Engr 421 – HW #3 Report

Firstly, I divided the data set to train and test, train contains first 150 data points and test contains the remaining 122 data points. In addition, I seperated the x and y values. I got number of test data as N_test. I set bin_width and origin parameters as 0.37 and 1.5 which are specified in the pdf. I calculated minimum and maximum values and using these values I calculated the data_interval. Using origin, bin_width, minimum, and maximum values I determined the left and right borders. Then, I applied the following formula to calculate p_hat for regressogram:

$$\hat{g}(x) = \frac{\sum_{t=1}^{N} b(x, x^t) r^t}{\sum_{t=1}^{N} b(x, x^t)}$$

where

$$b(x, x^t) = \begin{cases} 1 & \text{if } x^t \text{ is the same bin with } x \\ 0 & \text{otherwise} \end{cases}$$

Once I obtained the p_hat, I plotted the graph of regressogram. I calculated the rmse by using the formula given in the pdf.
Secondly, I applied the following formula to calculate p_hat for running mean smoother:

$$\hat{g}(x) = \frac{\sum_{t=1}^{N} w\left(\frac{x-x^t}{h}\right) r^t}{\sum_{t=1}^{N} w\left(\frac{x-x^t}{h}\right)}$$

where

$$w(u) = \begin{cases} 1 & \text{if } |u| < 1 \\ 0 & \text{otherwise} \end{cases}$$

Once I obtained the p_hat, I plotted the graph of running mean smoother. Again, I calculated the rmse by using the formula given in the pdf.
Lastly, I applied the following formula to calculate p_hat for kernel smoother:

$$\hat{g}(x) = \frac{\sum_{t} K\left(\frac{x-x^t}{h}\right) r^t}{\sum_{t} K\left(\frac{x-x^t}{h}\right)}$$

Once I obtained the p_hat, I plotted the graph of kernel smoother. Again, I calculated the rmse by using the formula given in the pdf.

Formulas above are taken from the textbook.