(x)$ 的位置由 μ 确定, 称 μ 为位置参数. 固定 μ 改变 σ, 则最大值 $f(\mu)=\dfrac{1}{\sqrt{2\pi}\sigma}$ 改变, σ越小, $f(x)$ 越陡峭, 相反则越平坦, 故称 σ 为 $f(x)$ 的尺度参数





- 实际问题中大量随机变量服从正态分布, 如
  - 人的身高，体重；
  - 农作物的收获量；
  - 炮弹的落点 等；

- 某工厂对其生产的 A 型号零件的重量进行抽样，测量了3805个样品的重量，计算不同重量出现的频率，并画出直方图；（均值56.94，标准差8.2 的正态分布符合得相当好）



(2) 标准正态分布:

当$\mu=0,\sigma=1$时, $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$, $\quad \varPhi(x)=\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{x}\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t$,

则称X服从标准正态分布，记为$X\sim N(0,1)$.

$\varPhi(x)$即标准正态分布函数，其表已列出供查用.

(3) 对于一般正态分布，分布函数怎么求？方法是建立与标准正态分布的转换关系；

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

证明：$Z = \dfrac{X-\mu}{\sigma}$的分布函数为：

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

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

$$\overset{令 u=\frac{t-\mu}{\sigma}}{=} \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{u^2}{2}} \mathrm{d}u = \Phi(x)$$

可知 $Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$

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

对于任意区间$(x_1,x_2]$，有

$$P\{x_1<X\le x_2\}=P\left\{\frac{x_1-\mu}{\sigma}<\boxed{\frac{X-\mu}{\sigma}}\le\frac{x_2-\mu}{\sigma}\right\}$$

$$=\Phi\left(\frac{x_2-\mu}{\sigma}\right)-\Phi\left(\frac{x_1-\mu}{\sigma}\right)$$

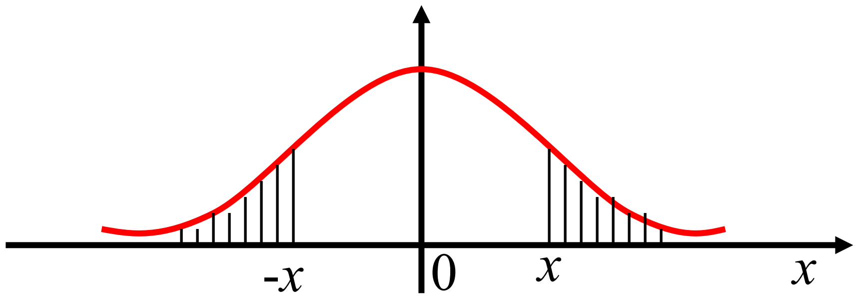
➢ **引理：**对于标准正态分布有 $\Phi(-x)=1-\Phi(x)$.

证明: 考虑 $x>0$ 的情形，由于标准正态密度 $\varphi(x)$ 是偶函数，作积分变换 u=-t, 有

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

$$=\frac{1}{\sqrt{2\pi}}\int_{x}^{+\infty}\mathrm{e}^{-\frac{u^2}{2}}\mathrm{d}u=1-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\mathrm{e}^{-\frac{u^2}{2}}\mathrm{d}u$$

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



例1: 设$X \sim N(1,4)$, 则 $P\{0 < X \le 1.6\} = ?$

$$P\{0 < X \le 1.6\} = \varPhi\left(\frac{1.6-1}{2}\right) - \varPhi\left(\frac{0-1}{2}\right)$$

$$= \varPhi(0.3) - \varPhi(-0.5)$$

$$= 0.6179 - [1 - \varPhi(0.5)]$$

$$= 0.6179 - 1 + 0.6915$$

$$= 0.3094.$$

例2. 公共汽车车门的高度是按男子与车门碰头机会在 0.01 以下设计的，设男子身高 $X \sim N(170, 6^2)$ (厘米），问车门高度应为多少？

解：设车门高度为 h, 按题意有

$$P\{X>h\}<0.01$$

$$P\{X > h\} = 1 - F(h) = 1 - \Phi(\frac{h-170}{6}) < 0.01$$

即 $\Phi(\frac{h-170}{6}) > 0.99$，查表可得

$$\frac{h-170}{6} \geq 2.33 \Rightarrow h \approx 184(\text{厘米})$$

例3. 若 $X \sim N(\mu, \sigma^2)$, 求 X 落入区间: $[\mu-\sigma, \mu+\sigma]$, $[\mu-2\sigma, \mu+2\sigma]$, $[\mu-3\sigma, \mu+3\sigma]$ 的概率为多少?

解: $P\{\mu-\sigma \le X \le \mu+\sigma\} = F(\mu+\sigma) - F(\mu-\sigma)$

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

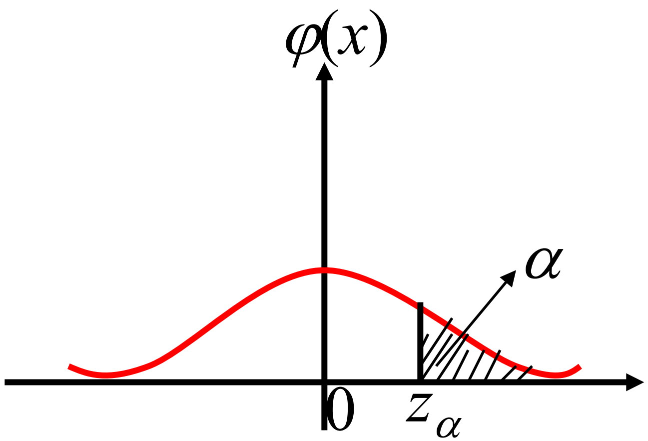
$$= \Phi(1) - \Phi(-1) \approx 0.6826$$

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

$$P\{\mu-3\sigma \le X \le \mu+3\sigma\} = \Phi(3) - \Phi(-3) = 0.997$$

➢ 由上三式可知, 服从正态分布 $N(\mu, \sigma^2)$ 的 r.v. X 之值基本上落入$[\mu-2\sigma, \mu+2\sigma]$之内, 几乎全部落入 $[\mu-3\sigma, \mu+3\sigma]$ 内.

(4) 标准正态分布的上 α 分位点:



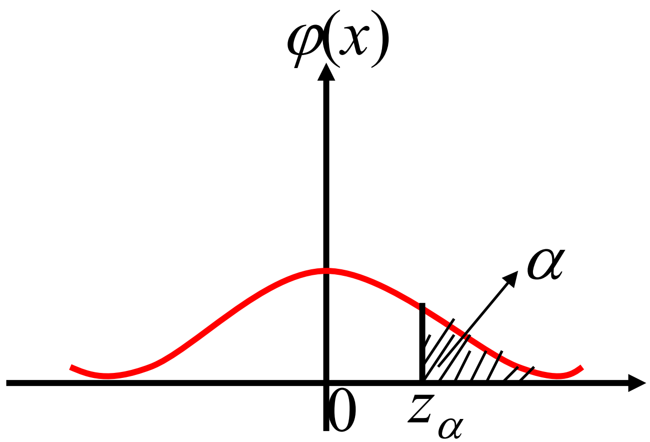
设$X \sim N(0,1)$，若z_α满足条件

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

则称点z_α为标准正态分布的上α 分位点；

由查标准正态分布函数表可知 $z_{0.05} = 1.645$， $z_{0.025} = 1.96$

(4) 标准正态分布的上 α 分位点:



由对称性可得: $z_{1-\alpha} = -z_\alpha$

$$z_{0.95} = -z_{0.05} = -1.645, z_{0.975} = -z_{0.025} = -1.96$$

例4 某校抽样调查结果表明，考生的概率论与数理统计成绩 近似地服从正态分布，平均成绩 72分(μ)，96分以上的占考生总数的 2.3%，求考生的概率统计成绩在 60 分至 84 分之间的概率。

解: $X \sim N(72, \sigma^2)$

$$P\{X > 96\} = 1 - P\{X \le 96\} = 1 - \Phi\left(\frac{96-72}{\sigma}\right) = 2.3\%$$

$\Phi\left(\frac{24}{\sigma}\right) = 0.977$，查表得 $\frac{24}{\sigma} = 2$，$\sigma = 12$

所求概率为:

$$P\{60 \le X \le 84\} = \Phi\left(\frac{84-72}{12}\right) - \Phi\left(\frac{60-72}{12}\right) = \Phi(1) - \Phi(-1)$$

$$= 2\Phi(1) - 1 = 0.682$$

例 5 VAR（Value at Risk）是财务核算中的一个核心概念，投资的 VAR 可以定义为一个值 v，满足投资的损失大于 v 的概率只有 1%。 如果投资收益 X 服从正态分布 $N(\mu,\sigma^2)$，那么，因为损失是收益的相反数，所以我们有

$$0.01=P\{-X>v\}$$

➢由 $-X$ 服从正态分布 $N(-\mu,\sigma^2)$，可得

$$0.01=P\left\{\frac{-X+\mu}{\sigma}>\frac{v+\mu}{\sigma}\right\}=1-\Phi\left(\frac{v+\mu}{\sigma}\right)$$

$$0.01 = P\left\{\frac{-X+\mu}{\sigma} > \frac{\nu+\mu}{\sigma}\right\} = 1-\Phi\left(\frac{\nu+\mu}{\sigma}\right)$$

– 根据正态分布表，$\Phi(2.33)=0.99$，于是我们有

$$\frac{\nu+\mu}{\sigma} = 2.33$$

- $\nu = VAR = 2.33\sigma - \mu$

- 　结论是，在所有收益服从正态分布的投资集合中，使 μ-2.33σ 达到最大值的投资风险最小.

(四) 伽玛分布:

1. 定义: 如果连续型随机变量 X 的概率密度为:

$$f(x)=\begin{cases}\dfrac{\lambda^{p}}{\Gamma(p)}x^{p-1}\mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x\le 0.\end{cases}$$

其中$\lambda>0$, $p>0$为参数，伽玛函数为$\Gamma(p)=\int_0^{+\infty}x^{p-1}\mathrm{e}^{-x}\mathrm{d}x$

则称 X 服从伽玛分布, 简记 $X\sim\Gamma(p,\lambda)$

2. 特例: $\Gamma(1, \lambda)$ 是参数为 $1/\lambda$ 的指数分布.

$$f(x)=\begin{cases}\dfrac{\lambda^p}{\Gamma(p)}x^{p-1}\mathrm{e}^{-\lambda x}, x>0,\\ 0, \qquad x\le 0.\end{cases} \qquad f(x)=\begin{cases}\lambda\mathrm{e}^{-\lambda x}, x>0,\\ 0, \qquad x\le 0,\end{cases}, \theta>0.$$

3. 伽玛函数的性质:

(i) $\Gamma(p+1)=p\Gamma(p)$;

(ii) 对于正整数 n, $\Gamma(n+1)=n!$;

(iii) $\Gamma(\frac{1}{2})=\sqrt{\pi}$.

$$\Gamma(\alpha)=\int_0^{\infty}\mathrm{e}^{-y}y^{\alpha-1}\mathrm{d}y$$

- 对 $\Gamma(\alpha)$ 分部积分可得

$$\Gamma(\alpha)=-\mathrm{e}^{-y}y^{\alpha-1}\Big|_0^{\infty}+\int_0^{\infty}\mathrm{e}^{-y}(\alpha-1)y^{\alpha-2}\mathrm{d}y$$

$$=(\alpha-1)\int_0^{\infty}\mathrm{e}^{-y}y^{\alpha-2}\mathrm{d}y=(\alpha-1)\Gamma(\alpha-1) \qquad (6.1)$$

- 对应于 α 的积分值，比如说 $\alpha=n$，重复利用式(6.1)得到

$$\Gamma(n)=(n-1)\Gamma(n-1)=(n-1)(n-2)\Gamma(n-2)=\cdots$$
$$=(n-1)\times(n-2)\times\cdots\times3\times2\times\Gamma(1)$$

- 又因为 $\Gamma(1)=\int_0^{\infty}\mathrm{e}^{-x}\mathrm{d}x=1$，可得 n 的积分值为：

$$\Gamma(n)=(n-1)!$$

Γ分布与 χ^2 分布

另外：

- $\lambda=1/2, \alpha=n/2$ 的 Γ 分布(n为正整数)称为自由度为n的χ^2 (读作“卡方”)分布.

- 实际中，卡方分布常出现在误差分布中. 例如，在 n 维空间中试图击中某一靶子，其中各坐标的偏差相互独立且为标准正态分布，则偏差的平方和服从自由度为 n 的 χ^2 分布.

回顾

- 随机变量
- 离散型随机变量及分布律
  - (0-1)分布；二项分布；泊松分布；几何分布；超几何分布
- 随机变量的分布函数
- 连续型随机变量及概率密度函数
  - 均匀分布；指数分布；正态分布；伽玛分布

3.5 随机变量的函数的分布

- 研究如何由已知的 r.v. X 的分布，求得它的函数 $Y=g(X)$ 的分布 (其中 $g(.)$ 是已知的连续函数), 分两种情形讨论:

(一) X 为离散型 r.v.

例1：设 X 具有以下的分布律，求 $Y=(X-1)^2$ 分布律:

X	-1	0	1	2
p_k	0.2	0.3	0.1	0.4

X	-1	0	1	2
p_k	0.2	0.3	0.1	0.4
Y	4	1	0	1

即 $P\{Y=0\}=P\{X=1\}=0.1$

$P\{Y=1\}=P\{X=0\}+P\{X=2\}=0.3+0.4=0.7$

$P\{Y=4\}=P\{X=-1\}=0.2$

即

Y	0	1	4
p_k	0.1	0.7	0.2

1. 离散 r.v. 函数的概率分布的求法:

设 X 的概率分布如下表:

X	x_1	x_2	$\cdots$	x_k	$\cdots$
$P\{X=x_i\}$	p_1	p_2	$\cdots$	p_k	$\cdots$

(1) 记 $y_i = g(x_i)$ ($i = 1, 2, \ldots$), y_i 的值也是互不相同的， 则 Y 的概率分布如下表:

Y	y_1	y_2	$\cdots$	y_k	$\cdots$
$P\{Y=y_i\}$	p_1	p_2	$\cdots$	p_k	$\cdots$

(2) 若$g(x_1), g(x_2), \ldots$中不是互不相等的，则应将那些相等的值分别合并，并根据概率加法公式把相应的 p_i 相加，就得到了 Y 的概率分布律.

(二) X 为连续型 r.v.

例2. 设 r.v. X具有概率密度 $f_X(x)$，求$Y=2X+8$ 的概率密度.

$$f_X(x)=\begin{cases} x/8, & 0<x<4 \\ 0, & \text{其它} \end{cases}$$

解： 先求 $Y=2X+8$ 的分布函数 $F_Y(y)$：

$$F_Y(y)=P\{Y\le y\}=P\{2X+8\le y\}$$

$$=P\{X\le \frac{y-8}{2}\} = \int_{-\infty}^{(y-8)/2} f_X(x)\mathrm{d}x$$

$$\left(=\frac{(y-8)^2}{64}, 8<y<16\right)$$

$$= \int_{-\infty}^{(y-8)/2} f_X(x)\mathrm{d}x \quad \left(= \frac{(y-8)^2}{64}, 8 < y < 16 \right)$$

$$\therefore f_Y(y) = F_Y'(y)$$

$$= f_X\left(\frac{y-8}{2}\right)\left(\frac{y-8}{2}\right)' = \begin{cases} \dfrac{y-8}{32}, & 8 < y < 16, \\ 0, & \text{其它}. \end{cases}$$

"分布函数法"：

(1) 先求出 Y 的分布函数: $F_Y(y)=P\{Y\le y\}=P\{g(X)\le y\}=P\{X\in G\}$，其中 $G=\{x: g(x)\le y\}$，转化为关于 X 的事件, 再利用 X 的分布函数表示.

(2) 对 y 求导得到 Y 的概率密度: $f_Y(y)=F_Y'(y)$.

例3. 设X的概率密度为 $f_X(x)$, $-\infty < x < +\infty$, 求$Y = X^2$ 的概率密度.

解： 当$y \le 0$时, $F_Y(y) = P\{Y \le y\} = 0.$

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

$$= P\{-\sqrt{y} \le X \le \sqrt{y}\} = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\mathrm{d}x$$

则 $$f_Y(y) = F_Y'(y) = \begin{cases} [f_X(\sqrt{y}) + f_X(-\sqrt{y})]\dfrac{1}{2\sqrt{y}}, & y > 0, \\ 0, & y \le 0. \end{cases}$$

则 $f_Y(y)=F_Y'(y)=\begin{cases}[f_X(\sqrt{y})+f_X(-\sqrt{y})]\dfrac{1}{2\sqrt{y}}, & y>0,\\ 0, & y\le 0.\end{cases}$

例如，设$X\sim N(0,1)$，其概率密度为 $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}, -\infty<x<+\infty,$

则$Y=X^2$ 的概率密度为 $f_Y(y)=\begin{cases}\dfrac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}\mathrm{e}^{-\frac{y}{2}}, & y>0,\\ 0, & y\le 0,\end{cases}$

此时称Y服从自由度为 1 的χ^2分布.

设随机变量X具有概率密度$f_X(x), -\infty < x < \infty$，又设函数$g(x)$处处可导且恒有$g'(x) > 0$(或恒有$g'(x) < 0$), 则$Y = g(X)$是连续型随机变量，其概率密度为

$$f_Y(y) = \begin{cases} f_X(h(y))\,|h'(y)|, \alpha < y < \beta \\ 0, & \text{其他} \end{cases}$$

其中$\alpha = \min\{g(-\infty), g(\infty)\}, \beta = \max\{g(-\infty), g(\infty)\}, h(y)$是$g(x)$的反函数。

$$f_Y(y)=\begin{cases} f_X(h(y))\,|h'(y)|, & \alpha<y<\beta \\ 0, & \text{其他} \end{cases}$$

例2. 设 r.v. X具有概率密度 $f_X(x)$，求$Y=2X+8$ 的概率密度.

$$f_X(x)=\begin{cases} x/8, & 0<x<4, \\ 0, & \text{其它}, \end{cases}$$

$$f_Y(y)=\begin{cases} \dfrac{y\text{-}8}{32}, & 8<y<16, \\ 0, & \text{其它}. \end{cases}$$

例3. 设随机变量 $X\sim N(\mu,\sigma^2)$, 求 $Y=aX+b\ (a>0)$ 的概率密度。

解：由$X\sim N(\mu,\sigma^2)$知$f_X(x)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

现在$y=g(x)=ax+b$,由这一式子解得 $x=h(y)=\dfrac{y-b}{a}$,且有$h'(y)=\dfrac{1}{a}$

由定理可得$Y=aX+b$ 的概率密度为

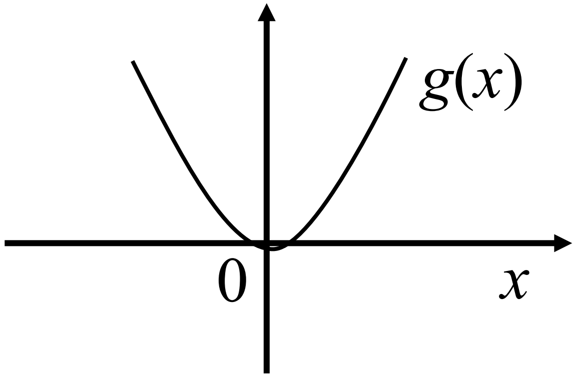
$$f_Y(y)=\frac{1}{|a|}f_X\left(\frac{y-b}{a}\right),-\infty<y<\infty$$

即 $$f_Y(y)=\frac{1}{|a|}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}}=\frac{1}{|a|\sigma\sqrt{2\pi}}e^{-\frac{[y-(b+a\mu)]^2}{2(a\sigma)^2}}$$

即有 $Y=aX+b\sim N(a\mu+b,(a\sigma)^2)$

设随机变量X具有概率密度$f_X(x)$, $-\infty < x < \infty$，又设函数$g(x)$在不相重叠的区间I_1，I_2, … 上逐段严格单调，其反函数分别为$h_1(y)$，$h_2(y)$,…，且$h'_1(y)$，$h'_2(y)$,…均为连续函数，那么Y是连续型随机变量，且密度函数为

$$f_Y(y) = f_X(h_1(y))\,|\,h_1'(y)\,| + f_X(h_2(y))\,|\,h_2'(y)\,| + \cdots$$



例4. 设X的概率密度为$f_X(x)$, $-\infty < x < +\infty$, 求$Y = X^2$ 的概率密度.

$X \sim N(0, 1)$, 其概率密度为 $\phi(x) = \dfrac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}, -\infty < x < +\infty,$

$$f_Y(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} \mathrm{e}^{-\frac{y}{2}}, & y > 0, \\ 0, & y \leq 0, \end{cases}$$

例5. 设电压$V = A\sin\theta$，其中A 是一个已知的正常数，相角Θ 是一个随机变量，且有$\Theta \sim U(-\pi/2，\pi/2)$，试求电压$V$ 的概率密度。

解： $v = g(\theta) = A\sin\theta$在（$-\pi/2$，$\pi/2$）上恒有$g'(\theta) = A\cos\theta > 0$,

且有反函数 $\theta = h(v) = \arcsin(v/A)$,

$$h'(v) = \frac{1}{\sqrt{A^2 - v^2}}$$

$$(\arcsin x)' = \frac{1}{\sqrt{1 - x^2}}$$

又Θ 的概率密度为$f(\theta)=\begin{cases}1/\pi, & -\pi/2<\theta<\pi/2 \\ 0, & \text{其他}\end{cases}$

由定理得$V=A\sin\Theta$ 的概率密度为$\psi(v)=\begin{cases}\dfrac{1}{\pi\sqrt{A^2-v^2}}, & -A<v<A \\ 0, & \text{其他}\end{cases}$