Joe DeMarco

ECE 3522: Stocastics, CA06

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

# Problem Statement

For part one of this assignment I need generate a set of random numbers in the range of 0 to 1. I then need to find the actual mean and variance of the set. Then I can calculate a running mean and variance and compare this plot to the plot of the actual mean.

For part two I will need to plot the histogram of the random set, normalize by the number of samples, and estimate a pdf for the set. I will then compare the estimated pdf to the actual pdf.

# Approach and Results

The red line is the actual mean (sum/set size). The blue line is a running mean.

for 1:n

running mean(n) = sum(0:n)/n

#

The same approach is used for the variance plot.



The cyan bar graph is a histogram of the set of random values divided by the number of values. The blue line is a uniform distribution plot. Which is just a line at 1/sample size.



# MATLAB Code

clc;clear;

lim = 10^4;

xn=rand(1,lim);

ux=(1/lim)\*sum(xn);%actual mean

vx=(1/lim)\*(sum((xn-ux).^2));%actual variance

uvect=ones(size(xn)).\*ux;

vvect=ones(size(xn)).\*vx;

urun=0\*xn;

vrun=0\*xn;

vsum=0;

xsum=0;

for i=1:lim

 xsum=xsum+xn(i);

 urun(i)=(1/i)\*(xsum);%running mean

 vsum=vsum+(xn(i)-urun(i))^2;%running variance

 vrun(i)=(1/i)\*vsum;

end

%this is a histogram of xn

figure

[a,b]=hist(xn,100);

bar(b,a/lim,'c');

title('histogram')

s=linspace(1,lim,lim);

hold on

%here is the pdf

line([0;1] ,[1/lim;1/lim])

figure

semilogx(s,uvect,'r')

title('mean')

hold on

semilogx(s,urun)

figure

semilogx(s,vvect,'r')

title('variance')

hold on

semilogx(s,vrun)

# Conclusions

The first plot shows the actual mean (red) and the running mean (blue). This shows that as I use more of the set to calculate the mean, the running mean approaches the actual mean. It also shows that the running mean becomes reasonably close to the actual mean at 1/10th of the total set size. This means that we can expect to get a reasonable estimate of the mean with a small sample of the set.

The same thing is true for the variance (second plot), except that it appears that the running variance converges even faster than the mean. 1/100th of the set seems adequate.

The histogram shows pretty much what I expected, which was a uniform distribution. I plotted the actual uniform distribution on this plot also (it is the blue line at the bottom). It seems like the histogram estimate is very poor. I think this could be corrected a bit by increasing the number of bins, but I think there will always be empty bins, so the estimate will always be off.