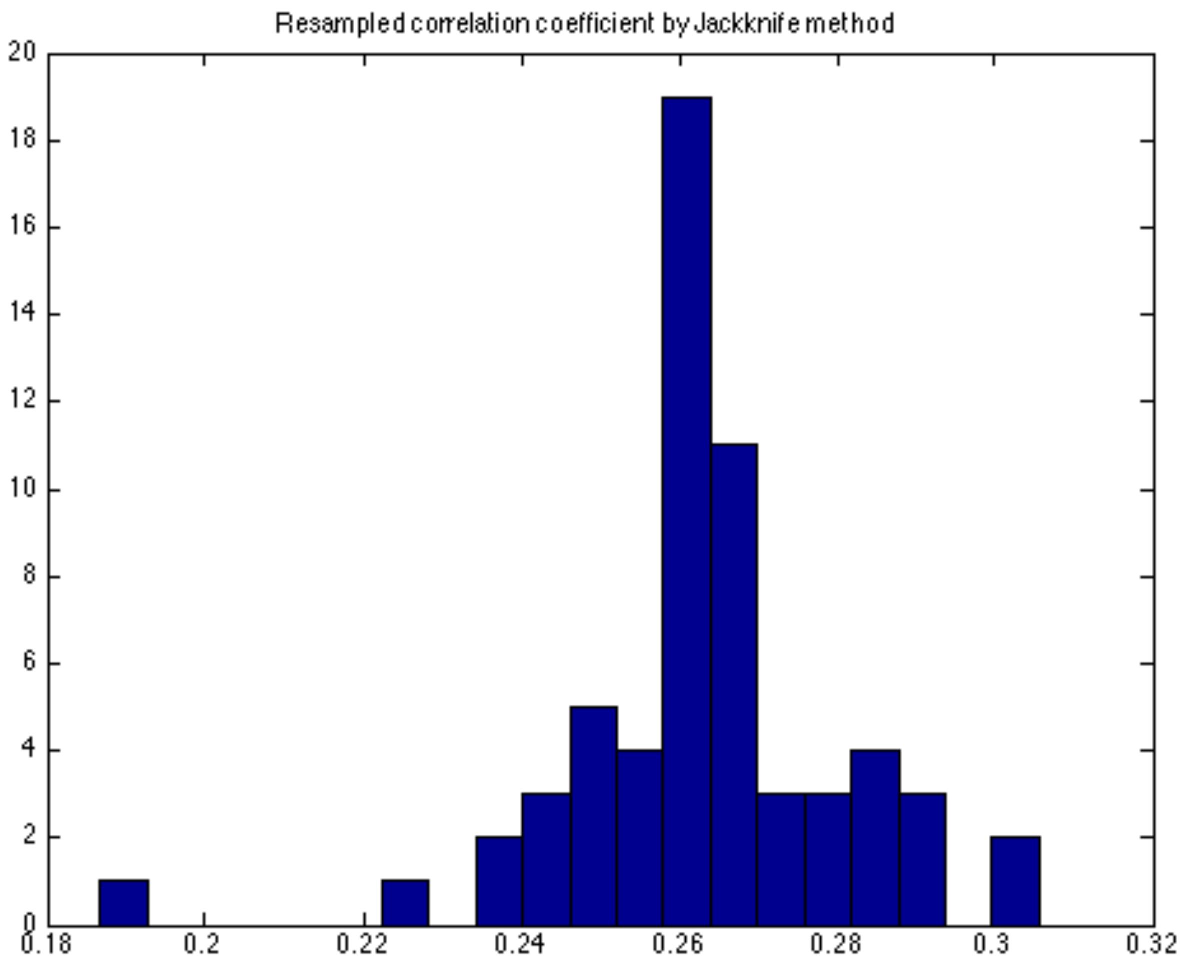
Namrata Kolla
March 30, 2015
Homework 8

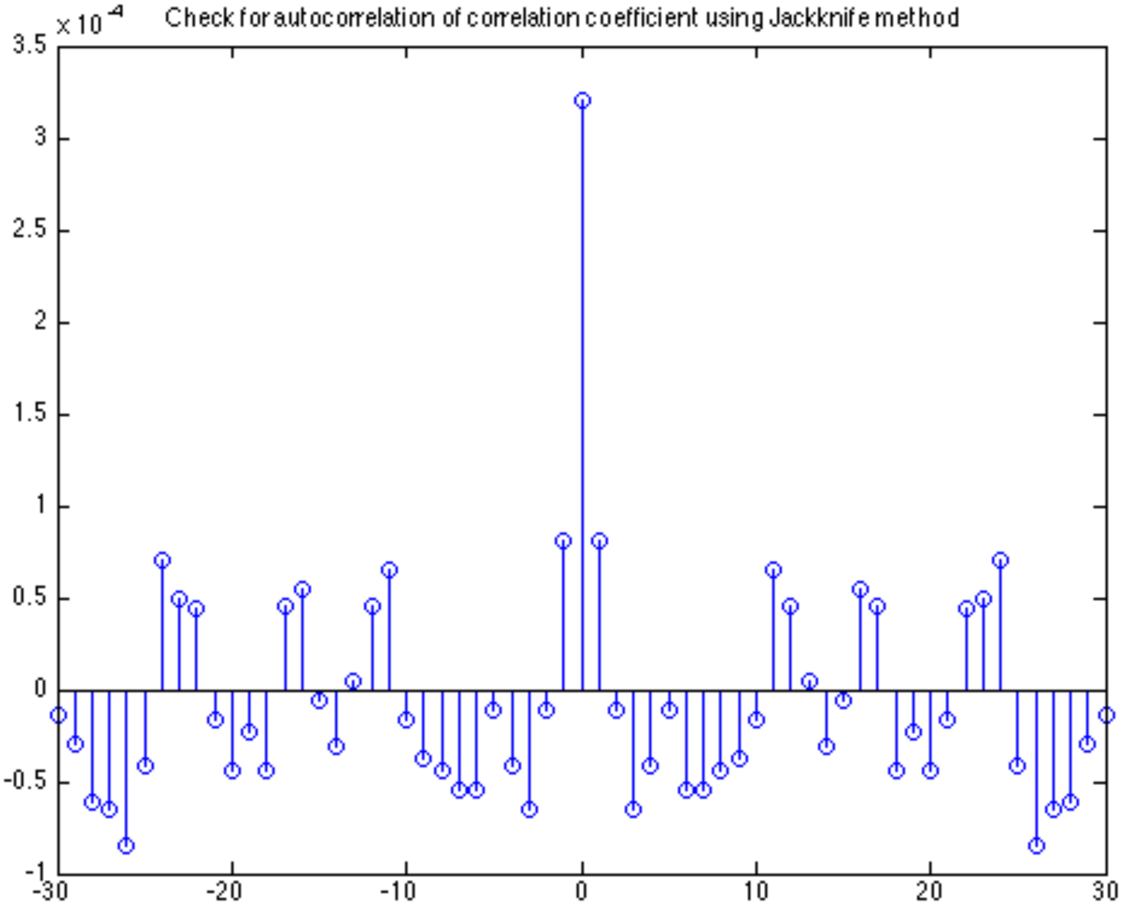
Measurements of ozone and carbon monoxide from Atlanta

Problem 1. Uncertainties of statistical parameters

#1: Jackknife method for correlation coefficient
Mean: 0.2631
Standard deviation: 0.0180



95% confidence interval of resampled parameter: [0.2255, 0.3019]
The Jackknife method assumes that observations are independent of each other. We can check if this is true by performing an autocorrelation analysis.

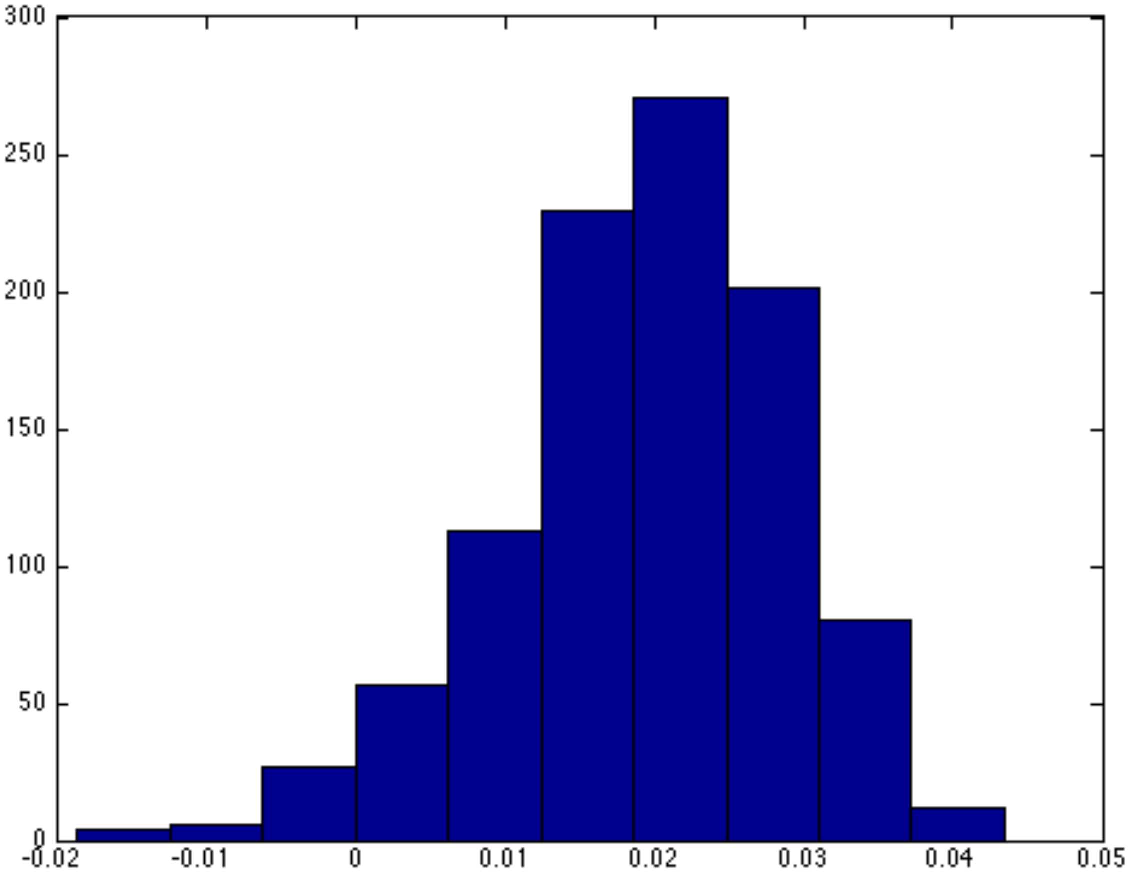


From this plot (lag = 30 days), the correlation coefficient does seem to be autocorrelated because the data is symmetrical and not random.

#2: Bootstrap method for least-squares regression slope

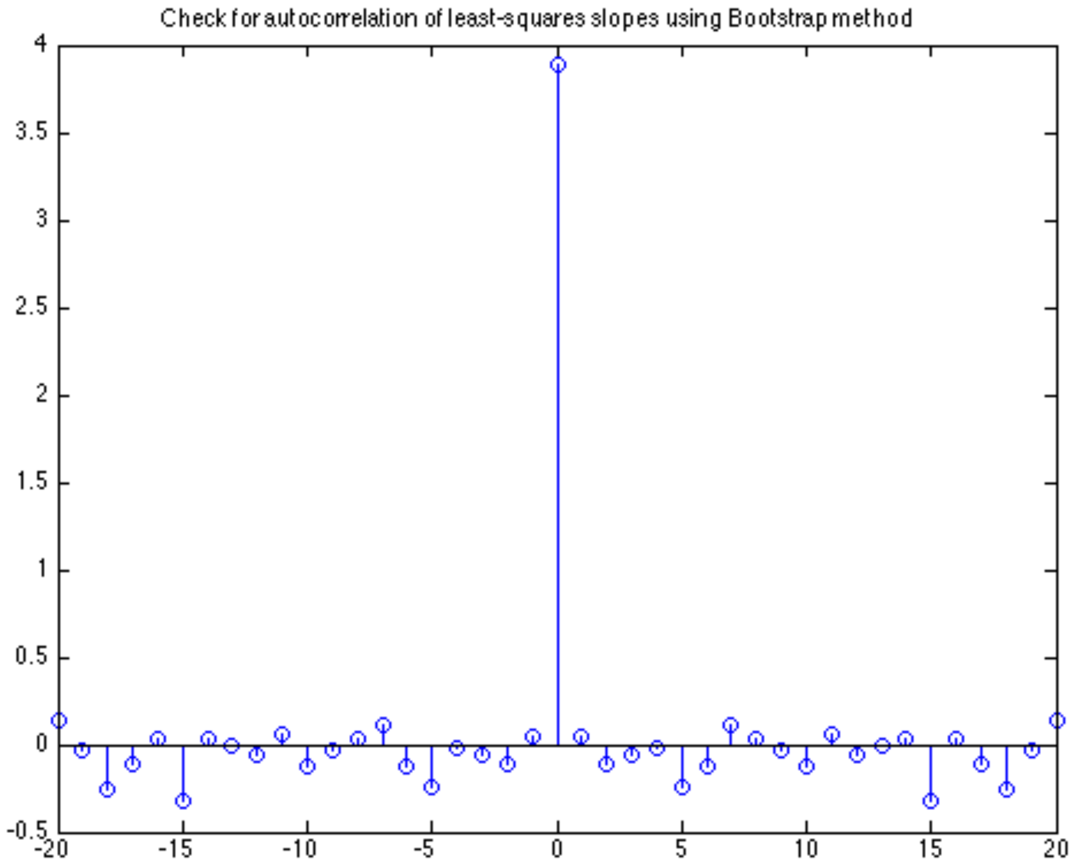
Mean: 0.0192

Standard deviation: 0.0095



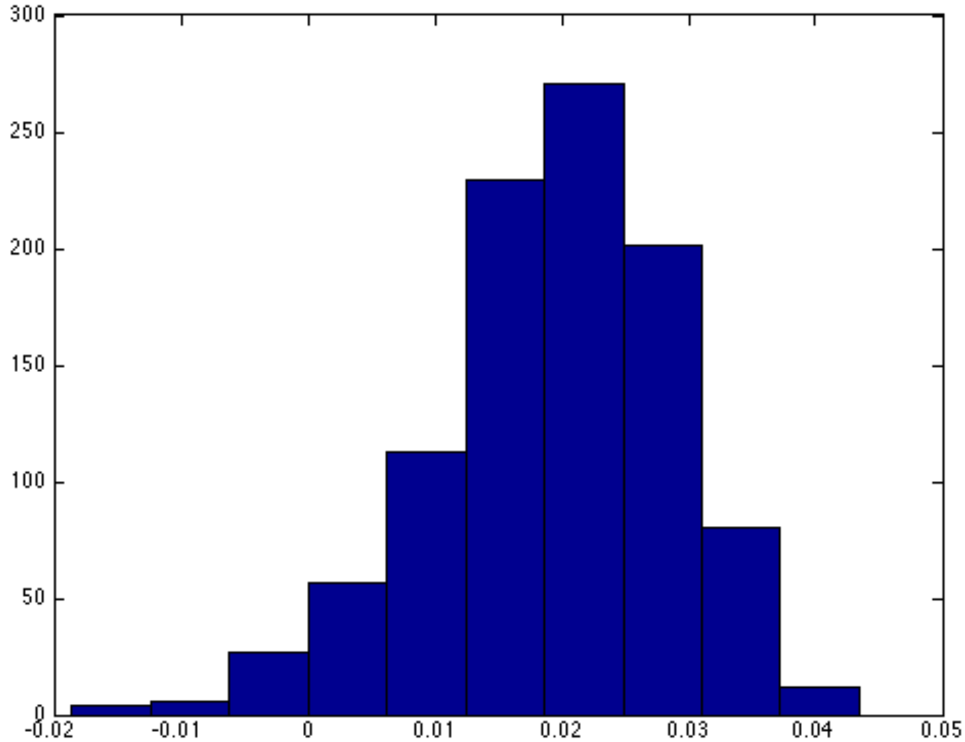
95% confidence interval of resampled parameter: [-0.0022 0.0357]

The Bootstrap method assumes that observations are independent of each other. We can check if this is true by performing an autocorrelation analysis.



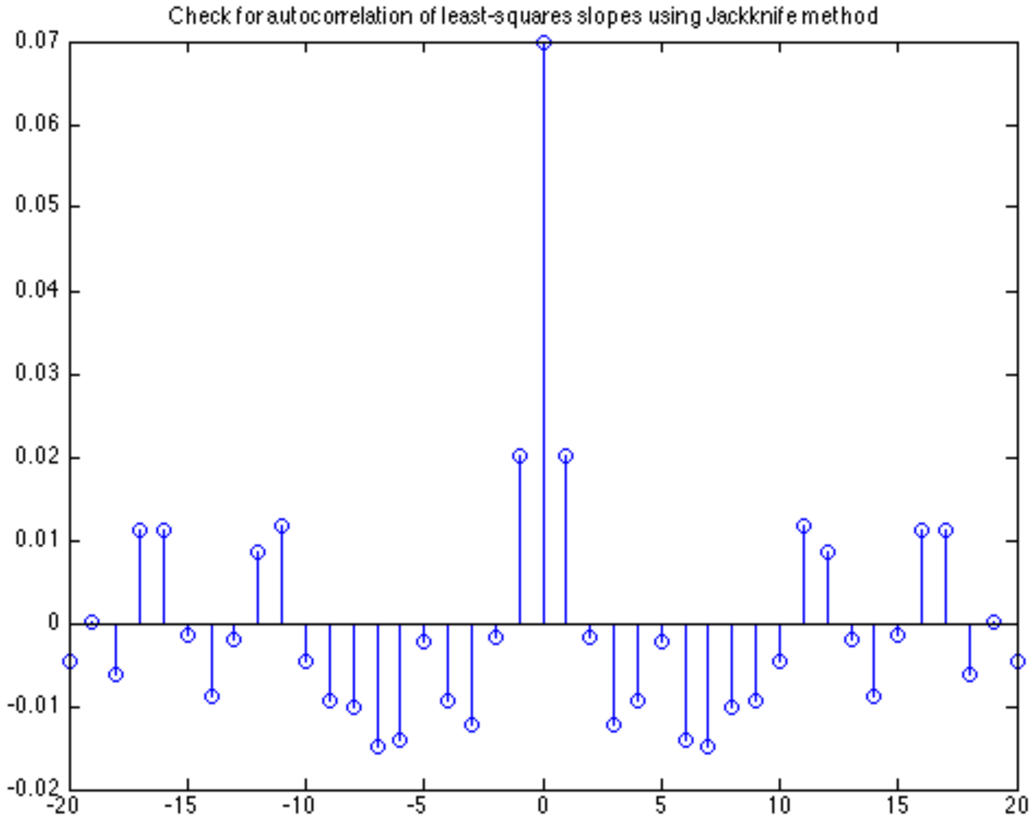
From this plot (lag = 30 days), the slope values do seem to be auto-correlated because the residuals are symmetrical and not random.

#3: Jackknife method for least-squares regression slope
Mean: 3.3701
Standard deviation: 0.2664



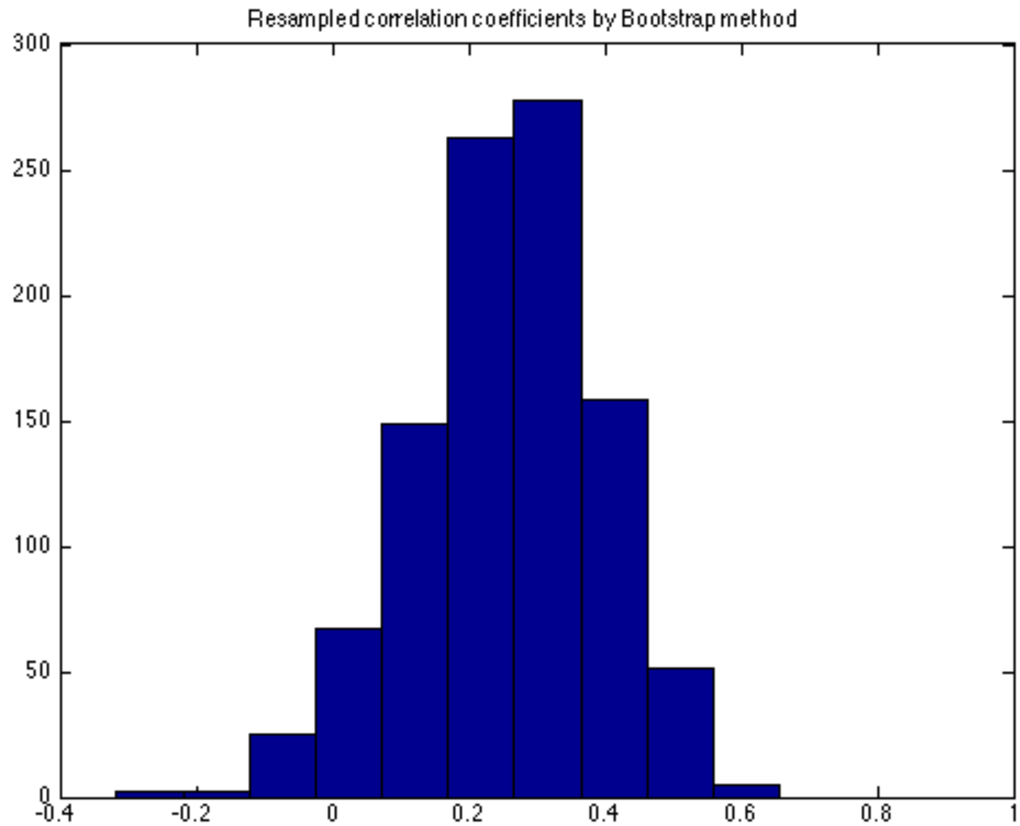
95% confidence interval of resampled parameter: [0.0184, 0.0233]
The Jackknife method assumes that observations are independent of each other. We can check if this is true by performing an autocorrelation analysis.

Autocorrelation graph (same as #1)



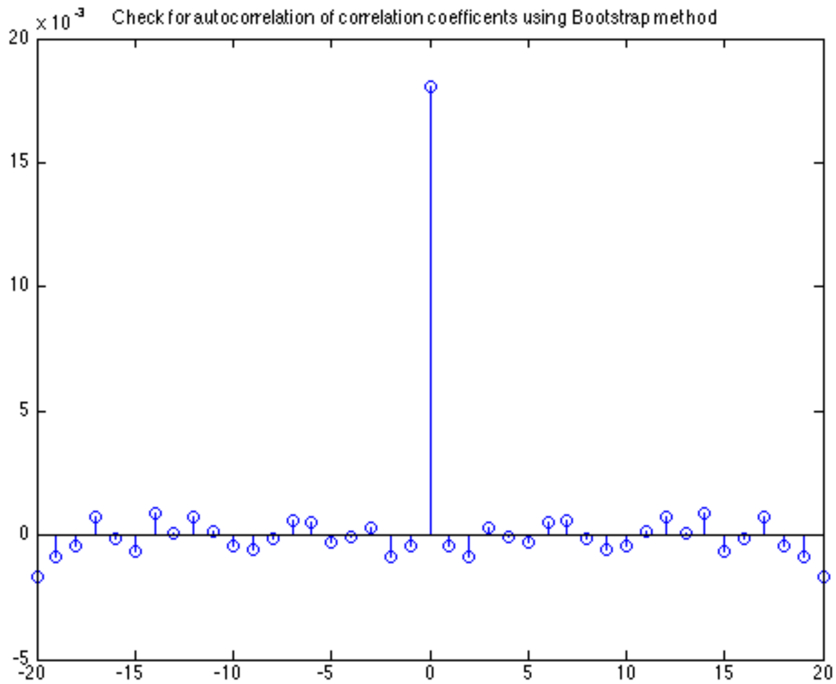
From this plot (lag = 30 days), the slope values do seem to be auto-correlated because the residuals are symmetrical and not random.

#4: Bootstrap method for correlation coefficient
Mean: 0.2600
Standard deviation: 0.1321



95% confidence interval of resampled parameter: [-0.0206, 0.5217]
The Bootstrap method assumes that observations are independent of each other. We can check if this is true by performing an autocorrelation analysis.

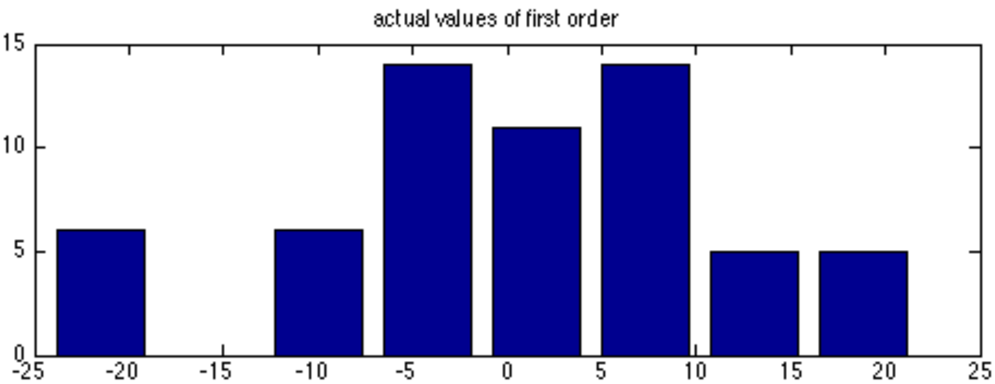
Autocorrelation graph (same as #2)

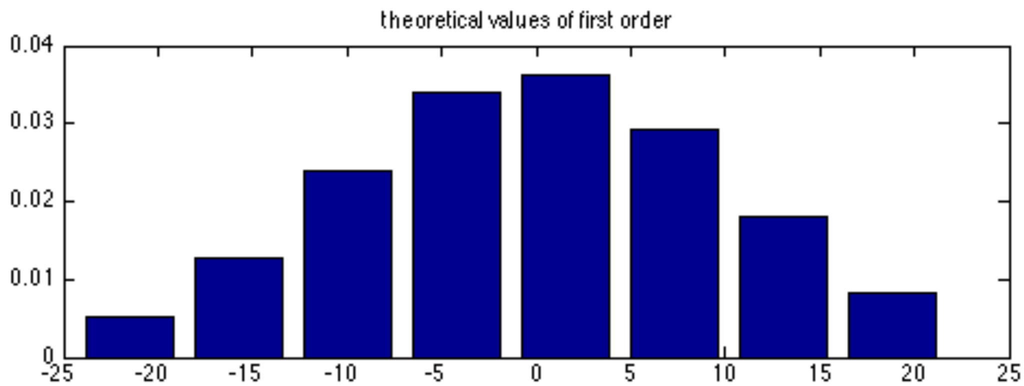


From this plot (lag = 30 days), the correlation coefficient values do seem to be auto-correlated because the residuals are symmetrical and not random.

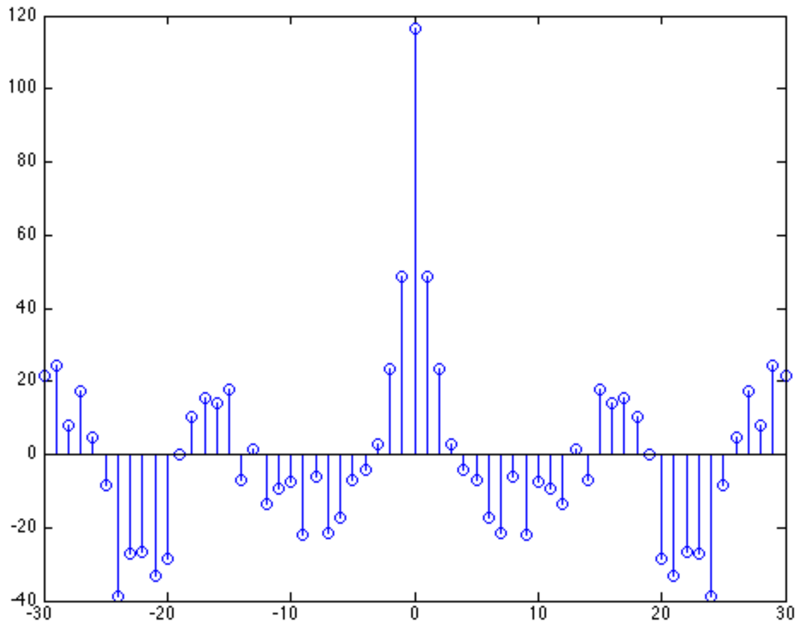
Problem 2. Fitting errors – Is a 1st or 3rd order polynomial fit better for the data?

<u>#1: First-order polynomial fit</u>
Checking if residuals are normally distributed:
Calculated chi-squared value is 17.8021; critical chi-squared value is 11.0705
Calculated value < critical value means the null hypothesis cannot be rejected
Thus, residuals are normally distributed

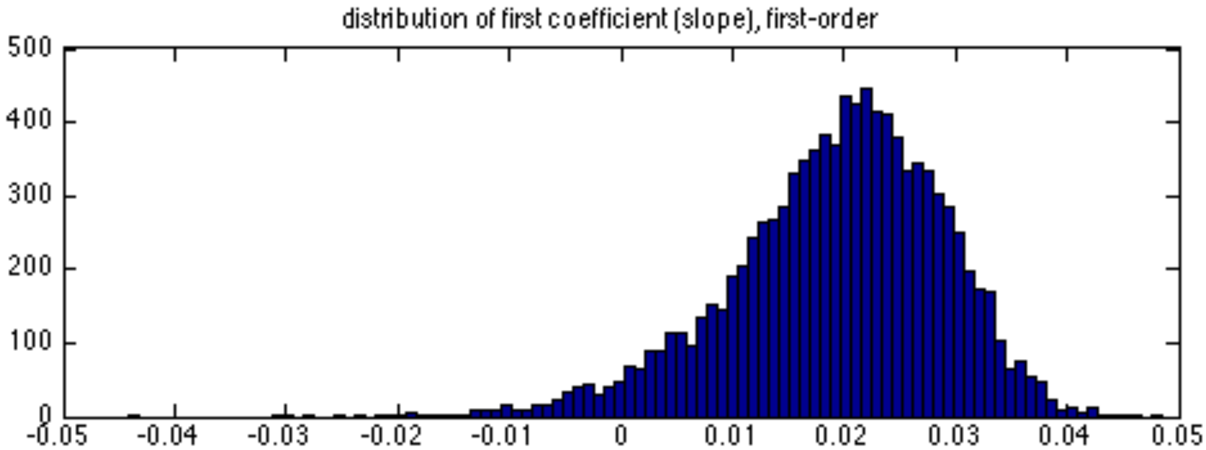


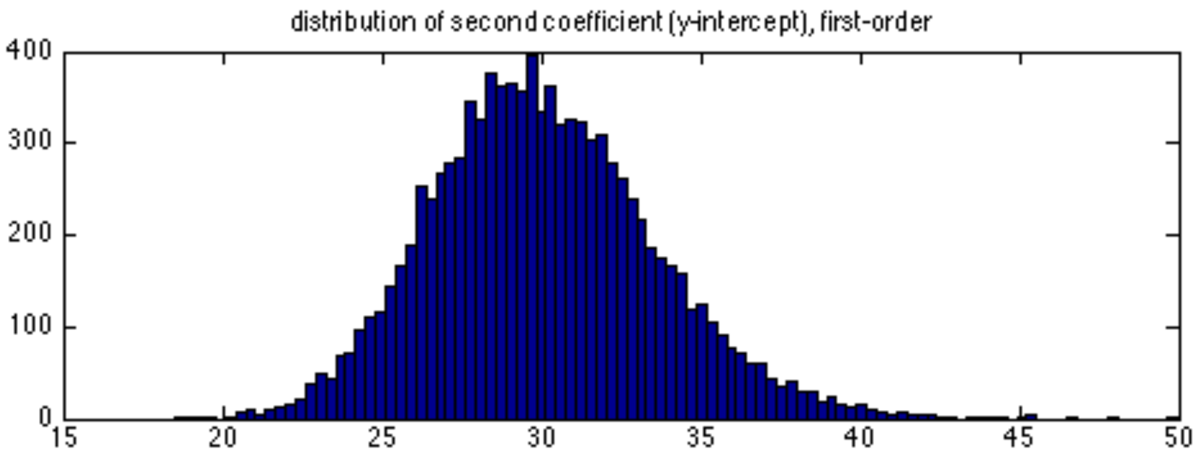


Autocorrelation analysis of residuals:

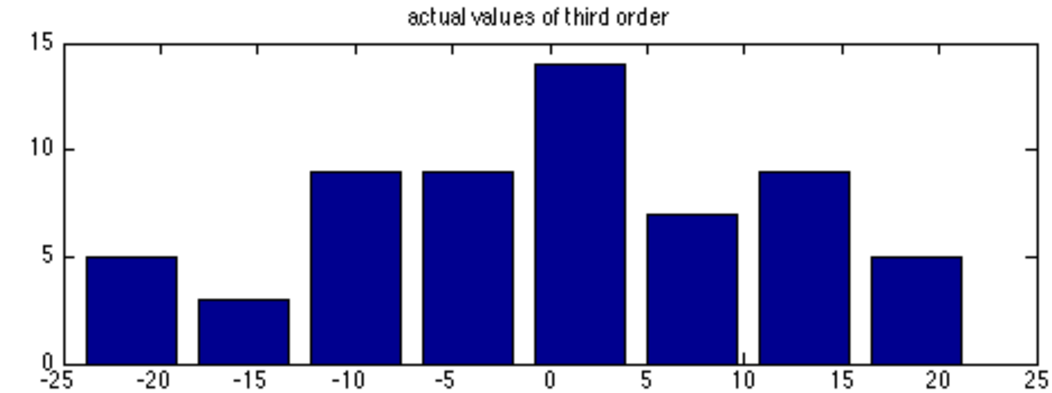


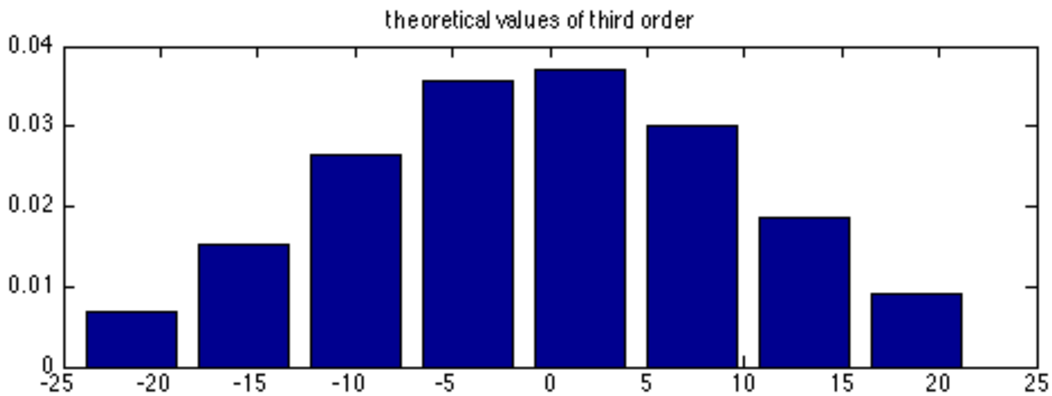
From this plot (lag = 30 days), the residuals for the first-order fit do seem to be auto-correlated because the results are symmetrical and not random.



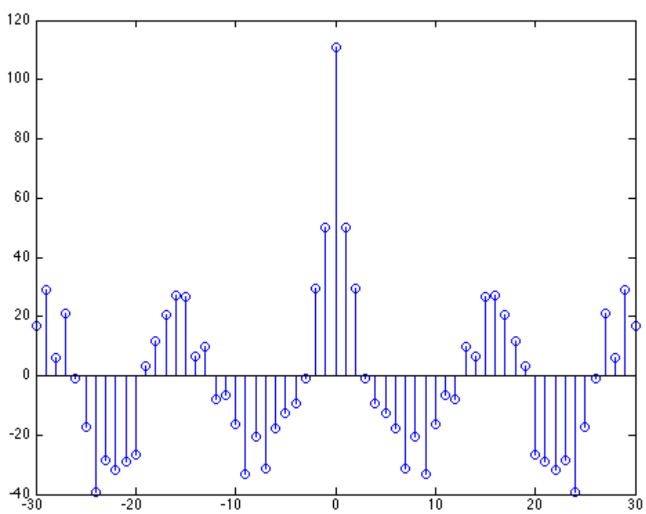


<u>#3: Third-order polynomial fit</u>
Checking if residuals are normally distributed:
Calculated chi-squared value is 22.0441; critical chi-squared value is 27.5871
Calculated value < critical value means the null hypothesis cannot be rejected
Thus, residuals are normally distributed

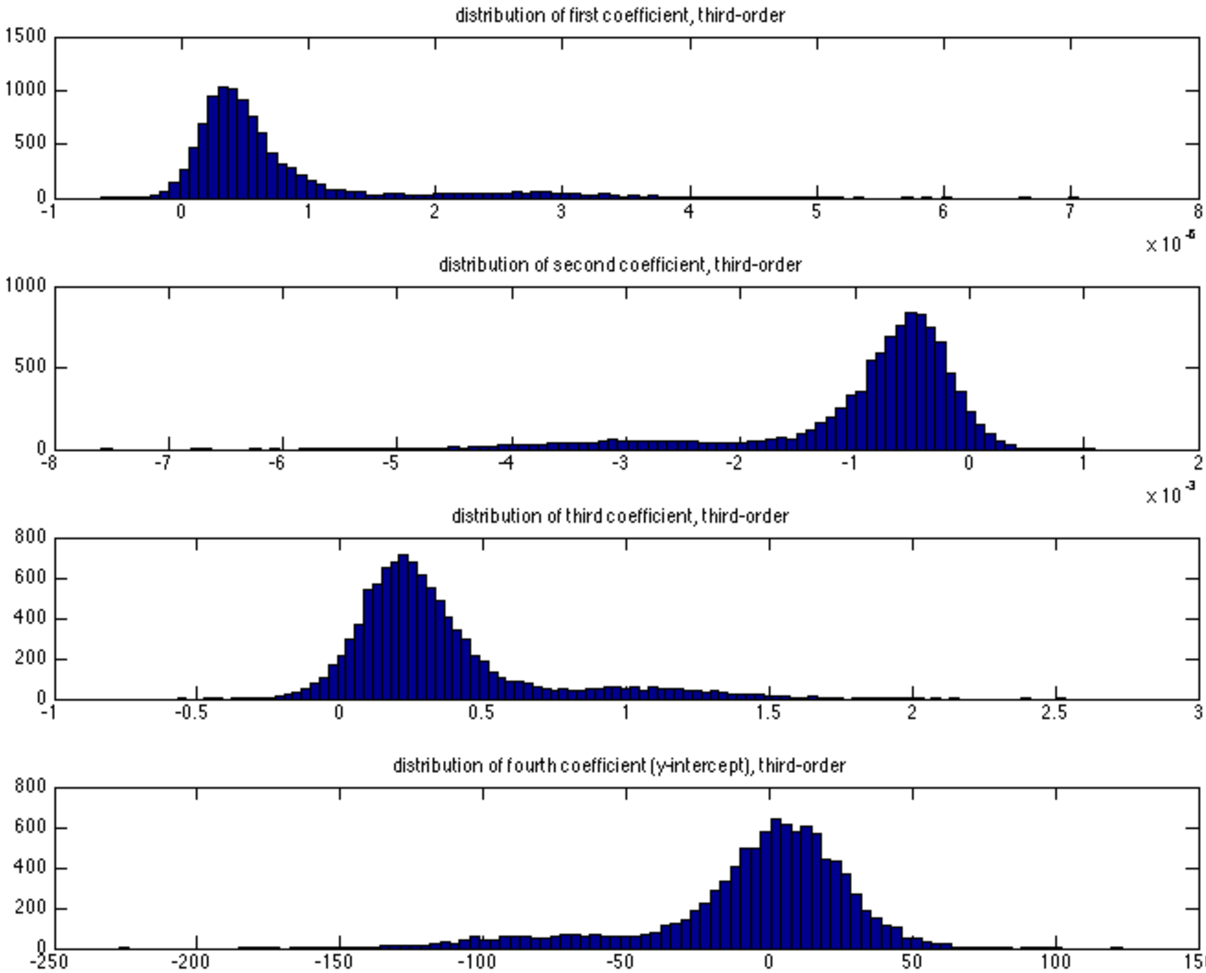




Autocorrelation analysis of residuals:



From this plot (lag = 20 days), the residuals for the third-order fit do seem to be auto-correlated because the residuals are symmetrical and not random.



The long tails for the bar graphs of the coefficients suggest that there are influential outlier points for the third-order polynomial.