Homework 6 Write-up

Sarah Vaughn

1 question 1

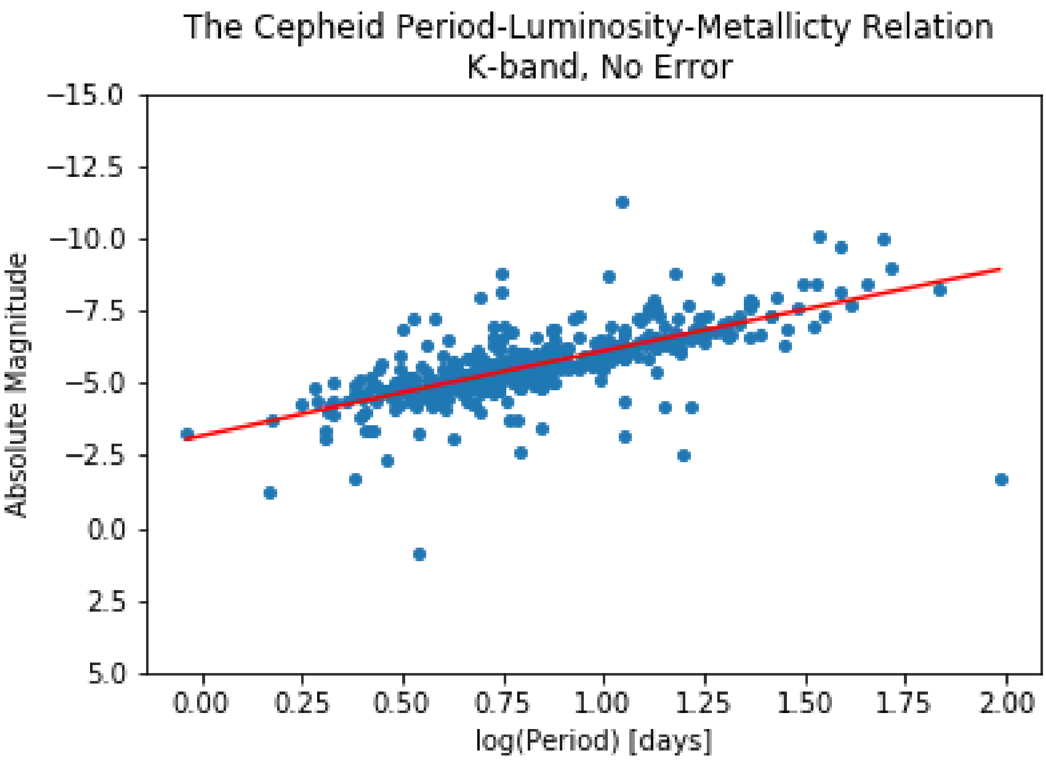
I chose the K-band Magnitude as the photometric band that I analysed and using linear Least squares to solve for the optimal values of α, β, and γ in the equation

$$M = \alpha + \beta * log(P) + \gamma * [Fe/H] \tag{1}$$

by using Matrix multiplication to solve for θ using $\theta = (X^{\dagger}X)^{-1}X^{\dagger}y$ and the Absolute Magnitude. Where θ is a matrix of the resulting values:
$\alpha = -3.26600659$, $\beta = -2.8343203$, $\gamma = -0.05965186$
The uncertainties are calculated using $\sigma^2 = (X^{\dagger}X)^{-1}$ and taking the diagonal values to get:
$\sigma_\alpha = 0.1430526, \sigma_\beta = 0.16764732, \sigma_\gamma = 0.25234996$

2 question 2

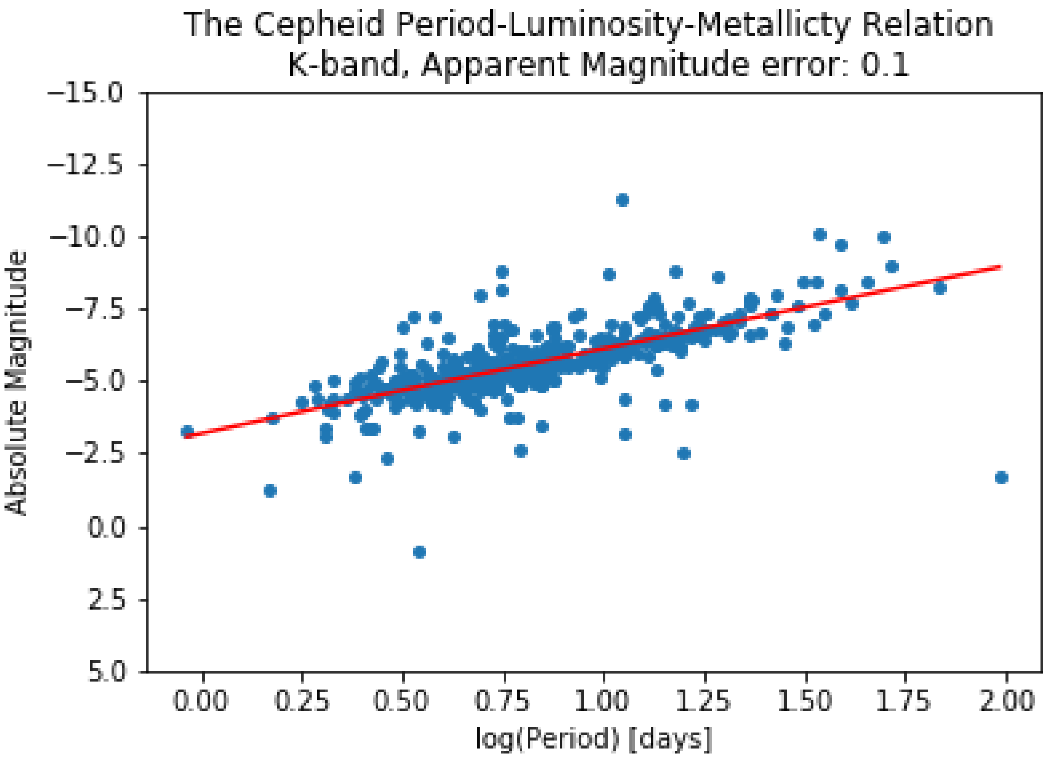
Plotting all the data points from the Cepheid text file and fitting equation 1 from above with the optimal values that were calculated in part 1 results in the red fitted line below:



3 question 3

Using the same photometric band K, and a similar process as part 1, I was able to calculate the optimal values again. This time assuming the errors on the apparent magnitudes are all $\sigma = 0.1$ mag and using

$\theta = (X^\dagger V^{-1}X)^{-1}X^\dagger V^{-1}y$ where V is the identity of the error and solving to get the values:
$\alpha = -3.26600659$ $\beta = -2.8343203$, $\gamma = -0.05965186$
These values arnt different from the first but this makes sense because the data is still the same and the optimal values for fitting the same data should stay the same. The equation to solve for the uncertainty is $\sigma^2 = (X^\dagger V^{-1}X)^{-1}$ and the resulting values are:
$\sigma_\alpha = 0.01430526$ $\sigma_\beta = 0.01676473$ $\sigma_\gamma = 0.025235$ again plotting all the data points from the Cepheid text file and fitting equation 1 from part one with these optimal values will get the fitted curve in red below:

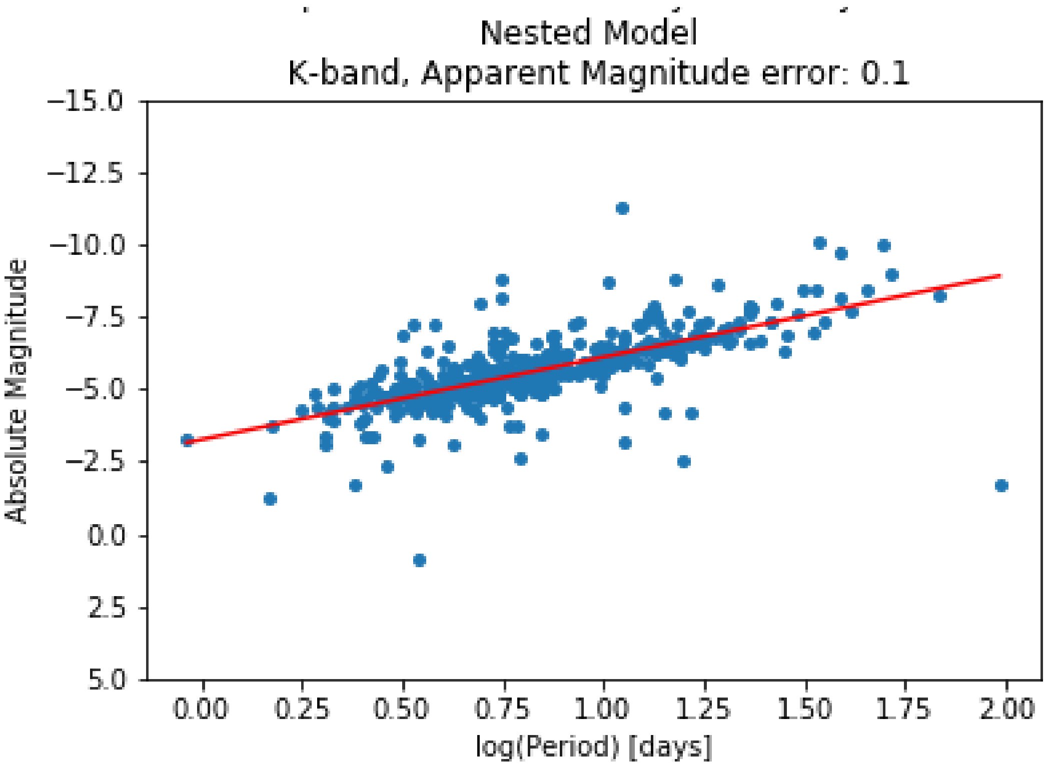


4 Bonus

This process is very similar to part 3 but the model equation doesn't include γ and the metallicity:

$$M = \alpha + \beta * log(P) \tag{2}$$

The values that I calculated to be the optimal values for equation 2 are:
$\alpha = -3.25766808$ $\beta = -2.84686528$
and the uncertainty values are:
$\sigma_\alpha = 0.01386352$ $\sigma_\beta = 0.01590258$
This is the resulting fitted plot:



Next to calculate the F statistic, I first calculated the χ^2 of both the nested model(equation 2) and the full model(equation 1) using the chisquare function from the scipy.stats package and the ν values where just the data size minus the number of parameters of each. Plugging these values into the equation given in the question I got 3.0679. to me this number seems high and when I calculated the cdf from the stats class I got 0.919. According to the notes, this would mean significant variation from just removing the last parameter and metallicity. However, by just looking at the values for α and β and their uncertainty values, they haven't changed that significantly and the F and cdf shouldn't be showing that there is such a change in the fit when not including the metallicity and the γ parameters. Even just looking at the graphs and the fitting lines on them they don't seem to have changed that significantly to me. I think that the code I wrote might have something flipped and it should show the opposite of what these numbers found.