# Exercises - Lectures 4.1 and 4.2

Harrison B. Prosper

6 June 2013

INFN SOS 2013

## I. PROBLEMS

1. Consider the maximum likelihood estimator (MLE) $\hat{\theta}$ of a parameter $\theta$. Show that, in general, MLEs are biased. Hint: consider $\hat{\alpha} = f(\hat{\theta})$, where $f$ is some non-linear function. For example, $\theta$ could be the Higgs boson mass while $\alpha$ is the predicted Higgs boson production cross section. Taylor expand $\hat{\alpha} = f(\hat{\theta} + h)$ about $\hat{\theta}$, average both sides of the equation and show that even if $\hat{\theta}$ is unbiased, that is, that $\overline{\hat{\theta}} = \theta$, in general, the estimator $\hat{\alpha}$ is not.

2. The discovery of the top quark by DØ and CDF was based on relatively small event samples. DØ found $N = 17$ events with a background estimate of $B = 3.8 \pm \delta B = 0.6$ events. Assuming the following likelihood function

$$p(N|s,b) = \frac{(s+b)^N e^{-(s+b)}}{N!} \frac{(bk)^Q e^{-bk}}{\Gamma(Q+1)},$$

where the effective count $Q$ and scale factor $k$ are defined by $B = Q/k$ and $\delta B = \sqrt{Q}/k$, show that the maximum likelihood estimate of $b$, call it $\hat{b}(s)$, for a given $s$ is given by

$$\hat{b}(s) = \frac{g + \sqrt{g^2 + 4(1+k)Qs}}{2(1+k)},$$

where $g \equiv D + Q - (1+k)s$. Then show that the solution of

$$\chi^2 = -2 \ln \frac{p_{PL}(17|s, \hat{b}(s))}{p(N|\hat{s}, \hat{b})} = 1,$$

is

$$s \in [9.4, 17.7] \quad @\ 68.3\%\ \text{C.L.},$$

where in the denominator $\hat{s} = N - B$ and $\hat{b} = B$.

## II. PROJECTS

Unpack `tutorials.tar.gz` using

```
tar zxvf tutorials.tar.gz
```

then

```
cd tutorials-cowan
python expFit.py
```

If this works, then do

```
cd ../tutorials-MVA
source setup.sh      (if you use a bash shell or source setup.csh otherwise)
```

Next

```
cd classification/higgs/TMVA
ln -s ../ntuples/*.root .                  (make a link to Root ntuple files)
```

1. **higgs**

   (a) Run the `train.py` program using the BDT method and then the `plot.py` program and determine what cut on the BDT yields the smallest classification error rate.

   (b) Modify `train.py` and `plot.py` and determine if there are 2 variables that are about as good, in terms of error rate, as `f_Z1mass` and `f_Z2mass`. Use either the BDT or MLP methods.

2. **iris**, **titanic**

   (a) Using the Higgs codes as examples, construct a BDT to discriminate between any two kinds or Irises or to discriminate between survivors and non-survivors on the Titanic. Be sure to use only a fraction of examples as training data (say 25 examples each for the Iris data and 300 each for the Titanic data).

   (b) Determine the correct classification rate as a function of the cut on the BDT and find the best Bayes classifier for the Irises and the Titanic survivors.