Table of Contents

1.1
2.1
2.1.1
2.1.2
2.1.3
2.1.4
3.1
3.1.1
3.1.2
3.1.3
3.1.4
4.1
4.2
5.1

The Ostap tutorials build passing

Ostap is a set of extensions/decorators and utilities over the basic PyROOT functionality (python wrapper for ROOT framework). These utilities greatly simplify the interactive manipulations with ROOT classes through python. The main ingredients of Ostap are

- preconfigured ipython script ostap, that can be invoked from the command line.
- *decoration* of the basic ROOT objects, like histograms, graphs etc.
 - o operations and operators
 - o iteration, element access, etc
 - o extended functionality
- decoration of many basic ROOT.ROOFit objects
- set of new useful fit models, components and operations
- other useful analysis utilities

Getting started

The main ingredients of Ostap are

• preconfigured ipython script ostap , that can be invoked from the command line.

ostap

Challenge

Invoke the script with -h option to get the whole list of all command line options and keys

Optionally one can specify the list of python files to be executed before appearance of the interactive command prompt:

```
ostap a.py b.py c.py d.py
```

The list of optional arguments can include also root-files, in this case the files will be opened and their handlers will be available via local list root_files

```
ostap a.py b.py c.py d.py file1.root file2.root e.py file3.root
```

Also ROOT macros can be specified on the command line

```
ostap a.py b.py c.py d.py file1.root q1.C file2.root q2.C e.py file3.root q4.C
```

The script automatically opens Tcanvas window (unless --no-canvas option is specified) with (a little bit modified) LHCb style. It also loads necessary decorators for ROOT classes. At last it executes the python scripts and opens root-files, specified as command line arguments.

Values with uncertanties: ValueWithError

One of the central object in ostap is C++ class Gaudi::Math::ValutWithError , accessible in python via shortcut VE . This class stands for r a combination of the value with uncertainties:

```
from Ostap.Core import VE
a = VE( 10 , 10 ) ## the value & squared uncertainty - 'variance'
b = VE( 20 , 20 ) ## the value & squared uncertainty - 'variance'
print "a=%s" % a
print "b=%s" % b
print 'Value of a is %s' % a.value()
print 'Effor of b is %s' % b.error()
print 'Variance of b is %s' % b.cov2 ()
```

A lot of math operations are predefined for VE -objects.

Challenge

Make a try with all binary operations (+, -, *, /, **) for the pair of VE objects and combinations of VE -objects with numbers, e.g.

```
a + b
a + 1
1 - b
2 ** a
a +=1
b += a
```

Compare the difference for following expresssions:

```
a/a ## <--- HERE
a/VE(a) ## <--- HERE
a-a ## <--- HERE
a-VE(a) ## <--- HERE
```

Note that for trivial cases the correlations are propertly taken into account

Additionally many math-functions are provided, carefully takes care on uncertainties

```
from LHCbMath.math_ve import *
sin(a)+cos(b)/tanh(b)
atan2(a,b)/log(a)
```

Simple operations with histograms

Historgam content

Ostap.PyRoUts module provides two ways to access the histogram content

- by bin index, using operator []: for 1D historgam index is a simple integer number, for 2D and 3D-histograms the bin index is a 2 or 3-element tuple
- using *functional* interface with operator () .

```
histo = ...
print histo[2] ## print the value/error associated with the 2nd bin
print histo(2.21) ## print the value/error at x=2.21
```

Note that the result in both cases is of type VE , *value+/-uncertainty*, and the interpolation is involved in the second case. The interpolation can be controlled using interpolation argument

```
print histo ( 2.1 , interpolation = 0 ) ## no interpolation
print histo ( 2.1 , interpolation = 1 ) ## linear interpolation
print histo ( 2.1 , interpolation = 2 ) ## parabolic interpolation
print histo ( 2.1 , interpolation = 3 ) ## cubic interpolation
```

Similarly for 2D and 3D cases, interpolation parameter is 2 or 3-element tuple, e.g. (2,2) (3,2,2), (3,0,0), ...

Set bin content

```
histo[1] = VE(10,10)
histo[2] = VE(20,20)
```

Loops over the histogram content:

```
for i in histo :
    print 'Bin# %s, the content%s' % ( i, histo[i] )
for entry in histo.iteritems() :
    print 'item ', entry
```

The reversed iterations are also supported

```
for i in reversed(histo) :
    print 'Bin# %s, the content%s' % ( i, histo[i] )
```

Histogram slicing

The slicing of 1D-historgam can be done easily using native slice in python

```
h1 = h[3:8]
```

For 2D and 3D-casss the slicing is less trivial, but still simple

```
histo2D = ...
h1 = histo2D.sliceX ( 1 )
h2 = histo2D.sliceY ( [1,3,5] )
h3 = histo2D.sliceY ( 3 )
h4 = histo2D.sliceY ( [3,4,5] )
```

Operators and operations

A lot of operators and operations are defined for histograms.

```
histo += 1
histo /= 10
histo = 1 + histo ## operations with constants
histo = histo + math.cos ## operations with functions
histo /= lambda x : 1 + x ## lambdas are also functions
```

Also binary operations are defined

```
h1 = ...

h2 = ...

h3 = h1 + h2

h4 = h1 / h2

h5 = h1 * h2

h6 = h1 - h2
```

For the binary operations the action is defiened accordinh to the rule

- the type of the result is defined by the first operand (type, and binning)
- for each bin i the result is estimated as a oper b , where:

```
o oper stands for corresponding operator ( +, -, *, /, **)
```

- a = h1[i] is a value of the first operand at bin i
- o b = h2(x), where x is a bin-center of bin i

More operations

There are many other useful opetations:

- abs : apply abs function bin-by-bin
- asym : equivalent to (h1-h2)/(h1+h2) with correct treatment of correlated uncertainties
- frac : equivalent to (h1)/(h1+h2) with correct treatment of correlated uncertainties
- average: make an average of two historgam
- chi2: bin-by-bin chi2-tension between two historgams
- ... and many more

Transformations

```
 h1 = histo.transform \ ( \ lambda \ x,y : y \ ) \ \# \ identical \ transformation \ (copy)   h2 = histo.transform \ ( \ lambda \ x,y : y^**^3 \ ) \ \# \ get \ the \ third \ power \ of \ the \ histogram \ content   h3 = histo.transform \ ( \ lambda \ x,y : y/x \ ) \ \# \ less \ trivial \ functional \ transformation
```

Math functions

The standard math-functions can be applied to the histoigram (bin-by-bin):

```
from LHCbMath.math_ve import *
h1 = sin ( histo )
h2 = exp ( histo )
h3 = exp ( abs ( histo ) )
...
```

Sampling

There is an easy way to sample the histograms according to their content, e.g. for toy-experiments:

```
h1 = histo.sample() ## make a random histogram with content sampled according to bin+-error in original histo h2 = histo.sample( accept = lambda s: s > 0 ) ##sample but require that sampled values are positive
```

Smearing/convolution with gaussian

It is very easy to smear 1D histogram according to gaussian resolution

```
h1 = histo.smear ( 0.015 ) ## apply "smearing" with sigma = 0.015
h2 = histo.smear ( sigma = lambda x : 0.1*x ) ## smear using 'running' sigma of 10% resolution
```

Rebin

```
original = ... ## the original historgam to be rebinned
template = ... ## historgams that deifned new binning scheme
rebin1 = original.rebinNumbers ( template ) ## compare it!
rebin2 = original.rebinFunction ( template ) ## compare it!
```

Note that there are two methods for rebin rebinNumbers and rebinFunction - they depends on the treatment of the histogram.

Challenge

Choose some initial histogram with non-uniform biuning, choose *template* historam with non-uniform binning and compare two methods: rebinNumbers and rebinFunction.

Integrals

There are severalintegral-like methods for (1D)histograms

• accumulate : useful for *numbers*-like histograms, only bin-content inn used for summation (unless the bin is effectively split in case of low/high summation edge does not coinside with bin edges)

• integrate : useful for *function*-like histograms, perform integration taking into account bin-width.

• integral it transform the histogram into ROOT.TF1 object and invokes ROOT.TF1.Integral

Running sums

and the efficiencies of cuts_

```
h1 = histo.sumv ()  ## increasing order: sum(first,x)
h2 = histo.sumv ( False ) ## decreasing order: sum(x,last )
```

Efficiency of the cut

Such functionality immediately allows to calculate efficiency historgrams using effic method:

```
h1 = histo.effic () ## efficiency of var<x cut
h2 = histo.effic ( False ) ## efficiency of var>x cut
```

Conversion to ROOT. TF(1, 2, 3)

Scaling

In addition to trivial scaling operations h *= 3 and h /= 10 there are seevral dedicated method for scaling

• scale it scales the historgam content to a given sum of *in-range* bins

```
print histo.accumulate()
histo.scale(10)
print histo.accumulate()
```

o rescale_bins : it allows the treatment of non-uniform histograms as density distributions. Essentially each bin i is rescaled according to the rule h[i] *= a / S , where a is specified factor and s is bin-area. such type of rescaling is important for histograms with non-uniform binning

Density

There is method density that converts the histgram into *density* histogram. The density histogram (being interpreted as *function*) has unit integral. It is different from the simple rescaling for histograms with non-uniform bins.

```
d = histo.density()
```

Statistics

There are many statistic functions

- mean
- rms
- kurtosis
- skewness
- moment
- centralMoment
- nEff: number of equivalent entries
- stat : statistical information about bin-to-bin content: mean, rms, minmax, ... in form of Gaudi::Math::StatEntity class

Figure-of-Merit evaluation and cut optimisation

If *figure-of-merit* is natural and equals to *sigma(S)/S* (note that it is equal to *sqrt(S+B)/S*):

```
signal = ... ## distribition for signal
fom1 = signal.FoM2 () ## FoM for var<x cut
fom2 = signal.FoM2 ( False ) ## FoM for var>x cut
```

Note that no explicit knowledge of background is needed here - it enters indirectly via the uncertainties in signal determination.

If figure-of-merit is defined as S/sqrt(S+alpha*B)

```
signal = ...
background = ...
alpha = ...
fom1 = signal.FoM1 ( background , alpha ) ## FoM for var<x cut
fom2 = signal.FoM1 ( background , alpha , False ) ## FoM for var>x cut
```

Solve equations

One can also solve equations h(x) = v

```
value = 3
solutions = histo.solve ( value )
for x in solutions : print x
```

Conversion to `ROOT.TF(1,2,3)

The conversion of histogram to ROOT.TF1 objects is straighforward

```
f = histo.tf1()
```

Optionally one can specify interpolate flag to define the interpolation rules.

The obtained TF1 object is defined with three parameters

- 1. normalization
- 2. bias
- 3. scale

It can be used e.g. for visualize interpolated historgam as function or e.g. in ROOT.TH1.Fit method for fitting of other historgams

Efficiencies

There are several special cases to get the efficiency-historgams

Binomial efficiencies

In addition to the methods described above, few more sophisticated treatments of binomial efficiencies are provided

```
accepted = ...
total = ...

eff1 = accepted.          zechEff ( total ) ## valid for all histograms, including sPlot-weighted
eff2 = accepted.          binomEff ( total ) ## only for natural histograms
eff3 = accepted.          wilsonEff ( total ) ## only for natural histograms
eff4 = accepted.agrestiCoullEff ( total ) ## only for natural histograms
```

For *natural* historgams only one can use even more sophisticated methods, that evaluates the interval. Each method returns *graph*, and the graphs can be visuzalised for comparison:

All of this functions have an optional argument interval that defines the confidence interval, the default value is interval=0.682689492137086 that corresponds to 1 sigma.

Optimal binning?

It is not a rare case when one needs to find the binbing of the histogram that ensures almost equal bin populations. This task could be solved using eqaul_bins method

```
very_fine_binned_histo = ... ## get the fine binned histograms
edges1 = fine_binned.equal_edges ( 10 ) ## try to fing binning with 10 almost equally populated bins
edges2 = fine_binned.equal_edges ( 10 , wmax = 5 ) ## try to fing binning with 10 almost equally populated bins, but avoid bin
s wider than "wmax"
```

Persistency

Ostap.ZipShelve

Ostap offers very nice&efficient way to store the objects in persistent dbase. This persistency is build around shelve module and differs in two way

- 1. the conntent of payload is compressed, using zlib module making the data base very compact
 - o (optionally) the whole database can ve further gzip 'ed using gzip module, if the extension .gz is provided. It makes data banse even more compact.
- 2. in addition to the native dict interface from shelve , more extensive interface with more methods is supported.

Create database and write objects to it:

```
a = ...
import Ostap.ZipShelve as DBASE
with DBASE.open ( 'my_dbase.db' ) as db : ## create DBASE
db.ls()
db['a'] = a
db['histo'] = ROOT.TH1D('h1','',10,0,1)
```

Reading objects from database

```
with DBASE.open ( 'my_dbase.db' , 'read') as db : ## read DBASE
db.ls()
b = db['a']
h2 = db['histo']
```

One can store in database all *pickable* objects, that means all python objects, all (serializeable) ROOT objects. All C++ objects with LCG/Reflex/Cint -dictionaries are also could be stored database. In practice, everything is storable, including complex combination of python&C++ objects, like dictionary of historgams and python classed, inherited from C++ -base classes.

Plain ROOT.TFile

Ostap adds some decorations even for the plain ROOT. TFile , making its interface more pythonic:

RootOnlyShelve

The module Ostap.RootShelve offers the thin wrapper over ROOT.TFile that implement shelve -interface. As a result one gets a light database build a top of underlying ROOT.TFile, where ROOT -objects could be stored:

```
from Ostap.RootShelve import RooOnlyShelf
db = RooOnlyShelf('mydb.root','c')
h1 = ...
db ['histogram'] = h1
db.ls()
```

RootShelve

The module Ostap.RootShelve offers also more sophisticated wrapper over ROOT.TFile that also implements shelve -interface and able to store ROOT and any other *pickable* objects

```
from Ostap.RootShelve import RooShelf
db = RooShelf('mydb.root','c')
h1 = ...
db ['histogram'] = h1
db ['histogramlist'] = h1,h2,h3
db.ls()
```

In details ...

For non-ROOT objects, database actually stores them as ROOT::TString objects carrying their pickle representation with on-flight removal/substitutions of some magic symbol sequences, since ROOT::TString is not a real BLOB.

Decorations

Ostap decorates many ROOT.ROOFit classes, adding more convinient methods to them.

RooArgList and RooArgSet

All these classes have got set of additional python-like methods for iteration, extension, addition, elemtn access checking the content etc...

Also several methods to provide more coherent interfaces (e.g. add vs Add) are added.

```
# Ostap.PyRoUts
                                 INFO
                                         Zillions of decorations for ROOT/RooFit objects
    Lengths are 2 2
     'a' : ( 0 +- 0 )
    'b' : ( -10 +- 0 )
    'b' : ( -10 +- 0 )
    'c' : ( 1 +- 0 )
     a in 1 ? True True
     b in 1 ? True True
     c in 1 ? False False
9
     a in 1 ? False False
     b in 1 ? True True
11
     c in 1 ? True True
13
     'a' : ( 0 +- 0 ) 'b' : ( -10 +- 0 )
     'b' : ( -10 +- 0 ) 'c' : ( 1 +- 0 )
14
              ['a:0.0', 'b:-10.0', 'a:0.0', 'b:-10.0']
    11+11 :
15
16
    11+12 :
              ['a:0.0', 'b:-10.0', 'b:-10.0', 'c:1.0']
17
    12+12 :
              ('b:-10.0', 'c:1.0')
              ('b:-10.0', 'c:1.0', 'a:0.0')
18
    12+11 :
              ['a:0.0', 'b:-10.0', 'c:1.0']
    l1+c :
19
              ('b:-10.0', 'c:1.0')
    12+c :
              ['a:0.0', 'b:-10.0', 'd:-1.0']
    l1+d :
21
    12+d :
              ('b:-10.0', 'c:1.0', 'd:-1.0')
22
              ['a:0.0', 'b:-10.0', 'c:1.0']
    c+l1 :
23
              ('b:-10.0', 'c:1.0')
    c+12 :
24
    d+l1 :
              ['a:0.0', 'b:-10.0', 'd:-1.0']
               ('b:-10.0', 'c:1.0', 'd:-1.0')
    d+12 :
output.txt hosted with ♥ by GitHub
                                                                                    view raw
```

```
import ROOT
import Ostap.PyRoUts

a = ROOT.RooRealVar ('a', 'a', -10, 10)
b = ROOT.RooRealVar ('b', 'b', -10)
c = ROOT.RooConstVar('c', 'c', 1)
d = ROOT.RooConstVar('d', 'd', -1)
```

```
11 = ROOT.RooArgList
                            (a,b)
                            (b,c)
    12 = ROOT.RooArgSet
11
    print 'Lengths are %s %s ' % ( len ( l1 ) , len( l2 ) )
12
13
14
    for i in l1 : print i
15
    for i in 12 : print i
16
17
    for 1 in ( 11 , 12 ) :
        print ' a in 1 ? %s %s ' % ( a in 1 , 'a' in 1 )
18
        print ' b in 1 ? %s %s ' % ( b in 1 , 'b' in 1 )
20
        print ' c in 1 ? %s %s ' % ( c in 1 , 'c' in 1 )
21
22
23
    print l1[0] , l1[1]
24
    print 12['b'] , 12['c']
25
26
    print 'l1+l1 :
                      %s' % ( l1 + l1 )
                      %s' % ( l1 + l2 )
    print 'l1+l2 :
28
    print '12+12 :
                    %s' % ( 12 + 12 )
29
    print 'l2+l1 :
                      %s' % ( 12 + 11 )
                     %s' % ( l1 + c )
    print 'l1+c :
    print '12+c :
                     %s' % ( 12 + c )
    print 'l1+d :
                     %s' % ( l1 + d )
                     %s' % ( 12 + d )
34
    print '12+d :
    print 'c+l1 :
                     %s' % ( c + l1 )
36
    print 'c+l2 :
                     %s' % ( c + 12 )
                     %s' % ( d + l1 )
38
    print 'd+l1 :
    print 'd+12 :
                     %s' % ( d + 12 )
roofit_lists.py hosted with ♥ by GitHub
                                                                                    view raw
```

RooAbsData and RooDataSet

These methods also have got the extended interface with many useful methods and operators, like e.g. concatenation of datasets a+b and merging them a*c.

RooDataSet class also has go many methods, that are similar to those of ROOT.TTree , in particular project and draw :

```
dataset = ...
dataset.draw('mass','pt>1')
histo = ...
dataset.project ( histo , 'mass', 'pt>1' )
```

 $Many other methonds \ like \ \ statVar \ , \ \ sumVar \ , \ \ statCov \ , \ \ vminmax \ \ are \ also \ the \ same \ as for \ \ ROOT.TTree \ , see \ above.$

```
s1 = dataset.statVar ('eff')
s2 = dataset.sumVar ('eff')
r = dataset.statCov ('eff', 'pt')
mn,mx = dataset.vminmax ('eff')
```

RooFitResult

The class RoofitResult get many decorations that allow to access fit results

```
result = ...

par1 = result.params() ## get all floating parameters

par2 = result.params( float_only = False ) ## all parameters

a,v = result.param ( 'a' ) ## par by name

a,v = result.param ( a ) ## par by RooFit object itself

p = result.a ## par as attribute

for par in result: print par ## iteration

for name,par in result.iteritems(): print par ## iteration

print result.cov ( 'a' , 'b' ) ## get the covariance submatrix

print result.corr ( 'a' , 'b' ) ## get the correlation coefficient
```

Also the simple math with fiting parameters is supported

RooRealVar & friends

Few simple operations are added to simplify the calculations with RooRealVar objects:

```
x = ROOT.RooRealVar( ... )
x + 10
x - 10
x * 10
x / 10
10 + x
10 - x
10 * x
10 / x
x += 2
x -= 2
x *= 2
x /= 2
x ** 3
```

PDFs and the basic models

Ostap provides set of useful wrapper and helper class that drastically simplify the construction and manipulations with RooAbsPdf - objects.

E.g. consider the simplest case - creation of the Gaussian PDF using the standard way the standard way:

```
x = ROOT.RooRealVar ('x' ,'x' ,2,3)
mean = ROOT.RooRealVar ('mean' ,'mean' ,3.100,3.080,3.120)
sigma = ROOT.RooRealVar ('sigma','sigma',0.015,0.010,0.025)
bare = ROOT.RooGaussian('Gauss','Gaussian', x , mean , sigma ) ## <--- HERE</pre>
```

In Ostap it can be done in a bit simpler way

How to define parameter?

There are may ways to define parameter

1. One can use the existing RooAbsReal object, e.g. RooRealVar or RooConstVar :

2. One can use the plain number value , 2- or 3-element tuple (minval, maxval) or (value, minval, maxval) . In this case the variable of the type RoorealVar will be automatically created using this specification. (In case of the plain number, the corresponding parameter will be fixed in the fit).

For all models, all known parameter are accessible (and documented) as python property

```
gauss = ...
help(gauss.xvar)
print gauss.sigma
help(gauss.mean)
```

There are many predefined models, accesible via Ostap.FitModels module:

```
import Ostap.FitModels as Models
help(Models)
```

Base class PDF

All Ostap-based fit models and PDFs (directy or indirectly) inherit from python base class PDF , that provides great additional functionality, in particular the methods fitTo and draw that simplfy the fitting procedure itself and visualization of the results:

The method fitTo

```
gauss = Gauss_pdf ( ... )
dataset = ....
result , frame = gauss.fitTo ( dataset , silent = True , reFit = 2 )
print 'FitResults: %s' % result
```

All the native Roofit *commands* can be specified as optional arguments, as well as many commands specific for Ostap, e.g. refit=2 above means *in case of fit failure*, *try to refit it (up to 2 times)*, and the meaning of silent=True is obvious.

The method draw

```
gauss = Gauss_pdf ( ... )
dataset = ....
result , frame = gauss.fitTo ( dataset , silent = True , reFit = 2 )
print 'FitResults: %s' % result
frame = gauss.draw ( dataset , nbins = 100 )
```

Fitting and vizualisation can be combined:

```
gauss = Gauss_pdf ( ... )
dataset = ....
result , frame = gauss.fitTo ( dataset , draw = True , nbins = 100 ) ## draw it after the fit
```

Access to the underlying RooAbsPdf object

The access to the underlying bare RooAbsPdf -object can be done (if needed) via the propety pdf

```
gauss = Gauss_pdf ( ... )
root_pdf = gauss.pdf
```

Other methods

PDF class is equipped with many other useful methods:

• fitHisto: The method fitTo can be blindly applied not only to RoodataSet -objects, but also to the histograms:

```
histo = ...
r, f = gauss.fitTo ( histo , draw = True )
```

However the dedicated method fithisto sometimes could be more usefu

```
histo = ...
gauss.fitHisto ( histo , draw = True )
```

draw_nll: vizualize NLL-scans and LL-profiles

• generate : tiny but useful wrapper for RooAbsPdf::generate

• minmax: make the estimates for the minimal and maximal values for the PDF. For some models it is done analytically or semianalitycally, for remaining models it is done using random shoots.

```
mn,mx = gauss.minmax( 500000 )
```

• __call__ : it allows to use PDF as simple function

```
gauss = ...
print gauss( 3.090 ), gauss( 3.100 ), gauss( 3.110 )
```

- Several statistical functions. For some models analytical orsemianalitycal calculations are used, for remnig models numerical
 estimations are performed using scipy
- rms : rms for the distribution
- o fwhm: full width at half maximum
- fwhm: full width at half maximum
- moment : the moment of the distribution
- o central moment : the central moment of the distribution
- o skewness : skewness for the distribution
- kurtosis : kurtosis for the distribution
- mode : the *mode* for the distribution
- median : median value for the distribution
- get_mean : *mean* value for the distribution
- cl_symm : symmetric confidence interval
- o cl_asymm : asymmetric confidence interval
- quantile : *quantile* value for the distribution
- o integral: integral for the distribution
- o derivative : derivative of the PDF at the given point

Convolution

Ostap provides helper class that simplify construction of fit models taking into accotun resolution functions:

As resolution one can specify

- 1. Any resolutuon model (RooAbsPdf)
- 2. simple number s, in this case the gaussian resolution model with sigma = s will be used
- 3. Any RooAbsReal objetct, it will be used as sigma for gaussian resoltuion model

There are several optional flags

- useFFT=True : use Fast-Fourier-Transform or plain numerical convolution ?
- nbins=100000 : sampling for Fast-Fourier-Transform
- buffer=0.25 : buffer size for Fast-Fourier-Transform, argument for setBufferFraction call
- nsigmas=6: window size for plain numeric convolution, the argument for setConvolutionWindow call

Generic Wrapper Generic1D_pdf

The bare RooAbsPdf could be easily converted to Ostap-form using the generic wrapper Generic1D_pdf:

```
bare = ROOT.RooGaussian('Gauss','Gaussian', x , mean , sigma )
gauss = Generic1D_pdf ( pdf = bare , xvar = x )
gauss.draw() ## one can immediately use the full power of Ostap-PDF
```

In a similar way there are generic wrappers for 2D and 3D -models:

```
bare2D = ...
bare3D = ...
ostap_2d = Generic2D_pdf ( pdf = bare2D , xvar = x , xvar = y )
ostap_3d = Generic2D_pdf ( pdf = bare3D , xvar = x , xvar = y , zvar = z )
```

2D and 3D -cases

For 2D and 3D cases there are base classes PDF2 and PDF3 that in turn inhetic from PDF and gets all the nice functionality. Of course several new method specific for 2D and 3D -cases are added and the behavior of some 1D -specific methods is fixed.

Compound fit models

1D -case

Ostap offers a very easy way to build the compound fit models from the individual components. E.g.the case of the trivial fit model that consists of one signal and one background components:

```
signal = ...
background = ...
model = Fit1D ( signal = signal , background = backround ) ## <--- HERE!
dataset = ...
result , frame = model.fitTo ( dataset , draw = True ) ## fit and vizualize</pre>
```

The fit model can contains several *signal* and *backround* components, and also *other* components :

In this case several *signal*, *backgrounds* and/or *others* components can be combined into single *signal*, *backround* and/or *others* components:

In practice it is very convinients approach is several signal/background/other componens are specified.

On default extended `RooAddPdf' fit model is created, however, one can force non-extended model:

```
model = Fit1D ( extended = False , ... )
```

In this case one can also instruct the class Fit1D to create recursive (default) or non-recursive fit fractions:

```
model = Fit1D ( extended = False , recursive = False , ... )
```

All components (signal/background/others) can be specified as Ostap-based models. Also one can provide them in a form of bare RooAbsPdf, but for this case one needs to provide also xvar -variable

```
mass = ROOT.RooRealVar('mass','mass',2,3)
gauss = ROOT.RooGaussian( 'Gauss', 'Gauss', mass , ... )
model = Fit1D( signal = gauss , xvar = mass , ... )
```

For background components there is also an alternative way to specify it:

- None : RooPolynomial of zero degree (uniform distribution) will be created and used as background component
 - Attention: background=None does *not* imply the absence of background component
- negative integer n : Ostap model PolyPos_pdf will be created and used as *background* component. This model corresponds to the *positive* polynomial of degree -n . The polinomial is constrained to be non-negative for the whole considered interval of xvar . This constraint allows rather robust and stable fits, especially for the low-statistics case.
- non-negative integer n: Ostap model Bkg_pdf, that is a product of the exponential function and the *positive* polynomial of degree
 n will be created and used as *background* component. Note:
 - The background=0 case corresponds to simple exponential backtround

- Since the polynomial is constrained to be *non-negative* this PDF is very stable and robust, especually for the low-statistic case,
- as RooAbsReal object, in this case it is interpreted as the exponental slope

Actually, the separation into signal, background and other components is a bit arbitrary. However it is helpful for

- to define the meaningful names for the fit parameters
- to separate different components for visualisation, since different styles (lines, colors, etc) are used for different categories

Access to the model components

The individual components can be accessed using python properties

```
gaudd = Gauss_pdf ( ...
model = Fit1D ( signal = gauss , ... )
print model.signal.sigma ## get sigma of Gauss
print model.signal.mean ## get mean of Gauss
```

Fit parameters

The parametters of the created RooAddPdf can be accessed via python properties, e.g. for extended fits:

```
gaudd = Gauss_pdf ( ...
model = Fit1D ( signal = gauss , ... )
print 'signal yield(s):' , model.S
print 'background(s):' , model.B
print 'others: ' , model.C
model.S = 100  ## set value of signal component to be 100 events
model.B.fix(50) ## fix the yield of the background component at 50 events
model.draw()
```

Depending on the number of corresponsing componens and flags <code>combine_signals</code> , <code>combine_backgrounds</code> , <code>combine_others</code> these properties can be <code>scalar</code> values or arrays/tuples.

For combine_signal=True, combine_backgrounds =True, combine_others=True cases oen also gets properties fs, fB and fc that corresponds to the fractions of individual signal/backgroud/others components for the compound signal/signal/backgroud/others.

For non-extended fits, the main parameters are fractions:

```
gaudd = Gauss_pdf ( ...
model = Fit1D ( signal = gauss , ... , extended = False )
...
print 'fractions:' , model.F
```

_Extended multi-component fit model

One can define some initial setting for fit-parameters:

```
model_ext1.S[0].value = 5000
model_ext1.S[1].value = 5000
model_ext1.S[2].value = 5000

model_ext1.B[0].value = 1700
model_ext1.B[1].value = 2300

model_ext1.C[0].value = 500
model_ext1.C[1].value = 400
```

The fit itself is trivial

```
r, f = model_ext1.fitTo ( dataset , draw = False , silent = True )
r, f = model_ext1.fitTo ( dataset , draw = False , silent = True )
```

And accessing fit results is also simple:

Extended fit model with compound components

Setting the initial values of fit parameters is trivial:

```
model_ext2.S = 5000
model_ext2.B = 4200
model_ext2.C = 700

model_ext2.fS[0].value = 0.33
model_ext2.fS[1].value = 0.50

model_ext2.fB[0].value = 0.40
model_ext2.fC[0].value = 0.60
```

The fit itself and access to fit parameters is the same as above.

Non-extended multi-component fit model with non-recursive fit-fractions

Setting the initial values of fit-fractions:

```
model_ne1.F[0].value = 0.25
model_ne1.F[1].value = 0.25
model_ne1.F[2].value = 0.25
model_ne1.F[3].value = 0.08
model_ne1.F[4].value = 0.12
model_ne1.F[5].value = 0.05
```

The fit itself and access to fit parameters is the same as above.

Non-extended multi-component fit model with recursive fit-fractions

Setting initial fit parameters:

```
model_ne2.F[0].value = 0.25
model_ne2.F[1].value = 0.33
model_ne2.F[2].value = 0.50
model_ne2.F[3].value = 0.37
model_ne2.F[4].value = 0.74
model_ne2.F[5].value = 0.50
```

The fit itself and access to fit parameters is the same as above.

Non-extended fit model with compound components and non-recursive fit-fractions

Setting the initial values:

```
model_ne3. F[0].value = 0.75
model_ne3. F[1].value = 0.30
model_ne3.fS[0].value = 0.33
model_ne3.fS[1].value = 0.50
model_ne3.fB[0].value = 0.41
model_ne3.fC[0].value = 0.58
```

The fit itself and access to fit parameters is the same as above.

Non-extended fit model with compound components and recursive fit-fractions

Setting the initial values:

```
model_ne4. F[0].value = 0.75

model_ne4. F[1].value = 0.80

model_ne4.fS[0].value = 0.33

model_ne4.fS[1].value = 0.50

model_ne4.fB[0].value = 0.41

model_ne4.fC[0].value = 0.50
```

The fit itself and access to fit parameters is the same as above.

All the ways to deal with Fit1D objects are illustrated here:

```
# tests for various multicomponents models
7
8
  # @author Vanya BELYAEV Ivan.Belyaeve@itep.ru
9
  # @date 2014-05-11
   """Tests for various multicomponent models
11
13
   14
   __version__ = "$Revision:"
15
   __author__ = "Vanya BELYAEV Ivan.Belyaev@itep.ru"
16
   __date__ = "2014-05-10"
17
   __all__
           = () ## nothing to be imported
   18
19
   import ROOT, random
   from
       Ostap.PyRoUts import *
   from
       Ostap.Utils import rooSilent
   # logging
   24
25
   from AnalysisPython.Logger import getLogger
26
   if '__main__' == __name__ or '__builtin__' == __name__ :
     logger = getLogger ( 'Ostap.TestComponents' )
   else :
28
29
      logger = getLogger ( __name__ )
   # ------
31
   logger.info ( 'Test for multi-component models from Analysis/Ostap')
   ## make simple test mass
34
        = ROOT.RooRealVar ( 'test_mass' , 'Some test mass' , 0 , 10 )
36
   ## book very simple data set
   varset = ROOT.RooArgSet ( mass )
37
38
   dataset = ROOT.RooDataSet ( dsID() , 'Test Data set-0' , varset )
40
   mmin, mmax = mass.minmax()
41
  ### fill it
42
   m1 = VE(3, 0.300**2)
43
   m2 = VE(5, 0.200**2)
44
45
   m3 = VE(7, 0.100**2)
46
47
   for i in xrange(0,5000) :
48
    for m in (m1, m2, m3) :
49
        mass.value = m.gauss ()
        dataset.add ( varset )
50
51
```

```
52
    for i in xrange(0,5000):
53
        mass.value = random.uniform ( *mass.minmax() )
54
        dataset.add ( varset
    logger.info ('Dataset: %s' % dataset )
57
    import Ostap.FitModels as Models
58
    signal_1 = Models.Gauss_pdf ( 'G1' , xvar = mass , mean = m1.value() , sigma = m1.errpr() )
59
    signal_2 = Models.Gauss_pdf ( 'G2'
                                       , xvar = mass , mean = m2.value() , sigma = m2.errpr() )
60
    signal_3 = Models.Gauss_pdf ( 'G3' , xvar = mass , mean = m3.value() , sigma = m3.errpr() )
62
63
    wide_1 = Models.Gauss_pdf ( 'GW1' , xvar = mass , mean = 1.0 , sigma = 2 )
    wide_2 = Models.Gauss_pdf ( 'GW2' , xvar = mass , mean = 9.0 , sigma = 3 )
64
65
    narrow_1 = Models.Gauss_pdf ( 'GN1' , xvar = mass , mean = 4.0 , sigma = 1 )
66
67
    narrow_2 = Models.Gauss_pdf ( 'GN2' , xvar = mass , mean = 6.0 , sigma = 1 )
68
69
    logger.info ( 'Test the extended fit with many components' )
    model_ext1 = Models.Fit1D (
71
72
        name
                        = 'EXT1'
        signal
                        = signal_1 ,
                       = [ signal_2 , signal_3 ] ,
74
        othersignals
        background
                        = wide_1
76
        otherbackgrounds = [ wide_2 ] ,
77
        others
                       = [ narrow_1 , narrow_2 ] ,
78
        )
    model_ext1.S[0].value = 5000
79
80
    model_ext1.S[1].value = 5000
81
    model_ext1.S[2].value = 5000
82
    model_ext1.B[0].value = 1700
83
84
    model_ext1.B[1].value = 2300
85
    model_ext1.C[0].value = 500
86
    model_ext1.C[1].value = 400
87
88
    r, f = model_ext1.fitTo ( dataset , draw = False , silent = True )
    r, f = model_ext1.fitTo ( dataset , draw = False , silent = True )
    logger.info ( 'Model %s Fit results:\n#%s ' % ( model_ext1.name , r ) )
                                [S]:', model_ext1.S
    print 'Signals
                                [B]:', model_ext1.B
    print 'Backgrounds
94
    print 'Components
                               [C]:', model_ext1.C
                                [F]:' , model_ext1.F
    print 'Fractions
                               [fS]:' , model_ext1.fS
    print 'Signal fractions
96
    print 'Background fractions [fB]:' , model_ext1.fB
97
```

```
print 'Component fractions [fC]:' , model_ext1.fC
98
                           [yields]:' , model_ext1.yields
99
     print 'Yields
                        [fractions]:' , model_ext1.fractions
     print 'Fractions
101
     104
     logger.info ( 'Test the extended fit with compound components' )
     model_ext2 = Models.Fit1D (
        name
                           = 'EXT2'
        signal
                              signal_1 ,
        othersignals
                          = [ signal_2 , signal_3 ] ,
109
        background
                              wide_1
110
        otherbackgrounds
                         = [ wide_2 ] ,
                          = [ narrow_1 , narrow_2 ] ,
111
        others
112
        combine_signals
                          = True
        combine_backgrounds = True
114
115
        combine_others
                          = True
116
         )
117
     model_ext2.S = 5000
118
     model_ext2.B = 4200
     model_ext2.C = 700
119
121
     model_ext2.fS[0].value = 0.33
     model_ext2.fS[1].value = 0.50
123
124
     model_ext2.fB[0].value = 0.40
125
     model_ext2.fC[0].value = 0.60
126
     r, f = model_ext2.fitTo ( dataset , draw = False , silent = True )
     r, f = model_ext2.fitTo ( dataset , draw = False , silent = True )
128
     logger.info ( 'Model %s Fit results:\n#%s ' % ( model_ext2.name , r ) )
129
     print 'Signals
130
                               [S]:', model_ext2.S
131
     print 'Backgrounds
                               [B]:', model_ext2.B
                               [C]:' , model_ext2.C
132
     print 'Components
                               [F]:' , model_ext2.F
     print 'Fractions
133
                              [fS]:' , model_ext2.fS
134
     print 'Signal fractions
     print 'Background fractions [fB]:' , model_ext2.fB
     print 'Component fractions [fC]:' , model_ext2.fC
137
     print 'Yields
                           [yields]:' , model_ext2.yields
138
     print 'Fractions
                        [fractions]:' , model_ext2.fractions
139
140
     141
     logger.info ( 'Test non-extended fit with all components, non-recursive' )
     model_ne1 = Models.Fit1D (
142
143
        name
                           = 'NE1'
```

```
144
         signal
                                signal 1
145
         othersignals
                            = [ signal_2 , signal_3 ] ,
146
         background
                                wide_1 ,
         otherbackgrounds
                            = [ wide_2 ] ,
147
         others
                            = [ narrow_1 , narrow_2 ] ,
148
149
         ##
         extended
                            = False ,
         recursive
                            = False
         )
154
     model_ne1.F[0].value = 0.25
155
     model_ne1.F[1].value = 0.25
     model_ne1.F[2].value = 0.25
156
157
     model_ne1.F[3].value = 0.08
     model_ne1.F[4].value = 0.12
158
159
     model_ne1.F[5].value = 0.05
161
     r, f = model_ne1.fitTo ( dataset , draw = False , silent = True )
     r, f = model_ne1.fitTo ( dataset , draw = False , silent = True )
162
163
     logger.info ( 'Model %s Fit results:\n#%s ' % ( model_ne1.name , r ) )
164
     print 'Signals
                                 [S]:', model_ne1.S
                                 [B]:', model_ne1.B
     print 'Backgrounds
     print 'Components
                                 [C]:', model_ne1.C
                                 [F]:' , model_ne1.F
     print 'Fractions
168
     print 'Signal fractions
                                [fS]:' , model_ne1.fS
169
     print 'Background fractions [fB]:' , model_ne1.fB
     print 'Component fractions [fC]:' , model_ne1.fC
170
171
     print 'Yields
                            [yields]:' , model_ne1.yields
172
     print 'Fractions
                         [fractions]:' , model_ne1.fractions
173
     174
175
     logger.info ( 'Test non-extended fit with all components, recursive' )
176
     model_ne2 = Models.Fit1D (
177
         name
                            = 'NE2'
178
         signal
                               signal_1
179
         othersignals
                            = [ signal_2 , signal_3 ] ,
                            = wide_1 ,
         background
         otherbackgrounds
                            = [ wide_2 ] ,
         others
                            = [ narrow_1 , narrow_2 ] ,
         ##
184
         extended
                            = False ,
         recursive
                            = True ,
186
         )
187
     model_ne2.F[0].value = 0.25
     model_ne2.F[1].value = 0.33
```

```
model ne2.F[2].value = 0.50
     model_ne2.F[3].value = 0.37
     model_ne2.F[4].value = 0.74
     model_ne2.F[5].value = 0.50
193
194
195
     r, f = model_ne2.fitTo ( dataset , draw = False , silent = True )
196
     r, f = model_ne2.fitTo ( dataset , draw = False , silent = True )
     logger.info ( 'Model %s Fit results:\n#%s ' % ( model_ne2.name , r ) )
198
     print 'Signals
                                [S]:', model_ne2.S
199
     print 'Backgrounds
                                [B]:', model_ne2.B
                                [C]:' , model_ne2.C
     print 'Components
201
     print 'Fractions
                                [F]:' , model_ne2.F
                                [fS]:' , model_ne2.fS
     print 'Signal fractions
     print 'Background fractions [fB]:' , model_ne2.fB
203
     print 'Component fractions [fC]:' , model_ne2.fC
204
     print 'Yields
                            [yields]:' , model_ne2.yields
                         [fractions]:' , model_ne2.fractions
     print 'Fractions
207
209
     210
     logger.info ( 'Test non-extended fit with compound components, non-recursive' )
     model_ne3 = Models.Fit1D (
212
         name
                            = 'NE2'
213
         signal
                            = signal_1
                            = [ signal_2 , signal_3 ] ,
214
         othersignals
215
         background
                                wide_1
                            =
216
         otherbackgrounds = [ wide_2 ] ,
         others
                            = [ narrow_1 , narrow_2 ] ,
218
219
         combine_signals
                            = True
220
         combine_backgrounds = True
         combine_others
                            = True
222
         ##
         extended
                            = False
224
         recursive
                            = False
225
         )
226
227
     model_ne3.F[0].value = 0.75
228
     model_ne3.F[1].value = 0.30
229
230
     model_ne3.fS[0].value = 0.33
231
     model_ne3.fS[1].value = 0.50
232
233
     model_ne3.fB[0].value = 0.41
     model_ne3.fC[0].value = 0.58
234
235
```

```
236
      r, f = model_ne3.fitTo ( dataset , draw = False , silent = True )
      r, f = model_ne3.fitTo ( dataset , draw = False , silent = True )
238
      logger.info ( 'Model %s Fit results:\n#%s ' % ( model_ne3.name , r ) )
      print 'Signals
                                   [S]:', model_ne3.S
239
                                   [B]:', model_ne3.B
     print 'Backgrounds
240
241
      print 'Components
                                   [C]:' , model_ne3.C
242
      print 'Fractions
                                   [F]:' , model_ne3.F
243
                                  [fS]:', model_ne3.fS
      print 'Signal fractions
244
      print 'Background fractions [fB]:' , model_ne3.fB
      print 'Component fractions [fC]:' , model_ne3.fC
245
      print 'Yields
                              [yields]:' , model_ne3.yields
246
247
      print 'Fractions
                           [fractions]:' , model_ne3.fractions
248
249
250
      logger.info ( 'Test non-extended fit with compound components, recursive' )
      model_ne4 = Models.Fit1D (
253
         name
                              = 'NE4'
         signal
                                  signal_1
255
         othersignals
                             = [ signal_2 , signal_3 ] ,
256
         background
                                 wide_1
         otherbackgrounds
                             = [ wide_2 ] ,
258
         others
                             = [ narrow_1 , narrow_2 ] ,
259
260
         combine_signals
                             = True
261
         combine_backgrounds = True
262
          combine_others
                              = True
263
         ##
264
         extended
                              = False
265
          recursive
                              = True
          )
267
268
      model_ne4.F[0].value = 0.75
     model_ne4.F[1].value = 0.80
270
271
      model_ne4.fS[0].value = 0.33
      model_ne4.fS[1].value = 0.50
274
      model_ne4.fB[0].value = 0.41
275
      model_ne4.fC[0].value = 0.50
276
277
      r, f = model_ne4.fitTo ( dataset , draw = False , silent = True )
278
      r, f = model_ne4.fitTo ( dataset , draw = False , silent = True )
      logger.info ( 'Model %s Fit results:\n#%s ' % ( model_ne4.name , r ) )
279
                                   [S]:', model_ne4.S
      print 'Signals
     print 'Backgrounds
                                   [B]:', model_ne4.B
```

```
print 'Components
                              [C]:' , model_ne4.C
                              [F]:', model_ne4.F
     print 'Fractions
                             [fS]:', model_ne4.fS
284
     print 'Signal fractions
     print 'Background fractions [fB]:' , model_ne4.fB
     print 'Component fractions [fC]:' , model_ne4.fC
286
     print 'Yields
                          [yields]:' , model_ne4.yields
287
     print 'Fractions
288
                       [fractions]:' , model_ne4.fractions
289
290
291
     292
     # The END
293
4
components.py hosted with ♥ by GitHub
                                                                        view raw
```

The corresponding output can be inspected here

2D -case

Generic 2D -case

Symmetric 2D -case

3D -case

Generic 3D -case

Symmetric 3D -case

Mixed-symmetry 3D -case

sPlot

Using sPlot is rather trivial in Ostap:

```
dataset = ...
model = Fit1D ( signal = ... , backgrund = ... )
model.fitTo ( dataset )
print datatset
model.sPlot ( dataset ) ## <--- HERE
print datatset ## <--- note appearence of new variables</pre>
```

Contributing

ostap-tutorials is an open source project, and we welcome contributions of all kinds:

- New lessons;
- Fixes to existing material;
- Bug reports; and
- Reviews of proposed changes.

By contributing, you are agreeing that we may redistribute your work under these licenses. You also agree to abide by our contributor code of conduct.

Getting Started

- 1. We use the fork and pull model to manage changes. More information about forking a repository and making a Pull Request.
- 2. To build the lessons please install the dependencies.
- 3. For our lessons, you should branch from and submit pull requests against the master branch.
- 4. When editing lesson pages, you need only commit changes to the Markdown source files.
- 5. If you're looking for things to work on, please see the list of issues for this repository. Comments on issues and reviews of pull requests are equally welcome.

Dependencies

To build the lessons locally, install the following:

1. Gitbook

Install the Gitbook plugins:

\$ gitbook install

Then (from the ostap-tutorials directory) build the pages and start a web server to host them:

\$ gitbook serve

You can see your local version by using a web-browser to navigate to http://localhost:4000 or wherever it says it's serving the book.

The title

Learning Objectives

- The starterkit lessons all start with objectives about the lesson
- Objective 2 with some *formatted* **text** *like* this

Basic formatting

You can make **bold**, *italic* and strikethrough text. Add relative links like this one and absolute links in a couple of different ways.

Have bulleted lists:

- Point 1
- Point 2
 - o Sub point
 - Sub point
 - Sub point
- Point 2

Use numbered lists:

- 1. First
- 2. Second
 - i. Second first
 - i. Second first first
 - ii. Second second
- 3. Third

LaTeX

You can use inline LaTeX maths such as talking about the decay $\$D^{*+} \subset D^0 \subset K^{-}$ \pi^{+} \right

Code highlighting

And have small lines of code inline like saying print("Hello world") or have multiple lines with syntax highlighting for python:

```
import sys

def stderr_print(string):
    sys.stderr.write(string)

stderr_print("Hello world")
```

bash:

```
lb-run Bender/latest $SHELL
dst_dump -f -n 100 my_file.dst 2>&1 | tee log.log
```

Callouts

• Summary point 1

Prequisites	
 Prequisite 1 Prequisite 2 	
Objectives	
 Objective 1 Objective 2 	
Challenge	
Set a challenge here, and the solution will remain hidden until it's clicked • How to print?	
Solution	
The answer is: print("Hello world")	
Extra details that are hidden by default	
Some extra details	
Keypoints	

Quotes

This was said by someone

Tables

Simple tables are possible

First Header	Second Header
Content from cell 1	Content from cell 2
Content in the first column	Content in the second column

Images



Section types

This is a section

Subsections

And a subsection

Subsubsections

And a subsubsection