

Zainab H. AlAli
SPSS 3522
Joseph Picone

Computer Assignment (2) REGRESSION AND HISTOGRAMS

# Image of butterfliesIntroduction:

In this computer assignment we are going to learn more about basic production functions with regression and histograms. The goal is to understand the graph function and what they mean. This computer assignment is based on the first computer assignment where the results are based on CA1. Moreover, this assignment is based on google stock data which has a frame size of 1 to 7.

# Problem Statement:

The tasks to be accomplished are:

1. Building on your results from CA #1, and using the Google stock data with a frame size of 1 and a window size of 7, plot the signal data and overlay a plot of the mean value. Next, research and implement a linear regression function. Compute one regression model for the entire data set and plot it over the waveform. How does this compare to the plot of the mean values?

Later in the course we will discuss regression models. You might have used these in Excel when analyzing data.

1. For the speech signal, plot a histogram of the amplitude of the signal. Use a range of +/-32767.0 and a bin size of 10. Plot the number of samples in each bin as a function of the middle value of the bin. Your first bin should span the range [-5,5]. Your second bin [5, 15] and so forth. Do the same for negative amplitude values (e.g., [-15, -5]). Normalize the number of samples in each bin by the total number of samples in the file.

In a separate plot, plot the cumulative distribution.

Describe these plots. Do they make sense? What do they imply about the speech signal?

## Results:

The original stock system:

 Figure (1)

Figure (2)

Figure (3)

Figure (4)

Figure (5)





### Conclusion & Appendex:

This assignment goal was to understand the matlab plot, and compute the mean value of the original signal, where we simply use the mean function and mean contour to find the involve functions and their sizes. Moreover, these information can helps us to demonstrate the array and the created arrays to create a plot that shows the linearity of the mean values of the closing stock, and outputs the coefficients of desired line’s polynomial as an array with respect to the sample space array and outputs the value of the polynomial at the different samples. However, I used the histogram function to show different ranges that represent the google stock.

function rms=SPSS\_ca2\_Pb1\_stock\_rms\_mean\_regress

clear;close all;clc;

M = [ 1 ];

N = [ 7 ];

google\_set = 'google\_v00.xlsx';

sig = xlsread(google\_set,'E2:EF2617');

rms = zeros(length(M), length(N), length(sig));

for m = 1:length(M)

 h1 = figure(...

 'name', 'RMS Frame 1 Window 7 Plot with Mean Value and Regression'...

 , 'numbertitle', 'off');

 for n = 1:length(N)

 rms(m,n,:) = compute\_rms(sig, M(m), N(n));

 figure(h1);

 str = sprintf(' RMS Plot with frame: %d, window: %d Also Mean and Regression', M(m), N(n));

 plot(squeeze(rms(m,n,:)));

 title(str);

 xlabel('time');

 ylabel('rms');

 end

 figure(2);

 plot(sig);

 title('Original Closing Data Signal');

 xlabel('time');

 ylabel('amplitude');

end

rms\_out = squeeze(rms);

google\_set = 'google\_v00.xlsx';

data = xlsread(google\_set,'E2:EF2617');

dspace = linspace(1,length(data),length(data));

x = dspace;

y(x) = data(x);

figure(1)

hold on

m = mean(y);

plot( x, ones(size(x))\*m, 'g--');

p = polyfit(x,y(x),1);

f1 = polyval(p,x);

hold on

plot(x,f1,'r--');

end

function rms\_full = compute\_rms(sig\_a, fdur\_a, wdur\_a)

%

sig\_wbuf = zeros(1, wdur\_a);

num\_samples = length(sig\_a);

num\_frames = 1+round(num\_samples / fdur\_a);

rms\_full = zeros(length(sig\_a),1);

%

for i = 1:num\_frames

 n\_center = (i - 1) \* fdur\_a + (fdur\_a / 2);

 n\_left = n\_center - (wdur\_a / 2);

 n\_right = n\_left + wdur\_a - 1;

 if( (n\_left < 0) || (n\_right > num\_samples) )

 sig\_wbuf = zeros(1, wdur\_a);

 end

 %

 for j = 1:wdur\_a

 index = n\_left + (j - 1);

 if ((index > 0) && (index <= num\_samples))

 sig\_wbuf(j) = sig\_a(index);

 end

 end

 %

 rms = sqrt( (1 / wdur\_a) \* sum(sig\_wbuf.^2));

 %

 for j = 1:fdur\_a

 index = n\_center + (j - 1) - (fdur\_a/2);

 if ((index > 0) && (index <= num\_samples))

 rms\_full(index) = rms;

 end

 end

end

end

clear;clf;close all;clc;

fp=fopen('rec\_01\_speech.raw','r');

%

sig = fread(fp,inf,'int16');

fclose(fp);

%histograms:

histogram(sig,-5:5)

title('Histogram of "rec 01 speech", -5:5')

for i=5: +10: 55

 figure

 histogram( sig, i:i+10 );

 title('Histogram of "rec 01 speech.raw"');

 str = sprintf('Histogram of "rec 01 speech", %d:%d',i,i+10);

 title(str);

end

for i=5: +10: 55

 figure

 histogram( sig, -i-10:-i );

 str = sprintf('Histogram of "rec 01 speech", %d:%d',-i-10,-i);

 title(str);

end

histogram(sig,-15000:15000)

%Cumulative Distribution

figure;

cdfplot(sig);

This assignment has demonstrated the importance of using different distribution fits appropriately for different types of data sets.