Random Number Generator

Mahyar Albalan

403201223

1. Uniform disribution Generator

This question is fairly simple. First, I wrote a loop and plotted the results for different Ns. The plots are as follows:

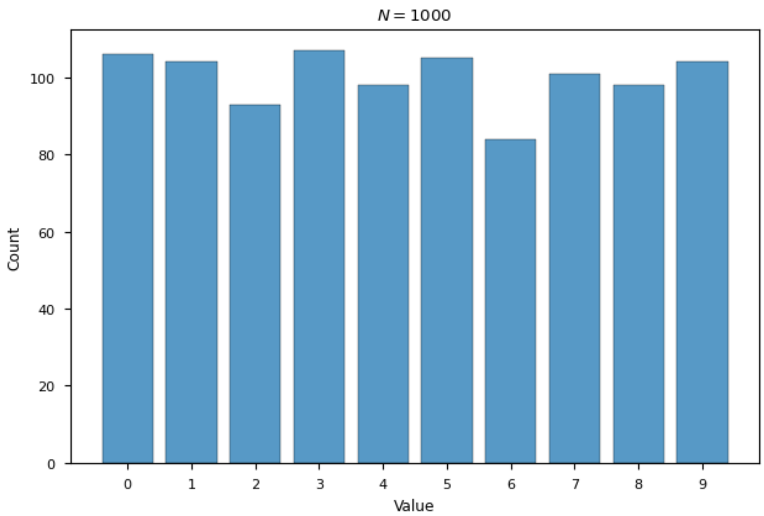

Figure 1: frequency plot for N = 1000

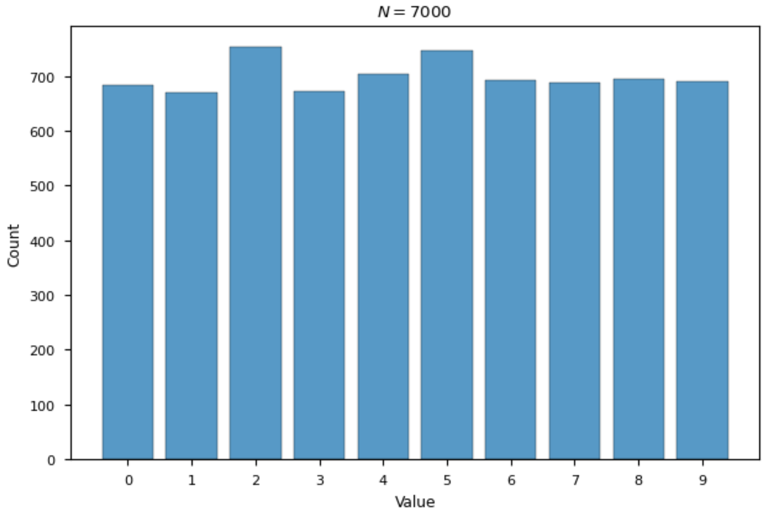

Figure 2: frequency plot for N = 7000

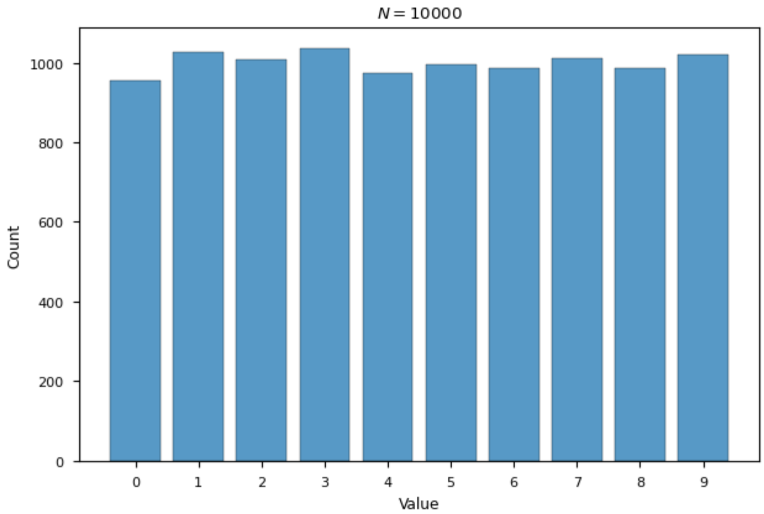

Figure 3: frequency plot for N = 10000

As expected, when N increases, the difference between each column will drop. I expect this distribution to be uniform. In statistical terms, the distribution of the population is uniform, but here we choose a subset of the population randomly (sample), and due to random chance, the sample distribution will be slightly different from the population. However, when the sample size increases (higher N), we expect this random error to decrease.

To check this, I plot σ/N versus $1/N^{1/2}$ (we expect their relation to be linear). (First, I plot σ versus N).

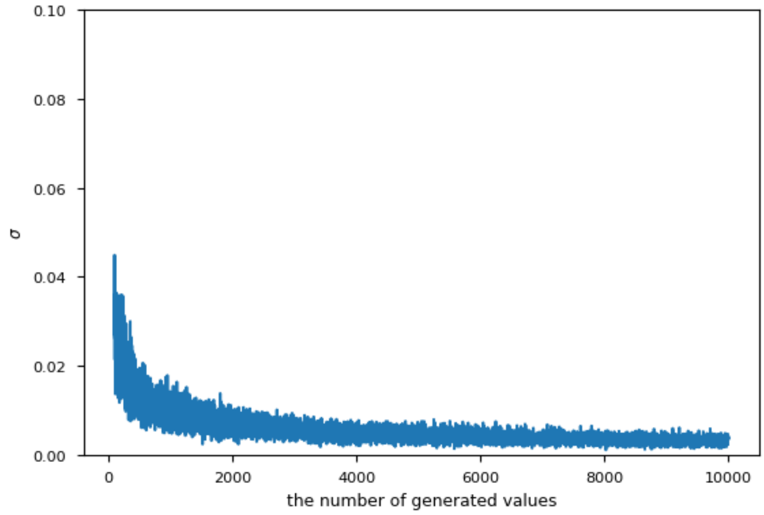

Figure 4: σ versus N

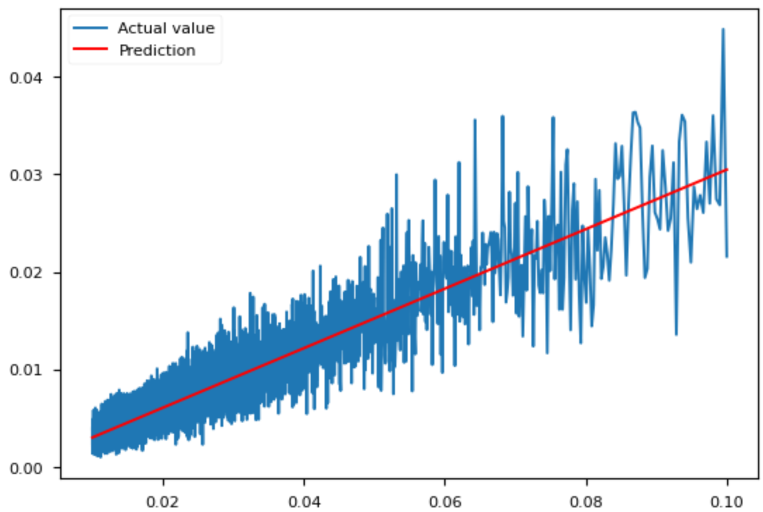

Figure 5: σ/N versus $1/N^{1/2}$

We can see that there is a linear relation, so the distribution is indeed uniform. To confirm this, I use the **Chi-Square** statistical test to check whether the distribution is uniform or not. The P-value is 0.92, which confirms that the distribution is uniform.
This problem is basically the random Ballistic Deposition. In that problem, we just have to create a random number between 0 and 200 with the same probability.

2.checking correlation

This problem is the same as the previous question. Here, I used Boolean indexing (mask) to take the target values. Everything is the same as before. The frequency plot and σ/N versus $1/N^{1/2}$ are as follows:

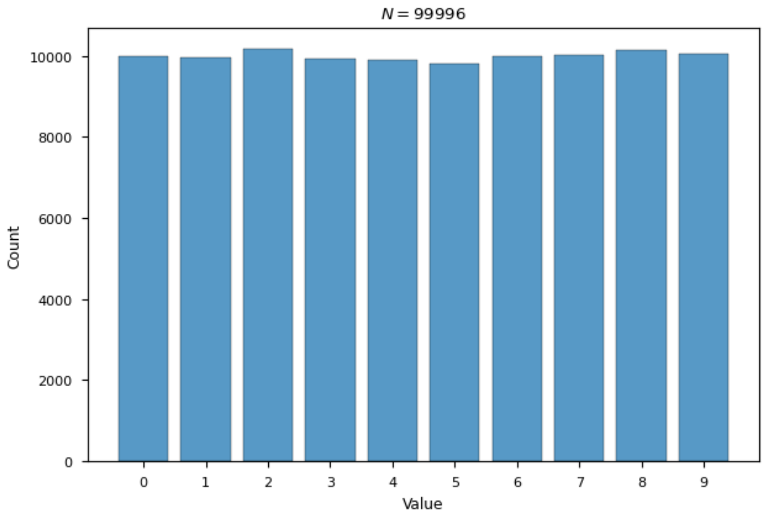

Figure 6: frequency plot for $N \approx 100000$

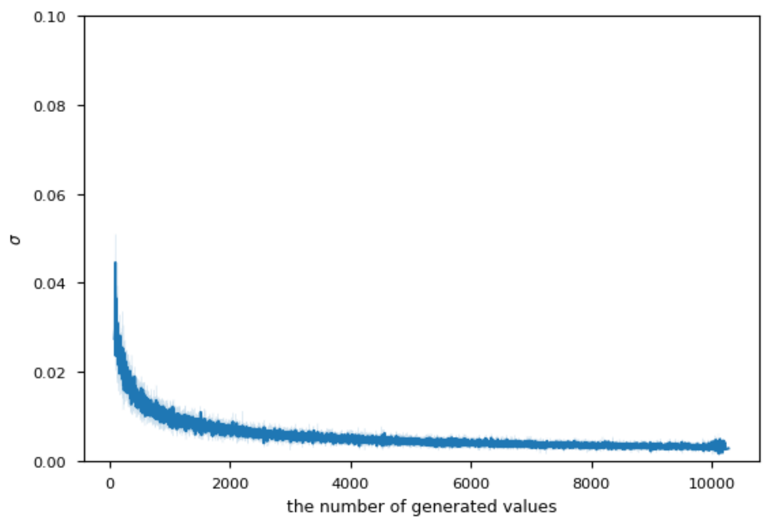

Figure 7: σ versus N

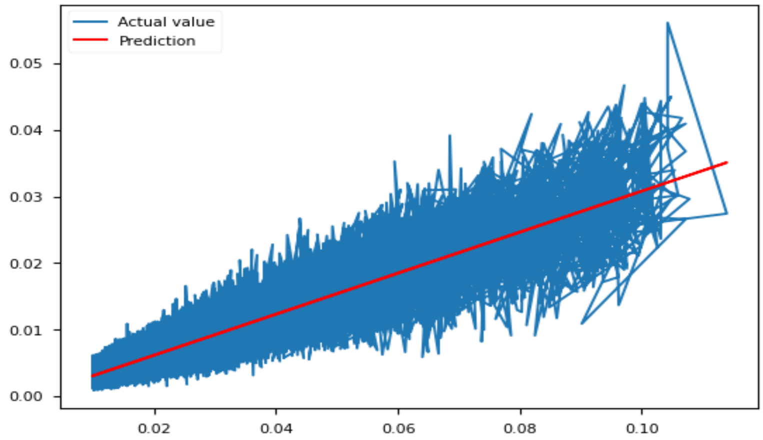

Figure 8: σ/N versus $1/N^{1/2}$

The plots are the same as in the previous question. There is no correlation, as expected.(You can check the MSE for this question and the previous question and see that they are in the same range,same for Chi-square test result.)

3. Uniform disribution Generator

I follow the LGC equation for creating random numbers (check the code for more details). The rest is the same as before. First, I will plot the frequency table, then I will plot σ/N versus $1/N^{1/2}$ (we expect their relation to be linear). (First, I plot σ versus N).

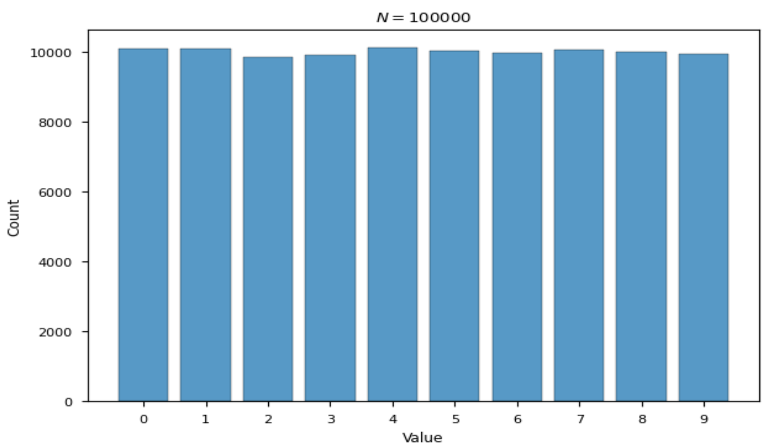

Figure 9: frequency plot for $N = 100000$

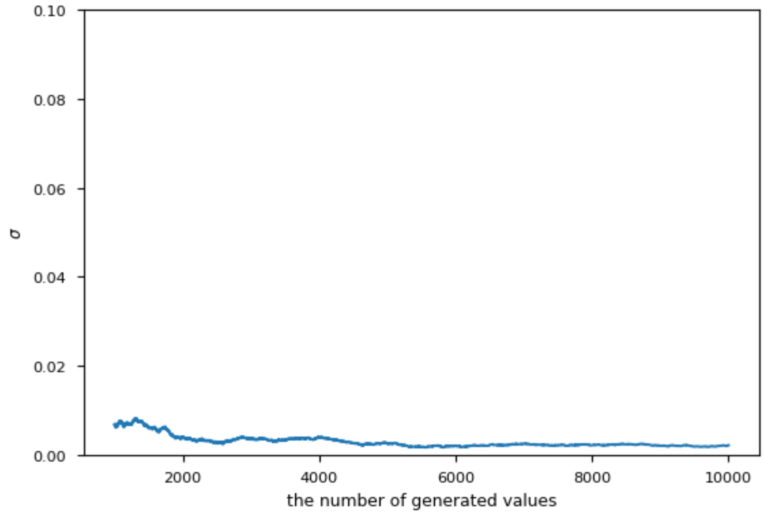

Figure 10: σ versus N

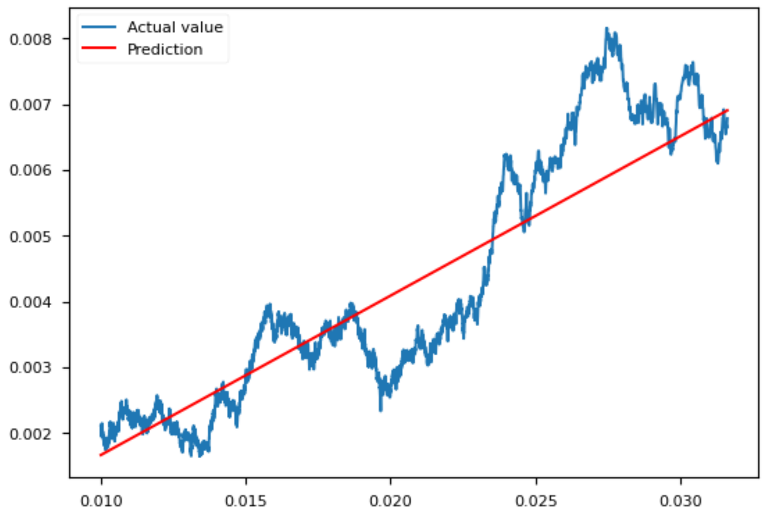

Figure 11: σ/N versus $1/N^{1/2}$

Again, we can see the linear relation as before, and from the plots, it is obvious that the distribution is normal. However, it is different from the previous plots, which is not surprising because, as mentioned in the text, there is probably another algorithm for creating random numbers in the *np.random.randint* function. Now, let's check the correlation. The plots are as follows:

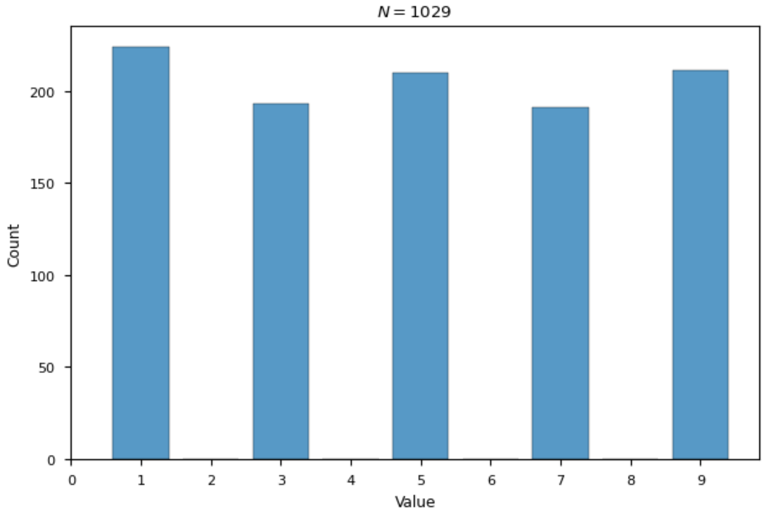

Figure 12: frequency plot for $N \approx 1000$

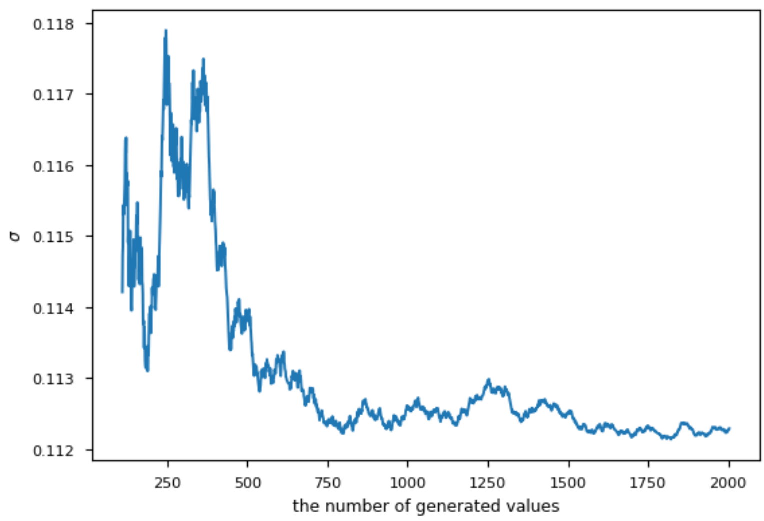

Figure 13: σ versus N

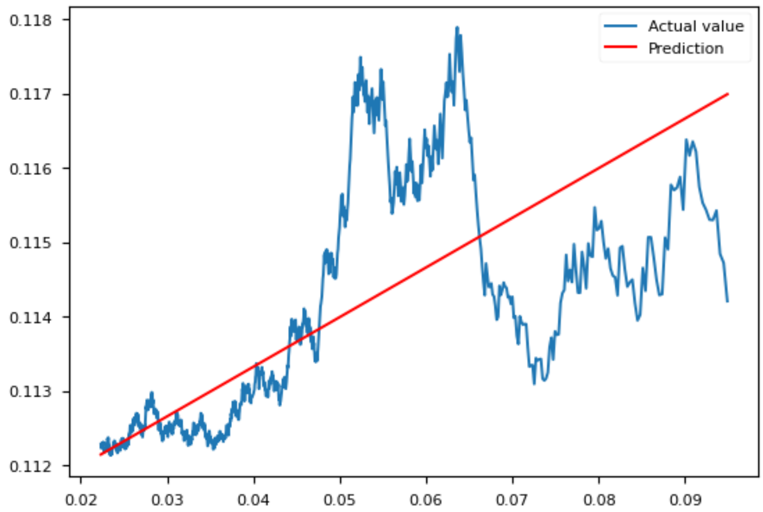

Figure 14: σ/N versus $1/N^{1/2}$

As expected (because here we have an equation to calculate these numbers), there is a correlation, and every value before 4s is odd, as mentioned in the text.

4.Centeral Limit Theorem

This problem is pretty straightforward. First, I create a set of N numbers, then choose k numbers from this set each time and calculate the sum of these selected numbers. The results for different values of k are as follows:

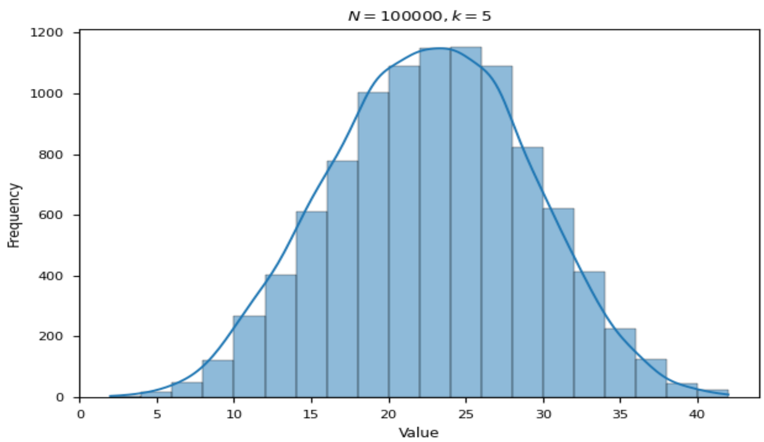

Figure 15: frequency plot for $k = 5$

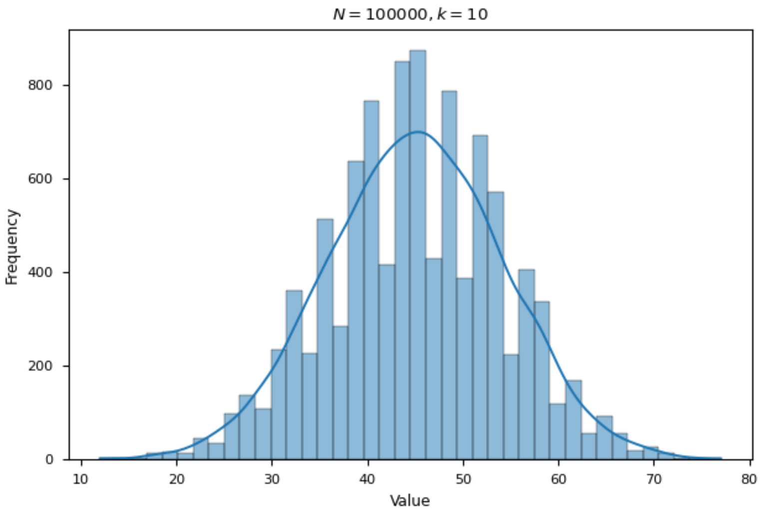

Figure 16: frequency plot for $k = 10$

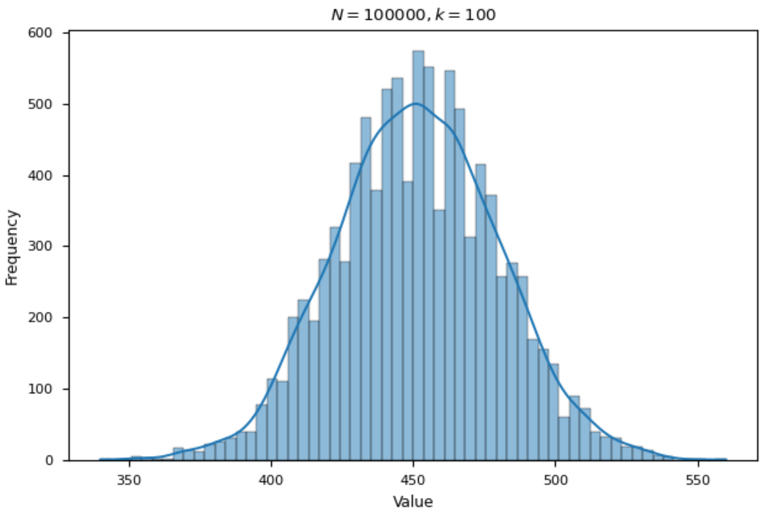

Figure 17: frequency plot for $k = 100$

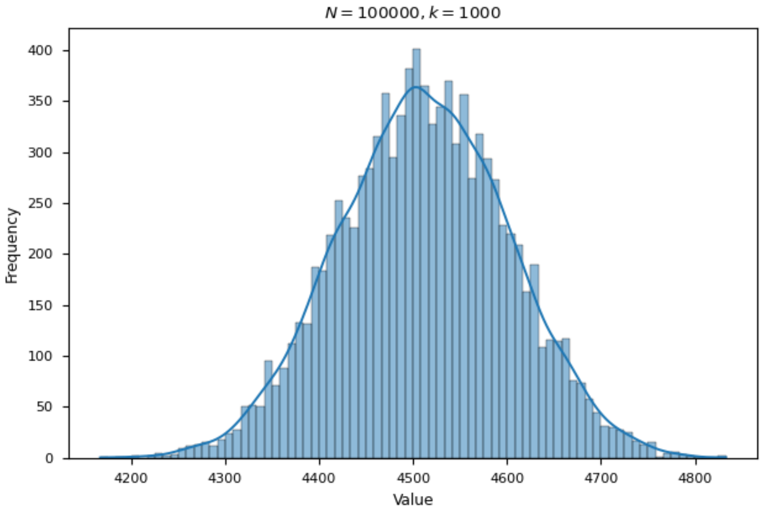

Figure 18: frequency plot for $k = 1000$

As expected, the distributions are normal. In the random walk problem, the distance from the starting point was calculated via a similar approach by adding random steps. As a result, the distance from the starting point follows a normal distribution. For random deposition, the total height of each column was the sum of several random events, so the heights followed a normal distribution. The growth rate of the standard deviation of heights, denoted as β, is 0.5, which is expected for a normal distribution.

5.Centeral Limit Theorem

I follow the method described in the text. By using two different uniform random number generators between 0 and 1, I generate ρ and θ with the given distributions. The resulting expressions for these two variables are:

$$\rho = \sqrt{-2\sigma^2 \ln(x)}$$

$$\theta = 2\pi x'$$

where x and x' are numbers generated from two independent uniform distributions. Now, let's check the plot for our Gaussian distribution generator.

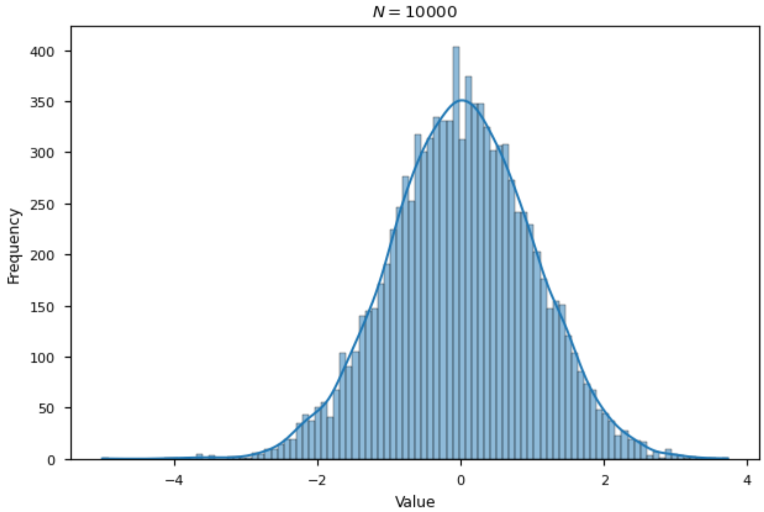

Figure 19: frequency plot for $k = 1000$

As you can see, the data follows a Gaussian distribution. To confirm this, I fitted a **GaussianMixture** model to the data, and the resulting distribution mean was 0.017, which is very close to the theoretical mean of 0.
Additionally, I set $\sigma = 1$ when generating this distribution, and the model's estimated standard deviation was 0.99, which is again very close to the expected value. (I only generated 10,000 numbers; if we generate more, these calculated values will converge even closer to the expected values.)