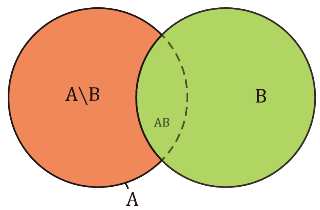
1 集合与概率

1.1 集合运算

减法的等效形式:

$$A \setminus B = A \setminus AB = A \cap \overline{B}$$



1.2 集合恒等式

交换律, 结合律, 分配律略. 记全集为 U,
幂等律:

$$A \cup A = A$$

$$A \cap A = A$$

零元:

$$A \cup U = U$$

$$A \cap \varnothing = \varnothing$$

单位元:

$$A \cup \varnothing = A$$

$$A \cap U = A$$

吸收律:

$$A \cup (A \cap B) = A$$

$$A \cap (A \cup B) = A$$

若 $A \subseteq A^+$:

$$A \cup A^+ = A^+$$

$$A \cap A^+ = A$$

即: 并集只可能增大集合，交集只可能减小集合

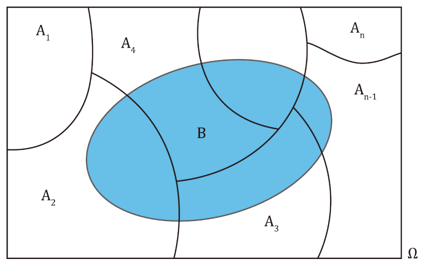
1.3 概率运算

减法公式:

$$P(A \setminus B) = P(A \setminus AB) = P(A) - P(AB)$$

容斥原理:
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
条件概率:
$$P(A \mid B) = \frac{P(AB)}{P(B)}$$
乘法公式:
$$P(AB) = P(A)P(B \mid A) = P(B)P(A \mid B)$$
全概率公式: A_i $(i = 1, \ldots, n)$ 构成样本空间 Ω 的完备事件组, 即 A_i 将 Ω 有限划分.

$$P(B) = \sum_i P(A_i)P(B|A_i)$$



贝叶斯公式:
$$P(B_i \mid A) = \frac{P(AB_i)}{P(A)} = \frac{P(B_i)P(A \mid B_i)}{\sum\limits_i P(B_i)P(A \mid B_i)}$$

1.4 独立与互斥

互斥: 两事件不会同时发生, 即: $A \cap B = \varnothing$.
独立: 一个事件的发生与否不影响另一个事件的发生, 即: $P(AB) = P(A)P(B)$.

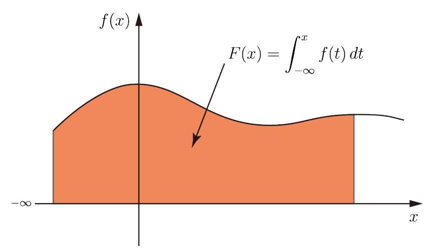
2 随机变量及其分布

2.1 一维

连续:

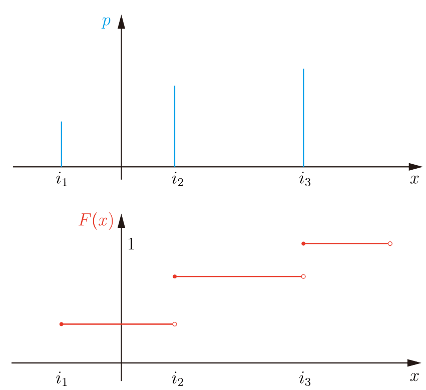
$$F_X(x) = P(X \leqslant x) = \int_{-\infty}^{x} f_X(t)\,dt$$

$$f_X(x) = \frac{d}{dx} F_X(x)$$



离散:

$$F_X(x) = P(X \leqslant x) = \sum_{i \leqslant x} P(X = i)$$



2.2 二维

2.2.1 连续

联合:

$$F(x,y) = P(X \leqslant x, Y \leqslant y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\,dv\,du$$

$$f(x,y) = \frac{\partial^2 f}{\partial x \partial y}$$

边缘:

$$F_X(x) = F(x, +\infty) = P(X \leqslant x) = \int_{-\infty}^{x} f_X(t)\,dt$$

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy$$

2.2.2 离散

联合: 记 $p_{ij} = P(X = i, Y = j)$

$$F(x, y) = P(X \leqslant x, Y \leqslant y) = \sum_{\substack{i \leqslant x \\ j \leqslant y}} p_{ij}$$

边缘: 记 $p_{i\cdot}$ 表示 $P(X = i)$ 而 Y 任意取值a

$$p_{i\cdot} = \sum_j p_{ij}$$

$$p_{\cdot j} = \sum_i p_{ij}$$

2.3 条件分布

$$F_{X|Y}(x \mid y) = P(X \leqslant x \mid Y = y)$$

离散:

$$P(X = x_i \mid Y = y_j) = p_{ij} / p_{\cdot j}$$

连续:

$$f_{X|Y}(x \mid y) = \frac{f(x, y)}{f_Y(y)}$$

$$F_{X|Y}(x \mid y) = \int_{-\infty}^{x} f_{X|Y}(u \mid y)\, du$$

2.4 独立

$$F(x, y) = F_X(x) F_Y(y) \Rightarrow \text{独立}$$

$$\text{独立} \Leftrightarrow f(x, y) = f_X(x) f_Y(y)$$

$$\text{独立} \Leftrightarrow p_{ij} = p_{i\cdot} p_{\cdot j}$$

$$X, Y \text{ 独立} \Rightarrow f(X) \text{ 与 } g(Y) \text{ 也独立}$$

3 随机变量的数字特征

3.1 期望

离散:

$$E(X) = \sum_{k=1}^{\infty} x_k p_k$$

连续:

$$E(X) = \int_{-\infty}^{+\infty} x f(x)\, dx$$

随机变量函数的期望:

$$E(g(X)) = \begin{cases} \displaystyle\sum_{k=1}^{\infty} g(x_k)p_k \\ \displaystyle\int_{-\infty}^{+\infty} g(x)f(x)\,dx \end{cases}$$

性质:

- 常数: $E(C) = C$
- 线性: $E(aX + b) = aE(X) + b$
- 若 X, Y 独立: $E(XY) = E(X)E(Y)$

3.2 方差

- 定义: $D(X) = E[X - E(X)]^2 = E(X^2) - E(X)^2$
- 常数: $D(C) = 0$
- 二次齐次, 平移不变性: $D(aX + b) = a^2 D(X)$
- 加/减法: $D(X \pm Y) = D(X) + D(Y) \pm 2\,\mathrm{Cov}(X, Y)$
- 协方差: $D(X) = \mathrm{Cov}(X, X)$

3.3 协方差

- 定义: $\mathrm{Cov}(X, Y) = E\big[[X - E(X)][Y - E(Y)]\big] = E(XY) - E(X)E(Y)$
- 交换律: $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
- 常数: $\mathrm{Cov}(X, C) = 0$
- 双线性: $\mathrm{Cov}(aX + a', bY + b') = ab\,\mathrm{Cov}(X, Y)$

  $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$
- X, Y 独立: $\mathrm{Cov}(X, Y) = 0$

协方差阵: 记 $\sigma_{ij} = \mathrm{Cov}(X_i, X_j)$:

$$\mathbf{V} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}$$

3.4 相关系数

$$R(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}}$$

性质:

- $|R(X,Y)|\leqslant 1$
- $|R(X,Y)|=1$ 的充要条件为: X, Y 线性相关
- $R(X,Y)=0$, X, Y 没有线性关系(可能有其他相关关系)

4 正态分布

4.1 标准正态分布

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$$

$$\Phi(x)=\int_{\infty}^{x}\varphi(t)\,dt$$

性质:
$\Phi(0)=\frac{1}{2}$
$\Phi(-x)=1-\Phi(x)$

4.2 正态分布

若 $\frac{X-\mu}{\sigma}\sim N(0,1)$, 则 $X\sim N(\mu,\sigma^2)$. 此时对于 X:

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

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

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

4.3 正态分布性质

$X\sim N(\mu,\sigma^2)$, $Y_i\sim N(\mu_i,\sigma_i^2)$:

- $E(X)=\mu$, $D(X)=\sigma^2$
- $aX+b\sim N(a\mu+b,a^2\sigma^2)$
- $Y_1\pm Y_2\sim N(\mu_1\pm\mu_2,\sigma_1^2+\sigma_2^2)$ (注意方差总是相加)

5 大数定律与中心极限定理

5.1 切比雪夫不等式

$$P\big(|X-E(X)|<\varepsilon\big)\geqslant 1-\frac{D(X)}{\varepsilon^2}$$

由此推导出:

随机变量序列 $E(X_n)=\mu_n$, $D(X_n)=\sigma_n^2$, $\lim\limits_{n\to\infty}\sigma_n^2=0$, 则:

$$X_n\xrightarrow{P}\mu_n$$

5.2 大数定律

5.2.1 切比雪夫大数律

<u>独立</u>随机变量序列 $\{X_k\}$, X_k 的期望方差均存在且<u>方差有界</u>, 则:

$$\frac{1}{n}\sum_{k=1}^{n}X_k-\frac{1}{n}\sum_{k=1}^{n}E(X_k)\xrightarrow{P}0$$

5.2.2 独立同分布大数律

$\{X_k\}$ <u>独立同分布</u>, $E(X_k)=\mu$, $D(X_k)=\sigma^2$:

$$\overline{X}\xrightarrow{P}\mu$$

5.3 中心极限定理

5.3.1 独立同分布

随机变量序列 $\{X_k\}$ <u>独立同分布</u>, $E(X_k)=\mu_k$, $D(X_k)=\sigma^2$.

则随机变量 $Y_n=\dfrac{\sum\limits_{k=1}^{n}X_k-n\mu}{\sqrt{n}\sigma}=\dfrac{\frac{1}{n}\sum\limits_{k=1}^{n}X_k-\mu}{\sigma/\sqrt{n}}$ 当 $n\to\infty$ 时的极限分布函数为标准正态分布函数 $\Phi(x)$.

即意味着, 当 n 充分大:

$$\sum_{k=1}^{n}X_k\overset{\cdot}{\sim}N(n\mu,n\sigma^2)$$

$$\overline{X}=\frac{1}{n}\sum_{k=1}^{n}X_k\overset{\cdot}{\sim}N\left(\mu,\frac{\sigma^2}{n}\right)$$

6 数理统计

6.1 三大分布

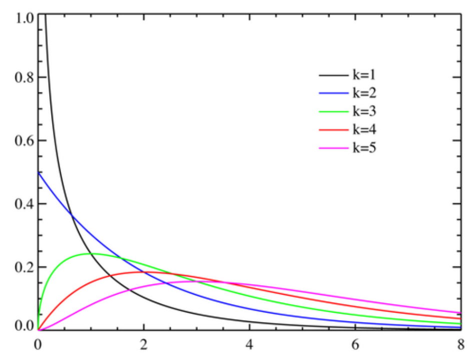
6.1.1 χ^2 分布

$X_1, \ldots, X_n$ 独立同服从 $N(0,1)$, 称随机变量 $\chi^2 = \sum_{i=1}^{n} X_i^2$ 服从自由度为 n 的 χ^2 分布, 记作:

$$\chi^2 \sim \chi^2(n)$$

性质:

1. $\chi^2(n) = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$
2. $\chi_1^2 \sim \chi^2(n)$, $\chi_2^2 \sim \chi^2(m)$, χ_1, χ_2 独立, 则 $\chi_1^2 + \chi_2^2 \sim \chi^2(n+m)$



6.1.2 t 分布

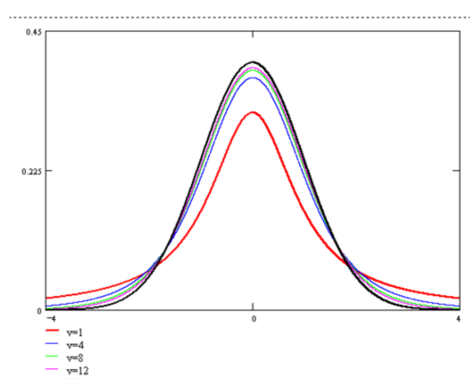
$X \sim N(0,1)$, $Y \sim \chi^2(n)$, X, Y 独立:

$$t = \frac{X}{\sqrt{Y/n}} \sim t(n)$$

称 $t(n)$ 自由度为 n

性质:

1. $n \to \infty$, 则 $t(n) \to N(0,1)$



6.1.3 F 分布

$X_1 \sim \chi^2(n_1)$, $X_2 \sim \chi^2(n_2)$, X_1, X_2 独立:

$$F = \frac{X_1/n_1}{X_2/n_2} \sim F(n_1, n_2)$$

称 $F(n_1, n_2)$ 自由度为 (n_1, n_2). F 分布的图形与 χ^2 分布类似.

性质:

1. $F \sim F(n_1, n_2)$, 则 $\dfrac{1}{F} \sim F(n_2, n_1)$
2. $X \sim t(n)$, 则 $X^2 \sim F(1, n)$

6.1.4 分位点

若有: $F(x_p) = P(X \leqslant x_p) = p$, 则称 x_p 为分布 F 的 p 分位点:

1. 标准正态分布密度函数为偶函数 $u_p + u_{1-p} = 0$
2. 当 $n > 45$ 时, χ^2 分布分位点: $\chi_p^2(n) \approx (u_p + \sqrt{2n-1})^2/2$
3. t 分布与标准正态分布密度函数近似: $-t_p = t_{1-p}$
4. F 分布有: $F_p(n_1, n_2) = \dfrac{1}{F_{1-p}(n_2, n_1)}$

6.2 统计量

样本均值:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

样本方差:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right)$$

样本标准差: $S = \sqrt{S^2}$

样本 k 阶(原点)矩:

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

样本 k 阶中心距:

$$B_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k$$

恒等式:

$$B_2 = A_2 - \overline{X}^2$$

$$S^2 = \frac{n}{n-1} B_2$$

6.3 抽样分布定理

6.3.1 一个正态分布总体

样本均值的分布: $X_1, X_2, \ldots, X_n$ 来自总体 $N(\mu, \sigma^2)$, 则:

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \qquad \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

样本方差和 2 阶中心距的分布:

$$\frac{(n-1)S^2}{\sigma^2} = \frac{nB_2}{\sigma^2} \sim \chi^2(n-1)$$

标准差的分布与样本均值:

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

6.3.2 任意总体

$X_1, \ldots, X_n$ 来自任何总体:

$$E(\overline{X}) = E(X) \qquad D(\overline{X}) = \frac{D(X)}{n}$$

中心极限定理: 当 n 充分大时, 近似有

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \overset{\cdot}{\sim} N(0,1)$$

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \overset{\cdot}{\sim} N(0,1)$$

7 点估计

7.1 矩估计法

7.2 极大似然估计法

似然函数:

$$L(\theta) = \prod_{i=1}^{n} p(x_i, \theta) \text{ 或 } L(\theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

要使似然函数取得最大值, 即求满足似然方程的 θ:

$$\frac{dL(\theta)}{d\theta} = 0$$

上式等价于

$$\frac{d \ln L(\theta)}{d\theta} = 0$$

7.3 估计量的评选标准

7.3.1 无偏性

若估计量 $\hat{\theta}$ 的期望等于被估计的参数, 则称 $\hat{\theta}$ 为无偏估计量, 否则为有偏估计量.

$$E(\hat{\theta}) = \theta$$

对有偏估计量 $\hat{\theta}$, 记 $b(\hat{\theta}) = E(\hat{\theta}) - \theta$. 若:

$$\lim_{n\to\infty} b(\hat{\theta}) = 0$$

则称 $\hat{\theta}$ 为渐进无偏估计量.

常见无偏估计量 若总体 X 服从任意分布, 则样本均值 $\overline{X}$, 样本 k 阶矩 A_k, 样本方差 S^2 分别是 $E(X)$, $E(X^k)$, $D(X)$ 的无偏估计量.

7.3.2 有效性

设 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 为两个无偏估计量, 若:

$$D(\hat{\theta}_1) \leqslant D(\hat{\theta}_2)$$

则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效.

8 区间估计

8.1 一个正态总体

总体 $X \sim N(\mu, \sigma^2)$, $X_1, \ldots, X_n$ 是来自总体的样本:

8.1.1 已知 $\sigma^2 = \sigma_0^2$, 求 μ 的置信区间

取样本函数:

$$\frac{\overline{X} - \mu}{\sigma_0/\sqrt{n}} \sim N(0,1)$$

8.1.2 σ^2 未知, 求 μ 的置信区间

取样本函数:

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$$

8.1.3 μ 未知, 求 σ^2 的置信区间

取样本函数:

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

9　假设检验

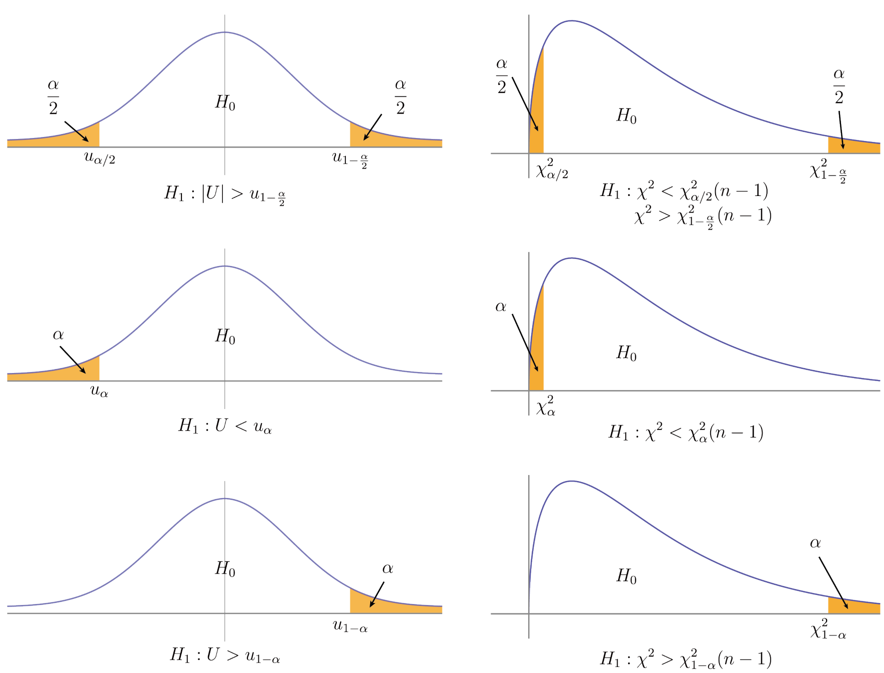
9.1　一个正态总体均值假设检验表

H_0	H_1	σ^2 已知	σ^2 未知
		显著性水平 α 下 H_0 的拒绝域	
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert U\rvert > u_{1-\frac{\alpha}{2}}$	$\lvert t\rvert > t_{1-\frac{\alpha}{2}}$
$\mu = \mu_0$	$\mu > \mu_0$	$U > u_{1-\alpha}$	$t > t_{1-\alpha}(n-1)$
$\mu = \mu_0$	$\mu < \mu_0$	$U < u_{\alpha}$	$t < t_{\alpha}(n-1)$

9.2　一个正态总体方差假设检验表

H_0	H_1	μ 未知, 显著性水平 α 下 H_0 的拒绝域
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi^2_{\alpha/2}(n-1)$ 或 $\chi^2 > \chi^2_{1-\frac{\alpha}{2}}(n-1)$
$\sigma^2 = \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_{1-\alpha}(n-1)$
$\sigma^2 = \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$\chi^2 < \chi^2_{\alpha}(n-1)$

下图从上到下依次为双侧检验, 左侧检验, 右侧检验的图示:



u 和 χ^2 检验: 黄色部分为拒绝域 H_1

t 分布概率密度函数图像与标准正态分布相似, t 检验可参考 u 检验.

10　离散型模型

名称	记号/参数	概率	期望	方差
二项分布	$B(n,k)$	$p_k=\binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$
几何分布	$G(p)$	$p_k=pq^{k-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
超几何分布	$H(n,M,N)$	$p_k=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$	$n\frac{M}{N}$	$n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}$
泊松分布	$P(\lambda)$	$p_k=\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$	λ	λ

11　连续型模型

名称	记号/参数	密度函数	期望	方差
均匀分布	$U(a,b),U(G)$	$f(x)=\begin{cases}\frac{1}{m(G)}=\frac{1}{b-a} & ,x\in(a,b)\\ 0 & ,x\notin(a,b)\end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Γ 分布	$\Gamma(\alpha,\beta)$	$f(x)=\begin{cases}\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\beta x} & ,x>0\\ 0 & ,x\leqslant 0\end{cases}$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$
指数分布	$e(\beta)=\Gamma(1,\beta)$	$f(x)=\beta\mathrm{e}^{-\beta x}\quad ,x>0$ $F(x)=1-\mathrm{e}^{-\beta x}\quad ,x>0$	$\frac{1}{\beta}$	$\frac{1}{\beta^2}$
χ^2 分布	$\chi^2(n)$	$f(x)=\begin{cases}\frac{x^{n/2-1}\mathrm{e}^{-x/2}}{2^{n/2}\Gamma(n/2)} & ,x>0\\ 0 & ,x\leqslant 0\end{cases}$	n	$2n$