Data Analysis Tutorials

Introduction

The following exercises will take you through the various steps of analysing data in nuclear physics. The tutorial will focus on the use of c++ ROOT, however, for individual steps you are welcome to use pen & paper, excel, radware, python etc, though ROOT use is encouraged.

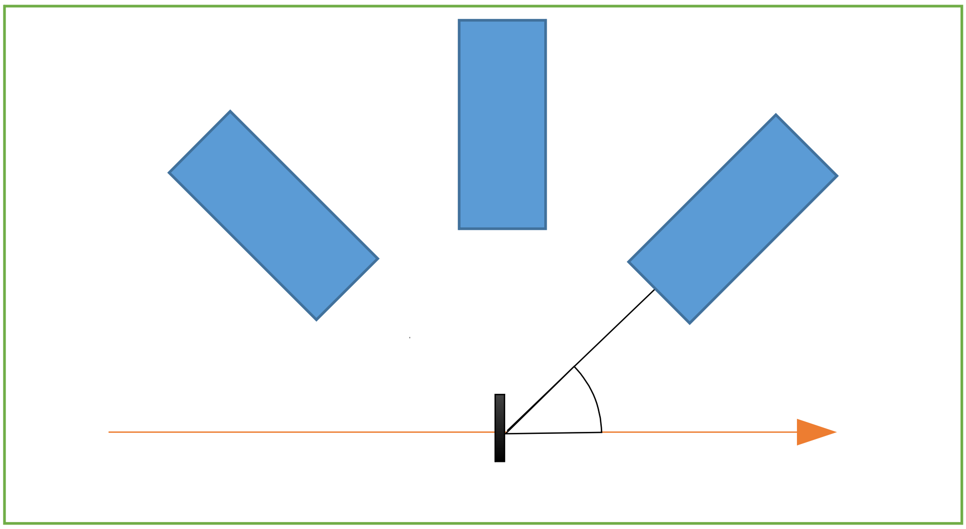
The majority of tasks in ROOT can be accomplished with the Graphical User Interface supplemented by entering commands on the command line. Or by executing a series of commands in a macro script .C file. Using the [tab] autocomplete functionality on the ROOT command line is an effective way of checking available functions and required inputs. The online ROOT class documentation is another invaluable source of information:
https://root.cern.ch/doc/master/classTH1.html
https://root.cern.ch/doc/master/classTF1.html

Do utilise the instructors and each other.

To get the most from these exercises, don’t read too far ahead.

The Experiment

Data are taken with an array consisting of 3 gamma-ray detectors positioned around a target position on an accelerator beam line. The 3 detectors are positioned, with respect to the target and beam, at angles of 45,90,135 degrees.



Data are stored in root “trees”, effectively long lists of individual events; an event being a single moment in time, in which the data acquisition computer detected an input and recorded all data from detectors during a short time window.

A series of .root data files are provided from the experiment:

Run1 – A source of ^{60}Co at the target position
Run2 – A source of ^{152}Eu at the target position
Run3 – A source of ^{133}Ba at the target position
Run4 – An A=36 beam with an energy of 250 MeV impinging on a thin A=40

target.
Run5 – A decaying source feeding the nucleus observed in Run4

Important note :
Before attempting to open the run files in a root session, run the command ".L Det.h" to load the data format library for the experiment. This is achieved with an include statement in the example scripts.

Exercise 1 – Look at the Data

Before moving on to complex tasks ensure you can access, view and sort the events. This can be done using the GUI by creating a instance of the TBrowser class on the command line and then navigating to the desired root file. Data "TTrees" can be interacted with in the TBrowser, or in more detail by right clicking and starting a TTreeViewer (StartViewer). Histograms and graphs in a TFile can be viewed directly in a TBrowser by double clicking.

A) How many events are there in Run1.root ?

…………………………………………………………………..

For more complex sorting using a script is advisable. More advanced users might prefer to use a TSelector, but this is not needed or recommended for beginner/intermediate users. An example sorting script is provided. Load it with the following command (within root)

```
.L ExampleSort.C
```

Then execute the sorting function with :

```
ExampleSort("RunFile.root")
```

Output histograms are saved to a new .root file which can then be viewed in a TBrowser.

B) Create and view a 1D histogram of the charge from the detector with ID=0 in Run1.root with an appropriate range and binning.

C) What is the integral of counts between 0 and the highest point of the spectrum?

…………………………………………………………………..

D) Over what range of time do the events of Run1.root occur?

…………………………………………………………………..

Exercise 2 – Calibrate Detectors

Detectors actually record an induced charge which is proportional to the incoming energy. Due to differences in detectors and amplifiers all channels in an experiment must have the conversion between charge and energy determined from data of known energy.

1. Search online for literature giving the energy of gamma radiation resulting from the decay of ^{60}Co

2. Create and fill and appropriately binned histogram of the charge detected by each detector in Run1
3. Looking at the charge spectra you have created, identify the largest peaks. The charge centroid of these peaks must correspond to the energy of the most intense gamma rays expected from the decay.
4. Determine the centroids of the peak by fitting the peak with a Gaussian. This can be done with the FitPanel GUI, on the command line or with a script such as ExampleFit.C
5. Calculate initial calibration parameters for the detectors.
6. Produce a histogram of energy in keV

E) What is the energy resolution of the detector array (combined spectra of all channels)?

………………………………………………………………………………

7. Sort an energy spectrum of Run2.root

F) What is the energy resolution of the 122 keV?
Give 2 reasons for the change from your answer to E.

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

Exercise 2 Extra

Use the ^{152}Eu data from Run2.root to improve your calibration.

Exercise 3 – Determine Array Efficiency

Using your ^{152}Eu sum spectra determine the <u>relative</u> efficiency of the array as a function of energy.

There are many possible functions to describe the efficiency curve of a gamma-ray detector (the shape of which is dominated by gamma-ray attenuation coefficients) the following is an example you can use which converges easily (some functions such as those used by radware and sigma require extra effort to ensure a successful fit). Start with N=2

$$y = e^{\left(\sum_{n=0}^{n=N} P_n log(x)^n\right)}$$

Whether determining peak areas by integration or by fitting, be sure to determine appropriate errors. It is recommended you make use of the TGraphErrors root class:
https://root.cern.ch/doc/master/classTGraph.html
https://root.cern.ch/doc/master/classTGraphErrors.html

G) What is the efficiency ratio between 300 and 1200 keV?

………………………………………………………………………………….

H) What is unphysical about this initial function?

…………………………………………………………………………………….

…………………………………………………………………………………….

…………………………………………………………………………………….

…………………………………………………………………………………….

EXTRA QUESTION 1) If the time units of the data are ns, and the activity of the ^{60}Co source is 5.4 KBq, what is the absolute efficiency at 400 keV?

…………………………………………………………………………….

ROOT TF1 Fitting Tips:

Fitting often requires setting the parameters of the TF1 to initial values. If you have no idea simply set a parameter to zero.

If the fit of a TF1 results in “STATUS=CALL LIMIT” you can simply run the fit multiple times until it converges or fails (you can also change the call limit but this can have knock on issues).

You can only calculate the full correlated errors if a fit result displays “ERROR MATRIX ACCURATE” if instead you get “ERR MATRIX NOT POS-DEF” you can try setting limits on parameters in a narrow band around their optimum values or fitting with the extra “E” option.

Exercise 4 – Determining Fit Uncertainties

Choose one of your efficiency functions parameters and produce a graph of the fit chi squared statistic as a function of the parameter (hint: Use the TF1 FixParameter function).

Adjust the range of parameter space until you can see the minimum chi squared value + 2

I) What shape does the graph form?

…………………………………………………………………………………….

J) What chi squared value corresponds to the parameter upper and lower error bounds as determined by the initial root fit?

…………………………………………………………………………………….

This relation between chi-squared and error is true for a Normally (Gaussian) distributed parameter. If you repeat the process for say a peak area (which is exactly Normally distributed) you would get a more precise result. This relation can be used to determine parameter errors from chi squared in situations where you are not simply fitting a function, such as adjusting a parameter in a simulation.

Do determine the uncertainty on the entire efficiency curve, not just a parameter we can apply the error formula:

$$V(f) = \sum_j \left(\frac{\partial f}{\partial x_j}\right)^2 V(x_j) + \sum_j \sum_{k \neq j} \left(\frac{\partial f}{\partial x_j}\right) \left(\frac{\partial f}{\partial x_k}\right) \mathrm{cov}(x_j, x_k)$$

ROOT’s TMinuit fitting functionality can do this for you. It is also useful to note you can extract covariance matrix elements and and gradients from a TF1 fit so that you can apply the above formula to a function that depends on a subset of parameters from your fit results, such as the uncertainty on the background of a background+peak fit. Note however the error/covariance matrix is not stored in the TF1 after a fit in same way parameters and their errors are, if you wish to access the off-diagonal matrix elements this must be done immediately following the fit.

For now simply use the following command (directly after fitting) to evaluate the uncertainty of your efficiency function at a series of points:

```
(TVirtualFitter::GetFitter())->GetConfidenceIntervals(hist,interval);
```

“hist” should be a pointer to a histogram with binning defining the points at which to calculate (a TGraphErrors can also be used as per example https://root.cern/doc/master/ConfidenceIntervals_8C.html) and interval should be a number giving the desired central probability interval.

Now the histogram has the fitted function values as the bin contents and the confidence intervals as bin errors.

Retrieve the data and plot the error bands alongside the ^{152}Eu data

K) What is the % uncertainty of the array efficiency at 1000 keV?

..

Exercise 4 Extra

Use the ^{133}Ba data from Run3.root to improve your calibration around the low energy turning point.

You will either need to extend your previous function to N>=3 or use an alternative functional form such as:

$$y = P_1 \cdot 10^{\left(P_2 log_{10}(x) + P_3 log_{10}(x)^2 + P_4/x^2\right)}$$

Exercise 5 – Kinematic Correction

Sort Run4.root, which is data from a simulated experiment in which an A=36 beam with an energy of 250 MeV impinges on a thin A=40 target.

You should immediately see the spectrum is worse than those seen before.

L) Why is the spectrum poor?

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

M) Why is there one narrow peak observed?

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

In such an experiment the energy is sufficient to overcome Coulomb repulsion and the fusion of the beam and target occurs. A small number of protons and neutrons are subsequently evaporated (a fusion-evaporation reaction). This leaves an excited compound-like nucleus which subsequently gamma-decays.

Using the information and data available, apply a correction to the data and produce and improved spectrum.

We wish now to explain every feature observed under 3 MeV in the singles spectra. You may attempt to begin now, but will need to follow the next exercises to complete this.

Exercise 6 – Coincidences

As multiple hits may occur approximately simultaneously, such as in a cascade, it is advantageous to look not only at the singles spectrum but also coincident hits in multiple detectors.

Produce a 2 dimensional histogram (sometimes refereed to as a matrix) of coincidence events. Be sure to diagonalise this as we have not specified anything special about any particular hit.

Now by projecting along one axis of this histogram at a given energy, you can see all the gamma-rays that were coincident with a gamma-ray of that energy.

Use the TH2 functions: ProjectionX, ProjectionY, SetShowProjectionX and SetShowProjectionY

The latter 2 being useful for looking at multiple projections in the GUI by moving the mouse. Beware as many arguments are taken in bin number, which depending on your chosen binning may not have a 1:1 correspondence with energy.

Look at the coincidences of the largest peaks

N) What coincidence do you observe that is surprising?

…………………………………………………………………………………….

……………………………………………………………………………………

Use the function Hit::GetTime() and produce a 1D histogram showing the time between coincident hits.

O) Explain the features of the time spectrum?

…………………………………………………………………………………….

……………………………………………………………………………………

……………………………………………………………………………………

……………………………………………………………………………………

……………………………………………………………………………………

From the timing spectrum determine a suitable timing coincidence requirement (a time gate).

P) Calculate the fraction of events selected by your requirement that are true coincidence events?

……………………………………………………………………………………

By making the reasonable assumption that false coincidences are randomly distributed with time, you should be able to determine a time-random contribution that you can subtract.

Now by looking at what transition are coincident (after subtracting your calculated background) attempt to construct a level scheme for the compound-like nucleus produced in the reaction.

Q) At what energy is the isomeric state?

……………………………………………………………………………………

EXTRA QUESTION 2) Using your efficiency curve, explain the features at ~2000 keV?

…………………………………………………………………………

Exercise 7 – Lifetimes

Sort the data from Run5.root, this simulated data set is from a decay run, hence there is no kinematic correction. Additionally a lower rate means the false coincidences are significantly fewer (i.e. can be neglected).

Using this higher-resolution data, confirm your level scheme and calculate branching ratios using your efficiency curve.

U) What is the nucleus? (note state is not isomeric in real nucleus)

…………………………………………………………………………

EXTRA QUESTION 3) What other change was made to the nucleus in the simulated data?

……………………………………………………………………………………

……………………………………………………………………………………

By looking at the time difference of the levels populating and de-populating the isomeric state you should be able resolve the lifetime of the state.

By fitting the isomeric tail of the timing distribution determine the lifetime.

V) What is the lifetime of the isomeric state?

……………………………………………………………………………

Try fitting with the “L” likelihood fitting option rather than the chi-squared default. Do you notice a difference? (the significance of the effect will depend on your binning). At very low counts the Pearson’s chi squared test becomes unreliable and binned maximum likelihood method is preferable for an accurate result.

Plot the same time spectra but in log(t) (with histograms bins uniformly spaced in log(t) instead of t) a peak should form with the maximum at the state lifetime. You can fit this with a Gaussian and see how it compares to the exponential method. This method (Often called the Schmidt method) is extremely useful in low statistics work where a reliable exponential fit is not viable.

EXTRA QUESTION 4) Assuming all transitions are pure E2. Correcting for efficiency. What is the B(E2) value of the highest energy transition in the nucleus

……………………………………………………………………………

EXTRA QUESTION 5) Calculate the relative population cross-sections of the different states of the nucleus.

……………………………………………………………………………

……………………………………………………………………………

……………………………………………………………………………

……………………………………………………………………………

……………………………………………………………………………

……………………………………………………………………………

……………………………………………………………………………