[Robert Irwin]

ECE 3522: Stochastics

Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 19121

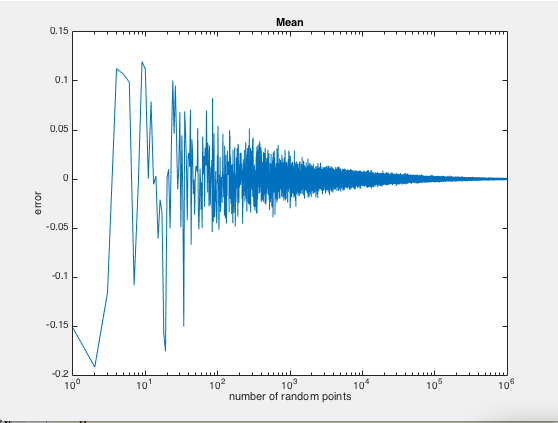
# Problem Statement

In this assignment we are looking into the statistics of random processes; in particular of a random number generator. We will generate one random number, compute the mean, variance, and plot the error of these values. This process will be repeated for 2 through 1 million random numbers. Then we will obtain the pdf for each set of random numbers and compute the mean squared error for each set. The mean squared error for all 1 million sets of random numbers will be plotted, along with the pdf that corresponds to the set of 10 random numbers, 1000 random numbers, and 1,000,000 random numbers.

# Approach and Results

To begin the experiment, it was important to understand what the expected mean and variance was for our particular random number generator. To find the mean, we first had to determine what the range of our random number generator was. It was [0,1]. From this information, we can conclude tat the expected mean was .5, (if the generator is unbiased), and the variance is ­, where a and b represent the upper and lower bound of the range of the random number generator. This yields a variance of . For each set of random numbers, the error is computed as, 

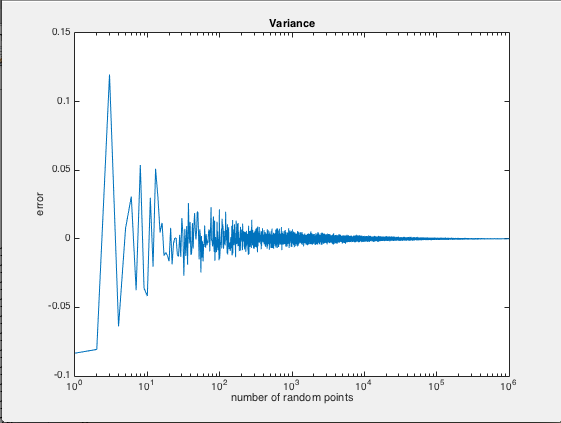
Implementing the above equation on the left, the error of the mean is shown below in Figure 1.



*Figure 1: Plot of the Error of the Mean*

As we can see above, when the number of random numbers is small, the error is significantly large, but as the size of the set of random numbers increases, the error of the mean converges to 0. This means the average of the random numbers converges to the expected value of .5. This is in fact the result we expect, and we will see this in a different way when we analyze the pdfs described in the Introduction.

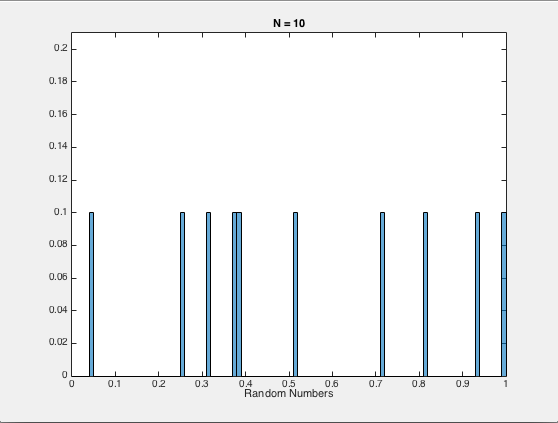
Now we will analyze the variance computed from the equation on the right from the same sets of random numbers. This is shown below in Figure 2.



*Figure 2: Plot of the Error of the Variance*

Again, we see that the variance is large when the set of random numbers is small, and as the size of the set increases, the variance increases. This means that the variance of the set converges to the expected value of .

Now we will analyze the pdfs of 10 random numbers, 1000 random numbers, and 1,000,000 random numbers.

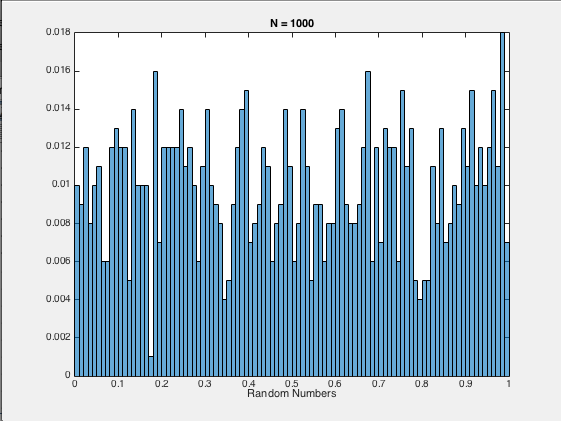


*Figure 3: PDF of 10 Random Numbers*

In Figure 3, we can see that for only 10 random numbers, the mean is above the expected value of .5. This is evident if we implement Equation 1, which is shown below. This is the equation for the expected value of a pdf.

It is also important to look at how close this distribution is to the normal distribution. For 100 bins between the value of 0 and 1, we expect that the probability of each of those bins should be .01. For only 10 numbers, we can see that the error will be extremely high. This is because each number that occurred has an estimated probability of .1, when we know the probability should be .01.

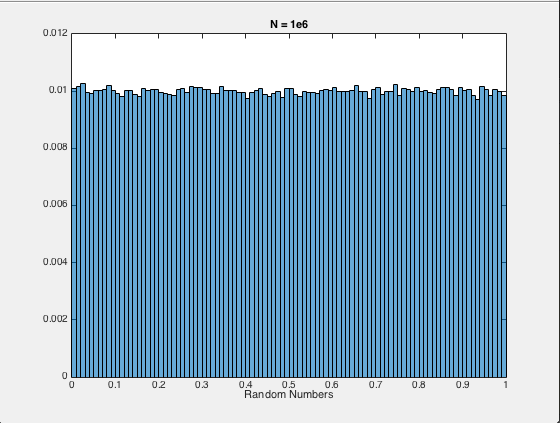
As we increase the size of the set of random numbers, we expect to see the pdf centered around .5, and the probability of each bin to be .01. The pdf of 1000 random numbers is shown below.



*Figure 4: PDF of 1000 Random Numbers*

In Figure 4, we see that the pdf is much closer to the expected values discussed above. The probabilities of each bin is closer to the expected value of .01, and the data seems to be more centered around .5.

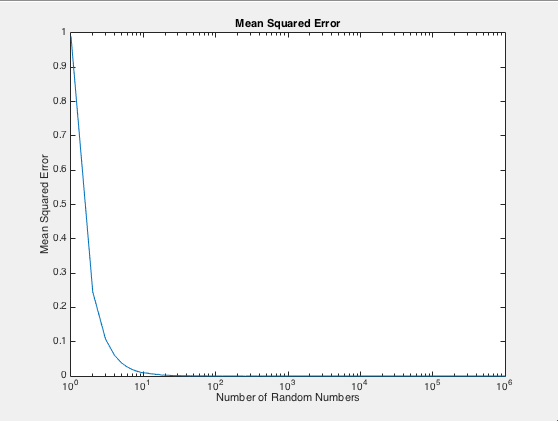
Now we will look at the pdf of 1,000,000 random numbers.



*Figure 5: Plot of the PDF of 1,000,000 Random Numbers*

This is evidence that as the size of the set of random numbers increases, the statistics of the random numbers converge to their expected values. We see the probability of each bin is almost exactly .01, and the data is almost perfectly symmetrical about .5.

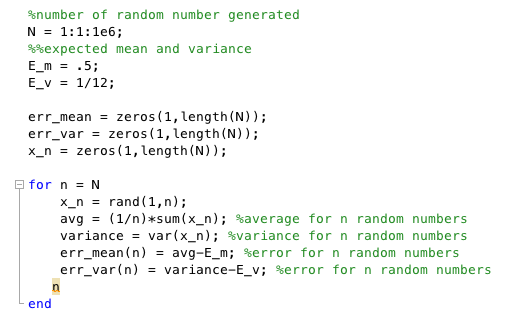
Now we will look at the plot of the mean squared error, which tells us how well the pdf fits a normal distribution.



*Figure 6: Mean Squared Error*

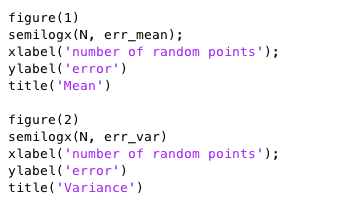
This plot supports all of the conclusions drawn from our previous results. When the number of random points is small, the data does not resemble a normal distribution, but as we increase the number of random numbers, the error between our distribution and a normal distribution goes to 0.

# MATLAB Code

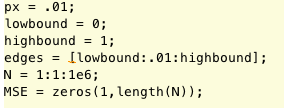


*Figure 7: Computing the Error of the Mean and Variance*

In the code above, we define the expected mean and the expected variance. Then we define our vector x(n), as the set of n random variables where n ranges from [1,1E6]. Then each time a set of random numbers is computed, we compute the mean and variance of the set, and the error between these two values. The errors are then stored in a vector that will be plotted later.

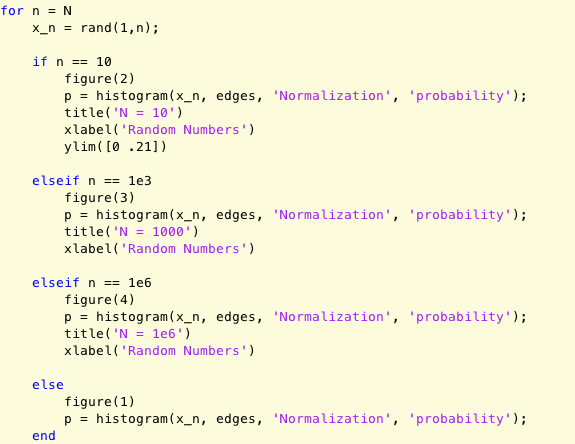


*Figure 8: Plotting the Error*

**

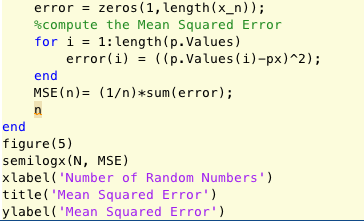
*Figure 9: Defining Constants*

In this piece of code we are defining the edges of our bins, and the expected probability of each bin.



*Figure 10: Computing the PDFs*

In the code above, we generate 1-1,000,000 random numbers, and then compute the pdf of each set. The if statements are there to keep the plot of the pdf of 10, 1000, and 1,000,000 random numbers. This is necessary because the Matlab macro “histogram” outputs a histogram each time it is called.



*Figure 11: Computing the Mean Squared Error and Plotting*

In the code above, we are computing the mean squared error for each of the pdfs computed above. The mean squared error is stored in a vector that is plotted after all iterations of N have been completed.

# Conclusions

In this assignment it is clear that the statistics of random variables are more unpredictable in the short term, but if enough trials are run, the statistics will converge to their expected values. If we compare Figures 3-5 to Figures 1, 2 and 6, we see this behavior. When n is small, i.e. 10, the variance, mean and mean squared error is extremely high. As we increase the size of the set, the data converges to the expected statistics.

I have just had an experience with this type of statistic in the casino. I was playing blackjack and one of the guys at the table was winning every hand. He got to be around $1500 dollars up. As he continued playing he began giving up everything he had just won. As blackjack is a game that has close to 1:1 odds, I knew the house would eventually get back up on him, which it did. After seeing that I took my $300 in winnings and walked away. This was a great example of short-term and long-term random statistics.