

计算物理 作业报告6

PB14203209 张静宁 2017.11.01

第六题

用 $\langle x^k \rangle$ 测试均匀性（取不同量级的 N 值，讨论偏差与 N 的关系）， $C(l)$ 测试其二维独立性（总点数 $N > 10^7$ ），与前面的 `randomz` 子程序进行比较（采用同样的常数以及单精度或双精度实数运算），总结和评价两个随机数产生器的随机性质量。

算法思路

生成随机数的独立性和均匀性，是判断随机数生成器质量的重要标准。针对生成的随机数序列 $\{x_n\}$ ，共 N 个随机数，可以用以下方法来检验伪随机数质量。

1. 用 $\langle x^k \rangle$ 测试均匀性

X 是 $[0, 1]$ 上均匀分布的随机变量，它的 k 阶原点矩理论值为 $\langle x^k \rangle = \int_0^1 x^k dx = \frac{1}{k+1}$ 。

计算伪随机序列的 k 阶矩 $\langle x^k \rangle = \sum_{n=1}^N x_n^k$ ，比较实际值和理论值的偏差。偏差越小，均匀性越好。

2. 用 $C(l)$ 测试独立性

讨论伪随机数序列独立性的一个方法是顺序相关法，用相邻两个随机数的相关系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。相距为 l 的相关系数为 $C(l) = \frac{\langle x_n x_{n+l} \rangle - \langle x_n \rangle^2}{\langle x_n^2 \rangle - \langle x_n \rangle^2}$ 。

其中平均值的定义是 $\langle x_n \rangle = \sum_{n=1}^N x_n / N$ 。当两个随机数序列 x_n 与 x_{n+l} 不相关时，相关系数为 0。

3. 用 χ^2 测试均匀性 - 频率性检验

将区间 $[0, 1]$ 分为 K 个子区间，统计随机数落在第 i 个子区间的实际频数 o_i ，它应该趋近于理论频数 $e_i = N/K$ 。

$$\langle \chi^2 \rangle = \sum_{i=1}^K \frac{(o_i - e_i)^2}{e_i}$$

如果 χ^2 值很大，表示实际值远远偏离理想值，因此要求 χ^2 尽可能小，通常求和中的每一项大小约为 1，因此 χ^2 的值约为 K 。若给定显著水平 α （或置信系数 $1 - \alpha$ ），查表可以求出，系统的自由度为 $\nu = K - 1$ ，此时 $P(\chi^2 < \chi_{\alpha}^2) = 1 - \alpha$ 的 χ_{α} 值。

程序说明

作业压缩包中含有 3 个 C 程序，和对应的 linux 下的可执行文件。

- kth-order-moment.c 计算 k 阶矩近似值和理论值的偏差，并输出结果
- correlation.c 计算距离为 l 的相关系数，并输出结果
- chi-square.c 频率测试均匀性，输出结果

结果分析

以下对比两个随机数生成器的质量：**16807** 随机数生成器和 `randomz` 子程序。这两个随机数生成器都采用线性同余法，随机数种子 *Seed* = 1。

不同的是用Schrage方法实现的16807随机数生成器：

$$a = 7^5 = 16807, b = 0, m = 2^{31} - 1 = 2147483647, q = 127773, r = 2836$$

而 randomz 子程序的参数为： $a = 329, b = 0, m = 100000001, q = 303951, r = 122$

1. 用 $\langle x^k \rangle$ 测试均匀性

以下计算了随机数个数 $N = 10^3, 10^4, 10^5, 10^6, 10^7$ 时，伪随机序列 k 阶矩 $\langle x^k \rangle = \sum_{n=1}^N x_n^k$ 实际值和理论值的偏差，这里 $k = \{1, 2, 3, 4, 5\}$ 。

16807 随机数生成器

$$N = 10^3$$

| k | kth moment | exact value | deviation |
|---|-------------------|--------------------|-------------------|
| 1 | 0.497961379353871 | 0.5000000000000000 | 0.002038620646129 |
| 2 | 0.326714332985556 | 0.3333333333333333 | 0.006619000347778 |
| 3 | 0.240649039774140 | 0.2500000000000000 | 0.009350960225860 |
| 4 | 0.189388583017933 | 0.2000000000000000 | 0.010611416982067 |
| 5 | 0.155662594227676 | 0.1666666666666667 | 0.011004072438990 |

$$N = 10^4$$

| k | kth moment | exact value | deviation |
|---|-------------------|--------------------|-------------------|
| 1 | 0.501826822208346 | 0.5000000000000000 | 0.001826822208346 |
| 2 | 0.335474181824202 | 0.3333333333333333 | 0.002140848490869 |
| 3 | 0.252227887119419 | 0.2500000000000000 | 0.002227887119419 |
| 4 | 0.202284192730505 | 0.2000000000000000 | 0.002284192730505 |
| 5 | 0.169010758675060 | 0.1666666666666667 | 0.002344092008393 |

$$N = 10^5$$

| k | kth moment | exact value | deviation |
|---|-------------------|--------------------|-------------------|
| 1 | 0.500284291040183 | 0.5000000000000000 | 0.000284291040183 |
| 2 | 0.333479083574409 | 0.3333333333333333 | 0.000145750241075 |
| 3 | 0.249986406269735 | 0.2500000000000000 | 0.000013593730265 |
| 4 | 0.199876901167479 | 0.2000000000000000 | 0.000123098832521 |
| 5 | 0.166480365252800 | 0.1666666666666667 | 0.000186301413866 |

$$N = 10^6$$

| k | kth moment | exact value | deviation |
|---|-------------------|--------------------|-------------------|
| 1 | 0.500030059810347 | 0.5000000000000000 | 0.000030059810347 |
| 2 | 0.333277476842541 | 0.3333333333333333 | 0.000055856490792 |
| 3 | 0.249887435959530 | 0.2500000000000000 | 0.000112564040470 |
| 4 | 0.199855340869154 | 0.2000000000000000 | 0.000144659130846 |
| 5 | 0.166506514811243 | 0.1666666666666667 | 0.000160151855423 |

$$N = 10^7$$

| k | kth moment | exact value | deviation |
|---|-------------------|--------------------|-------------------|
| 1 | 0.500018649529947 | 0.5000000000000000 | 0.000018649529947 |
| 2 | 0.333338565201802 | 0.3333333333333333 | 0.000005231868468 |
| 3 | 0.249999311568265 | 0.2500000000000000 | 0.000000688431735 |
| 4 | 0.199995011791548 | 0.2000000000000000 | 0.000004988208452 |
| 5 | 0.166657935933247 | 0.1666666666666667 | 0.000008730733420 |

randomz

$N = 10^3$

| k | kth moment | exact value | deviation |
|---|-------------------|--------------------|-------------------|
| 1 | 0.494168796918312 | 0.5000000000000000 | 0.005831203081688 |
| 2 | 0.327209749224882 | 0.3333333333333333 | 0.006123584108451 |
| 3 | 0.244058479602565 | 0.2500000000000000 | 0.005941520397435 |
| 4 | 0.194434481961693 | 0.2000000000000000 | 0.005565518038307 |
| 5 | 0.161564966711644 | 0.1666666666666667 | 0.005101699955022 |

$N = 10^4$

| k | kth moment | exact value | deviation |
|---|-------------------|--------------------|-------------------|
| 1 | 0.495326428114739 | 0.5000000000000000 | 0.004673571885261 |
| 2 | 0.328362662186440 | 0.3333333333333333 | 0.004970671146893 |
| 3 | 0.245384427836464 | 0.2500000000000000 | 0.004615572163536 |
| 4 | 0.195810631192975 | 0.2000000000000000 | 0.004189368807025 |
| 5 | 0.162868352006182 | 0.1666666666666667 | 0.003798314660485 |

$N = 10^5$

| k | kth moment | exact value | deviation |
|---|-------------------|--------------------|-------------------|
| 1 | 0.498766980902036 | 0.5000000000000000 | 0.001233019097964 |
| 2 | 0.332221741898782 | 0.3333333333333333 | 0.001111591434551 |
| 3 | 0.249083833772251 | 0.2500000000000000 | 0.000916166227749 |
| 4 | 0.199243379765067 | 0.2000000000000000 | 0.000756620234933 |
| 5 | 0.166029902961628 | 0.1666666666666667 | 0.000636763705039 |

$N = 10^6$

| k | kth moment | exact value | deviation |
|---|-------------------|--------------------|-------------------|
| 1 | 0.499548065658263 | 0.5000000000000000 | 0.000451934341737 |
| 2 | 0.332917404061433 | 0.3333333333333333 | 0.000415929271901 |
| 3 | 0.249677505682398 | 0.2500000000000000 | 0.000322494317602 |
| 4 | 0.199762070045909 | 0.2000000000000000 | 0.000237929954091 |
| 5 | 0.166494584040172 | 0.1666666666666667 | 0.000172082626494 |

$N = 10^7$

| k | kth moment | exact value | deviation |
|---|-------------------|--------------------|-------------------|
| 1 | 0.499829460155746 | 0.5000000000000000 | 0.000170539844254 |
| 2 | 0.333143689935933 | 0.3333333333333333 | 0.000189643397401 |
| 3 | 0.249813558420753 | 0.2500000000000000 | 0.000186441579247 |
| 4 | 0.199822656809137 | 0.2000000000000000 | 0.000177343190863 |
| 5 | 0.166500043746960 | 0.1666666666666667 | 0.000166622919707 |

可见随着随机数个数 N 的增加，两组随机数的 $\langle x^k \rangle$ 阶矩实验值和理论值偏差越来越小（比较相同阶的矩）。并且明显可以看出 **16807** 生成器 比 **randomz** 子程序均匀性更好一些。

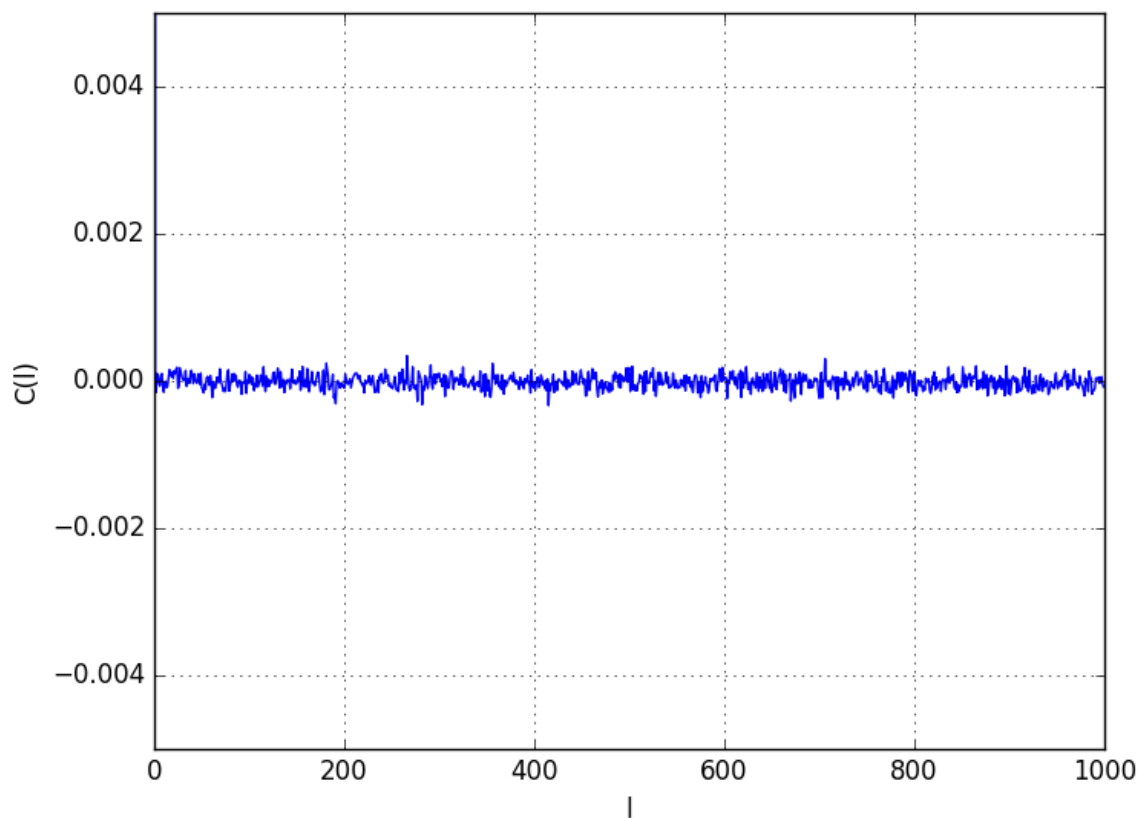
2. 用 $C(l)$ 测试独立性

以下测量随机数序列 $N = 10^8$ ， $Seed = 1$ ，相关系数为 $C(l) = \frac{\langle x_n x_{n+l} \rangle - \langle x_n \rangle^2}{\langle x_n^2 \rangle - \langle x_n \rangle^2}$ 。

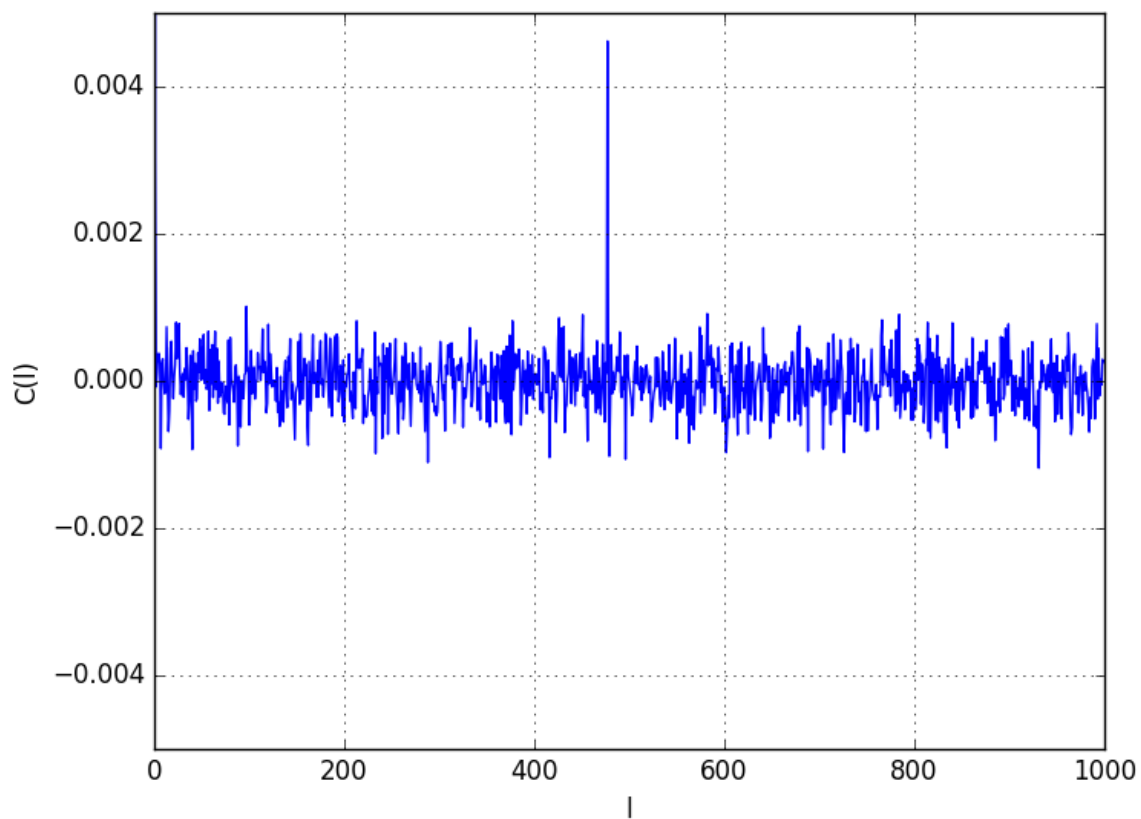
由以下表格和图片可以看出，这两组随机数 $C(l)$ 都非常小，故独立性都还不错。但 **16807** 随机数生成器 生成的随机数独立性更好一些！

| l | 16807 $C(l)$ | randomz $C(l)$ |
|-----|------------------------|-----------------------|
| 0 | 1.000000 | 1.000000 |
| 1 | 0.000048 | 0.003025 |
| 2 | -0.000151 | -0.000079 |
| 3 | 0.000103 | 0.000104 |
| 4 | 0.000071 | 0.000369 |
| 5 | -0.000037 | 0.000374 |
| 6 | -0.000067 | 0.000231 |
| 7 | 0.000074 | -0.000917 |
| 8 | 0.000101 | 0.000233 |
| 9 | -0.000065 | 0.000300 |

Correlation Coefficient of 16807(N=10⁸,Seed=1)



Correlation Coefficient of Randomz(N=10⁸,Seed=1)



3. 用 χ^2 测试均匀性 - 频率性检验

令 $D_i = \frac{(\alpha_i - e_i)^2}{e_i}$ ，测量随机数数列 $N = 3 * 10^7$, $K = 30$, $Seed = 1$.

16807随机数生成器

$\langle \chi^2 \rangle = 34.195560$

| Cell | Observed | Expected | Di |
|------|----------|-----------|-------|
| 0 | 999347 | 1000000.0 | 0.426 |
| 1 | 1000004 | 1000000.0 | 0.000 |
| 2 | 1001152 | 1000000.0 | 1.327 |
| 3 | 998556 | 1000000.0 | 2.085 |
| 4 | 1002066 | 1000000.0 | 4.268 |
| 5 | 1002188 | 1000000.0 | 4.787 |
| 6 | 999513 | 1000000.0 | 0.237 |
| 7 | 999208 | 1000000.0 | 0.627 |
| 8 | 998594 | 1000000.0 | 1.977 |
| 9 | 999594 | 1000000.0 | 0.165 |
| 10 | 1000060 | 1000000.0 | 0.004 |
| 11 | 1000244 | 1000000.0 | 0.060 |
| 12 | 999313 | 1000000.0 | 0.472 |
| 13 | 1000580 | 1000000.0 | 0.336 |
| 14 | 1001644 | 1000000.0 | 2.703 |
| 15 | 999581 | 1000000.0 | 0.176 |
| 16 | 999357 | 1000000.0 | 0.413 |
| 17 | 999648 | 1000000.0 | 0.124 |
| 18 | 998376 | 1000000.0 | 2.637 |
| 19 | 1000765 | 1000000.0 | 0.585 |
| 20 | 1001102 | 1000000.0 | 1.214 |
| 21 | 999050 | 1000000.0 | 0.902 |
| 22 | 1002044 | 1000000.0 | 4.178 |
| 23 | 999543 | 1000000.0 | 0.209 |
| 24 | 1000829 | 1000000.0 | 0.687 |
| 25 | 1000221 | 1000000.0 | 0.049 |
| 26 | 999935 | 1000000.0 | 0.004 |
| 27 | 998293 | 1000000.0 | 2.914 |
| 28 | 999985 | 1000000.0 | 0.000 |
| 29 | 999208 | 1000000.0 | 0.627 |

randomz 子程序

$\langle \chi^2 \rangle = 89.564122$

| Cell | Observed | Expected | Di |
|-------|----------|-----------|--------|
| 0 | 1002053 | 1000000.0 | 4.215 |
| 1 | 999830 | 1000000.0 | 0.029 |
| 2 | 1000938 | 1000000.0 | 0.880 |
| 3 | 1001217 | 1000000.0 | 1.481 |
| 4 | 1000499 | 1000000.0 | 0.249 |
| 5 | 1000538 | 1000000.0 | 0.289 |
| 6 | 1002664 | 1000000.0 | 7.097 |
| 7 | 998079 | 1000000.0 | 3.690 |
| 8 | 1001320 | 1000000.0 | 1.742 |
| 9 | 999684 | 1000000.0 | 0.100 |
| 10 | 1000741 | 1000000.0 | 0.549 |
| 11 | 999486 | 1000000.0 | 0.264 |
| 12 | 1001897 | 1000000.0 | 3.599 |
| 13 | 999791 | 1000000.0 | 0.044 |
| 14 | 998280 | 1000000.0 | 2.958 |
| 15 | 1001453 | 1000000.0 | 2.111 |
| 16 | 996590 | 1000000.0 | 11.628 |
| 17 | 1001774 | 1000000.0 | 3.147 |
| 18 | 1001344 | 1000000.0 | 1.806 |
| 19 | 996007 | 1000000.0 | 15.944 |
| 20 | 1001985 | 1000000.0 | 3.940 |
| 21 | 1001868 | 1000000.0 | 3.489 |
| 22 | 1000222 | 1000000.0 | 0.049 |
| 23 | 999344 | 1000000.0 | 0.430 |
| 24 | 998446 | 1000000.0 | 2.415 |
| 25 | 997899 | 1000000.0 | 4.414 |
| 26 | 1001261 | 1000000.0 | 1.590 |
| 27 | 997107 | 1000000.0 | 8.369 |
| 28 | 998418 | 1000000.0 | 2.503 |
| 29 | 999265 | 1000000.0 | |
| 0.540 | | | |

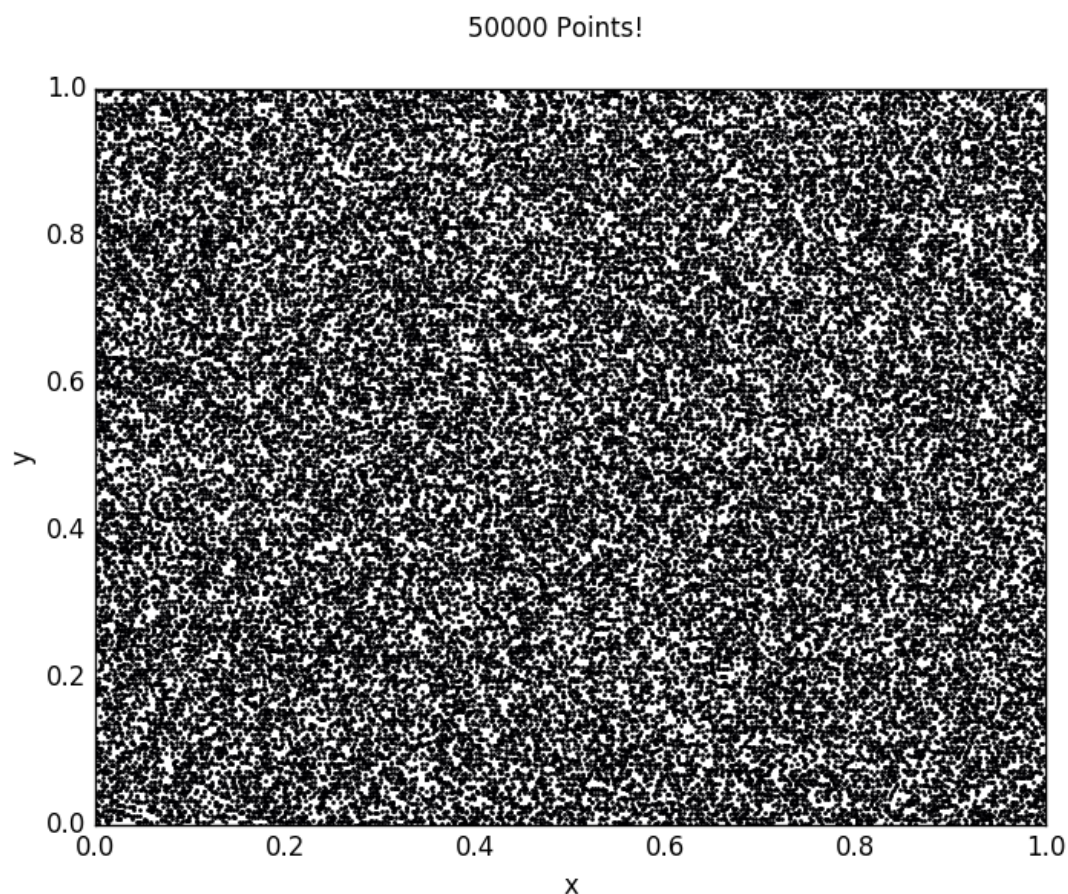
用卡方检验法，给定显著水平 $\alpha = 0.1$ （即置信系数0.9），系统的自由度为 $v = 29$ ，满足 $P(\chi^2 < \chi_{\alpha}^2) = 0.9$ 的 $\chi_{\alpha} = 39.087$ 。即 χ^2 有90%的概率小于39.087。

而上表得出： $< \chi_{16807}^2 > = 34.195560 < 39.087 < \chi_{randomz}^2 > = 89.564122 > 39.087$

故16807生成器满足该置信系数下的随机检验，而randomz子程序不满足。所以16807生成随机数的均匀性更好。

PS:用相邻随机数做图

randomz 子程序



总结评价

1. 可见随着随机数个数 N 的增加, 两组随机数的 $\langle x^k \rangle$ 阶矩实验值和理论值偏差越来越小 (比较相同阶的矩). 并且明显可以看出 **16807** 生成器 比 **randomz** 子程序均匀性更好一些.
2. 这两组随机数 $C(l)$ 都非常小, 故独立性都还不错. 但 **16807** 随机数生成器 生成的随机数独立性更好一些!
3. 用卡方检验法, 16807 生成器满足置信系数为 **0.9** 时的拟合优度检验, 而 randomz 子程序不满足. 所以 **16807** 生成随机数的均匀性更好.

综上, **16807** 随机数生成器, 生成的伪随机数均匀性和独立性, 都比 **randomz** 子程序的更好!