October 17, 2003

Astronomy 97

John Huchra's Problem for October 24th.

1. Use the least squares fitting technique to derive a Period-Luminosity relation for the galactic calibrating Cepheids. (extra credit — add a "color" term in B-V). Uusally the fit is done as Magnitude versus log Period. (Two logarithmic quantities, a log-log plot, as opposed to luminosity versus period directly).

Least squares is a very common (perhaps the most common) technique used in the sciences to fit data. The simplest version fits a straight line through x and y data points by minimizing the sum of the squares of the deviation of each point from the fitted line. The equation of the line is taken as

$$y(x) \; = \; a \; + \; bx$$

where a is the intercept, b is the slope and one is "fitting y on x." If there is no prior knowledge about errors of the measurements, or if its known that all the points are to be weighted equally, then the solution to the least squares fit for a and b is:

$$a \; = \; \frac{1}{\Delta}(\sum x_i^2 \sum y_i - \sum x_i \sum (x_i y_i))$$

$$b \; = \; \frac{1}{\Delta}(N \sum (x_i y_i) - \sum x_i \sum y_i)$$

where

$$\Delta \; = \; N \sum x_i^2 - (\sum x_i)^2 \qquad .$$

and where N is the number of (x_i, y_i) points.

A good reference for this is Bevington and Robinson's *Data Reduction and Error Analysis for the Physical Sciences*, (New York: McGraw Hill). In a complete analysis, one would include weighting by the errors in the individual points.

Attached is one example of a plot of how this is done. The data files you need can be found on my website:

$$\mathrm{http://cfa-www.harvard.edu/\sim huchra}$$

2. Use the P-L relation derived above to estimate a distance (in Megaparsecs) to the galaxies you have been assigned. (Find the offset between the Cepheids observed in that galaxy and the calibrating P-L relation — there are several ways to do this, you pick one!). Hint — it makes the most sense to measure the

offset where you have data, so don't merely take the zeropoint of the fit P-L relation, a, take the fit at some period like 20 days (log P = 1.3) as the fiducial.

3.. Look up each galaxy's radial velocity (best in the galactocentric or Local Groupocentric reference frames), and estimate the Hubble Constant

$$H_0 \;=\; \frac{\text{Velocity in km/s}}{\text{Distance in Magaparsecs}}$$

4. See if you can estimate and error in that distance by knowing the scatter in the calibrating P-L relation and combining that error with the scatter in the observed Cepheids. For the simple, unweighted fitting done above, the error in the intercept a, σ_a is given by

$$\sigma_a^2 \;=\; \frac{\sigma^2}{\Delta}\sum x_i^2$$

and the error in the slope, b is

$$\sigma_b^2 \;=\; N\frac{\sigma^2}{\Delta}$$

where σ^2 is the variance or the dispersion aroind the line given by

$$\sigma^2 \;=\; \frac{1}{(N-2)}\sum (y_i - a - bx_i)^2$$

where N is the number of data points in the fit. σ is an estimate of the the "standard deviation" of the parent population.

Extra Credit: 5. What's possibly wrong with these distances? List other sources of error, either systematic or random, that can effect any such very simple use of the Cepheid P-L relation — or other techniques, for that matter!