Report1

> Rainzor

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。再用 x^k 测试均匀性（取不同量级的N值，讨论偏差与N的关系）、 $C(l)$ 测试其2维独立性（总点数$N > 10^7$）

2.Method

2.1 随机数的生成

Schrage 方法主要解决的是线性同余法产生随机数时可能出现的越界问题。其递推关系为:

$$I_{n+1} = \begin{cases} a(I_n\ mod\ q) - r[I_n/q], & if\ I_{n+1} > 0 \\ a(I_n\ mod\ q) - r[I_n/q] + m, & otherwise \end{cases}$$

实际计算中取q=127773， r=2836，a=16807，m=2147483647。当提供了第一个小于 m 的初始值 I_0 之后，就可以 根据次递推公式得到后续的伪随机数。

代码中用 `Class Schrage16807`封装实现

2.2 种子生成

采用计算机中的时间，来生成初始值

$$I_0 = (i_y - 2000) + 70(i_m + 12\{i_d + 31[i_h + 23(i_n + 59i_s)]\})$$

其值所在区间为 $[0, 2^{31} - 1]$

代码中用 `seed_time()`函数封装实现

3.Experiment

3.1 随机数平面分布图

采用相邻间隔随机数 (l_n, l_{n+l}) 生成二维平面,取5000个数,画出散点图.

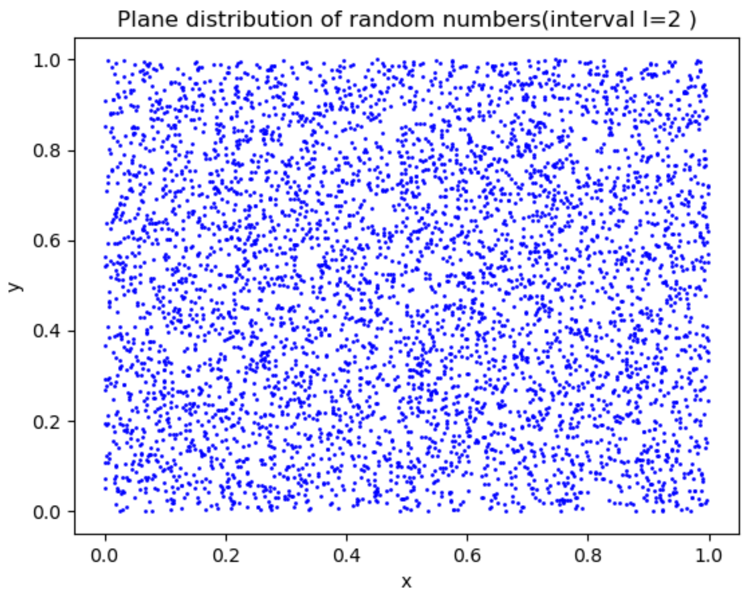

图1：随机数分布

图中取间隔为2,由图可见未有散点存在明显的分层分区.可见在此检验下,Schrage随机数生成器不具有明显相关性,性能较好.

3.2 检验随机数均匀性

本实验中,我采取了两种方式检验随机性:关于$< x^k >$k阶矩与期望值的比较以及卡方分布检验

3.2.1 k阶矩检验

随机数均匀性测试结果如图 2 所示。N表示随机数个数，k表示k阶矩，下图展现一部分结果

其期望值满足

$$E(x^k) = \int_0^1 x^k dx = \frac{1}{1+k}$$

```
N:1.00e+01, k:3, Average of xk:0.32124, Expectation:0.25000
N:1.00e+02, k:3, Average of xk:0.23816, Expectation:0.25000
N:1.00e+03, k:3, Average of xk:0.25646, Expectation:0.25000
N:1.00e+04, k:3, Average of xk:0.24877, Expectation:0.25000
N:1.00e+05, k:3, Average of xk:0.25162, Expectation:0.25000
N:1.00e+06, k:3, Average of xk:0.25004, Expectation:0.25000
N:1.00e+07, k:3, Average of xk:0.24995, Expectation:0.25000


N:1.00e+01, k:4, Average of xk:0.27613, Expectation:0.20000
N:1.00e+02, k:4, Average of xk:0.18875, Expectation:0.20000
N:1.00e+03, k:4, Average of xk:0.20577, Expectation:0.20000
N:1.00e+04, k:4, Average of xk:0.19888, Expectation:0.20000
N:1.00e+05, k:4, Average of xk:0.20143, Expectation:0.20000
N:1.00e+06, k:4, Average of xk:0.20001, Expectation:0.20000
N:1.00e+07, k:4, Average of xk:0.19996, Expectation:0.20000


N:1.00e+01, k:5, Average of xk:0.24654, Expectation:0.16667
N:1.00e+02, k:5, Average of xk:0.15604, Expectation:0.16667
N:1.00e+03, k:5, Average of xk:0.17170, Expectation:0.16667
N:1.00e+04, k:5, Average of xk:0.16571, Expectation:0.16667
N:1.00e+05, k:5, Average of xk:0.16796, Expectation:0.16667
N:1.00e+06, k:5, Average of xk:0.16666, Expectation:0.16667
N:1.00e+07, k:5, Average of xk:0.16664, Expectation:0.16667
```

图2：不同N，k下，实验值与理论值结果

总体上可见随着 N 阶 数的增加，各 k 阶矩的实验值越来越接近理论值。可见 Schrage 方法得到的随机数均匀性不错。

以k=5为例，如图3所示，样本矩和理想值的偏差大致满足满足 O（$1/\sqrt{N}$）的关系

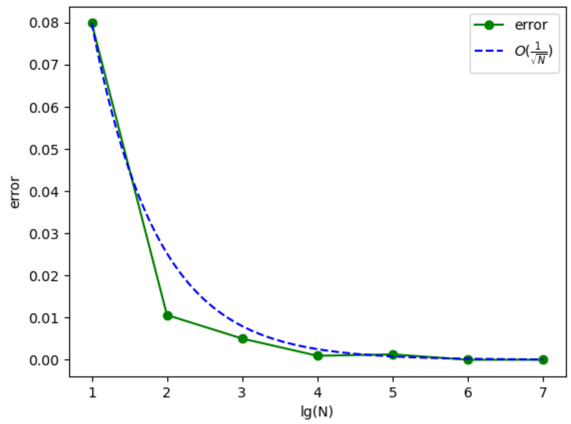

图3：5阶矩实验偏差与样本数N的关系

3.2.2 卡方检验

将区间[0,1] 分为 K 个子区间，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 $m_k = N/K, (k = 1, \ldots, K)$,令统计量为

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

χ^2服从自由度为K-1的卡方分布，在显著水平 $\alpha = 0.05$的条件下，对不同k，与N进行检验，得到图4结果（下图仅给出 $N \geq 10^5$的结果），N是随机数个数，k为子区间个数，statistics为检验统计量，Percent_point为分布点

```
N:1.00e+05, k:2, Statistics:4.84416, Percent_point:3.84146
N:1.00e+05, k:3, Statistics:4.34198, Percent_point:5.99146
N:1.00e+05, k:4, Statistics:5.32096, Percent_point:7.81473
N:1.00e+05, k:5, Statistics:3.81690, Percent_point:9.48773
N:1.00e+05, k:6, Statistics:6.72680, Percent_point:11.07050
N:1.00e+05, k:7, Statistics:6.08240, Percent_point:12.59159
N:1.00e+05, k:8, Statistics:6.51024, Percent_point:14.06714
N:1.00e+05, k:9, Statistics:12.94100, Percent_point:15.50731
N:1.00e+05, k:10, Statistics:8.36060, Percent_point:16.91898

N:1.00e+06, k:2, Statistics:0.64964, Percent_point:3.84146
N:1.00e+06, k:3, Statistics:0.95641, Percent_point:5.99146
N:1.00e+06, k:4, Statistics:2.00050, Percent_point:7.81473
N:1.00e+06, k:5, Statistics:0.53269, Percent_point:9.48773
N:1.00e+06, k:6, Statistics:2.24340, Percent_point:11.07050
N:1.00e+06, k:7, Statistics:0.62833, Percent_point:12.59159
N:1.00e+06, k:8, Statistics:3.77744, Percent_point:14.06714
N:1.00e+06, k:9, Statistics:3.50065, Percent_point:15.50731
N:1.00e+06, k:10, Statistics:2.20250, Percent_point:16.91898

N:1.00e+07, k:2, Statistics:2.33289, Percent_point:3.84146
N:1.00e+07, k:3, Statistics:0.77622, Percent_point:5.99146
N:1.00e+07, k:4, Statistics:3.44425, Percent_point:7.81473
N:1.00e+07, k:5, Statistics:2.31505, Percent_point:9.48773
N:1.00e+07, k:6, Statistics:7.46426, Percent_point:11.07050
N:1.00e+07, k:7, Statistics:7.65595, Percent_point:12.59159
N:1.00e+07, k:8, Statistics:8.97125, Percent_point:14.06714
N:1.00e+07, k:9, Statistics:9.63087, Percent_point:15.50731
N:1.00e+07, k:10, Statistics:5.19994, Percent_point:16.91898
```

图4：卡方分布检验

由图可见，在置信度 $1 - \alpha = 0.95$的条件下，满足随机数满足均匀性，性能很好

3.3 二维独立性检验

根据公式，得到检验统计量：

$$C(l)=\frac{<x_nx_{n+l}>-<x_n>^2}{<x^2>-<x_n>^2}$$

代码实现在 `covariance2` 函数中。

实验当中 N 取 2×10^7，在 l 处于 1 到 9 时，可见 C(l)接近于 0。即数据的二维独立性非常高，可见 Schrage 方法得到的 数据独立性不错。

```
检验关联性：

l = 1, C(l) = 0.0055892347
l = 2, C(l) = 0.0099037843
l = 3, C(l) = 0.0129640202
l = 4, C(l) = 0.0092159231
l = 5, C(l) = 0.0070320046
l = 6, C(l) = -0.0050442802
l = 7, C(l) = -0.0152519093
l = 8, C(l) = -0.0168483918
l = 9, C(l) = -0.0003001003
```

4 Summary

本实验掌握了随机数的产生方式和判断随机数产生器优劣的标准。

了解用 schrage 方法编写的 16807 产生器产生的随机数具有很弱的相关性以及很好的均匀性。