Kirchhoff-Insitut für Physik
Physikalisches Insitut

Winter semester 2013-14
KIP CIP Pool (1.401)

Exercises for Statistical Methods in Particle Physics

http://www.physi.uni-heidelberg.de/~nberger/teaching/ws13/statistics/statistics.php

Dr. Niklaus Berger (nberger@physi.uni-heidelberg.de)
Dr. Oleg Brandt (obrandt@kip.uni-heidelberg.de)

Exercise 6: Extracting parameters by maximising the likelihood

18. November 2013
Hand-in solutions by 14:00, 25. November 2013

Please send your solutions to `obrandt@kip.uni-heidelberg.de` by 25.11.2013, 16:00, punctually. Make sure that you use *SMIPP:Exercise06* as subject line. Please put each macro into one separate .C or .py file, which can easily be tested. If you are providing C++ code, please do so in compiled mode. If plots are requested, please include print statements to produce pdf files in your code, and provide the plots separately. Please add comments to your source code explaining the steps. Test macros and programs before sending them off...

In this exercise, we will extract a fundamental parameter of the Standard Model (SM), the top quark mass m_t, from the measured production cross section for $t\bar{t}$ pairs, $\sigma_{t\bar{t}}$, with the maximum likelihood method. For this, we will use the measurement of $\sigma_{t\bar{t}}$ in early LHC data of 35 pb^{-1} with the ATLAS experiment [1], and compare it to the world's first next-to-next-to-leading order (NNLO) prediction with next-to-leading logarithmic corrections (NLL) which became available earlier this year [2]. The exciting bit of this extraction is that it allows to determine m_t in a well-defined mass definition scheme, the pole scheme. In this scheme, m_t corresponds directly to the mass pole of the top propagator in the SM Lagrangian. In contrast, conventional, more precise methods use a theoretically less well-defined mass concept.

1 Graphical solution

First, we want to solve the problem graphically, i.e. by simply plotting the experimentally measured $\sigma_{t\bar{t}}^{\rm exp}(m_t)$ dependence alongside the theory prediction $\sigma_{t\bar{t}}^{\rm theo}(m_t)$, and reading off the most probable m_t value from the plot.

1.1 Plotting experimental data

First, we want to graphically represent the measured $\sigma_{t\bar{t}}^{\rm exp}(m_t)$ dependence given in Table 1. For this, we want to split the statistical and systematic uncertainty. For simplicity, we assume a uniform statistical uncertainty of 5%, a systematic uncertainty of 12% and a systematic uncertainty from the luminosity determination of 3%. The total systematic uncertainty is obtained by adding the latter two values quadratically. Analogously, the total experimental uncertainty can be calculated.
To start, we diplay the measured $\sigma_{t\bar{t}}^{\rm exp}(m_t)$ points with the total experimental uncertainty as error bars, singling out the statistical uncertainty. This can be done by e.g. drawing a horizontal line through the total uncertainty error bars (refer to the ROOT documentation of the `TGraphPainter` class). Technically, one can plot on top of each other two `TGraphError` objects

m_t [GeV]	$\sigma_{t\bar{t}}^{\exp}(m_t)$ [pb^{-1}]
140	279.6
150	240.7
160	219.0
170	200.4
172.5	186.3
180	185.5
190	173.2
200	159.7
210	154.9

Tabelle 1: The measured $\sigma_{t\bar{t}}^{\exp}(m_t)$ dependence from [1].

with identical central values, where one has only statistical uncertainties and the other the total uncertainty.
In the next step, we fit these experimental points with the equation:

$$\sigma_{t\bar{t}}(m_t) = \sigma_{t\bar{t}}(m_t^{\mathrm{ref}}) \left(\frac{m_t^{\mathrm{ref}}}{m_t}\right)^4 \left(1 + a_1 \frac{m_t - m_t^{\mathrm{ref}}}{m_t^{\mathrm{ref}}} + a_2 \left(\frac{m_t - m_t^{\mathrm{ref}}}{m_t^{\mathrm{ref}}}\right)^2\right), \tag{1}$$

where $m_t^{\mathrm{ref}} = 172.5$ GeV. In this fit, we consider the total uncertainty. To define the function in Eq. 1, use a `TF1` object and a custom function, as in the previous problem sheet. Please provide the fitted coefficients $\sigma_{t\bar{t}}(m_t^{\mathrm{ref}})$, a_1, and a_2. At this stage, no plot is requested.

1.2 Plotting the theory prediction

This step is "easy": we use the same parametrisation from Eq. 1 to display the $\sigma_{t\bar{t}}^{\mathrm{theo}}(m_t)$ dependence, with $\sigma_{t\bar{t}}(m_t^{\mathrm{ref}}) = 176.2$ GeV, $a_1 = -1.2149$, and $a_2 = 0.874646$. The uncertainty on the theory prediction is 6%. Draw the theory prediction as a `TGraphError` object with a hatched or lightly shaded area representing the uncertainty band (again, refer to the `TGraphPainter` documentation how to do this).

1.3 Piece it all together

Now our plot contains the best fit to experimental data and the theory prediction. Please provide the code `ex_6_1.C/.py` and the plot `ex_6_1.pdf`, and extract the best m_t value graphically (i.e. by eye).

2 Extracting the top quark mass with a maximum likelihood fit

Now that we have the parametrised $\sigma_{t\bar{t}}^{\exp}(m_t)$ and $\sigma_{t\bar{t}}^{\mathrm{theo}}(m_t)$ dependencies with their respective uncertainties $\delta\sigma_{t\bar{t}}(m_t)$ (note you will need to translate the relative uncertainty values into pb^{-1}!), we can proceed to extract m_t by maximising the likelihood:

$$L(m_t) \propto \int f_{\exp}(\sigma_{t\bar{t}}^{\exp}(m_t)|m_t) \cdot f_{\mathrm{theo}}(\sigma_{t\bar{t}}^{\mathrm{theo}}(m_t)|m_t) \mathrm{d}\sigma_{t\bar{t}}$$

where:

- $f_{\exp}(\sigma_{t\bar{t}}^{\exp}(m_t)|m_t)$ is the experimental probability density function constructed using a Gaussian likelihood function centered on the measured $\sigma_{t\bar{t}}^{\exp}$ and having as a width the total experimental uncertainty;

- $f_{\text{theo}}(\sigma_{t\bar{t}}^{\text{theo}}(m_t)|m_t)$ is is the theoretical probability density function constructed using a Gaussian likelihood function centered on the predicted $\sigma_{t\bar{t}}^{\text{theo}}$ and having as a width the associated theory uncertainty.

To perform the actual m_t extraction, we take $-\ln L$ and represent it as a `TF1` object. To find the value that maximises the likelihood, for our purposes it is sufficient to use the `TF1::GetMinimum` method. The 1 SD uncertainty on the extracted value is given by the $m_t^{\pm}$ values where

$$-\ln L(m_t^{\pm}) = \ln L(\hat{m}_t) + 0.5\,,$$

with $\hat{m}_t$ being the m_t value that maximises the likelihood. Represent your results graphically by plotting the $-\ln L$ function, mark $\hat{m}_t$ with an arrow (`TArrow`) and indicate the $\ln L(\hat{m}_t) + 0.5$ line with a dashed `TLine`. What do you obtain ($m_t{}^{+\delta m_t}_{-\delta m_t'}$)? Does it match with your graphical solution? Provide the code that performs the likelihood minimisation and graphically represents the results (`ex_6_2.C/.py`) and the plot (`ex_6_2.pdf`).

3 Cross-checks

3.1 Neglecting experimental uncertainty

A typical cross-check of the extracted result is to perform the extraction of m_t where the experimental uncertainties are neglected. Technically, this can be done by maximising the likelihood where $f_{\text{exp}}(\sigma_{t\bar{t}}^{\text{exp}}(m_t)|m_t)$ is defined with artificially very small, but non-zero experimental uncertainty such that $\delta\sigma_{t\bar{t}}(m_t)^{\text{exp}} \ll \delta\sigma_{t\bar{t}}(m_t)^{\text{theo}}$. What do you find for m_t? How does the likelihood look like qualitatively compared to Problem 6.2?
(provide `ex_6_3_1.C/.py` and `ex_6_3_1.pdf`).

3.2 Neglecting the systematic uncertainty

Another common cross-check to evaluate the impact of experimental systematic uncertainties (and to obtain only the statistical uncertainty on the final result) is to perform the extraction of m_t where only the statistical uncertainties are considered on the experimental side. What do you find for m_t now? Give the m_t value with uncertainties split by statistical and systematic components. How does the likelihood look like qualitatively?
(provide `ex_6_3_1.C/.py` and `ex_6_3_1.pdf`).

Literatur

[1] ATLAS collaboration, *Measurement of the top quark-pair cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV in the semileptonic channel using b-tagging*, ATLAS-CONF-2011-035 (2011).

[2] M. Czakon *et al.*, *The total top quark pair production cross-section at hadron colliders through $O(\alpha_S^4)$*, Phys. Rev. Lett. **110** (2013) 252004, arXiv:1303.6254 [hep-ph].