REPORT 1

潘硕 PB24020526

2025 年 9 月 19 日

1 Question

用 Schrage 方法编写随机数子程序; 用指定间隔 (非连续 $l > 1$) 两个随机数作为点的坐标值绘出若干点的平面分布图, 用 $< x^n >$ 测试均匀性 (取不同量级的 N 值, 讨论偏差与 N 的关系), 用 $C(l)$ 测试其 2 维独立性 (总点数 $N > 10^7$)

2 Method

2.1 随机数生成

简单的均匀随机数的产生法是线性同余法, 其递推关系为:

$$I_{n+1} = (aI_n + b) \mod m \tag{1}$$

有一个被认为是"最低标准"的产生器 (16807 产生器) 是:

$$a = 16807, b = 0, m = 2^{31} - 1 = 2147483647 \tag{2}$$

但是该方法在计算机上直接实现时,aI_n 可能超过整数范围, 会导致溢出错误, 所以需要 Schrage 方法来避免溢出, 即:

$$I_{n+1} = \begin{cases} a(I_n \mod q) - r\lfloor I_n/q \rfloor, if \geqslant 0 \\ a(I_n \mod q) - r\lfloor I_n/q \rfloor + m, otherwise \end{cases} \tag{3}$$

其中取 $q = 127773, r = 2836$.

代码实现:

```
a = 16807;m = 2147483647;q = 127773;r = 2836;
N = 20000000;l = 10;
I = zeros(1,N+l);
```

2.2 种子生成

使用计算机报告当前时间的整数值构造初始值:$i_y, i_m, i_d, i_h, i_n, i_s$
可以设种子值为 $I_0 = i_y + 70(i_m + 12(i_d + ...31(i_h + 23(i_n + 59i_s))))$;
可以保证 100 年内不重复.
代码实现:

```
cl = fix(clock());
I(1) = cl(1) + 70*(cl(2) + 12 * (cl(3) +...
31 * (cl(4) + 23 * (cl(5) + 59 * cl(6)))));
```

3 Experiment

3.1 随机数平面分布图

采用相邻间隔随机数 (x_n, x_{n+l}), 作为点的坐标生成散点图:

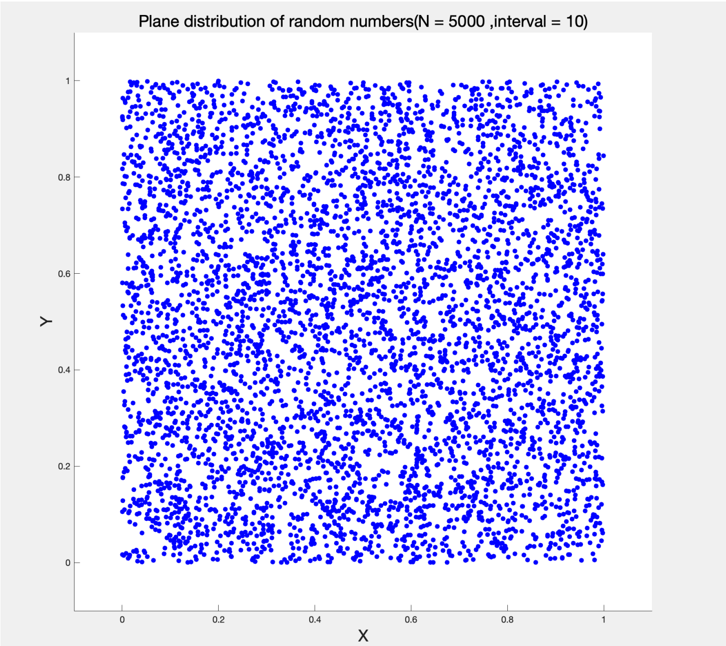

图 1: 随机数平面分布图

图中取间隔为 10, 点数 5000, 并未观察到明显的分层结构, 符合均匀分布.

3.2 均匀性测试

均匀性测试使用 $< x^n >$, 理论值为:

$$< x^n >= \int_0^1 x^n dx = \frac{1}{n+1} \tag{4}$$

差.

```
N_1 = 100
<x^1> = 0.50961 expectation = 0.5
<x^2> = 0.34501 expectation = 0.33333
<x^3> = 0.26231 expectation = 0.25
<x^4> = 0.21229 expectation = 0.2
<x^5> = 0.1784 expectation = 0.16667
N_1 = 1000
<x^1> = 0.50687 expectation = 0.5
<x^2> = 0.34186 expectation = 0.33333
<x^3> = 0.25724 expectation = 0.25
<x^4> = 0.20541 expectation = 0.2
<x^5> = 0.17028 expectation = 0.16667
N_1 = 10000
<x^1> = 0.49858 expectation = 0.5
<x^2> = 0.33179 expectation = 0.33333
<x^3> = 0.24865 expectation = 0.25
<x^4> = 0.19886 expectation = 0.2
<x^5> = 0.16569 expectation = 0.16667
N_1 = 100000
<x^1> = 0.49948 expectation = 0.5
<x^2> = 0.3329 expectation = 0.33333
<x^3> = 0.24969 expectation = 0.25
<x^4> = 0.19979 expectation = 0.2
<x^5> = 0.16653 expectation = 0.16667
N_1 = 1000000
<x^1> = 0.49935 expectation = 0.5
<x^2> = 0.33269 expectation = 0.33333
<x^3> = 0.24939 expectation = 0.25
<x^4> = 0.19943 expectation = 0.2
<x^5> = 0.16614 expectation = 0.16667
N_1 = 10000000
<x^1> = 0.49993 expectation = 0.5
<x^2> = 0.33328 expectation = 0.33333
<x^3> = 0.24996 expectation = 0.25
<x^4> = 0.19996 expectation = 0.2
<x^5> = 0.16664 expectation = 0.16667
```

图 2: 不同 N,k 下,x 的 k 阶矩计算

取用不同量级的 N, 计算偏差. 理论上, 偏差应正比于 $\sqrt{N}$.

以 $k=5$ 为例, 计算不同 N 下的偏差. 对 N 和偏差 err 取对数, 并进行线性拟合, 如下图所示:

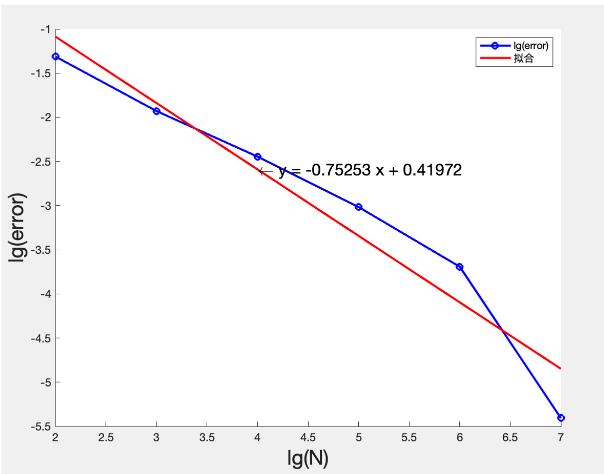

图 3: x 的 k 阶矩误差与 N 的关系

可以看出, 偏差随着 N 递减, 但拟合系数 $k=-0.75$, 与理论值 $k=-0.5$ 有一定偏差.

3.3 C(l) 二维独立性检验

可以用两个随机数的自相关函数来标识伪随机数序列的独立性情况, 相关系数越小, 独立性越好. 间距为 l 的自相关函数是:

$$C(l) = \frac{< x_n x_{n+l} > - < x_n >< x_{n+l} >}{< x_n^2 > - < x_n >^2} \tag{5}$$

当两个随机序列 x_n, x_{n+1} 不相关时,$< x_n x_{n+1} >=< x_n >< x_{n+1} >=< x_n >^2$, 根据理论推算,$C(l) = 0$.

取 $N = 20000000$, 计算不同 l 下的 $C(l)$:

```
C(2) = 0.00039136
C(3) = -0.00017696
C(4) = -1.2955e-05
C(5) = -3.0445e-05
C(6) = -0.0001123
C(7) = 0.0003609
C(8) = -7.068e-05
C(9) = -2.2816e-05
C(10) = 0.00014723
```

图 4: 不同 l 下,C(l) 计算

根据计算, 对 l 取 2 到 10,$C(l)$ 的数量级均在 10^{-3} 以下, 说明随机数序列具有较好的二维独立性.

3.4 卡方分布

将区间 $[0, 1]$ 分为 K 个小区间, 统计随机数落在第 k 个子区间的实际频数 n_k, 它应当趋近于理论频数 $m_k = N/K$

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_k - m_k)^2}{m_k} \tag{6}$$

通常求和中的每一项的大小约为 1, 因此 χ^2 的值约为 K. 概率论中的 Pearson 定理说明,(6) 式的极限概率分布是 χ^2 分布, 即:

$$P(\chi^2 \leqslant x|\upsilon) = \frac{1}{2^{\upsilon/2}\Gamma(\upsilon/2)} \int_0^x t^{\upsilon/2-1} e^{-t/2}\, dt \tag{7}$$

其中 $\upsilon = K - 1$ 为自由度.

取 $N = 20000000$, 计算不同 K 下的 χ^2 值及其对应的 $P(\chi^2 \leqslant x|\upsilon)$, 如下图所示:

对于 K 值取 2 到 10, 在置信度 $1 - \alpha = 0.95$ 的条件下, 随机数满足均匀性, 通过卡方分布检验.

```
K = 2      chi_2 = 0.0098618     P = 0.079105
K = 3      chi_2 = 0.13032     P = 0.063081
K = 4      chi_2 = 3.0372     P = 0.61408
K = 5      chi_2 = 2.1269     P = 0.28756
K = 6      chi_2 = 1.4644     P = 0.082862
K = 7      chi_2 = 6.5766     P = 0.63822
K = 8      chi_2 = 8.5255     P = 0.71146
K = 9      chi_2 = 3.3215     P = 0.08741
K = 10      chi_2 = 8.6743     P = 0.53213
```

图 5: 不同 K 值下, 卡方分布计算

4 Summary

本实验使用 16807 产生器, 用 Schrage 方法实现了随机数生成, 并使用随机数的 k 阶矩, 计算 $C(l)$ 以及卡方分布, 对生成的随机数进行了均匀性和独立性的检验, 结果表明生成的随机数具有较好的均匀性和独立性, 可以满足一般的计算需求