Computational Physics Homework Report

Fateme RamezanZade-99100693 Problem Set 5

1. Probability Distribution Function Check

1.1 Function Description

s_err(x,y,slope,y0):

This function calculates the slope error given the data and slope and y-intercept of the best fit line using this formula:

$$\Delta s = \frac{1}{\sqrt{n-2}} \frac{\sqrt{\sum_{i=0}^{n} (y_i - y)^2}}{\sqrt{\sum_{i=0}^{n} (x_i - \bar{x})^2}}$$

1.2 Main Code

Part 1:

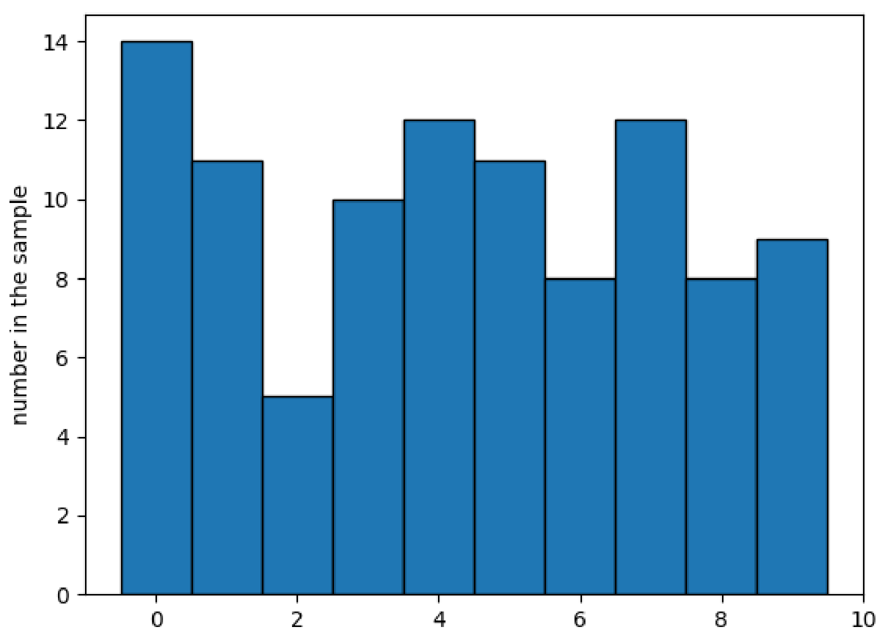
Firstly, an array of N random numbers in range 0 to 9 is created. Then the histogram of each number's reaccurance is plotted using hist function from matplot.lib library.

Part 2:

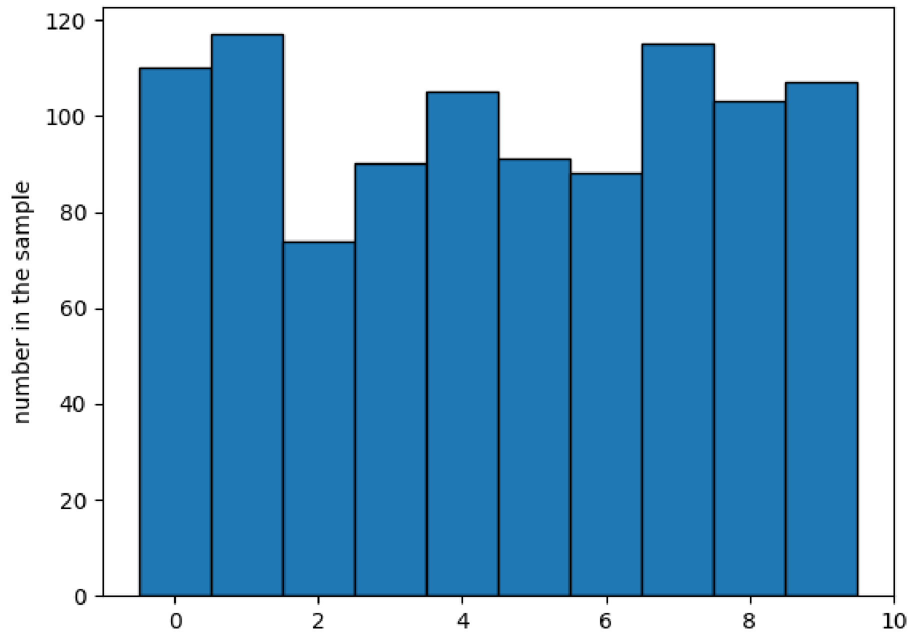
The previous procedure is repeated, this time for N ranging from 10 to 100. For each N, the standard derivative of height of the bars divided by N is calculated 100 times and the average amount is considered in order to improve accuracy. The standard derivative of these 100 amounts is reported as the error bar. Then all the data is converted to logarithmic scale. Now these points are plotted with x amount being ln(N^(-1/2)). Then a least squares line is fitted to the data, and the slope error is calculated using s_err function. Lastly, the graph is shown.

1.3 Results

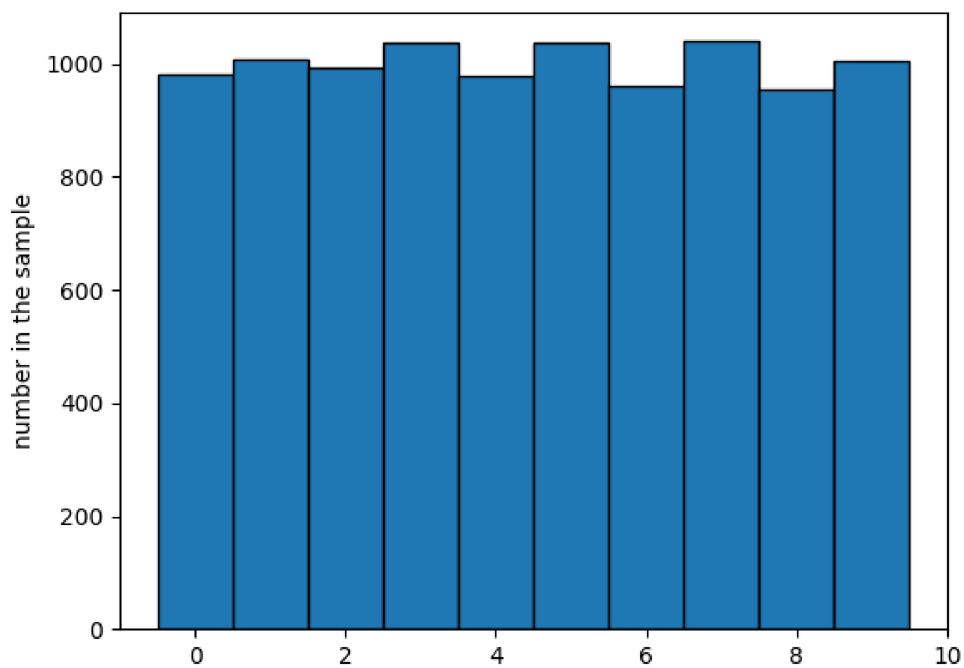
Part 1:



N=100



N=1000



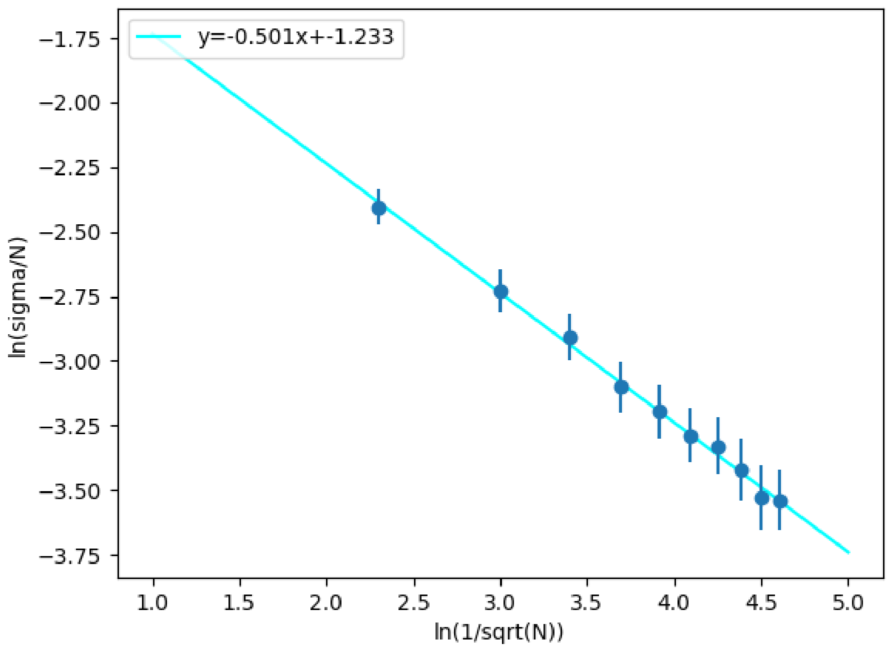
N=10000

We can see the similarities between these histograms and random deposition graph. In both cases, the outer edge gets steeper as the number of samples increases. The relation between standard deviation and N is shown in the next part.

Part 2

We can see that the slope of the fitted line equals -0.501 ± 0.010; which agrees to our expectation:

$$\frac{\sigma}{N} = \frac{1}{\sqrt{N}}$$



2. Correlation Check

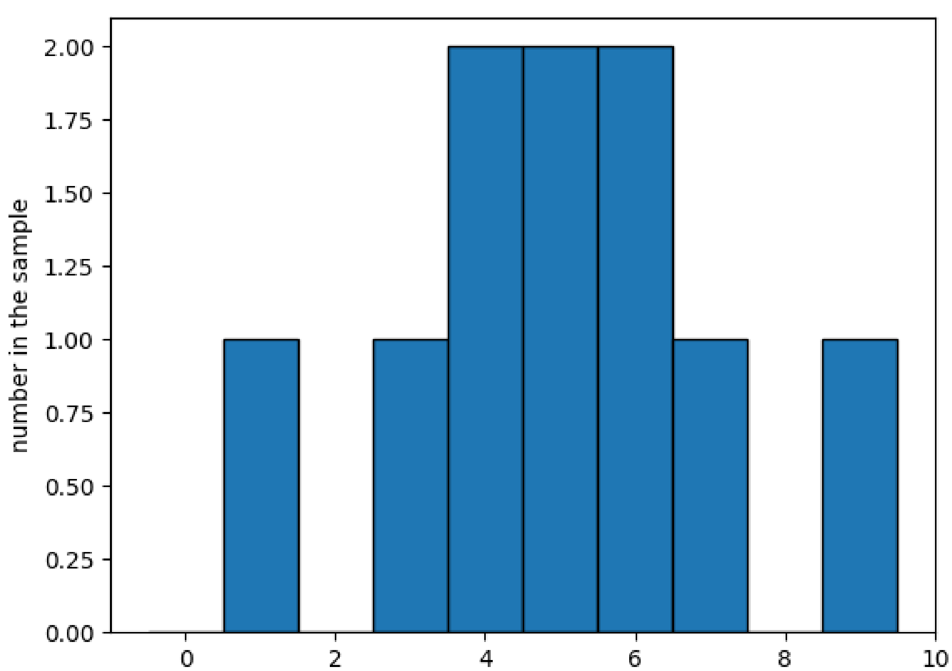
2.1 Functions Description

Same as problem 1.

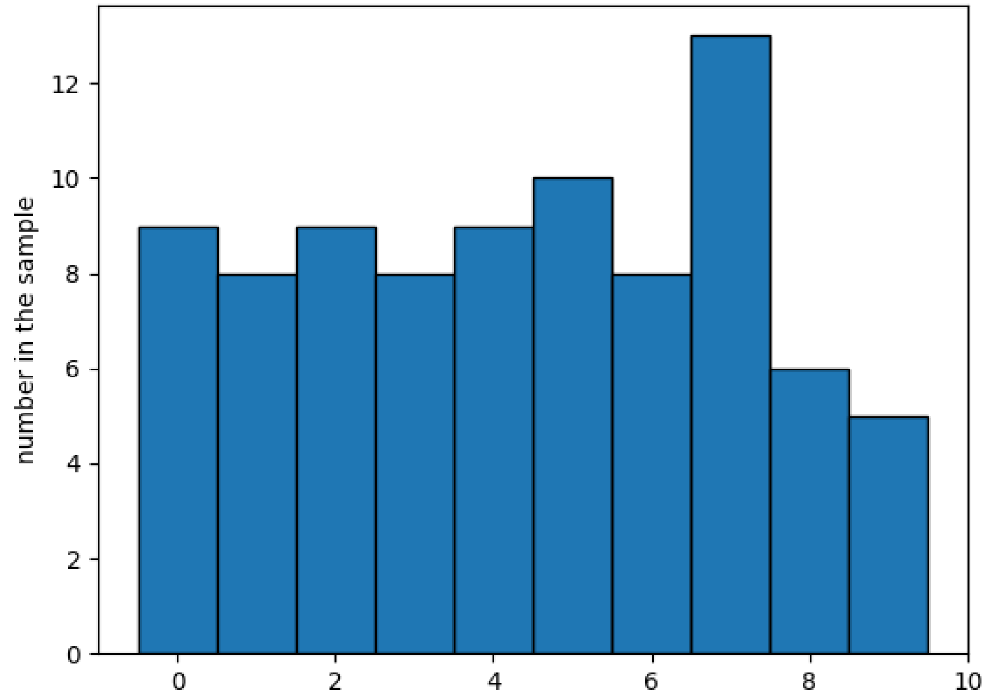
2.2 Main Code

This is also the same as problem 1, unless I did not calculate the average over a number of runs, because this time the size of the sample is the number of numbers that appeared before 4 in a N element array, which is not predictable. Therefore errorbars are not calculated neither.

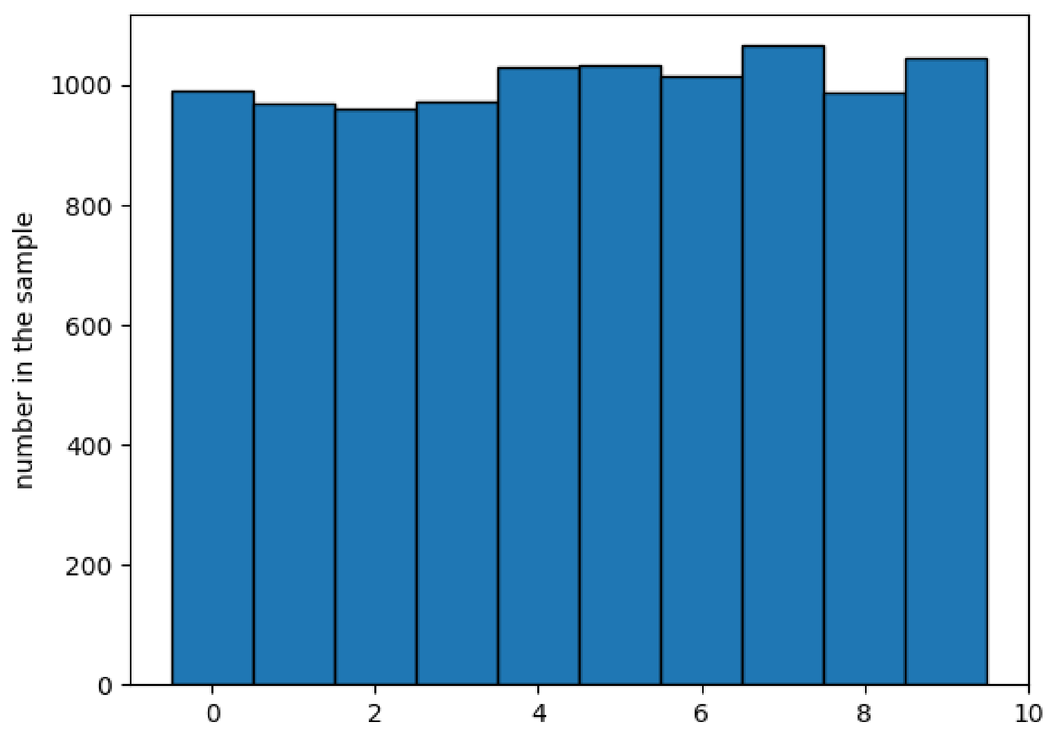
2.3 Results



N=100 , sapmle size=10

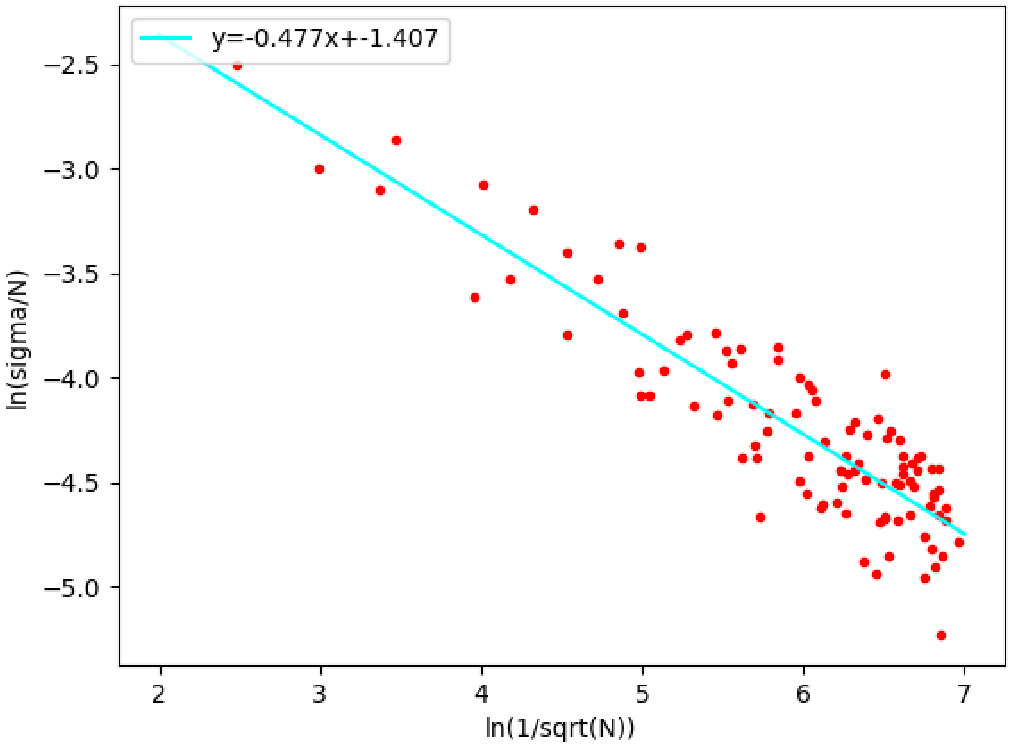


N=1000 , sapmle size=85



N=100000 , sapmle size=10068

Just like the previous problem, the outer edge gets steeper as the number of samples increases. Now we check the standard deviation:



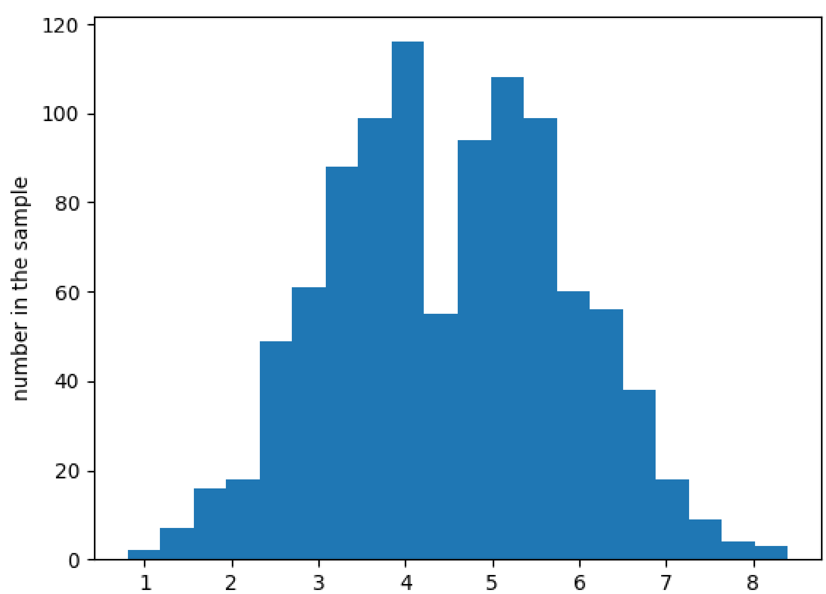
We can see that the slope equals 0.477 ± 0.23. the result obviously has larger error compared to the previous one, because only one sample is considered rather than the average. But is still meets our expectation and we can conclude that the sample has a uniform distribution function thus there is no correlation in the original sample.

3. Central Limit Theorem

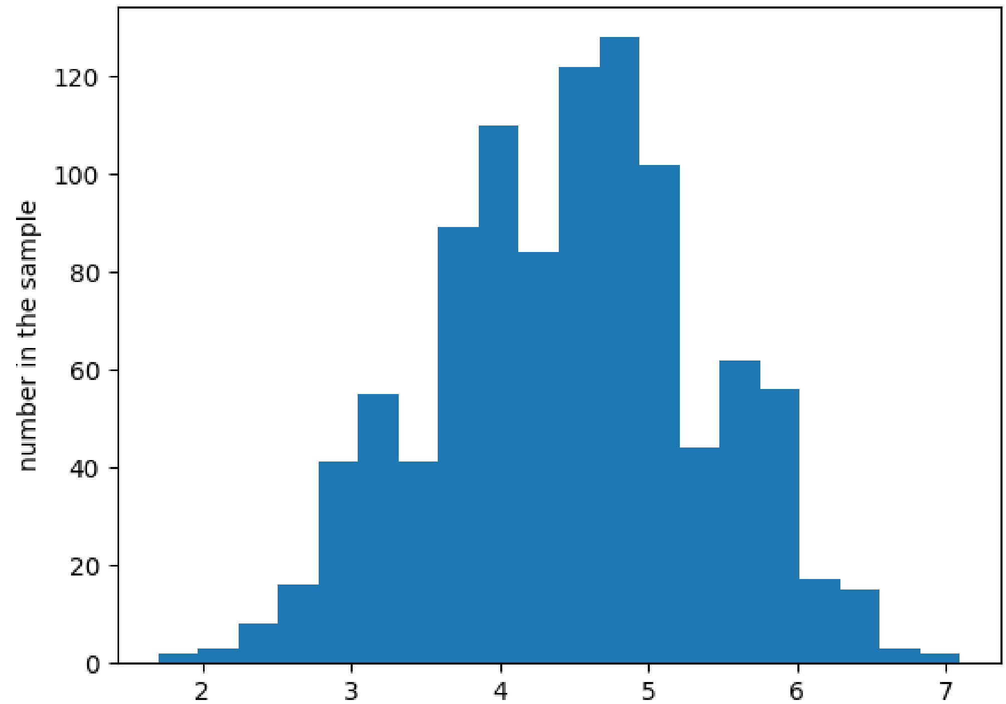
3.1 Main Code

As discussed in the book, we want to check that the distribution function of $y = \frac{1}{N}\sum_{i=0}^{N} x_i$ is a gaussian one, where x_i is a random number. Thus I calculated y for 1000 samples with length N and plotted the histogram.

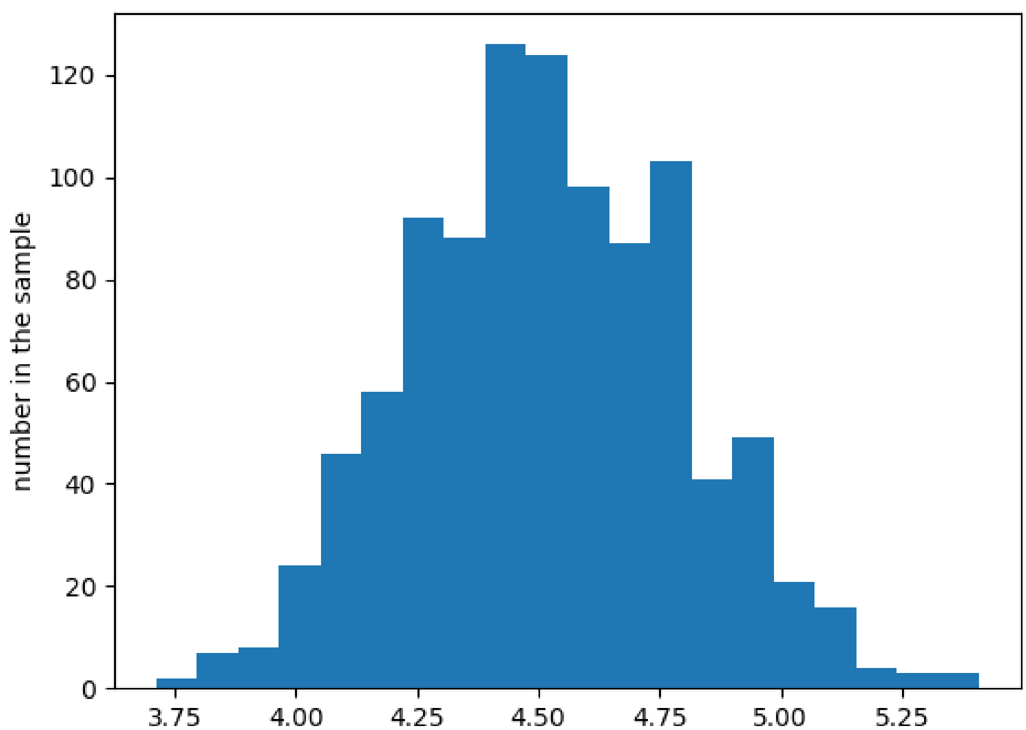
3.2 Results



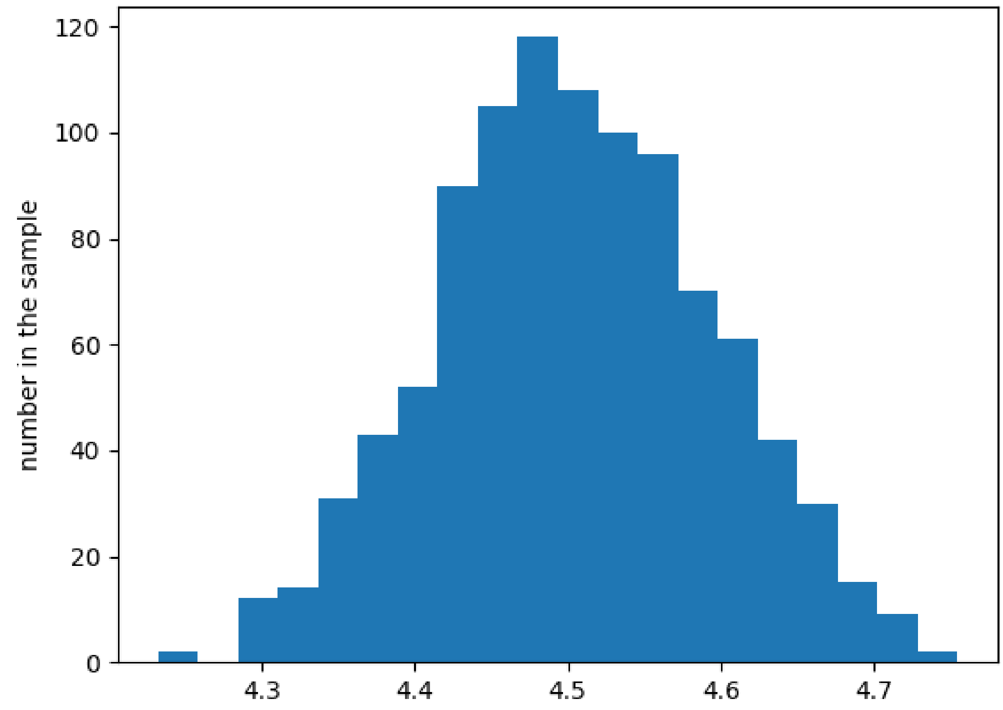
N=5


number in the sample
0
20
40
60
80
100
120
2
3
4
5
6
7

N=10


number in the sample
0
20
40
60
80
100
120
3.75
4.00
4.25
4.50
4.75
5.00
5.25

N=100



N=1000

We can see that all of the graphs resemble a gaussian curve, getting closer as N increases.

4. Using Transform Function

4.1 Theory

Continueing from equations 32 and 24 in the book, we will have:

$$y_1 = \int_0^r \frac{1}{\sigma^2} e^{-\frac{\rho^2}{2\sigma^2}} \rho \, d\rho \quad , \quad y_2 = \int_0^\theta \frac{d\theta}{2\pi}$$

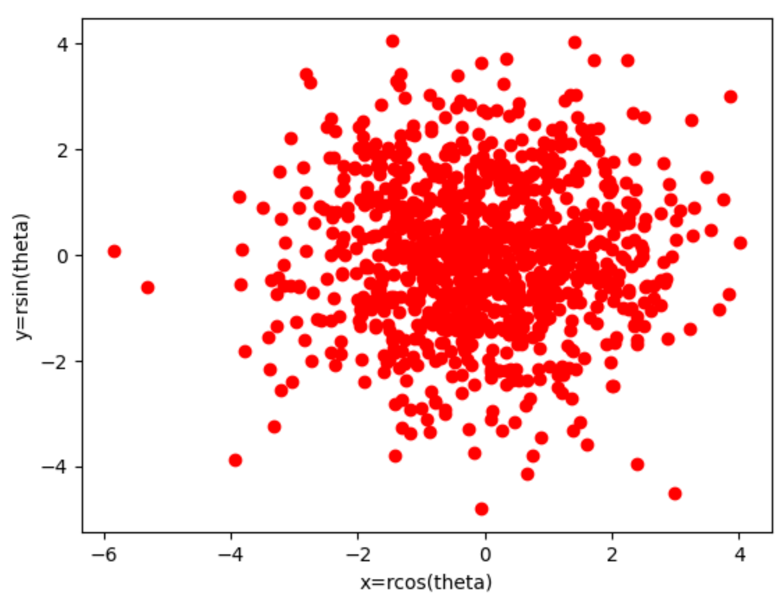
$$\Rightarrow r = \sqrt{-2\sigma \ln (1 - y_1)} \quad , \quad \theta = 2\pi y_2$$

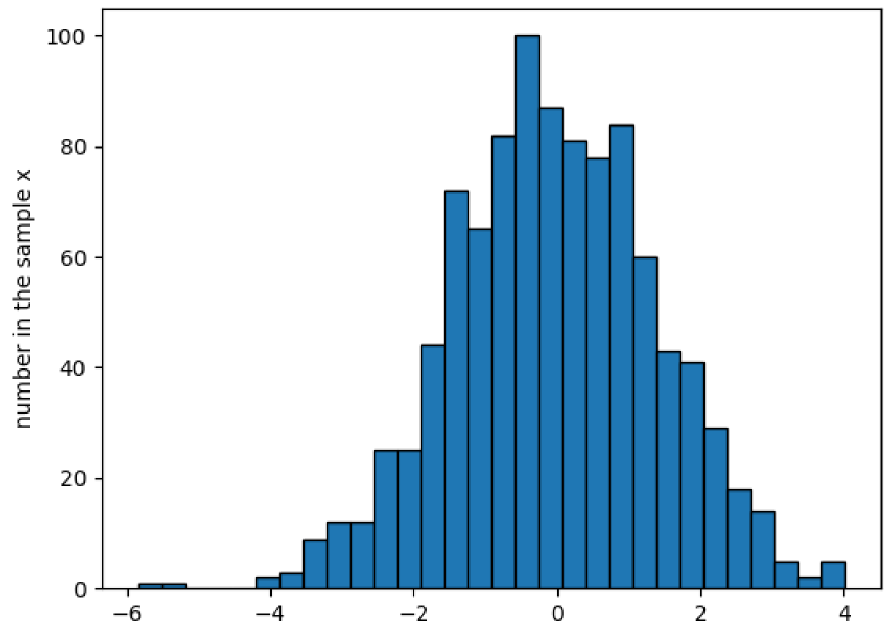
4.2 Main Code

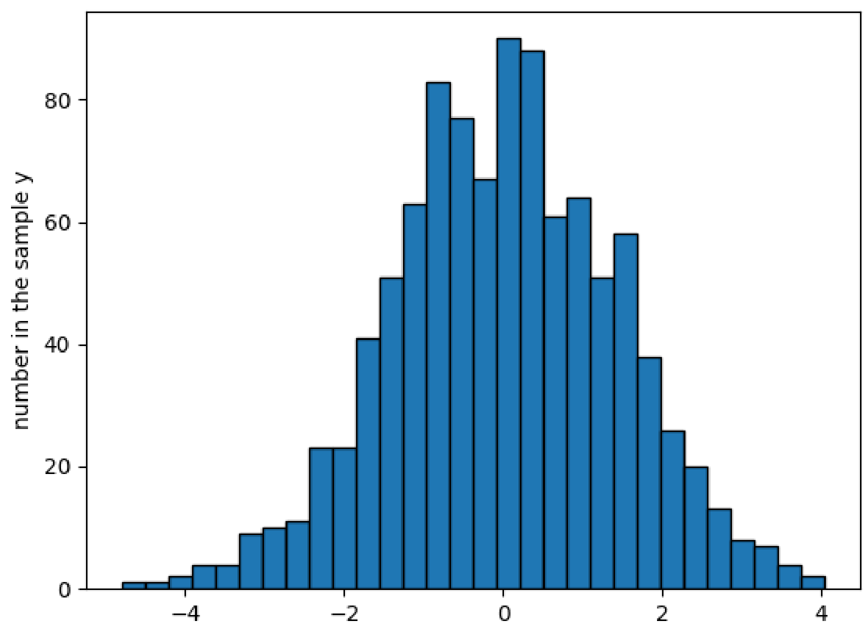
Using the results above, I defined two arrays for r and θ and fill them with amounts calculated for some random elements. Then converted them to cartesian coordinates and plotted them on an x-y plane. Then plotted two histograms for x and y amounts.

4.3 Result

We can see that the data is symmetrically distributed around the origin with more density closer to it.







We can see that the distributions are gaussian, as expected. These results used 1000 pair of random samples. Increasing the number would develop the approximation.

5. Integration

5.1 Theory and Functions Description

For the important sampling method, firstly, I calculate the transform function, knowing $p(x) = 1/2$ in range (0,2) and $g(y) = e^{-y}$:

$$\int_0^x \frac{dx}{2} = \int_0^y e^{-y}\, dy \Rightarrow y = -\ln\,(1 - \frac{x}{2})$$

Then the integral would be:

$$I = \int_{-\infty}^{\infty} e^{-x}\, dx\ \langle \frac{e^{-x^2}}{e^{-x}} \rangle$$

Where the samples for average term come from $g(x)$.

The functions func, dist and g each correspond to $f(x)$, $g^{-1}(x)$ and $g(x)$.

5.2 Main Code

Part 1:

In this part, I used the simple sampling method with 10000 random numbers to find the average. I repeated the process for 100 times to find the standard derivative. Lastly printed the mean value and the standard derivative.

Part 2:

This part is the same as the previous one, except that I have used the important sampling method using the relations I explained before.

5.3 Results

	I	σ	$\delta I(I = 0.8820813907624217)$
Simple sampling	0.8826434895056374	0.006573094869377035	0.0637%
Important sampling	0.8819004056659915	0.0028346776632665544	0.020%

We can see that in the important sampling model, the data is more efficient (smaller sigma) and the relative error is smaller.