Measurement Errors

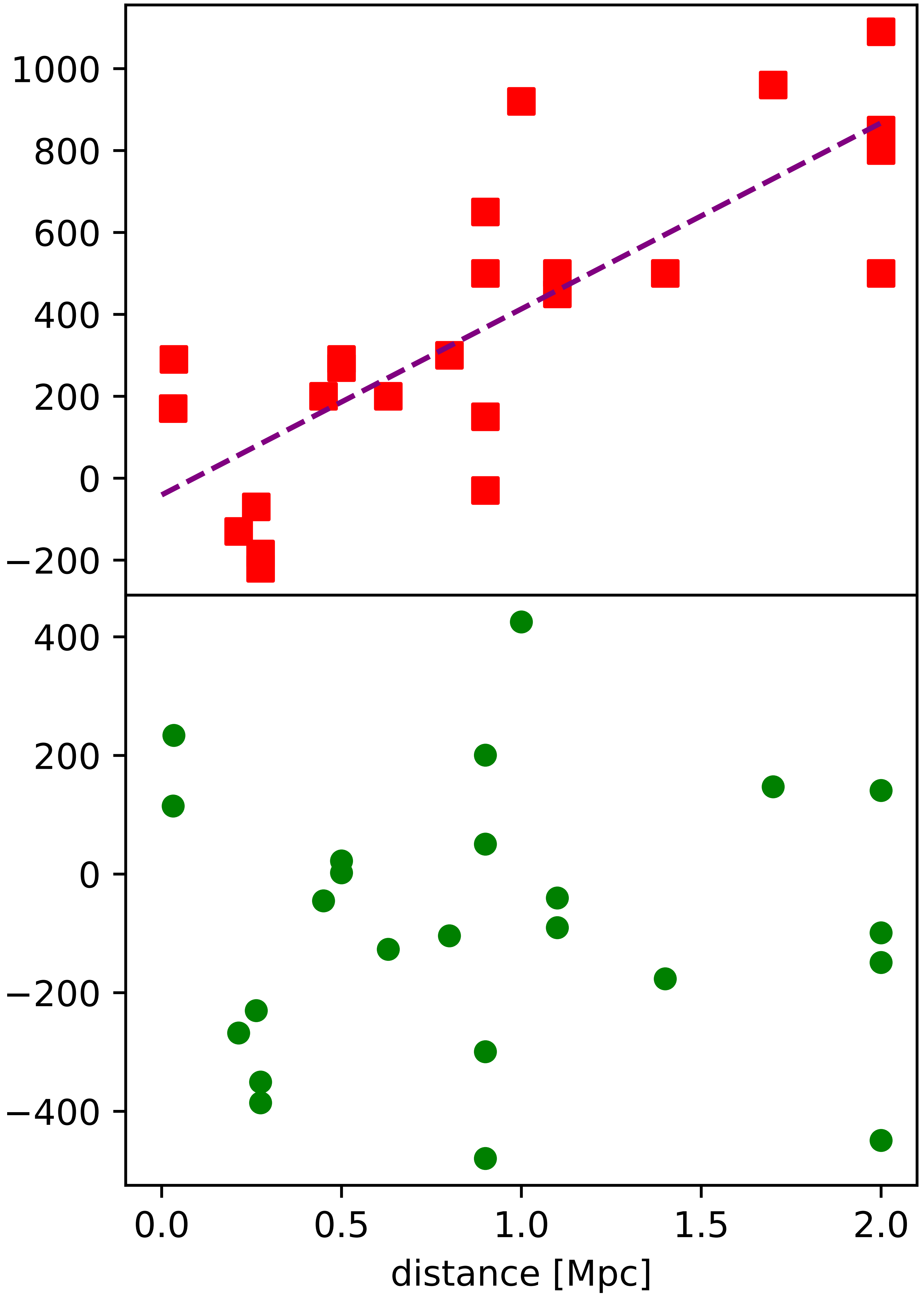
PHYS 2601

Residuals

- As we have seen when fitting a line to the Hubble 1929 data, real data will always differ from a simple mathematical formula.
- The difference between the data and a model are called the residuals and given by

$$y_i - f(x_i)$$

- Note that residuals are defined (and different) for every model, not just a best fit line, or a line.



Errors

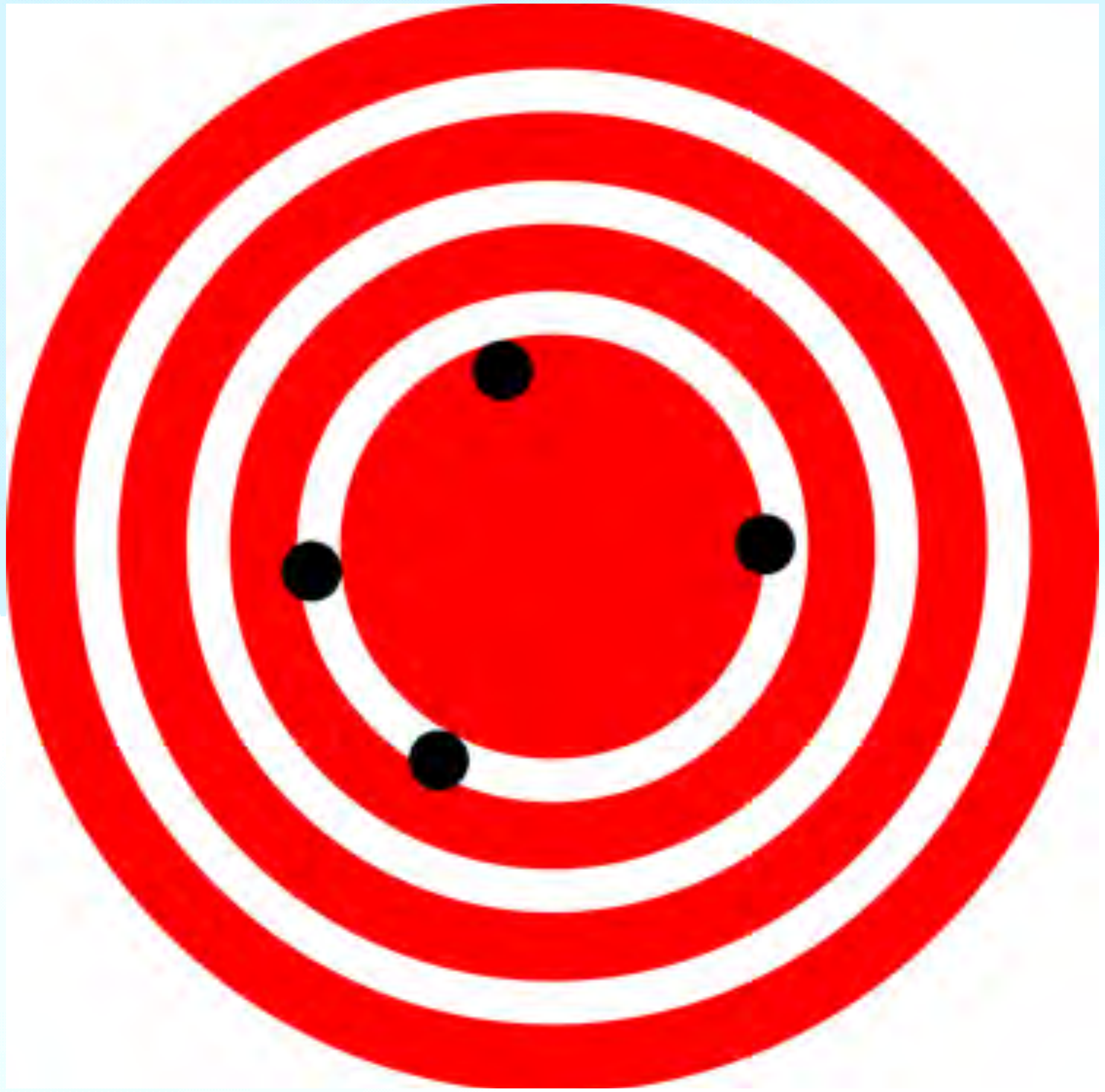
- The reason there are always residuals is because real data always contains errors. The reasons for these errors will vary, but overall we can group them in a few ways.
    - Measurement Errors: Are due to there be a finite accuracy to any measuring device. Furthermore, there may be calibration errors and many quantities are measured indirectly, so the inferred values can have errors because of incorrect assumptions or modeling.
    - Modelling Errors: In addition, there are almost always errors in the model one is trying to fit to the data, either because some physics has not been included or is must be treated with finite accuracy.

Measurement Errors

- All measurements will have errors because any measuring device can only work with finite precision. Thus, a ruler, a clock, a scale can only measure to some level of accuracy. Furthermore, these devices must be calibrated, so that a minute on one clock is the same as a minute on another clock. There can be errors in the calibration and over time measuring devices will tend to become less calibrated so that there will be systematic offsets in its measurements.

- Most measurements aren't even direct measurements but are instead indirect. In this case, errors will occur when converting the measured property into the property of interest. These will be both errors that decrease the precision of the measurement and the accuracy.

Accuracy and Precision



Accuracy is how close a measurement is to the expected value.
Precision is how close different measurements are to one another.

- Errors can also be thought of as being of two types; Statistical Errors and Systematic Errors.

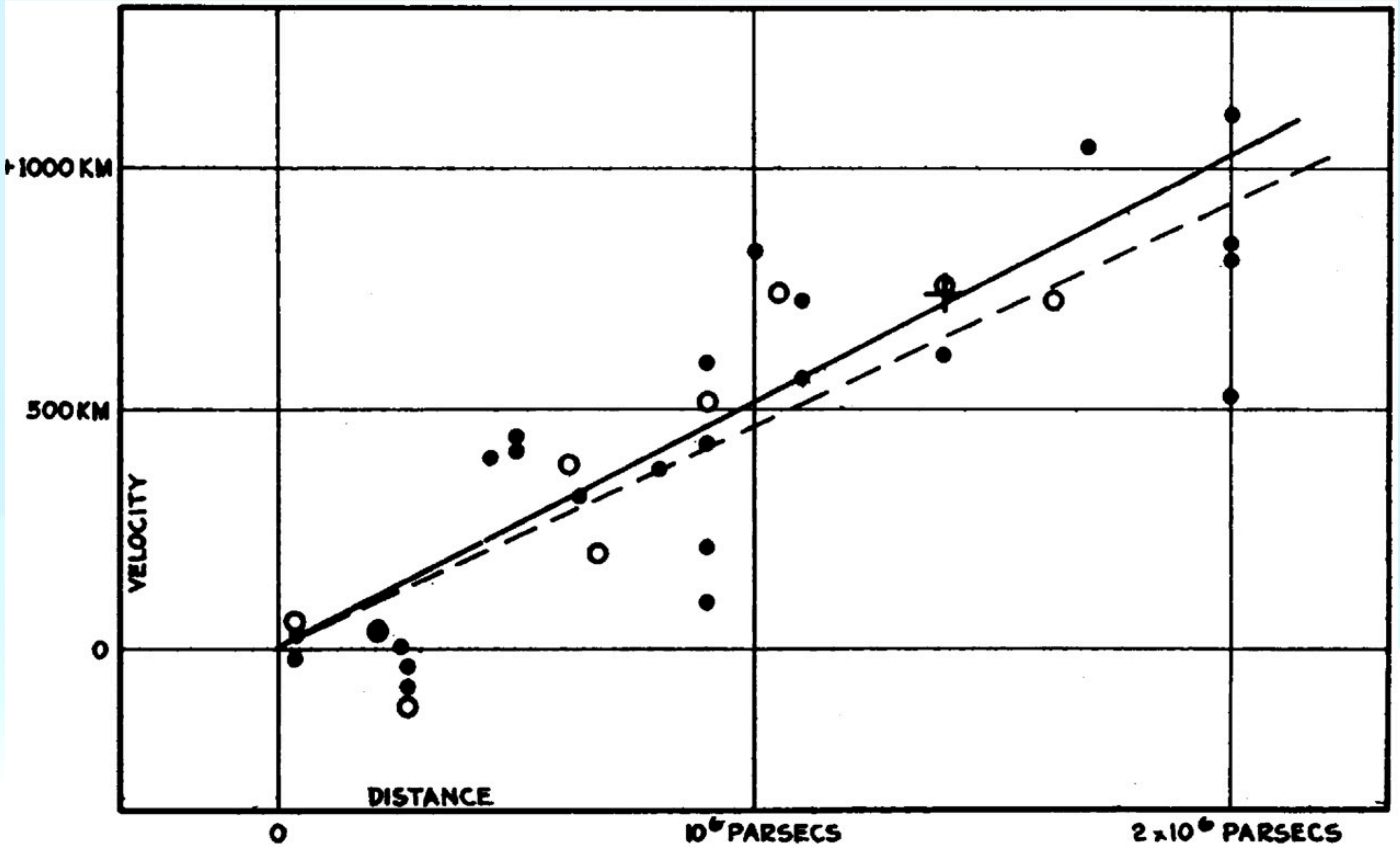
- Statistical errors, refer to the error associated with the different values one gets from making the same measurement multiple times. Thus statistical errors are related to precision, the more precise your measurement the smaller are the statistical errors. The more measurements one can take the smaller the statistical error becomes

- Systematic errors, refer to an error that offsets the measurements from the true correct value. That is accuracy. Systematic errors will often offset the measurements in one direction, then one can think of them as a calibration errror. Increasing measurements won't improve systematic errors.

Modeling Errors

- Besides the errors in the measurements, there are also errors associated with modeling. In general, we will try to create a simple model of some data, so that we can understand it, but that simplicity means we are ignoring some things.
- Even if we try to include everything in our model there will be things we didn’t consider, or our modeling of these things will be in exact.
- However, the goal should only be that our modeling errors are smaller than our measurement errors.

Hubble's Errors



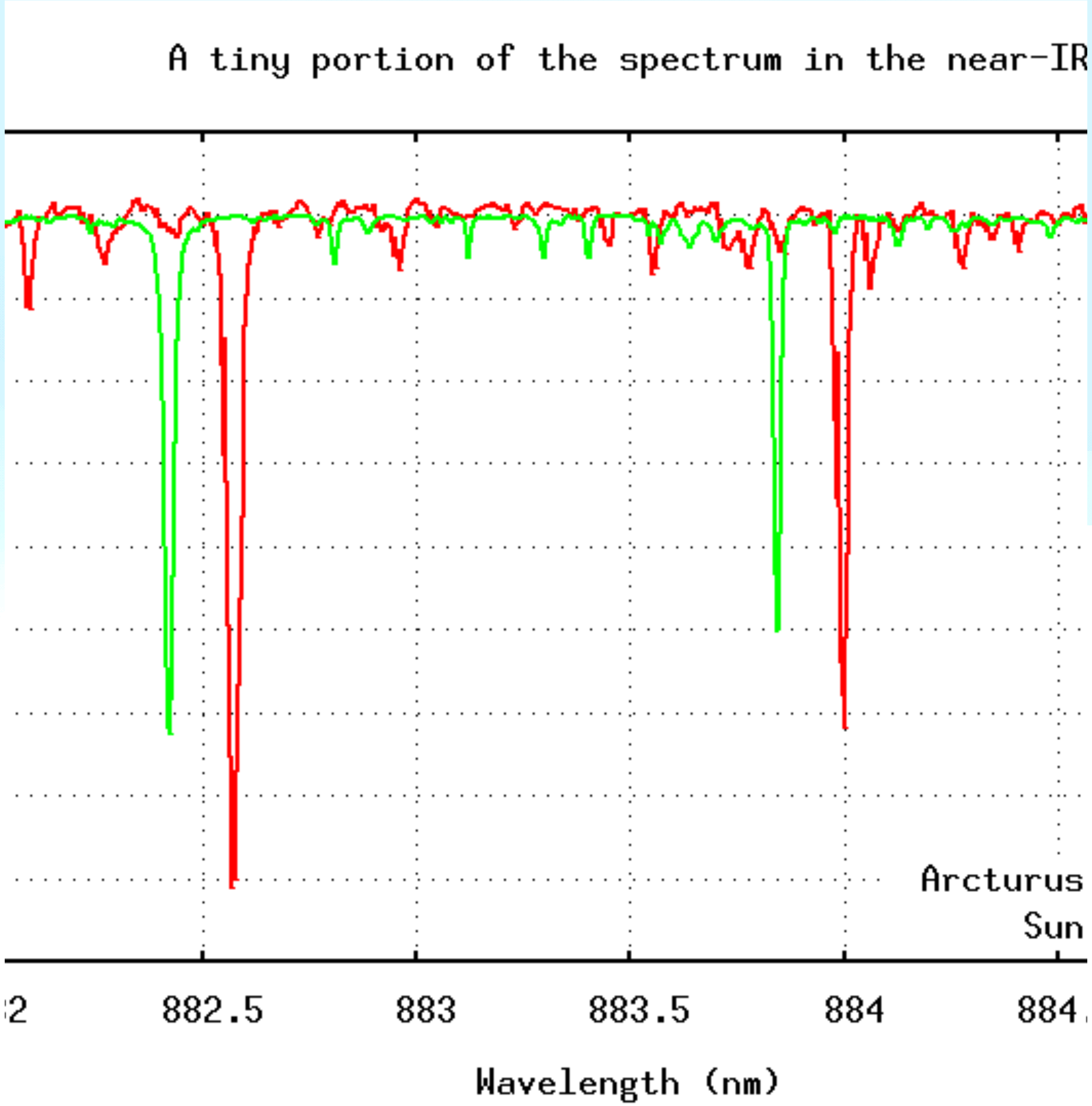
- Let's examine possible errors in Hubble's observations to have an example of measurement and model errors.

- The two measurements are velocities determined from the doppler shift of spectral lines.

- And distance by measuring brightnesses of 'standard candles'.

Radial Velocities

- In astronomy, radial velocities are measured by the doppler shift of spectra.
- These measurement is reasonable accurate, one can usually identify some spectral lines, but the exact amount of the shift has some uncertainty.
- For Hubble, this had a relatively small uncertainty, especially compared to the distance.



Distances

- The easiest way to measure distances is astronomy is using standard candles. That is if you know how luminous some object is and you measure its brightness you can find the distance from

$$B = \frac{L}{4\pi r^2}$$

- But how can one tell how luminous is an object. Hubble used two methods. One there are variable stars who change their luminosity over a few to hundred of days. Thus knowing that a star is variable, tells you its luminosity. The second is that the brightest stars are not that different in luminosity, so if one can identify the brightest star in a galaxy then one has a reasonable guess of its distance.

Model Uncertaintes

- If the Universe is expanding, then galaxies will be moving away from us and the farther away they are the faster they will be traveling.
- But this relationship need not be a line, and we will see later that it actually isn’t a line because gravity causes the rate of expansion to slow down while the cosmological constant causes it to speed up. But at the distances Hubble could probe a line turns out to be correct.
- However, there are other reasons a galaxy might have velocity.
  - The sun moves around our galaxy, which gives an apparent motion to all other galaxies. Hubble took this out by adding this term to all the measured velocities and then solving for the sun’s motion. You didn’t do this in your plots.
  - In addition, galaxies are pulled by the gravity of other galaxies and large scale structure that gives them what astronomers call peculiar velocities, velocities in addition to the expansion of the universe.

Hubble’s Errors

- In turns out Hubble’s measurement where actually off by a lot, he got a slope of the line, called Hubble’s constant of 500 km/s/Mpc. Today we think the value is about 70 km/s/Mpc. At the time people thought his value might be wrong because it implies the universe is 2 billion years old, but geologist had already determined that the Earth was 4 billion years old.
- It turns out that Hubble’s distances are way off. When Hubble thought he was looking at the brightest star in a galaxy, he didn’t know about star clusters. Star clusters are groupings of stars that can be much brighter than a single star and are so close together that they can look like a single star. This made his distances off by about a factor of ~7.
- Because this was a systemic error, it just changed the value of the slope, but not the overall linear relationship, so he was lucky to get the basic picture right even with data that had very large errors.