Astronomy 8824: Problem Set 7

Due Friday, December 6, 2019

Systematics and Nuisance Parameters

Parts 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.

Part 1. Best-fit slope and intercept with correlated errors

Download the file ps7data.tgz from the class website. This contains ten data sets generated with five different covariance matrices, with two random number seeds per covariance matrix (and a data file for Part 3). The five covariance matrices are the ones from Part 4 of Problem Set 5.

a) Write an analytic expression for the best-fit intercept θ_1 and best-fit slope θ_2 in terms of the Fisher matrix and the covariance matrix?

b) For each of these ten realizations compute the best-fit slope and intercept, using the appropriate covariance matrix for each case. Also report the uncertainties on both quantities.

Hint: see Stats Notes #4, page 6.

Part 2: Calibration errors in measurement of H_0, covariance matrix approach

The goal of this exercise is to estimate the Hubble constant using Type Ia supernova distances to galaxies and quantify the uncertainties.

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 $y_{i,obs}$ will be Gaussian distributed with dispersion σ about $y_{i,true}$ + Δ, 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 $\sigma^2_\Delta = \langle\Delta^2\rangle$. (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.

Optional question: Why is this the case?

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:

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

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

(c) For N = 16, σ = 0.08, σ_Δ = 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, e.g. points i = 1,8 from telescope 1 and i = 9,16 from telescope 2, each with its own calibration uncertainty $\sigma_{\Delta,1}$ or $\sigma_{\Delta,2}$. 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 $\sigma_{\Delta,1} = \sigma_{\Delta,2} = 0.01$?

Part 3. Calibration errors in H0 measurement, treated via marginalization

The dataset h0.data is also in the file ps7data.tgz. This has 16 data points and columns 2 and 3 are ln(v/km s^{-1}) and ln(d/Mpc).

Data points 1 − 8 come from Telescope 1 with calibration uncertainty $\sigma_{\Delta,1}$ and points 9 − 16 from Telescope 2 with calibration uncertainty $\sigma_{\Delta,2}$.

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

Treat $\sigma_{\Delta,1}$ and $\sigma_{\Delta,2}$ as nuisance parameters, and adopt Gaussian priors on their values:

$$p(\Delta) = \frac{1}{\sqrt{2\pi\sigma_\Delta^2}} \exp - \frac{\Delta^2}{2\sigma_\Delta^2}$$

with σ^2_Δ = 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:

$$p(\ln H_0, \Delta_1, \Delta_2) \propto \exp(-\chi^2/2).$$

What is an analytic expression for χ^2?

(b) Write an MCMC program for the 3-dimensional parameter space $\ln H_0$, Δ_1, and 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_1, and Δ_1, vs Δ_2. For example, you may use `corner`.

(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?

Note: This problem set was developed by David Weinberg.