MATH 512 - Project 2

Wasif Ahmed, Haoxiang Deng, Jacob Fein-Ashley, Kanav Malhotra

Question 1 (a)

We use the *Kolmogorov-Smirnov* test to test for the uniformity of the random numbers generated by the LCG. The test statistic is given by

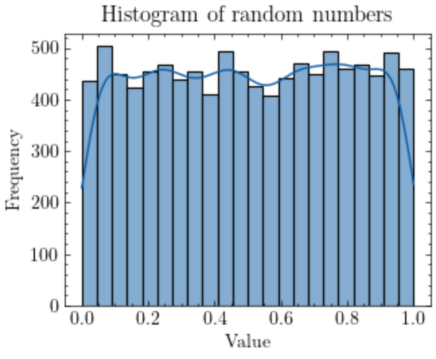
$$D_n = \max_{1\leqslant i\leqslant n}\left(\frac{i}{n} - U_{(i)}\right) \vee \max_{1\leqslant i\leqslant n}\left(U_{(i)} - \frac{i-1}{n}\right)$$

where $U_{(i)}$ is the i-th order statistic of the U_i's.

- $H_\circ$: the random numbers are uniformly distributed.

We find that $D_n = 0.0069$ and a p-value of 0.708. This means that we fail to reject $H_\circ$ at the 5% significance level and conclude that the random numbers are uniformly distributed.

Question 1 (a)


Histogram of random numbers
Frequency
Value
0
100
200
300
400
500
0.0
0.2
0.4
0.6
0.8
1.0

Question 1 (b)

Parameters: $a = 6, m = 11, x_0 = 3, c = 1$ and $n = 10$

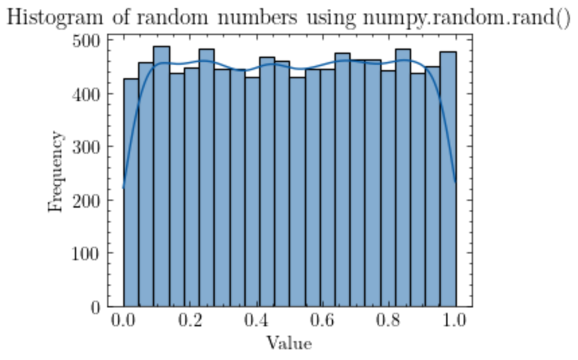
- The sequence is $\{3, 7, 9, 10, 5, 8, 4, 2, 1, 6\}$
- The period is 1

Parameters: $a = 6, m = 10, x_0 = 3, c = 1$ and $n = 10$

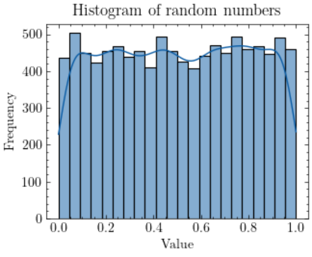
- The sequence is $\{3, 8, 8, 8, 8, 8, 8, 8, 8, 8\}$
- The period is 2

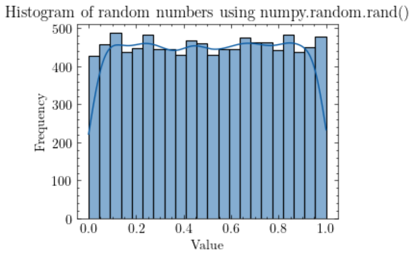
We notice that even a *small* change in the parameters results in a seemingly non-random sample.

Question 1 (c)


Histogram of random numbers using numpy.random.rand()
Frequency
0
100
200
300
400
500
0.0
0.2
0.4
0.6
0.8
1.0
Value

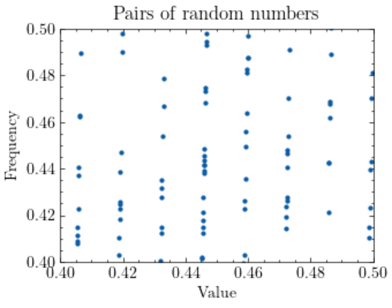
Question 1 (d)





- The two histograms look relatively similar, meaning both look relatively uniform.

Question 1 (e)



- The pairs of numbers seem to have a random pattern.

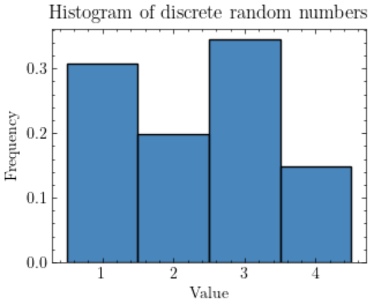
Question 1 (f)

Disadvantages of LCG:

- It can appear random with the right set of parameters, but as we saw, it can get "stuck" in a loop.
- The randomness depends on the choice of parameters.

Question 2

$$P(X = k) = \begin{cases} 0.3 & \text{for } k = 1 \\ 0.2 & \text{for } k = 2 \\ 0.35 & \text{for } k = 3 \\ 0.15 & \text{for } k = 4 \end{cases}$$

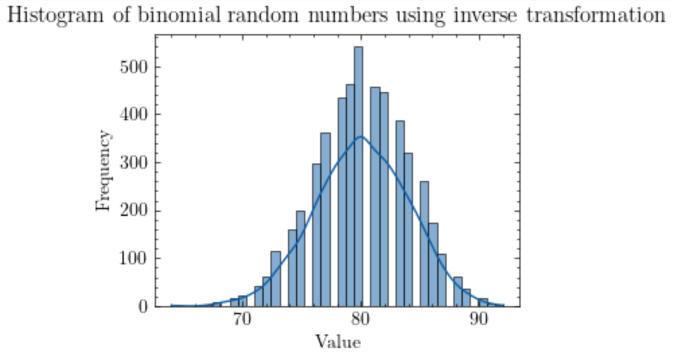


Question 3(a)

- Time for generation: 0.0349 seconds using NumPy's built-in function.
- Probability that $X < 50$: 0.00000000.

$$P(X < 50) = \sum_{k=0}^{49} \binom{100}{k} (0.8)^k (0.2)^{100-k} = 0.00000000$$

Question 3(b)

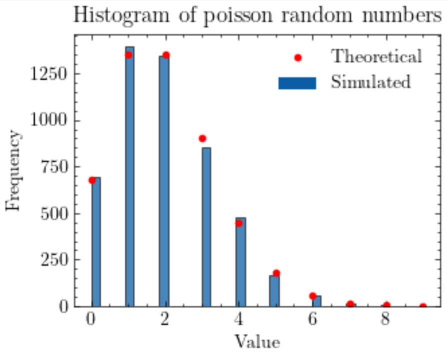


- Time for generation: 0.0350 seconds using the the inverse transform method.

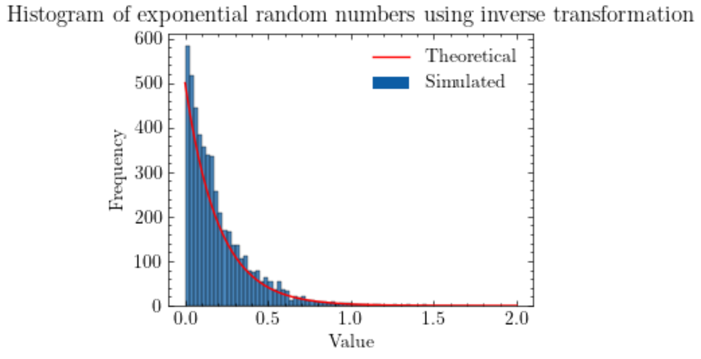
Question 3(c)

The histograms and the times for generation for the inverse method and NumPy's random number generator are very similar. We are comparing 0.0349 seconds to 0.0350 seconds, which is a very small difference.

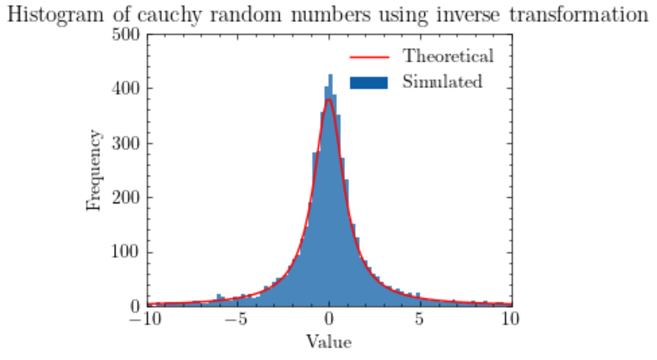
Question 4


Histogram of poisson random numbers
Theoretical
Simulated
Frequency
0
250
500
750
1000
1250
0
2
4
6
8
Value

Question 5


Histogram of exponential random numbers using inverse transformation
Theoretical
Simulated
Frequency
Value
0
100
200
300
400
500
600
0.0
0.5
1.0
1.5
2.0

Question 6


Histogram of cauchy random numbers using inverse transformation
Theoretical
Simulated
Frequency
Value
500
400
300
200
100
0
−10
−5
0
5
10

With a repetition of 10000 simulations, we find that the expectation is 1.00 and the variance is 1.01.

We can calculate $\mathbb{E}[X]$ and $\mathbb{V}[X]$ exactly and compare them with our estimates. We have

$$\mathbb{E}[X] = \sum_{i=1}^{100} i \cdot \frac{1}{100} = 1$$

$$\mathbb{V}[X] = \sum_{i=1}^{100} (i-1)^2 \cdot \frac{1}{100} - 1 = 1$$

The estimates are very close to the exact values. This is expected because the number of simulations is large.

1. Model the number of rolls needed as a geometric r.v.

$$X \sim \text{Geometric}(p)$$

   where p is the probability that all the possible outcomes have occurred at least once.

2. Calculate the probability

$$p = \frac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{6^6} = \frac{720}{46656} \approx 0.0154$$

3. Calculate the expected value of X as

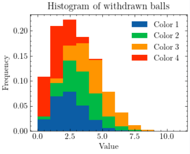
$$\mathbb{E}[X] = \frac{1}{p} = \frac{46656}{720} \approx 64.8$$

   The simulation gives us an estimate of $\mathbb{E}[X] \approx 64.8$.

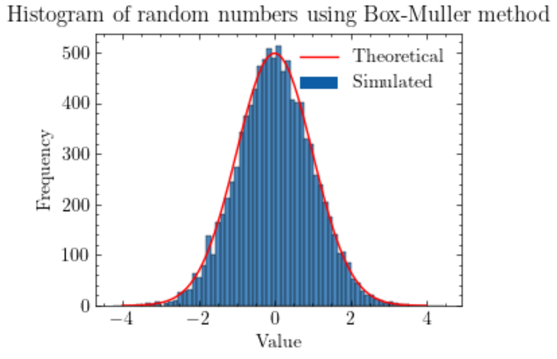
4. Then, the variance of X is

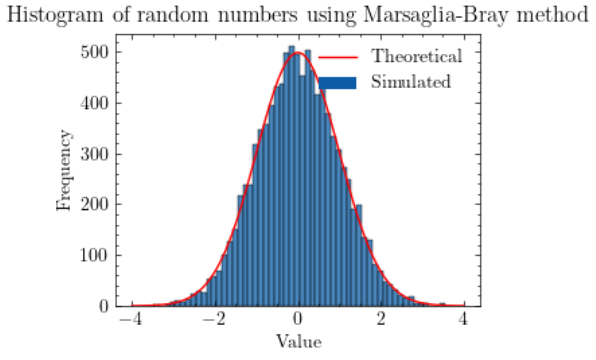
$$\mathbb{V}[X] = \frac{1-p}{p^2} = \frac{1-0.0154}{0.0154^2} \approx 4151.63$$

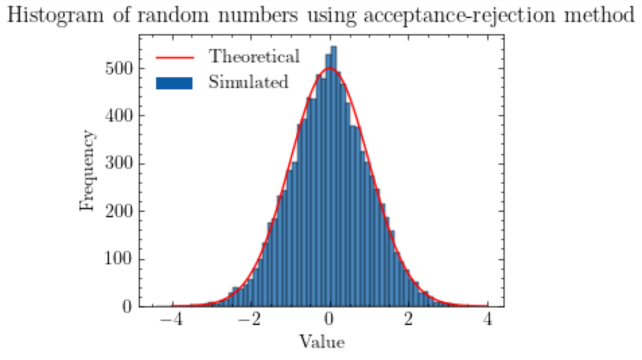
Question 9

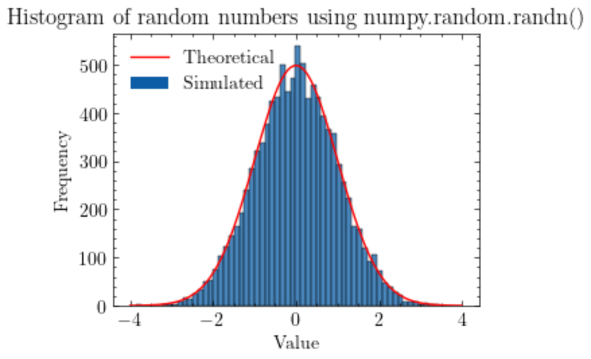

Histogram of withdrawn balls
Color 1
Color 2
Color 3
Color 4
Frequency
0.00
0.05
0.10
0.15
0.20
0.0
2.5
5.0
7.5
10.0
Value

Question 10


Histogram of random numbers using Box-Muller method
Theoretical
Simulated
Frequency
500
400
300
200
100
0
−4
−2
0
2
4
Value


Histogram of random numbers using Marsaglia-Bray method
Theoretical
Simulated
Frequency
Value
500
400
300
200
100
0
−4
−2
0
2
4


Histogram of random numbers using acceptance-rejection method
Theoretical
Simulated
Frequency
Value
0
100
200
300
400
500
−4
−2
0
2
4



All of the methods are very close to the theoretical normal distribution.

Questions?