STATISTICAL INFERENCE FOR ASTRO AND PARTICLE PHYSICS WORKSHOP
WEIZMANN INSTITUTE MARCH 2015

ALAN HEAVENS

1. Introduction: Supernova Cosmology

The purpose of this preparatory exercise is to ensure you are are set up on your laptop to do a few of the basic things which will be needed for the hands-on exercises in the workshop. Basically this means you have some programming language installed and you are comfortable using it (it's your choice what you use; if you want to try R, which is used extensively in the statistics community, download from http://www.r-project.org/). You will need to do some programming, read in data files, plot data, and draw some random numbers from uniform and gaussian distributions; if you don't do these, you won't be able to take advantage of the hands-on part of the workshop (or you will be spending your time in the workshop catching up). The tasks themselves are pretty straightforward - the theory section is here for completeness, but actually all that is needed from this section is to be able to program a couple of the formulae (specifically equations 1-3).

2. Background

This exercise is based around the relationship between the apparent brightness and distance of objects of a standard brightness (so-called 'standard candles'). It is useful as the theoretical relation depends on the amount of Dark Matter, Dark Energy and on the curvature of the Universe. Type I supernova are thought to arise from collapse of white dwarf stars when the Chandrasekhar limit is reached, and are thus[1] standard candles, and the resulting "Supernova Hubble Diagram" was famously used to infer the acceleration of the Universe, culminating in the award of the Nobel prize to Schmidt, Perlmutter and Riess. We will build on this example in the hands-on exercises.

3. Data

The data file (from the Riess et al 2007 'gold' sample) consists of measurements of the redshift (assumed precisely accurate), and a distance modulus μ_i (in magnitudes), with associated measurement errors σ_i. The sample ASCII file (SN.txt) contains one header line and there are 291 supernovae in total, with $z < 1.8$.

---

[1]sweeping many complications under the carpet.

SN	z	μ	σ	Quality
SN90O	0.030	35.90	0.21	Gold
SN90T	0.040	36.38	0.20	Gold
SN90af	0.050	36.84	0.22	Gold
...				

4. Theory

For the purpose of this exercise, we just give you a formula to code[2]. It is the theoretical function $\mu(z)$ for different values of the cosmological parameters Ω_m and H_0, which are the density parameter (today) in matter and the Hubble constant today. In the workshop, a major goal will be to obtain probabilities of these parameters given the data set.

Theoretically, the distance modulus μ is related to the 'luminosity distance' by which is the value of m if the source is at a distance 10pc:

$$(1) \qquad \mu = 25 + 5\log_{10}\left(\frac{D_L}{\text{Mpc}}\right)$$

For a flat Universe, we can use an accurate fitting formula, given by U.-L. Pen, ApJS, 120:4950, 1999:

$$(2) \qquad D_L(z) = \frac{c}{H_0}(1+z)\left[\eta(1,\Omega_m) - \eta\left(\frac{1}{1+z},\Omega_m\right)\right]$$

where the fitting function η is given by

$$(3) \quad \eta(a,\Omega_m) \equiv 2\sqrt{s^3+1}\left[\frac{1}{a^4} - 0.1540\frac{s}{a^3} + 0.4304\frac{s^2}{a^2} + 0.19097\frac{s^3}{a} + 0.066941s^4\right]^{-1/8}$$

and $s^3 \equiv (1 - \Omega_m)/\Omega_m$. This is claimed to be accurate to better than 0.4% for $0.2 \leq \Omega_m \leq 1$. You will see that μ is a function of the redshifts z (which we know), and the parameters Ω_m, H_0, which we would like to know.

If we have measurements of μ, then we can use Bayesian arguments to estimate the parameters Ω_m, h, where $h \equiv H_0/(100\text{km}\,\text{s}^{-1}\text{Mpc}^{-1})$. For anyone unfamiliar with cosmology, these numbers are somewhere between 0 and 1.

5. Exercises

(1) Plot distance modulus vs redshift for $\Omega_m = 0.2$, 0.3, 0.4, 0.5 for the range $0 < z < 2$, for $H_0 = 70\,\text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}$.
(2) Plot the observational data in the attached file (SN.txt) on the graph.
(3) Generate a random set of 20 supernovae with redshifts drawn uniformly from $0 < z < 2$ and with gaussian errors in μ with r.m.s. 0.1 mag, for $\Omega_m = 0.3$ in a flat Universe.
(4) Plot these data on the graph, with error bars.
(5) Plot a histogram of the redshift distribution of the supernovae.

---

[2]For specialists: there is a simplification in the exercise here; we assume we know what the absolute magnitude (or luminosity) of type I supernovae are, but in fact unless we have supernovae with known distances, we don't. M and h are degenerate, since M is set from low-redshift supernovae where we assume Hubble's law to give us the distance. For the purpose of this exercise, we will cheat.