第 6 章　随机模拟方法

司守奎
烟台市，海军航空大学
Email：sishoukui@163.com
微信：sishoukui

随机模拟方法亦称 Monte Carlo 方法，是一种基于“随机数”的计算方法。很早以前人们已发现和利用其基本思想，例如，19 世纪数学家蒲丰（Buffon）用投针实验的方法来计算圆周率。

6. 1　Matlab 产生随机数的函数

Matlab 工具箱提供了 30 多种随机数发生函数，我们列举出一些主要函数见表 6.1。

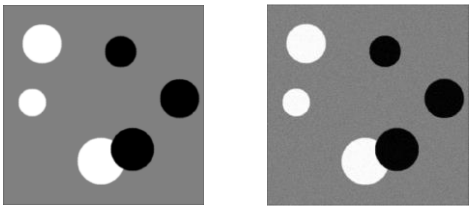
表 6.1　随机数产生函数

函数名称	函数说明	调用格式
betarnd	β 分布的随机数	R=betarnd(A,B,m,n)
binornd	二项分布随机数	R=binornd(N,P,m,n)
chi2rnd	χ^2 分布随机数	R=chi2rnd(V,m,n)
exprnd	指数分布随机数	R=exprnd(MU,m,n)
frnd	F 分布随机数	R=frnd(V1,V2,m,n)
gamrnd	γ 分布随机数	R=gamrnd(A,B,m,n)
geornd	几何分布随机数	R=geornd(P,m,n)
hygernd	超几何分布随机数	R=hygernd(M,K,N,m,n)
normrnd	正态分布随机数	R=normrnd(MU,SIGMA,m,n)
lognrnd	对数正态分布随机数	R=lognrnd(MU,SIGMA,m,n)
nbinrnd	负二项分布随机数	R=nbinrnd(R,P,m,n)
ncfrnd	非中心 F 分布随机数	R=ncfrnd(NU1,NU2,DELTA,m,n)
nctrnd	非中心 t 分布	R=nctrnd(V,DELTA,m,n)
ncx2rnd	非中心 χ^2 分布随机数	R=ncx2rnd(V,DELTA,m,n)
poissrnd	泊松分布随机数	R=poissrnd(LAMBDA,m,n)
rand	(0,1)区间上均匀分布随机数	R=rand(m,n)
randi	均匀分布的伪随机整数	R=randi([imin,imax],m,n)
randn	标准正态分布的随机数	R=randn(m,n)
raylrnd	Rayleigh 分布随机数	R=raylrnd(B,m,n)
trnd	t 分布随机数	R=trnd(V,m,n)
unidrnd	离散均匀分布随机数	R=unidrnd(N,m,n)
unifrnd	连续均匀分布随机数	R=unifrnd(A,B,m,n)
wblrnd	Weibull 分布随机数	R=weibrnd(A,B,m,n)

例 6.1　对 Matlab 图像处理工具箱的灰度图像 circlesBrightDark.png 添加高斯白噪声，其中均值为 0，标准差为 10。

```
clc, clear, close all
A=double(imread('circlesBrightDark.png'));
mu=0; sigma=10;
N=normrnd(mu,sigma,size(A));          %高斯白噪声矩阵
AN=A+N;                               %加噪声后的图像
subplot(121), imshow(uint8(A))
subplot(122), imshow(uint8(AN))
```

输出图像如图 6.1 所示。



（A）原始图像　　　（B）噪声图像

图 6.1　circlesBrightDark.png **加噪对比图像**

若 d 维随机向量 $(X_1, X_2, \cdots, X_d)$ 的密度函数为

$$f(\boldsymbol{x}) = \frac{1}{(2\pi)^{d/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu}) \right],$$

则称 $(X_1, X_2, \cdots, X_d)$ 服从均值向量为 $\boldsymbol{\mu}$ 、协方差矩阵为 $\boldsymbol{\Sigma}$ 的 d 维正态分布，记为 $(X_1, X_2, \cdots, X_d) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$。在上述密度函数中，$|\boldsymbol{\Sigma}|$ 表示 $\boldsymbol{\Sigma}$ 的行列式，且 $\boldsymbol{\Sigma}$ 为对称的正定矩阵。

在 Matlab 中，产生多维正态分布随机数的函数为 mvnrnd，其调用格式为

```
R = mvnrnd(mu,sigma,n)
```

其中，mu 为均值向量，sigma 为协方差矩阵，输出的矩阵共有 n 行。

例 6.2　分别生成服从 $N(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ 和 $N(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ 分布的 500 个数据点，其中

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{bmatrix} 1 & 1.5 \\ 1.5 & 3 \end{bmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{bmatrix} 2 & 1.2 \\ 1.2 & 3 \end{bmatrix}.$$

绘出上述数据点的散点图，要求第一个正态总体用加号表示点，第二个正态总体用圆圈表示点。

```
clc, clear, close all
rng(3)        %进行一致性比较
mu1=[1;2]; s1=[1,1.5;1.5,3];
mu2=[6;2]; s2=[2,1.2;1.2,3];
r1=mvnrnd(mu1,s1,500); r2=mvnrnd(mu2,s2,500);
plot(r1(:,1),r1(:,2),'+',r2(:,1),r2(:,2),'o')
legend({'正态总体一','正态总体二'})
```

绘出的散点图如图 6.2 所示。

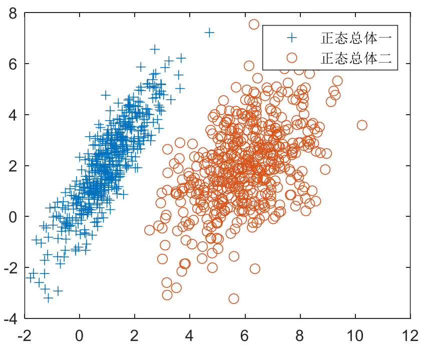


图 6.2　**两个二维正态分布的散点图**

6.2　随机模拟举例

例 6.3　设计随机试验求 π 的近似值。

在单位正方形中取 1000000 个随机点 (x_i, y_i)，$i = 1, 2, \cdots, 1000000$，统计点落在 $x^2 + y^2 \le 1$ 内的频数 n。则由几何概率知，任取单位正方形内一点，落在单位圆内部（图 6.3 第一象限部分）的概率为 $p = \frac{\pi}{4}$，由于试验次数充分多，频率近似于概率，有 $\frac{n}{1000000} \approx \frac{\pi}{4}$，所以 $\pi \approx \frac{4n}{1000000}$。

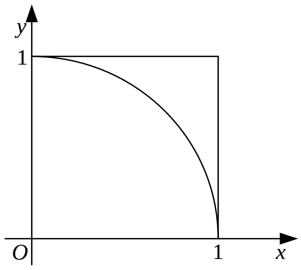


图 6.3　求几何概率的示意图

模拟的 Matlab 程序如下：

```
clc, clear, N=10^6; rng(1)   %进行一致性比较
x=rand(1,N); y=rand(1,N); %生成随机点的 x,y 坐标
n=sum(x.^2+y.^2<=1); %统计落在单位圆内部的点数
s=4*n/N   %计算 pi 的近似值
```

求得 π 的近似值为 3.1436。

例 6.4　蒲丰投针问题

蒲丰（Buffon）是法国著名学者，于 1977 年提出了用随机投针试验求圆周率 π 的方法。在平面上画有等距离为 a 的一些平行直线，向平面上随机投掷一长为 l（$l < a$）的针。设投针次数为 n，针与平行线相交次数为 m。试求针与一平行线相交的概率 p，并利用计算机模拟求 π 的近似值。

（1）问题分析与数学模型：

令 M 表示针的中点，针投在平面上时，x 表示点 M 与最近一条平行线的距离，φ 表示针与平行线的交角，如图 6.4 所示。显然 $0 \le x \le a/2$，$0 \le \varphi \le \pi$。

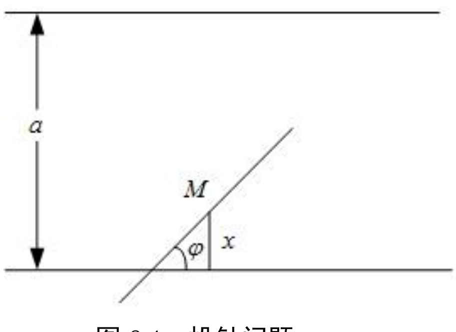


图 6.4　投针问题

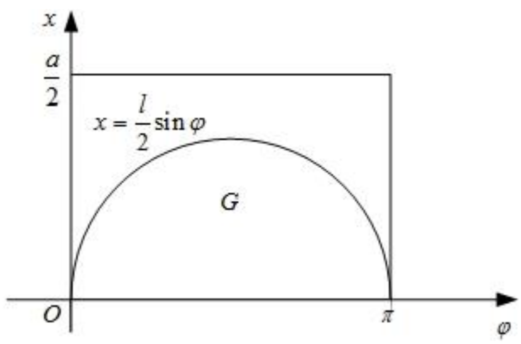


图 6.5　样本空间及事件的几何表示

随机投针的概率含义是：针的中点 M 与平行线的距离 x 均匀地分布于区间 $[0, a/2]$ 内，针与平行线交角 φ 均匀分布于区间 $[0, \pi]$ 内，x 与 φ 是相互独立的。而针与平行线相交的充分必要条件是 $x \le \frac{l}{2}\sin\varphi$。

将针投掷到平面上理解为向样本空间$\Omega=\{(x,\varphi)\,|\,0\le x\le \frac{a}{2},0\le\varphi\le\pi\}$（如图 6.5 所示）内均匀分布地投掷点，求针与一平行线相交的概率 p，即求点(x,φ)落在

$$G=\{(x,\varphi)\,|\,0\le x\le \frac{l}{2}\sin\varphi,0\le\varphi\le\pi\}$$

中的概率，显然，这一概率为

$$p=\frac{\int_0^{\pi}\frac{l}{2}\sin\varphi d\varphi}{\frac{a}{2}\pi}=\frac{2l}{a\pi}.$$

这表明，可以利用投针试验计算π值。当投针次数n充分大，且针与平行线相交m次，可用频率m/n作为概率 p 的估计值，因此可求得π的估计值为

$$\pi\approx\frac{2nl}{am}.$$

历史上曾经有一些学者做了随机投针试验，并得到了π的估计值。表 6.2 列出了两个最详细的试验情况。

表 6.2　历史上蒲丰投针试验

试验者	a	l	投针次数n	相交次数m	π的近似值
Wolf（1853）	45	36	5000	2532	3.1596
Lazzarini（1911）	3	2.5	3408	1808	3.1415929

（2）蒲丰随机投针试验的计算机模拟

真正使用随机投针试验方法来计算π值，需要作大量的试验才能完成。可以把蒲丰随机投针试验交给计算机来模拟实现，具体做法如下：

①产生互相独立的随机变量Φ和X的抽样序列$\{(\varphi_i,x_i)\,|\,i=1,\cdots,n\}$，其中$\Phi\sim U(0,\pi)$，$X\sim U(0,a/2)$。

②检验条件$x_i\le\frac{l}{2}\sin\varphi_i$（$i=1,\cdots,n$）是否成立，若上述条件成立，则表示第$i$次试验成功，即针与平行线相交（$(\varphi_i,x_i)\in G$），如果在$n$次试验中成功次数为$m$，则$\pi$的估计值为$\frac{2nl}{am}$。

下面是蒲丰投针的 Matlab 程序，其中的a、l、n的取值与 Wolf 实验相同。

```
clc, clear, rng(2)              %进行一致性比较
a=45; L=36; n=5000;
x=unifrnd(0,a/2,1,n);           %产生 n 个[0,a/2]区间上均匀分布的随机数
phi=unifrnd(0,pi,1,n);          %产生 n 个[0,pi]区间上均匀分布的随机数
m=sum(x<=L*sin(phi)/2);         %统计满足 x<=L*sin(phi)/2 的次数
pis=2*n*L/(a*m)                 %计算近似值
```

求得π的近似值为 3.1311。

（3）说明

随机模拟方法是一种具有独特风格的数值计算方法。这一方法是以概率统计理论为主要基础，以随机抽样为主要手段的广义的数值计算方法。它用随机数进行统计试验，把得到的统计特征（均值和概率等）作为所求问题的数值解。

例 6.5　敌坦克分队对我方阵地实施突袭，其到达规律服从泊松分布，平均每分钟到达4 辆。试模拟：

（1）敌坦克在 6 分钟内到达目标区的数量，以及在每分钟内各到达几辆坦克？

（2）在6分钟内每辆敌坦克的到达时刻。

解　（1）由题意知泊松分布的参数$\lambda=4$。使用poissrnd命令生成6个随机数，即代表各分钟内到达的坦克两数，分别为

4　5　0　3　2　3，

6分钟内到达目标区的坦克总辆数为17。

（2）两辆相邻到达的坦克的时间间隔服从参数为$\mu=1/4$的指数分布。模拟得到在6分钟内来到了24辆坦克，具体的到达时刻见下面程序运行结果。

```
clc, clear, rng(1)          %进行一致性比较
lambda=4; mu=1/4;
a=poissrnd(lambda,1,6)      %生成6个随机数
b=sum(a)                    %6分钟内到达的坦克总辆数

t=exprnd(mu,1,30);          %生成30辆坦克的到达时间间隔
T=cumsum(t)                 %计算30辆坦克的到达时刻
ind=find(T<=6)
TT=T(ind)                   %提取6分钟之内坦克到达时刻
```

例 6.6　使用随机模拟方法计算$I=\int_0^2 \sin(x^2)dx$。

解　令随机变量$X \sim U(0,2)$，X的概率密度函数

$$f(x)=\begin{cases}\dfrac{1}{2}, & x\in[0,2],\\ 0, & 其他.\end{cases}$$

则

$$I=2\int_0^2 \sin(x^2)f(x)dx=2E[\sin(X^2)].$$

根据大数定律，可以用均值来近似期望，即I的近似值

$$\hat{I}=2\frac{1}{N}\sum_{i=1}^{N}\sin(x_i^2),$$

其中，$x_1,x_2,\cdots,x_N$为服从$U(0,2)$分布的随机数。

计算得到I的近似值为0.8047，与积分的数值解0.8048相差很少。

```
clc, clear, rng(1)               %进行一致性比较
N=4e6                            %生成随机数的个数
x=unifrnd(0,2,1,N);              %生成区间[0,2]上均匀分布的随机数
I1=2*mean(sin(x.^2))             %计算积分的模拟值
I2=integral(@(x)sin(x.^2),0,2)   %计算积分的数值解
```

例 6.7（作战打击模拟）　在我方某前沿防守阵地，敌人以一个炮排（含两门火炮）为单位对我方进行干扰和破坏。为躲避我方打击，敌方对其阵地进行了伪装并经常变换射击地点。经过长期观察发现，我们指挥所对敌方目标的指示有 50%是准确的；而我方火力单位在指示正确时，有$1/3$的射击效果能毁伤敌人一门火炮，有$1/6$的射击效果能全部消灭敌人（即毁伤两门火炮）。随机模拟对敌人实施的 20 次打击结果，并确定有效射击的比率及平均每次毁伤敌方火炮的平均值。

解　准确发现敌人目标的可能性为 50%，用生成区间[0,1]上随机数命令 rand 来实现，

当 rand≤0.5 时，表示准确发现目标；否则，没有正确发现目标。在指示正确时，击中敌人零门、一门、两门火炮的可能性分别为3/6 、2/6 、1/6 ，也用生成区间[0,1]上随机数命令 rand 来实现，当 rand≤0.5 时，击中敌人零门火炮；当 0.5<rand≤5/6时，击中敌人一门火炮；当 5/6<rand≤1 ，击中敌人两门火炮。

```
clc, clear, rng(3)                %进行一致性比较
a=rand(20,1);                     %模拟是否准备发现目标
R=cell(21,5);                     %将模拟结果保存到元胞数组中
R(1,:)={'实验序号','生成随机数','指示正确','生成随机数','毁伤火炮数'};
R([2:end],1)=num2cell([1:20]');
R([2:end],2)=num2cell(a);
ind1=find(a<=0.5); ind2=find(a>0.5);
R(ind1+1,3)=cellstr('Yes'); R(ind2+1,3)=cellstr('No');
b=rand(size(ind1));               %模拟毁伤火炮的随机数
c=ones(size(b));                  %毁伤火炮数初始化
c(b<=0.5)=0; c(b>5/6)=2;          %模拟有效射击毁伤火炮数
R(ind1+1,4)=num2cell(b);
R(ind1+1,5)=num2cell(c)
writecell(R,'ti6_7.xlsx')
E1=sum(c>0)/20                    %计算有效射击的比率
E2=sum(c)/20                      %计算平均每次毁伤敌方火炮的门数
```

输出有效射击的比率E1=0.25，平均每次毁伤敌方火炮的门数E2=0.35。20次作战的模拟结果保存在Excel文件ti6_7.xlsx中，结果显示如表6.3所示。

表 6.3 作战打击模拟

实验序号	生成随机数	指示正确	生成随机数	毁伤火炮数
1	0.5508	No		
2	0.7081	No		
3	0.2909	Yes	0.2835	0
4	0.5108	No		
5	0.8929	No		
6	0.8963	No		
7	0.1256	Yes	0.6931	1
8	0.2072	Yes	0.4405	0
9	0.0515	Yes	0.1569	0
10	0.4408	Yes	0.5446	1
11	0.0299	Yes	0.7803	1
12	0.4568	Yes	0.3064	0
13	0.6491	No		
14	0.2785	Yes	0.2220	0
15	0.6763	No		
16	0.5909	No		
17	0.0240	Yes	0.3880	0
18	0.5589	No		
19	0.2593	Yes	0.9364	2
20	0.4151	Yes	0.9760	2

例 6.8 某报童以每份 0.3 元的价格买进报纸，以 0.5 元的价格出售。根据长期统计，报纸每天的销售量及概率如？所列。已知当天销售不出去的报纸，将以每份 0.15 的价格退还报社。试用随机模拟方法确定报童每天买进多少份报纸，才能使平均总收入最大？

表 6.4 每天的销售量及概率

销售量	200	210	220	230	240	250

概率	0.1	0.2	0.4	0.15	0.1	0.05

解　假设报童每天买进报纸数量为n，显然n的取值范围为$200\sim 250$。每天报纸的销量X（市场需求量）有6种情形。由表6.4中的概率，得到累积概率和生成随机数与对应事件间关系如表6.5所列。

表 6.5　累积概率和生成随机数与对应事件间关系

累积概率	0.1	0.3	0.7	0.85	0.95	1
随机数区间	$[0,0.1]$	$(0.1,03]$	$(0.3,0.7]$	$(0.7,0.85]$	$(0.85,0.9]$	$(0.9,1]$
对应销售量	200	210	220	230	240	250

报童每天的收入

$$Y=\begin{cases}0.2n, & X\ge n,\\ 0.2X-0.15(n-X), & X<n.\end{cases}$$

我们模拟一年365天，求365天收入的均值，确定出最佳的进货量为$n=220$，平均每天收入的最大值为42.6192元。

```
clc, clear, rng(1)                  %进行一致性比较
n=200:10:250;                       %报童每天买进报纸的数量
p=[0.1,0.2,0.4,0.15,0.1,0.05];
pp=cumsum(p);
a=0.3; b=0.5; c=0.15; N=365;  %模拟的天数
Y=[];                               %所有天数的收入记录
for i=1:N
    r=rand;                         %生成一个随机数
    ind=find(r<pp,1);               %确定哪种需求量
    x=n(ind);                       %确定当天的需求量
    y=[(b-a)*n(1:ind),(b-a)*x-(a-c)*(n(ind+1:end)-x)];
    Y=[Y;y];
end
my=mean(Y)                          %求每天的平均收入
ind2=find(my==max(my))              %求哪种购进量收入最大
nx=n(ind2)                          %求对应的购进量
```

习题6

1. 炮弹射击的目标为一椭圆$\frac{x^2}{60^2}+\frac{y^2}{80^2}=1$所围成的区域的中心，当瞄准目标的中心发射时，受到各种因素的影响，炮弹着地点与目标中心有随机偏差。设炮弹着地点围绕目标中心呈二维正态分布，且偏差的标准差在x和y方向均为100米，并相互独立，用Monte Carlo法计算炮弹落在椭圆区域内的概率，并与数值积分计算的概率进行比较。

2. 利用Monte Carlo方法，模拟掷骰子各面出现的概率。

3. 利用 Monte Carlo 方法，求积分$\int_1^2 \frac{\sin x}{x}dx$，并与数值解的结果进行比较。

4. 使用随机模拟方法计算积分$I=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{2}\exp\left(-\frac{x^2}{2}\right)dx$.

5. 使用 Monte Carlo 方法，求椭球面$\frac{x^2}{3}+\frac{y^2}{6}+\frac{z^2}{8}=1$所围立体的体积。

6. 分别随机生成服从$N(\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1)$和$N(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_2)$的 100 个数据点，并绘制散点图，其中

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & -1.5 \\ 1 & -1.5 & 4 \end{bmatrix}, \quad \boldsymbol{\Sigma}_3 = \begin{bmatrix} 1.5 & -0.5 & 0.5 \\ -0.5 & 2 & 0.5 \\ 0.5 & 0.5 & 2.5 \end{bmatrix}.$$