Higgs ID Explorer

Sam Hewamanage

July 26, 2014

```
[1] "/Users/samantha/Documents/Personal/Samantha/GITSandbox/KaggleHiggsBosonID/rawcode"
```

```
'data.frame':   250000 obs. of  33 variables:
 $ EventId                   : int  100000 100001 100002 100003 100004 100005 100006 100007 100008 100
 $ DER_mass_MMC              : num  138 161 -999 144 176 ...
 $ DER_mass_transverse_met_lep: num  51.7 68.8 162.2 81.4 16.9 ...
 $ DER_mass_vis              : num  97.8 103.2 126 80.9 134.8 ...
 $ DER_pt_h                  : num  27.98 48.146 35.635 0.414 16.405 ...
 $ DER_deltaeta_jet_jet      : num  0.91 -999 -999 -999 -999 ...
 $ DER_mass_jet_jet          : num  125 -999 -999 -999 -999 ...
 $ DER_prodeta_jet_jet       : num  2.67 -999 -999 -999 -999 ...
 $ DER_deltar_tau_lep        : num  3.06 3.47 3.15 3.31 3.89 ...
 $ DER_pt_tot                : num  41.928 2.078 9.336 0.414 16.405 ...
 $ DER_sum_pt                : num  198 125 198 76 58 ...
 $ DER_pt_ratio_lep_tau      : num  1.582 0.879 3.776 2.354 1.056 ...
 $ DER_met_phi_centrality    : num  1.4 1.41 1.41 -1.28 -1.39 ...
 $ DER_lep_eta_centrality    : num  0.2 -999 -999 -999 -999 0.975 0.791 -999 -999 -999 ...
 $ PRI_tau_pt                : num  32.6 42 32.2 22.6 28.2 ...
 $ PRI_tau_eta               : num  1.017 2.039 -0.705 -1.655 -2.197 ...
 $ PRI_tau_phi               : num  0.381 -3.011 -2.093 0.01 -2.231 ...
 $ PRI_lep_pt                : num  51.6 36.9 121.4 53.3 29.8 ...
 $ PRI_lep_eta               : num  2.273 0.501 -0.953 -0.522 0.798 ...
 $ PRI_lep_phi               : num  -2.414 0.103 1.052 -3.1 1.569 ...
 $ PRI_met                   : num  16.82 44.7 54.28 31.08 2.72 ...
 $ PRI_met_phi               : num  -0.277 -1.916 -2.186 0.06 -0.871 ...
 $ PRI_met_sumet             : num  258.7 164.5 260.4 86.1 53.1 ...
 $ PRI_jet_num               : int  2 1 1 0 0 3 2 1 0 1 ...
 $ PRI_jet_leading_pt        : num  67.4 46.2 44.3 -999 -999 ...
 $ PRI_jet_leading_eta       : num  2.15 0.725 2.053 -999 -999 ...
 $ PRI_jet_leading_phi       : num  0.444 1.158 -2.028 -999 -999 ...
 $ PRI_jet_subleading_pt     : num  46.1 -999 -999 -999 -999 ...
 $ PRI_jet_subleading_eta    : num  1.24 -999 -999 -999 -999 0.224 0.131 -999 -999 -999 ...
 $ PRI_jet_subleading_phi    : num  -2.48 -999 -999 -999 -999 ...
 $ PRI_jet_all_pt            : num  113.5 46.2 44.3 0 0 ...
 $ Weight                    : num  0.00265 2.23358 2.34739 5.44638 6.24533 ...
 $ Label                     : Factor w/ 2 levels "b","s": 2 1 1 1 1 1 2 2 1 2 ...
```

```
   EventId        DER_mass_MMC    DER_mass_transverse_met_lep
 Min.   :100000   Min.   :-999.0   Min.   :  0.0
 1st Qu.:162500   1st Qu.:  78.1   1st Qu.: 19.2
 Median :225000   Median : 105.0   Median : 46.5
 Mean   :225000   Mean   : -49.0   Mean   : 49.2
 3rd Qu.:287499   3rd Qu.: 130.6   3rd Qu.: 73.6
 Max.   :349999   Max.   :1192.0   Max.   :690.1
  DER_mass_vis       DER_pt_h      DER_deltaeta_jet_jet DER_mass_jet_jet
 Min.   :   6.3   Min.   :   0.0   Min.   :-999.0       Min.   :-999
 1st Qu.:  59.4   1st Qu.:  14.1   1st Qu.:-999.0       1st Qu.:-999
 Median :  73.8   Median :  38.5   Median :-999.0       Median :-999
```

```
Mean   :  81.2   Mean   :  57.9   Mean   :-708.4       Mean   :-601
3rd Qu.:  92.3   3rd Qu.:  79.2   3rd Qu.:   0.5       3rd Qu.:  83
Max.   :1349.4   Max.   :2835.0   Max.   :   8.5       Max.   :4975
DER_prodeta_jet_jet DER_deltar_tau_lep   DER_pt_tot       DER_sum_pt
Min.   :-999.0      Min.   :0.208      Min.   :   0.0   Min.   :  46.1
1st Qu.:-999.0      1st Qu.:1.810      1st Qu.:   2.8   1st Qu.:  77.5
Median :-999.0      Median :2.491      Median :  12.3   Median : 120.7
Mean   :-709.4      Mean   :2.373      Mean   :  18.9   Mean   : 158.4
3rd Qu.:  -4.6      3rd Qu.:2.961      3rd Qu.:  27.6   3rd Qu.: 200.5
Max.   :  16.7      Max.   :5.684      Max.   :2835.0   Max.   :1852.5
DER_pt_ratio_lep_tau DER_met_phi_centrality DER_lep_eta_centrality
Min.   : 0.047       Min.   :-1.414         Min.   :-999
1st Qu.: 0.883       1st Qu.:-1.371         1st Qu.:-999
Median : 1.280       Median :-0.356         Median :-999
Mean   : 1.438       Mean   :-0.128         Mean   :-709
3rd Qu.: 1.777       3rd Qu.: 1.225         3rd Qu.:   0
Max.   :19.773       Max.   : 1.414         Max.   :   1
  PRI_tau_pt     PRI_tau_eta      PRI_tau_phi        PRI_lep_pt
Min.   : 20.0   Min.   :-2.499   Min.   :-3.1420   Min.   : 26.0
1st Qu.: 24.6   1st Qu.:-0.925   1st Qu.:-1.5750   1st Qu.: 32.4
Median : 31.8   Median :-0.023   Median :-0.0330   Median : 40.5
Mean   : 38.7   Mean   :-0.011   Mean   :-0.0082   Mean   : 46.7
3rd Qu.: 45.0   3rd Qu.: 0.898   3rd Qu.: 1.5650   3rd Qu.: 53.4
Max.   :764.4   Max.   : 2.497   Max.   : 3.1420   Max.   :560.3
 PRI_lep_eta       PRI_lep_phi         PRI_met        PRI_met_phi
Min.   :-2.5050   Min.   :-3.1420   Min.   :   0.1   Min.   :-3.1420
1st Qu.:-1.0140   1st Qu.:-1.5220   1st Qu.:  21.4   1st Qu.:-1.5750
Median :-0.0450   Median : 0.0860   Median :  34.8   Median :-0.0240
Mean   :-0.0195   Mean   : 0.0435   Mean   :  41.7   Mean   :-0.0101
3rd Qu.: 0.9590   3rd Qu.: 1.6180   3rd Qu.:  51.9   3rd Qu.: 1.5610
Max.   : 2.5030   Max.   : 3.1420   Max.   :2842.6   Max.   : 3.1420
PRI_met_sumet     PRI_jet_num    PRI_jet_leading_pt PRI_jet_leading_eta
Min.   :  13.7   Min.   :0.000   Min.   :-999.0     Min.   :-999.0
1st Qu.: 123.0   1st Qu.:0.000   1st Qu.:-999.0     1st Qu.:-999.0
Median : 179.7   Median :1.000   Median :  39.0     Median :  -1.9
Mean   : 209.8   Mean   :0.979   Mean   :-348.3     Mean   :-399.3
3rd Qu.: 263.4   3rd Qu.:2.000   3rd Qu.:  75.3     3rd Qu.:   0.4
Max.   :2004.0   Max.   :3.000   Max.   :1120.6     Max.   :   4.5
PRI_jet_leading_phi PRI_jet_subleading_pt PRI_jet_subleading_eta
Min.   :-999.0      Min.   :-999.0        Min.   :-999.0
1st Qu.:-999.0      1st Qu.:-999.0        1st Qu.:-999.0
Median :  -2.1      Median :-999.0        Median :-999.0
Mean   :-399.3      Mean   :-692.4        Mean   :-709.1
3rd Qu.:   0.5      3rd Qu.:  33.7        3rd Qu.:  -2.5
Max.   :   3.1      Max.   : 721.5        Max.   :   4.5
PRI_jet_subleading_phi PRI_jet_all_pt       Weight      Label
Min.   :-999.0         Min.   :   0.0   Min.   :0.002   b:164333
1st Qu.:-999.0         1st Qu.:   0.0   1st Qu.:0.019   s: 85667
Median :-999.0         Median :  40.5   Median :1.156
Mean   :-709.1         Mean   :  73.1   Mean   :1.647
3rd Qu.:  -2.3         3rd Qu.: 109.9   3rd Qu.:2.404
Max.   :   3.1         Max.   :1633.4   Max.   :7.823
```

Mean and Variance of Derived variables for Signal and Background

```
Loading required package: survival
Loading required package: splines
Loading required package: MASS
```

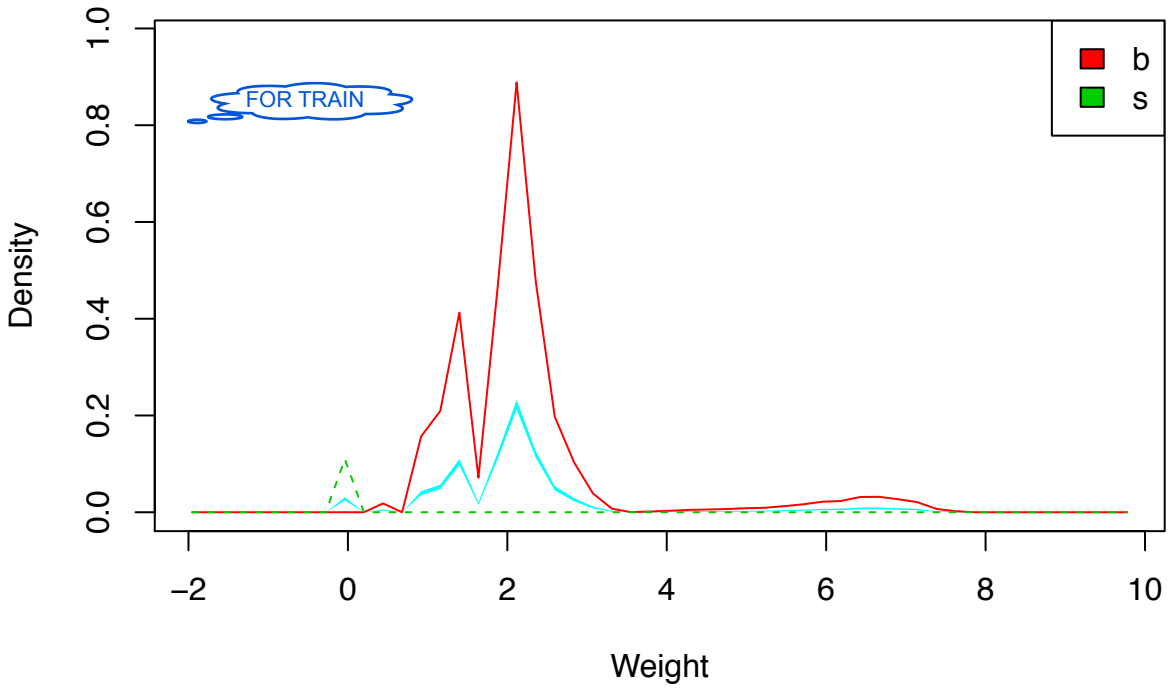
```
  Label Weight.mean DER_mass_MMC.mean DER_mass_transverse_met_lep.mean
1     b    2.501018          -119.19                            58.21
2     s    0.008078            85.57                            32.04
  DER_mass_vis.mean DER_pt_h.mean DER_deltaeta_jet_jet.mean
1             81.60         49.05                    -754.9
2             80.39         74.87                    -619.3
  DER_mass_jet_jet.mean DER_prodeta_jet_jet.mean DER_deltar_tau_lep.mean
1                -692.3                   -755.3                   2.366
2                -426.5                   -621.2                   2.386
  DER_pt_tot.mean DER_sum_pt.mean DER_pt_ratio_lep_tau.mean
1           19.16           145.6                     1.557
2           18.45           183.0                     1.209
  DER_met_phi_centrality.mean DER_lep_eta_centrality.mean Weight.var
1                     -0.3625                     -755.3  3.219e+00
2                      0.3209                     -620.2  6.775e-05
  DER_mass_MMC.var DER_mass_transverse_met_lep.var DER_mass_vis.var
1           215455                          1191.7          2295.5
2            40997                           909.5           460.4
  DER_pt_h.var DER_deltaeta_jet_jet.var DER_mass_jet_jet.var
1         3061                   184734                305469
2         5515                   236316                630974
  DER_prodeta_jet_jet.var DER_deltar_tau_lep.var DER_pt_tot.var
1                  184086                 0.6849          464.6
2                  233953                 0.4747          556.3
  DER_sum_pt.var DER_pt_ratio_lep_tau.var DER_met_phi_centrality.var
1          12507                   0.7154                      1.288
2          14160                   0.6306                      1.380
  DER_lep_eta_centrality.var
1                     184165
2                     235163
```

Density plots of 'Weight' Variables for Signal & Background events

```
library(sm)
# create value labels
sb.f <- factor(df$Label, levels= c(1,2), labels = c("b", "s"))
sm.density.compare(df$Weight, df$Label, xlab="Weight", model="equal")
```

```
##
## Test of equal densities:  p-value =  0
```

```
  colfill<-c(2:(2+length(levels(sb.f))))
  legend("topright", levels(sb.f), fill=colfill)
```

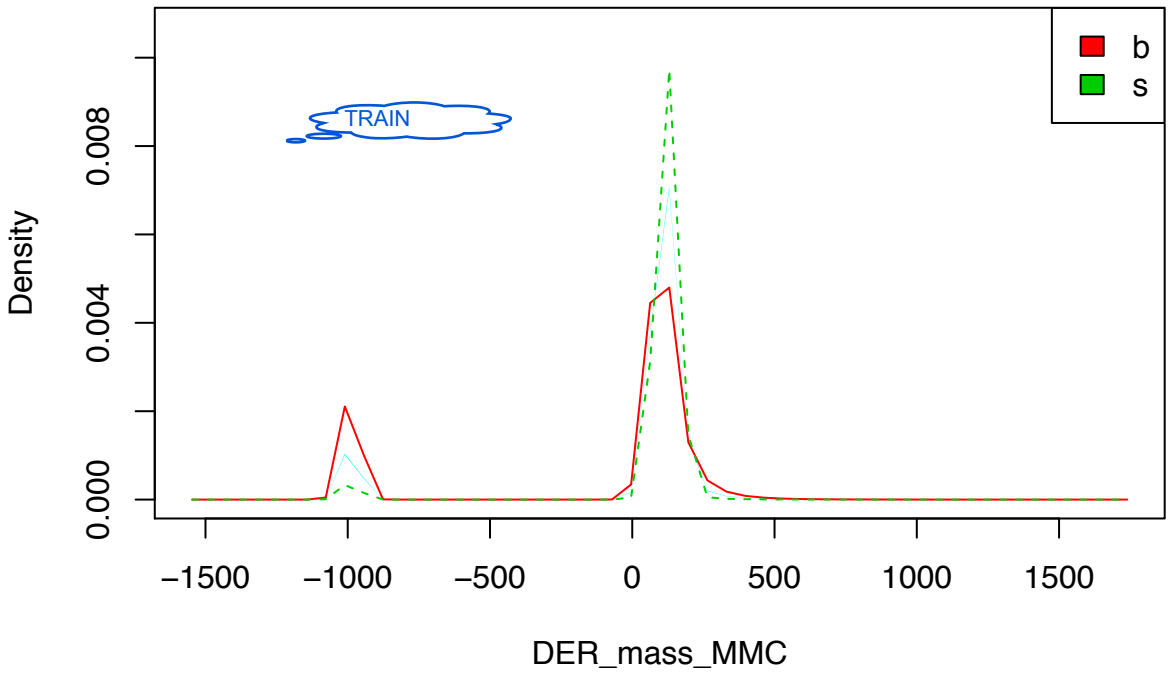


Density plots of Derived Variables for Signal & Background events

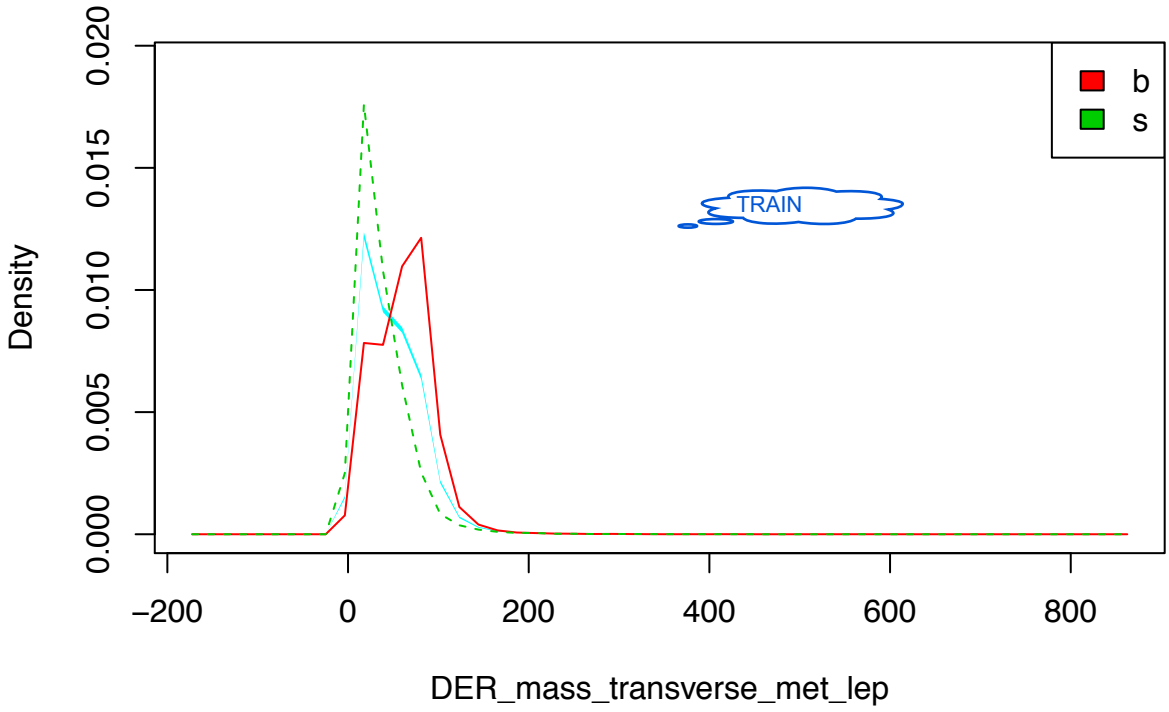
```
library(sm)

for(var in der.vars){
  sm.density.compare(df[,var], df$Label, xlab=var, model="equal")
  # add legend via mouse click
  #legend(locator(1), levels(sb.f), fill=colfill)
  legend("topright", levels(sb.f), fill=colfill)
}
```

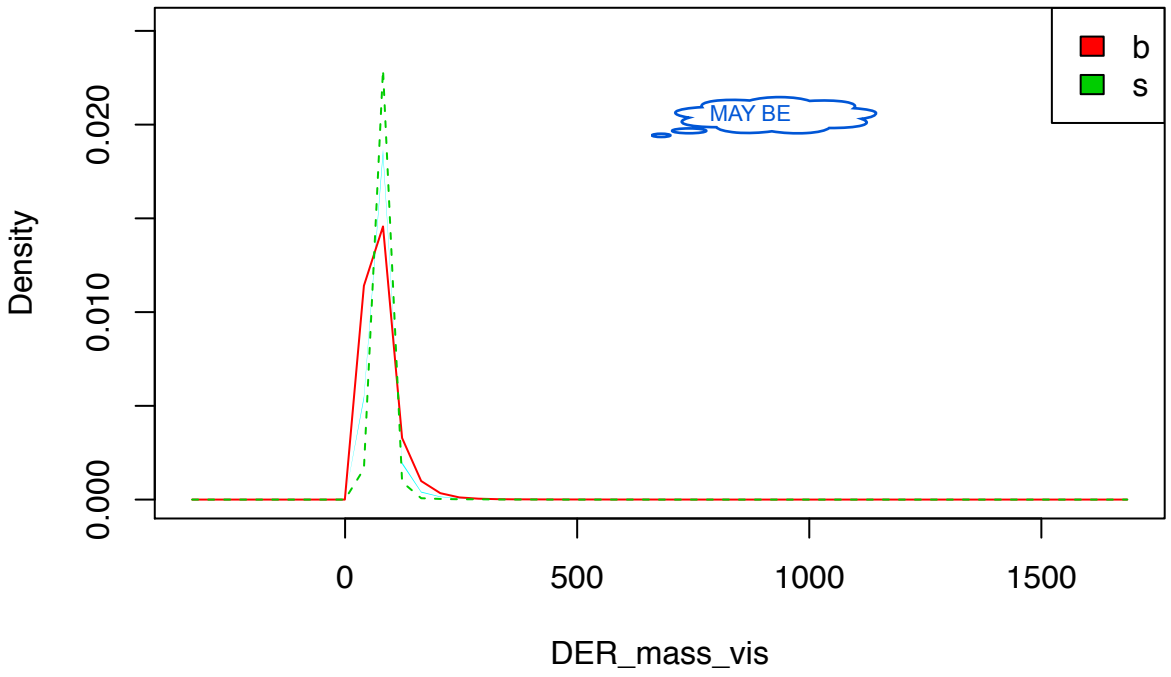
```
##
## Test of equal densities:  p-value =  0
```



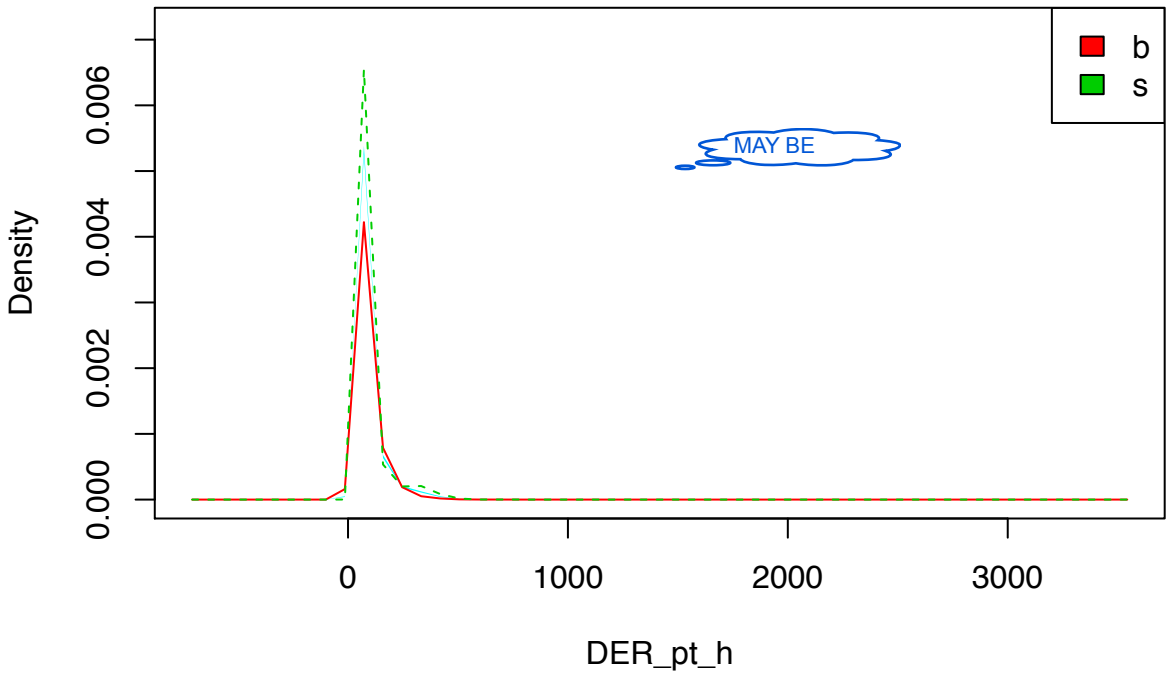

```
##
## Test of equal densities:  p-value =  0
```



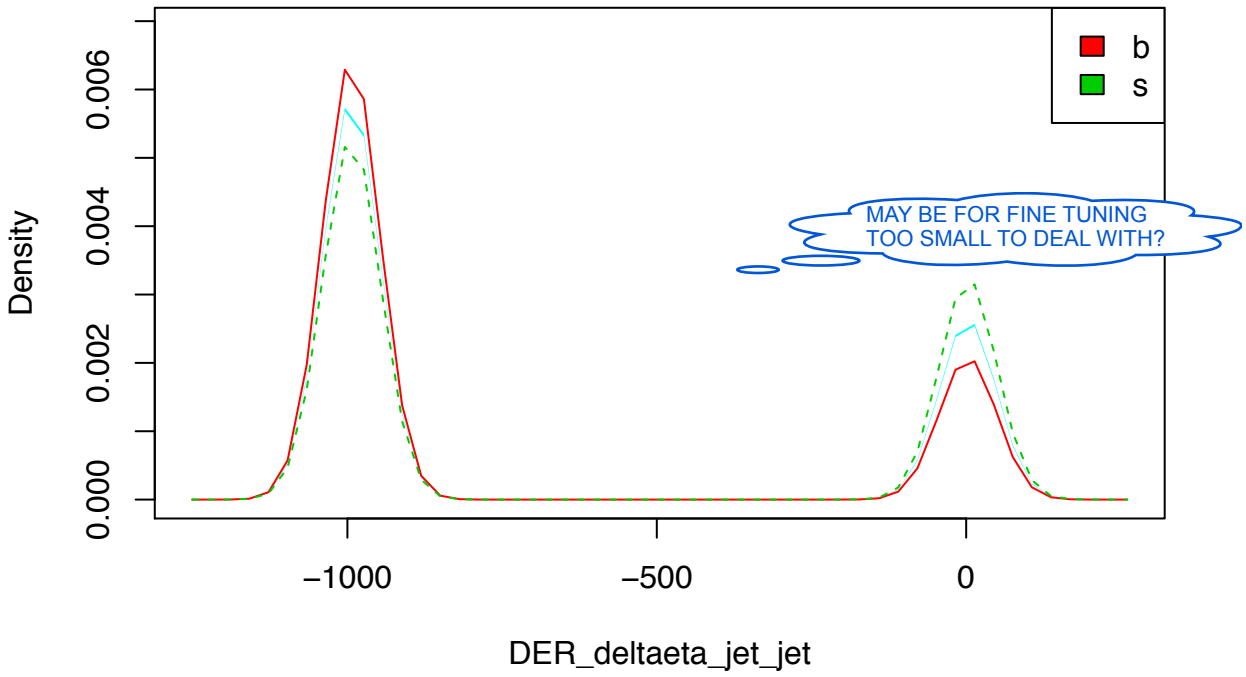

```
##
## Test of equal densities:  p-value =  0.43
```



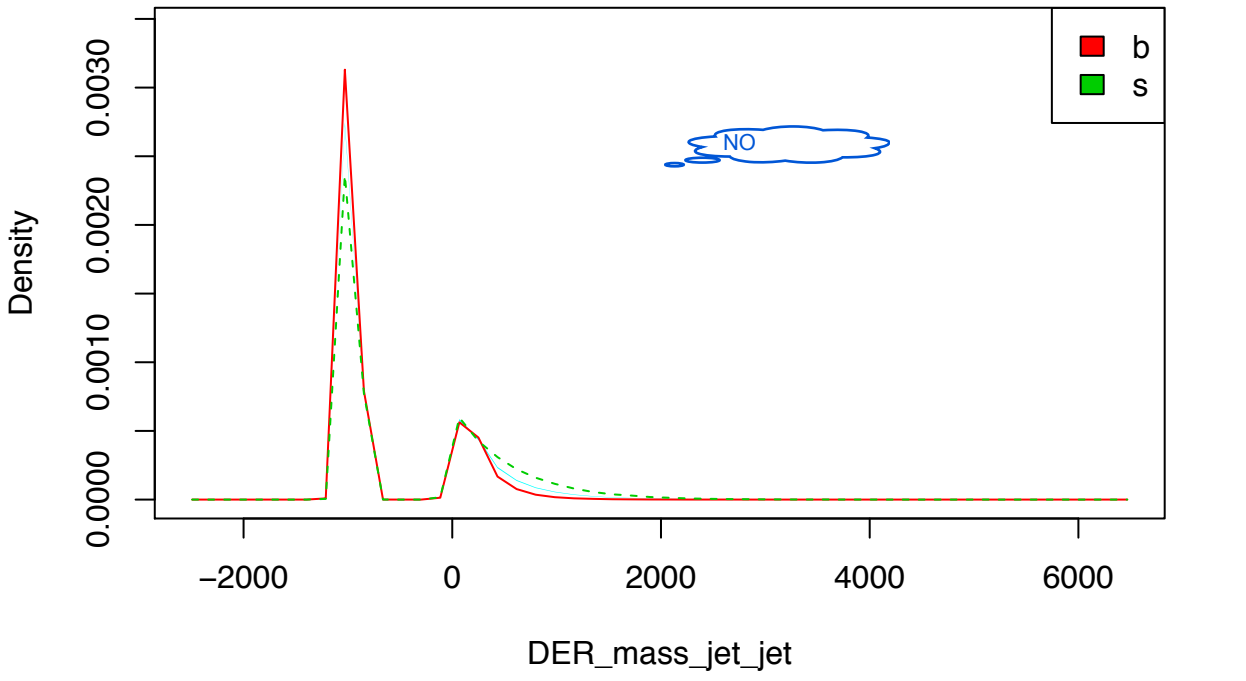

```
##
## Test of equal densities:  p-value =  0.61
```



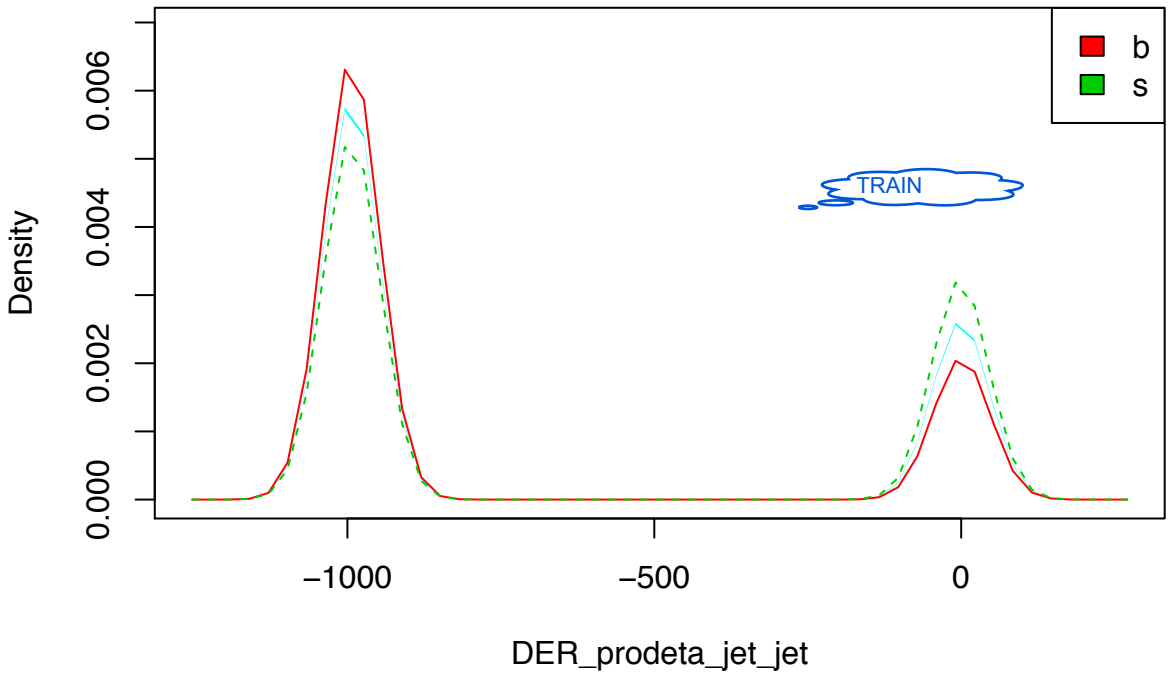

```
##
## Test of equal densities:  p-value =  0
```



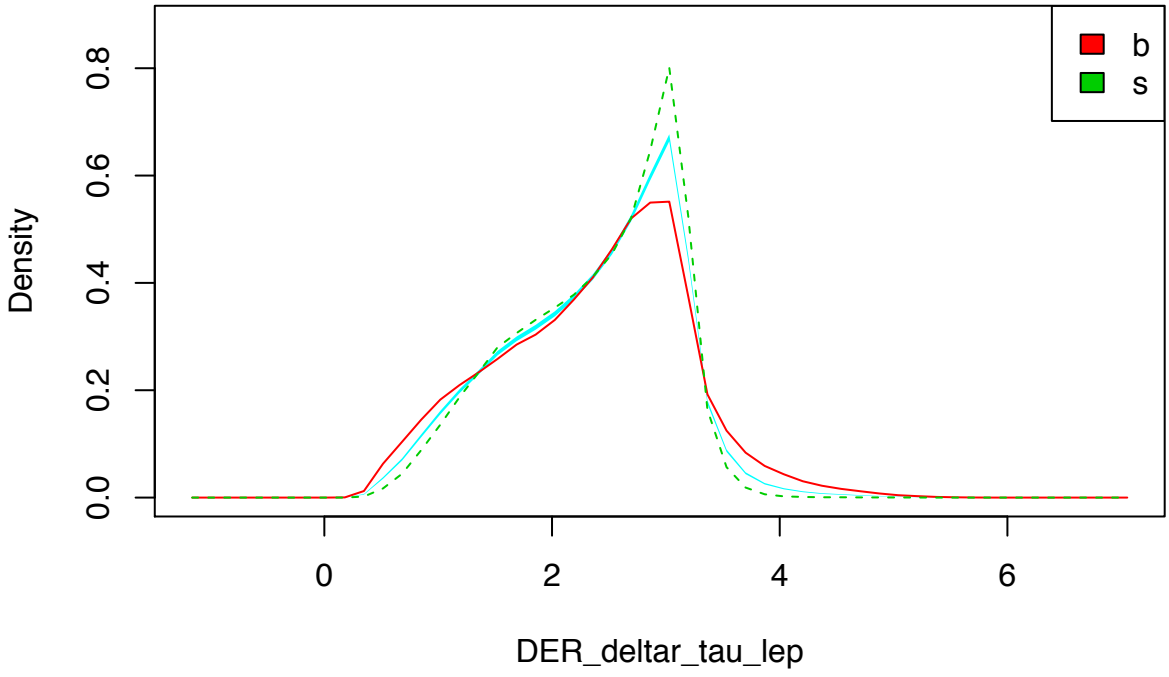

```
##
## Test of equal densities:  p-value =  0
```



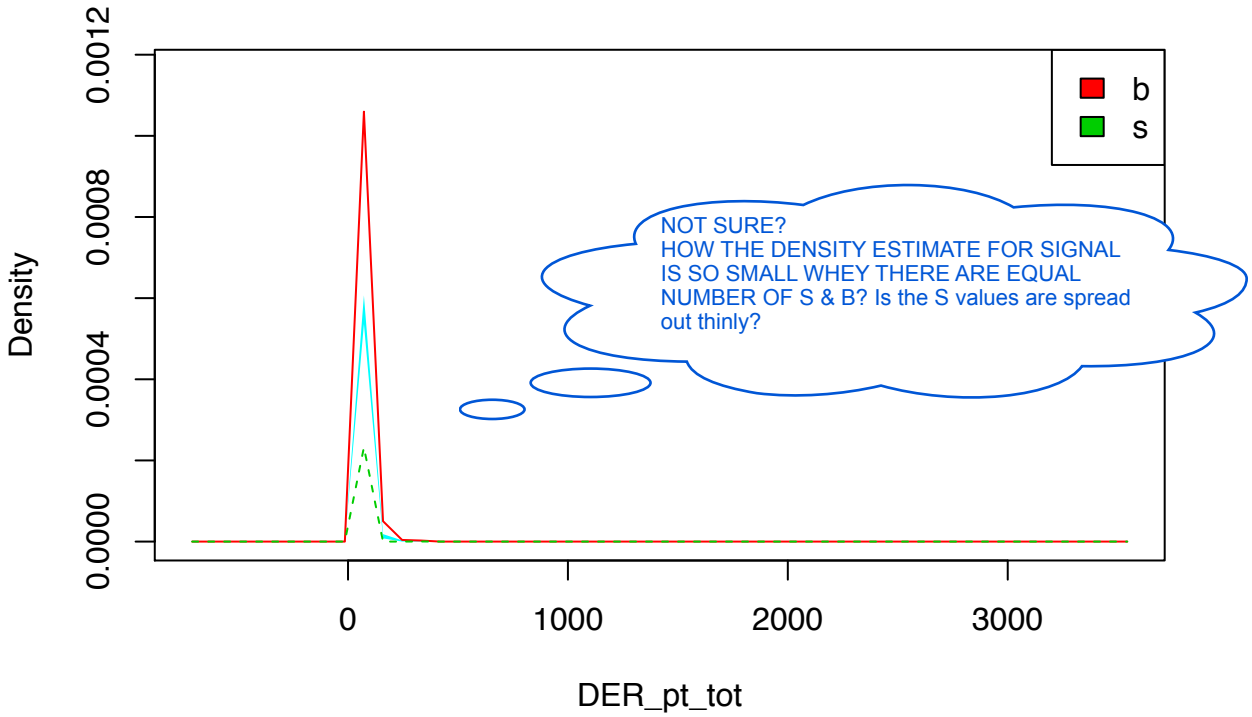

```
##
## Test of equal densities:  p-value =  0
```



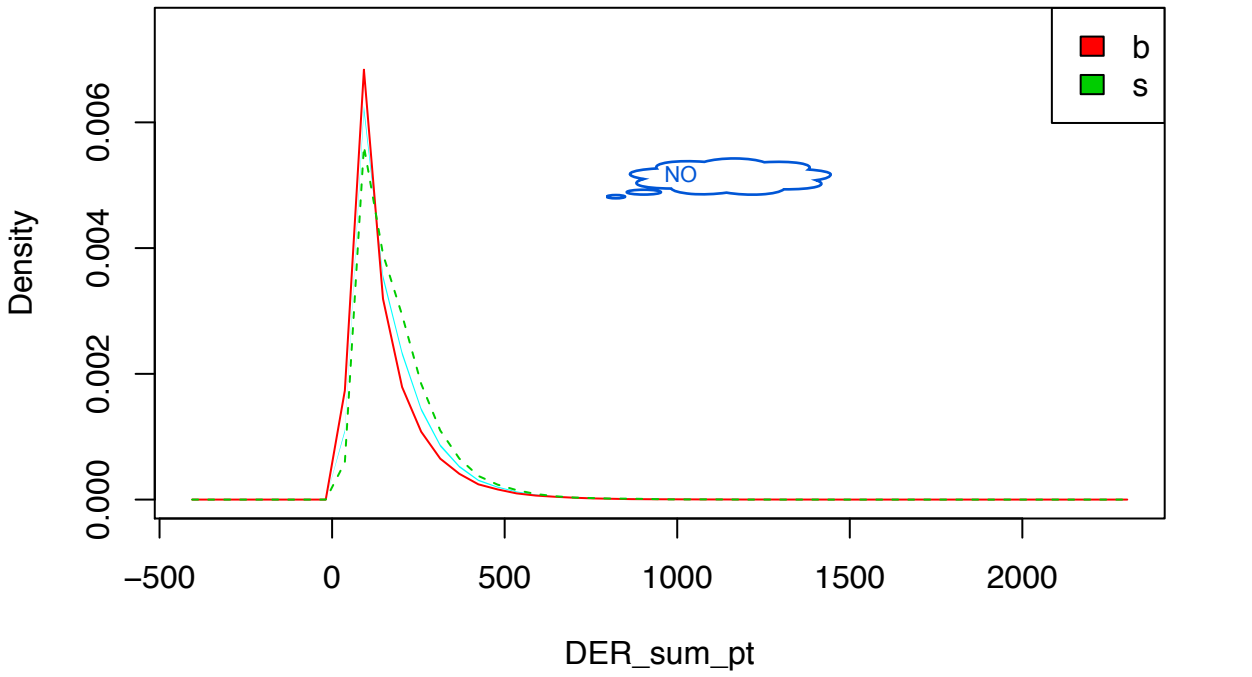

```
##
## Test of equal densities:  p-value =  0
```



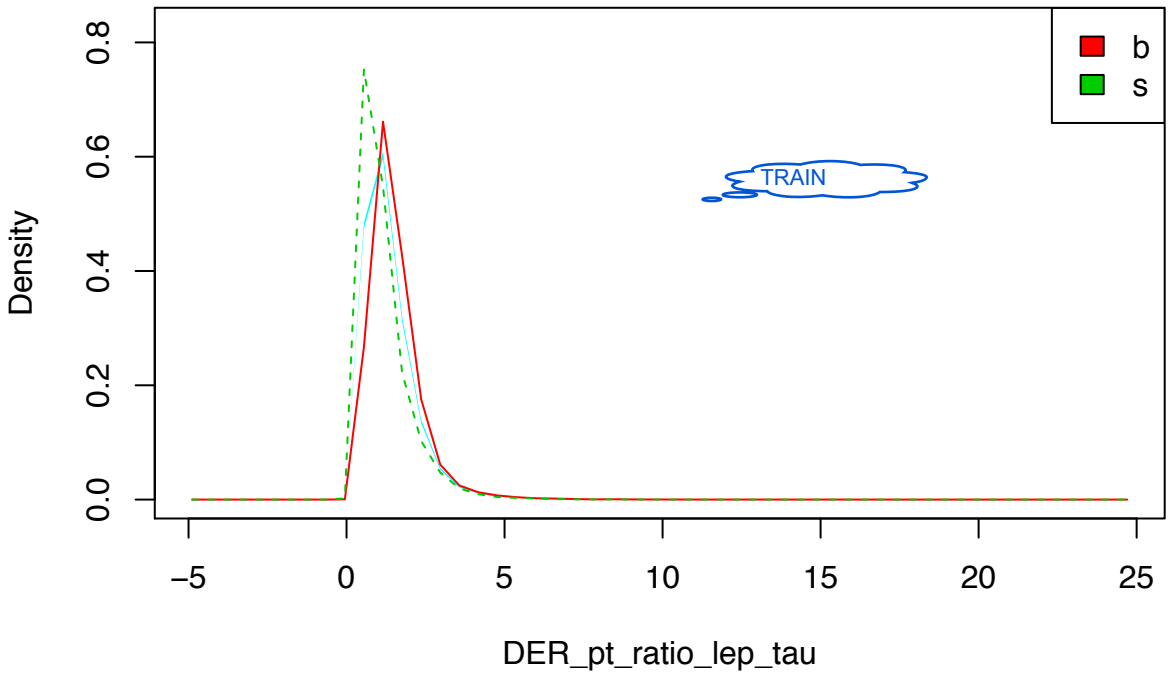

```
##
## Test of equal densities:  p-value =  0.61
```



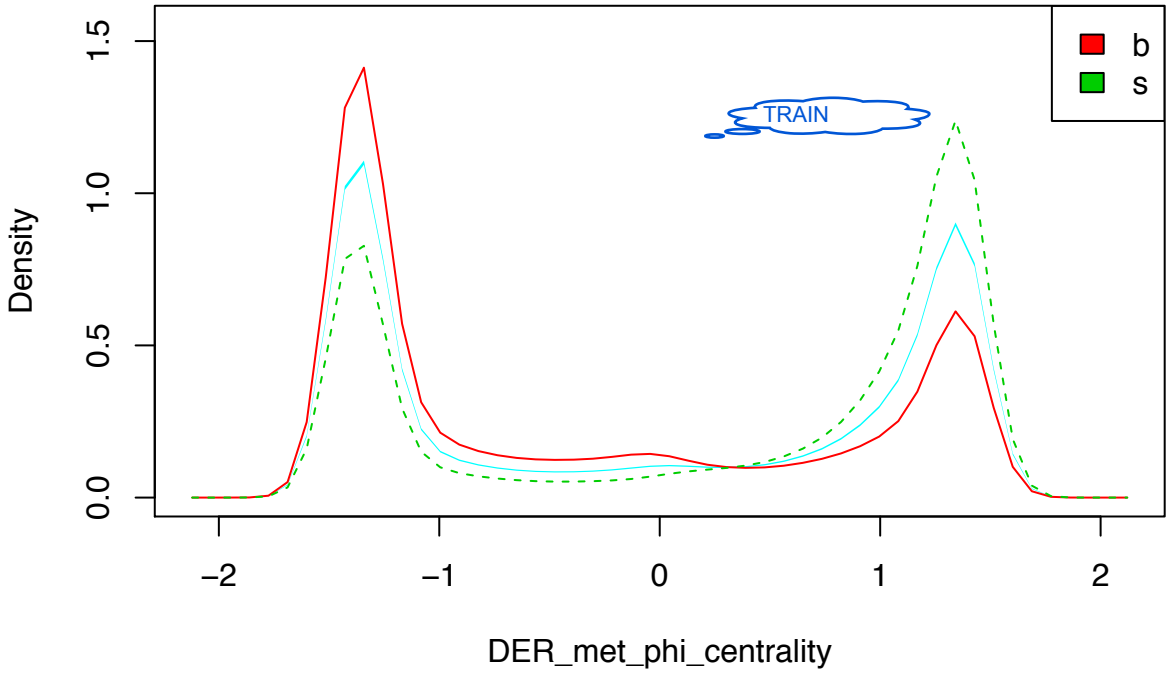

```
##
## Test of equal densities:  p-value =  0
```



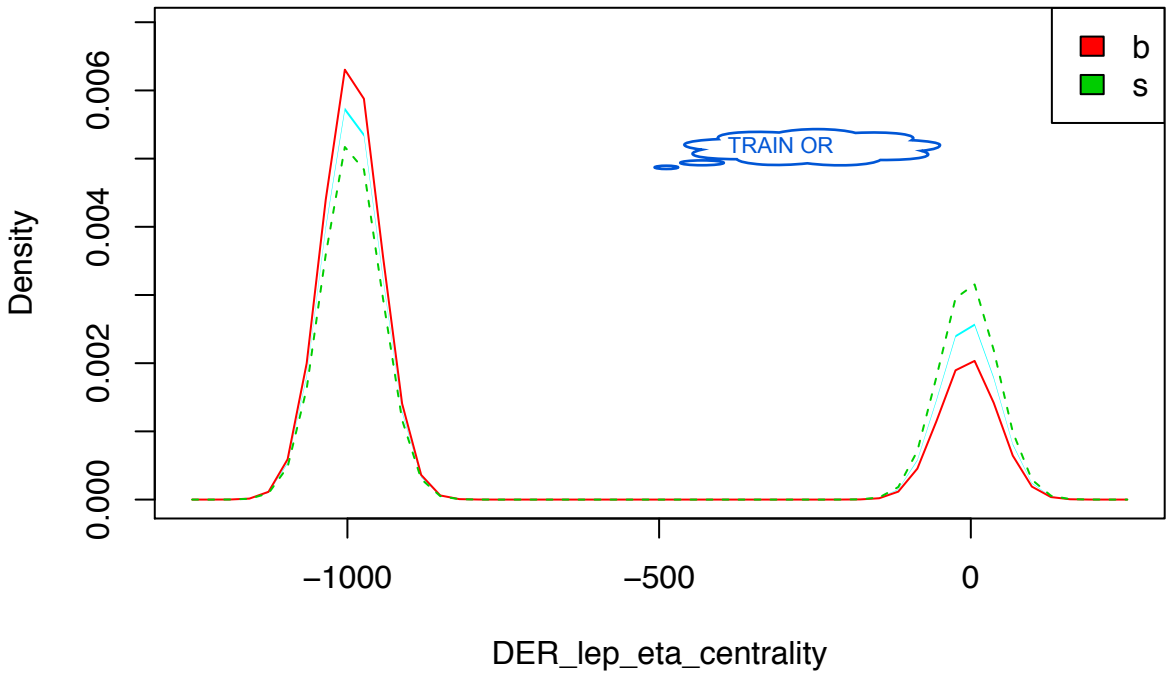

```
##
## Test of equal densities:  p-value =  0
```




```
##
## Test of equal densities:  p-value =  0
```



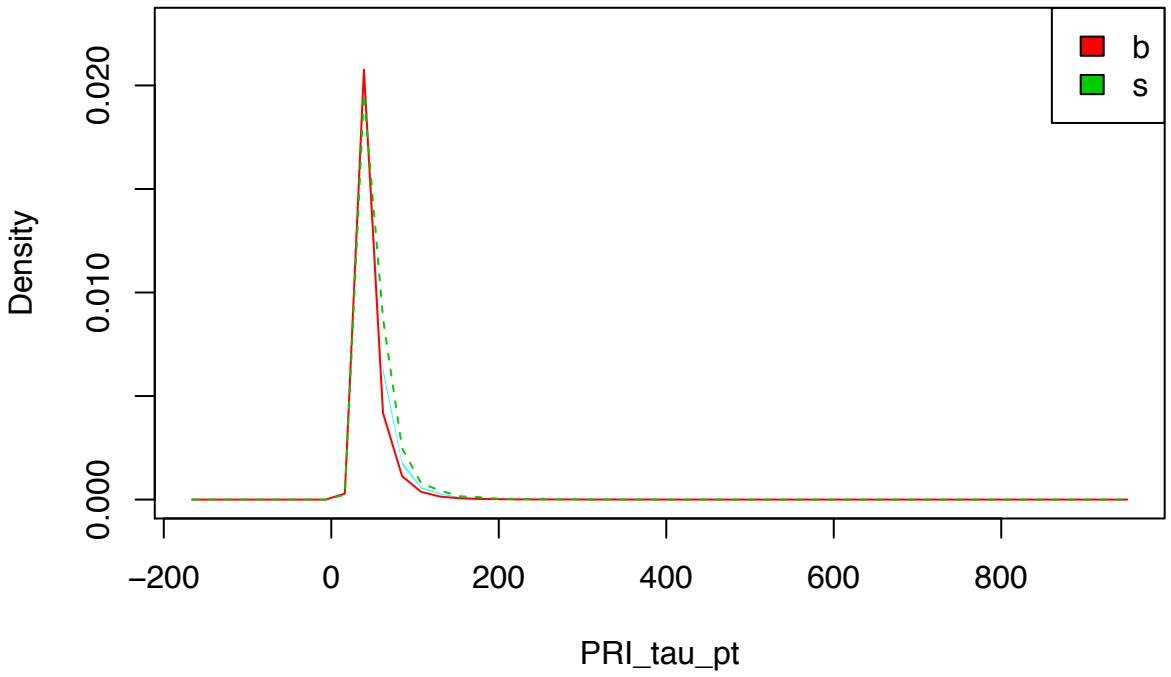

```
##
## Test of equal densities:  p-value =  0
```



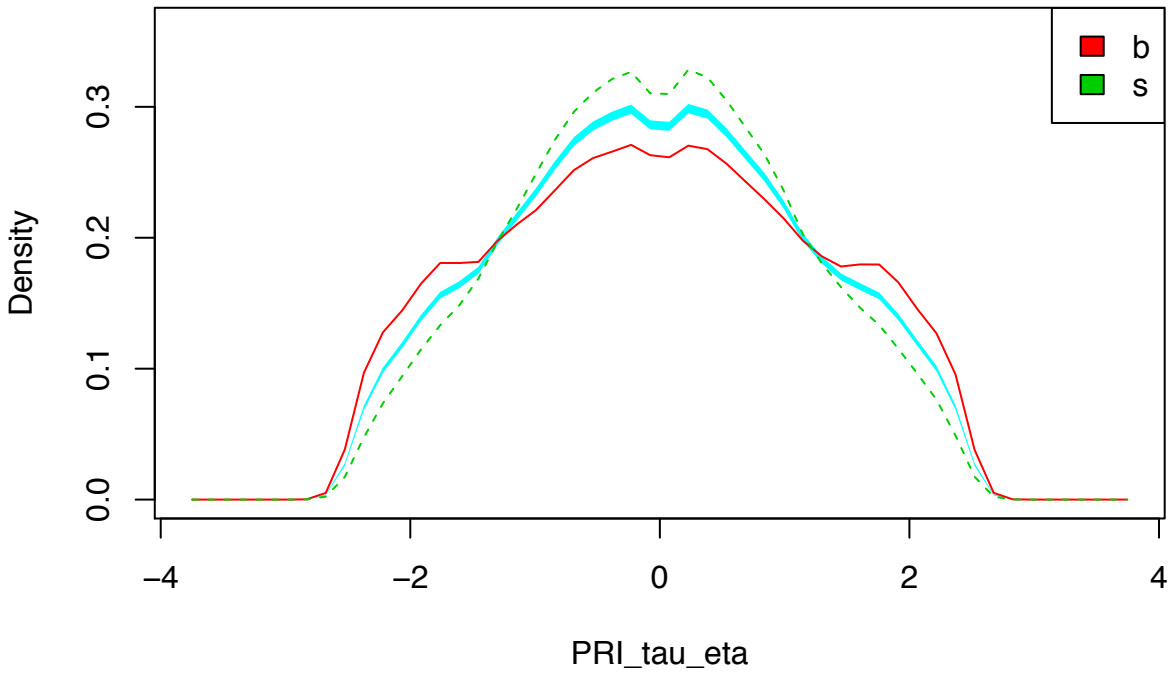
Density plots of Primary Variables for Signal & Background events

```
for(var in pri.vars){
  sm.density.compare(df[,var], df$Label, xlab=var,model="equal")
  legend("topright", levels(sb.f), fill=colfill)
}
```

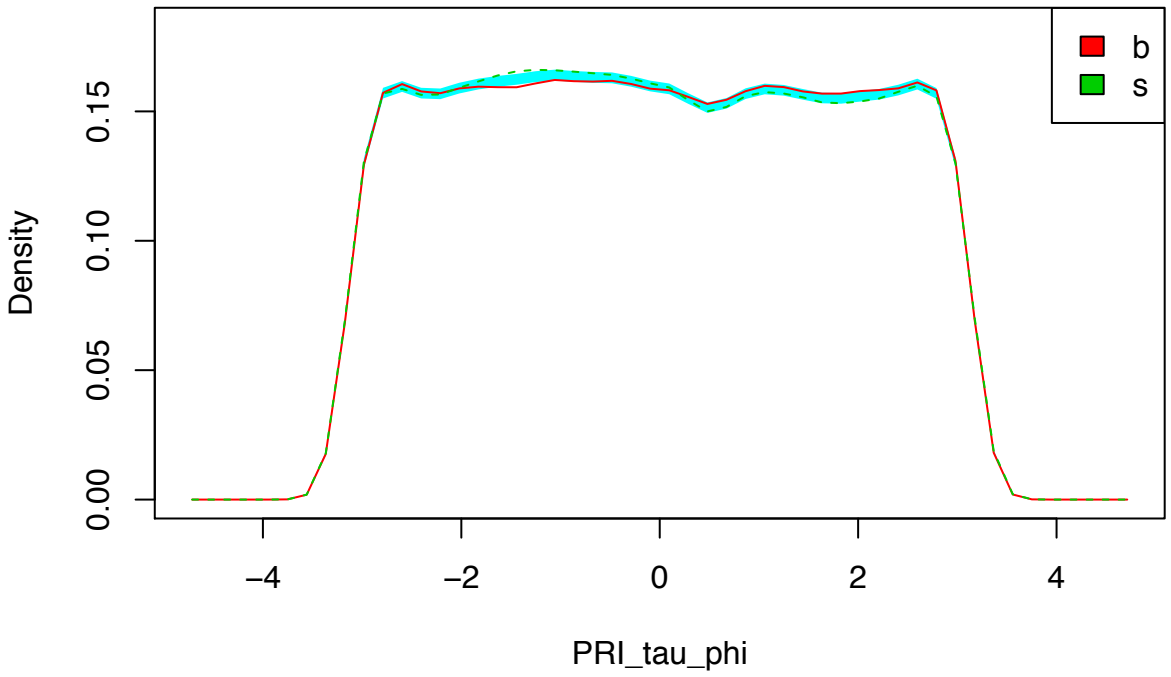
```
##
## Test of equal densities:  p-value =  0.87
```



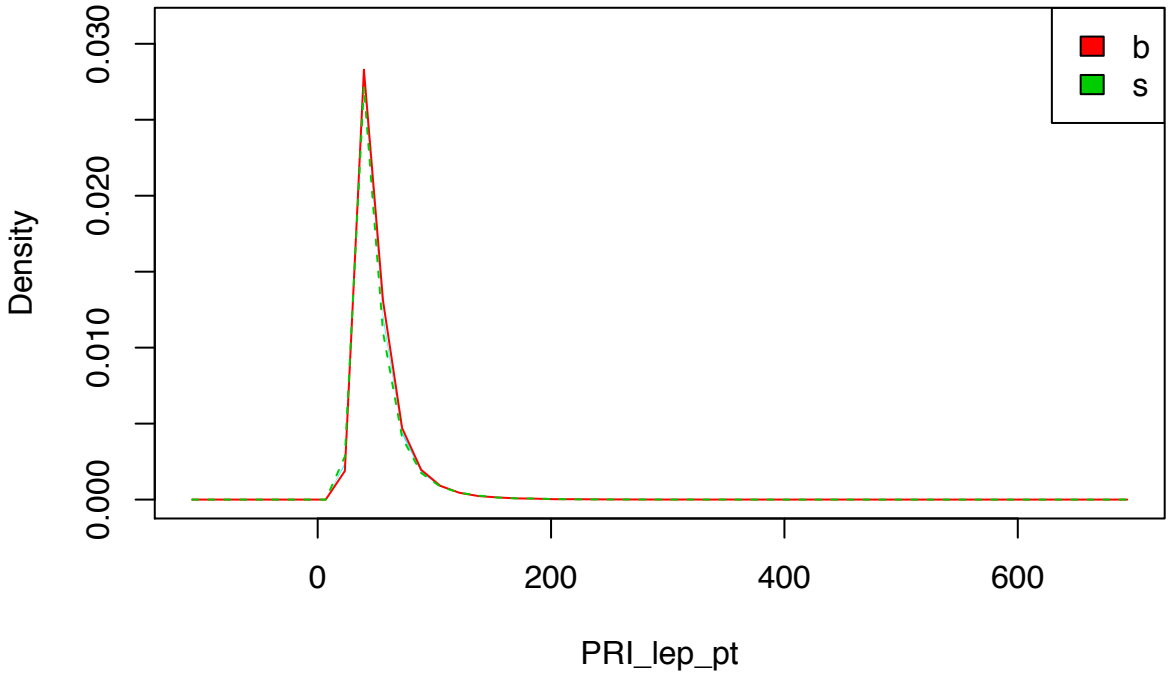

```
##
## Test of equal densities:  p-value =  0
```



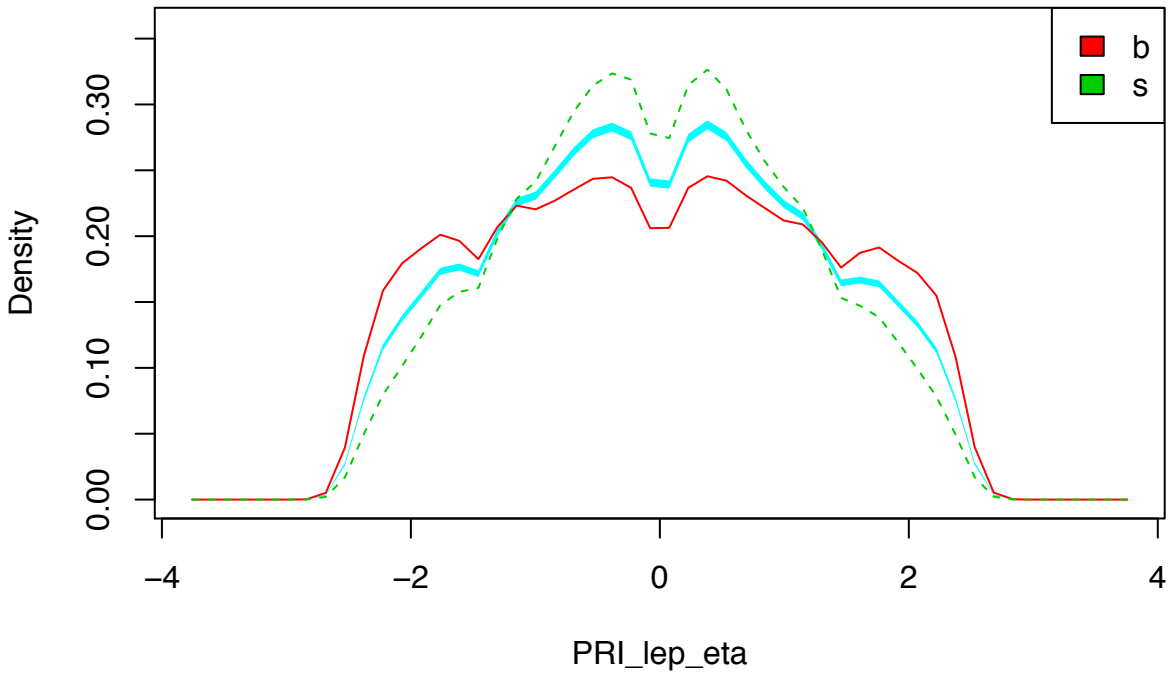

```
##
## Test of equal densities:  p-value =  0.01
```



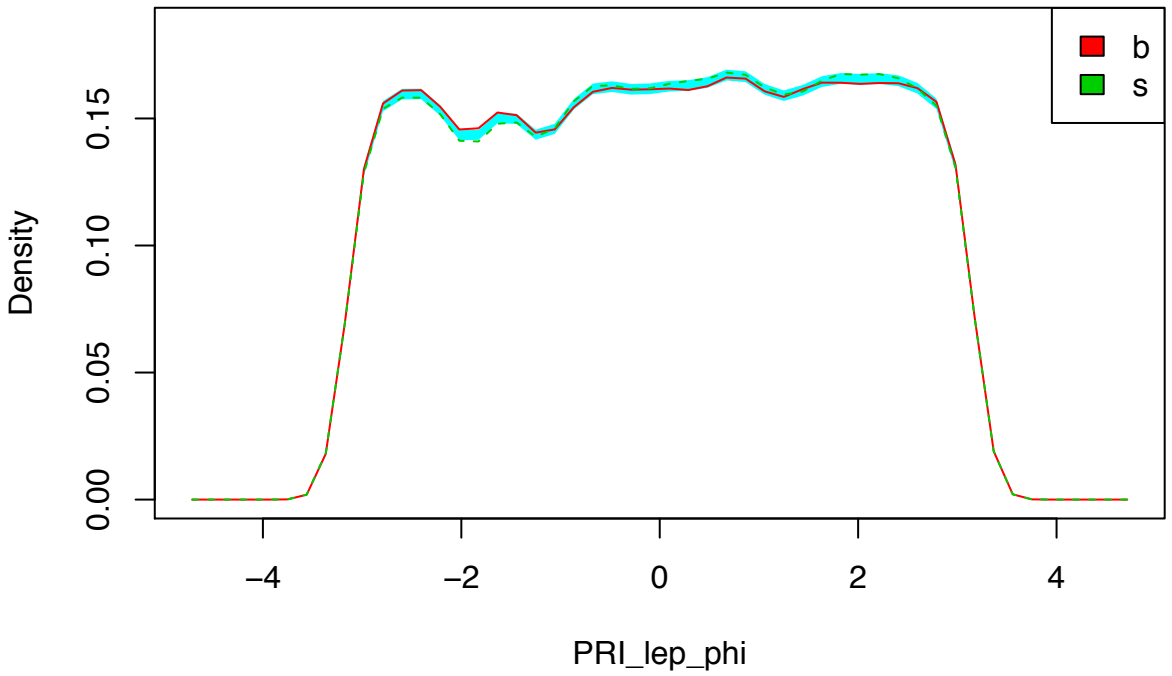

```
##
## Test of equal densities:  p-value =  0.04
```



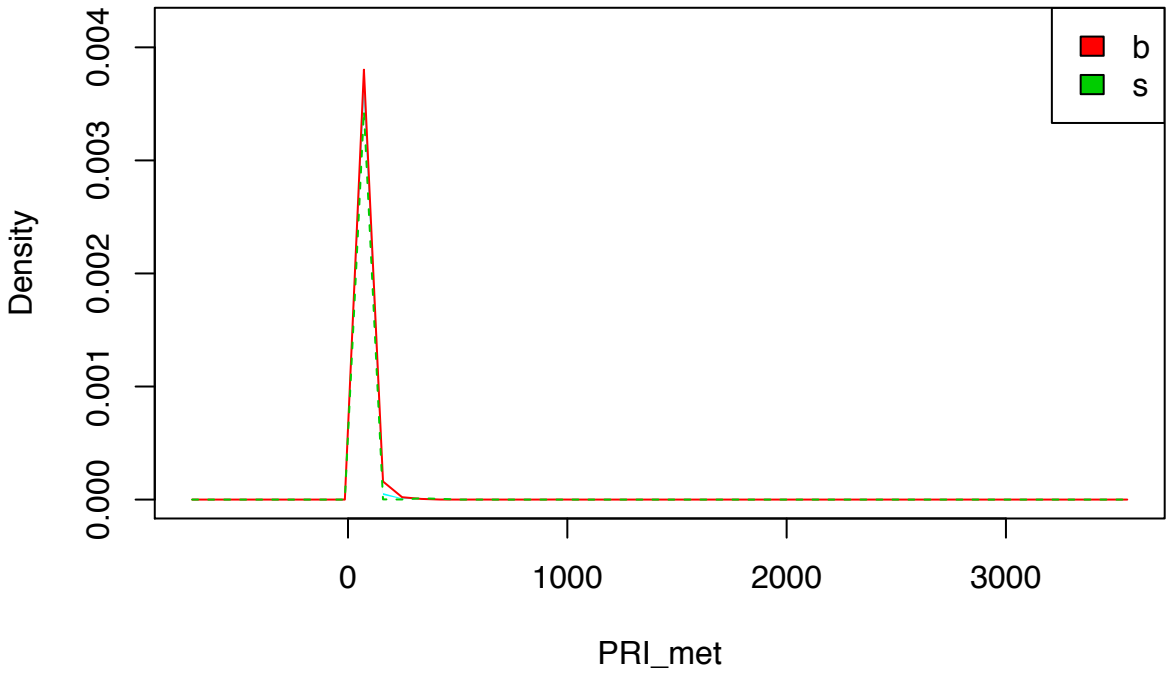

```
##
## Test of equal densities:  p-value =  0
```



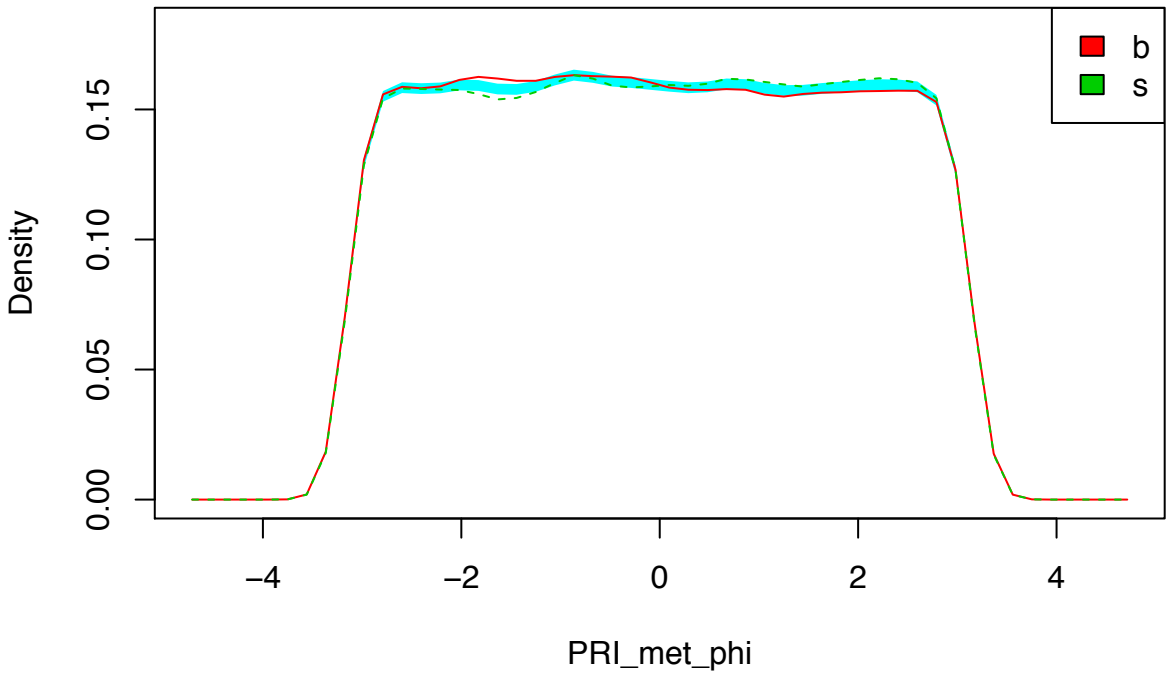

```
##
## Test of equal densities:  p-value =  0.1
```



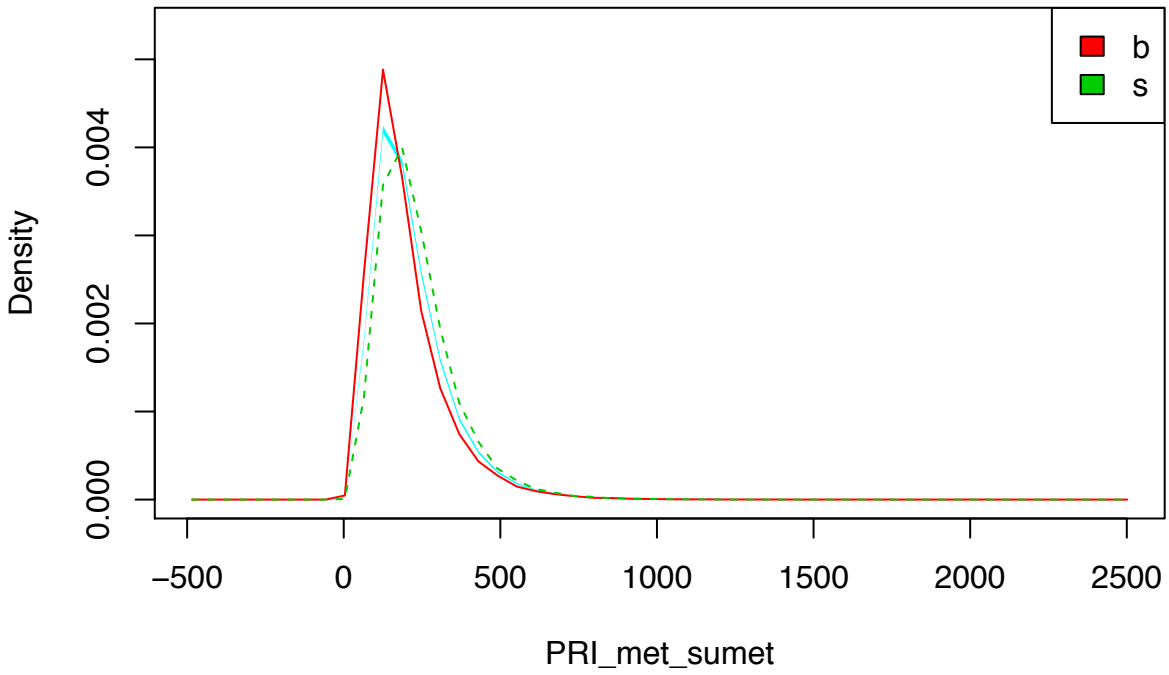

```
##
## Test of equal densities:  p-value =  1
```



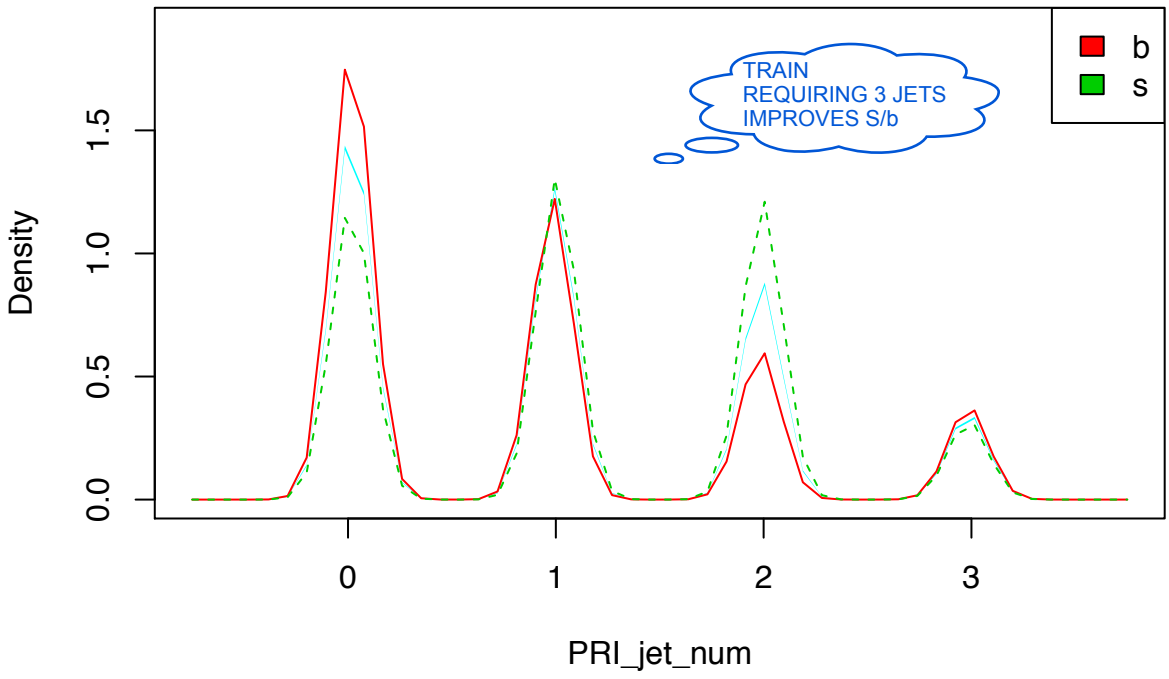

```
##
## Test of equal densities:  p-value =  0
```



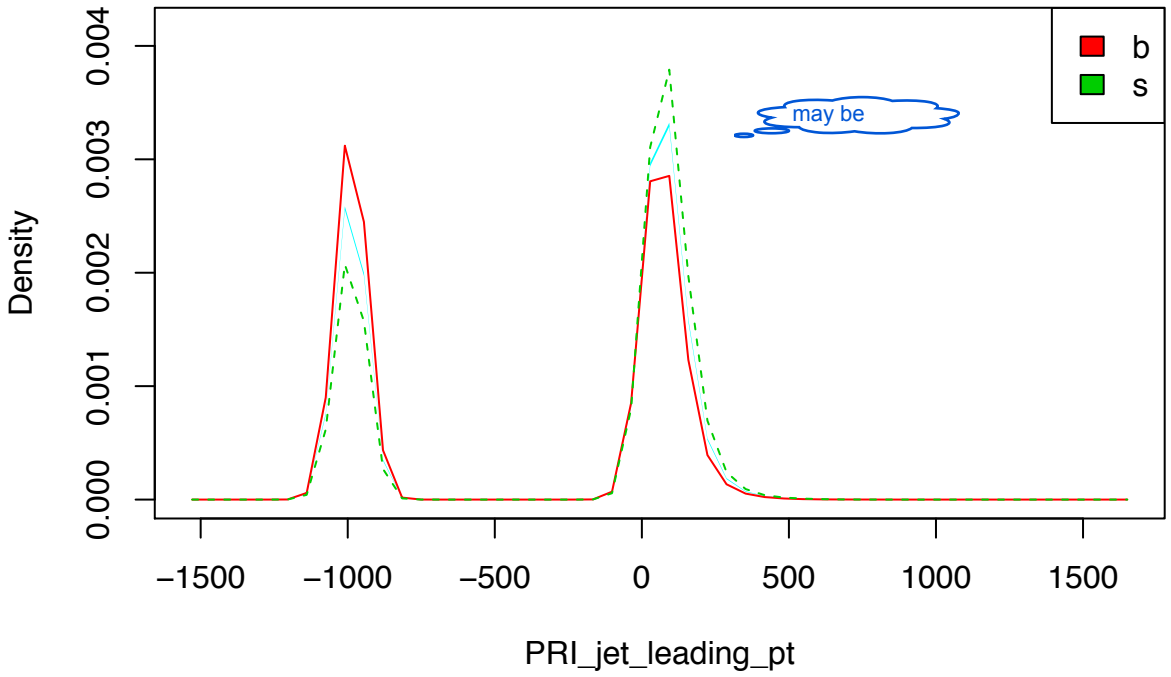

```
##
## Test of equal densities:  p-value =  0
```



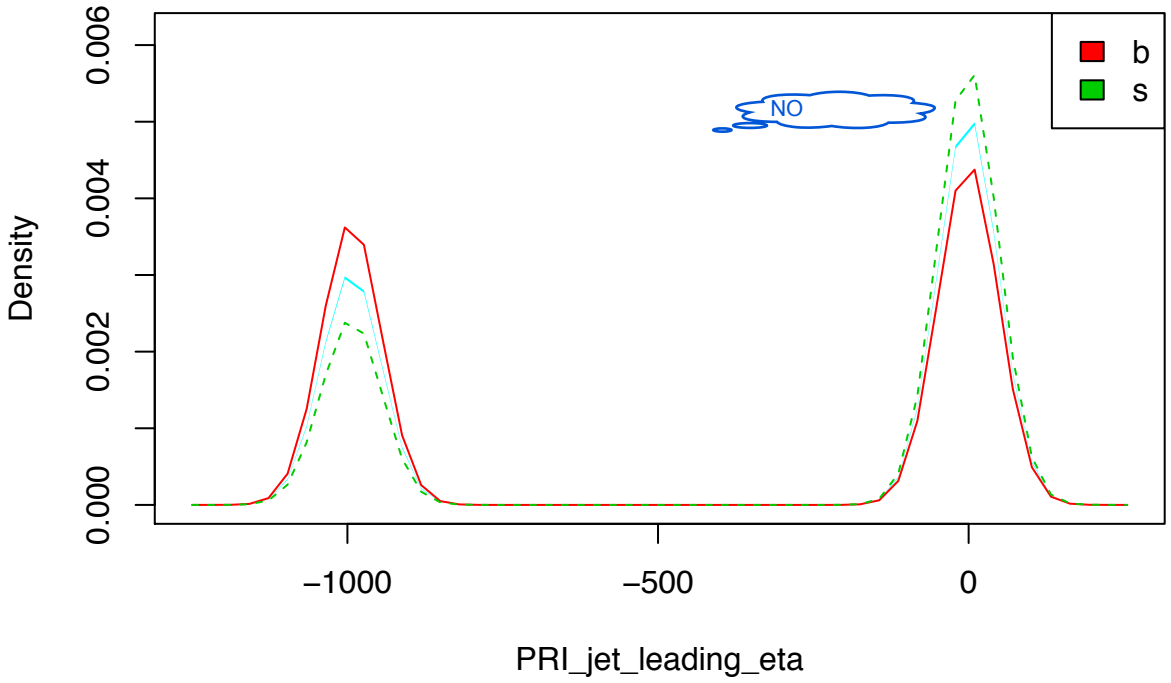

```
##
## Test of equal densities:  p-value =  0
```



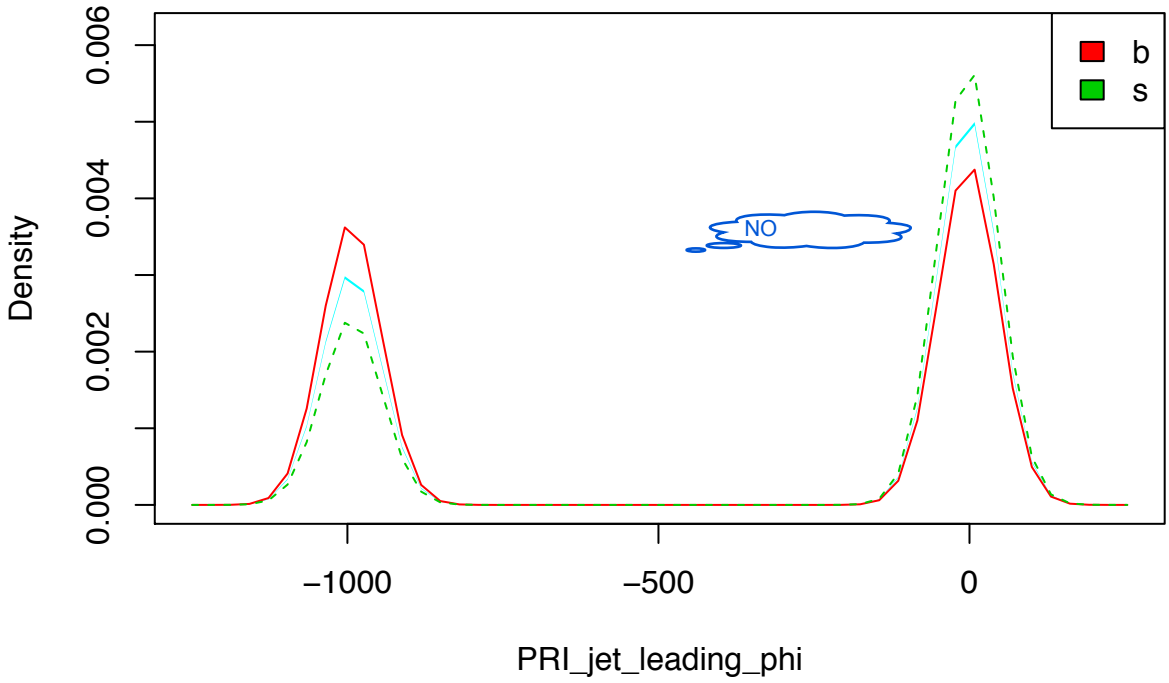

```
##
## Test of equal densities:  p-value =  0
```



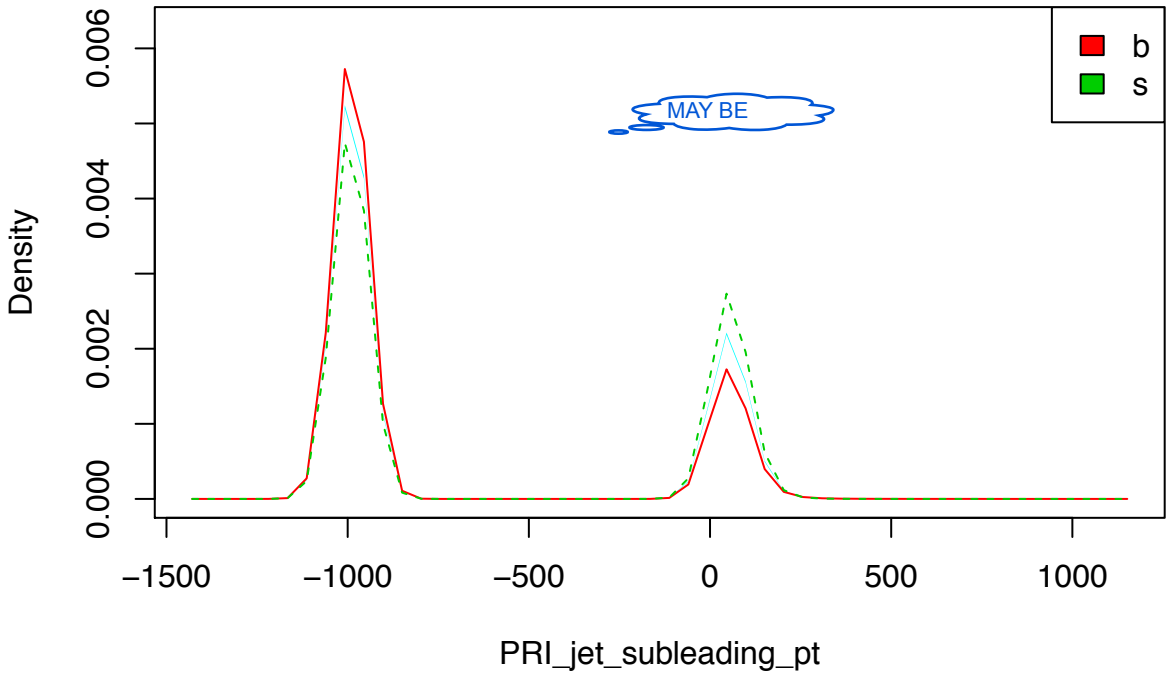

```
##
## Test of equal densities:  p-value =  0
```



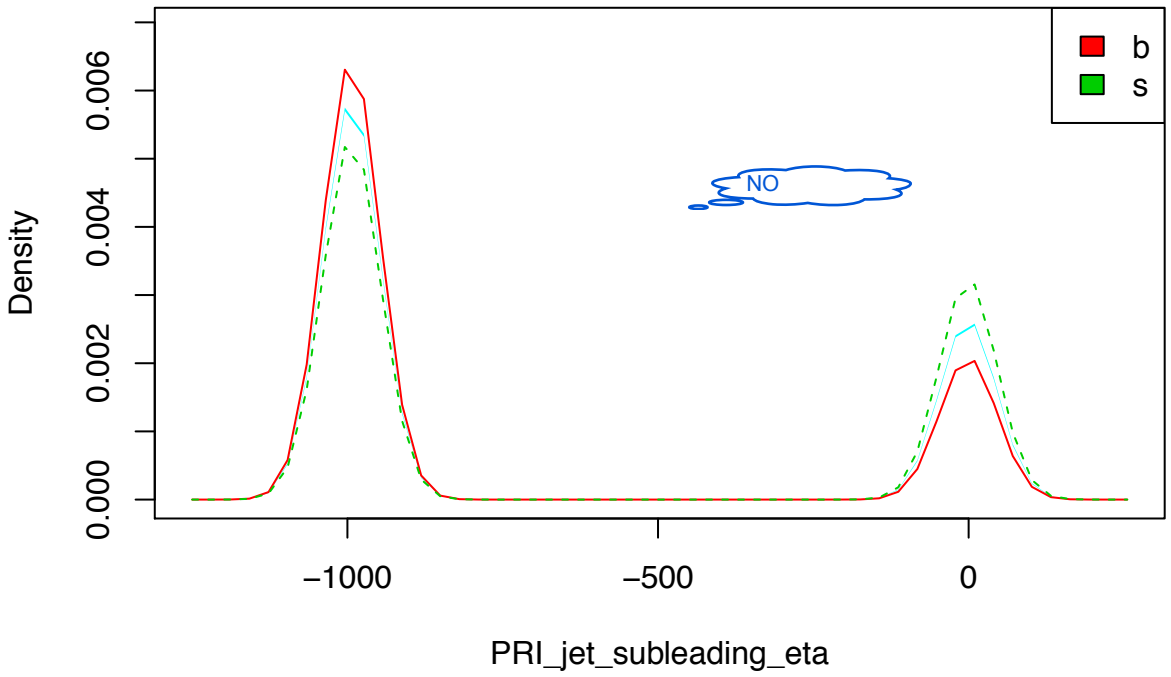

```
##
## Test of equal densities:  p-value =  0
```



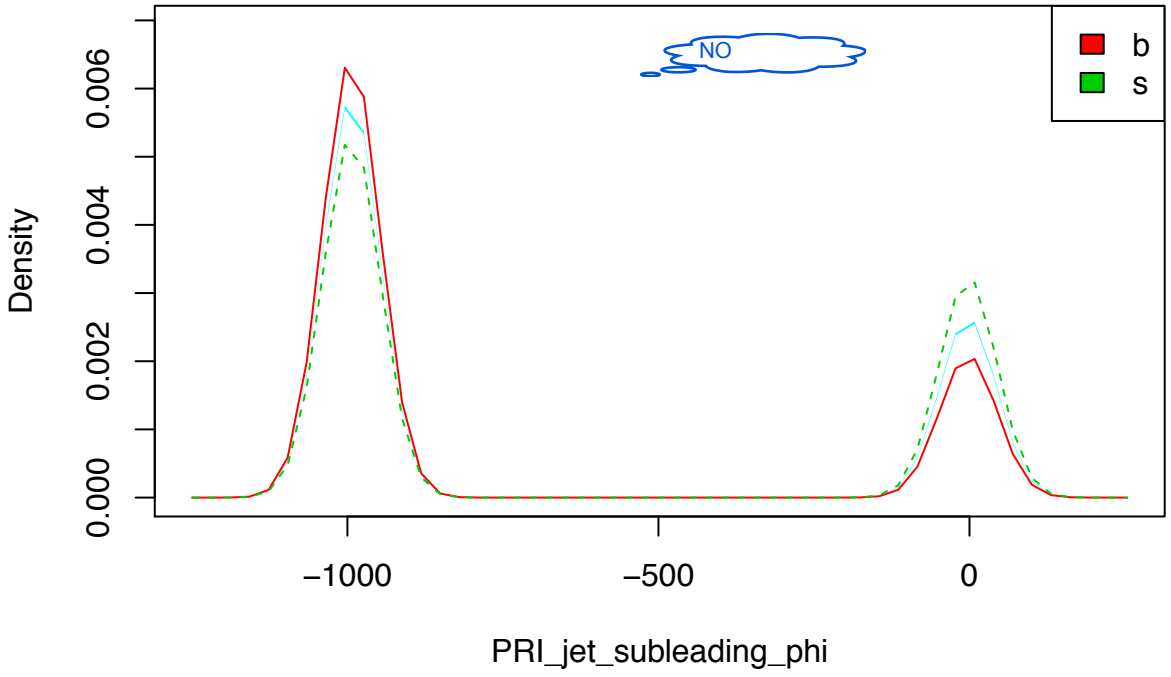

```
##
## Test of equal densities:  p-value =  0
```



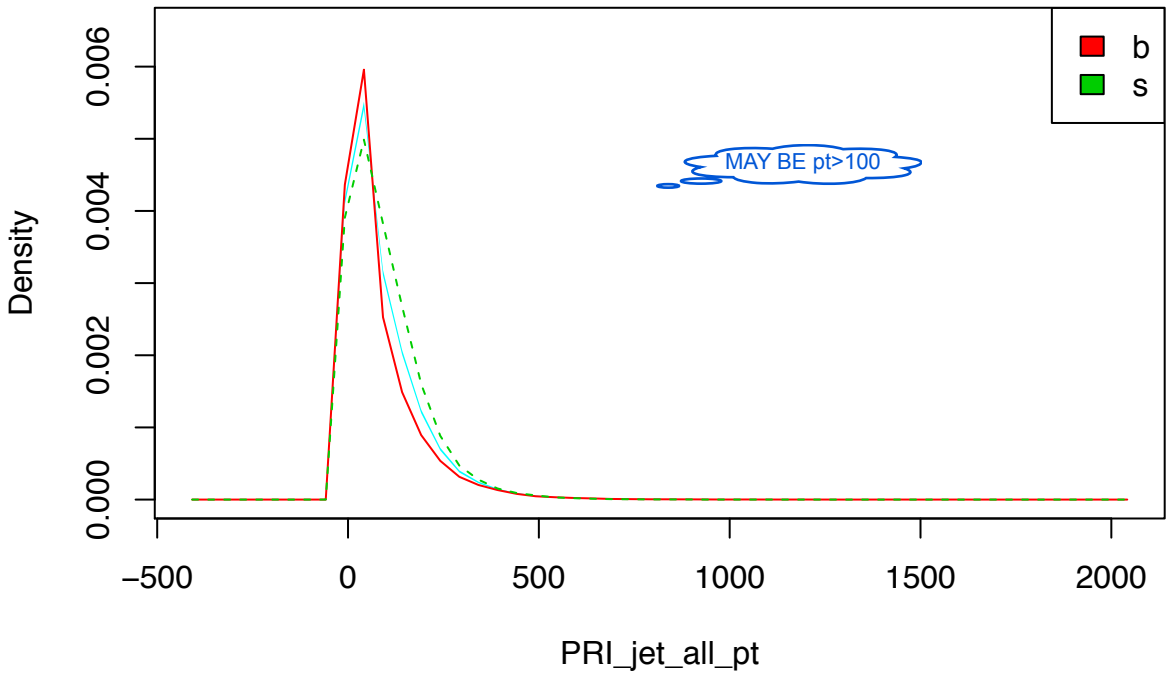

```
##
## Test of equal densities:  p-value =  0
```




```
##
## Test of equal densities:  p-value =  0
```




```
##
## Test of equal densities:  p-value =  0
```


b
s
MAY BE pt>100
Density
0.000
0.002
0.004
0.006
−500
0
500
1000
1500
2000
PRI_jet_all_pt