Project 1: A log-likelihood fit for extracting neutrino oscillation parameters

Key Topic: Functional Minimisation

1 Background

Neutrinos are one of the fundamental particles in the Standard Model of Particle Physics. They are the neutral partners to the charged leptons (electrons, e, muons, μ, and taus, τ) and come in three different 'flavours' - ν_e, ν_μ and ν_τ. When neutrinos interact they create their associated charged lepton (*i.e.* a ν_e will produced an electron), allowing us to identify the flavour of the neutrino.

In the Standard Model neutrinos are massless particles. However, at the turn of the millennium the Super-Kamiokande and SNO experiments observed deficits in the expected rate of neutrino interactions from the atmosphere and the Sun. This was the discovery of neutrino oscillations, a quantum mechanical phenomenon that causes neutrinos to oscillate between the flavours as they travel. Neutrino oscillations are only possible if neutrinos have non-zero masses, and is therefore the first observation of physics beyond the Standard Model.

Eqn. 1 is the first-order approximation to the 'survival probability' of a muon neutrino of energy E (GeV) as it travels a distance L (km). This is the probability that the muon neutrino will be observed as a muon neutrino and will not have oscillated into a tau neutrino.

$$P(\nu_\mu \longrightarrow \nu_\mu) = 1 - \sin^2(2\theta_{23})\sin^2\left(\frac{1.267\Delta m^2_{23} L}{E}\right). \tag{1}$$

Here θ_{23} is the 'mixing angle', the parameter that determines the amplitude of the neutrino oscillation probability, and Δm^2_{23} (eV^2) is the difference between the squared masses of the two neutrinos, which determines the frequency of the oscillations. Long-baseline neutrino experiments measure the rate of muon neutrino events as a function of energy (at a fixed distance L) in order to measure these two oscillation parameters.

Neutrinos only interact through the weak nuclear force, which as the name might suggest is exceedingly weak. To put this in context, the neutrino interaction length is most easily measured in light years, compared to millimetres or centimetres for photons. Neutrino experiments therefore use extremely intense sources of neutrinos and extremely large detectors to gather data. One such experiment is T2K, which uses a 500 kW neutrino beam and a 40,000 tonne detector to measure neutrino oscillations. Even so, the weak interaction strength means that T2K is a statistically limited experiment, with $O(100)$ observed neutrino events. Given the low statistics, the number of events is best represented by a Poisson distribution, which naturally leads to the use of a *Negative Log Likelihood* (NLL) fit to extract the oscillation parameters.

2 Negative Log Likelihood fits

Let us consider a probability density function $\mathcal{P}$. In an NLL fit we evaluate the *likelihood* of a given measurement m to come from $\mathcal{P}$ by simply evaluating $\mathcal{P}(m)$. The combined likelihood of n independent measurements $\mathbf{m}$ is given as:

$$\mathcal{L} = \prod_{i=1}^{n} \mathcal{P}(m_i). \tag{2}$$

If we now consider an ensemble of probability density functions $\mathcal{P}(\mathbf{u})$, where $\mathbf{u}$ is the set of unknown parameters we want to estimate, we can calculate the likelihood as a function of $\mathbf{u}$:

$$\mathcal{L}(\mathbf{u}) = \prod_{i=1}^{n} \mathcal{P}(\mathbf{u}; m_i). \tag{3}$$

The value $\mathbf{u}_m$ where $\mathcal{L}(\mathbf{u})$ takes the maximal value will represent the best fit between the data (the set of $\mathbf{m}$) and the probability density function, and we can take the $\mathbf{u}_m$ as estimators of the true values of $\mathbf{u}$. Instead of finding the maximum of Eqn. (3), we will find the minimum of the *Negative Log Likelihood* (NLL)

$$\mathrm{NLL}(\mathbf{u}) = -\ln\left(\prod_{i=1}^{n} \mathcal{P}(\mathbf{u}; m_i)\right) = -\sum_{i=1}^{n} \ln\left(\mathcal{P}(\mathbf{u}; m_i)\right). \tag{4}$$

where $\ln(x)$ means the natural log of x.

2.1 NLL fits for Poisson distributed variables

In High Energy Physics experiments, the number of entries in each bin of a data histogram can be treated as a discrete measurement whose probability follows the Poisson distribution. For one bin, the probability density function for a Poisson-distributed integer number of entries m is

$$\mathcal{P}(m) = \frac{\lambda^m e^{-\lambda}}{m!}. \tag{5}$$

where λ is the expected average number. For many bins, m_i denotes the observed number of neutrino events in bin i. Simulations are used to predict the expected number of events λ_i in bin i. This prediction will depend on the set of unknown parameters, $\mathbf{u}$, giving $\lambda_i(\mathbf{u})$.

The general NLL formula above for the Poisson case becomes

$$\text{NLL}(\mathbf{u}) = -\ln\left(\prod_{i=1}^{n} \frac{\lambda_i^{m_i}(\mathbf{u})\, e^{-\lambda_i(\mathbf{u})}}{m_i!}\right) = \sum_{i=1}^{n}\left[\lambda_i(\mathbf{u}) - m_i + m_i \ln\left(\frac{m_i}{\lambda_i(\mathbf{u})}\right)\right]. \tag{6}$$

where Stirling's approximation $\ln(m!) \approx m\ln(m) - m$ has been used.

3 Project

The aim of the project is to create an NLL fit to extract the neutrino oscillation parameters of Eqn. 1 from a set of simulated T2K data. We can apply Eqn. 6 to our situation with $\mathbf{m}$ being measurements of the muon neutrino event numbers, binned as a function of energy, and $\mathbf{u} = (\theta_{23}, \Delta m^2_{23})$.

3.1 The data

Create a function to read the experimental data and the unoscillated event rate prediction from the supplied data file. Personalised data files can be downloaded from:

`https://www.hep.ph.ic.ac.uk/~ms2609/CompPhys/neutrino_data/username.txt`,

where you should replace `username` with your college username (xyz123, for example). Please note that copying and pasting this text from the PDF to a browser will not work, due to different text encodings - please type the link out in your browser!

Each data file contains the observed number of muon neutrino events from 0–10 GeV in energy, with each new line representing a single 'bin' of energy. There are 200 energy bins in total, so, for example, the first number in the file therefore represents the number of events observed with an energy between 0 GeV and 0.05 GeV.

Create a few histograms of the data to make sure you understand what you are looking at.

3.2 Fit function

Code the oscillation probability $P(\nu_\mu \longrightarrow \nu_\mu)$ using Eqn. 1, and create some plots as a function of energy, E, to make sure you understand the effect of changing the oscillation parameters. Use the following initial values for the variables: $\theta_{23} = \frac{\pi}{4}$, $\Delta m^2_{23} = 2.4 \times 10^{-3}$ and $L = 295$.

The data file contains both the data and the simulated event number prediction assuming the muon neutrinos do not oscillate. Code a function to apply the oscillation probability from Eqn. 1 to the simulated event rate to produce $\lambda_i(\mathbf{u})$, the oscillated event rate prediction.

Comparing the plots to the data by eye should give you a rough estimate of the oscillation parameters.

3.3 Likelihood function

Write a function to calculate the negative log likelihood, Eqn. 6. Creating a graph of the NLL as a function of θ_{23} should allow you to find the approximate position of the minimum.

3.4 Minimise

Code a parabolic minimiser as presented in the lectures. The interval you search for the minimum in should be based on your rough estimate of the value for θ_{23}. Test the minimiser on the oscillation probability function (or equivalent validation function) first to make sure it works as expected. You will need to develop a criterion for when the method has converged. Now go on to find the value of θ_{23} that minimises the NLL.

3.5 Find accuracy of fit result

The accuracy of the fit results will depend on the available statistics of the input. In general for an NLL fit we can estimate the error by looking at the shape of the NLL function around its minimum. The values θ_{23}^{+} and θ_{23}^{-} where the NLL is changed by 0.5 (i.e. in absolute units, not a factor 2) compared to the value at the minimum correspond to a shift of one standard deviation in the positive and negative direction. You can also estimate the error from the curvature of the last parabolic estimate. Do this and compare to the result from the estimate based on the change in the NLL. Comment on the relative merit of the two methods.

Create a method to automatically find the standard deviation. The error can be estimated from the scans of the function close to the minimum if you cannot get an automatic method to work. You should at this stage be able to quote a measurement of θ_{23} with an error based on the complete dataset. Comment on how you would expect the error on this measurement to change as θ_{23} approaches $\frac{\pi}{4}$.

4 Two-dimensional minimisation

4.1 The univariate method

As you know from Sec. 3.2, the oscillation probability depends upon both θ_{23} and Δm_{23}^2. Changes in Δm_{23}^2 could be confused with changes to θ_{23}, so in practice a two-dimensional fit is used to extract values of both parameters.

Implement the 'Univariate' multi-dimensional fitting method discussed in the lectures to find the best fit values of both θ_{23} and Δm_{23}^2. Consider the order in which you should minimise the variables to get the best result.

4.2 Simultaneous minimisation

A more correct way to approach this problem is to simultaneously minimise the NLL with respect to both θ_{23} and Δm_{23}^2. Develop your own simultaneous two-dimensional minimiser and validate it using an appropriate function. Use this to extract the values of θ_{23} and Δm_{23}^2 from your dataset, including their uncertainty. Comment on any difference between these results and what you measured previously.

5 Neutrino interaction cross-section

The neutrino interaction cross-section governs the expected rate of neutrino events - if the cross-section increases then the number of observed events per neutrino should increase as well. So far we have assumed that the neutrino interaction cross-section is constant as a function of neutrino energy. In reality the cross-section increases approximately linearly with energy, and the rate of increase is not well known.

- Code up a new function to predict $\lambda_i(\mathbf{u})$, including the effect of the cross-section, so that the expected number of events is scaled in proportion to the neutrino energy. Since the cross-section is not well known it should be included as another free parameter in your NLL minimiser. The new predicted number of events in a histogram bin should look something like:

$$\lambda_i^{\text{new}}(\mathbf{u}) = \lambda_i^{\text{old}}(\mathbf{u}) \cdot \alpha \cdot E_\nu \tag{7}$$

  where α is the free parameter.

- Update your minimiser to handle this new parameter and then produce an estimate of θ_{23}, Δm_{23}^2 and the rate of increase of the cross-section with neutrino energy. Comment on your results.