Astronomy 8824: Problem Set 7
Due Monday, December 11, BY NOON

Systematics and Nuisance Parameters

There will not be a final exam for the course.

Problems 2 and 3 illustrate two *different* ways of dealing with calibration errors and their impact on a measurement of H_0. You'll compare results at the end, but don't automatically carry your ideas from Problem 2 over to Problem 3.

1. Best-fit slope and intercept with correlated errors

In Part 4 of PS 5 you generated 10 sets of data points using five different covariance matrices and two random number seeds for each.

For each of these ten realizations compute the best-fit slope and intercept, using the appropriate covariance matrix for each case.

(Hint: Stats Notes 4, p. 6)

2. Calibration errors in H_0 measurement, treated via error covariance

You would like to estimate the Hubble constant using Type Ia supernova distances to galaxies.

Assume (unrealistically) that you have calibrated the mean absolute magnitude (at peak luminosity) of Type Ia SNe with no uncertainty, from local galaxies whose distance is known by other means but which are too close to estimate H_0 because of peculiar velocities.

By comparing the peak apparent magnitude of SNe found in distant galaxies to this absolute magnitude, you get an estimate of $\ln d$ to each of these galaxies. Assume that the error in $\ln d$ has a constant value σ for all of your sample galaxies, which may be realistic if the error is dominated by the intrinsic scatter of supernova luminosities rather than by measurement uncertainties. I'm using $\ln d$ rather than d because a realistic error distribution is closer to Gaussian in $\ln d$.

You also measure the recession velocity v for each galaxy, with negligible uncertainty (which requires your galaxies to be distant enough that peculiar velocities can be ignored).

Hubble's law, $v = H_0 d$, can thus be written $\ln d = \ln v - \ln H_0$. If we think of the velocities as our independent variables x_i and the distance measurements as our data values y_i, then inferring H_0 comes down to determining the intercept of $y = x + b$, where the slope is fixed to unity because we are assuming that Hubble's law is correct for some value of H_0.

(a) For 16 measurements each with $\sigma = 0.08$, what is the expected fractional uncertainty in H_0?

(b) Now throw in a (realistic) wrinkle: the distant supernovae are observed with a different telescope and filter set from the local calibrator sample, so there is an uncertainty in $\ln d$ that affects *all* of the measurements in the same way.

Specifically, if the calibration error is Δ, then the observed value $y_{i,\mathrm{obs}}$ will be Gaussian distributed with dispersion σ about $y_{i,\mathrm{true}} + \Delta$, where $y_i = \ln d_i$.

We don't know Δ, of course, or we would just remove it and calibrate our data to the same system. However, we may know the plausible range of Δ — i.e., the calibration uncertainty $\sigma_\Delta^2 = \langle \Delta^2 \rangle$. (We've done the best we can on calibration, so $\langle \Delta \rangle = 0$.)

The value of σ_Δ is just about half the uncertainty of the photometric calibration in magnitudes. (Why?) A realistic value for good observations might be $\sigma_\Delta \approx 0.01 - 0.02$.

Give a mathematical argument that the covariance matrix of the errors in this case is

$$C_{ij} = \sigma^2 \delta_{ij} + \sigma_\Delta^2,$$

where δ_{ij} is the Kronecker-delta. (Hint: go back to the basic definition of C_{ij}, and think about what happens when you take expectation values.)

(c) For $N = 16$, $\sigma = 0.08$, $\sigma_\Delta = 0.02$, what is the uncertainty in H_0?

(d) More generally, for what conditions on N, σ, and σ_Δ does the calibration uncertainty make an important contribution to the overall uncertainty in H_0?

(e) Suppose that the sample of 16 comes from two different telescopes, $i = 1, 8$ from telescope 1 and $i = 9, 16$ from telescope 2, each with its own calibration uncertainty $\sigma_{\Delta,1}$ and $\sigma_{\Delta,2}$. Assume that the two calibration errors are uncorrelated with each other. What is the covariance matrix for the full data set?

(f) What is the uncertainty in H_0 for $\sigma_{\Delta,1} = \sigma_{\Delta,2} = 0.02$? For $\sigma_{\Delta,1} = \sigma_{\Delta,2} = 0.04$? For $\sigma_{\Delta,1} = \sigma_{\Delta,2} = 0.01$?

3. Calibration errors in H_0 measurement, treated via marginalization

Consider the following data set, available on the web page as `h0.data`, for $\ln(v/\mathrm{km\,s^{-1}})$, $\ln(d/\mathrm{Mpc})$:

1 8.775 4.399
2 8.583 4.327
3 9.098 4.930
4 8.972 4.818
5 8.556 4.328
6 8.576 4.296
7 9.140 4.928
8 8.986 4.880
9 8.912 4.536
10 8.963 4.582
11 8.851 4.523
12 8.652 4.465
13 8.658 4.368
14 8.774 4.463
15 8.698 4.422
16 8.596 4.398

Data points $1-8$ come from Telescope 1 with calibration uncertainty Δ_1 and points $9-16$ from Telescope 2 with calibration uncertainty Δ_2.

Assume that apart from the calibration uncertainty the errors σ in $\ln d$ are Gaussian with dispersion 0.08.

Treat Δ_1 and Δ_2 as nuisance parameters, and adopt Gaussian priors on their values:
$p(\Delta) = (2\pi\sigma_\Delta^2)^{-1/2}\exp(-\Delta^2/2\sigma_\Delta^2)$ with $\sigma_\Delta = 0.02$ for both calibration offsets.

Adopt a flat prior on $\ln H_0$.

(a) The probability of a given set of data points depends on H_0, Δ_1, and Δ_2:
$p(\mathrm{Data}|\ln H_0, \Delta_1, \Delta_2) \propto \exp(-\chi^2/2)$. What is the expression for χ^2?

(b) Write an MCMC program for the 3-dimensional parameter space $\ln H_0$, Δ_1, Δ_2, using the data points above.

From your MCMC chain, plot distributions in the three parameter planes $\ln H_0$ vs. Δ_1, $\ln H_0$ vs. Δ_2, Δ_1 vs. Δ_2.

(c) What is your estimate of H_0 and its fractional error, marginalized over Δ_1 and Δ_2?

What can you infer from your data about the *relative* values of the calibration errors Δ_1 and Δ_2?

(d) If you widen your prior on the calibrations to $\sigma_\Delta = 0.04$, how does your fractional error on H_0 change?

If you sharpen your prior on the calibrations to $\sigma_\Delta = 0.01$, how does your fractional error on H_0 change?

(e) How do the uncertainties in H_0 that you find from this marginalization approach compare to the ones you computed via the covariance matrix approach in Problem 2?