Questions for Discussion

6.1

> Explain this plot (and how you would use it to estimate the Hubble constant) in your own words.

This plot focuses on Type 1A supernovae; it plots their distance to Earth and the speed at which they're moving away from Earth. To estimate the Hubble constant H_0, we could find the slope of the line of best fit.

7.1

> What can we learn just by looking at the chart?

Empirically measured values of the Hubble constant do not always exactly agree; they vary across a range of values. However, it's probably somewhere in the range 68-77 km/s/Mpc.

7.2

> Just by looking at the chart, what is your best guess as to the true value of the Hubble constant?

My best guess for the true value of the Hubble constant is 73 km/s/Mpc.

7.3

> Just by looking at the chart, what would you estimate the uncertainty of the Hubble constant to be?

If we use the standard deviation to quantify the uncertainty in measurements, I'd estimate the standard deviation to be around 3 km/s/Mpc.

8.1

> Which of these statistics would you think gives the "best" estimate of the true value of the Hubble parameter?

I think the *median* would yield the "best" estimate of the Hubble parameter. This is because the mean is too susceptible to outliers and the mode assumes that science is a democracy; by the process of elimination, this leaves the median.

9.1

> Which of these estimates would you use to characterize the uncertainty on the Hubble parameter? Why?

I would use the standard deviation to characterize the uncertainty. This is because the range is susceptible to both high and low outliers, the standard error doesn't seem to be an appropriate statistic (we're not trying to estimate a population mean here), and I don't think the average deviation penalizes large deviations from the sample mean enough.

9.2

> To some extent, this depends on agreeing on what we mean when we say "uncertainty". What do you think a reasonable convention for defining the "uncertainty" of a measurement might be?

I think that, in addition to quoting the mean of a set of measurements, we should quote an interval centered around that mean within which a "good proportion" of measurements lie. This is exactly what the standard deviation does: it tells us an interval centered around the mean within which roughly 68% of the measurements lie.

9.3

> Sometimes it may make sense to quote more than one of these estimates when describing data. What might be some reasons for that?

Each estimate captures a different perspective on the data and allow us to make different inferences. Depending on the situation and the particular claims we want to make, each of these statistics have a time and place.

10.1

> What should we do about this new measurement?

We should probably ignore this measurement.

10.2

> What does this suggest about using the mean or the median to summarize a set of measurements? What about which statistic we might use to characterize the uncertainty?

If we know there are outliers in the data, then using the median would be better than the mean because it's less susceptible to the effects of outliers. We should prefer the standard deviation over the range for the same reasons.