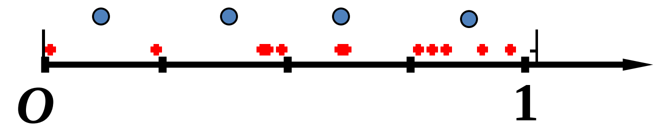
第五章 概率统计应用实验

- 随机数与统计直方图
- 相遇问题及其统计试验
- 贝努里试验与二项分布
- 正态随机数及应用
- 计算面积的蒙特卡罗方法

均匀分布随机数



*MATLAB*产生均匀随机数方法： rand(m，n)

产生m×n个 0,1 之间均匀随机数.随机数等可能落入区间[0，1]内长度相等子区间中。

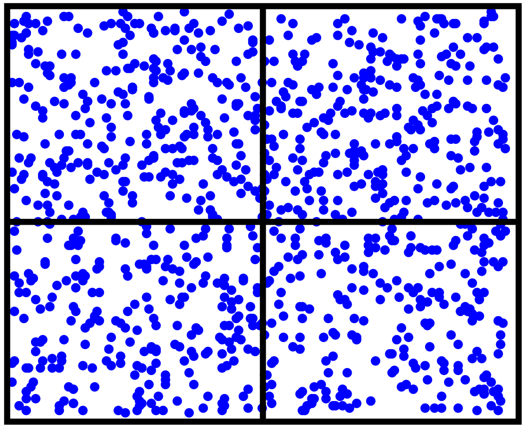
引例1. 观察12个1—4之间整型随机数情况

1+ fix(4*rand(1,12))

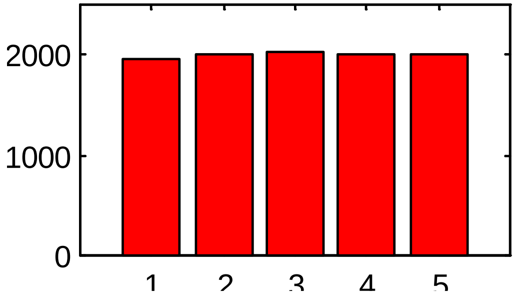
ans= 4 1 3 2 4 4 2 1 4 2 3 4

引例2. 观察1000个随机点分布情况

```
P=rand(2,1000);
x=P(1,:);y=P(2,:);
plot(x,y,'b.')
length(find(x>0.5 & y>0.5))/1000
```



统计直方图



直方图绘图命令：hist(data，n)

其中,data是需要处理的数据块,

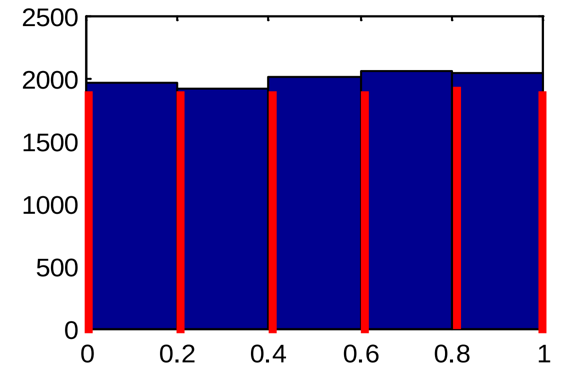
绘图原理:利用data中最小数和最大数构成一区间,将区间等分为n个小区间，统计落入每个小区间的数据量。以数据量为高度绘小矩形，形成直方图。如果省略参数n，MATLAB将n的默认值取为10。

直方图也可以用于统计计算

N=hist(data，n)

计算结果N是n个数的一维数组，分别表示data中各个小区间的数据量。这种方式只计算而不绘图。

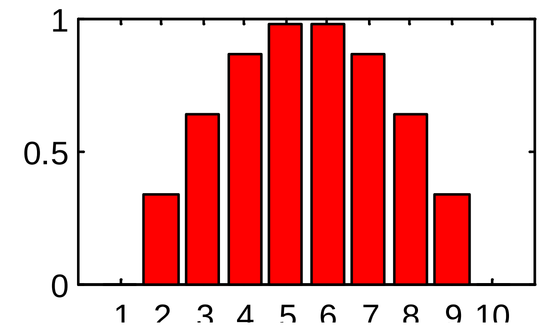
例5.1 统计10000个均匀随机数在五个小区间的分布 。



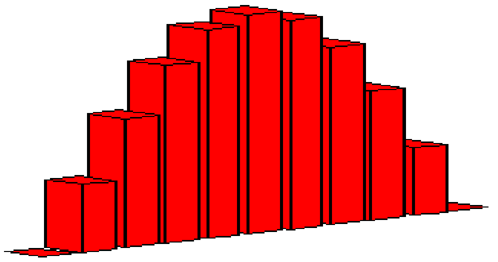

```
data=rand(10000,1);
    hist(data,5)
    N5=hist(data,5)
```

N5 = 1969   2010   2018   1999   2004

条形图是根据数据绘小矩形或小柱体。使用格式: bar(data)或bar3(data)




```
x=linspace(0,pi,10);
y=sin(x);
bar(y,'r')
bar3(y,'r')
```



例5.2 **相遇问题** 甲、乙两船在24小时内独立地随机到达码头，如果甲船到达码头后停留2小时,乙船到达码头后停留1小时. 问两船相遇的概率有多大?

分析：设两船到达码头时刻分别为 X 和 Y，则有

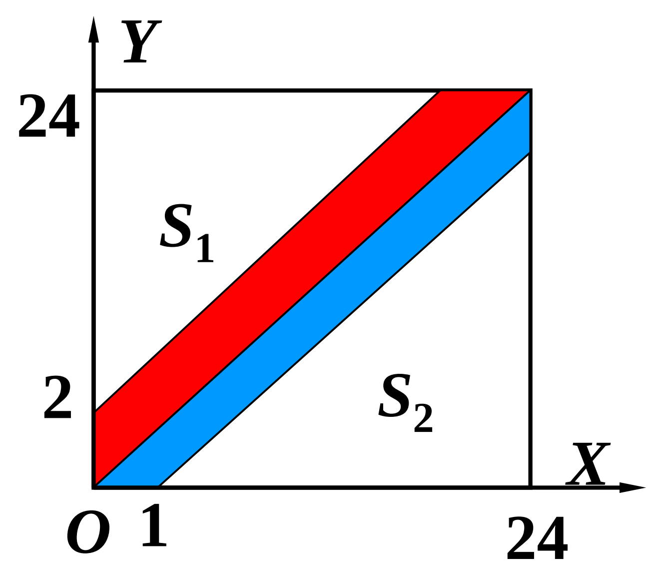
随机变量 $X \sim U(0\,,24),\ Y \sim U(0\,,24)$

$$D=\{(x,y)\,|\,x-1\le y\le x+2, 0<x,y<24\}$$

$$P\{(X,Y)\in D\}=\frac{24^2-S_1-S_2}{24^2}$$

$$S_2=0.5\times 23^2$$

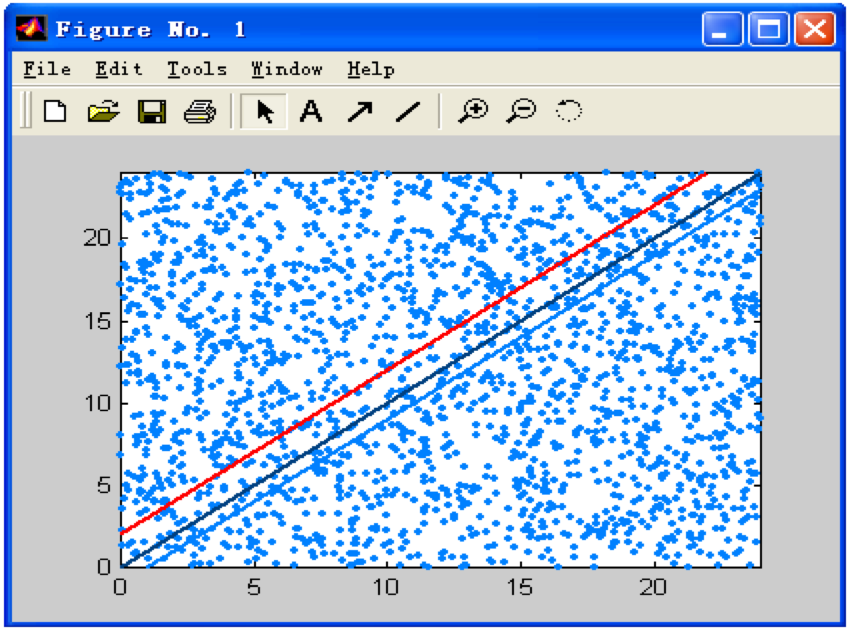
$$S_1=0.5\times 22^2$$



相遇问题的统计试验

$$D = \{(x, y) \mid x - 1 \le y \le x + 2,\ 0 < x, y < 24\}$$

```
function F=shipmeet(N)
if nargin==0,N=2000;end
P=24*rand(2,N);
X=P(1,:);Y= P(2,:);
I=find(X-1<=Y&Y<=X+2);
F=(length(I))/N
plot(X,Y,'b.') ,hold on
```



F =   0.1185

$$S_1 = 0.5 \times 22^2 \qquad S_2 = 0.5 \times 23^2$$

$$P\{(X, Y) \in D\} = \frac{24^2 - S_1 - S_2}{24^2} = 0.1207$$

贝努里概型

与贝努里试验

X	0	1
P	0.5	0.5

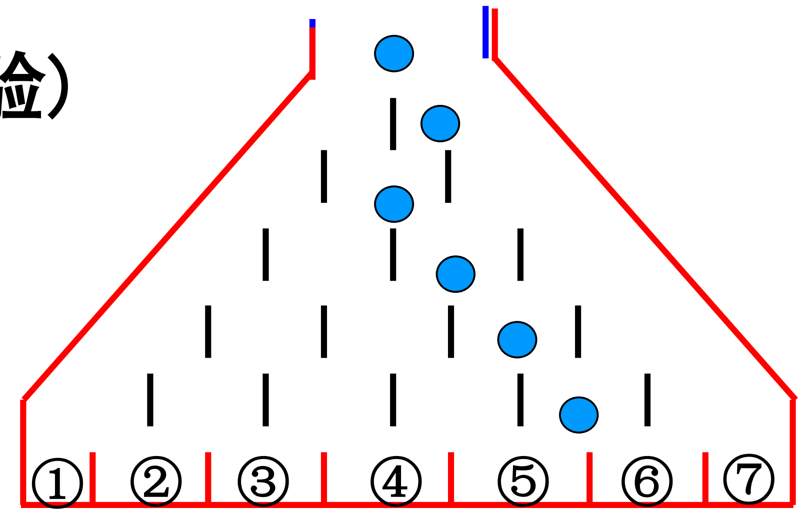
Bernoulli,1654--1705

例5.3设事件A出现的概率为p=0.5。模拟100次贝努里试验，统计实验结果中“0”出现的次数和“1”出现的次数。

```
data=fix(2*rand(100,1));
N=hist(data,2)
```

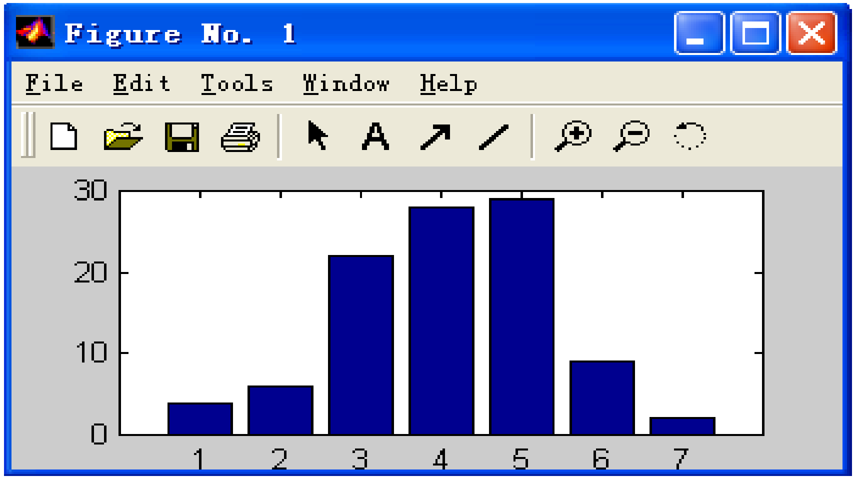
实验序号	1	2	3	4	5
0出现次数	50	52	52	61	54
1出现次数	50	48	48	39	46

六层Galton板（六重贝努里试验）



例5.4 小球自顶部落下,在每一层遭遇隔板,以1/2的概率向右（左）下落,底部六个隔板,形成七个槽.模拟100个小球依次落下,统计Galton板底部各槽中小球数

```
X=fix(2*rand(6,100));
Y=sum(X)+1;
N=hist(Y,7)   %统计
bar(N)
```



N= 4 6 22 28 29 9 2

记 $Y = X_1 + X_2 + X_3 + X_4 + X_5 + X_6$

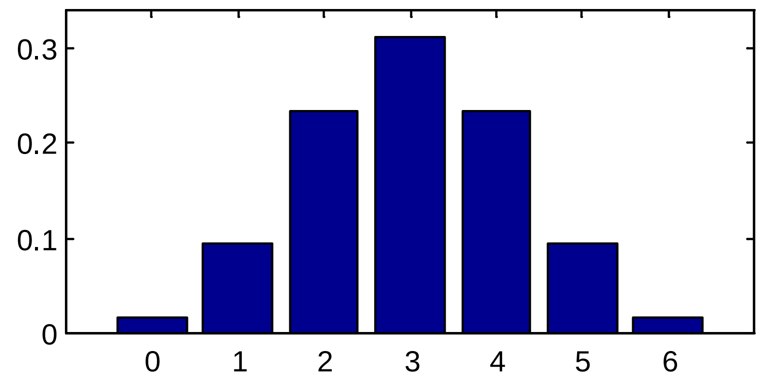
Y 服从n=6的二项分布 $Y \sim B(\text{n}, p)$ p=0.5

$$P\{Y = k\} = C_6^k p^k (1-p)^{n-k} \qquad k = 0,1,2,\ldots,6$$

二项分布概率计算函数: binopdf(x,n,p)

x 是n重贝努里试验中事件A出现的次数.

```
%计算Galton试验板分布律
n=6;x=0:n;
Y=binopdf(x,n,0.5)
bar(x,Y)
 ans =
    0.0156    0.0938    0.2344    0.3125    0.2344    0.0938    0.0156
```



计算二项分布随机变量*X*=*k*的命令使用格式为

Pk=binopdf(k，n，p)

其中，k是随机变量取值，n是贝努里试验的重数，p为n重贝努里试验中事件A发生的概率。

对于二项分布随机变量X，计算累加概率$P\{X \le k\}$的MATLAB命令使用格式为

P = binocdf(k，n，p)

MATLAB的二项分布随机数发生器使用格式为

R= binornd(n，p，L，M)

产生 L×M 个二项分布随机数。

例5.5 有一千名以上的小学生参加保险公司的平安保险, 参加保险的小学生每人一年交保险费50元. 若一年内出现意外事故, 保险公司赔付一万元。统计表明, 每年一千名小学生中平均有两名学生出事故。保险公司赔本的概率有多大？利用二项分布随机数进行模拟

分析：小学生出意外事故的概率为p=0.002，设随机变量X为一年内出事故的小学生人数。X服从二项分布$B(n，p)$，其中n为投保人数。由于对出事故的小学生，保险公司一次性赔付一万元，所以每年保险公司赔付费为：X（万元）。一年中保险公司赔付费不超过总的保险收费则会获利，如果赔付费超过总的保险收费将会赔本。每年保险公司所获利润为总保险收费减去总的赔付费。

```
function [P1,profits]=prob1(N)
p=0.002;
join=50;pay=10000;
all=join*N;
X1=fix(all/pay);                  %赔付最大承受人数
P1=1-binocdf(X1,N,p);             %赔本概率
puples=binornd(N,p,1,8);          %八年出事故人数模拟
Pays=pay*puples;                  %八年赔付金模拟
profits=all-Pays;                 %八年利润模拟
```

[P,p]=prob1(1500)➔

P =   0.0118

55000 65000 15000 45000 45000   5000 35000 55000

正态分布变量X的数学期望μ，方差σ^2，密度函数

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp[-(x-\mu)^2/2\sigma^2) \qquad x\in(-\infty, +\infty)$$

计算命令：y = normpdf(x，mu，sigma)

累积分布函数，即积分上限函数

$$P\{X<x\}=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}\exp[-(t-\mu)^2/2\sigma^2]dt$$

计算命令　：p = normcdf(x，mu，sigma)

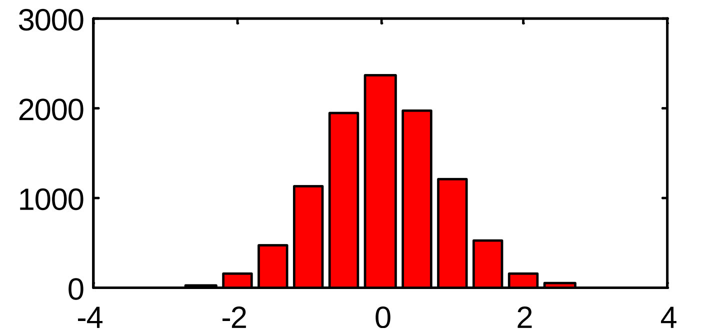
逆累积分布函数值，即已知概率值p，求z 使得

$$P\{X<z\}=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{z}\exp[-(t-\mu)^2/2\sigma^2]dt=p$$

计算命令　：z = norminv(p，mu，sigma)

产生正态分布随机数的函数为randn()，使用格式为

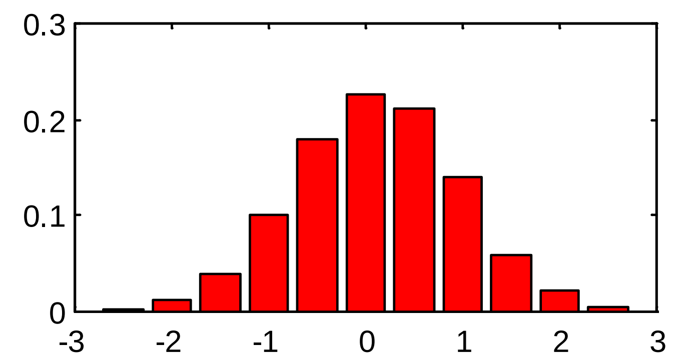
R=randn(m，n)



产生m×n阶矩阵R, 矩阵中元素都是区间(– 3，3)内的正态随机数。

例5.6 创建10000个正态随机数，将区间[–3，3]分为十三个小区间，分别绘频数和频率直方图。

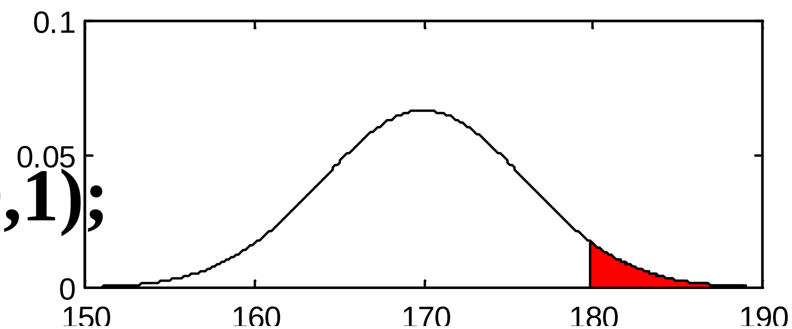
```
data=randn(10000,1);
N=hist(data,13);
figure(1),bar([-3:0.5:3],N,'r')
figure(2),M=N/10000;
bar([-3:0.5:3],M,'r')
```



例5.8 某城市中99%男子身高介于1.52米到1.88米，如果男子上公交车时头与车门相碰的概率小于5%，公交车门的高度应该是多少？

分析：设身高为正态分布随机变量X，170(cm)为X的数学期望，方差取为36。计算z，使得$P\{X > z\} = 0.05$。即求逆累积函数在 x=0.95 处的值

```
mu=170;sigam=6;
z=norminv(0.95,mu,sigam)
data=mu+sigam*randn(1000,1);
II=find(data>=z);
F=length(II)/10000
```



z = 179.8691

F = 0.0047

蒙特卡罗方法，或称计算机随机模拟方法，是一种基于“随机统计”的计算方法。方法源于美国在第二次世界大战中研制原子弹的“**曼哈顿计划**”。

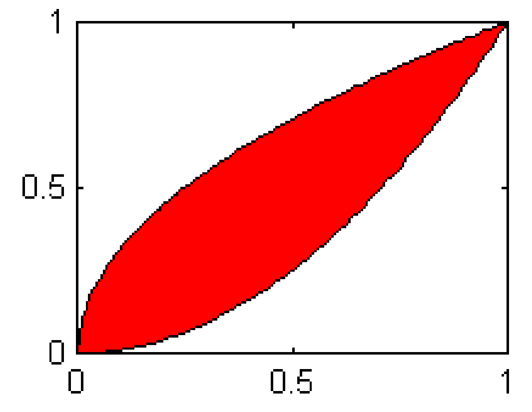
例5.13计算两条抛物线 $y = x^2$，$x = y^2$ 所围图形的面积.

在正方形区域D内投入N个点，统计坐标满足

$$x^2 \le y \le \sqrt{x}$$

的点P(x，y)的数目M。面积近似计算公式为：S=M/N

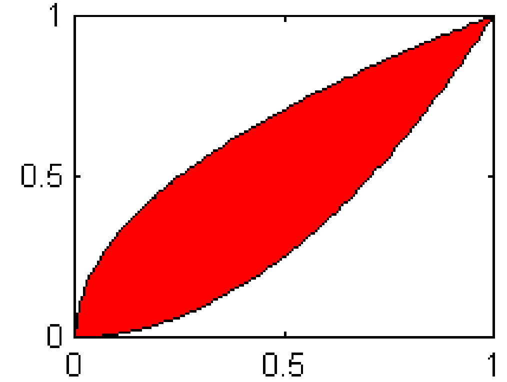
```
data=rand(N,2);
x=data(:,1);y=data(:,2);
II=find(y<=sqrt(x)&y>=x.^2);
M=length(II);
S=M/N
```



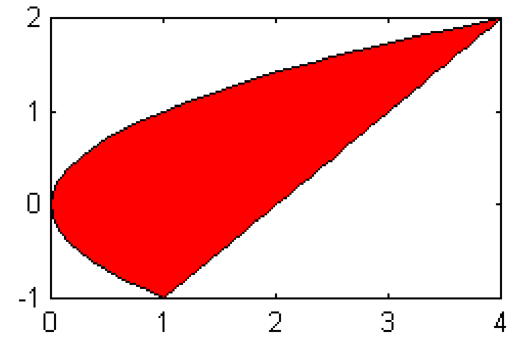
S = 0.3276

填充图绘制方法

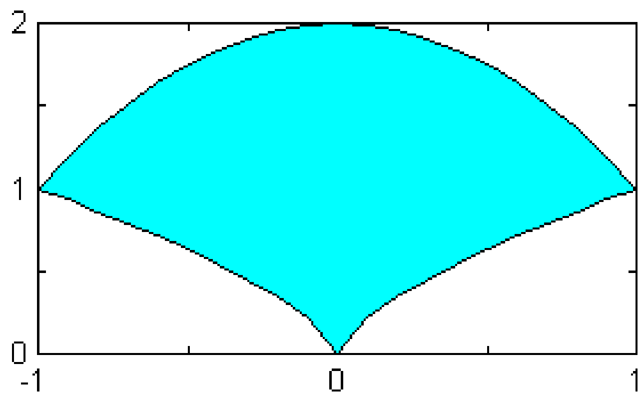
```
x1=0:.01:1;y1=sqrt(x1);
x2=1:-.01:0;y2=x2.^2;
fill([x1,x2],[y1,y2],'r')
```




```
y1=-1:.1:2;y2=2:-.1:-1;
x11=y1.*y1;x22=y2+2;
fill([x11,x22],[y1,y2],'r')
```



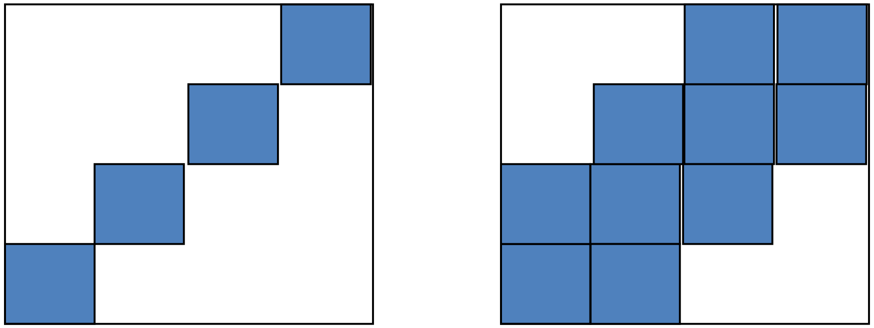

```
x1=-1:0.1:1; y1=x1.^2.^(1/3);
x2=1:-0.1:-1; y2=2-x2.^2;
fill([x1,x2],[y1,y2],'c')
```



练习与思考题

1.美国总统选举前民意测验的抽样调查与计算面积的蒙特卡罗方法有何相同之处?

2．甲、乙两人在下午1点到2点之间独立地随机到达汽车站,这段时间内有四趟班车,开车时间分别为1:15, 1:30，1:45，2:00；问在: (1)见车就乘, (2)最多等一趟车;两种情况下, 两人同乘一辆车的概率多大?



3. 贝努里概型取 $p = 0.6$
如何设计随机实验

X	0	1
P	$1-p$	p