Report - Question 4

Atishay Jain & Megh Gohil

August 22, 2022

1 M-shaped Probability Density Function (PDF)

1.1 Generating Independent Draws from $P_X(.)$

PDF i.e probability distribution function is a function that assigns probability to each point on $\mathbb{R}$
While CDF i.e. cumulative distribution function integrates the value of PDF over $-\infty \to x$ for given x.
So the $\int_{-\infty}^{\infty} P_X(x)dx = 1$

The method used to generate random numbers for a given PDF is inverse CDF method, where we provide a randomly generated number from uniform distribution to the inverse CDF and it gives out a number.
$f_X : \mathbb{R} \to [0, 1]$ is the CDF function with given properties.

- It is right continuous
- It is monotonically increasing function.

Now define a function

$$f^{-1} : [0, 1] \to \mathbb{R} \tag{1}$$

in such a way that,

$$f_X^{-1} = min\{x : f_X(x) \geq y\}, y \in [0, 1] \tag{2}$$

why we used minimum :
Suppose for a given y there is a interval $[x_i, x_j]$ such that $\forall x \in [x_i, x_j] f_X(x) = y$. Now since CDF is inetgration of PDF and if integration over an interval is constant (here it is y) then the $P_X(x) = 0 \forall x \in (x_i, x_j] \, and \, P_X(x_i) \neq 0$
So for every value of $y \in [0, 1]$ there is a unique x.
And this unique value is what is the randomly generated number from inverse CDF function.This function takes input a randomly generated y from uniform distribution from interval [0,1]

1.2 Histogram

APPROACH :

- First generate a uniform random number.
- inverse CDf function will give a number correspondin to it.
- generate such 10^5 numbers and plot the histogram which clearly represents the PDF

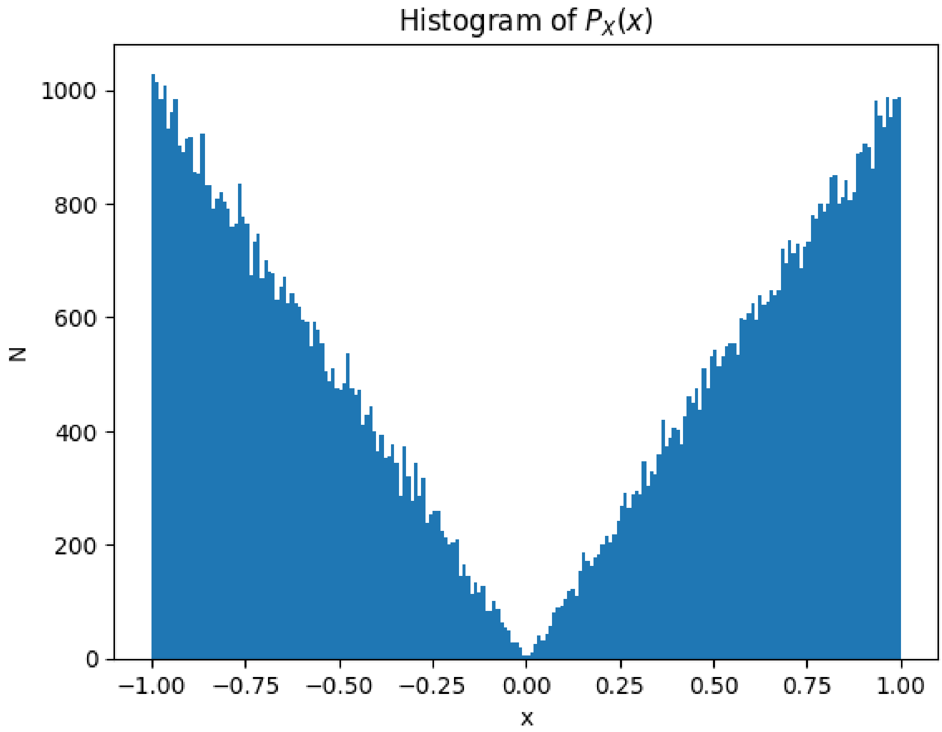

Figure 1: Histogram (with 200 bins) using M := 10^5 draws from $P_X(.)$

1.3 Cumulative Distribution Function (CDF)

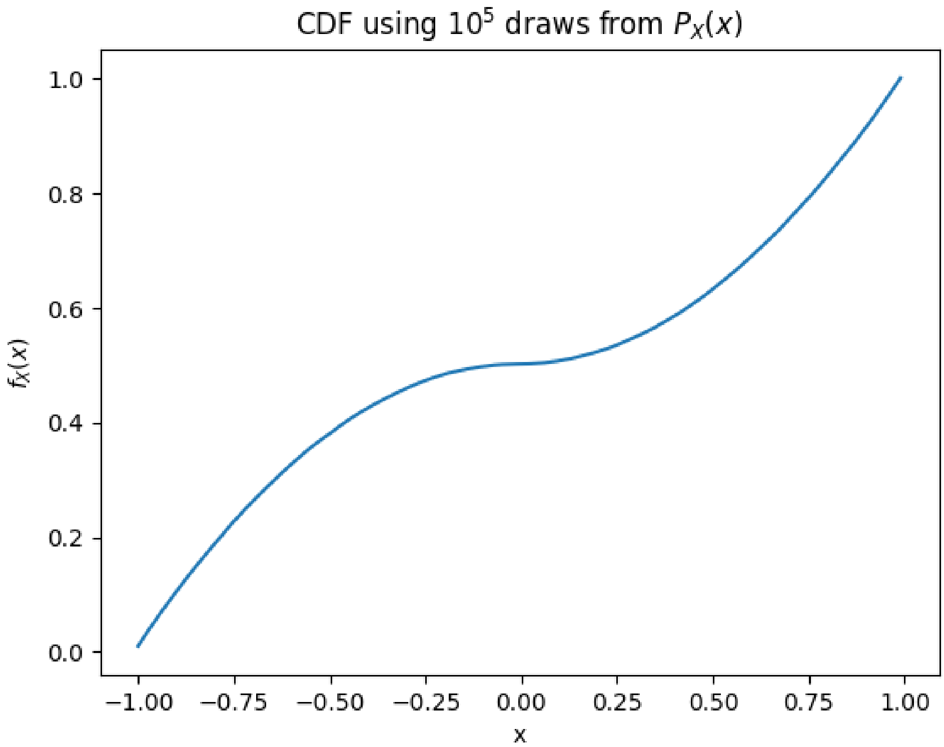

Figure 2: CDF using M := 10^5 draws from $P_X(.)$

2 Independent & Identical Random Variables, $P_{Y_N}(.)$

2.1 Generating Independent Draws from $P_{Y_N} := \frac{\sum_{i=0}^{N} X_i}{N}$ using previous code

For any value of N we generated the random values using previous code.
APPROACH : To do this we first produced N elements randomly using previous random number generator and then took its average value. This new value is our random number for this generator.
The histogram of 10^4 generated numbers from above code for different values of N are shown below.
CONCLUSION : And it can be seen that the distribution tends to Gaussian distribution as the value of N increases from 1 to 64.

2.2 Histograms

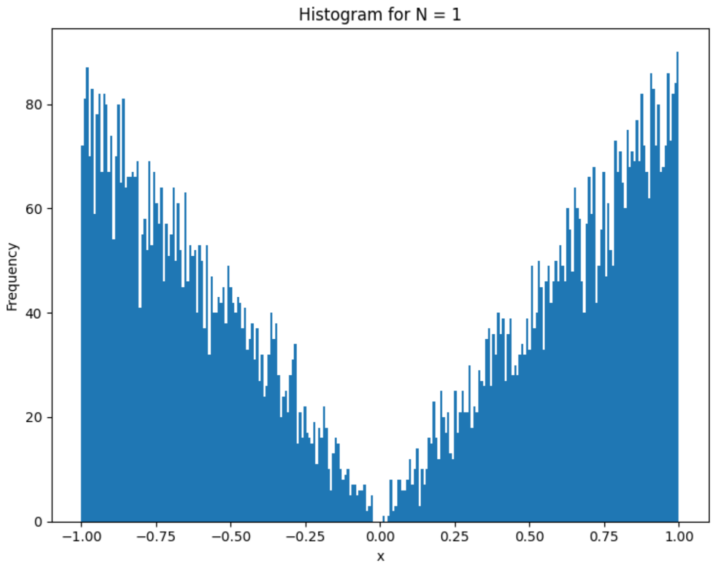

Figure 3: Histogram for N = 1 using M := 10^4 draws from $P_{Y_N}(.)$

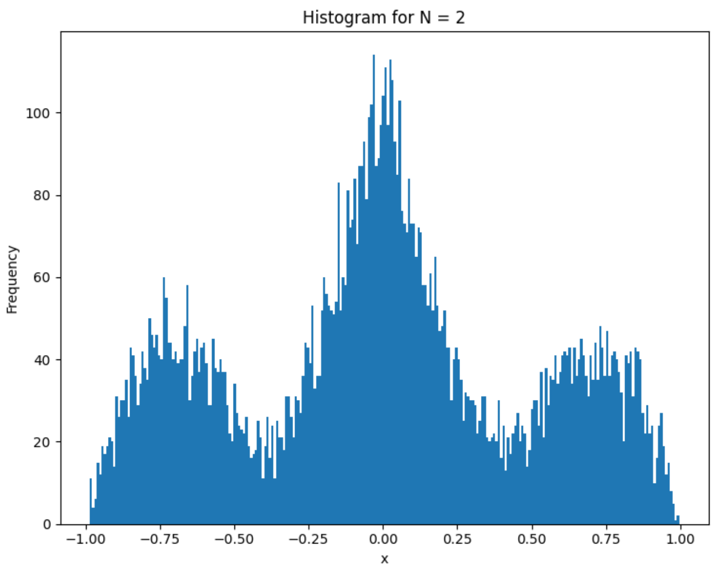

Figure 4: Histogram for N = 2 using M := 10^4 draws from $P_{Y_N}(.)$

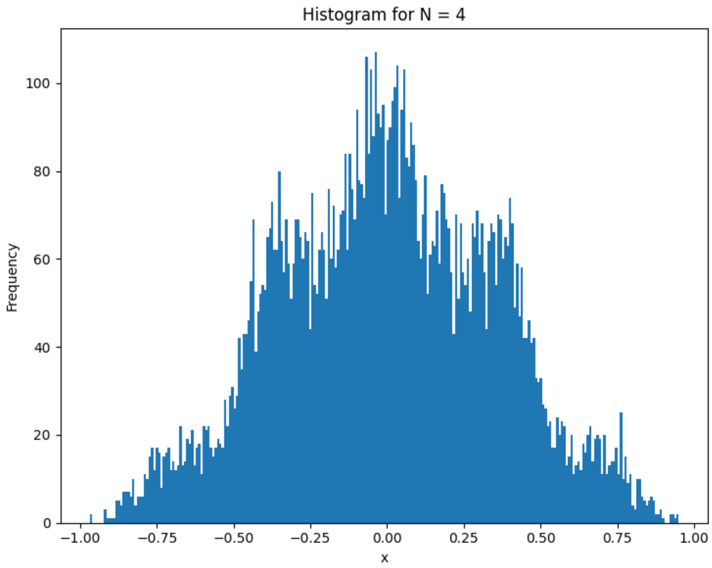

Figure 5: Histogram for N = 4 using M := 10^4 draws from $P_{Y_N}(.)$

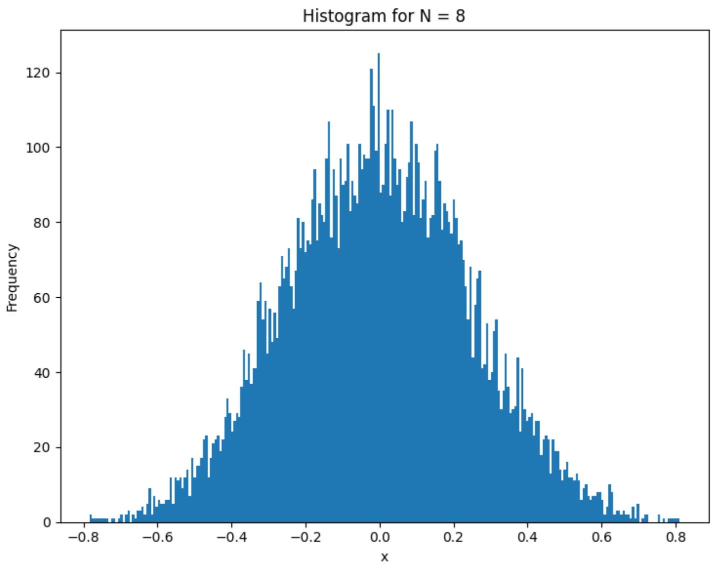

Figure 6: Histogram for N = 8 using M := 10^4 draws from $P_{Y_N}(.)$

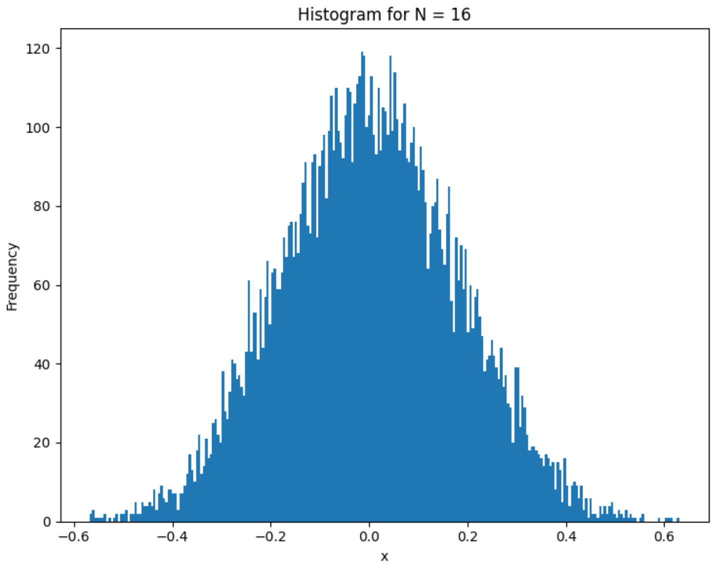

Figure 7: Histogram for N = 16 using M := 10^4 draws from $P_{Y_N}(.)$

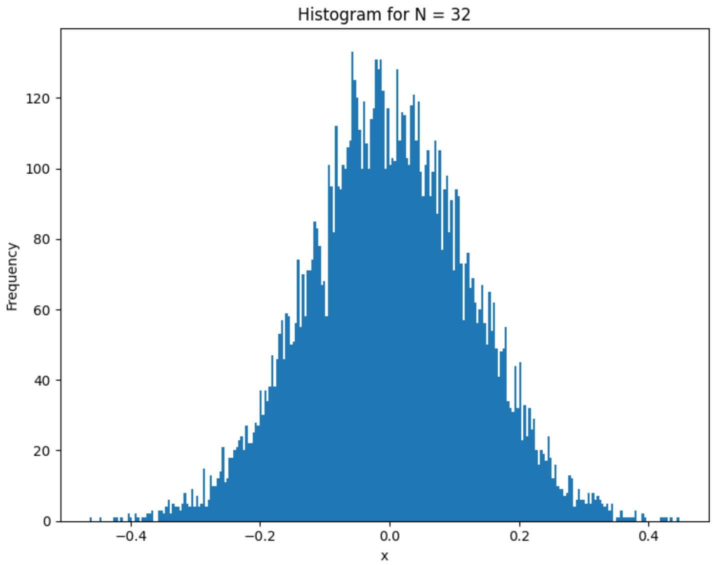

Figure 8: Histogram for N = 32 using M := 10^4 draws from $P_{Y_N}(.)$

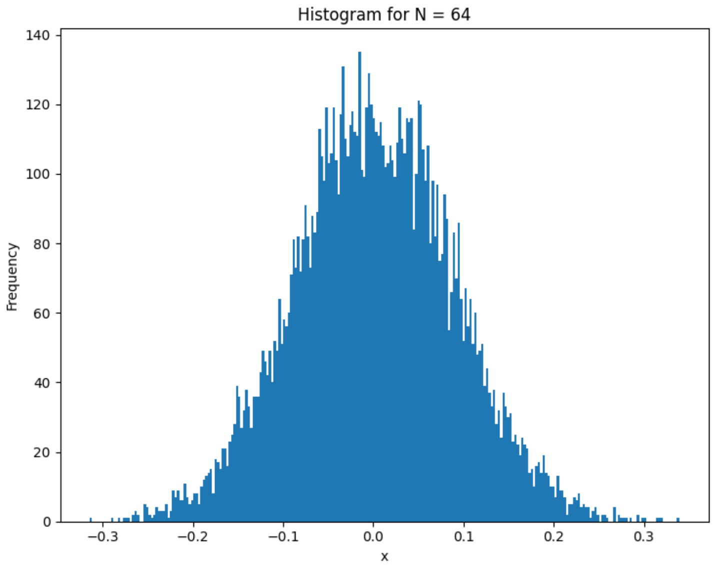

Figure 9: Histogram for N = 64 using M := 10^4 draws from $P_{Y_N}(.)$

2.3 CDFs associated with Y_N

APPROACH : The approach is pretty simple, use the histogram from previous code and find the cumulative sum of it and draw the graph using matplotlib.pyplot.plot function.
CONCLUSION : From all CDFs it can be confirmed that for large values of N the slop of CDF becomes steeper and steeper around 0 which means the values of PDF around this area is much higher than other areas. And this can be seen from the histogram of previous graphs.

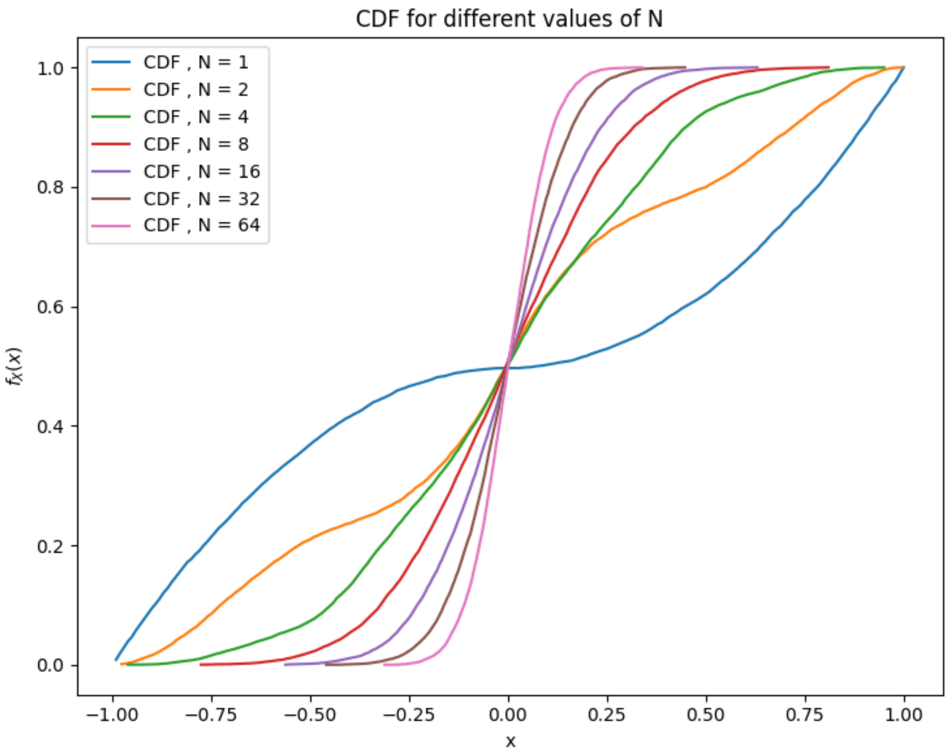

Figure 10: Histogram (with 200 bins) using M := 10^5 draws from $P_X(.)$