

Solution for Q5

Kavya Gupta

September 3, 2023

Instructions to Run the Code

The MATLAB code for this analysis is provided in the file `A2Q5.m` present in main zip. This code generates histograms, empirical cumulative distribution functions (ECDFs), and compares them to Gaussian cumulative distribution functions (CDFs) for different sample sizes N . The code also calculates and prints the Mean Absolute Deviation (MAD) values for each N .

Important Note: MATLAB Online displays max of 20 images at once, so there maybe some warning at 21st image, please ignore that.

The zip also contains a folder `Q5` which contains all the pictures of histograms and cdfs related with this question along with MAD vs N .

Results

For $N = 5$:

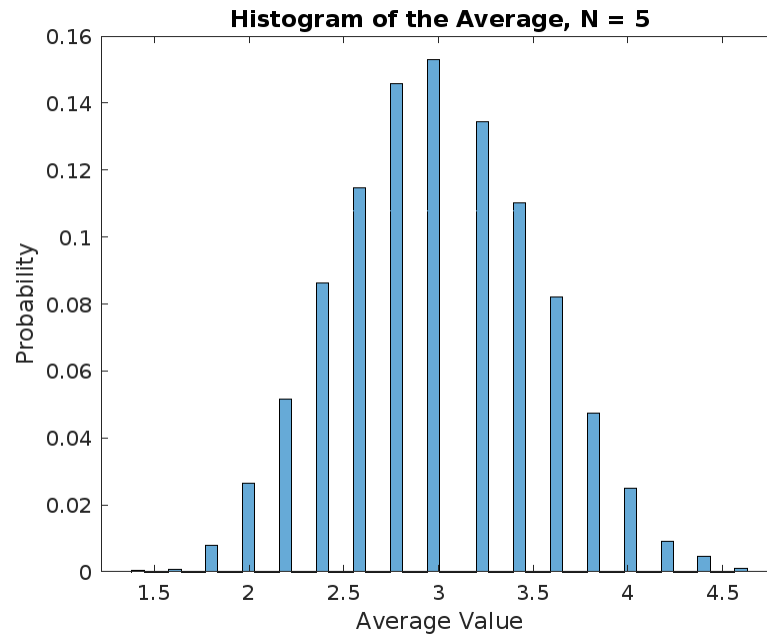


Figure 1: Histogram of the Average

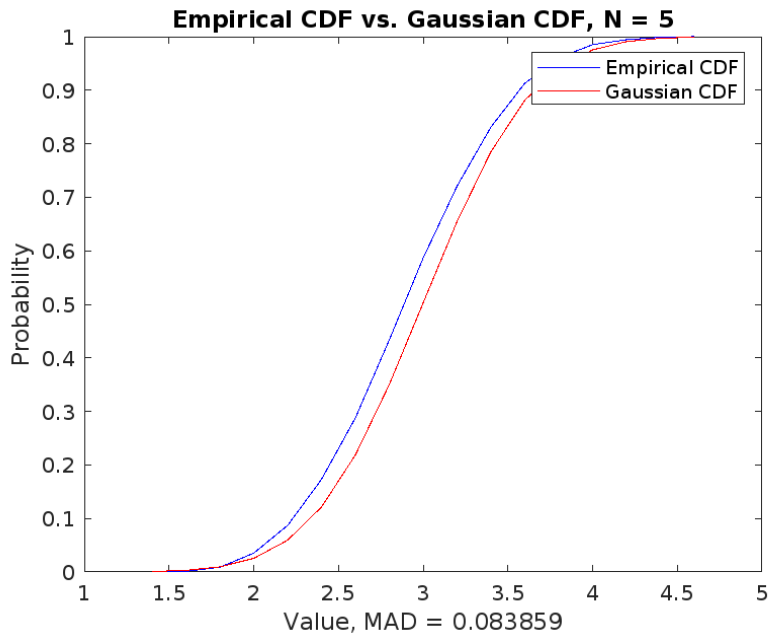


Figure 2: Empirical CDF vs. Gaussian CDF

The MAD for $N = 5$ is 0.083859.

For $N = 10$:

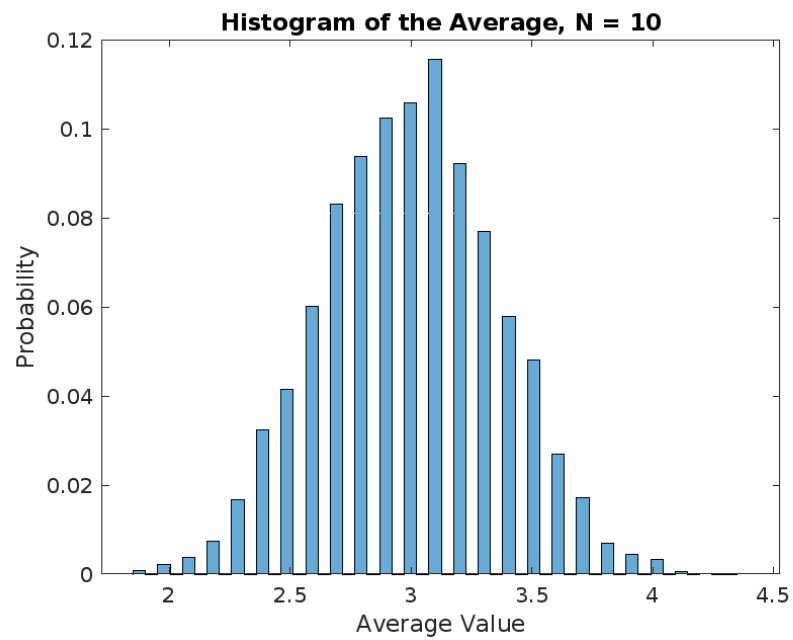


Figure 3: Histogram of the Average

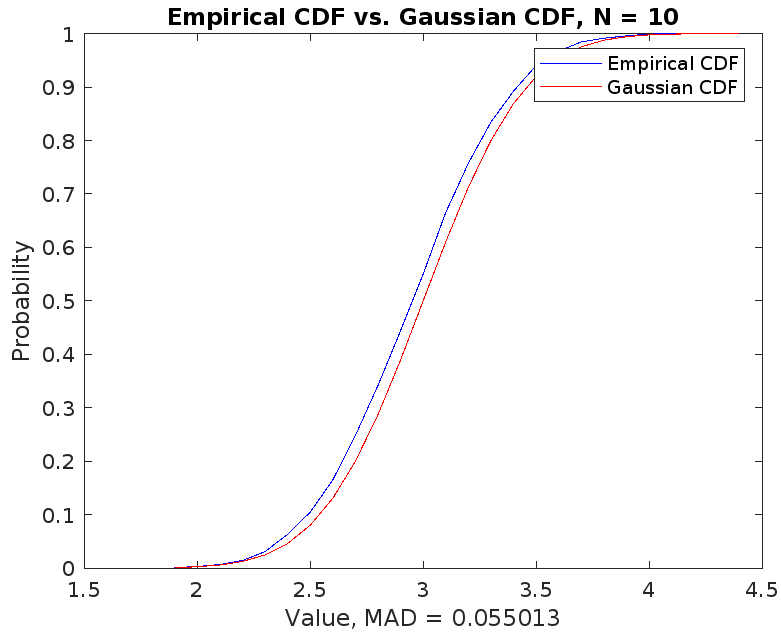


Figure 4: Empirical CDF vs. Gaussian CDF

The MAD for $N = 10$ is 0.055013.

For $N = 20$:

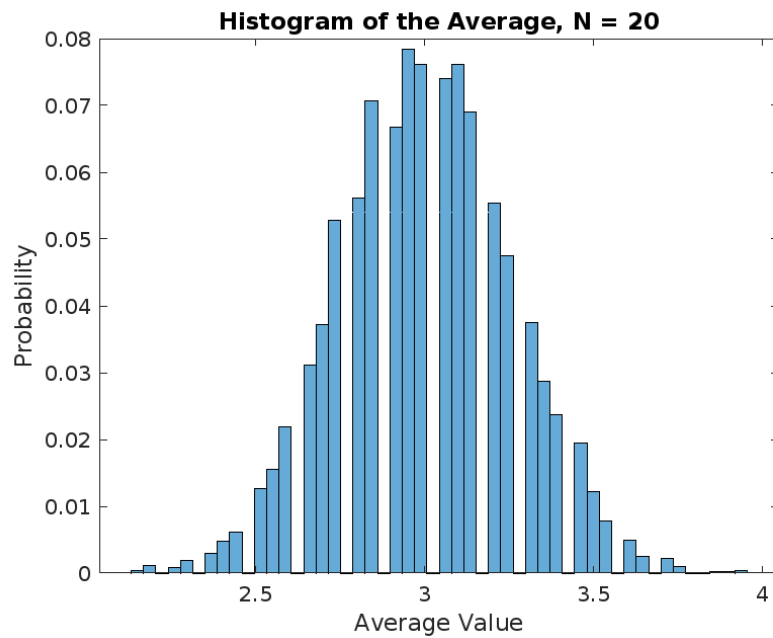


Figure 5: Histogram of the Average

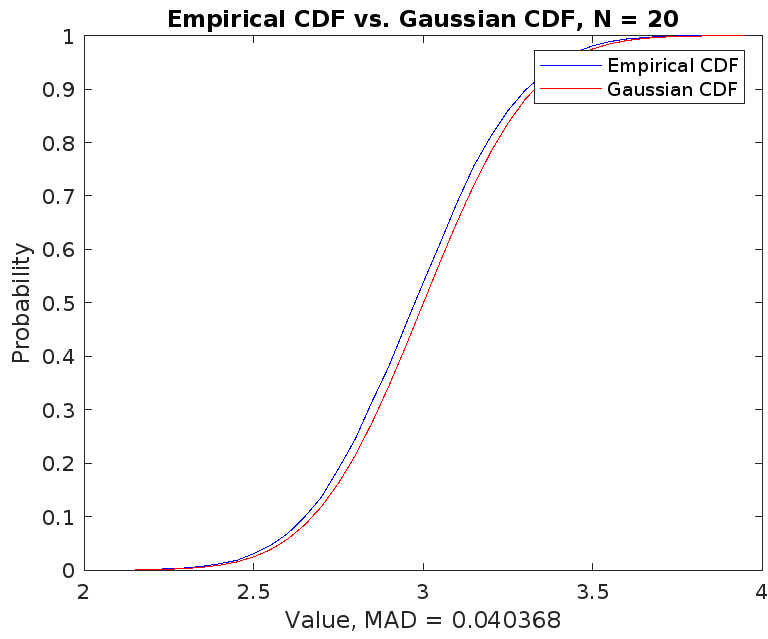


Figure 6: Empirical CDF vs. Gaussian CDF

The MAD for $N = 20$ is 0.040368.

For $N = 50$:

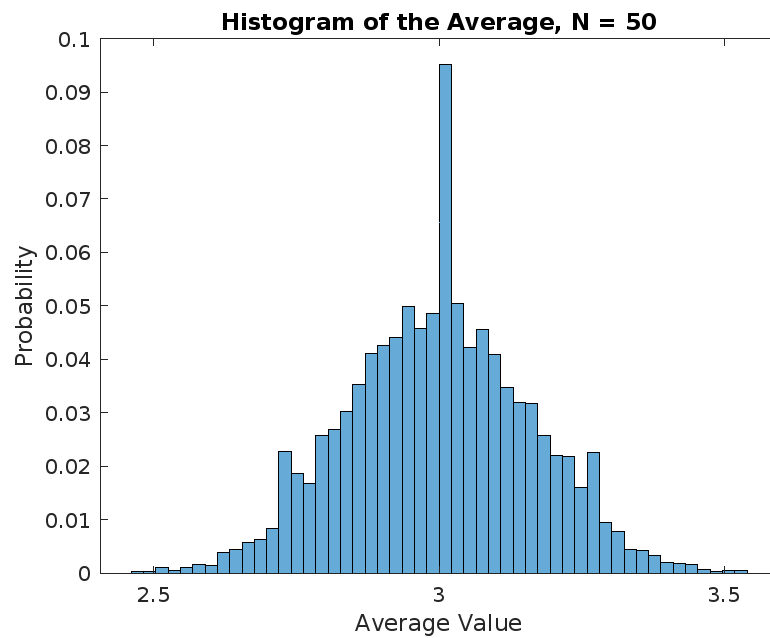


Figure 7: Histogram of the Average

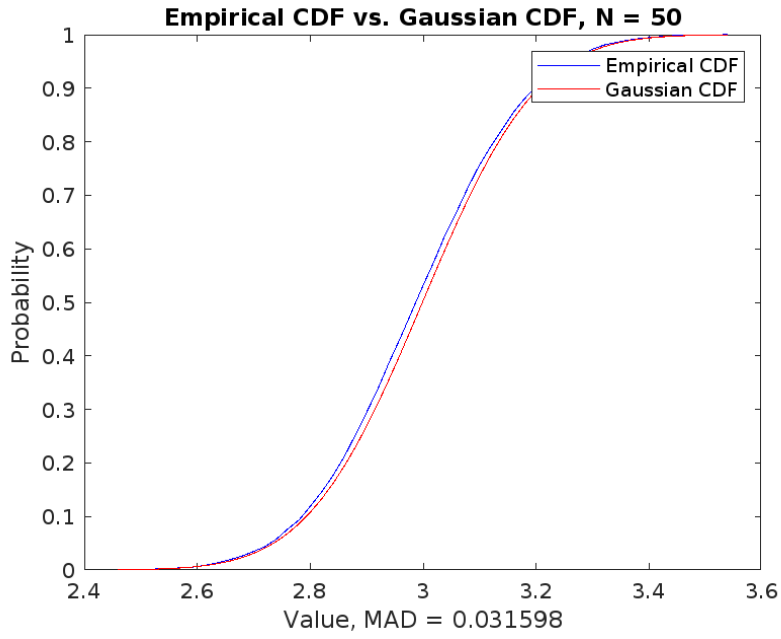


Figure 8: Empirical CDF vs. Gaussian CDF

The MAD for $N = 50$ is 0.031598.

For $N = 100$:

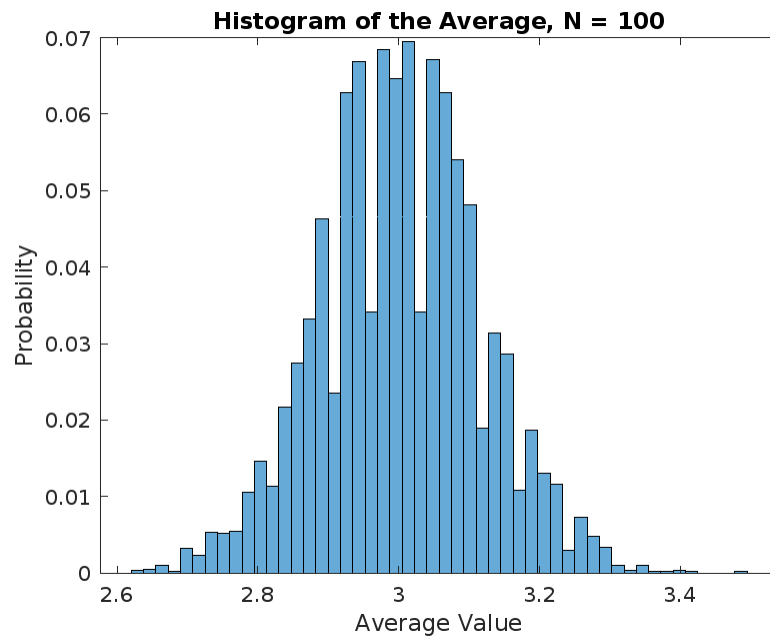


Figure 9: Histogram of the Average

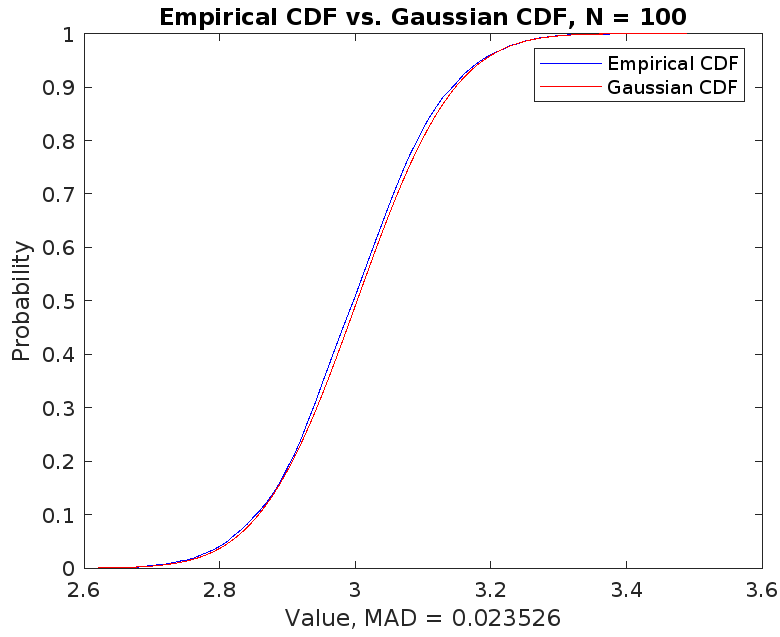


Figure 10: Empirical CDF vs. Gaussian CDF

The MAD for $N = 100$ is 0.023526.

For $N = 200$:

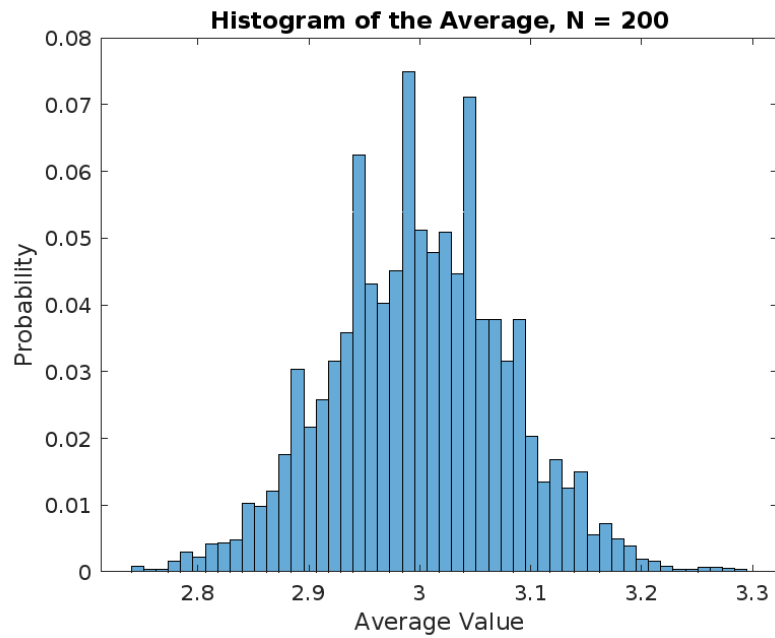


Figure 11: Histogram of the Average

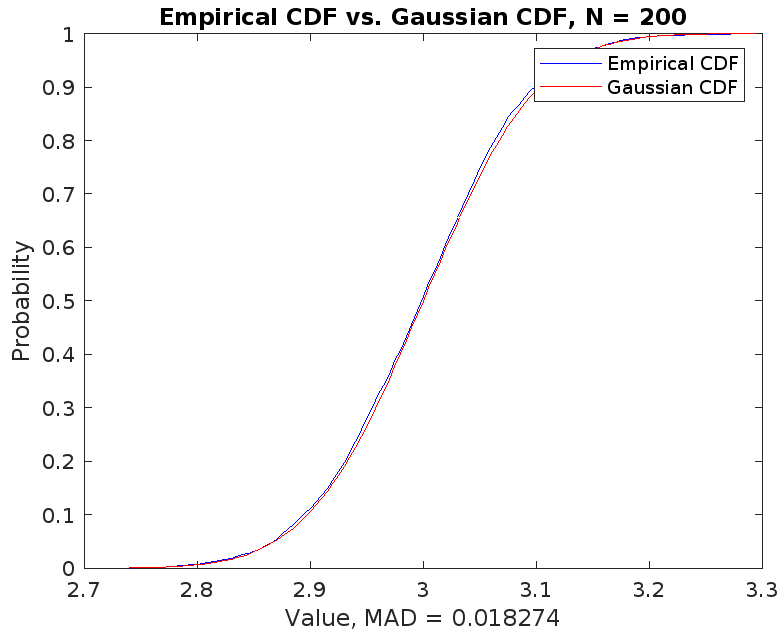


Figure 12: Empirical CDF vs. Gaussian CDF

The MAD for $N = 200$ is 0.018274.

For $N = 500$:

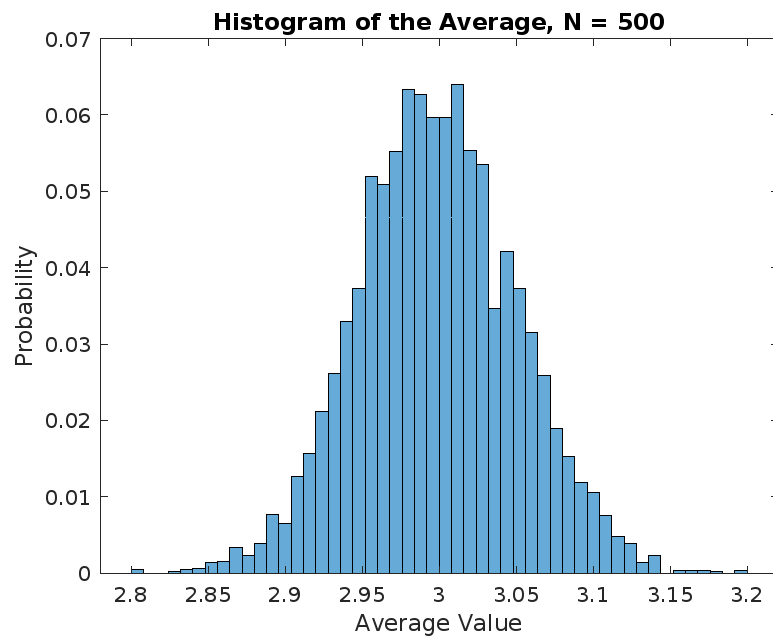


Figure 13: Histogram of the Average

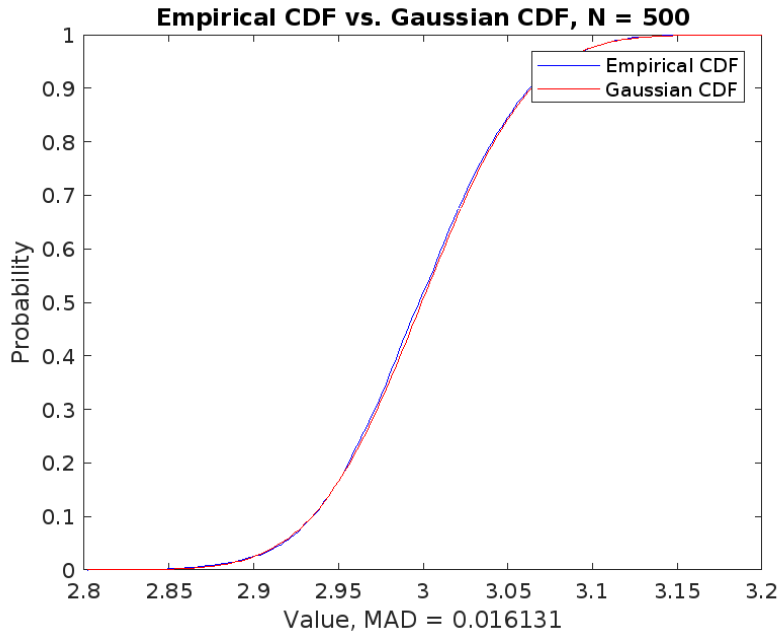


Figure 14: Empirical CDF vs. Gaussian CDF

The MAD for $N = 500$ is 0.016131.

For $N = 1000$:

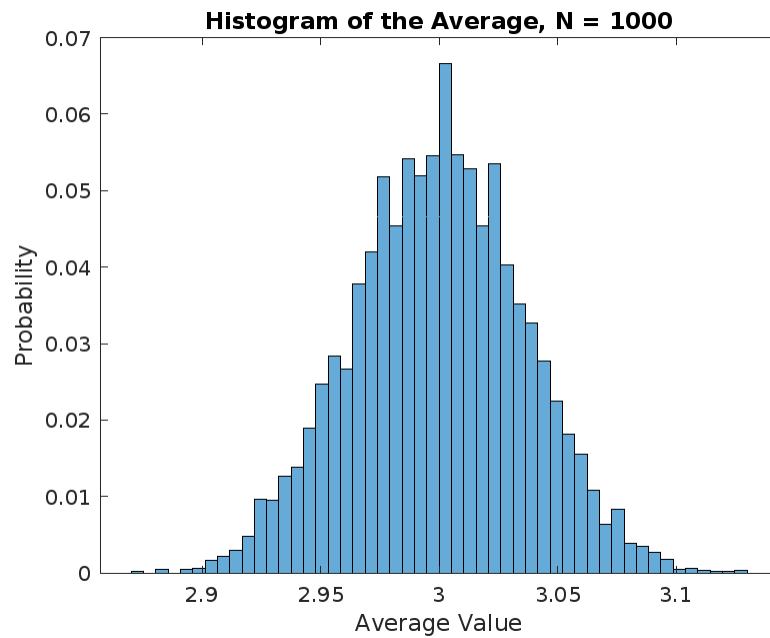


Figure 15: Histogram of the Average

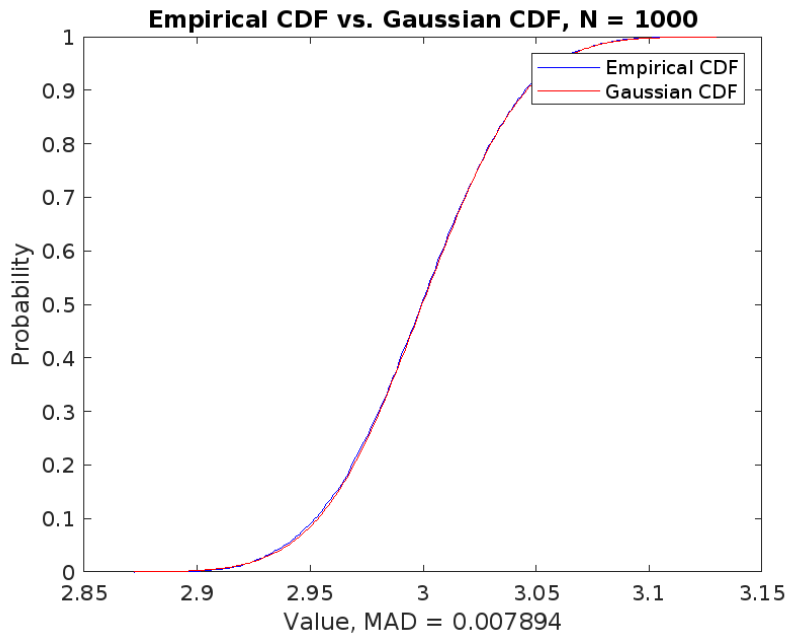


Figure 16: Empirical CDF vs. Gaussian CDF

The MAD for $N = 1000$ is 0.007894.

For $N = 5000$:

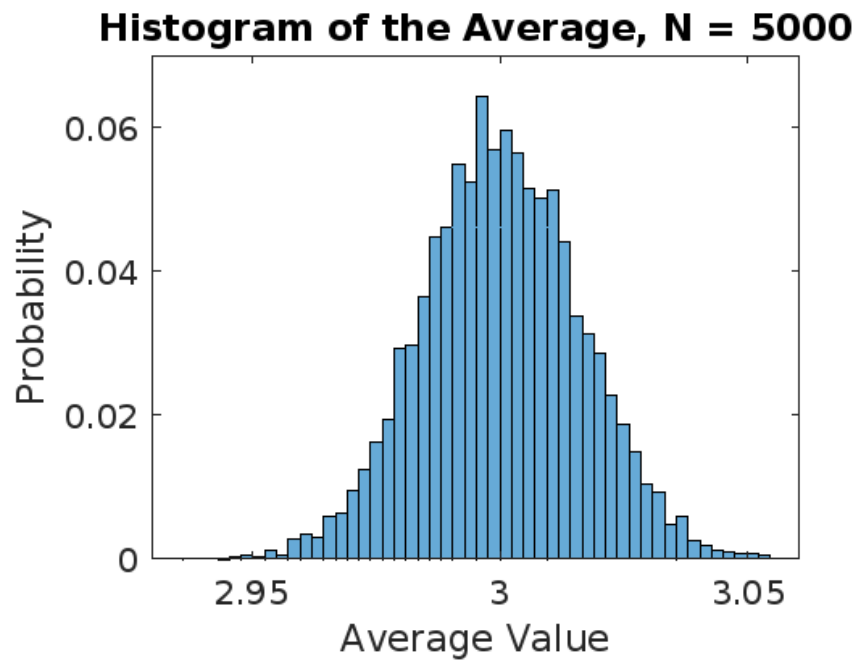


Figure 17: Histogram of the Average

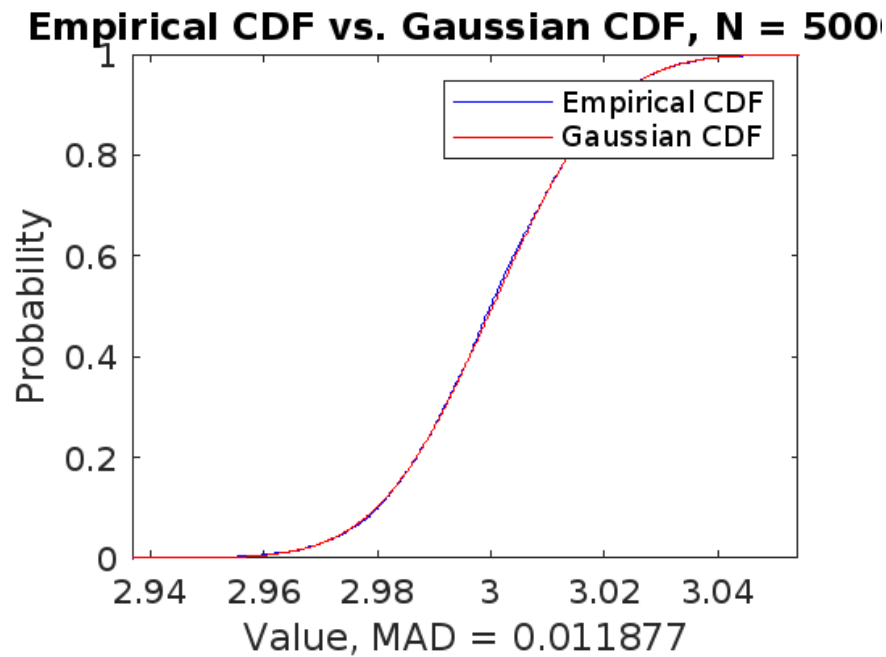


Figure 18: Empirical CDF vs. Gaussian CDF

The MAD for $N = 5000$ is 0.011877.

For $N = 10000$:

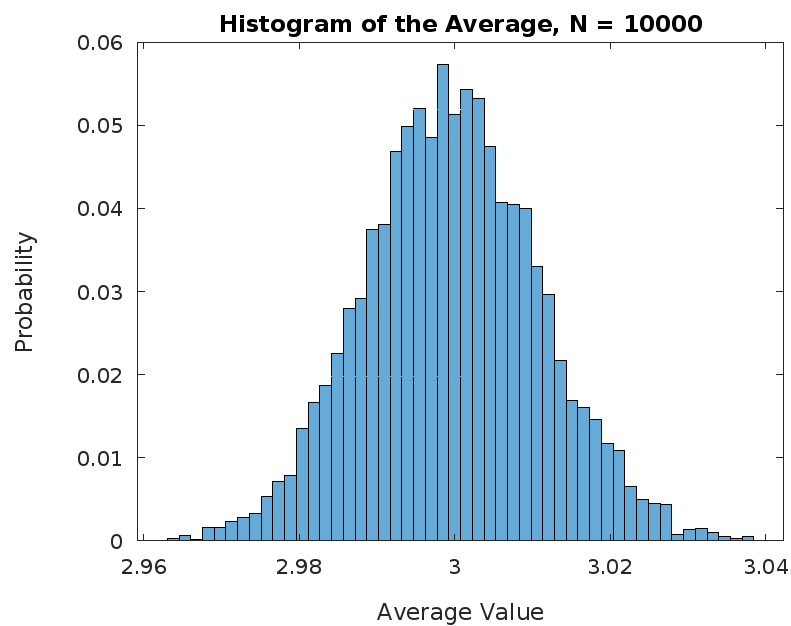


Figure 19: Histogram of the Average

Empirical CDF vs. Gaussian CDF, $N = 1000$

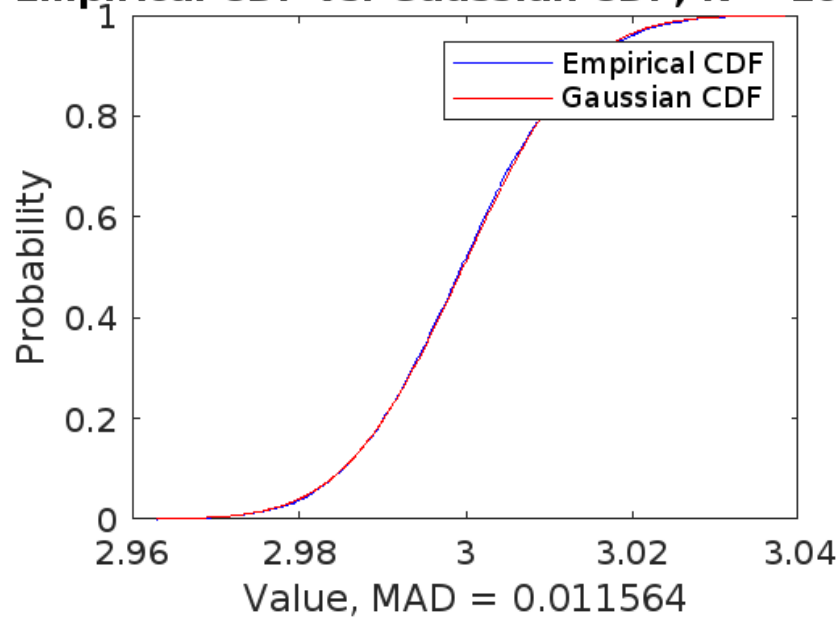


Figure 20: Empirical CDF vs. Gaussian CDF

The MAD for $N = 10000$ is 0.011564.

MAD vs. Sample Size (N)

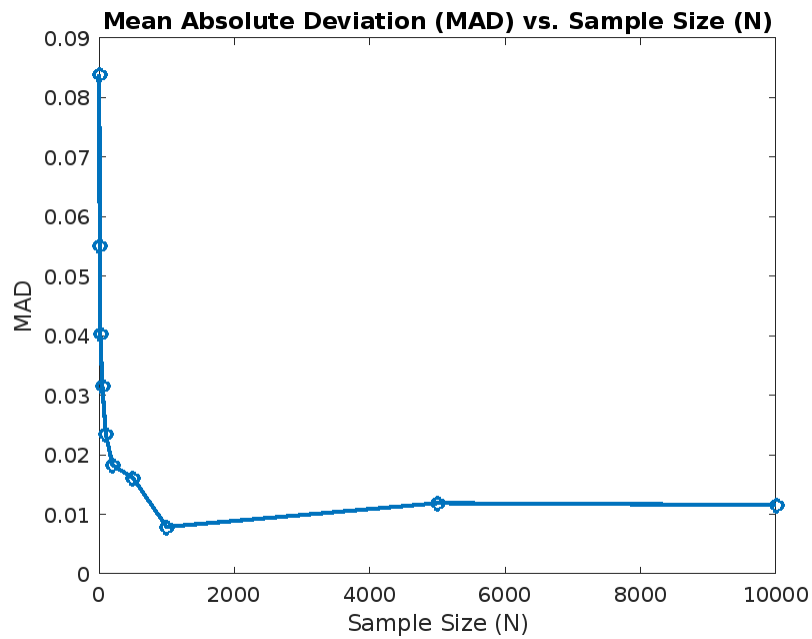


Figure 21: Mean Absolute Deviation (MAD) vs. Sample Size (N)

The plot above shows how the Mean Absolute Deviation (MAD) changes with increasing sample size (N). As N increases, the empirical cumulative distribution function (ECDF) gets closer to the Gaussian cumulative distribution function (CDF), resulting in lower MAD values. This was an important observation owing to **Central Limit Theorem**.