Table of Contents

The Ostap tutorials	1.1
Getting started	2.1
Values with uncertainties	2.1.1
Simple operations with histograms	2.1.2
Simple operations with trees	2.1.3
Persistency	2.1.4
Using RooFit	3.1
Useful decorations	3.1.1
Models	3.1.2
Contributing	4.1
Examples of formatting	4.2
Download PDF	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 indirectty 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 covariance submatrix

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

Also the simple math with fiting parameters is supported

RooRealVar & friends

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

```
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
```

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 $\$D^{*+} \subset D^0 \subset K^{-}$ yiéhtarrow $K^{-} \in K^{-}$

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