EuCAPT School 2023

Hands On Problems

Andrew Jaffe — Statistical Methods in Cosmology

These are all parameter inference problems, using Metropolis Hastings MCMC. If you wish to try Simulation-Based Inference (SBI) instead, further instructions will come!

1 Testing gravity with gravitational lensing

The aim of this exercise is to write an MCMC code to analyse photographic plates from Eddington's 1919 eclipse expedition, to determine whether the data favour Newtonian gravity or Einstein's General Theory of Relativity. Light from stars that passes close to the Sun is deflected, and during an eclipse, these stars can be detected and their displacements measured, when compared with photographs taken when the Sun is far away.

General Relativity predicts that light passing a mass M at distance r will be bent through an angle

$$\theta_{\rm GR}(r) = \frac{4GM}{rc^2}, \tag{1}$$

whereas an argument based on Newtonian gravity gives half this:

$$\theta_{\rm N}(r) = \frac{2GM}{rc^2}. \tag{2}$$

We can either treat this as a *parameter inference* problem, modelling the bending as

$$\theta_{\rm N}(r) = \frac{\alpha GM}{rc^2}, \tag{3}$$

and inferring α, or as a *model comparison* problem. For this exercise we will do the former. For GR, the bending is 1.75 arcsec for light passing close to the limb of the Sun.

We might conclude that equation (3) is the *model*, with a single *parameter* α. We assume that GM, c and r are known perfectly, from measurements of planet orbits etc. However, the experiment has some other parameters as well, which we investigate below.

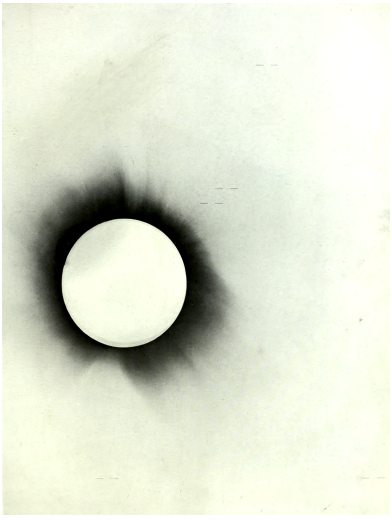

Figure 1: One of the photographs from the expedition.

19. The values of Dx and Dy were equated to expressions of the form

$$ax + by + c + aE_x (= Dx)$$

and

$$dx + ey + f + aE_y (= Dy),$$

where x, y are the co-ordinates of the stars given in Table I., and E_x, E_y are coefficients of the gravitational displacement.

The quantities c and f are corrections to zero, depending on the setting of the scale plate on the plate measured, a and e are differences of scale value, while b and d depend mainly on the orientation of the two plates. The quantity a denotes the deflection at unit distance (*i.e.*, 50′ from the sun's centre), so that aE_x and aE_y are the deflection in R.A. and Decl. respectively of a star whose co-ordinates are x and y.

The left-hand sides of the equation for the seven stars shown are :—

No.	Right Ascension.	Declination.
11	$c-0{\cdot}160b-1{\cdot}261a-0{\cdot}587\alpha$	$f-1{\cdot}261d-0{\cdot}160e+0{\cdot}036\alpha$
5	$c-1{\cdot}107b-0{\cdot}160a-0{\cdot}557\alpha$	$f-0{\cdot}160d-1{\cdot}107e-0{\cdot}789\alpha$
4	$c+0{\cdot}472b+0{\cdot}334a-0{\cdot}186\alpha$	$f+0{\cdot}334d+0{\cdot}472e+1{\cdot}336\alpha$
3	$c+0{\cdot}360b+0{\cdot}348a-0{\cdot}222\alpha$	$f+0{\cdot}348d+0{\cdot}360e+1{\cdot}574\alpha$
6	$c+1{\cdot}099b+0{\cdot}587a+0{\cdot}080\alpha$	$f+0{\cdot}587d+1{\cdot}099e+0{\cdot}726\alpha$
10	$c+1{\cdot}321b+0{\cdot}860a+0{\cdot}158\alpha$	$f+0{\cdot}860d+1{\cdot}321e+0{\cdot}589\alpha$
2	$c-0{\cdot}328b+1{\cdot}079a+1{\cdot}540\alpha$	$f+1{\cdot}079d-0{\cdot}328e-0{\cdot}156\alpha$

Figure 2: The data model.

1.1 Data

The Dyson, Eddington and Davidson 1920 paper (Phil. Trans. R. Society, 220, 571) is open access at `http://rsta.royalsocietypublishing.org/content/roypta/220/571-581/291.full.pdf` and contains the data we need. The figures in this document are taken from there.

1.2 Data model and model parameters

A displacement of the star images will be caused not just by bending of light, but by changes in the scale of the photographic plate (caused e.g., by changes in temperature), rotation of the plate with respect to the comparison plate, and offsets. We are not very interested in these effects, which introduce *nuisance parameters*, but we need to include them. The data model is then given by the equations in Fig. 2, i.e.,

$$\begin{aligned} Dx &= ax + by + c + \alpha E_x \\ Dy &= dx + ey + f + \alpha E_y. \end{aligned} \tag{4}$$

Thus there are 4 parameters for the displacement in x, and 4 for the y displacement. Only the light bending parameter, α is common. The data use units measured on the plate, and you will do the same. Dyson et al. do not give errors in the displacements. Assume they are 0.05 in the units given.

The positions of the stars are shown in Fig. 3, or you can read them in the table in Fig. 2, which also gives the expected bend angle due to General Relativity for each star, as the coefficient of α. This is in a strange unit, so your number α will need to be rescaled. **To translate the inferred value of α from the units they use, you need to multiply it by 45.3.**

The displacements are measured and shown in Fig. 4, for 7 plates (I,II,...VIII; VI was not used.)

TABLE I.

No.	Names.	Photog. Mag.	Co-ordinates. Unit = 50′.		Gravitational displacement.			
					Sobral.		Principe.	
			x.	y.	x.	y.	x.	y.
		m.			″ .	″	″	″
1	B.D., 21°, 641	7·0	+0·026	−0·200	−1·31	+0·20	−1·04	+0·09
2	Piazzi, IV, 82	5·8	+1·079	−0·328	+0·85	−0·09	+1·02	−0·16
3	κ^2 Tauri	5·5	+0·348	+0·360	−0·12	+0·87	−0·28	+0·81
4	κ^1 Tauri	4·5	+0·334	+0·472	−0·10	+0·73	−0·21	+0·70
5	Piazzi, IV, 61	6·0	−0·160	−1·107	−0·31	−0·43	−0·31	−0·38
6	υ Tauri	4·5	+0·587	+1·099	+0·04	+0·40	+0·01	+0·41
7	B.D., 20°, 741	7·0	−0·707	−0·864	−0·38	−0·20	−0·35	−0·17
8	B.D., 20°, 740	7·0	−0·727	−1·040	−0·33	−0·22	−0·29	−0·20
9	Piazzi, IV, 53	7·0	−0·483	−1·303	−0·26	−0·30	−0·26	−0·27
10	72 Tauri	5·5	+0·860	+1·321	+0·09	+0·32	+0·07	+0·34
11	66 Tauri	5·5	−1·261	−0·160	−0·32	+0·02	−0·30	+0·01
12	53 Tauri	5·5	−1·311	−0·918	−0·28	−0·10	−0·26	−0·09
13	B.D., 22°, 688	8·0	+0·089	+1·007	−0·17	+0·40	−0·14	+0·39

Figure 3: The x, y positions of the stars measured. Note that only seven stars are used, numbers $11, 5, 4, 3, 6, 10, 2$.

TABLE II.—Eclipse Plates—Scale.

No. of Star.	I.		II.		III.		IV.		V.		VII.		VIII.	
	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.
	r	r	r	r	r	r	r	r	r	r	r	r	r	r
11	−1·411	−0·554	−1·416	−1·324	+0·592	+0·956	+0·563	+1·238	+0·406	+0·970	−1·456	+0·964	−1·285	−1·195
5	−1·048	−0·338	−1·221	−1·312	+0·756	+0·843	+0·683	+1·226	+0·468	+0·861	−1·267	+0·777	−1·152	−1·332
4	−1·216	+0·114	−1·054	−0·944	+0·979	+1·172	+0·849	+1·524	+0·721	+1·167	−1·028	+1·142	−0·927	−0·930
3	−1·237	+0·150	−1·079	−0·862	+0·958	+1·244	+0·861	+1·587	+0·733	+1·234	−1·010	+1·185	−0·897	−0·894
6	−1·342	+0·124	−1·012	−0·932	+1·052	+1·197	+0·894	+1·564	+0·798	+1·130	−0·888	+1·125	−0·838	−0·937
10	−1·289	+0·205	−0·999	−0·948	+1·157	+1·211	+0·934	+1·522	+0·864	+1·119	−0·820	+1·072	−0·768	−0·964
2	−0·789	+0·109	−0·733	−1·019	+1·256	+0·924	+1·177	+1·373	+0·995	+0·935	−0·768	+0·892	−0·585	−1·166
	−1·500*	−0·554	−1·500	−1·324	+0·500	+0·843	+0·500	+1·226	+0·400	+0·861	−1·500	+0·777	−1·300	−1·322

Figure 4: The displacements Dx and Dy, as measured on 7 plates (I,II,... VII). Note that you have to subtract the number at the bottom of each column - e.g., add 1.5 to the numbers in the first column.

COMPARISON Plates—Scale.

No. of Star.	14_{2a}.		14_{2b}.		15_1.		15_2.		17_1.		17_2.		18_2.	
	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.	Dx.	Dy.
	r	r	r	r	r	r	r	r	r	r	r	r	r	r
11	−0·478	−0·109	+0·967	+1·170	+1·098	+1·228	+0·725	+0·830	−1·073	−1·330	+1·242	−0·302	−1·188	−1·572
5	−0·544	−0·204	+1·013	+1·192	+0·899	+1·232	+0·692	+0·938	−1·072	−1·075	+1·161	−0·224	−1·195	−1·432
4	−0·368	−0·136	+1·030	+1·249	+1·133	+1·086	+0·725	+0·854	−1·296	−1·031	+1·354	−0·281	−1·165	−1·454
3	−0·350	−0·073	+1·044	+1·305	+1·164	+1·114	+0·732	+0·893	−1·278	−1·014	+1·342	−0·261	−1·178	−1·394
6	−0·317	−0·144	+0·980	+1·319	+1·244	+1·012	+0·714	+0·824	−1·375	−1·052	+1·363	−0·390	−1·165	−1·473
10	−0·272	−0·146	+0·997	+1·327	+1·249	+0·960	+0·722	+0·831	−1·424	−1·038	+1·370	−0·423	−1·164	−1·476
2	−0·396	−0·182	+1·102	+1·289	+0·969	+1·052	+0·734	+0·941	−1·236	−0·909	+1·278	−0·328	−1·164	−1·335
	−0·552*	−0·206	+0·967	+1·170	+0·899	+0·960	+0·690	+0·824	−1·424	−1·330	+1·161	−0·423	−1·195	−1·572

* The numbers −1·500, −0·554, &c., given below the line, were taken out to make the values of Dx, Dy small and positive for arithmetical convenience.

Figure 5: The displacements of the comparison plate to the scale plate. As before, subtract the number at the bottom of each column.

However, there is a complication, in that, for technical reasons, it was not possible to compare an eclipse plate directly with the comparison plate, taken with the Sun elsewhere in the sky. Instead, both plates were compared with a reference (so-called 'scale') plate, so to these displacements it is necessary to subtract the displacements of the comparison plate to the scale plate. These are given in Fig. 5. So, for example, if you are using Plate II, and comparison plate 14_{2a}, the displacement that should be entered into your code for Dx for star 11 is $(-1.416 + 1.500) - (-0.478 + 0.552)$.

1.3 Exercise

Write an MCMC code to sample from the posterior joint probability of a, b, c and α, using the Dx displacement data for Plate II. The units are weird, so multiply α by 45.3 to translate it to the correct factor, and compare with the Newtonian result (2) and the GR result (4). Note that Dyson et al. make a correction of about 10% because of errors arising from measuring indirectly the displacements, but we will ignore that here.

- Assume uniform priors on the parameters (so you compute only the likelihood). What do you assume for the form of the likelihood?
- You might like to start with a very simple 'top-hat' proposal distribution, where the new point is selected from a 4D 'rectangular' region centred on the old point. For this you will need a simple random number generator. Alternatively use a gaussian for each parameter.
- Explore visually 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?
- Show how the acceptance probability changes as you change the size of the proposal distribution from very small (say 0.0001) to very large (say 10).
- 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.
- Marginalise over the nuisance parameters and plot the posterior distribution of α.

- Now see what happens if you don't consider the nuisance parameters (i.e., set them all to zero, or include a gaussian prior on each one, that keeps them close to zero). What would you conclude about the gravity theory?

1.4 Optional Extensions

- Write and apply a Gelman-Rubin convergence test, and deduce roughly how long the chains should be for convergence.

2 Cosmology with Supernovae Ia

The aim of this exercise is to write an MCMC code to infer cosmological parameters from supernova Ia data. 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.

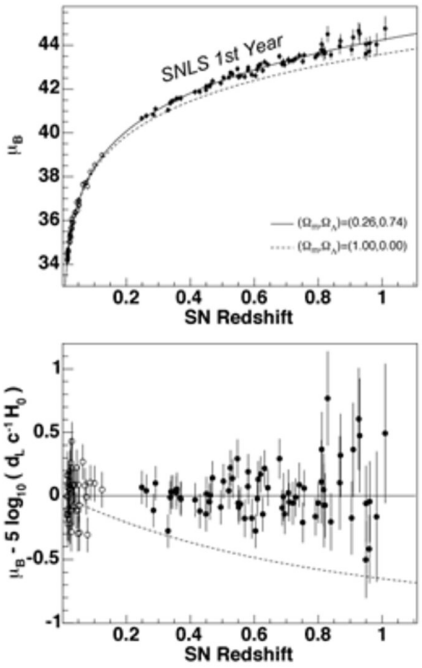

Figure 6: Supernova faintness-distance relation from SNLS.

2.1 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

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

and $S_k(r) = \sin r, r, \sinh r$, depending on whether $\Omega \equiv \Omega_{\rm m}+\Omega_{\rm v}$ is $>1, =1$, or <1, and z is the observed redshift of the supernova (known very precisely). (For avoidance of doubt, note that you have to cancel the $\sqrt{|1-\Omega|}$ factors when $k=0$.) $\Omega_{\rm m}, \Omega_{\rm v}$ 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

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

where $H_0 = 100h{\rm km\ s^{-1}\ Mpc^{-1}}$. 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*, $m = -2.5\log_{10}F + {\rm constant}$. 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 10 pc. The distance modulus is then the logarithm of the ratio of the observed flux to the (assumed known) flux that the object would have at this distance. With the inverse-square law for fluxes used to define the luminosity distacne, and D_L measured in Mpc[1], this is

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

The Hubble constant has been factored out of D_L: $D_L^* \equiv D_L(h=1)$, hence the separate term involving $h = H_0/(100\ {\rm km\ s^{-1}\ Mpc^{-1}})$.

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

2.2 Data

The data file is from the Pantheon+/SH0ES Supernova Sample (Scolnic et al, The Pantheon+ Analysis: The Full Data Set and Light-curve Release, ApJ 938 (2022) 113 [arXiv:2112.03863]) of about 1700 SNIa. The dataset includes a catalogue file with data for each supernova, in particular including the distance modulus μ for each one, using the (inferred) value of the standard candle absolute magnitude M, along with a separate file giving error covariance information. These files are available on the slack channel, along with some python code for reading the files. (Note that this is a different file from the preliminary exercise.)

[1]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. In fact 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.

2.3 Exercise

Write an MCMC code to infer h and $\Omega_{\rm m}$ from the supernova dataset, assuming the Universe is flat and the errors are gaussian[2], i.e., assume that the likelihood is

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

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). The observed distance modulus is μ_i and C_{ij} is the covariance matrix of the data. For clarity, we have not written the full dependence of $\mu_{\rm th}$; we should write $\mu_{\rm th}(z;\Omega_{\rm m},h)$, and indeed it also depends on the (LCDM) model M.

- h and Ω_m are positive, and have values of the rough order of unity
- Assume uniform priors on the parameters (so you compute 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 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?
- 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.
- Optionally, generalise to non-flat Universes and include $\Omega_{\rm v}$ as an independent parameter. You will need to perform the integral for D_L^* numerically.

2.3.1 Tips

If you are estimating h, $\Omega_{\rm m}$ and $\Omega_{\rm v}$, you can precalculate D_L for $h = 1$ as a function of $\Omega_{\rm m}$ and $\Omega_{\rm v}$, and do a bilinear interpolation when you are running the chains (and divide by h). 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.

2.4 Optional Extensions

- Write and apply a Gelman-Rubin convergence test, and deduce roughly how long the chains should be for convergence.

[2]This is largely justified by the averaging over large numbers of supernovae, and using the central limit theorem, but is a simplification.

3 Higgs Boson

The discovery of the Higgs boson was a fantastic example of both experimental science (in its grandest form, the Large Hadron Collider, the LHC) and theoretical physics (as the subsequent Nobel Prize went to the theoreticians who predicted the existence of such a particle). However, this discovery was also fundamentally statistical in nature, as the evidence for the Higgs boson took the form of a small bump in an otherwise smooth but noisy spectrum of measured collision energies.

This project is to implement a simplified version of the LHC Higgs analysis, applying Bayesian methods to a simulated data-set. This is given in the file `spectrum_lhc.dat`, and consists of the number of recorded collisions, $N_{1:B}$, in each of $B = 60$ bins with central energies $E_{1:B}$ (given in GeV). The bins all have an energy width of $\Delta E = 1$ GeV and the data corresponds to LHC runs of total duration $T = 3$ yr.

The overall energy-dependent rate of collisions is given by the sum of the two contributions as

$$R(E) = R_{\text{bkg}}(E) + R_{\text{H}}(E),$$

where $R_{\text{bkg}}(E)$ is the smooth background and $R_{\text{H}}(E)$ is the (possible) Higgs contribution. These rate $R(E)$ is defined such that the average/expected number of collisions in an infinitessimal time period $\mathrm{d}T$ and an infinitessimal energy range $\mathrm{d}E$ is $R(E)\,\mathrm{d}T\,\mathrm{d}E$.

The background spectrum can be taken to be of the form

$$R_{\text{bkg}}(E) = \frac{\Gamma_{\text{bkg}}}{E_0}\,\left(\frac{E}{E_0}\right)^{\alpha},$$

where Γ_{bkg} is the overall background rate (with units of inverse time), $E_0 = 100$ GeV is an (arbitrary) reference energy, and $\alpha < -1$ is the logarithmic slope of the background spectrum.

If the Higgs boson exists, a proposition denoted as H, it would produce a contribution of the form

$$R_{\text{H}}(E) = \Gamma_{\text{H}}\,\mathrm{N}(E; E_{\text{H}}, \sigma_{\text{H}}^2) = \frac{\Gamma_{\text{H}}}{(2\pi)^{1/2}\,\sigma_{\text{H}}}\,\exp\left[-\frac{1}{2}\,\left(\frac{E - E_{\text{H}}}{\sigma_{\text{H}}}\right)^2\right],$$

where $E_{\text{H}} = m_{\text{H}}\,c^2$ is the rest-mass energy of the Higgs boson, Γ_{H} is the associated rate of collisions, and σ_{H} characterises the spread in energy of the Higgs collision products.

1. Read in and make a plot of the LHC data, *i.e.*, number of collisions vs. energy. Add a plausible smooth background-only power-law spectrum to the plot, and hence make an informal assessment of i) whether there is evidence for a Higgs boson in the data and ii) what energy range is plausible for the Higgs boson (under the assumption that the Higgs is present).

2. Show that it is reasonable to assume the probability of recording N_b collisions in bin b, with central energy E_b, is

   $$\mathbb{P}(N_b|\lambda_b) = \Theta(N_b)\,\frac{{\lambda_b}^{N_b}\,e^{-\lambda_b}}{N_b!},$$

   where $\lambda_b = R(E_b)\,T\,\Delta E$ and $\Theta(\cdot)$ is the Heaviside step function.

   Also give an argument why it is reasonable to assume that the bins can be treated independently, and hence use this to write down an expression for the log-likelihood, $\log[\mathbb{P}(N_i, i = 1 \ldots B|E_{\text{H}}, \Gamma_{\text{H}}, \sigma, \Gamma_{\text{bkg}}, \hat{\alpha})]$. (This will be most easily done by again using λ_b as a shorthand.)

3. Fit a background-only model to the data by calculating the posterior $\mathbb{P}(\Gamma_{\text{bkg}}, \alpha|N_{1:B}, \neg\mathsf{H}, \mathsf{K})$, where $\neg\mathsf{H}$ means 'no Higgs'. State explicitly whatever prior information, K, is being assumed.

   Plot the joint posterior distribution in Γ_{bkg} and α, showing 30%, 60%, 90% highest posterior density credible regions.

   Comment on whether the resultant best-fit model can explain the data-set.

4. Adopting the fixed background model that $\hat{\Gamma}_{\rm bkg} = 2000\ {\rm yr}^{-1}$ and $\hat{\alpha} = -4$, use the Metropolis algorithm to draw samples from the posterior distribution of the Higgs parameters, $\mathbb{P}(E_{\rm H}, \Gamma_{\rm H}, \sigma | N_{1:B}, \hat{\Gamma}_{\rm bkg}, \hat{\alpha}, \mathsf{H}, \mathsf{K})$. Provide evidence that the posterior has been sampled effectively, using graphical methods (*e.g.*, trace plots, auto-correlation functions, corner plots) and/or numerical methods (*e.g.*, acceptance fractions, convergence statistics).

   Plot the resultant two-parameter posterior density $\mathbb{P}(E_{\rm H}, \Gamma_{\rm H} | N_{1:B}, \hat{\Gamma}_{\rm bkg}, \hat{\alpha}, \mathsf{H}, \mathsf{K})$ and the resultant one-parameter posterior density $\mathbb{P}(E_{\rm H} | N_{1:B}, \hat{\Gamma}_{\rm bkg}, \hat{\alpha}, \mathsf{H}, \mathsf{K})$ obtained from the samples. Summarise the constraints on the Higgs mass and whether this constitutes any evidence for a detection.

5. Extend the above Metropolis code to simultaneously infer both the background model and the Higgs parameters, by sampling from the joint background+source posterior distribution $\mathbb{P}(E_{\rm H}, \Gamma_{\rm H}, \sigma, \Gamma_{\rm bkg}, \hat{\alpha} | N_{1:B}, \mathsf{H}, \mathsf{K})$.

   Plot the joint constraints in i) the two background parameters and ii) the Higgs energy and rate, and compare these joint results to those obtained previously.

6. Explain how you could use Bayesian model comparison methods to use assess the degree to which this data provides evidence of the Higgs boson. In particular, state how you would deal with the fact that both models (*i.e.*, $\neg\mathsf{H}$ and H) have unspecified internal parameters.