Introduction to Computational Physics (PHYS105)

Lecture 3: Functions and Graphs

Carl Gwilliam
(C.Gwilliam@liverpool.ac.uk)

- **RECORD + Attendance Code!!**
- Morning everyone
  - **Please don't forget to register your attendance**
- How did you find last week's problems class?
  - Is it starting to make sense? Would anyone like to give any input?
- This week we are going to discuss functions in python and go more into detail on plotting

Lecture 2: Recap

- Last week we learnt about various python data structures
  - Both sequences (index) and mappings (key:value)
- Built-in structures
  - Lists `[]`: mutable sequences of variables of any type
  - Tuples `()`: like lists but immutable (can't change elements)
  - Strings `""` or `''` : immutable sequences of characters
  - Dicts `{}`: mapping of keys to values
- NumPy arrays
  - More efficient structures for numerical calculations on the same type of value in 1 or more dimensions
  - Can be used to read data from file and write back out
- All have similar syntax for accessing/changing the elements
  - Can use negative indices to count from the back
  - And slices to get a range of elements [lo:hi)

```
# List
mylist = [1,"two", 3.0]
print (mylist[1])
print (mylist[0:2])

# Tuple
mytuple = (1, "two", 3.0)
print(mytuple[-1])

# Dict
mydict = {"one": 1,
          "two" : "deux",
          "three" : 3.0}
print(mydict["one"])

# Numpy arrays
import numpy as np
mynp = np.array([[1,2,3], [10,20,30]])
print(mynp[1,2])
```

```
two
[1, 'two']
3.0
1
30
```

2

- Before we do so, let's recap last week, where we …
  - **Sequences = set of values referred to using position, starting from 0, as an index**
  - **Mappings = instead of an index, relate a key to a value (like a directory)**
- **Refer to example as go through**
- List: assign elements with square brackets; Tuple: assign elements with round brackets
- Saw strings were also data structures containing an immutable sequence of characters in quotes (single or double)
- Dict: created with with curly brackets and contain key:value pairs (with a colon) separated by commas
- **Due to flexibility of mixed types these quickly get slow** -> Use Numerical python -> numpy arrays (very like lists)
- Sequences: Access using index starting from 0.
  - Can use –ve index to go from back or slice lo:hi to give range; **In 2D just use two indices (either in separate square brackets or separated by a comma**
- **Mappings: same except use the key instead**

Lecture 2: Formative problem solutions

- Exercise 1: Print last item of shopping list then add two missing items to it

```
# Create list
shopping = ["cheese", "butter", "bread", "apples", "oranges"]

# Print last item (-1 easier)
print("Last item = ", shopping[-1])

# Make list of missing items
missing = ["chocolate", "wine"]

# Can concatenate lists using '+'
# (just like we saw with nums/strs)
shopping = shopping + missing
print(shopping)
```

```
Last item =  oranges
['cheese', 'butter', 'bread', 'apples', 'oranges', 'chocolate', 'wine']
```

  - For the last item could use `4` or `-1` as index
  - For lists the `+` operator performs concatenation
    - Could also have used `append()` to add items one-by-one or `extend()` to add new list
  - A few people mistakenly created a list of lists by

```
print([shopping + missing])
```

```
[['cheese', 'butter', 'bread', 'apples', 'oranges', 'chocolate', 'wine']]
```

← Two brackets = nested list

- Exercise 2: Extend the particle mass dict
  - Adding W, Z and H bosons

```
print(mydict)

# Assign value to key (in MeV)
mydict["W"] = 80e3
mydict["Z"] = 91e3
mydict["H"] = 125e3

print(mydict)
```

```
{'electron': 0.511, 'muon': 105.7, 'tau': 1777}
{'electron': 0.511, 'muon': 105.7, 'tau': 1777,
'W': 80000.0, 'Z': 91000.0, 'H': 125000.0}
```

  - Could also use `update()` with 2nd dict
  - Most did this correctly but a few created a new dict with all the items from scratch
    - Rather than adding new items to existing dict
  - Also, many put these masses in GeV
    - Rather than consistently in MeV

3

- Like last week, we will now look at the solutions to the formative problems and some common issues
- Ex 1: All managed to create list
  - Say why -1 better: don't need to change if increases in size
  - + for numbers, strings, lists
  - append takes single element **(if give a list will get a list within a list**); fif you want to give a list you need to use extend takes list (same as +)
  - Can see list of lists due to two brackets **(shows items of list can contain other lists -> nested data structures) [show]**
- Ex2: We saw that can change an element by assigning to the corresponding key
  - If the key doesn't exist it creates a new element with this value
  - E.g. assign 80 to the new key "W" using square brackets**, just like simple variable (could also use the update function)**
  - May not always be able to change the initial dict if someone else created it or don't have all values at the start

Lecture 2: Formative problem solutions (2)

- Exercise 3: Create array using `np.linspace`, printing the array and **number** of steps
  - Note: you can assign multiple variables on one line using commas

```
import numpy as np

# Define parameters
start, stop, step = 10.0, 22.0, 3

# Number of streps is one less than entries
nEntries= int((stop - start)/step) + 1
print("Number of steps",nEntries - 1)

# Linspace takes number of entries
myArray = np.linspace(start, stop, nEntries)
print("myArray =",myArray)
```

```
Number of steps 4
myArray = [10. 13. 16. 19. 22.]
```

  - Some people entered nEntries = 5 by hand
  - Many also printed the step size or number of entries rather than the number of steps
    - Which is the number of entries minus 1

- Exercise 4: Print out 2x2 lower right matrix
  - Remember slices indices are inclusive on lower bound but exclusive on upper one

```
[100]: # Oritinal matrix
       print("matrix = \n",matrix)
```

```
matrix =
 [[11. 12. 13. 14.]
 [21. 22. 23. 24.]
 [31. 32. 33. 34.]]
```

```
[99]: # Slice matrix to get submatrix
      submatrix = matrix[1:3, 2:4]
      print("matrix[1:3, 2:4] =\n", submatrix)
```

```
matrix[1:3, 2:4] =
 [[23. 24.]
 [33. 34.]]
```

Must use form with single square brackets [x, y] for this, not [x][y]

  - NB: As go to last index, don't need upper bound

```
[123]: submatrix = matrix[1:, 2:]
       print("matrix[1:3, 2:4] =\n", submatrix)
```

```
matrix[1:3, 2:4] =
 [[23. 24.]
 [33. 34.]]
```

  - Some people created a new array of the elements from scratch rather than slicing

4

- Ex 3: **Didn't cover in last week's lecture but one of most useful**
  - **linspace creates and array from start to stop with n linear-spaced (as opposed to log-spaced) entries. Hence the name.**
  - Now in the question you were told the step **size** but not what takes -> needs the nEntries
  - … by hand … Can calculate entries from # steps + 1(need to do so for more complex problems)
    - N steps = difference between start and end in terms of number of steps, as int then + 1
  - Once have this create linspace
  - Many also … number of steps = nEntries – 1 = 4 as see from output (not 3 which is the size of the step)
- Ex 4: Go through slice indices**, first row and then column**, reminding upper is exclusive
  - if no upper index in slice then goes to end, **similar for lower index -> will start from the beginning**

Lecture 2: Formative problem solutions (3)

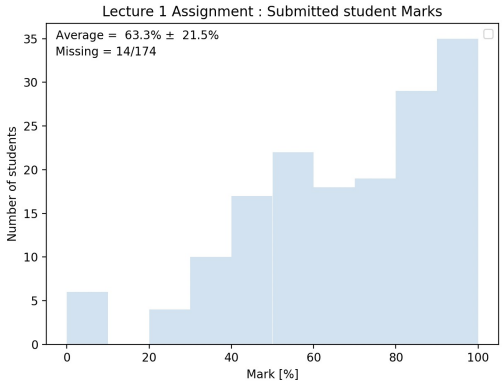
- Exercise 5: Swap the matrix rows and cols
  - Using `np.transpose()`

```
[117]: print("Submatrix =\n", submatrix)
       print("Transpose =\n",np.transpose(submatrix))
       print("Submatrix =\n", submatrix)

       Submatrix =
        [[23. 24.]
        [33. 34.]]
       Transpose =
        [[23. 33.]
        [24. 34.]]
       Submatrix =
        [[23. 24.]
        [33. 34.]]
```

  - Note: transpose function doesn't change original but makes a transposed copy
    - Can assign result to a variable

- Lecture 2 summative: marks released in Canvas
  - Analysed using python with matplotlib ☺



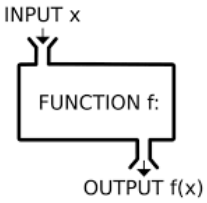
5

- Ex 5: Use submatrix that assigned on previous slide
  - Transpose takes an array and returns a transposed copy
  - **Show help/docs, esp. example part**
  - Need to call this with np.
  - Doesn't change original, which can see if print that out after, but rather a copy which you can assign to a variable
- **QUESTIONS on formative problems?**

- Have released the marks and feedback on the first week assignment in canvas
  - Analysed the summative marks
  - 8% missing -> won't pass
  - Average = 63% and peaked at high values -> most above pass mark
  - If you are below this mark and struggling please get in touch

Functions

- To be able to manipulate our data we need to look at functions in more detail
- Functions are self-contained pieces of code that, when called, accomplish a specific task
  - They take in data as comma-separated arguments, process it, and return some result
- We have already used several functions
  - Directly available functions e.g `print()` and `type()`
    - https://docs.python.org/3/library/functions.html
  - Functions in built-in modules e.g. `math.sin()`
  - Functions from third-party modules e.g. `np.linspace()`



- The power of functions is that, once written, they can be used over and over again
  - Performing the same task on different inputs to get new results (without rewriting code)
- This week we will look at writing and using our own functions
  - Generally, if you write code that you will use (or think you will use) more than once, make a function!
  - You will lose marks throughout the course if you don't make appropriate use of functions

6

- **Now that we can store our data we want to be able to manipulate it**
  - **Which we do using functions as we have seen. This week we will look at these in more detail**
- **Point to diagram**
  - **Similar to functions in maths**
- Modules:
  - precede with name followed by dot
  - Alternatively, import just functions want (e.g. from math import sin) -> no longer need
- **Performing …, which is more efficient as don't have to rewrite code**

NumPy functions

- Before doing so, we will briefly revisit NumPy functions to see a particularly powerful feature
- Often, we want to apply a function to a whole array
  - E.g. perform a math operation like exponentiation
- Since the built-in `math.exp` function only takes one input variable we'd need to call it for each element
- However, the equivalent NumPy function, `np.exp`, can also take arrays as arguments:

```
[128]: np.exp(x_array)
[128]: array([3.00416602e+00, 1.27410596e+03, 5.40364937e+05])
```

  - This is significantly faster so we'll use the np version
- The same is true of pretty much all NumPy functions
  - And math operators also work on whole NumPy arrays

```
[125]: x_array = np.linspace(1.1, 13.2, 3)

[130]: import math
       print(math.exp(x_array[0]))
       print(math.exp(x_array[1]))
       print(math.exp(x_array[2]))

3.0041660239464334
1274.105955173454
540364.9372466916

[127]: math.exp(x_array)

TypeError
Traceback (most recent call last)
<ipython-input-127-5676efe574fc> in <module>
----> 1 math.exp(x_array)

TypeError: only size-1 arrays can be converted to Python scalars

print(x_array)
print(x_array * 2)

[ 1.1   7.15 13.2 ]
[ 2.2 14.3 26.4]
```

7

- … such as the np.transpose function you used last week
- When have array of data want to apply function (e.g. exponentiation) to all the data i.e. the whole array
- … each element -> go over example
- **This quickly gets tiresome as array gets larger and, even once we know how to use a loop to repeat (next lecture) this is not efficient**
  - **Solution = equivalent numpy function np.exp**
- **As well as single value can take an array, performing array-at-a-time operations rather than element at a time and returning a new array**
  - Get error if try with math.exp (+ explain how to read again)
  - **More efficient -> faster -> use np rather than math ones when dealing with arrays**
- **Will play with this in the notebook this week**
- **QUESTIONS on array-at-a-time function usage?**

Writing our own functions

- We define a python function using the `def` keyword, followed by the name of the function
  - After this we have a comma-separated list of any arguments in brackets and a terminating colon (`:`)
- The body of the function does the work
  - It consists of indented lines of code
  - An optional `return` statement to give any result(s)

```
def function_name(arg1, arg2, ...):
    code
    return result1, result2, ...
```

- Let's look at an example for calculating the area of a triangle:

```
[133]: # Function taking two arguments
       def triangle_area(base, height):
           # Calculate the area
           area = 0.5*base*height
           # Return the result
           return area
```

```
[135]: b = 4 # mm
       h = 5 # mm

       result = triangle_area(b, h)

       print("Area (mm^2) = ", result)
```

```
Area (mm^2) =  10.0
```

- Once we have defined it, we call it as with the built-in functions
  - Passing the arguments in brackets and assigning the return value to a variable
  - Note: the names of the variables outside the function don't need to match those inside

8

- Point to the example (pseudo-code)
- **Must have a colon**
- **Indentation is how python knows which code lines are part of the function**
  - **Doesn't matter how much but you need to be consistent**
- Go over example [show] – **defined but not used so no output**:
  - def then sensible name
  - 2 arguments: base and height
  - performs calculation (can be many lines) – note indentation
  - Returns the area
- **We use the function by calling it by name and passing our particular values as arguments, in correct order, like any other function**
  - It returns the area, which we assign to a variable that we use like any other variables
  - **Note: names don't have to match -> those inside the function arguments are just placeholders that are assigned to our particular inputs when we call the function**
- **QUESTIONS on writing functions so far?**

Arguments and return values

- We can return multiple values from a function
  - The result is a tuple
  - Each element can be assigned to a separate variable

```
[150]: def circle_properties(radius):
           area = np.pi*radius**2
           perimeter = 2*np.pi*radius
           # Return both values
           return area, perimeter
```

```
[153]: result = circle_properties(10)
       print("Area =", result[0], "Perim =", result[1])
```

```
Area = 314.1592653589793 Perim = 62.83185307179586
```

```
[154]: area, perim = circle_properties(10)
       print("Area =", area, "Perim =", perim)
```

```
Area = 314.1592653589793 Perim = 62.83185307179586
```

- Arguments can also have default values
  - By assigning to a value in definition
  - If argument left out when called, get default

```
[158]: def circle_properties(radius, pi = np.pi):
           area = pi*radius**2
           perimeter = 2*pi*radius
           # Return both values
           return area, perimeter
```

```
area, perim = circle_properties(10)
print("Area =", area, "Perim =", perim)
```

```
Area = 314.1592653589793 Perim = 62.83185307179586
```

```
area, perim = circle_properties(10, 6.28)
print("Area =", area, "Perim =", perim)
```

```
Area = 628.0 Perim = 125.60000000000001
```

9

- Go through example
  - takes in one argument now, calcs area and perimeter -> both lines indented, return both separated by comma
- If call, we get a tuple, and can access each element by index
  - But easier to assign directly to multiple values (separated by a comma)
  - **Automatically unpacks the tuple. Of course, must have correct number of vars for how many elements returned**
- Let's give it an extra argument, which is the value of pi
  - Obviously, this is generally 3.14 … so we can assign it to this by default, exactly like assigning a normal var
  - **Then use this new arg (pi) rather than np.pi inside the function**
  - Since has a default -> don't need to provide it
  - If call with no 2nd argument, uses this default -> same result
  - But if we pass it a value (e.g. a different geometry) this overwrites the default

Named arguments

- So far, when calling a function we have passed in the arguments in the order they are defined
  - Python uses the order to work out which argument is which
  - In general, calling with a different order doesn't make sense

```
[141]: def happy_birthday(name, age):
           print ("Happy", age, "birthday", name, "!")
```

```
[142]: happy_birthday("Carl", 41)
Happy 41 birthday Carl !
```

```
[143]: happy_birthday(41, "Carl")
Happy Carl birthday 41 !
```

- These are called positional arguments, but you can also specify the arguments by name
  - And in this case the order doesn't matter

```
[144]: happy_birthday(age = 41, name = "Carl")
Happy 41 birthday Carl !
```

- When combined with default arguments, these allow us to call a function in different ways
  - E.g. can leave out the first argument unlike with positional args

```
[145]: def happy_birthday(name = "to you", age = 21):
           print ("Happy", age, "birthday", name, "!")
```

```
[148]: happy_birthday(age = 41)
Happy 41 birthday to you !
```

```
[149]: happy_birthday(41)
Happy 21 birthday 41 !
```

  - This is useful when have many optional arguments as we shall see

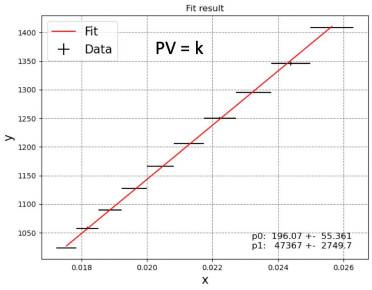
10

- So far … Suppose we have a happy bday function that takes a name and age and prints out a message
  - **You can see here that we don't have to return anything**
- Python uses order … Go over correct call and how it knows which is which based on order + wrong assignment if call in other order
- These … by name, using the argument names we gave when we defined the function (again like any variable assignment)
  - **Although we passed in the age first, python knows what to do since we have told it explicitly that this is the age -> not using position**
  - **NB: Don't confuse with defaults, which are given when defining the function, not calling it.**
- You might wonder why this is useful? First, more explicit I,e, easier to read and second when combined…
- Change function to give defaults: If we don't know the name we just say "to you" and if we don't know the age we assume 21
  - If call positionally then the we have to give the args in order -> no way to just give the age although default for name-> would be assigned to the name -> doesn't make sense

    - When using named arguments, we can just give age = and not the name
- Call in multiple ways -> useful when have many arguments **since allows some arguments to be optional**
- **QUESTIONS on arguments and return values?**

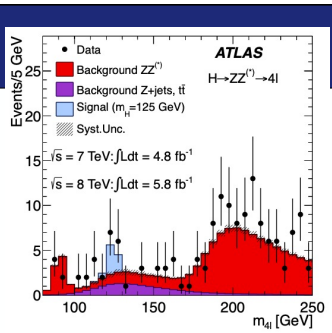
Plotting

- As mentioned last time, we often want to visualise data e.g.
  - Experimental results, results of calculations
  - Comparisons of the two
- This will be the case in your PHYS106 laboratories
  - Will plot the data taken and fit them e.g.

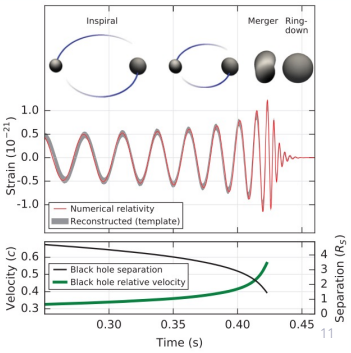


- It is also the case in almost all scientific papers
  - And in many other fields too

Higgs Discovery



Gravitational wave discovery



11

- Now we know more about how functions work we can go back to the plotting that we started looking at last time
- Scientific papers, such as the plots I showed previously for …

Matplotlib Histograms

- We have seen how matplotlib can plot hists
- We create a `figure()`, giving the size tuple
  - And set the title + axis labels
- We plot the histogram with `hist()`
  - We can explicitly specify the bin edges
- It also takes many other params
  - These are named args with default values

```
hist(x, bins=None, range=None, density=False, weights=None,
     cumulative=False, bottom=None, histtype='bar', align='mid',
     orientation='vertical', rwidth=None, log=False, color=None,
     label=None, stacked=False, *, data=None, **kwargs)
```

- We can also add a `legend()`
  - Which shows the hist labels

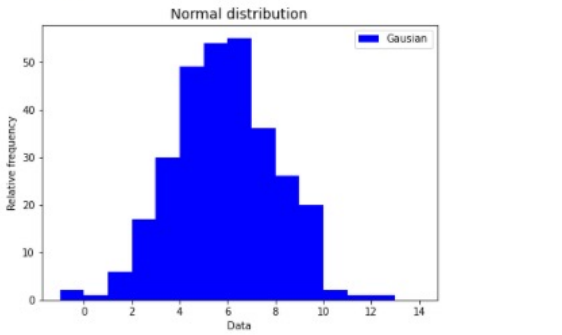
```
import matplotlib.pyplot as plt
%matplotlib inline

# Construct a figure with a title and axis labels
plt.figure(figsize = (7, 5))
plt.title('Normal distribution', fontsize = 14)
plt.xlabel('Data')
plt.ylabel('Relative frequency')

# Set the bin edges
bin_lo, bin_hi, nbins = -1, 14, 15
bin_edges = np.linspace(bin_lo, bin_hi, nbins+1)

# Make a histogram and display it, with legend
plt.hist(gaussArr, bins = bin_edges, color = 'b', label = "Gausian")
plt.legend()
plt.show()
```

edges = # bins + 1



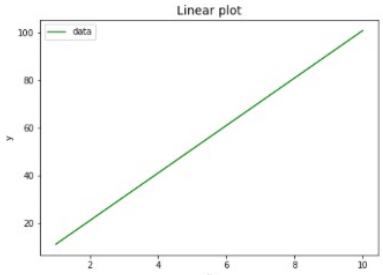
- Import
- Go over example …
- We create …
- **Now understand how we specify the figsize or fontsize -> named arguments with default values -> optional**
- Hist takes the data as a numpy array
  - **Now add explicitly bin edges (easiest to use linspace -> go through -> bins+1 as number of edges) and a label as named parameters**
- Takes many other params (do before call hist and show help)…
  - **Looks scarry but mostly only use a subset**
- Before showing … call legend function makes a legend from labels for each plot -> important if have several plots on a given figure (as we'll see)

Other Matplotlib plots

- It also has many, many other types of plots, from simple to complex. Two important ones are:
  - Data as lines or markers: `plot(x, y, …)`
  - Data with errors: `errorbar(x, y, yerr =, …)`

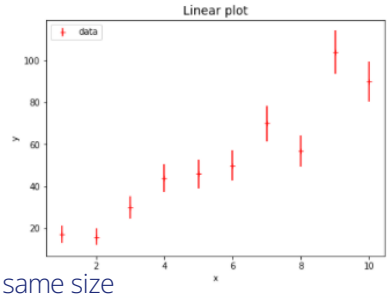
[197]:
```
# Construct a figure with a title and axis labels
plt.figure(figsize = (7, 5))
plt.title('Linear plot', fontsize = 14)
plt.xlabel('x')
plt.ylabel('y')

# Make a line plot
plt.plot(xdata, ydata, linestyle = "-", color = 'g', label = 'data')
plt.legend(loc = 2)
plt.show()
```



[198]:
```
# Construct a figure with a title and axis labels
plt.figure(figsize = (7, 5))
plt.title('Linear plot', fontsize = 14)
plt.xlabel('x')
plt.ylabel('y')

# Make a plot with error bars
plt.errorbar(x, y, yerr = yerrors, marker = '+',
             linestyle = "", color = 'r', label = 'data')
plt.legend(loc = 2)
plt.show()
```



- In each case `x`, `y` (and `yerr`) are NumPy arrays of the same size
  - Many of the other options you can pass are the same for all e.g. color

13

- Two important ones (which we will use a lot): plot + errorbar
  - Same (as see) except for the function that defines what is plotted (e.g. hist part)
- Plot takes numpy arrays of x and y data -> line
  - along with optional linestyle etc (show different ones)
  - legend also has an optional param to specify where to plot it that we didn't use on previous page (2 -> top left rather than default top right which would overlap)
- errorbar (saw last week in notebook) **e.g. for data**, takes the same but optionally np arrays of x (saw in last week's nb for RMS or FHWM) or y errors (here) or indeed both
  - along with optional marker style (show different ones)
  - **By default joins points up which we don't want for data -> empty linestyle**
  - **Errorbar is a bit of a misleading name: it draws the points AND the associated error bars**
  - **Show example of 10% xerr**
- **QUESTIONS on plotting?**

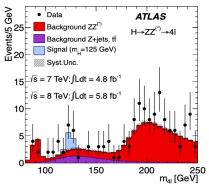
Complete Plot

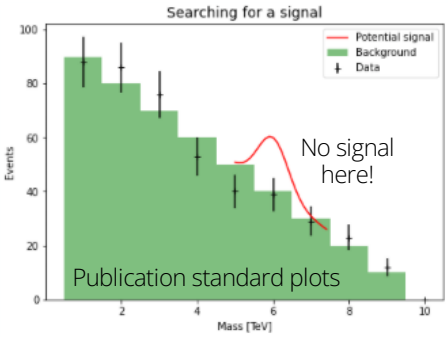
- You can plot multiple results on one figure
  - Simply call the appropriate function for each set of data before showing the figure
- The results can be plotted as different types
  - We can mix `hist()`, `plot()` and `errorbar()`
- E.g.: search for a new 6 TeV particle
  - Plot background simulation as hist
  - Plot potential signal as a line
  - Plot data as points with error bars
- Can also save the figure to a file
  - Using `savefig()` with the output name
  - Useful for inputting figures to lab reports etc

```
plt.figure(figsize = (7, 5))
plt.title('Searching for a signal', fontsize = 14)
plt.xlabel('Mass [TeV]')

plt.ylabel('Events')

plt.hist(background, bins = np.linspace(0.5, 9.5, 10), color = "g",
         alpha=0.5, label = "Background")
plt.errorbar(xdata, ydata, yerr = yerror, color = 'black',
             marker = "+", linestyle = "", label = "Data")
plt.plot(xsignal, ysignal, color = "r", label = "Potential signal")
plt.legend()
plt.savefig("MassPlot.png")
plt.show()
```





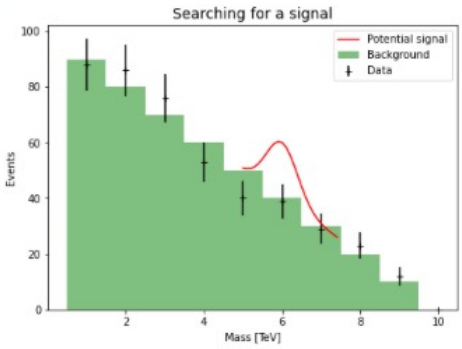
- **Create figure as normal but create multiple plots by calling hist/plot/errorbar multiple times before calling show**
- Example from my field, like the higgs search I showed in L1 and earlier
- Search for a new particle at 6 TeV
  - Theoretically predicted background -> hist
  - A hypothetical signal -> gaussian as a line centred at 6 TeV
  - We then compare this to the data which we plot as markers with error bars of course
  - Each are simply numpy arrays
- **Publication standard plots!**
- **It if often useful to save these to an output file, rather than just displaying them in line,**
  - E.g. for input to lab repots
  - You can do this using savefig() (BEFORE show) with the output name as an argument (don't just use a screen

shot!)

- Unfortunately, data follows background -> no sign of a signal -> unlike Higgs, no Nobel prize for us
- **QUESTIONS: Any questions on plotting various types (individually or together?)**

Summary

- This week we looked at how to write and use our own functions
  - Allow us to perform calculations multiple times without rewriting code
- You should get into the practice of writing functions
  - Will save you work and time in the long run!
- We also looked in more detail at plotting
  - Allowing us to visualise and compare results
- Saw different types of visualisation
  - Lines, points, histograms
  - And how to combine these
- Along the way we had more practice with NumPy
- You will exercise these and see more features in the problem class



15

- which I released the notebook for in cocalc this morning and opened the assignment in canvas.
  - Again, please do try the formative questions
- **QUESTIONS: any final questions?**