Data Analysis

Summary

Part 1, finding the Period-Luminosity relation in cepheid variable stars.
From data on cepheid variable stars in the milky way, I obtain a model for the relationship between period and absolute magnitude of a cepheid variable star.

Part 2, finding distance to ncg4527.
Using data from many cepheid variable stars in the galaxy ncg4527, and the previously worked out relationship, work out the distance to ncg4527 and then, knowing the recession velocity of the galaxy, work out the Hubble constant given by that single galaxy.

Part 3, estimate the Hubble constant using many galaxies.
Using the data for ncg4527 and many other galaxies, estimate a value for the Hubble constant.

Part 1, finding the Period-Luminosity relation in cepheid variable stars.

The equation we are using to model the relationship in cepheid variable stars is:

$$M = \alpha \cdot \log(P) + \beta$$

where M is absolute magnitude, P is period, and α and β are parameters to be found for the model to work. On a graph of M against log(P), α is the gradient and β is the y-intercept.

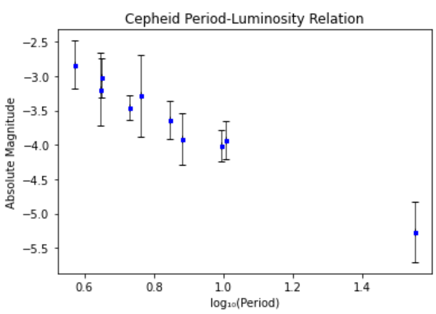

Fig 1. Cepheid's in the milky way plotted by their period's and absolute magnitudes.

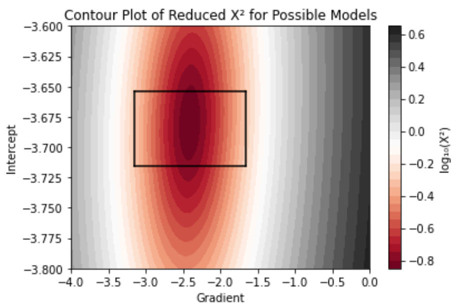

Fig 2. Reduced χ^2 graph for the gradient and y-intercept parameters with error shown.

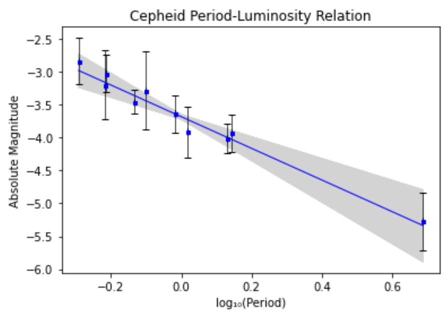

Fig 3. Cepheid's in the milky way with shifted x-axis, with model line and error shown.

Firstly, I converted parallax measurements into distance to each cepheid. Then using the equation $M = m - 5\log(d) + 5 - A$, where m is apparent magnitude, d is distance in parsecs, and A is the extinction, I found the absolute magnitude of each cepheid variable star, and it's error.

The plot in figure 1 shows that there is a negative correlation between magnitude and the log of period. In order to remove correlation between the variables, I shifted the points to the left by the mean of the log(period) of the dataset. This doesn't affect the gradient but does affect the intercept which was accounted for when taking the results.

I then ran a reduced χ^2 test shown in figure 2, to find the best fitting model for the data, and the error using the standard deviation of reduced χ^2. Plotting that back onto the original graph along with the error in both parameters gives figure 3. The goodness of fit is very good: reduced $\chi^2 = 1.15$

My results gave α to be -2.41 ± 0.7 and β to be -1.6 ± 0.03.

Some values of these are -2.43 and -1.62 from the Hubble telescope, meaning my values are sane.

Part 2, finding distance to ncg4527.

Using data for many cepheids in the galaxy ncg4527, their periods and apparent magnitudes, we should be able to tell the distance to the galaxy.

Finding the absolute magnitude of each cepheid variable star in ncg4527 and the error using the model found in the previous part of the analysis then lets us find the distance to each cepheid variable star. Using the same equation as before, this time rearranged to find d in parsecs:

$$d = 10^{\frac{-M+5-A+m}{5}}$$

We find the distance to each star and the error in each star's distance. This data is plotted in figure 4.

As you can see in figure 4, there is a single obvious outlier in the data which I have excluded from the calculations: C1-V11. This could be due to some anomoly in recording or some other type of object mistaken for the kind of cepheid variable star we are looking for. Excluding the outlier, the average distance to the cepheids in ncg4527 is 13.9 ± 2 Mpc

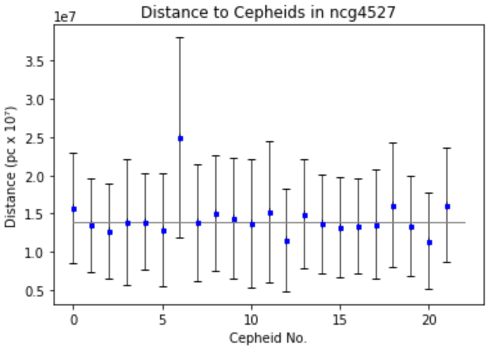

Fig 4. Graph of each Cepheid in ncg4527 and its distance. Grey line is the mean distance excluding the outlier.

We know that the recession velocity of ncg4527 is 1152 km/s. The Hubble constant equation is $v = H_0 D$ where v is recession velocity in km/s, D is distance in Mpc, and H_0 is the Hubble constant in km/s/Mpc. Using only ncg4527, we find that the Hubble constant is 81.7 ± 9 km/s/Mpc. Since the Hubble equation is a model for something we observe and not a natural law, data from a single galaxy may not give an accurate value for the Hubble constant.

Part 3, estimate the Hubble constant using many galaxies.

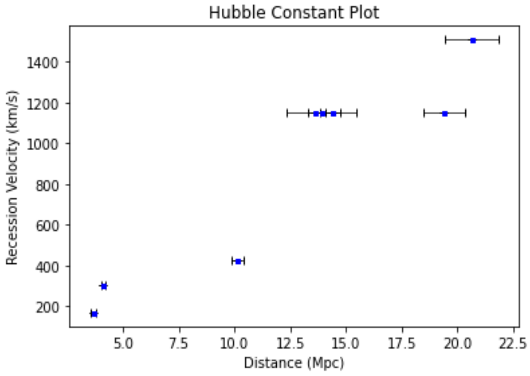

Fig 5. Hubble plot of the given galaxies including ncg4527 without added intrinsic scatter. The trend is clearly visible.

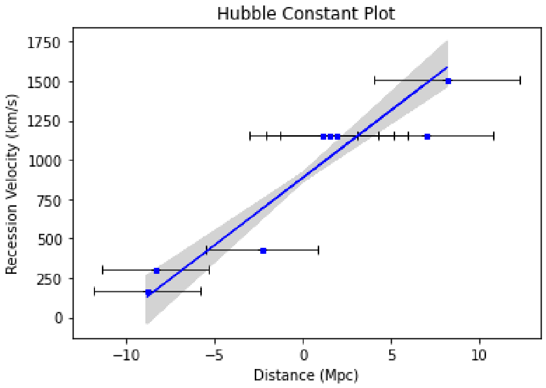

Fig 6. Hubble plot of the given galaxies including ncg4527 with added intrinsic scatter and the model with error.

A Hubble plot is a plot of the recession velocities and distances to many galaxies with recession velocity on the y-axis and distance on the x-axis. The Hubble constant is therefore given by the gradient. Figure 5 is my hubble plot.

Again, since the hubble equation is a model for an observation and not a law, galaxies will naturally not follow the predictions perfectly, therefore to increase the goodness-of-fit of the model I have added intrinsic scatter to each distance error. The value I have chosen is 2.9 Mpc which brings the reduced χ^2 value of the line of best fit into the acceptable range, to 2.8. Running the reduced χ^2 goodness-of-fit test to find the best model for the graph gives us the gradient, which is the hubble constant, and it's error.

The value for the hubble constant given by the graph in figure 6 is 85.8 ± 14.4 km/s/Mpc. Out of curiosity, without the data from ncg4527 the result is 84.4 ± 12.0 with a better goodness of fit value, likely because my ncg4527 values have a large error.

References:

Chakraborty, B. (2023) Hubble Constant

Chakraborty, B. (2023) Reminders Tips Tricks

Benedict, G.F., McArthur, B.E., Feast, M.W., Barnes, T.G., Harrison, T.E., Patterson, R.J., Menzies, J.W., Bean, J.L. and Freedman, W.L. (2007). Hubble Space Telescope Fine Guidance Sensor Parallaxes of Galactic Cepheid Variable Stars: Period-Luminosity Relations. *The Astronomical Journal*, 133(4), pp.1810–1827. doi:https://doi.org/10.1086/511980.