Computer Assignment (CA) No. 6:
Mean and Variance Revisited

Truc Le

ECE 3512: Signals – Continuous and Discrete

Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 1912

# Problem Statement

The overall purpose of this assignment was to find the error between the measured values of a uniform distribution and the expected values. We would be calculating the error in the mean, variance, and the mean-squared error of the PDF’s that varied in the number of samples which from the range of [1,106].

# Approach and Results

$$μ= \frac{b-a}{2}= \frac{1}{2}=0.5 σ\_{x}^{2}=\frac{(b-a)^{2}}{12}= \frac{1}{12}=0.0833$$

Above are the equations for the (Mean & Variance) of a uniform distribution.

In assignment 1, I would have to write function in Matlab code to perform the calculations and return the results which would be stored in vectors and then it’s plotted. Figure 1 shows the error between the expected values and the measured values in both the mean and variance. The expected value for the mean is 0.5, and for the variance is 0.0833. The errors could be found by subtracting the expected value from the measured values shown below Figure 1.


Figure 1

Figure 2 to Figure 4 below shows the pdf’s of a three random signals that are uniformly distributed. Figure 2 is the pdf of a signal that has 10 points, Figure 3 is the pdf of a signal that has 1,000 points, and Figure 4 shows the pdf of a signal that has 1,000,000 points. Each pdf is represented through a histogram and utilizes a bin size of 0.01 to generate 100 bins. And Figure 5 shows the mean-squared error, MSE of each signal.


Figure 2: PDF of a signal that has 10 points


Figure 3: PDF of a signal that has 1,000 points


Figure 4: PDF of a signal that has 1,000,000 points


Figure 5: MSE of each signal

# MATLAB Code

clear all; close all; clc;

%generate random numbers and set up vectors to be used

N = randi([1,10^6],1,10^3);

x\_mu = zeros(1,length(N));

x\_sigmaX = zeros(1,length(N));

sigmaX = 1/12;

mu = 0.5;

%Find the mean and variance calculated using different points

%from N. Store in another vector

for i = 1:length(N)

 [N\_mu, N\_sigmaX] = get\_Mean\_and\_Var(N(i));

 x\_mu(i) = N\_mu;

 x\_sigmaX(i) = N\_sigmaX;

end

%find the error

MuE = x\_mu - mu;

SigmaXE = x\_sigmaX - sigmaX;

%plot everything using a scatter plot

figure(1);

subplot(2,1,1);

semilogx(N,MuE,'k.');

title('Error in Mean vs N Points');

xlabel('Number of points used in calculation');

ylabel('Mean Error');

subplot(2,1,2);

semilogx(N,SigmaXE,'r.');

title('Error in Variance vs N Points');

xlabel('Number of points used in calculations');

ylabel('Variance Error');

%Find MSE using vector N2

N2 = [10 1000 10^6];

MSEn = zeros(1,length(N2));

 k = 2;

for i = 1:length(N2)

 [MSEn(i),k] = meanSquaredE(N2(i),k);

end

%Plot MSE vs points used

figure(k);

semilogx(N2,MSEn);

title('Mean-Squared Error');

ylabel('MSE');

xlabel('Number of Points Used');

- The above code would create a vector of 1000 random sizes to be used in order to create the random signals. The sizes were passed into a function one at a time, which returns the variance and mean for the random signal of that size. Each mean and variance is stored in another vector. The errors is then calculated from these vectors and plotted. The same methodology is used for the second part, but the function this time estimates the pdf’s through histograms and calculates and returns the MSE, which is then stored in a vector. The variable, k, is used for defining figures inside the function.

# Conclusions

To sum up, in part 1 of the assignment, as the number of samples in a signal increase, the error in both the mean and variance would approached zero, shown in Figure 1. For the next problem, we see that with a low number of samples, not every bin is filled. So if there are less samples than bins, there are not enough samples to fill each and every bin. As the number of samples increase, two things can be observed. One is that all the bins become filled, the other is that the pdf will begin to approach the expected value, which in the uniform case is that each value for the random variable has approximately the same probability. This convergence to the distribution makes sense because as each sample is taken, its probability is based on the pdf of the distribution. So as more samples are taken, the more the pdf will resemble the distribution.

For the mean-squared error, MSE, we noticed that the error falls as the number of samples increases. This agrees with the result found in the first problem, where increasing the number of samples minimizes the error. Since the MSE calculates the average error squared, it is relative to the error itself, and as the error decreases, so the mean squared error would also tends to decrease.