ROOT Hands-on Analysis Exercise for Students

Brian Krar

EIEIOO

May 11th 2020

“Give a man a fish, you feed him for a day. Teach a man to code, he’ll die hungry and alone” – Sun Tzu, probably

Now It's Your Turn!

After a certain point, there's only so much you can learn by reading how to code and watching others (me) try to code

- In order to understand ROOT, you must use ROOT

Now that you've had some exposure to ROOT – through lectures and the hands-on tutorials – it's time to try doing a physics analysis yourself

Attached is a practical data analysis exercise meant to test/apply your skills

- This is divided up into 5 subsections, and will use several of the concepts you've seen earlier
- It is meant to help you get more comfortable with ROOT by working on an analysis somewhat comparable to something you might see this summer

If you feel like you are unable to complete certain parts, brush up on ROOT some more by going over the ROOT references/tutorials and trying the exercises from the previous session

Guided Exercise #1a – HEP Based Problem: H-> γγ

Problem: We want to perform a search for H -> γγ using the invariant diphoton mass
The H -> γγ channel has a tiny branching ratio (10^{-3}), but this is very pure and has a great mass resolution (due to good energy resolution and a distinct signature of the two photons)

Attached you should find three files:

PseudoData_Histogram_100fb.root, Signal_1fb.root, and Background_1fb.root

- You'll need to upload these to your guest directory on neutrino

Part 1: Plot the diphoton mass data

PseudoData_Histogram_100fb.root is the measured data corresponding to 100 1 fb^{-1}
Inside the data file is a TH1D histogram called "signal", which shows the invariant diphoton mass

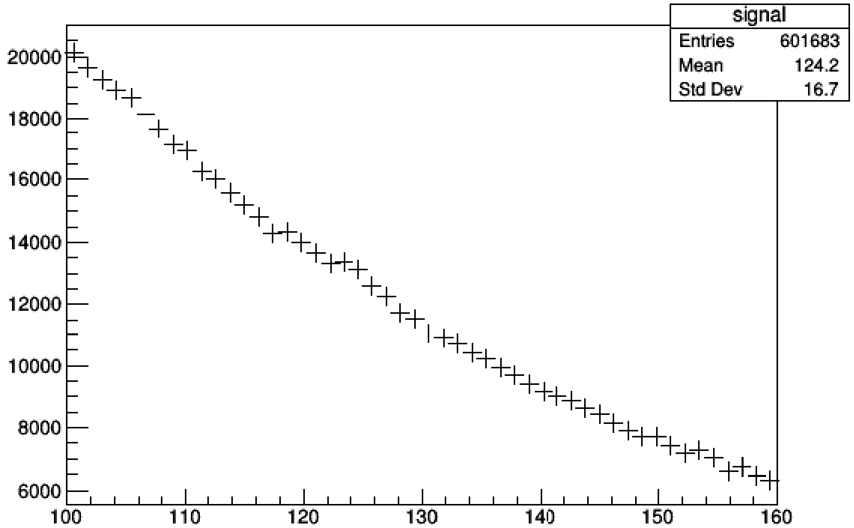
- **Plot the "signal" histogram on a canvas**
- **Find the number of events in the data**

Guided Exercise #1a – Answer: H-> γγ

Hints:

- **To read in the data, recall the functionality of "TFile"**
- **To get the number of events in the data, recall "Integral()"**

Part 1a Output: The Signal Histogram:



Guided Exercise #1b – HEP Based Problem: H-> γγ

Problem: We want to perform a search for H -> γγ using the invariant diphoton mass
The H -> γγ channel has a tiny branching ratio (10^{-3}), but this is very pure and has a great mass resolution (due to good energy resolution and a distinct signature of the two photons)

Among the Provided Files:

Signal_1fb.root and Background_1fb.root are the signal and background MC simulations corresponding to 1 fb^{-1} (note that this is 0.01 of the data, so the MC must be re-scaled accordingly)

- Inside these simulation files, you'll find a TTree is stored called "tree", which contains two variables: invariantMass and eventWeight (the invariant diphoton mass and the event weight, respectively)

Part 2: Create MC Histograms from Trees

- **Create histograms and fill them with the invariantMass weighted by the eventWeight (Hint: use an event-by-event loop)**
- **Scale these simulation histograms to the correct integrated luminosity of the data (so by x100)**
- **Find the number of signal and BG events before and after the rescaling**
- **Compare the data to the background-only hypothesis (Plot only the reweighted BG and the Data)**

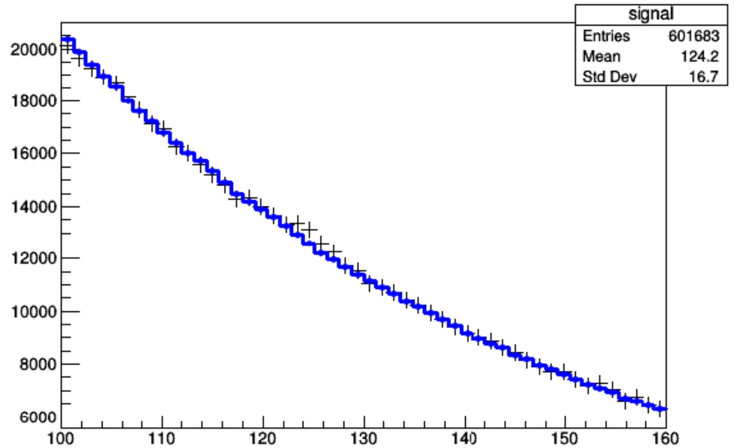
Guided Exercise #1b – Answer: H-> γγ

Hints:

- **Inside the trees, there are two variables:**

  `invariantMass and eventWeight`

The Signal Histogram and Reweighted BG (in Blue):




```
Before reweighting:
Signal:          16.8333
Background:      6000
After reweighting:
Signal:          1683.33
Background:      600000
```

Guided Exercise #1c – HEP Based Problem: H-> γγ

Problem: We want to perform a search for H -> γγ using the invariant diphoton mass
The H -> γγ channel has a tiny branching ratio (10^{-3}), but this is very pure and has a great mass resolution (due to good energy resolution and a distinct signature of the two photons)

Part 3: Perform Fit to the Data with the BG Only ("BG-only" hypothesis test)

To neglect fluctuations, we want to fit the background now to get a good background model

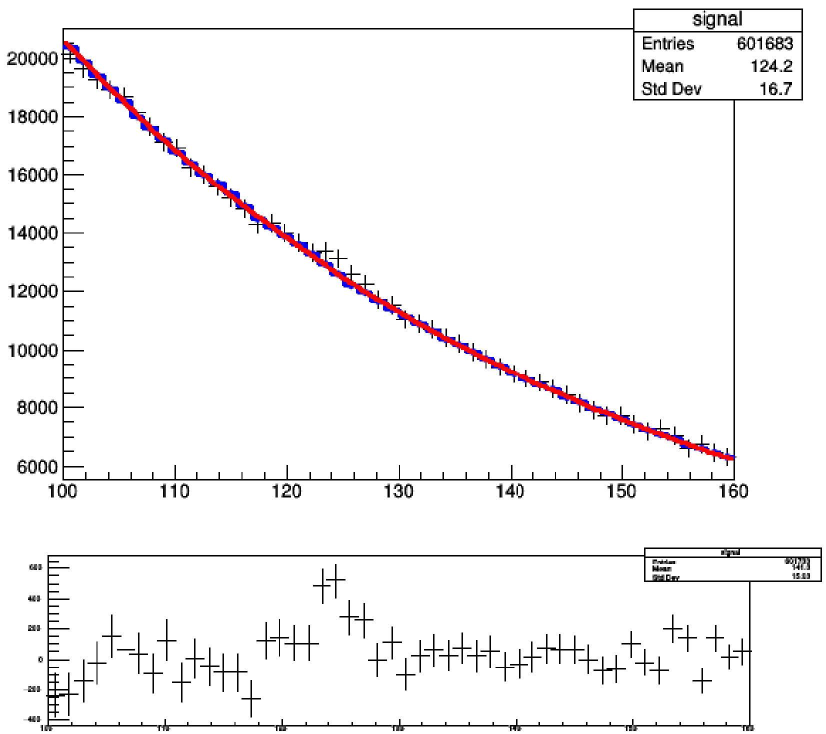
- **Perform a fit to the background**
- **Compare the data with the background model by plotting the difference between them in a sub-plot below the main plot**

Guided Exercise #1c – Answer: H-> γγ

Hint:

- **What would be a good choice for a model to the Background?**
  - The background looks very exponential (hint,hint)

Sample Output:



Guided Exercise #1d – HEP Based Problem: H-> γγ

Problem: We want to perform a search for H -> γγ using the invariant diphoton mass
The H -> γγ channel has a tiny branching ratio (10^{-3}), but this is very pure and has a great mass resolution (due to good energy resolution and a distinct signature of the two photons)

Here is where you get to show off you plotting skills

Part 4: Finalize the plot (BG Only Model Fit)

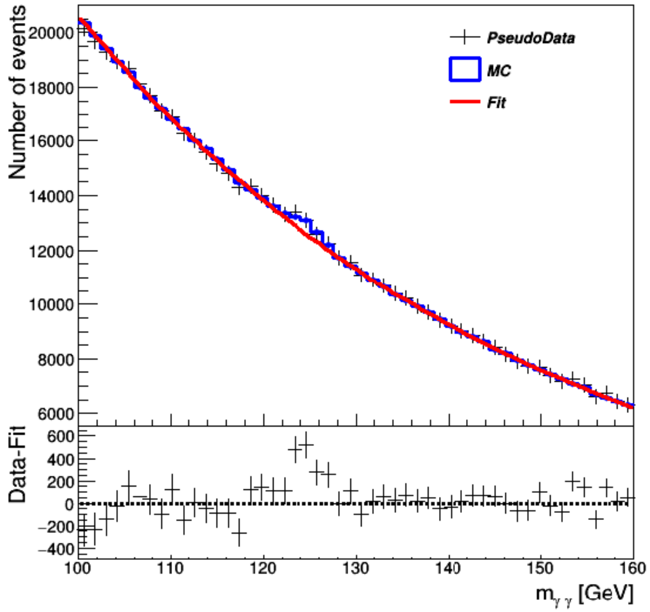
- **Add the signal MC histogram to the background histogram**
- **Make your plots look *nice***

Guided Exercise #1d – Answer: H-> γγ

Hint:

- **Recommended Features**
  - Axis labels
  - Plot Label
  - Legend
  - Dotted line through 0 in the Data-MC plot (for readability)
  - Adding parameters from BG Fit to Stats box
- **Don't let the best be the enemy of the good!**

Sample Output:



Guided Exercise #1e – HEP Based Problem: H-> $\gamma\gamma$

Problem: We want to perform a search for H -> $\gamma\gamma$ using the invariant diphoton mass
The H -> $\gamma\gamma$ channel has a tiny branching ratio (10^{-3}), but this is very pure and has a great mass resolution (due to good energy resolution and a distinct signature of the two photons)

Part 5: Final Analysis – Full model (Perform a signal + background fit)

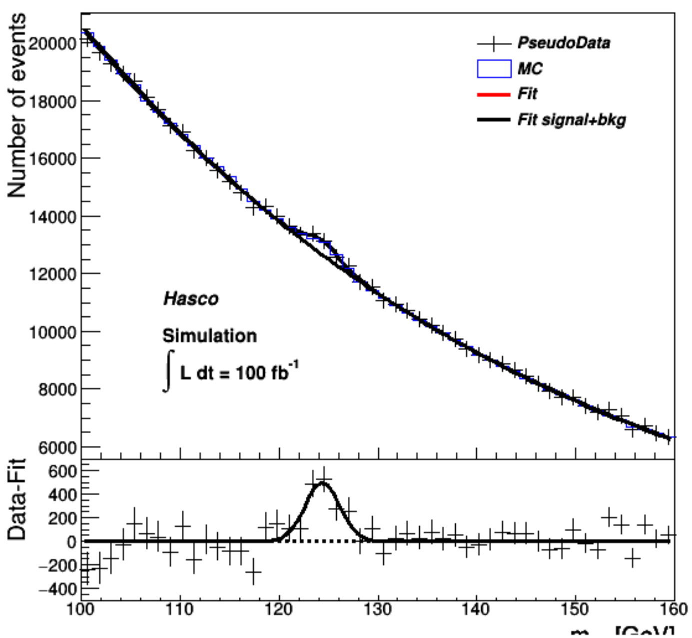
- **Perform a signal + background fit**
  - For stable fit behaviour, set the initial fit values of the background part to those obtained from the previous fit, and the signal ones to values that seem sensible
- **From the signal component of this fit, determine the number of signal events, and compare to the number we expect from the signal simulation**
  - The ratio extracted/expected is called the signal strength
- **Finalize the plot (Sig+BG Model Fit)**
  - Add the signal simulation to the background histogram.
  - Plot the signal component of the signal + background fit in the sub-plot

Guided Exercise #1e – Answer: H-> γγ

Hint:

- **For the minimizer to yield reasonable fit values, logical initial estimates (guesses) on the parameters need to be provided**
  - Based on your graph from part d., what would make sense to use as best guesses on the gaussian mean and sigma of the peak? (Hint – also think about the fit range)

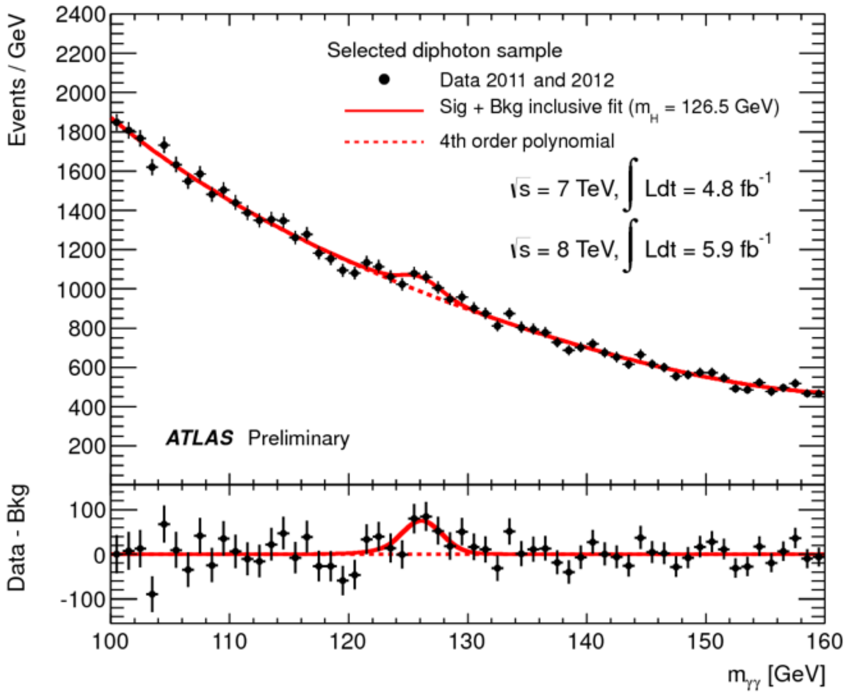
Sample Output:



So… What Does This All Mean?

Congratulations! You just found the Higg's!*

- **If you've completed this exercise on your own, then you should be pretty well equipped to tackle the projects you will be working on this summer!**
- **Of course, this doesn't cover everything ROOT has to offer. So keep practicing, and don't be afraid to look up specific functionality as needed**



***Well sort of. In reality, the actual analysis was a bit more complicated!**