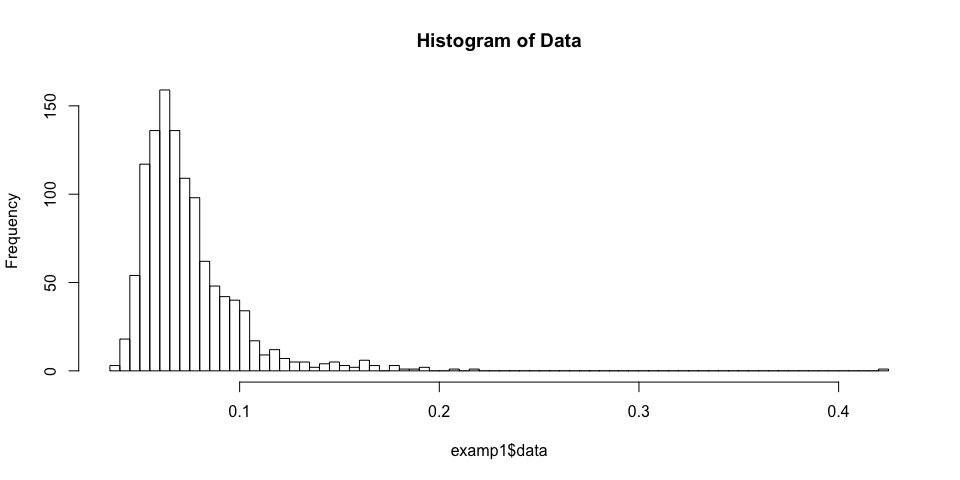
Mixture Modeling Paper Results

Sarah Sullivan

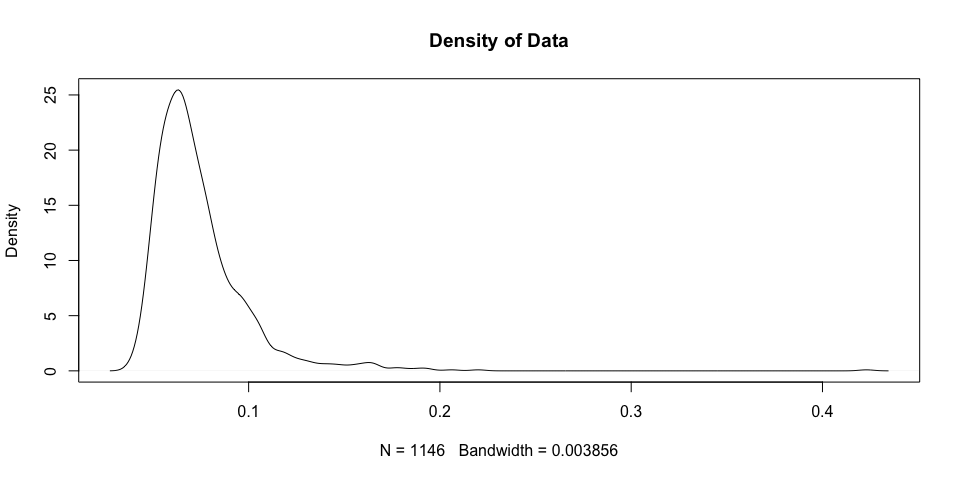
March 8, 2017

# Example 1

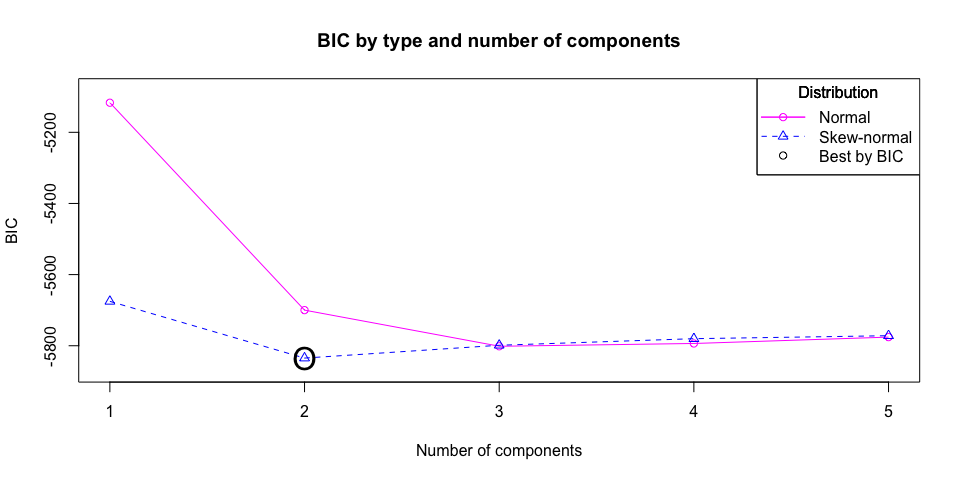
examp1<-read.csv(file="/Users/sarahsullivan/Documents/SullivanDocs/Rwd/cutoff/ex1.csv", colClasses=c("NULL", NA, NA))  
  
hist(examp1$data, breaks=100, main="Histogram of Data")



dens1<-density(examp1$data)  
plot(dens1, main="Density of Data")



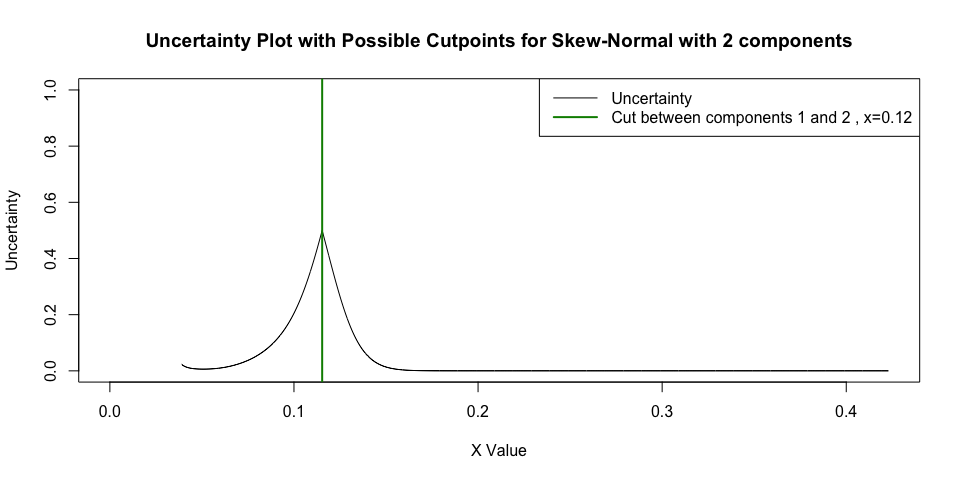
#fitloopsobject1<-fitloops(datawithids=examp2)  
 #might not run due to 5 component solution not being easily optimized, if it doesn't run either load the full file, or run fitloopsobject1<-fitloops(datawithids=examp2,maxcp=4):   
load(file="/Users/sarahsullivan/Documents/SullivanDocs/Rwd/cutoff/ex1fitloops")  
  
bicgraph(fitobj=fitloopsobject1)



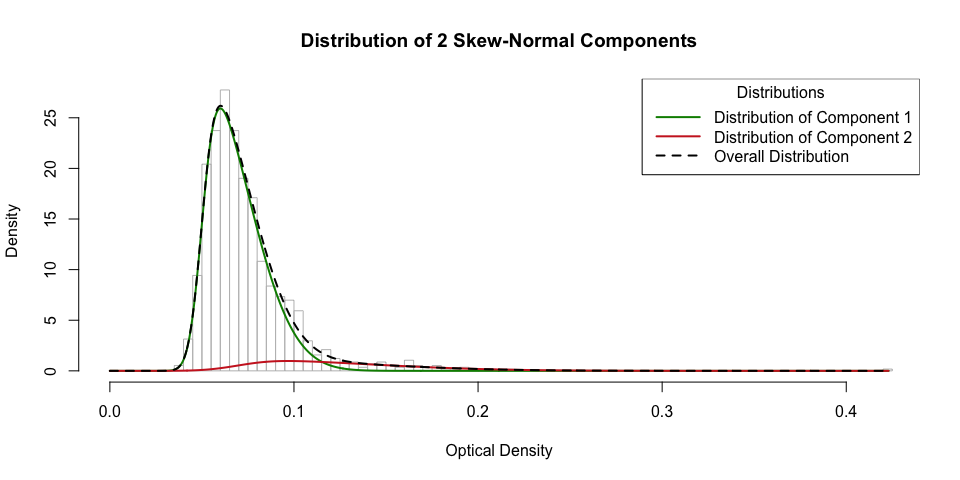
modelpickobject1<-modelpick(fitloopsobject1)

## Number Negative Number Positive  
## Cut between components 1 and 2 1083 63  
## Percent Negative Percent Positive  
## Cut between components 1 and 2 94.5% 5.5%

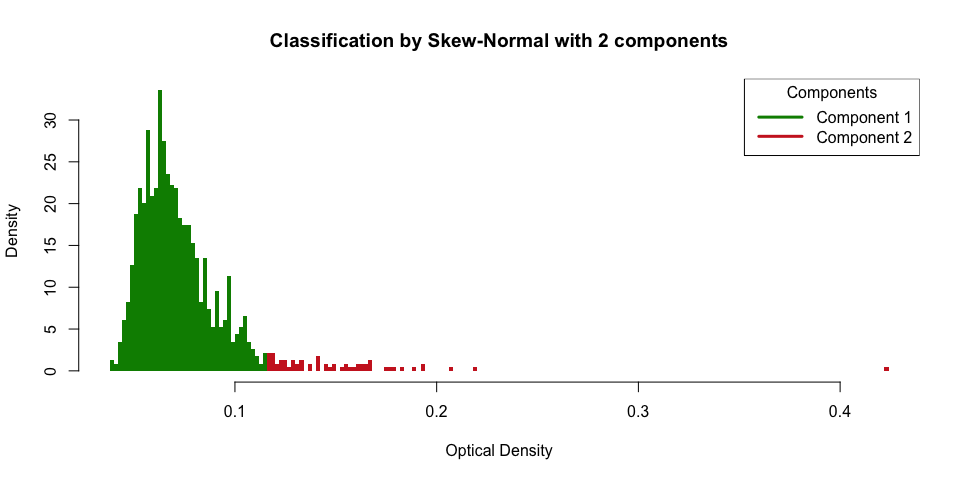
rawuncertgraph(modelpickobject1)



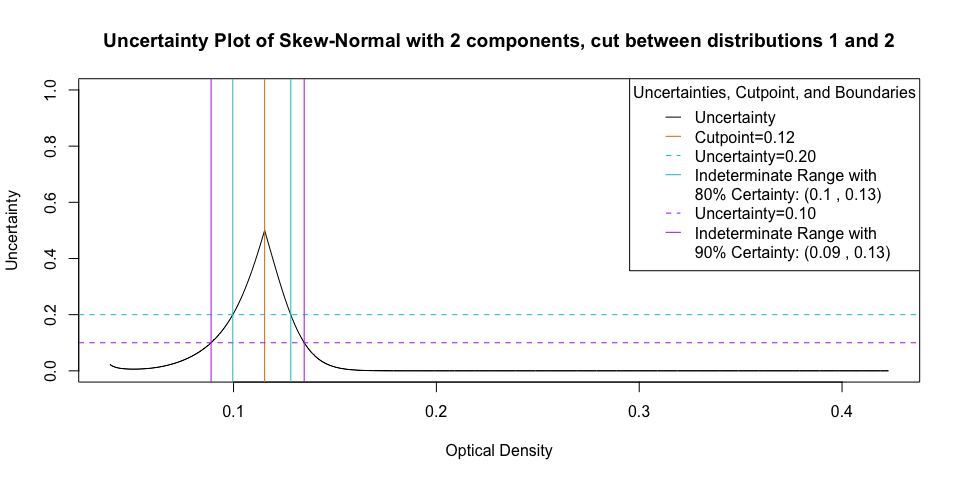
rawdistgraph(modelpickobject1)



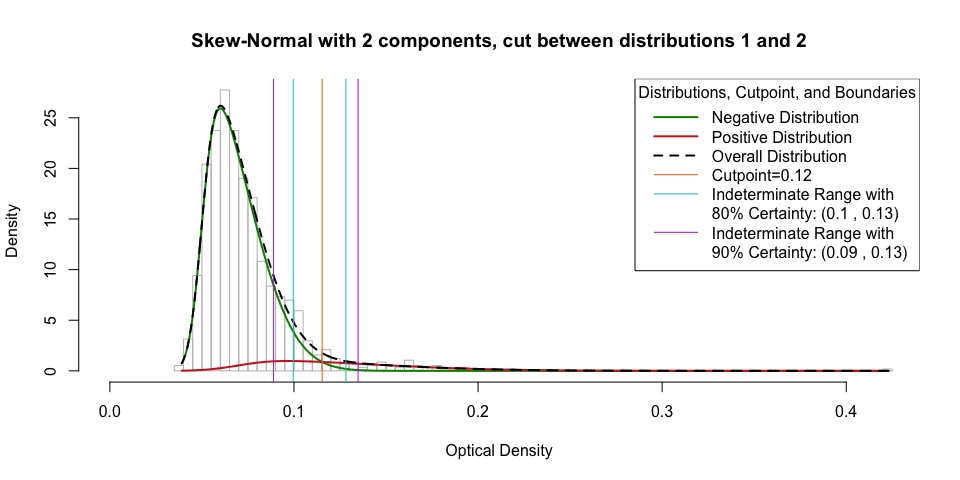
rawhistcuts(modelpickobject1)



cutoffobject1<-cutoff(modelpickobj = modelpickobject1)  
  
cutuncertgraph(cutobj = cutoffobject1)



cutdistgraph(cutobj = cutoffobject1)

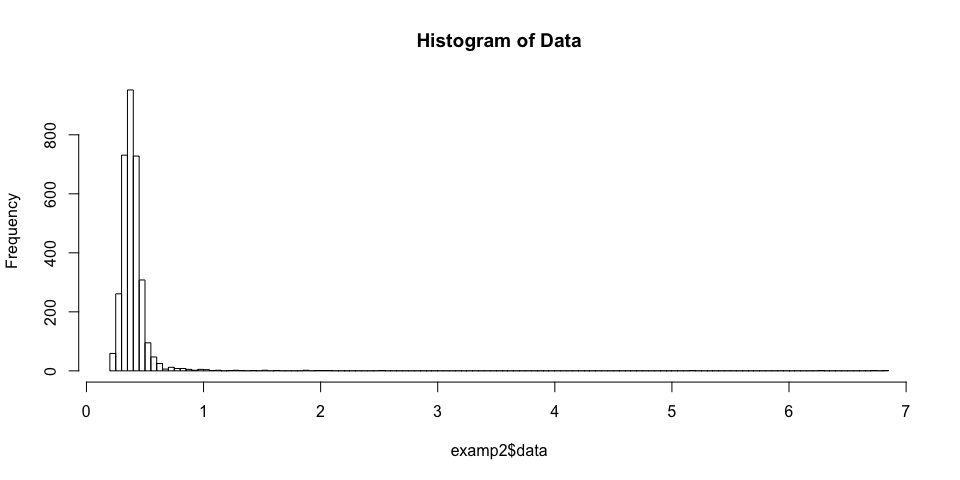


summaryobject1<-summaryout(cutobj = cutoffobject1)  
summaryobject1$outtab

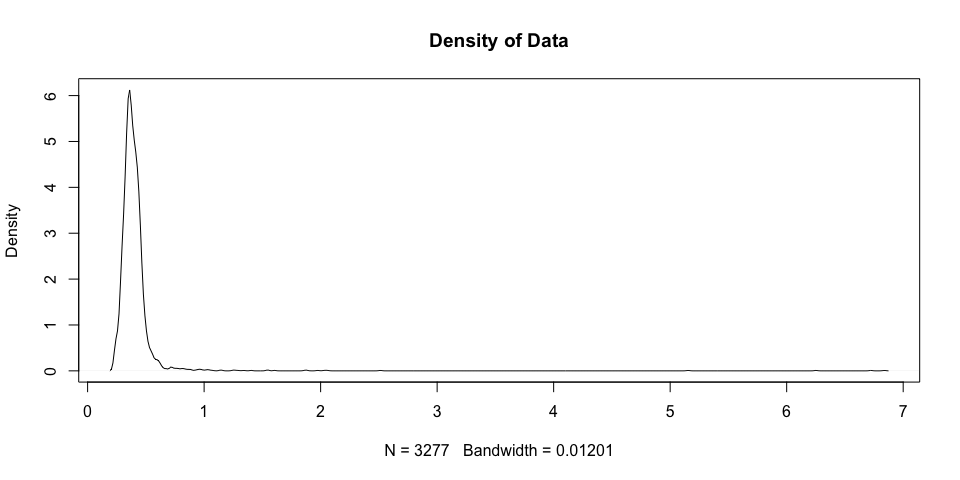
## No. Neg No. Indet No. Pos % Neg % Indet % Pos  
## Raw Cutoff 1083 0 63 94.5% 0% 5.5%  
## Cutoff with 80% Certainty 1022 82 42 89.18% 7.16% 3.66%  
## Cutoff with 90% Certainty 933 178 35 81.41% 15.53% 3.05%

# Example 2

examp2<-read.csv(file="/Users/sarahsullivan/Documents/SullivanDocs/Rwd/cutoff/ex2.csv", stringsAsFactors = F, colClasses=c("NULL", NA, NA))  
examp2<-examp2[,c("id", "data")]  
  
hist(examp2$data, breaks=100, main="Histogram of Data")



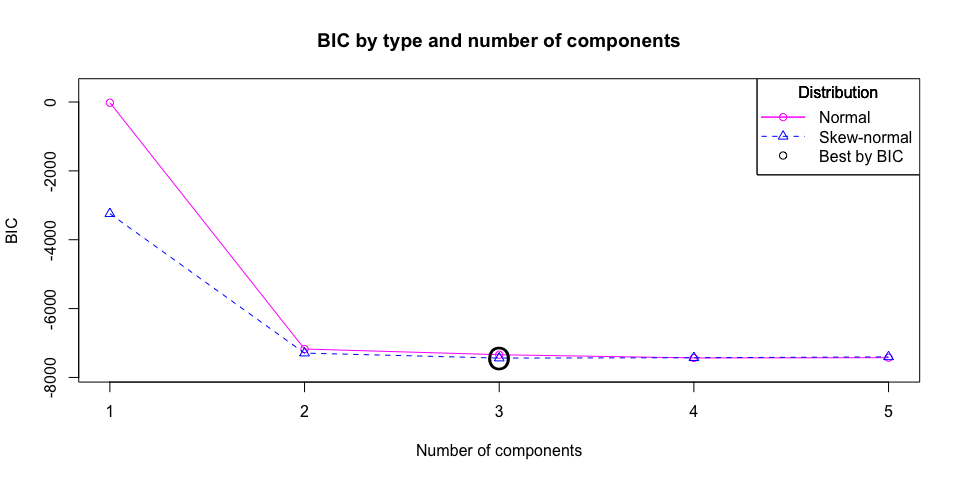
dens2<-density(examp2$data)  
plot(dens2, main="Density of Data")



fitloopsobject2<-fitloops(datawithids = examp2)

## [1] ""  
## [1] "The best model by BIC is Skew-normal with 3 components"  
## [1] ""  
## [1] "Bayesian Information Criterion (BIC) Matrix"  
## [1] "BIC should be minimized and a difference of 10 BIC indicates strong evidence that the model with lower BIC is superior"  
## Normal Skew-normal  
## 1 Component(s): -20.46063 -3242.631  
## 2 Component(s): -7174.86787 -7293.374  
## 3 Component(s): -7340.69321 -7437.852  
## 4 Component(s): -7435.03427 -7427.607  
## 5 Component(s): -7426.57499 -7403.177  
## [1] "Below is a table of the BIC's of the most common distribution and number of component combinations to base a cutpoint on"  
## Description BIC   
## Best Overall Skew-normal with 3 components -7437.852  
## Best Skew-normal Skew-normal with 3 components -7437.852  
## Best Normal Normal with 4 components -7435.034  
## Two Skew-normal Skew-normal with 2 components -7293.374  
## Two Normal Normal with 2 components -7174.868

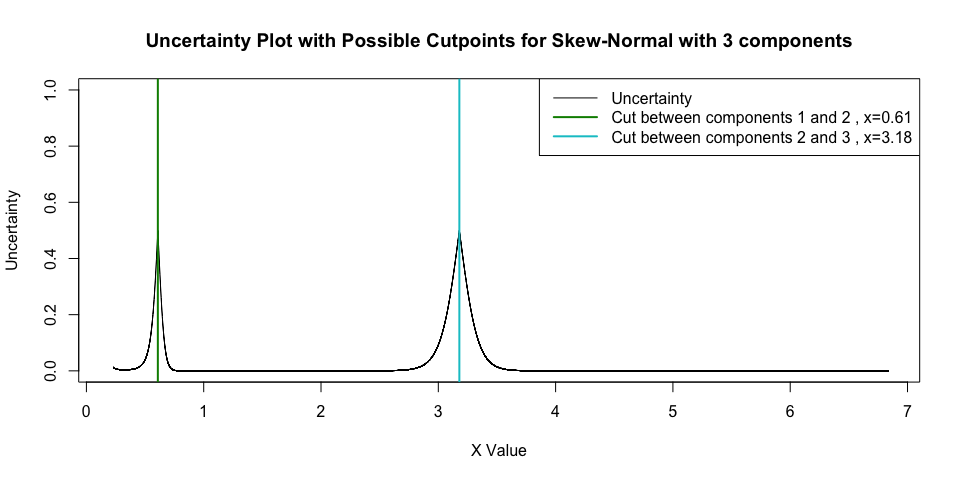
bicgraph(fitobj = fitloopsobject2)



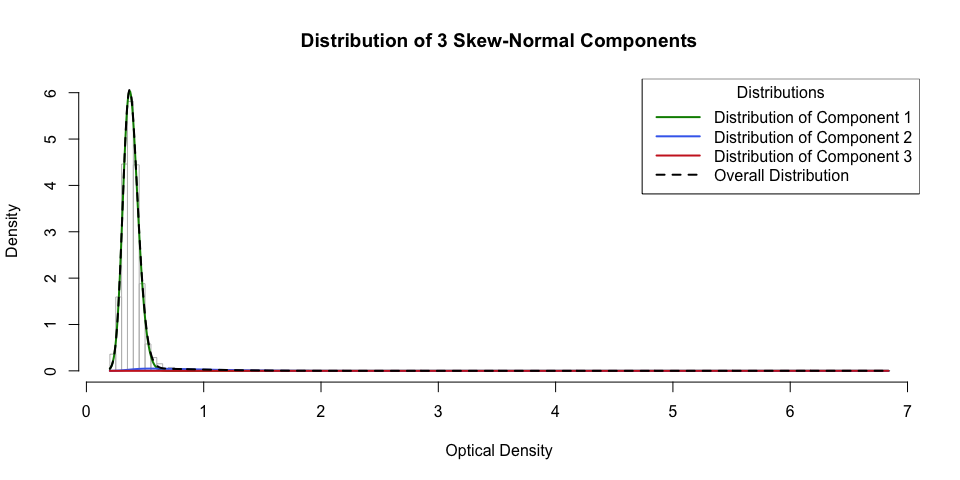
modelpickobject2<-modelpick(fitobj = fitloopsobject2)

## Number Negative Number Positive  
## Cut between components 1 and 2 3187 90  
## Cut between components 2 and 3 3273 4  
## Percent Negative Percent Positive  
## Cut between components 1 and 2 97.25% 2.75%  
## Cut between components 2 and 3 99.88% 0.12%

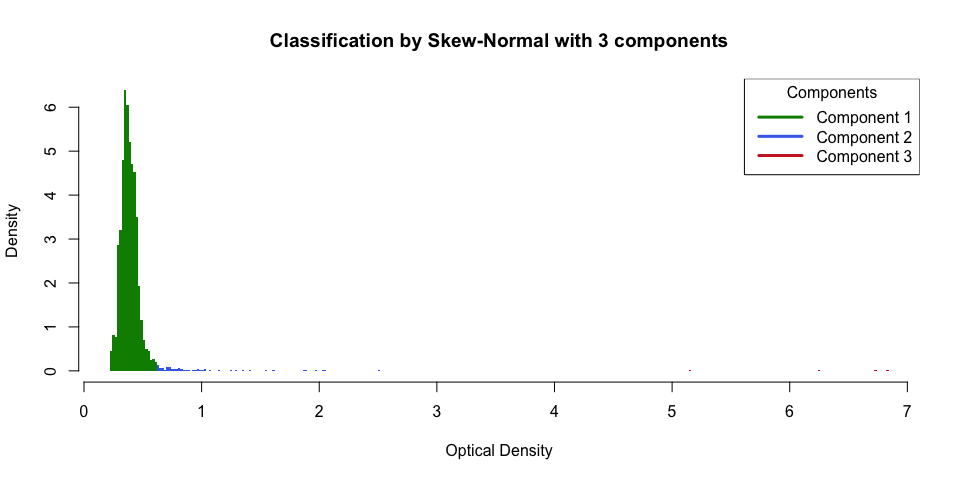
rawuncertgraph(modelpickobj = modelpickobject2)



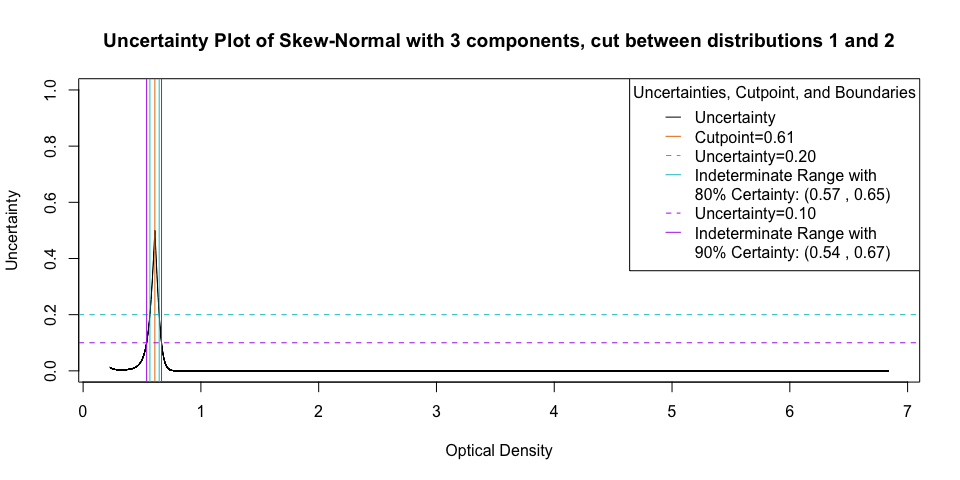
rawdistgraph(modelpickobj = modelpickobject2)



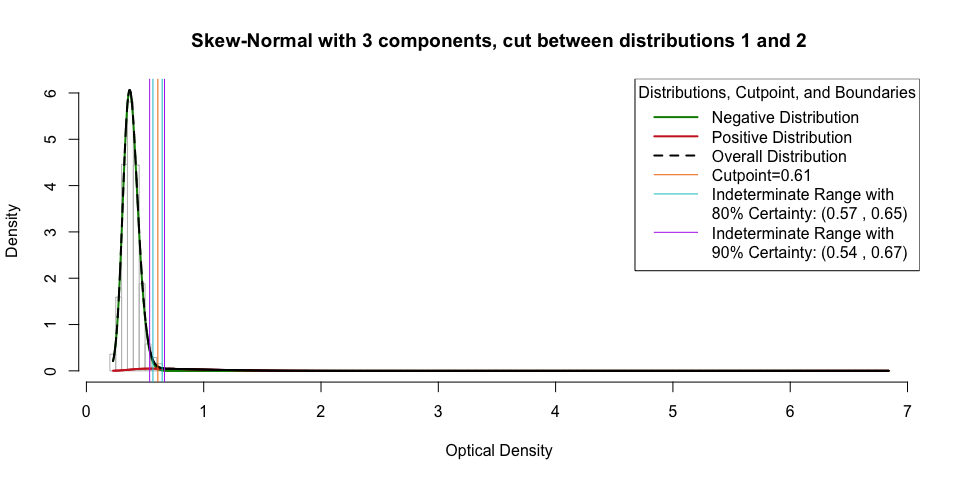
rawhistcuts(modelpickobj = modelpickobject2)



cutoffobject2<-cutoff(modelpickobj = modelpickobject2, cutcomp = 1)  
  
cutuncertgraph(cutobj = cutoffobject2)



cutdistgraph(cutobj = cutoffobject2)



summaryobject2<-summaryout(cutobj = cutoffobject2)  
summaryobject2$outtab

## No. Neg No. Indet No. Pos % Neg % Indet % Pos  
## Raw Cutoff 3187 0 90 97.25% 0% 2.75%  
## Cutoff with 80% Certainty 3156 49 72 96.31% 1.5% 2.2%  
## Cutoff with 90% Certainty 3117 90 70 95.12% 2.75% 2.14%