Assessed Problem Sheet 4

Alan Heavens, ICIC, Imperial College London

MRes Bayesian Statistics

Hand in by 5 pm Thursday 14th December. Please submit a project-like report with the results, and a copy of your code included. There are no page limits.

1 Cosmology with Supernovae Ia

The aim of this exercise is to write an MCMC code to infer cosmological parameters from supernova Ia data. Do not use a package for the main code, but you may use packages (e.g. corner.py) to display results.

Supernova Ia are standard candles (or can be made so), so can be used to measure the contents of the Universe.

As standard candles, the apparent brightness (or faintness, represented by a quantity μ) should depend on distance (or redshift z) in a parameter-dependent way, and is illustrated with the SNLS dataset in Fig. 1.

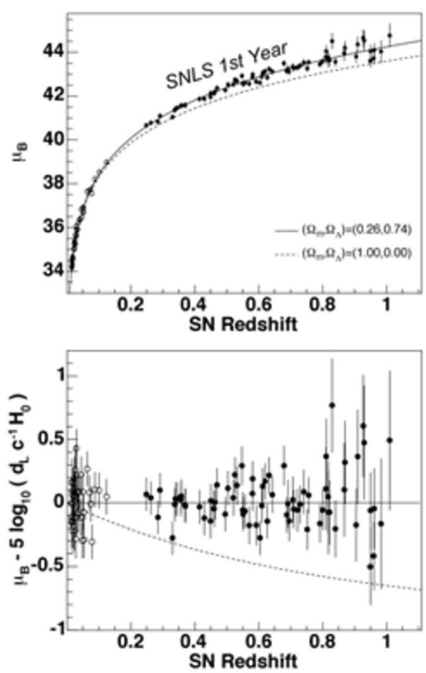

Figure 1: Supernova faintness-distance relation from SNLS.

1.1 Theory and parameters

The flux from a supernova of luminosity L is given by

$$f = \frac{L}{4\pi D_L^2}$$

where D_L is the *Luminosity Distance*. In Big Bang cosmology it is given by

$$D_L = \frac{(1+z)c}{H_0\sqrt{|1-\Omega|}} S_k(r),$$

where

$$r(z) = \sqrt{|1-\Omega|}\int_0^z \frac{dz'}{\sqrt{\Omega_\mathrm{m}(1+z')^3 + \Omega_\mathrm{v} + (1-\Omega)(1+z')^2}}.$$

and $S_k(r) = \sin r, r, \sinh r$, depending on whether $\Omega \equiv \Omega_\mathrm{m} + \Omega_\mathrm{v}$ is $> 1, = 1$, or < 1, and z is the observed redshift of the supernova (known very precisely). $\Omega_\mathrm{m}, \Omega_\mathrm{v}$ and H_0 are the density parameters (today) in matter, vacuum energy, and the Hubble constant. It is beyond the scope of these notes to derive this, but it is standard material for an undergraduate cosmology course.

For a flat Universe ($\Omega = 1$), this simplifies to

$$D_L(z) = 3000h^{-1}(1+z)\int_0^z \frac{dz'}{\sqrt{\Omega_\mathrm{m}(1+z')^3 + 1 - \Omega_\mathrm{m}}}\ \mathrm{Mpc},$$

where $H_0 = 100h\mathrm{km\ s^{-1}\ Mpc^{-1}}$. To avoid evaluating integrals to calculate D_L, we can use an accurate fitting formula (valid for flat universes only), given by U.-L. Pen, ApJS, 120, 49 (1999):

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

where

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

and $s^3 \equiv (1-\Omega_m)/\Omega_m$. This is accurate to better than 0.4% for $0.2 \le \Omega_m \le 1$.

Fluxes are usually expressed on a logarithmic scale, in terms of *apparent magnitudes*, $m = -2.5\log_{10} F$+constant. The *distance modulus* is defined as $\mu = m - M$, where M is the *absolute magnitude* (a logarithmic measure of power output), which is the value of m if the source is at a distance 10pc. With D_L in Mpc, this is

$$\mu = 25 - 5\log_{10} h + 5\log_{10}\left(\frac{D_L^*}{\mathrm{Mpc}}\right)$$

The Hubble constant has been factored out of D_L: $D_L^* \equiv D_L(h=1)$ (so it does not depend on h).

If we have measurements of μ, then we can use Bayesian arguments to infer the parameters $\Omega_\mathrm{m}, \Omega_\mathrm{v}, h$. For anyone unfamiliar with cosmology, these numbers are somewhere between 0 and 1.

1.2 Data

The data file (from the 'Pantheon' sample - see `https://arxiv.org/abs/2112.03863` for more detail) consists of data from 1701 supernovae, with a redshift and distance modulus μ for each supernova. The file (`Pantheon+SHOES.dat` from Teams or Blackboard) contains the data.

1.3 Exercise

Write an MCMC code to infer h and Ω_m from the supernova dataset, assuming the Universe is flat and the errors are gaussian[1], i.e. assume that the likelihood is

$$L \propto \exp\left[-\frac{1}{2}\sum_{i,j=1}^{n}[\mu_i - \mu_\mathrm{th}(z_i)]C_{ij}^{-1}[\mu_j - \mu_\mathrm{th}(z_j)]\right]$$

[1]This is a simplification

where μ_{th} is the theoretical value of the distance modulus, for which you will need to compute the integral for D_L^* numerically, using the fitting formula (for a flat Universe). For clarity, we have not written the full dependence of μ_{th}; we should write $\mu_{\mathrm{th}}(z; \Omega_{\mathrm{m}}, h)$, and indeed it also depends on the (LCDM) model M.

C is the covariance matrix of the data, provided as a list of numbers in an obvious order, from the Teams or Blackboard site, in the file `Pantheon+SHOES_STAT+SYS.cov.txt`.

- Nearby supernovae have redshifts that are significantly altered by 'peculiar velocities', not associated with the general expansion of the Universe. Discard supernovae with $z < 0.01$.
- h and Ω_m are positive, and have values of the rough order of unity
- Assume uniform priors on the parameters (so you will sample from the likelihood)
- You might like to start with a very simple 'top-hat' proposal distribution, where the new point is selected from a rectangular region centred on the old point. For this you will need a simple random number generator. Or use a gaussian for each parameter.
- Explore visually (with trace plots) the chain when you have (a) a very small proposal distribution, and (b) a very large proposal distribution, for a maximum of 1000 trials. What do you conclude?
- Choose a suitable burn-in and say why you chose it.
- Show how the acceptance probability changes as you change the size of the proposal distribution from very small (say 0.001) to very large (say 100).
- Once you have settled on a 'reasonable' proposal distribution, compute the average value of the parameters under the posterior distribution, and their variances and covariance.
- Generalise to non-flat Universes and include Ω_{v} as an independent parameter. You will need to perform the integral for D_L^* numerically.

1.3.1 Tips

If you are estimating h, Ω_{m} and Ω_{v}, you can precalculate D_L for $h = 1$ as a function of Ω_{m} and Ω_{v}, and do a bilinear interpolation when you are running the chains (and divide by h). You only need to do this at the redshifts of the supernovae. This will be much faster than computing D_L every time you change parameters. This is *not* necessary if your parameters are h and Ω_m only.

1.4 Optional Extensions

- Write and apply a Gelman-Rubin convergence test, and deduce roughly how long the chains should be for convergence.
- Compute and display the correlation function of the chain, and show how it changes as the proposal distribution size changes. Calculate the effective number of samples, using the formula in the lectures. Does it reach a maximum for a certain proposal size?