<u>MINERvA 101 Activity, 2017</u>

Systematic Uncertainty Calculation

(Example: Charged-Current Inclusive)

Deepika Jena and Nuruzzaman

Goals of the activity

When you complete this activity, you will be able to:

- Plot a histogram of a physics quantity from one of our MINERvA ntuples
- Calculate the error bands associated some of the standard MINERvA systematic uncertainties
- Make a plot showing the fractional uncertainties associated with each systematic
- Draw your data and Monte Carlo distributions with the systematic error band included

Before you start

1. Run `kinit` and then log onto on of the Minerva virtual machines minervagpvm0x
2. You will need to be able to use ROOT.
3. Copy the setup script to your cmtuser area from: `/minerva/data/users/Minerva101/`
4. Run `source cmtuser/setup_Minerva101_2017_Sys.sh`
   This may take a minute if it's the first time you are setting up this environment. (This script is setting a new area `Minerva_v10r8p6_101`, getting `PlotUtils` and copying the files necessary for the exercises. If you don't want to use the script the relevant files are at: `/minerva/data/users/Minerva101/SystematicsActivity2017`)
5. You should have all the packages and files needed on your area now.

Exercise 1: Plot a histogram of muon energy

For this activity, we will be using ntuples from the charged-current inclusive study (that is, an analysis that looks at all neutrino-nucleon interactions producing a charged lepton). We will be using interactive ROOT to analyze the data and produce histograms.

We will be looking at ntuples from two files:

- Monte Carlo file:
  `/minerva/data3/users/minervapro/mc_ana_minerva1_Dec04_v10r6p13_ccinc_CCCohPion_CCPionInc/grid/central_value/minerva/ana/v10r6p13/00/01/02/00/SIM_minerva_00010200_Subruns_0001-0002-0003-0004_CCInclusiveReco_Ana_Tuple_v10r6p13.root`
- Data file:
  `/minerva/data/users/minervapro/New_CCInclusiveReco_POTbugFixed/data_minerva1/grid/minerva/ana/numibeam/v10r6p13/00/00/20/00/MV_00002000_Subruns_0001-0002-0003-0004-0005_CCInclusiveReco_AnaData_Tuple_v10r6p13.root`

The branches we need are in the tree **`CCInclusiveReco`**

1. Open the data file in ROOT:
   `root -l <datafile.root>`

2. Open a TBrowser
   `new TBrowser`
   Locate the CCInclusiveReco tree and try to work out which variable represents the muon energy.

   The muon energy variable is ____

   Note: the muon energy variable is represented as a 4-vector, containing 3 momentum components and an energy component. In terms of code, that means that it is a double [4] and its components correspond to $\{p_x, p_y, p_z, E\}$. You need to use the [3] component of this array to get the muon energy.

3. Use the ROOT command line to make a histogram of the muon energy.

   What would be good limits for the histogram? _______________ to ______________

   What do you think the units are? ____

   Use the "`E`" drawing option to draw the plot with error bars. Currently, these only include statistical uncertainties.

   What formula relates statistical uncertainty to number of events?_________________

   Does that formula seem to hold? ____

4. Now open the Monte Carlo file and plot the muon energy.

   Compare number of events and statistical uncertainty between data and MC.

Exercise 2: Calculate some systematic error bands

Start by taking a look at the folder `/SystematicsActivity/` in your area, you should have the templates and solutions for this and the next exercises.

Building executables

Your code for this exercise is all in files with the extension .cxx - C++ source files. In order to build these into executables, type the command make. This will generate an executable for each source file, whose name will be the same as the source file, but without the .cxx extension. For example, the code in the file `Systematics_Template2.cxx` can be run by typing `./Systematics_Template2` at the command line. The instructions to the C++

compiler for making these executables are in the file `Makefile` - but we don't need to worry about that for this exercise.

The template file

The template file should generate a central value histogram and create a single error band.

1. Execute the template program in ROOT.

   ```
   ./Systematics_Template2
   ```

   It should return the number of protons on target (POT) for your data and Monte Carlo files. This number is proportional to the number of incoming neutrinos corresponding to each file, and we will use it later to make a comparison of data and Monte Carlo.

   Data POT ______________________ Monte Carlo POT _______________________

   Do these numbers appear to correspond with the relative sizes of the distributions you saw in exercise 1?

2. The template file should have produced an output file called `mySystematics.root`. Open that file in ROOT and look at it in a `TBrowser`.

Identify:

The data histogram ____

The central-value Monte Carlo histogram _____________________________________

An error band with 100 universes ____

Is the error band on the data or the MC ? _____________________________________

Is it a vertical or a lateral error band? ______________________________________

Use the methods GetNVertErrorBands and GetNLatErrorBands from the MnvH1D object to help you answer this.
http://nusoft.fnal.gov/minerva/minervadat/software_doxygen/HEAD/MINERVA/classPlotUtils_1_1MnvH1D.html

3. Now open the template file Systematics_Template2.cxx in your favorite text editor. This code initializes the variables we will load from the ntuple, and then loops through the Monte Carlo ntuple, filling the central value histogram and each of the weighted universes. It then loops through the data ntuple, which it fills with just a central value. Identify where in the template code we:

Declare the branch object for the muon energy _________________________________

Declare the branch object for the 100 universe weights ___________________________

Declare a variable into which the energy will be loaded ___________________________

Tell the ntuple to link the branch to the variable _________________________________

Create an error band to hold the 100 reweighted histograms ______________________

Loop through the Monte Carlo events ____

Fill the central value histogram with the central weight ____________________________

Fill the error band with the 100 shifted weights __________________________________

4. Now try to add a flux beam focusing error band with 100 universes. The ntuple already contains the weights for this error band, and the correct branches have been loaded in the template.

5. Build your executable by running make (correct any syntax errors discovered by the compiler until it builds) and then run it with ./Systematics_Template2. Once you run the code, check again in `mySystematics.root` to see that your new error band plots appear.

6. **Optional:** add the flux NA49 error band with 100 universes. This time, you will need to look in the ntuple to find the name of the branch to load, and you will have to load the branch yourself before filling the error band.

7. If you get stuck, you can look at `/SystematicsActivity/Systematics_Solution2.cxx` for a complete solution to exercises 2 and 3.

Making a lateral error band

The information we need to calculate the muon MINOS energy shift uncertainty is also saved in our ntuple. This time, instead of looking at 100 event weights, we want to look at what happens when we shift the muon energy – the quantity we are plotting – up and down by a given amount. This corresponds to an uncertainty in how well we can measure this energy using the MINOS detector. In this case, we will have two universes – one where we shift the energy up, and one where we shift it down. Because we are shifting the quantity that we are plotting, this is a lateral error band.

1. Continue to work with the code you used to make the vertical error bands. If you had problems with that and want to try calculating just the lateral error bands separately, go back to the template file `Systematics_Template2.cxx`

2. The branch we need to use to calculate these energy shifts has already been declared in the template.

   Its name is _________________________________of type ______________________

3. The CCInclusiveReco_sys_muon_energy_shift branch has two components, corresponding to a downward energy shift ([0] component) and an upward energy shift ([1] component). Use these components as the shiftDown and shiftUp to fill a lateral error band, using the FillLatErrorBand method http://goo.gl/tnKRuu . Make sure you are filling correctly the error band.

Remember you will have to create your error band before you can fill it.

The error band should have ____________ universes

4. Once you build (make) and run the code, check again in mySystematics.root to see that your new error band plots appear. Your error band should consist of a central value and 2 shifted histograms. Use the MnvH1D::Print("ALL") method to print the values in the central value and the two shifted histograms. Note the values in bin 10:

CV histogram ____________ Universe 0 ___________ Universe 1 ____________

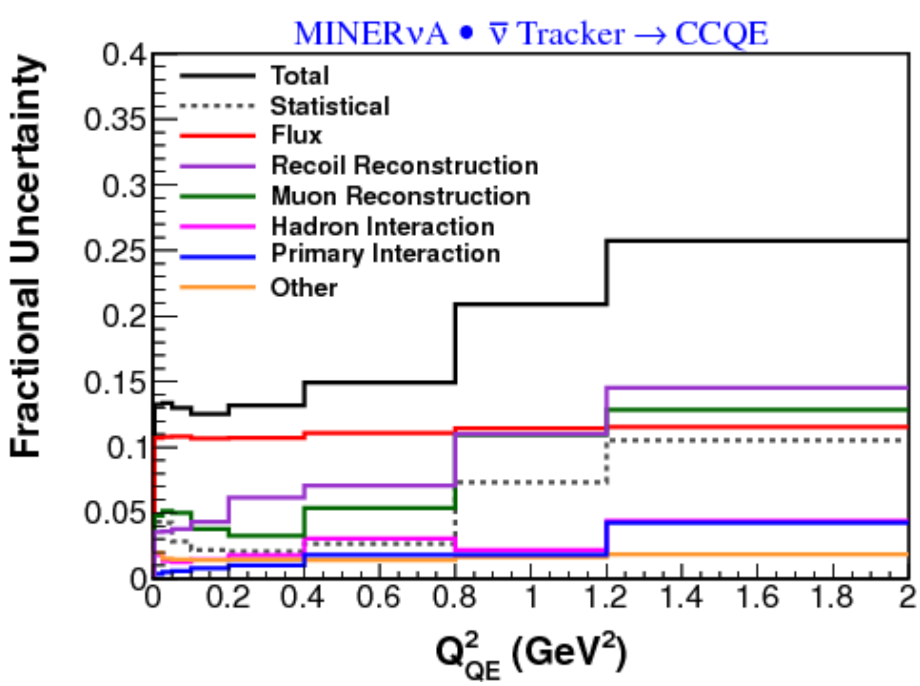
What is the uncertainty due to muon energy shift in bin 10? ___________________

What is the fractional uncertainty due to muon energy shift in bin 10? ___________

Exercise 3: Plot your results

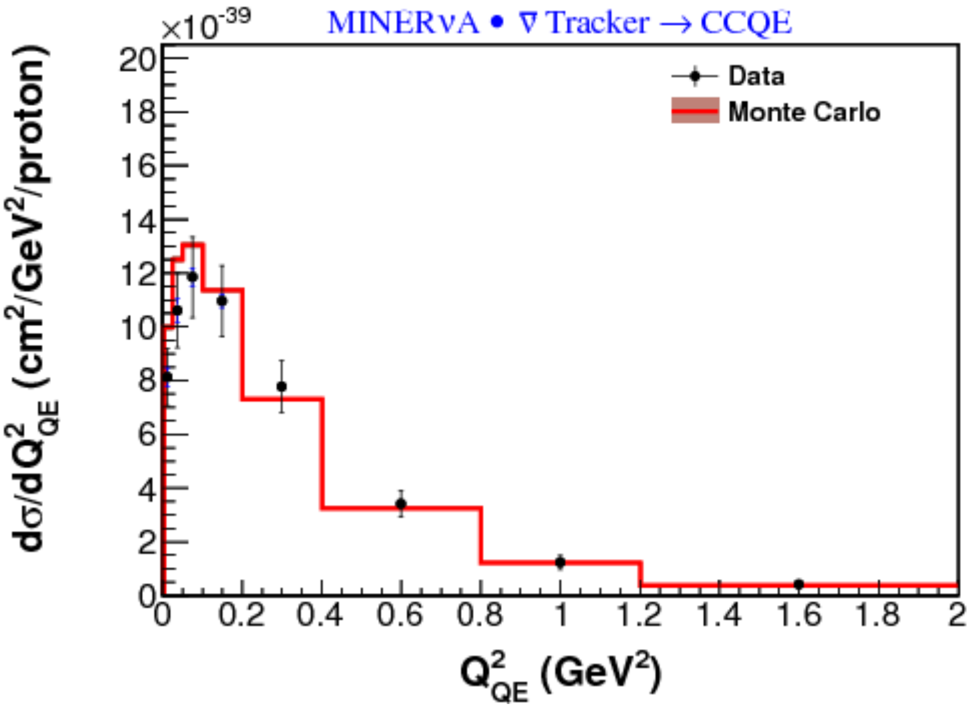
In this exercise, you will make two important plots to demonstrate the systematic errors you calculated in Exercises 2 and 3. The first of these will show the fractional uncertainty due to each systematic you calculated. The second will compare data and Monte Carlo, with full systematic error bands.

Plot 1 – fractional uncertainties



The example above shows a plot of fractional uncertainties for the charged-current quasi-elastic antineutrino analysis. You're going to make a plot like this showing the uncertainties you just calculated.

Plot 2 – data and Monte Carlo comparison



The plot above shows data and Monte Carlo cross-sections with uncertainties for the charged-current quasi-elastic antineutrino analysis. We will also be making a plot like this for our own sample.

Again, we will start from a template ROOT file:
`/SystematicsActivity/Systematics_Template4.cxx`

The plotting code takes as an input the mySystematics.root file you made in examples 2 and 3. You can either use your own, or you can use the original one we made at:
`/minerva/data/users/Minerva101/SystematicsActivity/mySystematics.root`

1. Take a look at the Systematics_Template4.cxx file and make sure you understand the reason for the code in the template. Try running the template code:
   `./Systematics_Template4`

2. Create your fractional uncertainty plot using the `DrawDataMCErrorSummary` method of MnvPlotter. Use this documentation to help you: http://goo.gl/FX4zUp and save your result in multiple image formats using the MultiPrint method. Remember you need to make the executable before you run it, in order to pick up your changes.

3. Create your data and Monte Carlo error summary plot using the DrawDataMCWithErrorBand method of MnvPlotter. Use this documentation to help you: http://goo.gl/s5MYfX and save your result in multiple image formats using MultiPrint.

4. Look at your output images. To view a png file on the Minerva virtual machines from the command line, try: display <myImage.png>

5. In the last exercise, you calculated the fractional uncertainty in bin 10. Look at your plot of fractional uncertainties. Does the plot agree with what you calculated?

Fractional uncertainty in bin 10 from the plot ________________________________

In the data/MC comparison plot, look at the values in bin 10 for the data, Monte Carlo and for the uncertainties on each. Do they match the fractional uncertainties from the other plot?

Data value in bin 10 ___________ uncertainty ____________ fraction ___________

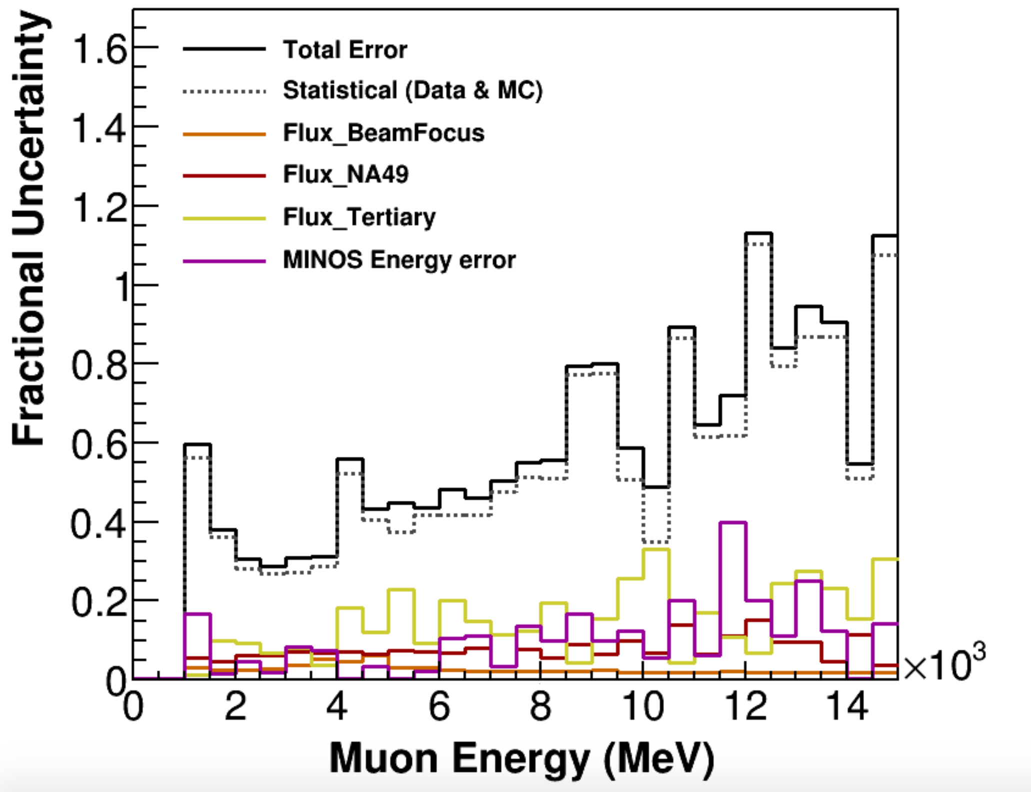
MC value in bin 10 ___________ uncertainty ____________ fraction ____________

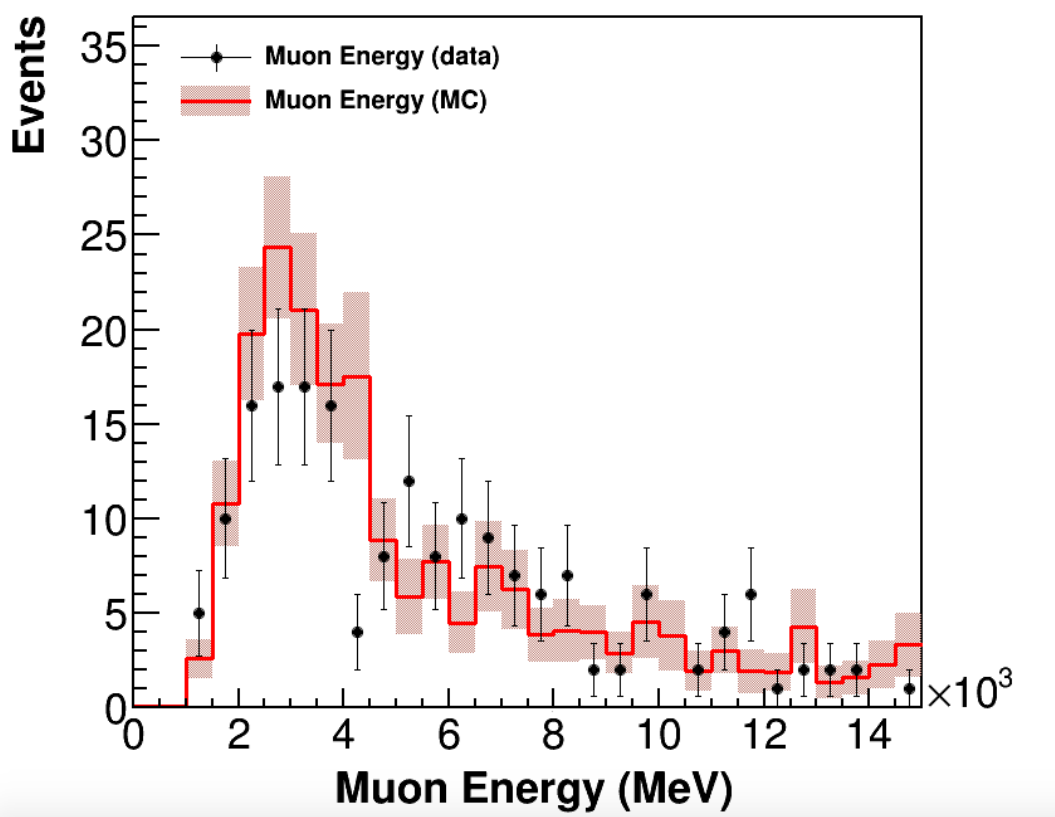
6. You may notice that your data and Monte Carlo give very different distributions. That is because the input files you are using are for different numbers of protons on target (POT). Using the POT values printed out in exercise 2, scale ("POT normalize") your Monte Carlo so that it represents the same number of POT as the data. How is the agreement now?

Does the POT normalization change the fractional uncertainty plot? _____________

7. If you want help or get stuck, look at Systematics_Solution4.cxx

8. **Optional**:
   a. Try making a data-Monte Carlo ratio plot
   b. The SetupStyles function does some basic formatting of the plot. Try changing the formatting to get a style you like, using the MnvPlotter documentation to help you.
   c. Add a "POT-normalized" box to your data plots
   d. Try making a shape-only uncertainty plot


Fractional Uncertainty
Total Error
Statistical (Data & MC)
Flux_BeamFocus
Flux_NA49
Flux_Tertiary
MINOS Energy error
×10^3
Muon Energy (MeV)


Events
Muon Energy (data)
Muon Energy (MC)
×10^3
Muon Energy (MeV)