

8.811 Stats Assignment

Code, data, and additional figures for this assignment can be found here: https://github.com/julius-heitkoetter/particle_physics_statistics

Problem 1 (*Collaborators: None*)

NOTE: Unless otherwise specified, the data used in this part is the *small* dataset. Instead of showing my work in the writeup, I carefully documented the code in "stats_assignment_part1.py"

Part 1(a)

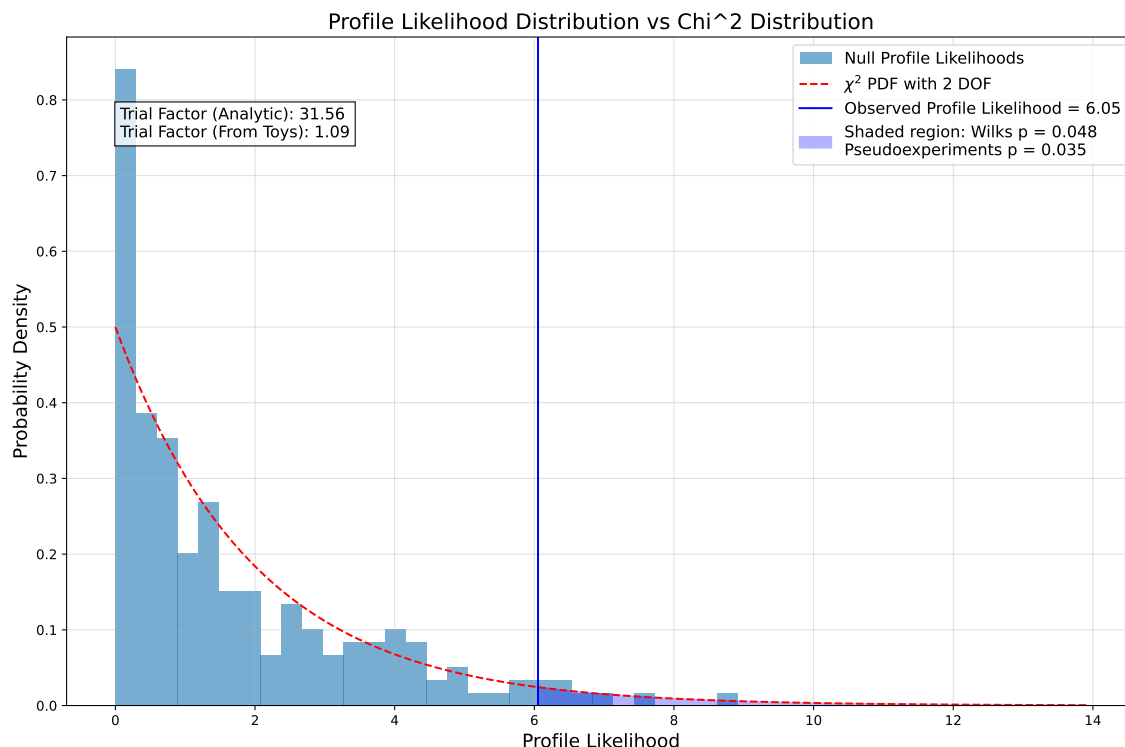
The local p-value of the peak that I see using the likelihood profile ratio as the test statistic and Wilkes theorem is 0.0485.

Part 1(b)

The trial factor I expect using the equations we have seen is class is: 31.56.

Part 1(c)

i Below is the distribution of the profile likelihood under the null hypothesis as well as a blue part denoting the value of the profile likelihood found in data. I've also overlayed a red dotted line displaying a χ^2 distribution with 2 degrees of freedom to demonstrate that the distribution looks like I would expect under Wilkes.



ii As can be read from the figure above, the p-value I got from my pseudoexperiment is 0.035. Additionally, assuming the experiment is in the asymptotic limit, the trial factor according to the pseudo-experiments is 1.09. This can also be read off of the plot above.

Part 1(d)

The new mass range experiment should have half the trial factor of the old mass range experiment.

When I ran the code, I got a ratio of p-values of roughly $5/8$, which is fairly close to the $1/2$ I would have predicted.

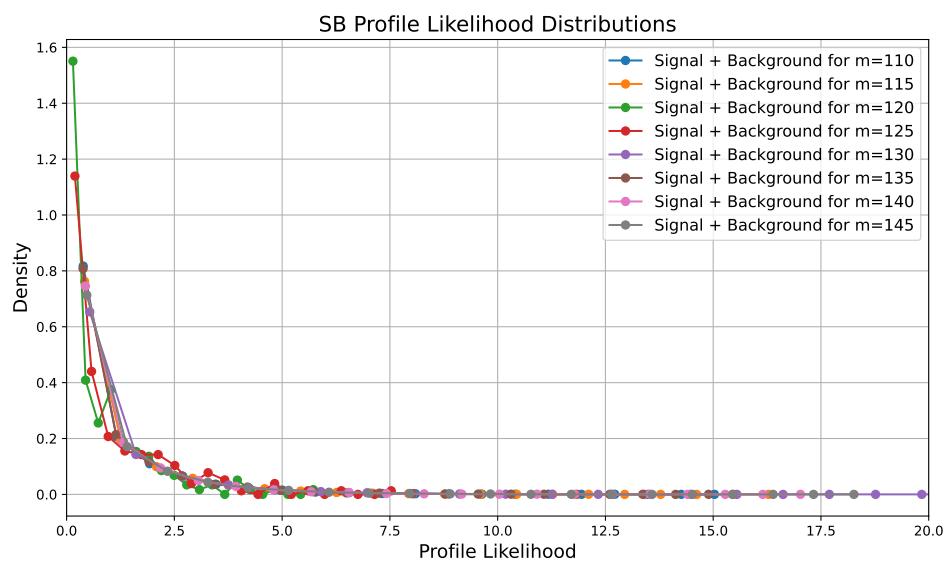
Problem 2 (Collaborators: None)

NOTE: Unless otherwise specified, the data used in this part is the *small* dataset. Additionally, because the peak is smaller, I found it appropriate to adjust the efficiencies to be given by something less than $N_s = 45 \times \mu$. After doing an initial couple fits, I decided to

instead test for $N_s = 22 \times \mu$. However, you can find in the git repo the figures associate with $N_s = 45 \times \mu$ under the "figures" directory.

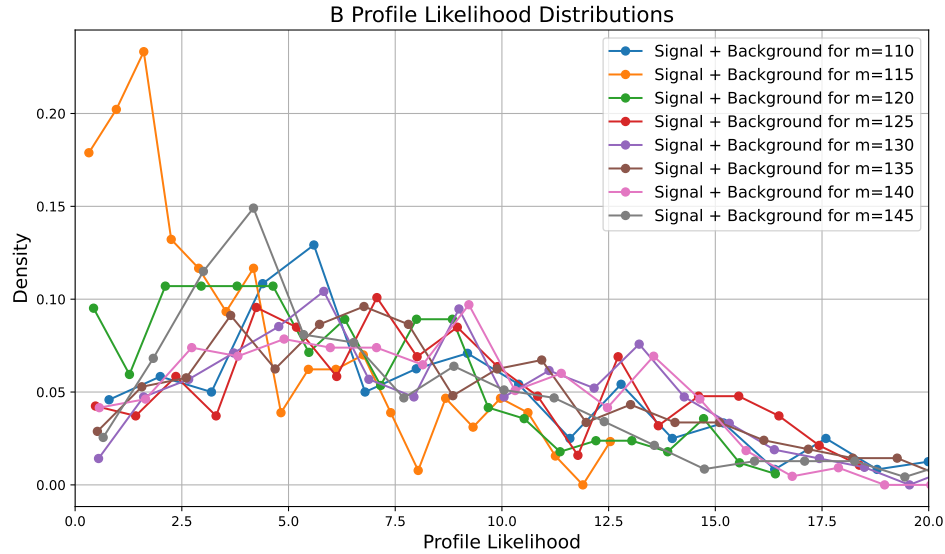
Part 2(a)

Below is the plot of the profile likelihood test statistic under the signal and background hypothesis under several different mass bins. 200 toys were used.



Part 2(b)

Below is the plot of the profile likelihood test statistic under the background only hypothesis under several different mass bins.

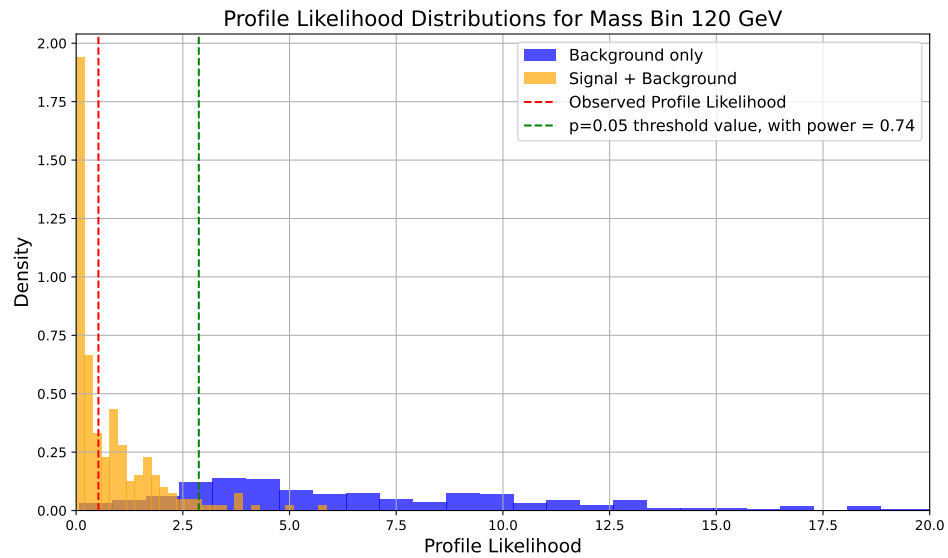


As you can see, the test statistic generated under the background only hypothesis is not dependent on the mass, m_{SM} , because the fit parameters are independent of the mass parameter. It is, however, quite noisy, even though 200 toys were used.

Part 2(c)

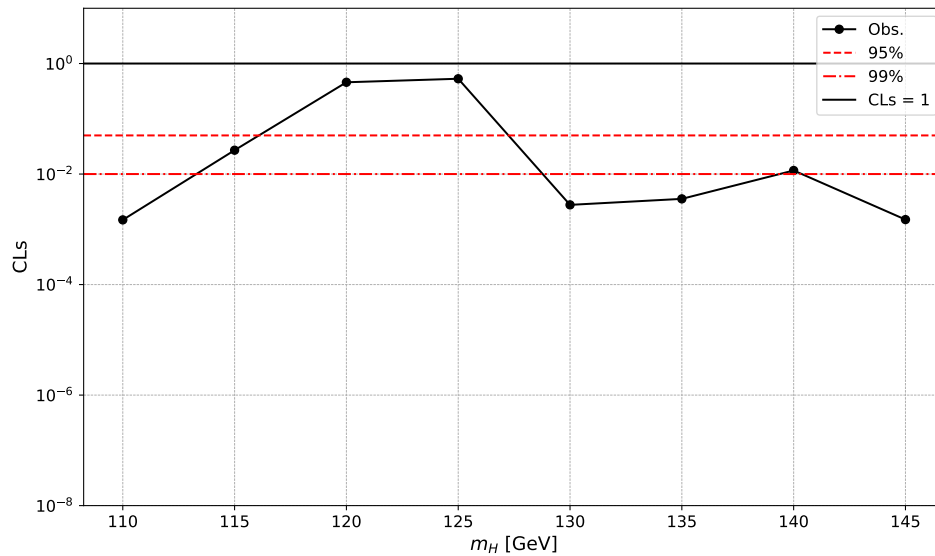
Below you can see the profile likelihood distribution for the background only and signal+background hypothesis, for the $m_{SM} = 120$ GeV bin.

The green line represents a confidence level of 0.05 with respect to the signal+background hypothesis. We then use the background only hypothesis to calculate the power as 0.74.



Part 2(d)

Below is the CLs plot generated using 200 toys

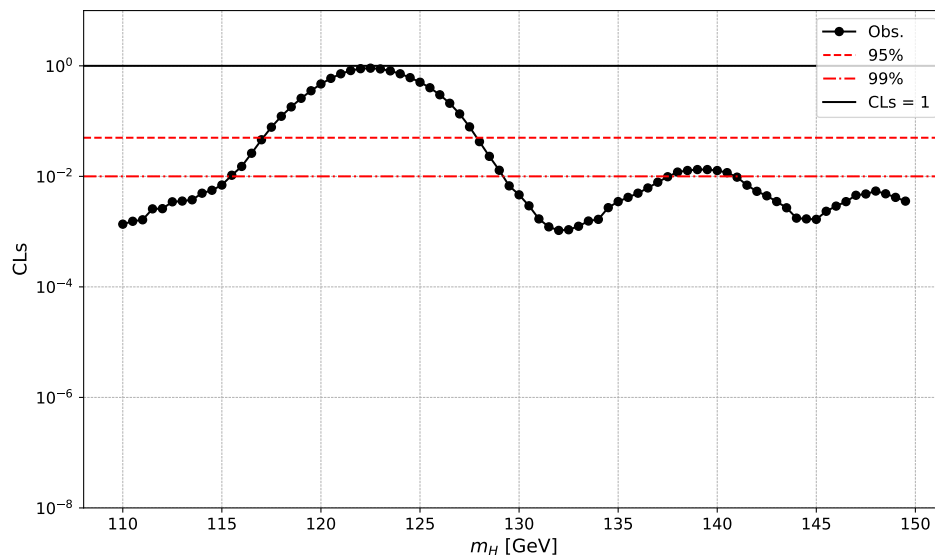


Part 2(e)

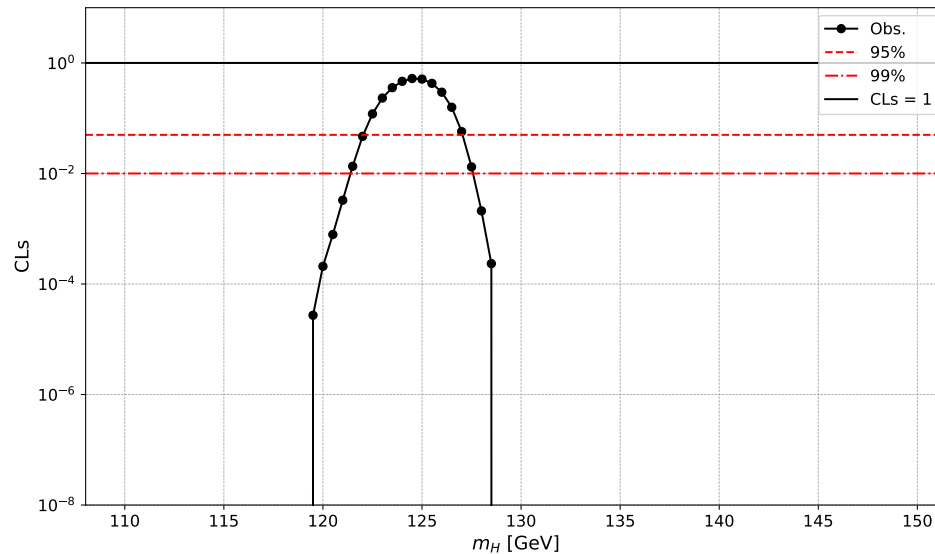
The regions we can exclude the signal+background hypothesis safely with 95% CL are: 110-115 GeV and 130-145 GeV. We cannot exclude the signal+background hypothesis safely with 95% CL in the range of 115-130 GeV. This is generally the plot we'd expect since the higgs mass is around 125 GeV.

One way to improve this plot would be with more mass bins. One way to do this efficiently, is to notice that the signal+background hypothesis follows Wilk's Theorem, and so simply can be approximated with a χ^2 distribution, so we don't need to make toys for them. Furthermore, the background only hypothesis shouldn't (and isn't by inspection) dependent on the mass bin itself, so we just need to make one set of toys. Using this, we only need to fit to the observed data per mass bin, which means it becomes much cheaper to interpolate more mass bins.

Doing this (and using 400 toys for the background instead) gives us:



Furthermore, it becomes much easier to do the larger peak dataset with $N_s = 45 \times \mu$, which gives us:



Part 2(f)

The CLs method is a conservative limit on an exclusion confidence interval that modifies the confidence interval to not falsely exclude a region in the case where the experiment is simply just not sensitive to the signal.

It is not a confidence interval because it is not purely frequentist, but instead weighted by the power of the test statistic. Instead, it is just an upper limit.

Part 2(g)

When the number of observed signal events is large enough, the likelihood function will become a Gaussian and the CLs coverage will correctly cover a confidence interval.