蒙特卡洛方法

目录

1 • 蒙特卡洛方法概述

2 • 蒙特卡洛方法求解问题的一般步骤

3 • 蒙特卡洛方法应用举例

4 • 蒙特卡洛仿真中的统计精度分析

1

蒙特卡洛方法概述

蒙特卡洛方法概述

定义

- 蒙特卡洛(**Monte Carlo**)方法，又称**随机抽样或统计模拟方法**
- 泛指所有基于统计采样进行数值计算的方法
- 以概率统计理论为基础
- 通过对随机样本进行统计计算得到问题的解

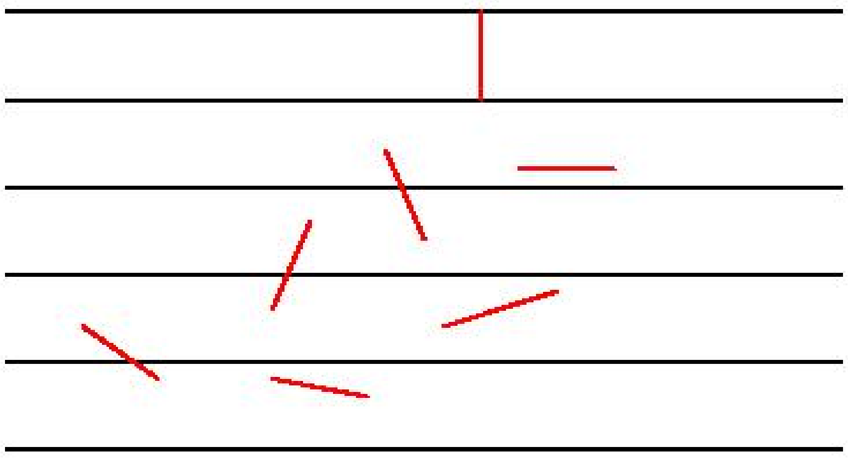
蒙特卡洛方法概述

历史

- 早在17世纪，人们就知道用事件发生的频率来确定事件发生的概率。
- 18世纪后半叶：著名的布冯(Buffon)投针实验
- 20世纪40年代：电子计算机的发展使得大量随机实验成为可能
- **Monte Carlo**：摩纳哥的一座城市，著名赌城之一

蒙特卡洛方法概述

Buffon投针实验



1．平行线间距＝针长＝s

（特例，原实验针长可小于或等于间距 ）

2．针与平行线垂直方向夹角为α，则相交概率为：

$$P=\frac{s|\cos\alpha|}{s}=|\cos\alpha|$$

3．各向同性随机投针，则夹角α在$[0,\pi]$均匀分布，所以：

$$E(|\cos\alpha|)=\int_0^{\pi}|\cos\alpha|\,f(\alpha)d\alpha=\frac{2}{\pi}$$

4．设投针N次，相交次数为M，则相交概率的期望值：

$$E(|\cos\alpha|)=\frac{2}{\pi}\equiv\frac{M}{N}\Rightarrow\pi=\frac{2N}{M}$$

$N\to\infty$大数定理

蒙特卡洛方法概述

Buffon投针实验

1．平行线间

（特例，

2．针与平

试验者	时间	投掷次数	相交次数	圆周率估计值
Wolf	1850年	5000	2532	3.1596
Smith	1855年	3204	1218	3.1554
C.De Morgan	1860年	600	382	3.137
Fox	1884年	1030	489	3.1595
Lazzerini	1901年	3408	1808	3.1415929
Reina	1925年	2520	859	3.1795

3．各向同性

$$E(|\cos\alpha|)=\int_0^{\pi}|\cos\alpha|\,f(\alpha)d\alpha=\frac{2}{\pi}$$

4

Buffon投针实验是第一个用几何形式表达概率问题的例子，他首次使用随机实验处理确定性数学问题，为概率论的发展起到一定的推动作用

$$E(|\cos\alpha|)=\frac{}{\pi}\approx\frac{}{N}\Rightarrow\pi=\frac{}{M}$$

$N\to\infty$大数定理

基本思想与理论依据

➢基本思想

- □针对待求问题，根据物理现象本身的**概率模型**，或**人为构造合适的依赖随机变量的概率模型**，使某些随机变量的统计量为待求问题的解

➢理论依据

- □**大数定理：当样本数足够大时，样本的平均值趋近于随机变量期望值**
- □**中心极限定理：实验次数足够多时，独立同分布随机变量的和趋于高斯分布**

蒙特卡洛方法概述

收敛性:大数定理

- 蒙特卡洛方法可以看作由随机变量X的样本X_1，X_2，…，X_N的**算术平均值**：

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

作为所求解期望的近似值。由**大数定理**可知，如果X_1，X_2，…，X_N独立同分布，且具有有限期望值（$E(X)<\infty$），则

$$P\left(\lim_{N\to\infty}\overline{X}_N = E(X)\right) = 1$$

- **即随机变量X的样本的算术平均值$\overline{X}_N$，当样本数N充分大时，以概率1收敛于它的期望值$E(X)$。**

收敛性:中心极限定理

- 蒙特卡洛方法的**近似值与真值的误差**问题，概率论的中心极限定理给出了答案。该定理指出，如果随机变量序列X_1，X_2，…，X_N独立同分布，且具有有限非零的方差σ^2，即

$$0 \neq \sigma^2 = \int (x - E(X))^2 f(x)dx < \infty$$

当N足够大时

$$\frac{\sum_{i=1}^{N} X_i - NE(x)}{\sigma\sqrt{N}}$$ 满足正态分布：$\mathcal{N}(0,1)$

整理得：

$$\frac{\sum_{i=1}^{N} X_i - NE(x)}{\sigma\sqrt{N}} = \frac{\sqrt{N}\ \left(\overline{X_N} - E(x)\right)}{\sigma}$$

求解概率：

$$P(|\overline{X_N} - E(x)| < \frac{\lambda_\alpha \sigma}{\sqrt{N}}) = P(|\frac{\sqrt{N}\left(\overline{X_N} - E(x)\right)}{\sigma}| < \lambda_\alpha)$$

蒙特卡洛方法概述

收敛性:中心极限定理

- 则当N充分大时，有如下的近似式：

$$P\left(\left|\overline{X}_N - E(X)\right| < \frac{\lambda_\alpha \sigma}{\sqrt{N}}\right) \approx \frac{2}{\sqrt{2\pi}} \int_0^{\lambda_\alpha} e^{-t^2/2} dt = 1-\alpha$$

- 其中1-α称为置信度。

$$\alpha = 2\mathrm{Q}(\lambda_\alpha) = \mathrm{erfc}(\frac{\lambda_\alpha}{\sqrt{2}})$$

$$Q(x) = \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} exp(-\frac{1}{2}t^2)dt$$

$$Q(x) = 1 - \phi(x) = \frac{1}{2} erfc(\frac{1}{\sqrt{2}}x)$$

- 这表明，不等式

$$\left|\overline{X}_N - E(X)\right| < \frac{\lambda_\alpha \sigma}{\sqrt{N}}$$

近似地以概率1—α成立，且误差收敛速度的阶为：$O(N^{-1/2})$

蒙特卡洛方法概述

蒙特卡洛误差

$$\text{MC方法的误差：}\varepsilon = \frac{\lambda_\alpha \sigma}{\sqrt{N}}$$

- 对于该误差，λ_α与置信度1-α是一一对应的，根据问题的要求确定出置信水平后，查标准正态分布表，就可以确定出λ_α。

给出几个常用的α与λ_α的数值：

$$\alpha = 2Q(\lambda_\alpha)$$

α	**0.5**	**0.05**	**0.003**
λ_α	**0.6745**	**1.96**	**3**

(1)MC方法的误差为**概率误差**，这与其他数值计算方法是有区别的

(2) 均方差σ是未知的，必须用**估计值**来代替，在计算所求量的同时，可计算出估计值

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} X_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} X_i\right)^2}$$

减小误差

- 显然当给定置信度1-α（λ_α）后，误差ε由σ和N决定。

$$\varepsilon = \frac{\lambda_\alpha \sigma}{\sqrt{N}}$$

要减小ε：

（1）增大试验次数N。在σ固定的情况下，要把精度提高一个数量级，试验次数N需增加两个数量级。因此，单纯增大N不是一个有效的办法。

（2）减小估计的均方差σ，比如降低一半，那误差就减小一半，这相当于N增大四倍的效果。

蒙特卡洛方法概述

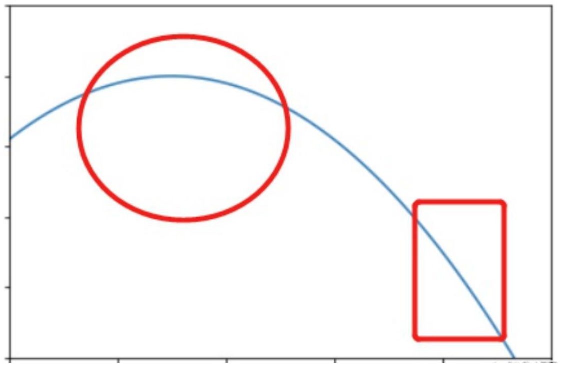
减小误差

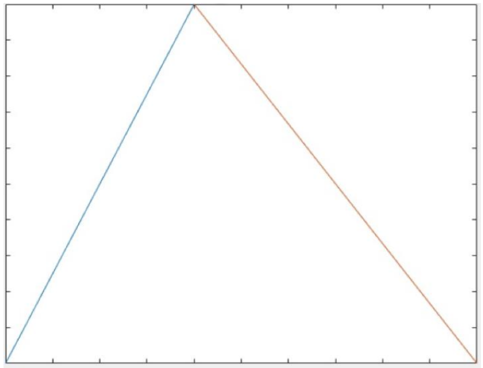
降低方差σ的方法：

（1）重要性采样

- 带有重要性权重的采样
- 不对原随机分布采样，而是对一个偏向“重要样本”的新的分布进行采样
- 重要性采样不只适用于仿真，在人工智能算法研究中同样重要。

举例





- 左图信号PDF $\pi(x)$：圆形区域的函数值对积分的贡献比方形区域大，在抽样时以更大概率抽取圆形区域的样本，以提高估计的准确度。
- 右图采样分布PDF $p(x)$
- 重要性权重：$\dfrac{\pi(x_i)}{p(x_i)}$

蒙特卡洛方法概述

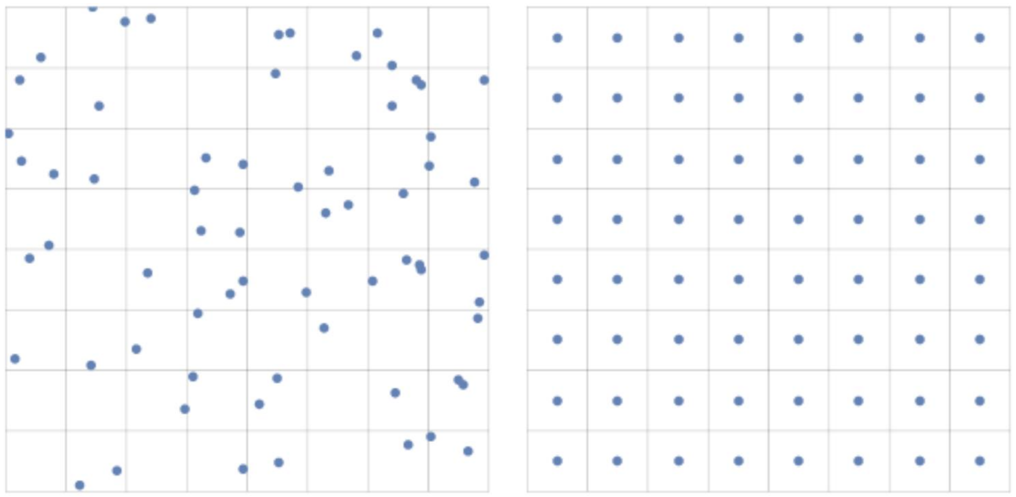
减小误差

降低方差σ的方法：

（2）分层采样

- 将样本空间分为若干个无交叠的子区域
- 先按不同区域的概率选择区域，再根据区域内的概率分布获得样本 。

举例



- 左图：纯随机采样。
- 右图：分层采样。假设在每份中取一个采样点

两个方法的目的，都是在采样次数有限的情况下减小估计的方差

蒙特卡洛方法概述

MC优点（1）

（1）能够比较逼真地描述具有随机性质的事物的特点

- 用蒙特卡洛方法解决实际问题，可以直接从**实际问题本身**出发，而不从方程或数学表达式出发
- 具有直观、形象的特点

（2）适合多维复杂问题

- 在具有随机性质的问题中，如果考虑的系统很复杂，同时存在多个维度，难以用一般数值方法求解
- 使用蒙特卡洛方法则不会有原则上的困难

MC优点（2）

（3）收敛速度与问题的维数无关

- 在给定置信度情况下，MC方法的误差为$O(N^{-1/2})$，与**问题本身的维数**无关。
- 维数的变化，只引起抽样时间及估计量计算时间的变化，不影响误差

一般数值方法，比如计算定积分时，计算时间随维数的幂次方而增加，而且由于点数与维数的幂次方成正比，需占用**相当数量的计算机内存**，这些都是一般数值方法计算高维积分时难以克服的问题。

蒙特卡洛方法概述

MC优点（2）

（4）具有同时计算多个方案与多个未知量的能力

- 对于那些需要计算多个方案的问题，使用蒙特卡洛方法有时不需要像常规方法那样逐个计算，而可以同时计算所有的方案，其全部计算量几乎与计算一个方案的计算量相当
- 使用蒙特卡洛方法还可以同时得到若干个所求量

（5）误差容易确定

- 根据蒙特卡洛方法的误差公式，可以在计算所求量的同时计算出误差

蒙特卡洛方法概述

MC缺点

（1）收敛速度慢

- 蒙特卡洛方法的**收敛速度为**$O(N^{-1/2})$，一般不容易得到**精确度较高的近似结果**。

（2）误差有随机性

- 由于蒙特卡洛方法的误差是在一定置信水平下估计的，所以它的**误差具有随机性**，而不是一般意义下的误差。

2

蒙塔卡洛方法的一般步骤

蒙塔卡洛方法的一般步骤

适用类型和步骤

- 蒙特卡洛方法适用于两类问题：
  - 本身就具有随机性的问题；
  - 能够**转化为概率模型进行求解**的确定性问题。
- 蒙特卡洛方法求解问题的三个步骤：
  - 构造或描述概率过程
  - 从已知概率分布抽样
  - 建立估计量

蒙塔卡洛方法的一般步骤

构造或描述概率过程

- 构造或描述概率过程实际上就是**建立随机试验模型**
  - 构造概率过程是对确定性问题而言的（比如pi的估算）
  - 描述概率过程是对随机性问题而言的
  - 不同的问题所需要建立的随机试验模型各不相同

蒙塔卡洛方法的一般步骤

从已知概率分布抽样

- 所谓的从已知概率分布抽样指的是随机试验过程
- 随机模型中必须要包含某些已知概率分布的随机变量或随机过程作为输入
- 进行随机试验的过程就是对这些随机变量的样本或随机过程的**样本函数作为输入产生相应输出的过程**
- 如何**产生已知分布的随机变量或随机过程**是蒙特卡洛方法中的一个关键问题

蒙塔卡洛方法的一般步骤

获得估计量

- 蒙特卡洛方法所得到的问题的解总是对真实解的一个估计；
- 本身也是一个随机变量；
- 这个随机变量是由随机试验模型输出通过统计处理得到的。

3

蒙塔卡洛方法应用举例

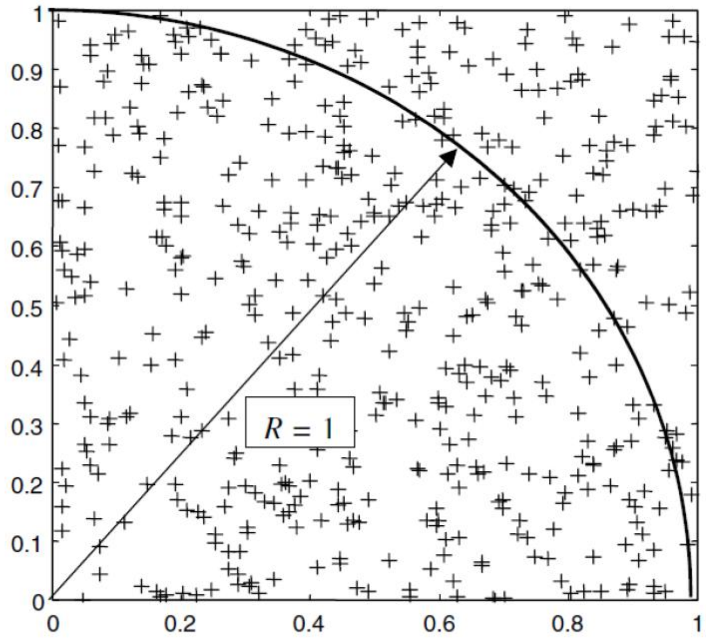
蒙特卡洛方法应用举例

π 值的估算（1）

1、构造随机试验模型

$$A_{quarter\_cycle} = \frac{1}{4}(\pi R^2)|_{R=1} = \frac{\pi}{4}$$

$$\pi = \frac{4A_{quarter\_cycle}}{A_{quarter}} \quad \text{（对于任意}R\text{）}$$



可以将π值估计问题归结为对四分之一圆面积的估计问题。

随机试验模型：在正方形区域中随机撒点，统计落在1/4圆区域内的点数

蒙特卡洛方法应用举例

π 值的估算（2）

2、从已知概率分布抽样

- 根据实验模型，需要在正方形区域内均匀撒点。
- 这实际上就是要对一个二维均匀分布进行采样。
- 二维均匀分布可以通过对**两个独立均匀分布**采样来实现。
- 在计算机上实现，每次采样就是**产生两个互相独立的[0,1]区间上均匀分布的随机数**

蒙特卡洛方法应用举例

π 值的估算（3）

3、获得统计量

$$\hat{\pi} = \frac{4N_B}{N} \approx \frac{4A_{quarter\_cycle}}{A_{quarter}} = 4P_B$$

蒙特卡洛方法应用举例

二进制反极性基带通信链路仿真（1）

➢用反极性信号波形 $s_0(t)=s(t)$ 和 $s_1(t)=-s(t)$ 来发送二进制信息0和1，$s(t)$是能量为E_b的任意波形，通过AWGN信道之后接收信号波形可以表示为

$$r(t)=\pm s(t)+n(t) \quad 0\le t\le T_b$$

➢这里$n(t)$是功率谱密度为 $\frac{N_0}{2}$ 的高斯白噪声。

➢用最佳接收机接收到的信号为

$$r=\pm E_b+n$$

➢这里 $n=\int_0^{T_b} n(t)s(t)dt$

蒙特卡洛方法应用举例

二进制反极性基带通信链路仿真（2）

- 其数学期望$E(n)=0$
- 方差：

$$
\begin{aligned}
\sigma^2 &= E(n^2) \\
&= \int_0^{T_b}\int_0^{T_b} E[n(t)n(\tau)]s(t)s(\tau)dtd\tau \\
&= \frac{N_0}{2}\int_0^{T_b}\int_0^{T_b} \delta(t-\tau)s(t)s(\tau)dtd\tau \\
&= \frac{N_0}{2}\int_0^{T_b} s^2(t)dt = \frac{N_0E}{2}
\end{aligned}
$$

- 通信原理中的结论：$P_e = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) = \frac{1}{2}erfc\left(\sqrt{\frac{E_b}{N_0}}\right)$

蒙特卡洛方法应用举例

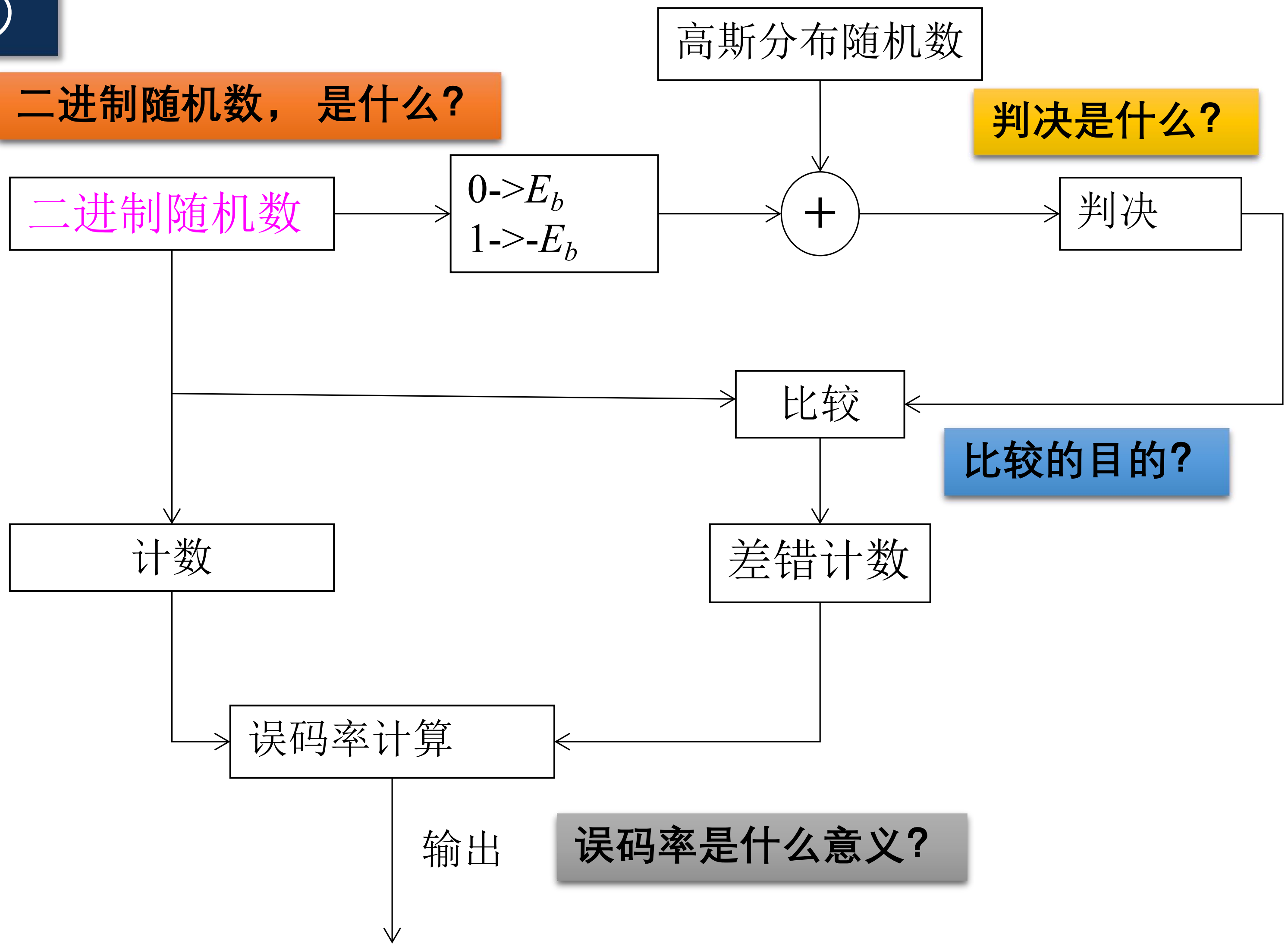
二进制反极性基带通信链路仿真（3）

- 用蒙特卡洛仿真方法来估计该系统的误码率
- 首先，**建立仿真模型**
- 由于假设接收机为理想的最佳接收机，因此没必要模拟整个信号的传输过程，而只关心最佳接收机输出端的信号组成。

高斯分布随机数，是什么？

二进制随机数，是什么？

判决是什么？



比较的目的？

误码率是什么意义？

蒙特卡洛方法应用举例

二进制反极性基带通信链路仿真（4）

- 该模型中包括对两个随机过程的采样
  - 模拟二进制信源的贝努利分布
  - 模拟最佳接收机输出端所包含的高斯噪声。
- 假设：
  - 信源和信道均无记忆

    无记忆是什么？
  - 信源和信道相互独立

    相互独立是什么？

二进制反极性基带通信链路仿真（4）

- 对这两个随机过程进行采样，只需要产生一系列相互独立、满足概率分布的随机变量即可
- 所关心的事件：**判决结果与实际传送比特相反**
- 建立的估计量：使用错误事件的出现相对频率近似系统的误码率，即：

$$\hat{P}_e = \frac{N_E}{N}$$

蒙特卡洛方法应用举例

二进制反极性基带通信链路仿真（5）

```
%Monte Carlo Simulation to estimate bit error rate performance of a binary
%antipodal communication system
clc
clear
N = [100, 1000, 10000, 100000,1000000];
EbN0_indB = -2:2:8;
Ebn0 = 10.^(EbN0_indB/10);
E=1;
times = 50;
err_rate = zeros(length(Ebn0),length(N),times);
```

蒙特卡洛方法应用举例

二进制反极性基带通信链路仿真（6）

```
for iiii=1:length(N)
    for iii=1:length(Ebn0)
        sigma = E/sqrt(2*Ebn0(iii));
        for ii = 1:times
            r = rand(1,N(iiii));
            source = double(r>=0.5);
            x=1-2*source;
            noise = randn(1,N(iiii))*sigma;
            y=E*x+noise;
            result = double(y<=0);
            error_num = sum(abs(result-source)>1.e-6);
            err_rate(iii,iiii,ii) = error_num/N(iiii);
        end
    end
end
```

```
pe_theory = 0.5*erfc(sqrt(Ebn0));
variance = zeros(length(Ebn0),length(N));
for iii = 1:length(Ebn0)
    for iiii = 1:length(N)
        ber = err_rate(iii,iiii,:);
        ber = reshape(ber,1,times);
        relative_error = (ber-
pe_theory(iii))/pe_theory(iii);
        variance(iii,iiii) = var(relative_error);
    end
end
```

蒙特卡洛方法应用举例

二进制反极性基带通信链路仿真（6）

```
figure1 = figure();
axes1 = axes('Parent',figure1,'YScale','log', ...
        'YMinorTick','on',...
        'YMinorGrid','on',...
        'Position',[0.1274 0.1123 0.775 0.815],...
        'FontSize',14);
box('on');
grid('on');
hold('all');
% Create multiple lines using matrix input to semilogy
semilogy1 = semilogy(EbN0_indB,variance,'Parent',
            axes1,'LineWidth',1,...
            'Color',[0 0 0]);
set(semilogy1(1),'Marker','o','DisplayName','100bits');
set(semilogy1(2),'Marker','^','DisplayName','1000bits');
set(semilogy1(3),'Marker','square','DisplayName','10000bit
s');
set(semilogy1(4),'MarkerSize',8,'Marker','x','DisplayName',
'100000bits');
set(semilogy1(5),'Marker','diamond','DisplayName','10000
0bits');
% Create xlabel
xlabel({'Eb/N0(dB)'},'FontSize',14);
% Create ylabel
ylabel({'误差方差'},'FontSize',14);
% Create legend
legend1 = legend(axes1,'show');
set(legend1,'Position',[0.1286 0.7088 0.1964 0.2177]);
```

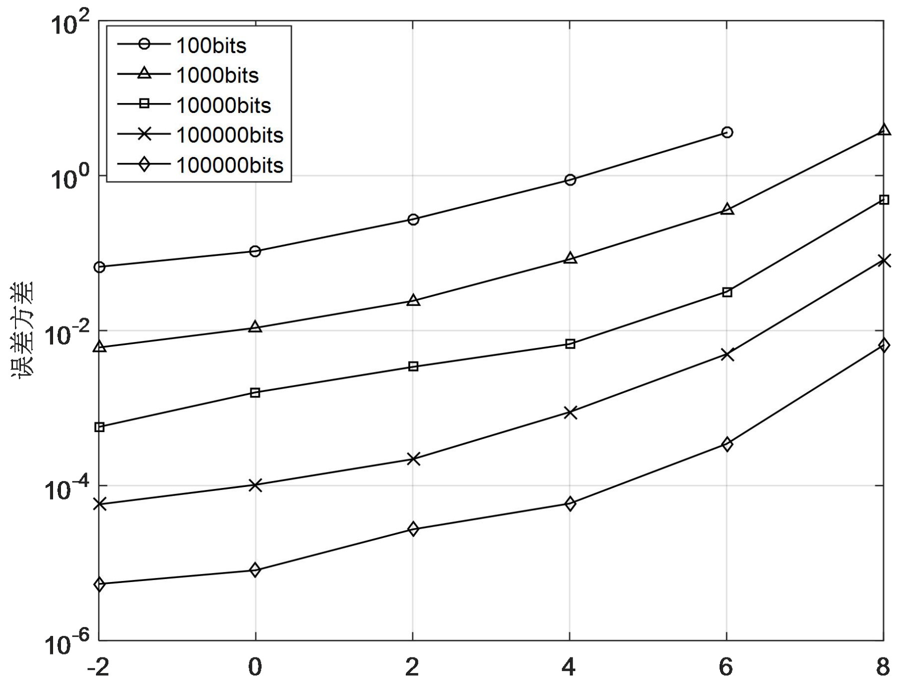
蒙特卡洛方法应用举例

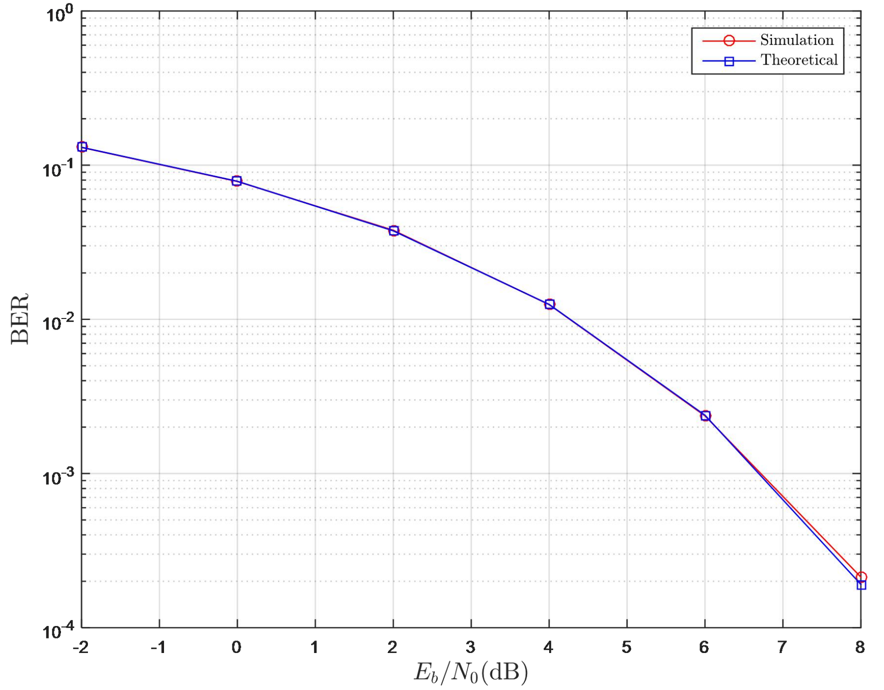
```
BER1 = err_rate(:,5,1);
Pe_theory = 0.5*erfc(sqrt(Ebn0));
figure;emilogy(EbN0_indB,BER1,’-ro’);
hold on; grid on;
semilogy(EbN0_indB,pe_theory,'-bs');
xlabel({'Eb/N0(dB)'},'FontSize',14);
ylabel({'BER'},'FontSize',14);
hl = legend('Simulation', 'Theoretical');
set(hl,'interpreter','latex');
hx = xlabel({'$E_b/N_0$(dB)'},'FontSize',14);
set(hx,'interpreter','latex');
hy = ylabel({'BER'},'FontSize',14);
set(hy,'interpreter','latex');
```

蒙特卡洛方法应用举例

二进制反极性基带通信链路仿真（6）

- 在同一信噪比条件下，仿真量越大相对误差的偏差越小
- 如果固定仿真量，相对误差的方差则随着信噪比的提高而迅速增大
- **问题：**
  - **对于这种问题至少要做多少次仿真才能得到接近于真实解的估计值呢？**





4

蒙特卡洛仿真中的统计精度分析

仿真量——精度的定量分析

- 蒙特卡洛仿真方法的本质是在**计算机上进行的随机试验和结果统计分析过程**，试验次数越多，得到的数据样本就越多，根据这些样本所得到的**统计结果精度和可信程度就越高**。
- **仿真量**是我们在进行蒙特卡洛仿真方案设计时不能回避的一个问题。
  - 仿真量**太小**，会造成仿真精度不能满足要求。
  - 仿真量**过大**，需要消耗大量的计算资源。
- 因此往往需要**根据仿真精度要求确定仿真量**，即需要解决仿真精度的定量分析问题。

仿真精度的衡量指标

- 仿真精度的两种衡量指标

  --**绝对精度**和**相对精度**

- 设数据的精确值是x_0，通过仿真得到的估计值为$\hat{x}$，$\hat{x}$是一个服从某种分布的随机变量。如果估计值$\hat{x}$有概率1-α落在某一区间$[x_0-\Delta, x_0+\Delta]$上，称区间$[x_0-\Delta, x_0+\Delta]$为置信度1-α的置信区间
- **绝对精度**：置信区间长度的一半，即Δ
- **相对精度**：绝对精度Δ与真值x_0的比值

蒙特卡洛仿真中的统计精度分析

由置信度和绝对精度确定仿真次数（1）

➢在很多仿真场景中，每次蒙特卡洛采样可以看做一次独立的**贝努利试验**

- 通信中传输一个数据符号，传输可能是正确的也可能是错误的
- 每次电话呼叫有可能接通也有可能呼叫阻塞
- 随机试验法求π时每次投下的点可能在四分之一圆内也可能在圆外

➢设一次独立的贝努利试验中事件E发生的概率为p，那么n次独立的贝努利试验中E发生的次数k服从二项分布，事件E发生的次数恰好为k次的概率是：

$$P_k(n,p)=\binom{n}{k}p^k(1-p)^{n-k}=\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$$

由置信度和绝对精度确定仿真次数（2）

- 如果以相对频率作为对概率p的估计，绝对精度要求小于δ，即

$$\left|\frac{k}{n}-p\right|<\delta$$

- k落在这个区间范围内的概率即置信度：

$$p_\delta = P(np-n\delta<k<np+n\delta)=\sum_{k=np-n\delta}^{np+n\delta} P_k(n,p)$$

- 给定置信度和绝对精度，可计算出所需要仿真的**最少次数*n***
- 但是，直接按上式计算比较复杂，特别是当需要的次数n比较大时，式中的组合数计算难以实现。可通过近似方法进行计算

由置信度和绝对精度确定仿真次数（3）

- 近似为正态分布
- 根据大数定理，当试验次数$n \to \infty$，试验中事件发生次数k服从均值为np方差为$np(1-p)$的正态分布，即

$$p\left(\left|\frac{k}{n}-p\right|<\delta\right) \approx \frac{1}{\sqrt{2\pi}}\int_{-a}^{a}\exp\left(-\frac{x^2}{2}\right)dx = 2\Phi(a)$$

- 其中

$$a = \frac{n\delta}{\sqrt{np(1-p)}}$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{0}^{x}\exp\left(-\frac{x^2}{2}\right)dx = \frac{1}{2}erf\left(\frac{x}{\sqrt{2}}\right)$$

由置信度和绝对精度确定仿真次数（4）

➢这样，给定置信度1-α和绝对精度以及事件的概率值p，就可以求解方程

$$erf\left(\frac{n\delta}{\sqrt{2np(1-p)}}\right)=1-\alpha$$

➢得出最小仿真次数n。如果概率值p未知，可用频率估计代替。

由置信度和绝对精度确定仿真次数（5）

➢已知某通信链路的设计传输错误概率为0.001，为了至少有95%的把握使仿真得到的错误概率与真值之间的误差在2×10^{-4}之内，问至少需要多少次仿真（即传输多少个独立符号）？

```
delta=2e-4; %绝对误差
p=1e-3;
alpha=0.05; %显著性水平
n=ceil(2*p*(1-p)/delta^2*(erfinv(1-alpha)^2))
```

n=95941

由置信度和绝对精度确定仿真次数（6）

- 除了利用正态分布来近似之外，还可以用泊松分布来近似。
- 泊松定理指出，在随机试验中事件的发生概率很小，随机试验次数很多的情况下，试验中事件的发生次数k近似服从参数为 $\lambda = np$ 的泊松分布，即

$$P_k(n,p) \approx \frac{(np)^k}{k!}\exp(-np)$$

$$P\left(\left|\frac{k}{n}-p\right|<\delta\right) \approx \sum_{k=\lfloor np-n\delta \rfloor}^{\lfloor np+n\delta \rfloor} \frac{(np)^k}{k!}\exp(-np) = F(np+n\delta)-F(np-n\delta)$$

- $F(x)$是参数为np的泊松分布的分布函数

由置信度和绝对精度确定仿真次数（7）

- 在前一例子的仿真系统中，设计传输错误率为10^{-3}，置信区为 $10^{-3}\pm2\times10^{-4}$，若总试验次数（独立传输符号数）为95941次，求仿真结果的置信度（分别用泊松分布和正态分布近似）。

```
delta=2e-4;%绝对误差
p=1e-3;
n=95941;
sigma = sqrt(n*p*(1-p));
P_delta_poiss=poisscdf(n*p+n*delta,n*p)-poisscdf(n*p-n*delta,n*p)
P_delta_norm=normcdf(n*p+n*delta,n*p,sigma)-normcdf(n*p-n*delta,n*p,sigma)
```

P_delta_poiss=0.953804688094842

P_delta_norm= 0.950000675727642

蒙特卡洛仿真中的统计精度分析

由置信度和相对精度确定仿真次数（1）

- 在很多情况下，相对精度比绝对精度更有意义，因此在实际仿真中往往需要通过置信度和相对精度确定最小仿真次数。
- 给定仿真的相对精度要求 $r=\delta/p$，则 $\delta=pr$，将之代入到

$$erf\left(\frac{n\delta}{\sqrt{2np(1-p)}}\right)=1-\alpha$$

- 并整理得到已知相对精度条件下的最小仿真次数为

$$n=\frac{2(1-p)}{pr^2}(erf^{-1}(1-\alpha))^2$$

由置信度和相对精度确定仿真次数（2）

- 反过来，如果给定仿真次数和置信度，则仿真结果的**相对精度**也可以计算出来

$$r=\sqrt{\frac{2(1-p)}{pn}}erf^{-1}(1-\alpha)\overset{p<<1}{\approx}\sqrt{\frac{2}{pn}}erf^{-1}(1-\alpha)$$

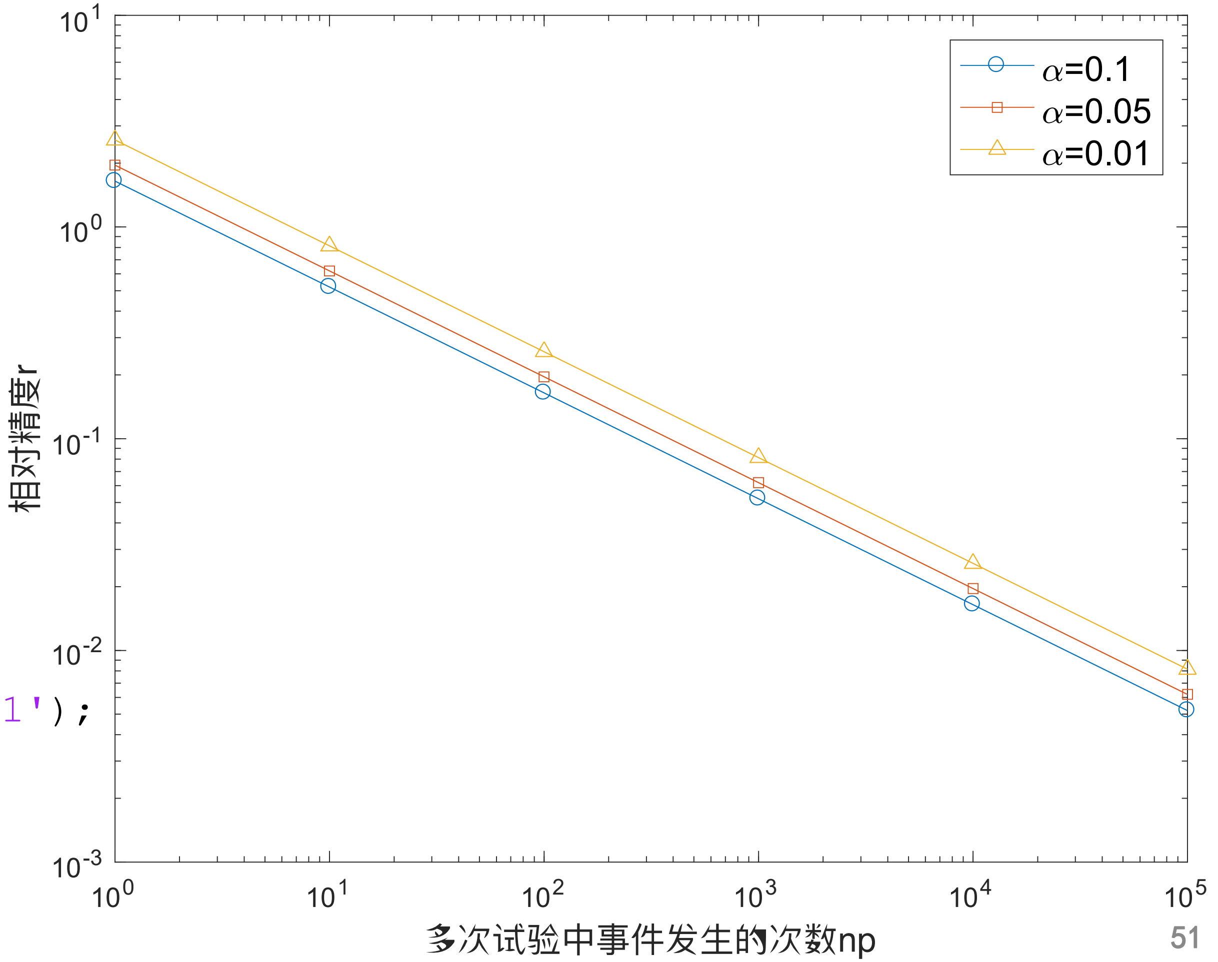
其中pn的物理意义是n次试验中事件出现的平均次数。

- 可以看出，我们关注的稀有事件(Rare events)出现次数越多，统计结果的相对精度越高（***r*值越小**）。

由置信度和相对精度确定仿真次数（3）

- 画出置信度为90%、95%和99%条件下，试验中事件发生次数*np*与相对精度*r*之间的关系曲线。

```
clear;
clc
close all
alpha=[0.1,0.05,0.01];
pn=[1 10 100 1000 10000 100000]';
for i=1:3
    r(:,i)=sqrt(2./pn).*erfinv(1-alpha(i));
end
f1=loglog(pn,r);
grid on
set(f1(1),'Marker','o');
set(f1(2),'Marker','square');
set(f1(3),'Marker','^');
hl = legend('\alpha=0.1', '\alpha=0.05', '\alpha=0.01');
hl = set(hl,'FontSize',14);
hx = xlabel('多次试验中事件发生的次数np');
set(hx,'FontSize',14);
hy = ylabel('相对精度r');
set(hy,'FontSize',14);
```



由置信度和相对精度确定仿真次数（4）

➢分析

- 如果要求试验结果的相对精度提高，那么就要使试验中观察得到事件发生次数成平方数量级增加。
- 在事件发生概率较小的情况下，将导致总试验次数大大增多。

➢这就是蒙特卡洛通常都需要较大数量仿真的原因。

由置信度和相对精度确定仿真次数（5）

- 一个通信系统，设传输错误概率很小，如果在仿真中每观察到10个、100个和1000个误码就进行一次误码率的统计，问得到的结果在95%置信度条件下的相对精度是多少？

- 以统计次数代替平均出错次数

```
alpha=0.05;
err_num=[10,100,1000];
r=sqrt(2./err_num)*erfinv(1-alpha)
```

r=[0.6198   0.1960   0.0620];

结论

- 可见，对于稀有事件仿真统计而言，只有稀有事件出现次数大于100，才能将相对误差不超过20%的可能性控制在5%以内
- 因此，在实际稀有事件仿真中，通常要求出错个数超过100
- 比如对传输误码率为10^{-3}的通信系统进行仿真时，要求仿真量大于10^{5}。

谢谢