

# Extended Maximum Likelihood Exercise

November 2021

Consider the binned probability distribution functions for signal and background for some toy multivariate analysis and the toy data histogram shown in Fig. 1. Perform a binned extended maximum likelihood fit to the data set to extract the number of signal events, including the effects of the uncertainties in the background shape.

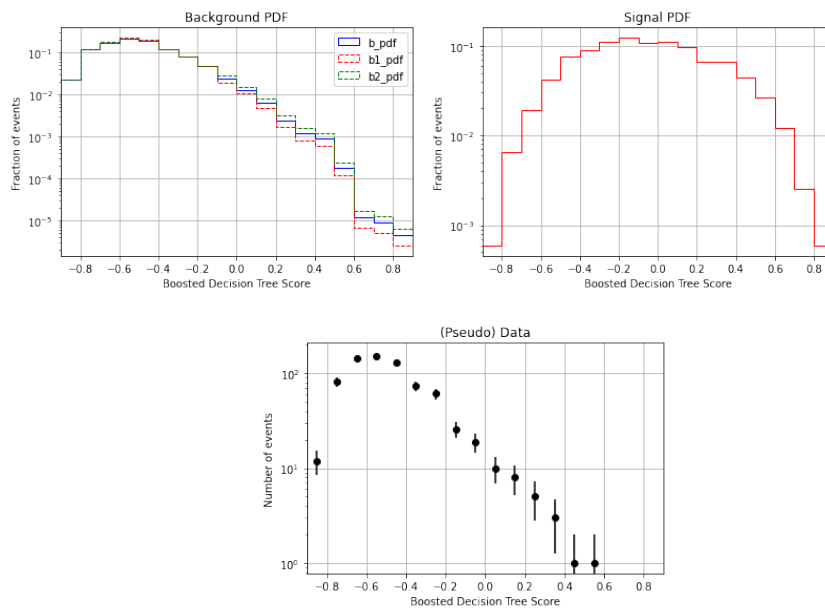


Figure 1: Top left: the background pdf, including one sigma up and down systematic variations. Top right: the signal pdf. Bottom: the data sample.

You can retrieve the information in the three plots from the *pickle* file <https://tinyurl.com/2njp74dd/histForMinuitFit.npz> using this snippet of code in python

```
import numpy as np
# Read the arrays that were prepared separately
```

```

# b_pdf      = the default background hist pdf
# b1_pdf     = 1st alternative to hist b_pdf
# b2_pdf     = 2nd alternative to hist b_pdf
# s_pdf      = the hist pdf for signal
# d          = the hist for the data
# binCen     = the center of the hist bins
# binEdges   = the edges of the bins
# The arrays were saved with this command:
# np.savez("histForMinuitFit.npz", b_pdf, b1_pdf, b2_pdf,
#      s_pdf, d, binCen, binEdges,
#      b_pdf=b_pdf, b1_pdf=b1_pdf, b2_pdf=b2_pdf,
#      s_pdf=s_pdf, d=d,
#      binCen=binCen, binEdges=binEdges)
npzfile = np.load("histForMinuitFit.npz")
b_pdf   = npzfile['b_pdf']
b1_pdf  = npzfile['b1_pdf']
b2_pdf  = npzfile['b2_pdf']
s_pdf   = npzfile['s_pdf']
d       = npzfile['d']
binCen  = npzfile['binCen']
binEdges = npzfile['binEdges']

```

In case you are wondering the toy data set was obtained by drawing 700 and 25 events from the background and signal pdfs, respectively.