

Activity Worksheet: Evaluating Uncertainties

Cheryl Patrick, UCL

6 July, 2021

Many thanks to the master of CAFAna, Chris Backhouse (UCL), who helped to create the exercises for this activity.

In the last couple of years, we've all had to deal with plenty of uncertainty in our personal lives — but for this week's activity, we'll be looking at the effects of uncertainty on our physics measurements.

For this activity, we'll be looking at simulation from the DUNE near detectors - specifically, the ND-LAr, which will use liquid argon as its neutrino target and detection medium.

In this worksheet, the commands you need to type will be formatted like this. Anything formatted *like this* is something you need to substitute – with something like your login name, or the server you're using.

1 Getting started

This section is copied from the previous activity.

As for the previous activity, you'll need a Fermilab account with access to the DUNE virtual machines. If you are in the hands-on class and don't have an account, you'll have to follow along with your team for now – you can try the exercises for yourself once you get an account. Here's how to get set up:

<https://indico.fnal.gov/event/48900/page/2795-requirements>.

Now you have your permissions, let's connect to the Fermilab server:

On Linux or Mac:

1. Open a terminal window.
2. Authenticate with

```
kinit -f <your Fermilab username>
```

and enter your Kerberos password when prompted.

3. Connect to the DUNE server: `ssh -XY dunegpvmxx.fnal.gov` (The DUNE servers are numbered from 02 to 15, substitute any of those numbers for `xx`. If we all use the same one, it will be slow!)

On Windows:

1. Start Xming
2. Get a Kerberos ticket following the *Getting a Ticket* instructions here:
https://fermi.servicenowservices.com/kb_view.do?sysparm_article=KB0011316.
3. Connect to the server with PuTTY following the instructions in *Connect to Kerberized Host using SSH Profile* on the same page. Note that you can save your profile to make this easier! Your server will be `dunegpvmxx.fnal.gov` (The DUNE servers are numbered from 02 to 15, substitute any of those numbers for `xx`. If we all use the same one, it will be slow!)
4. When your terminal opens on the Fermilab server, get another kerberos ticket on there, with the best chance possible of working: `kinit -f -p -A -r 8d -l 26h your_username@FNAL.GOV`. Be sure to put `FNAL.GOV` in capitals.

5. You should now be logged onto the Fermilab servers. From now on, the instructions should be the same whether you're using a Windows, Linux or Mac computer.

Now log onto your work area (check https://wiki.dunescience.org/wiki/Summer_School_Software_Setup if you have trouble):

Do this every time you connect!

1. Go to your working area
 - On the DUNE machines: `cd /dune/app/users/${USER}`
 - On the GENIE machines: `cd /genie/app/users/${USER}`
2. Switch to the folder we're using for this school: `cd DUNESchool/`
3. Set up the analysis framework: `source lblpwgtools/CAFAna/build/Linux/CAFAnaEnv.sh`

2 Scripts for this session

Copy the example scripts to your own home area (you should be in your DUNESchool directory – if you aren't, switch to it now. Then get the scripts from GitHub:

```
git clone https://github.com/cherylepatrick/DUNESchoolSystematics.git
cd DUNESchoolSystematics/
```

3 Statistical uncertainty

This section will help you understand the effects of statistical uncertainty on a probabilistic distribution. It doesn't use a lot of new CAFAna techniques, so if you are experienced in particle-physics analysis and feel you understand this well already, you might want to skip to the next section. If you're newer to this kind of analysis, you might find it useful, and a good revision of concepts from the previous class.

When we measure some physical quantity — a number like the CP-violating phase, a distribution like a differential scattering cross-section, or just something as simple as measuring a time with a stop-watch — it's important to understand not only the value we measure, but how well we know that value.

As neutrino-interaction processes are probabilistic, a key contributor to the uncertainty in our measurements is statistics - are the events we've observed really typical of what happens *on average*, or have we just happened to see an anomaly? Let's say we flip a coin - it's tails! Does that mean the coin will always show tails? Of course not! And if we flip it 10 times, will we get 5 tails results? Maybe, but it wouldn't be surprising if we got, 4, or 6, or even 7. But on the other hand, if we flipped 10,000 times - would we expect to see 7000 tails results? Probably not! But maybe we'd get something like 5020. And maybe if we did the whole flipping business again, next time we'd get 4990. Keep repeating these 10,000-flip experiments and — in addition to very sore fingers! — we'd end up with a distribution of measurements centered around a best-fit value (hopefully around 5000 tails), but also with some associated width (according to statistics, we'd expect the square root of 5000), indicating the variation between our individual counts. The central value and width of the distribution depend on the number of flips (10,000) in each experiment - our confidence in how well we know the central value and width depend on how many experiments we run - the more times we repeat the experiment, the better we'll be able to extract these values correctly.

So what does this mean for neutrino-interaction events? Just like coin-flips, neutrino interactions are probabilistic; the cross sections we measure are probability distributions. And because neutrino interactions are rare, to fully understand the subtleties in our distributions, we need to keep our experiments running for a long time — often several years. If we just ran for a few days, our measurements would be overwhelmed by statistical uncertainty, and we wouldn't be able to draw any conclusions.

Similarly, the Monte Carlo generators that produce our simulation are based around making random choices based on probability distributions. That means that if we run our simulation twice, we won't get exactly the same results both times. Let's take a look at what this means for our uncertainty. As in the previous hands-on activity, we'll do that by editing some example CAFAna scripts to analyze simulated data.

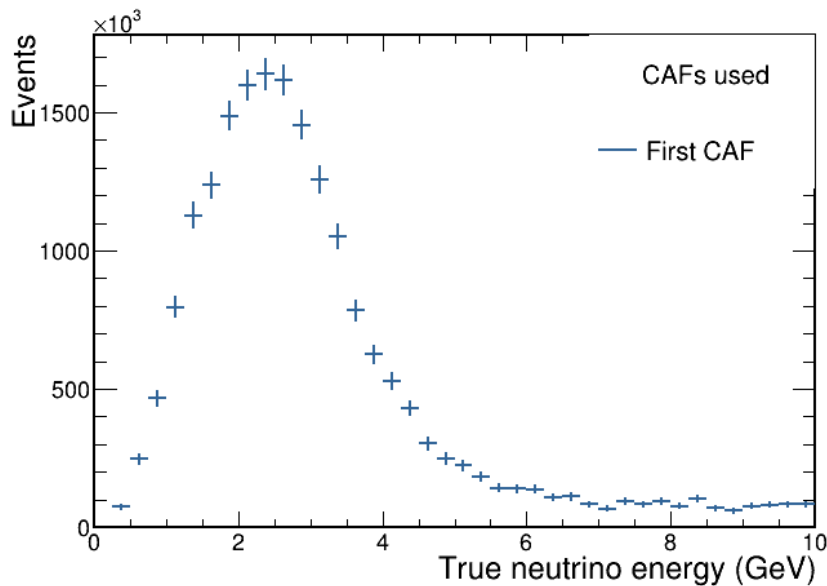


Figure 1: *True neutrino energy for ν_μ charged current interactions in one CAF, with statistical uncertainty.*

Running CAFAna scripts

Once you've done the setup listed earlier, running a CAFAna script is simple.

To run a script, type `cafe scriptname.C` at the command line. That will work for all the scripts in this tutorial. When you are finished, quit the CAFAna environment with `.q`

By default, CAFAna will print any plots you make to the screen. You can speed it up by running it without graphics (any images you save to disk will still be created): `cafe -bq scriptname.C`. In this case, CAFAna will automatically quit when it finishes your script.

In future, when you're developing and testing a script, you might want to speed it up by running over only a fraction of your CAFs. To do this, we use the `--stride` option. For example, to test with only 1/5 of your CAFs, you'd type `cafe --stride 5 scriptname.C`

1. If you haven't done so already, get the scripts for the session following the instructions in section 2.
You'll notice solutions for each exercise - no peeking until you've tried them!
2. We'll start by editing the script `Systematics1.C` (Advice on text editors [here](#).) It shows something you'll remember from last week - the true neutrino energy distribution. This time, we're plotting it for all charged-current muon-neutrino scattering events in a single CAF file. As a refresher from last week, look through it for the lines that define:
 - The `SpectrumLoader` that loads the data from the CAFs. What's our source this time?
 - The `Var` describing the variable your plot
 - The `Binning` that shows how we'll divide up our sample for the histogram
 - The `Cut` specifying the event selection (cuts)
 - The `HistAxis` defining the title, and linking the bins and variable details
 - The `Spectrum` that links the cuts with the axis
 - The code that sets the simulated beam exposure (POT, or protons on target) that we're mimicking.
 - The code to plot the histogram, add it to the legend, and save it to a file.
3. This will draw the distribution in histogram mode "HIST". Change the code to instead draw it with error bars (the draw option is "E"). Run the code with `cafe Systematics1.C`. You should see a plot like figure 1.
4. Is the uncertainty bigger for bins with lots of events, or for bins with fewer events? Is that what you would expect? Why?

5. There's a commented-out line in the code to Print the histogram. That will let you see the value and uncertainty ('error') of each bin. Un-comment that and use the numbers to calculate the fractional uncertainty (uncertainty / value) in a high-statistics bin (bin 10) and a low-statistics bin (bin 2). Which has the higher fractional uncertainty? Is that as you would expect? Why?

*Note: In particle physics, you will often hear uncertainties referred to as 'errors', with terms like 'error bar' being used. Don't confuse this with a phrase like 'there is an error in this formula', which would mean that the formula was actually **wrong**. For exercises like the ones you're doing today, 'errors' and 'uncertainties' are the same thing.*

- Bin 2 value _____ error _____ fractional uncertainty _____
- Bin 10 value _____ error _____ fractional uncertainty _____

6. Now let's see what the distribution looks like for a different CAF. Add a second line to your plot. Again, we'll load from a single CAF, but a different one. Edit the file to add:

- A second SpectrumLoader, loading from the file defined by SECOND_CAF.
- A Spectrum with the same binning and axis as before, but using your new loader
- Code to activate your loader
- Code to draw your histogram (use draw options "E SAME" to put in on the same axes). Choose a new color for it, and add it to your legend.

7. Look at the plot showing both distributions. The two CAFs you looked at each represent very similar total exposures to the neutrino beam.

- Are the two distributions exactly the same? Would you expect them to be?
- Do you think they are the same within statistical uncertainty? The error bars represent a 1σ uncertainty - that means that if you did the same experiment repeatedly, you'd expect 68% of results to be within the error bar.
- Do the error bars on the two distributions look similar?

8. I've also defined a wildcard path TEN_FILES that points to a set of ten CAFs, each around the size of the two you looked at so far. Add a third line to your graph corresponding to a Spectrum created with this set of files.

- You have 10 times the statistics for this third sample. Is your distribution 10 times the height? If not, why not? (Hint - remember that we scale our distributions to a fixed exposure (POT))
- Do you think the three plots represent the same distribution?
- Look at the error bars - how do they compare to the other two distributions. Print the new histogram and write down the values and errors for bins 2 and 10 again:

- Bin 2 value _____ error _____ fractional uncertainty _____
- Bin 10 value _____ error _____ fractional uncertainty _____

Compare them to the values with the smaller dataset. How do the errors compare? What would you expect the ratio of the uncertainties on the bigger and smaller samples to be? (Hint, remember the uncertainty on N counts is \sqrt{N} . What's the ratio of the counts in the two samples?)

- If you're using these distributions to calculate a quantity like a cross section, which do you think is more important - the absolute uncertainty, or the fractional uncertainty? What can we do to minimize the effect statistical uncertainty has on our physics measurements?

9. Optional bonus questions (no solution available)

- Try changing the POT for your histograms from $1e20$ to $1e21$. What changes do you expect to see to:
 - The total height of the plots?
 - The uncertainties?
 - The fractional uncertainties?
- Try changing the binning from 40 bins to 10 bins. What difference do you expect it to make to:
 - The total height of the plots?
 - The uncertainties?
 - The fractional uncertainties?

When we decide how long to run an experiment for, our goal is typically to become *systematics limited* - to run for long enough that our fractional statistical uncertainties are smaller than our fractional systematic uncertainties. As it's much cheaper to generate Monte Carlo simulation than to run a beam / detector, we make sure to generate our simulation with higher statistics than our data (perhaps ten times), so our measurement's precision is never limited by Monte Carlo statistics.

Statistical uncertainty is an important consideration when we choose how to bin our data. If we have distributions like the one we just studied, where we have much higher statistics in some regions of phase space, we sometimes choose variable-width bins to give ourselves sufficient statistics in every bin.

4 Systematic uncertainty

If you skipped the previous section, and you haven't done so already, get the scripts for the session following the instructions in section 2.

In contrast to statistical uncertainty, which can be reduced by using bigger volumes of simulation or data, systematic uncertainty is a measure of how well we understand our data. This could be due to the resolution (the number of 'significant figures' to which we can measure) or accuracy (any systematic offset or bias in our measurements) when we measure or reconstruct a quantity. Or it could be due to an uncertainty in the models we use for our simulation - how well do we really know the shape of the neutrino flux? How sure are we of the relative numbers of quasi-elastic vs. resonant scattering events?

The biggest challenge with these uncertainties is understanding them - how do we quantify our ability to reconstruct muon energies? Does it depend on the muon's energy, its direction, or where in the detector it was created? Those are physics questions, and they're for another day. For now, we'll look at how to model these uncertainties in CAFAna, how they can affect our distributions, and how to combine them.

We'll do this by considering three different example sources of uncertainty: a bias in a measurement, an uncertainty on reconstruction precision, and an uncertainty in one of our neutrino-interaction models.

4.1 Uncertainty due to a bias

Our first example is an uncertainty on our energy calibration. One way to measure the energy of a particle in our detector, is by summing the energy deposits from its track (there are also other tricks we can use like a track's curvature or range... but let's not worry about those for now). To measure the energy of a track we add the voltages received by our electronic sensors. Then we have to convert these measurements (in volts) into a particle energy (in GeV). Let's say we only know this *calibration* to the nearest 20%. It's possible that, in the worst case, all our muon energy measurements are 20% too high, or 20% too low. (DUNE will do much better than that; I've exaggerated it to show what's going on.) Let's see what that would mean...

1. The file *Systematics2.C* will draw the reconstructed muon energy distribution for $CC0\pi$ events (look at the previous activity for a refresher on what that means and why we pick that cut.) Run it draw the spectrum: `cafe Systematics2.C`
2. Edit the *Systematics2.C* file - you'll find that there's some extra code in there to model the effect of all reconstructed muon energies being shifted upwards by 20%. Read and understand the code - you'll find it makes a shifted `Spectrum` object, which we've called `sScaleUp`. Add this `Spectrum` to your plot (Hint - make a histogram object for the new spectrum and draw it with the `HIST SAME` option. Don't forget to add it to the legend!)
3. Now add another `Spectrum` to represent a shift down of 20%. You can use the same shifting class as before, but you'll need to make a new `SystShifts` object for the downwards shift, and then produce a `Spectrum` with your new shift. Add it to your plot and legend.
 - Do you see the effects of an increase and decrease in the energy as you'd expect?
 - Is the difference from the central value more pronounced at higher or lower energies? Why do you think that is?
 - Your central value's drawn with statistical error bars. With this systematic uncertainty and the statistics in this file, are we statistics-limited, or systematics-limited?

4.2 Uncertainty on measurement resolution

Sometimes, rather than having an unknown bias in our data, we might just have an uncertainty on the resolution to which we can measure it. A common way of understanding the effect of that is to **smear** our measured quantity. By this, we mean that we multiply our central value measurement by a random number from a Gaussian distribution centered on 1, with a width corresponding to our 1-standard-deviation uncertainty. Let's try it.

1. In your code, copy the `EMuScale` class and change the name of the copied class to `EMuSmear`.
2. Edit the code of the new `EMuSmear` class so that, instead of scaling the reconstructed muon energy, you instead smear the reconstructed muon energy by 20%. (Hint: To smear a number by $\sigma \times 20\%$, multiply it by $1 + \sigma * \text{gRandom} \rightarrow \text{Gaus}(0, 0.2)$)
3. Because this is a smearing effect, it will make some energies higher, and others lower. That means there's no reason to have a shift up and a shift down. So you just need to add one new uncertainty line, shifting by 1 sigma.
 - Does a 20% smearing have a bigger or a smaller effect than a 20% shift? Why do you think that is?
 - The effect of smearing is typically to 'smooth out' fine detail in your spectrum (look at the lecture on cross section measurements if you don't know why that is). Do you see this effect here?

4.3 A model uncertainty

The uncertainties we've looked at so far have been to do with how well we might be able to measure or reconstruct particle information. But as we're looking at simulated distributions, another thing we might be uncertain about is our interaction models. Think back to the lecture on cross section measurements and how we generate Monte Carlo simulation. At one point we use the relative probabilities of the different interaction modes to decide what type of interaction we'll simulate: quasi-elastic, resonant, DIS etc. What if we don't know those relative probabilities? For the next uncertainty, we'll see what would happen if resonant events were 50% more or less likely than we expected.

1. Add one more systematic shifting class called `ResNorm` - *Resonant event normalization*. For this class:
 - There's no need to shift any variable in the standard record, so you can skip the line that does that
 - As you aren't shifting anything, you don't need a `restore` line either
 - However, you do need to change the `weight` - which corresponds to the probability of the event occurring. To shift it by 50%, `weight *= 1 + 0.5*sigma;`
 - As weights are multiplicative, coding the weighting like this can allow us to combine multiple weights.
 - Again, by changing `sigma` between ± 1 , we can shift up and down
 - Don't forget we only want to apply this weight change for resonant events! You'll need to use one of the standard record variables - look them up [here](#) (I defined `MODE_RES` for you in the file).
2. Add lines to your plot to shift the resonant probability up and down by 50%. Even though your shift in the probability was dramatic, it hasn't made a huge difference to your plot. What do you think would have happened if you'd changed the probability of quasi-elastic events instead? (Feel free to test your theory).

4.4 Effects on other distributions

You've so far been shifting muon energy - and looking at the effect on muon energy distributions. But those can affect other things as well. With `CAFAna`, you can use your shifted values (the muon energy) to calculate other things — like neutrino energy — see the effect on those.

In *Exercise3.C*, we've copied the shifts you implemented in the previous exercise. I've also added in a function for the quasi-elastic formula that we used in the previous activity to reconstruct the neutrino energy from muon kinematics. Let's see what happens to the reconstructed neutrino energy when we shift our muon energy.

1. In *Systematics3.C*, you'll find my solution to the previous exercise - run it, and you should get a plot like the one you just made. (I've removed the uncertainty due to the resonant normalization, to make it a bit easier to see what's happening).

- Let's add one more uncertainty! Code up a smearing of the muon angle, with a sigma of $\pm 30^\circ$ (check the wiki for the CAF variable for the muon angle... and don't forget to check the units). Again, I have exaggerated this effect — DUNE can do much better than this!.

What effect does this have on your muon energy distribution? Is what you see as you would expect?

- I've also added code to implement the quasi-elastic reconstruction formula for neutrino energy:

$$E_{\nu}^{QE} = \frac{m_p^2 - (m_n - E_b)^2 - m_{\mu}^2 + 2(m_n - E_b)E_{\mu}}{2(m_n - E_b - E_{\mu} + p_{\mu} \cos \theta_{\mu})} \quad (1)$$

You'll notice that this formula depends on both the muon energy and the muon angle. CAFAna has the ability to propagate the shifts in the energy and angle into the calculation of the reconstructed neutrino energy. Change *Systematics3.C* so that instead of plotting the muon energy, you plot the **reconstructed neutrino energy**, and run it to see the effects of your uncertainties on the distribution.

- How does the plot compare to your muon-energy plot from the previous exercise? What looks similar? What looks different?
- The shifts not only affect the variables we plot (Var objects) but also the Cut that decides which events to include in the plot. Look at the cuts I'm using in the spectra - I've included some code to ensure that we're only including things that seem physical. This could mean that some events, which are normally included, are cut out when certain shifts are made. Can you see any evidence of that? Some combinations of reconstructed energy and angle could produce nonphysical values of neutrino energy when entered into our formula.

5 Optional bonus: Fractional uncertainty

When we think about our uncertainties, we often want to know their fractional effect on our measured value. That means the difference between the shifted and the central value in each bin, divided by the central value. For those confident in ROOT, or ready for a challenge, try making this plot using the TH1D histograms you made in the previous exercise.

- Copy the shifted histograms, and use the histogram addition/division operations described in the [ROOT TH1 documentation](#) to calculate the fractional shifts.
- Plot all the shifts. Where does the central value appear on the plot? Where in the spectrum do you see the biggest fractional effect?
- Find a way to draw the effect of the statistical uncertainty. How does it compare?