(Day 3)

Exercise 13 – A Bayesian interval

- In this exercise we will construct a Bayesian interval for an example measurement.

  – Copy input file ex13.C, have a look at the first block of code and run it.

  – You will see a plot of a Voigtian distribution, which is the convolution of a Breit-Wigner shape (=$1/(x^2+\Gamma^2)$) which is typical for mass resonances and a Gaussian, which is typical the detector resolution. The convolution of these two pdfs describes what a mass resonance will look like after detection, and is analytically calculated in class RooVoigtian. The plot shows three cases: the pure Breit-Wigner distribution in red, the Voigtian distribution with $\Gamma=3$ and $\sigma=1$ (blue) and with $\Gamma=3$ and $\sigma=3$ (dashed).

  – Now uncomment BLOCK 1 and rerun. We generate a toy dataset at $\Gamma=3$ and $\sigma=1$ and fit that with our model. Note the strong correlation between Γ and σ in the fit (which is typical if σ and Γ are of similar magnitude)

- A simple Bayesian interval

  – In this exercise Γ(width) is the parameter of interest and σ is a nuisance parameter.

  – We first run a scenario in which we assume that σ is known exactly (i.e. there are no nuisance parameters). Now uncomment BLOCK 2 and rerun. Here the –log(L) function is constructed and plotted versus Γ

Exercise 13 – A Bayesian interval

- – Now uncomment Block 3 and rerun. This code constructs the Likelihood function from the –log(L) function (through exponentiation) and constructs a Bayesian posterior function from the likelihood by multiplying it with a flat prior function in Γ. The posterior is then plotted. A 68% central Bayesian interval can now be defined by an area that integrates 68% of the likelihood and leaves 16% on each side.

- – Now uncomment Block 4. The additional code will create the cumulative distribution function (cdf) of the Bayesian posterior, i.e. $\int_\Gamma^\infty P(\Gamma')d\Gamma'$, which simplifies the calculation of the interval: The interval is now delimited by the values of Γ where the CDF crosses values 0.16 and 0.84. Determine what the Bayesian interval is (use can select cross-hairs on the canvas to simplify this task: click the right mouse button in the area above the plot frame and select the item SetCrosshairs). Write down the interval for future comparison.

- • A Bayesian interval with a nuisance parameter

  - – We now consider the case where we don't know σ a priori and need to constrain it from the data. Thus σ is now a nuisance parameter.

Exercise 13 – A Bayesian interval

- Uncomment Block 5 and rerun. The code that is added will show the –log(L) and the L distribution versus (Γ,σ). (You can again see from the L distribution that there is a significant correlation between Γ and σ.

- Uncomment Block 6 and rerun. We now follow the Bayesian procedure for the elimination of nuisance parameters: we integrate the posterior distribution (takes as the likelihood times a flat prior in both σ and Γ) over the nuisance parameter σ and plot the resulting posterior in Γ as well as the corresponding cumulative distribution (the required integrations may take a few minutes). Compare the distribution of this posterior to that of the previous scenario (with fixed σ). Calculate the 68% central interval from the c.d.f. and write down the values for future reference

- A Bayesian interval with a nuisance parameter with non-uniform prior

  - For illustration we can also choose a non-uniform prior for the σ parameter, e.g. a Gaussian with mean 1 and width 0.1. This represent a knowledge on the value of σ that is about three times as precise as what can be inferred from the data. Uncomment Block 7, rerun (will again take a few minutes) and calculate the 68% central interval for this case.

Exercise 14 – Profile Likelihood interval

- This exercise aims to calculate a interval on Γ for the same problem as Ex13, but now using a Profile Likelihood method.

  - Copy file ex14.C, look at the code and run it. The code in this exercise sets up the same model and data as Ex13, constructs the likelihood and plots the likelihood as function of Γ, assuming a known fixed value of σ=1.

  - Calculate the 68% likelihood interval by finding the values of Γ that correspond to LR=+0.5. Write down the value and compare it to the corresponding 68% Bayesian interval

  - Uncomment Block1 and rerun. Now we move to the scenario where must constrain σ from the data and will consider it a nuisance parameter.

  - The procedure to eliminate nuisance parameters in the profile likelihood method is to make a scan of the likelihood in Γ where we plot for each point in Γ the best (=lowest) value of the LR for any value σ (instead of the value of the LR at σ=1). [This will invariable *widen* the distribution. Try to understand why that is the case]

Exercise 14 – Profile Likelihood interval

- The newly added code creates the profile likelihood function (which is represented by a function object in RooFit as well (which will internally call MINUIT to perform the minimization of the likelihood w.r.t. the nuisance parameters everytime it is called)

- Calculate the 68% profile likelihood interval by finding the values of Γ that correspond to PLR=+0.5. Write down the value and compare it to the corresponding 68% Bayesian interval

- Finally uncomment Block 3 and run again. The newly added code will provide a visualization of how the model shape changes in the profile likelihood as function of the parameter of interest Γ. A canvas with 9 pads is created, which correspond to the situation at Γ=0.5,1,...,4.5. Each pad shows the data (which is always the same), and in red the model with σ=1, and in blue the model with the value of σ that gives the best fit (the likelihood of this best fit is used in the profile likelihhod)

Exercise 15 – Testing Wilks theorem

- This exercise aims to test the validity of Wilks theorem for the example analysis used in Ex13 and Ex14.

  - Wilks theorem states that the distribution of the 2 times the Likelihood ratio will be that of a chi-squared distribution in the asymptotic case (i.e. $N \to \infty$).

  - In this exercise we will generate a series of toy Monte Carlo samples, calculate the likelihood ratio for each of them, plot the distribution and compare that distribution to that of the asymptotic chi-squared distribution.

  - Copy file ex15.C, look at the code and run it (takes 1-2 minutes). Does the data look like the asymptotic distribution? You will have to put the y-axis in a log scale – to do so right-click just above the plot frame and select option SetLogY.

  - Up to what value of the LLR do you have enough statistics to claim agreement. What is the corresponding level of significance? (remember $LLR = 0.5 \cdot Z^2$)

  - The provided code generate the LLR distribution for L(width=3)/L(width=best). Change the code so that it generates the distribution for L(width=1)/L(width=best) and rerun. Does the distribution look different?

Exercise 15 – Testing Wilks theorem

- So far we have checked the behavior of the likelihood ratio with σ fixed to 1. Next we check if the *profile* likelihood also follows the prediction of Wilks theorem. To do so we need to make σ a floating parameter of the model. Change 'sigma[1]' into 'sigma[1,0.1,3]' in the factory string and rerun.

- How many toy data samples do you need to generate and fit to validate Wilks theorem up to Z=5? (i.e. calculate first the corresponding probability, then take 100/prob as a rough estimate of the number of toys you need).

- How long would it take you to do that check? (based on the time it took you to run 1000 toys). What if you had 1000 CPUs available?