ASTR 600: Problem Set 3

X	Y	yerr1	yerr2	yerr3
0	2.24	0.92	0.92	0.92
1	3.06	0.92	0.92	0.92
2	7.67	0.92	0.92	0.92
3	10.22	0.92	0.92	0.13
4	15.54	0.92	0.13	0.92

1) Use the table above for all the problems.

a) Find a linear, 1st order fit between the x and y points using equation 7.6. DO NOT USE R. Instead, make a table showing your step-by-step process. The first two columns should be the x values and the y values. The next two columns will be quantities X_i-X_{bar} and Y_i-Y_{bar}. The final two columns will be $(X_i\text{-}X_{bar})^2$ and $(X_i\text{-}X_{bar})(Y_i\text{-}Y_{bar})$.

b) Now you can use R. Use the same x and y value, but now use the uniform uncertainties from yerr1. Compare this fit to the one you got in a. What are the uncertainties on the slope and intercept?

c) Use the same x and y value, but use the uncertainties from yerr2 and then yerr2. Compare these fits to the one you got in b. How does decreasing the uncertainty of an end point compare with decreasing the uncertainty of a middle one?

d) Use Theil-Sen to find a fit for these points. (Hint: look up mblm in R.)

2) One way of determining spectral types is by taking the ratio of flux at one point in the spectrum – in this case in an absorption band – to another place in the spectrum which is often continuum. This ratio is called the spectral index. I have created a new spectral index, which needs to be converted into spectral types. The current spectral types are determined by the standard spectral index for this temperature regime. They range from -2 to 7, where 0 to 7 are M subclass (i.e. a 2 would be M2), -1 is K7, and -2 is a K5.

a) Find the best 1st order fit and the best 2nd order fit. Plot the data and both fits on the same graph. Determine which one is better by…

b) using the adjusted R^2. Note that equation 7.62 from the book should be:

$$R_a^2 = 1 - (1 - R^2)\frac{n-1}{n-p}$$

c) using the Bayesian Information Criteria. (See page 62 in the textbook.)

d) Make a normal quantile-quantile plot to examine the distribution of errors for the 2nd order fit.

The data set for this has 141 lines, so it is in an attached file.