SOLUTION FOR HOMEWORK ASSIGNMENT NO. 09

Nils Hoyer, Maurice Morgenthaler

Exercise 9.1

In this exercise we have to do a simple hypothesis testing on our own. The test statistic to use is defined by the Neyman-Pearson lemma,

$$\Lambda(\{x\}) = \frac{\mathcal{L}(\{x\} \mid H_1)}{\mathcal{L}(\{x\} \mid H_0)} \tag{1}$$

where $\{x\}$ stands for the data set, H_0 for the hypothesis for *background-only* signal and H_1 for the *background+signal* hypothesis.
The theory predicts the following parameters:

The particle mass of $751\,\mathrm{GeV}$,
a peak-width of $30\,\mathrm{GeV}$,
an exponential background $e^{-a\cdot x}$ with $a = 1 \times 10^{-3}\,\mathrm{GeV}^{-1}$ and
an production rate of 3 signal events for 10 background events.

First, we generate Monte Carlo data. As said before we use an exponential function for the background and an exponential function plus a gaussian function for background and signal

$$f_{\mathrm{bkg}} = A_{\mathrm{bkg}} \cdot \exp(-a \cdot x) \tag{2}$$

$$f_{\mathrm{sig}} = A_{\mathrm{sig}} \cdot \exp\left(\frac{(x-\mu)^2}{\sigma}\right) \tag{3}$$

with the amplitudes A_{bkg} and A_{sig}.
We generate 1000 times 1×10^5 datapoints for the *background-only* and the *background+signal* hypothesis. You can find a visualisation of the first iteration of datapoints in figure 1.

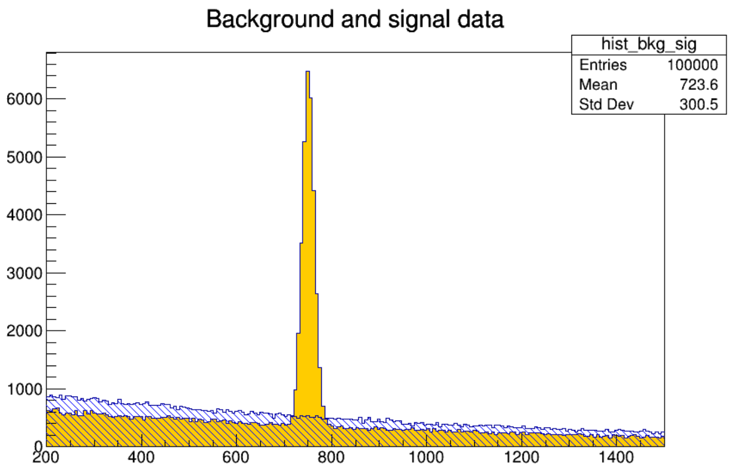

Figure 1: Visualisation of datapoints for the *background-only* (blue, dashed) and *background+signal* (orange, filled) hypothesis. Each histogram contains 1×10^5 datapoints.

Next, we calculate the test statistic Λ for every iteration. This is done via

$$\Lambda = \frac{\sum_{i=0}^{1000} \ln\left(A_{\text{bkg}}\right) - a \cdot x_i + \ln\left(A_{\text{sig}}\right) - \frac{(x_i-\mu)^2}{\sigma}}{\sum_{i=0}^{1000} \ln\left(A_{\text{bkg}}\right) - a \cdot x_i}. \tag{4}$$

The histograms for H_0 and H_1 can be found in figure 2.

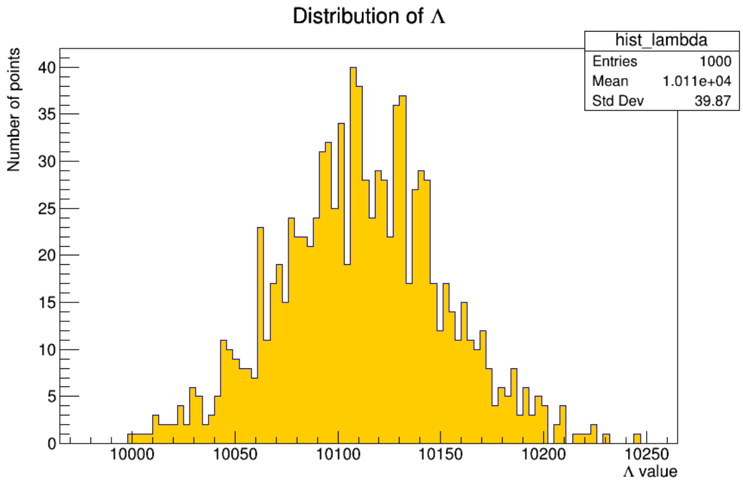

Figure 2: Visualisation distribution of Λ.