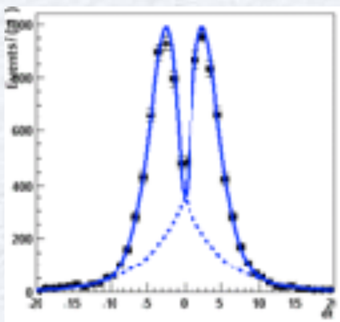
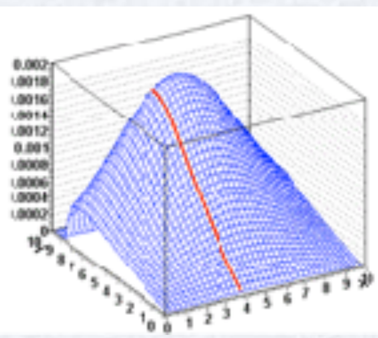
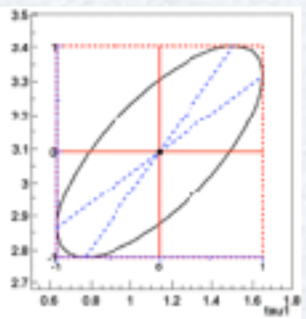
Statistics for Particle Physics Analyses: Introduction to Computing Examples

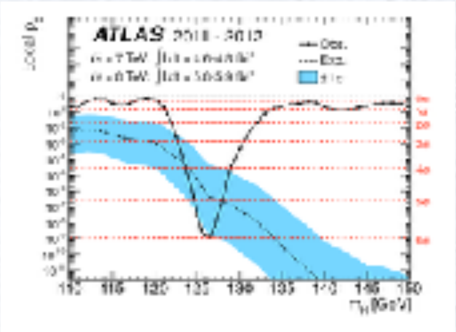
Louis Lyons (Imperial College), Lorenzo Moneta (CERN)

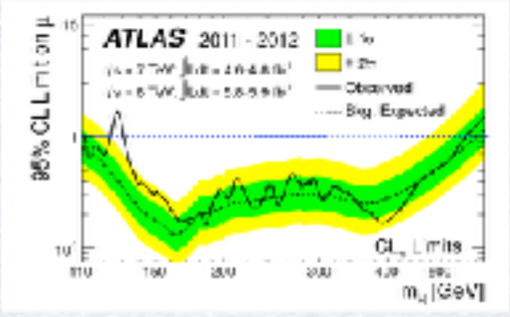
CERN Academic Training November 2016











Computing Examples

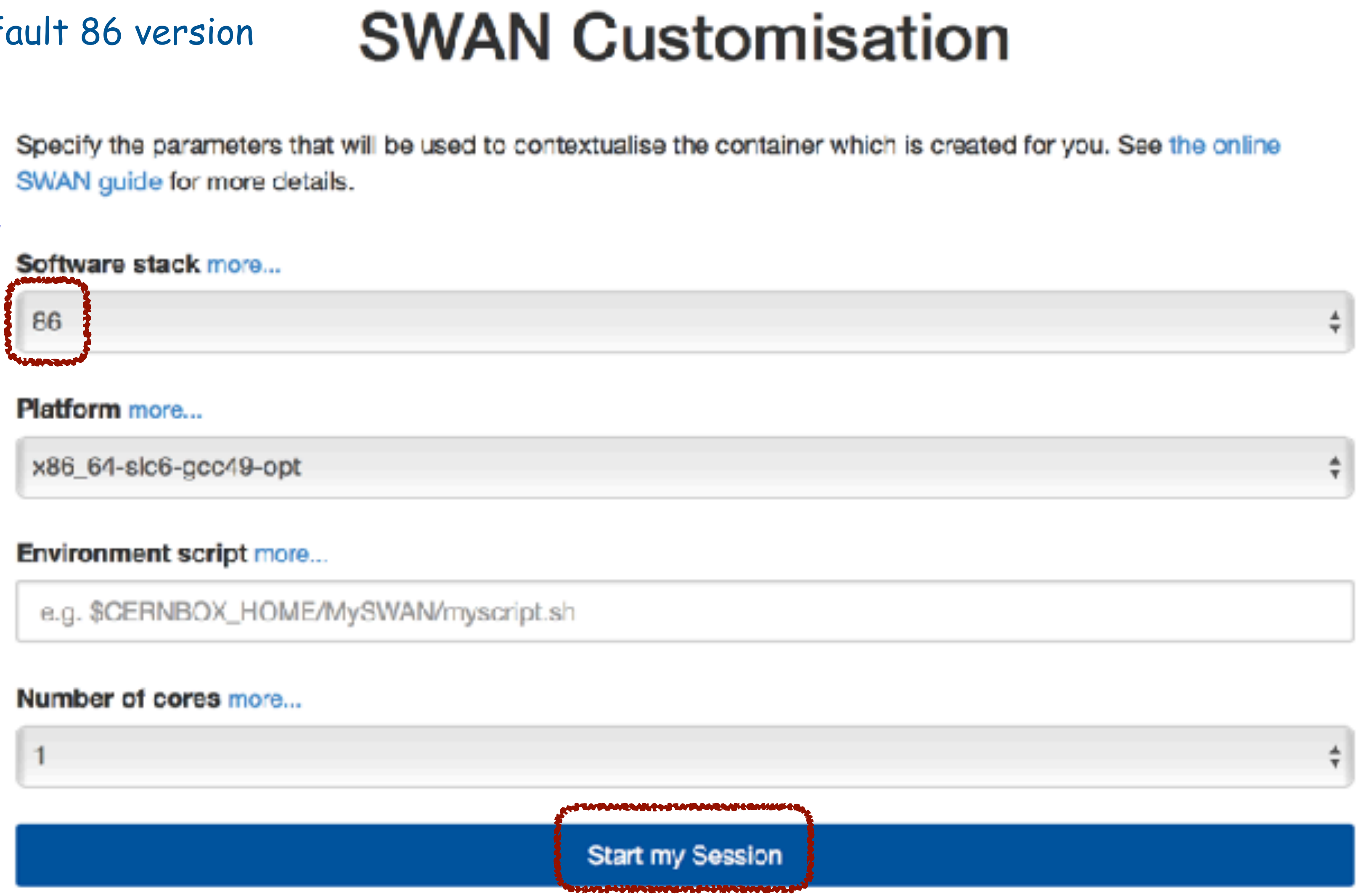
- Work on computing examples based on ROOT including RooFit and RooStats
- Examples are interactive demos made of notebooks
- Running an a new CERN notebook service
  - **SWAN** (swan.cern.ch)
  - files will be stored in **cernbox**
- Log-in into SWAN using your cern credential
  - if you don't have **cernbox**, a link to create it is automatically provided
  - if you don’t have a CERN computing account, some special accounts are provided for the course

SWAN

- swan.cern.ch

Starting SWAN

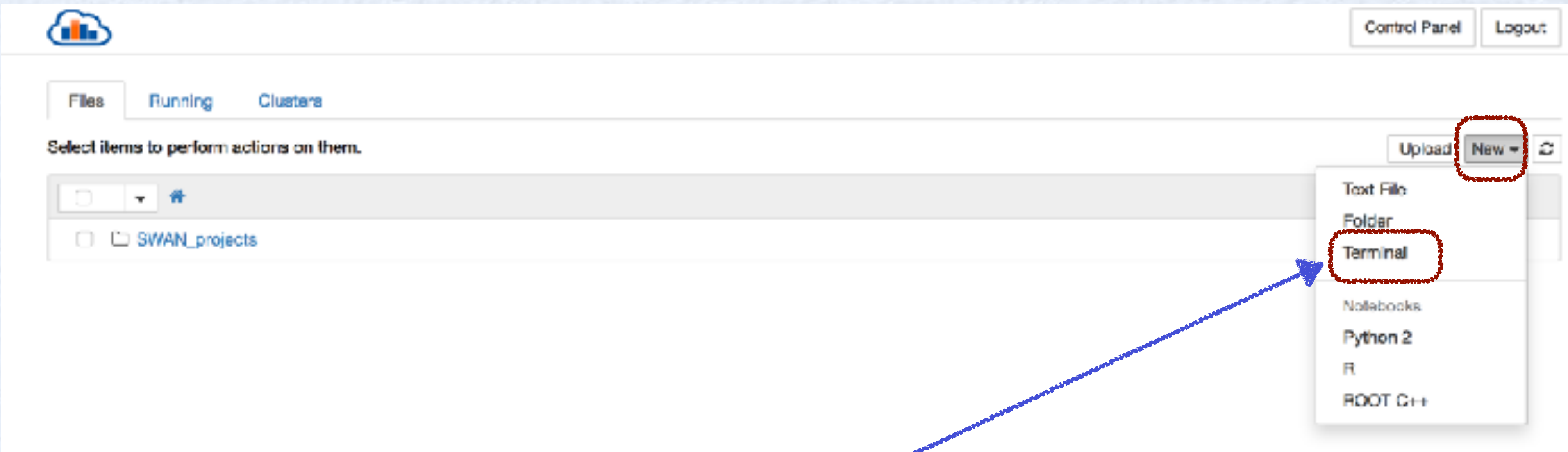
make sure to have selected
the default 86 version



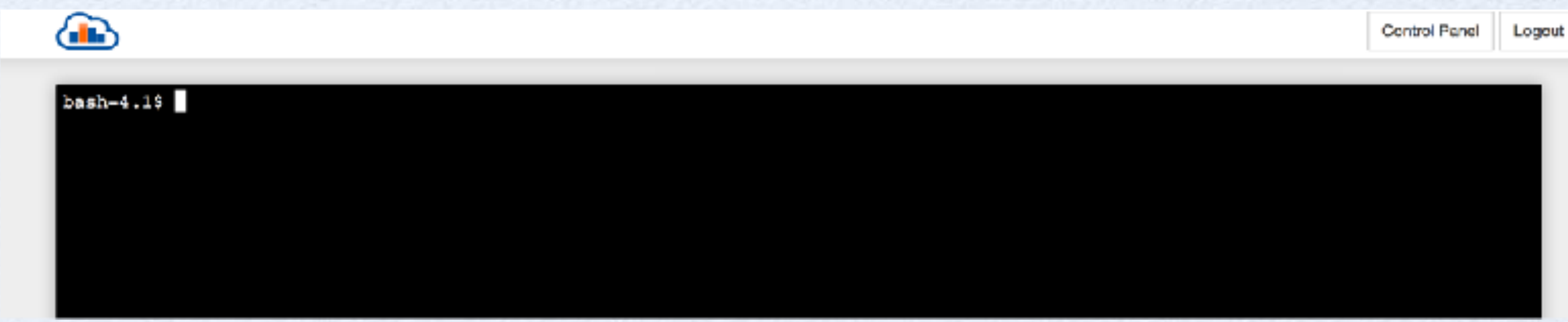
click here to start

Starting a Terminal in SWAN

After login cernbox home directory will be visible



Start a terminal window



Downloading the Notebooks

- After starting the terminal, download the notebooks for today's session.
- Download from my Web page using *wget*:

wget http://www.cern.ch/moneta/swan/Day1.tar.gz

```
bash-4.1$ wget -nv http://www.cern.ch/moneta/swan/Day1.tar.gz
2016-11-27 12:30:53 URL:http://moneta.web.cern.ch/moneta/swan/Day1.tar.gz [200496/200496] -> "Day1.tar.gz" [1]
bash-4.1$ ls
Day1.tar.gz  SWAN_projects
bash-4.1$ tar -zxvf Day1.tar.gz
Day1/
Day1/ExtendedMLFit.ipynb
Day1/LikelihoodAsGoF.ipynb
Day1/LikelihoodFit.ipynb
Day1/PoissonCoverage.ipynb
bash-4.1$
```

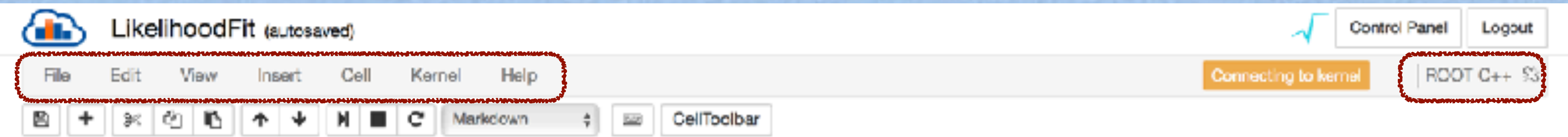
- untar the downloaded archive file

tar -zxvf Day1.tar.gz

Starting a Notebook

- After downloaded and extracted the notebooks, go back to SWAN home page and select the desired notebook

Starting a Notebook



Likelihood fit to exponential data

This notebook shows how to perform a likelihood fit to a data sample distributed according to an exponential distribution. The data sample consis of n events (e.g. n=20) generated randomly using the ROOT exponential random number generator. A negative log-likelihood function is constructed using the given data set. The log-likelihood function is a function of the fit parameter `tau`, which is the slope of the exponential distribution and the minimum of the function is found using simple tools provided by ROOT.
The example shows also how to extimate the error of the parameter using both the second derivative of the log-likelihood function or the `DeltaL=0.5` rule. By increasing the number of generated events one can see how the likelihood shape approaches a parabola and the errors estimated in the two methods become equal.

1. Event generation

First we generate n events using an exponential distribution with mean life tau=2

In [1]:
```
int n = 20;
double tau = 2;
std::vector<double> x(n);
TRandom3 r(0);
for (int i = 0; i < n; ++i)
    x[i] = r.Exp(tau)
```

In [2]:
```
// comment this line for avoiding printing the random generated values
x
```

```
(std::vector<double> &) { 0.736121, 0.257743, 0.211415, 2.53096, 0.656154, 0.237043, 2.76202, 0.672238, 1.22943, 0.15
9545, 8.65069, 0.146272, 2.50028, 3.00072, 0.492201, 0.766004, 2.56399, 1.05053, 0.126264, 0.492685 }
```

References

- Presentation about Jupiter notebooks and SWAN
  - https://cernbox.cern.ch/index.php/s/cpqdIOaNIwrlzJu
- Basics of ROOT for fitting
  - https://cernbox.cern.ch/index.php/s/SjJnYSkJrY9Kv8y
- RooFit tutorial slides
  - https://cernbox.cern.ch/index.php/s/kzGv4406IrpFpVM
- RooStats tutorial slides
  - https://cernbox.cern.ch/index.php/s/9DWJCObingTWB3r