HOMEWORK 1: Exercises for Monte Carlo Methods

> Student ID:
>
> Student Name:
>
> 注：代码见code文件夹

Exercise 1.

The Monte Carlo method can be used to generate an approximate value of pi. The figure below shows a unit square with a quarter of a circle inscribed. The area of the square is 1 and the area of the quarter circle is pi/4. Write a script to generate random points that are distributed uniformly in the unit square. The ratio between the number of points that fall inside the circle (red points) and the total number of points thrown (red and green points) gives an approximation to the value of pi/4. This process is a Monte Carlo simulation approximating pi. Let N be the total number of points thrown. When N=50, 100, 200, 300, 500, 1000, 5000, what are the estimated pi values, respectively? For each N, repeat the throwing process 100 times, and report the mean and variance. Record the means and the corresponding variances in a table.

蒙特卡洛方法可以用于产生接近pi的近似值。图1显示了一个带有1/4内切圆在内的边长为1的正方形。正方形的面积是1，该1/4圆的面积为pi/4。通过编程实现在这个正方形中产生均匀分布的点。落在圈内（红点）的点和总的投在正方形（红和绿点）上的点的比率给出了pi/4的近似值。这一过程称为使用蒙特卡洛方法来仿真逼近pi实际值。令N表示总的投在正方形的点。当投点个数分别是20, 50, 100, 200, 300, 500, 1000, 5000时，pi值分别是多少？对于每个N，每次实验算出pi值，重复这个过程100次，并在表中记下均值和方差。

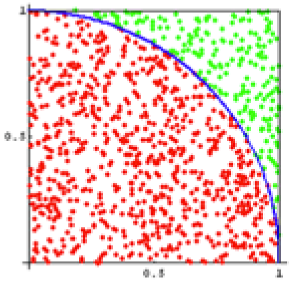

Figure 1 蒙特卡洛方法求解pi

解:

结果:

投点个数	均值	方差
20	3.0900000000000007	0.16590000000000005
50	3.1679999999999997	0.06156799999999984
100	3.1556	0.02962864
200	3.128400000000001	0.011289439999999998
300	3.141466666666666	0.008678737777777775
500	3.15024	0.003902182400000007
1000	3.139120000000001	0.002780025600000004
5000	3.1375840000000004	0.0004927485439999994

Exercise 2.

We are now trying to integrate the another function by Monte Carlo method:

$$\int_0^1 x^3$$

A simple analytic solution exists here: . If you compute this integration using Monte Carlo method, what distribution do you use to sample x? How good do you get when N = 5, 10, 20, 30, 40, 50, 60, 70, 80, 100, respectively? For each N, repeat the Monte Carlo process 100 times, and report the mean and variance of the integrate in a table.

我们现在尝试通过蒙特卡洛的方法求解如下的积分：

$$\int_0^1 x^3$$

该积分的求解我们可以直接求解，即有。如果你用蒙特卡洛的方法求解该积分，你认为x可以通过什么分布采样获得？如果采样次数是分别是N = 5, 10, 20, 30, 40, 50, 60, 70, 80, 100，积分结果有多好？对于每个采样次数N，重复蒙特卡洛过程100次，求出均值和方差，然后在表格中记录对应的均值和方差。

解：

1. x 可以通过均匀分布采样获得
2. 随着采样次数增加，积分结果会越来越接近真实值
3. 结果：

```
+----------+--------------------+----------------------+
| 投点个数 |        均值        |         方差         |
+----------+--------------------+----------------------+
|    5     |       0.244        | 0.043663999999999994 |
|    10    | 0.25199999999999995 |       0.015296       |
|    20    |       0.233        | 0.008310999999999999 |
|    30    |  0.2333333333333333 | 0.005488888888888888 |
|    40    | 0.24650000000000002 | 0.003837749999999999 |
|    50    | 0.25220000000000004 | 0.002991160000000005 |
|    60    |  0.2358333333333333 | 0.003129861111111106 |
|    70    |  0.2502857142857142 | 0.002401959183673469 |
|    80    | 0.25037499999999996 |     0.002388921875     |
```

```
|   100    |        0.248        | 0.0014740000000000003 |
+----------+---------------------+-----------------------+
```

Exercise 3.

We are now trying to integrate a more difficult function by Monte Carlo method that may not be analytically computed:

$$\int_{x=2}^{4}\int_{y=-1}^{1} f(x,y) = \frac{y^2 * e^{-y^2} + x^4 * e^{-x^2}}{x * e^{-x^2}}$$

Can you compute the above integration analytically? If you compute this integration using Monte Carlo method, what distribution do you use to sample (x,y)? How good do you get when the sample sizes are N = 5, 10, 20, 30, 40, 50, 60, 70, 80, 100, 200 respectively? For each N, repeat the Monte Carlo process 100 times, and report the mean and variance of the integrate.

我们现在尝试通过蒙特卡洛的方法求解如下的更复杂的积分：

$$\int_{x=2}^{4}\int_{y=-1}^{1} f(x,y) = \frac{y^2 * e^{-y^2} + x^4 * e^{-x^2}}{x * e^{-x^2}}$$

你能够通过公式直接求解上述的积分吗？如果你用蒙特卡洛的方法求解该积分，你认为(x, y)可以通过什么分布采样获得？如果点（x, y）的采样次数是分别是N = 10, 20, 30, 40, 50, 60, 70, 80, 100, 200, 500, 积分结果有多好？对于每个采样次数N，重复蒙特卡洛过程100次，求出均值和方差，然后在表格中记录对应的均值和方差。

解：

1. 难以获得原函数
2. x 可以通过均匀分布采样获得
3. 随着采样次数增加，积分结果会越来越接近真实值
4. 结果：

   采样公式： $F_n(x) = \frac{1}{n}\sum_{k=1}^{n}\frac{f(x_k)}{pdf(x_k)}$

```
+----------+--------------------+--------------------+
|  投点个数 |         均值         |         方差         |
+----------+--------------------+--------------------+
|    10    | 118026.86114027457 | 14292540974.595875 |
|    20    | 113608.0894653929  | 5062368079.943775  |
|    30    | 120618.94362821664 | 4637832474.967341  |
|    40    | 116021.7907955739  | 3620170346.163431  |
|    50    | 111198.40400887006 | 1680965189.0731325 |
|    60    | 114657.35018627212 | 2691426876.8992248 |
|    70    | 115411.67969029001 | 1845303738.8642554 |
|    80    | 114036.01428840049 | 1557689036.855965  |
|   100    | 116003.63165405788 | 1247610808.6733193 |
```

```
|    200    | 112088.72559289538 | 728424193.9103771   |
|    500    | 111060.30217425052 | 189390415.88236827 |
+----------+-------------------+--------------------+
```

Exercise 4.

An ant is trying to get from point A to point B in a grid. The coordinates of point A is (1,1) (this is top left corner), and the coordinates of point B is (n,n) (this is bottom right corner, n is the size of the grid).

Once the ant starts moving, there are four options, it can go left, right, up or down (no diagonal movement allowed). If any of these four options satisfy the following:

(a) The new point should still be within the boundaries of the n×n grid

(b) Only the center point (4, 4) is allowed to be visited zero, one or two times, while the remainder points should not be visited previously (are allowed to be visited zero or one time).

If P is the probability of the ant reaching point B for a 7×7 grid, use Monte Carlo simulation to compute P. Pick the answer closest to P in value (assume 20,000 simulations are sufficient enough to compute P).

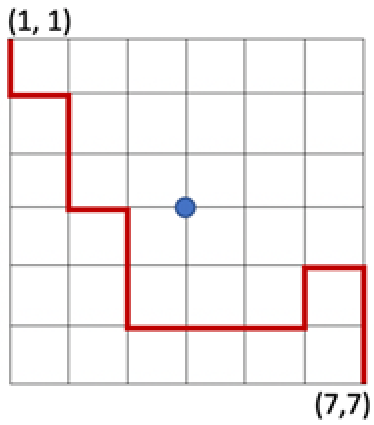

Figure 2 An ant is trying to get from point A (1,1) to point B (7,7) in a grid.

不能越界。走到不能再走了，或者到达 （7，7）这个点就结束了

除了中心点可以重复到过两次，其他点不能重复到即最多只能到过一次

第一次作业的第4题中pick the answer closest to P是什么意思，好像没有给可选项，只需要输出模拟20000次后的P值就行

到边界外那个动作是不允许，所有此刻只有三种可能的采样值

解：

输出模拟20000次后，运行结果：P:0.2552

Exercise 5.

Given a system made of discrete components with known reliability, what is the reliability of the overall system? For example, suppose we have a system that can be described with the following high-level diagram:

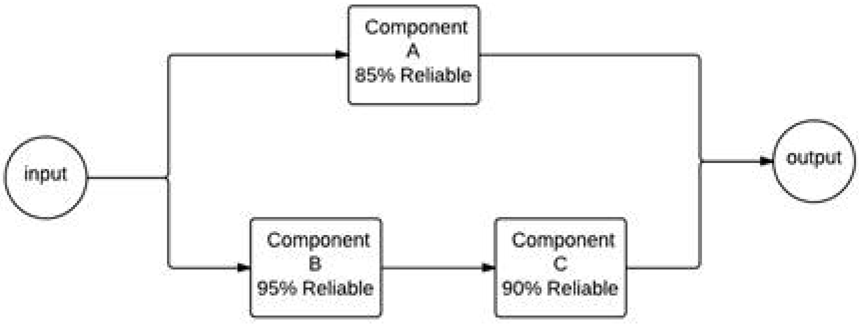

Figure 3 A system made of discrete components.

When given an input to the system, that input flows through component A or through components B and C, each of which has a certain reliability of correctness. Probability theory tells us the following:

$$reliability_{BC} = 0.95 * 0.90 = 0.855$$

$$reliability_{A} = 0.85$$

And the overall reliability of the system is:

$$\begin{aligned} reliability_{sys} &= 1.0 - [(1.0 - 0.85) * (1.0 - 0.855)] \\ &= 0.97825 \end{aligned}$$

Create a simulation of this system where half the time the input travels through component A. To simulate its reliability, generate a number between 0 and 1. If the number is 0.85 or below, component A succeeded, and the system works. The other half of the time, the input would travel on the lower half of the diagram. To simulate this, you will generate two numbers between 0 and 1. If the number for component B is less than 0.95 and the number for component C is less than 0.90, then the system also succeeds. Run many trials to see if you converge on the same reliability as predicted by probability theory.

解:

测试1000次，每次测试时产生3个数A、B、C，分别记录上下两部分的成功次数和最终的成功次数，最后将成功次数除以测试总次数即可获得概率。
最终结果如下，可以看出测试算得的概率与理论概率较为接近

```
上半部分成功次数：853;        下半部分成功次数：857;        总成功次数：979
上半部分成功概率：0.8530;     下半部分成功概率：0.8570;     总成功概率：0.9790
```