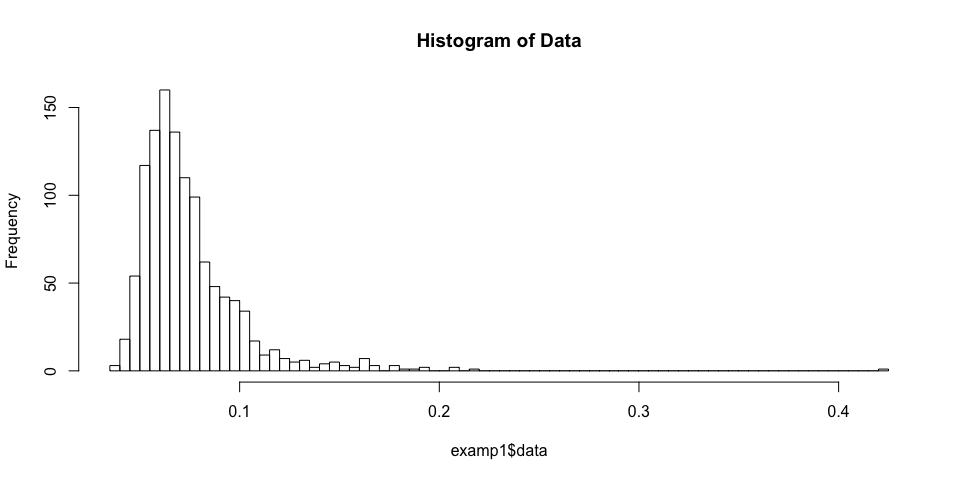
Mixture Modeling Paper Results

Sarah Sullivan

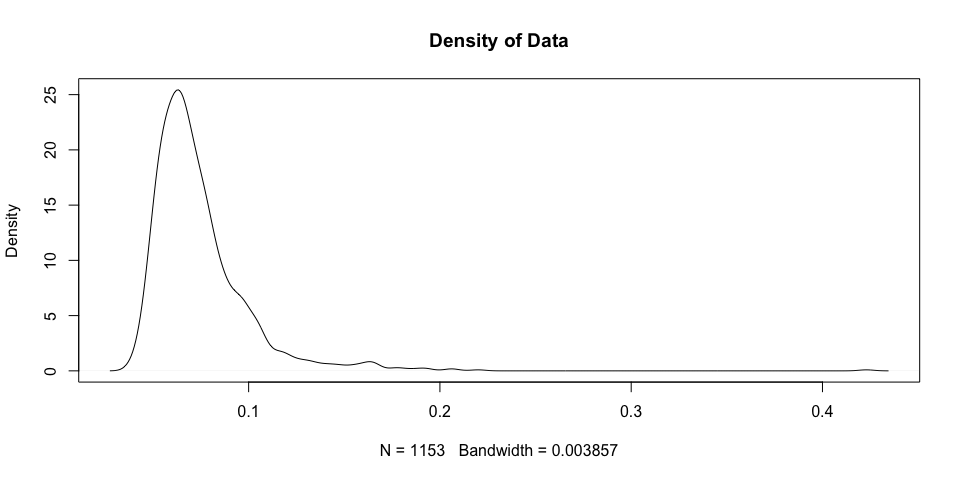
March 8, 2017

# 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")



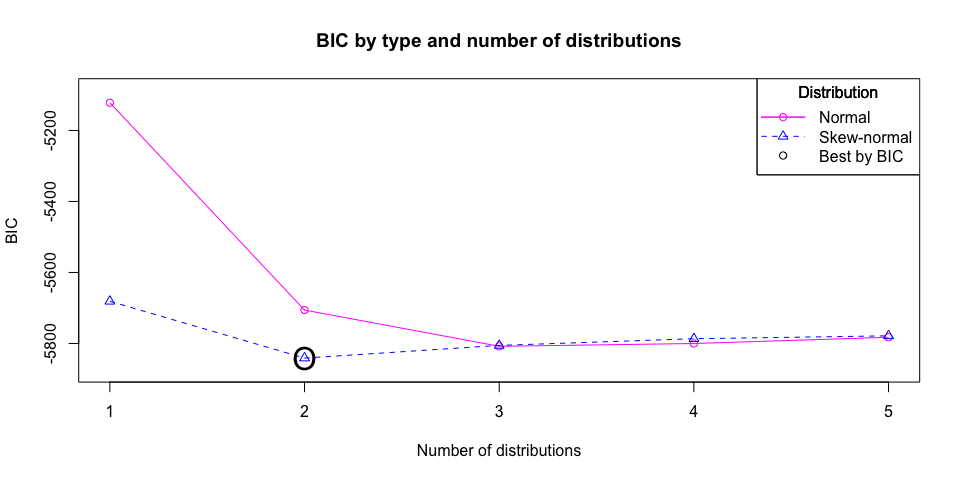
dens1<-density(examp1$data)  
plot(dens1, main="Density of Data")



#fitloopsobject<-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 fitloopsobject<-fitloops(datawithids=examp2,maxcp=4):   
load(file="/Users/sarahsullivan/Documents/SullivanDocs/Rwd/cutoff/ex1fitloops")  
  
fitloopsobject1<-ex1  
  
bestfitsobject1<-bestfits(fitres=fitloopsobject1)

## [1] "The best model by BIC is Skew-normal with 2 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): -5121.933 -5681.232  
## 2 Component(s): -5705.934 -5841.175  
## 3 Component(s): -5808.031 -5805.496  
## 4 Component(s): -5800.123 -5786.655  
## 5 Component(s): -5782.535 -5778.538  
## [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 2 components -5841.175  
## Best Skew-normal Skew-normal with 2 components -5841.175  
## Best Normal Normal with 3 components -5808.031  
## Two Skew-normal Skew-normal with 2 components -5841.175  
## Two Normal Normal with 2 components -5705.934

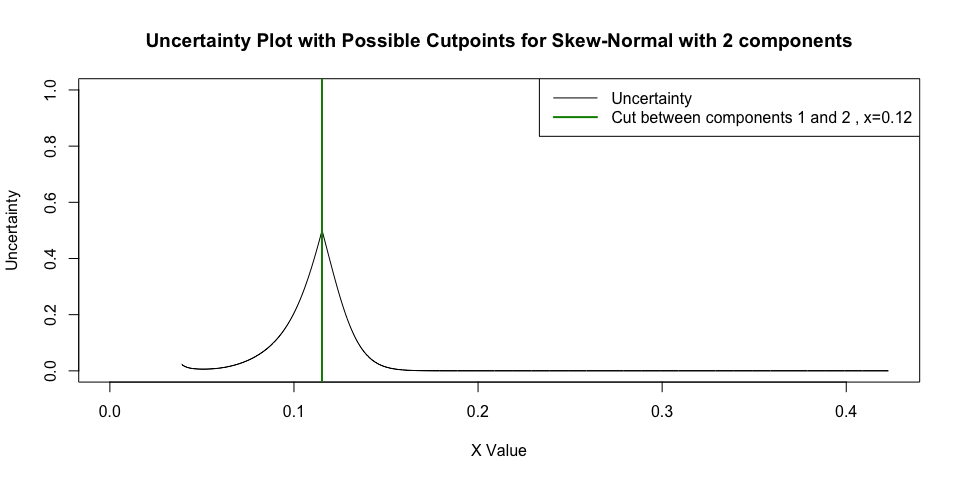
bicgraph(model=bestfitsobject1)



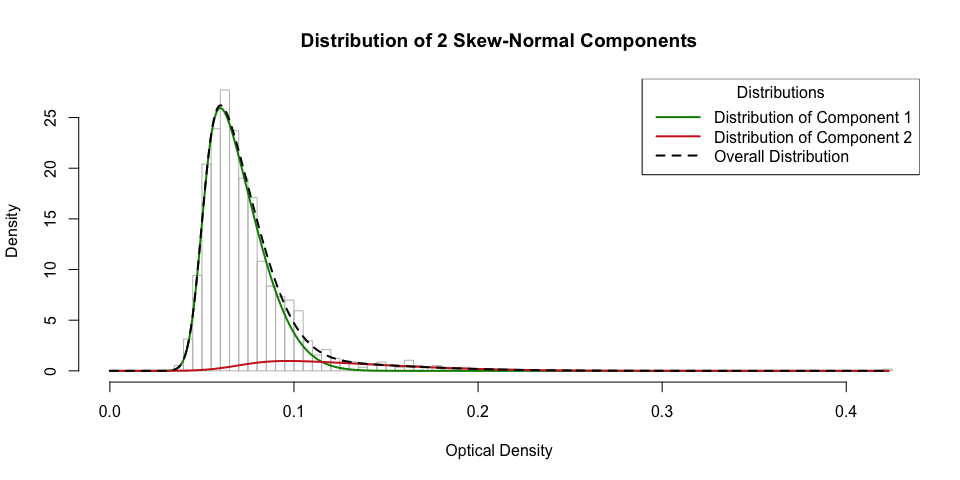
fitpickobject1<-fitpick(fitres=fitloopsobject1,bestfits=bestfitsobject1)  
  
uncertobject1<-uncert(fitpickobj=fitpickobject1)  
multcutobject1<-multcut(uncertobj=uncertobject1)

## Number Negative Number Positive  
## Cut between components 1 and 2 1084 63  
## Percent Negative Percent Positive  
## Cut between components 1 and 2 94.51% 5.49%

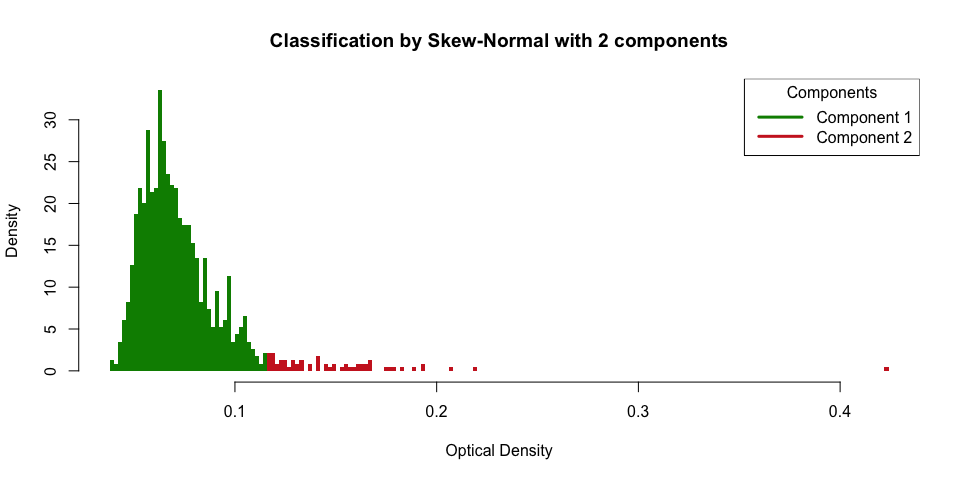
rawuncertgraph(uncertobj=uncertobject1,multcutobj=multcutobject1)



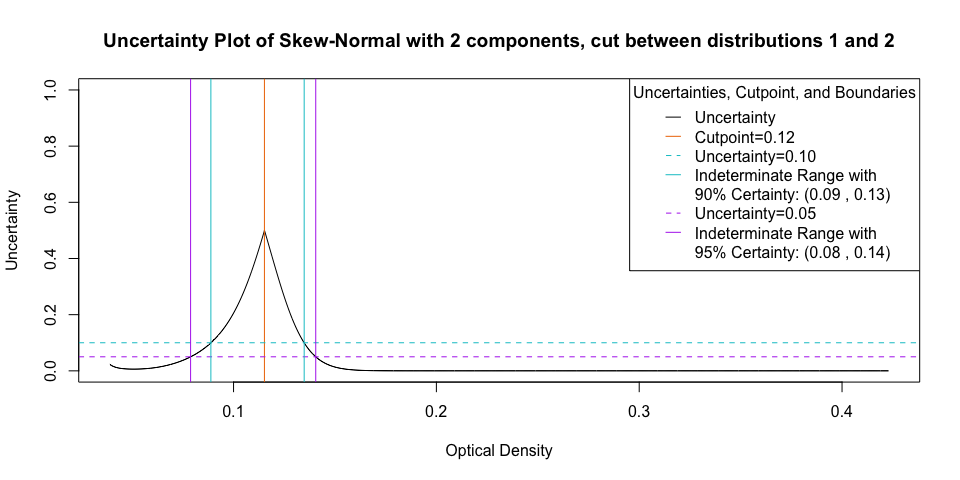
rawdistgraph(pickobj=fitpickobject1)



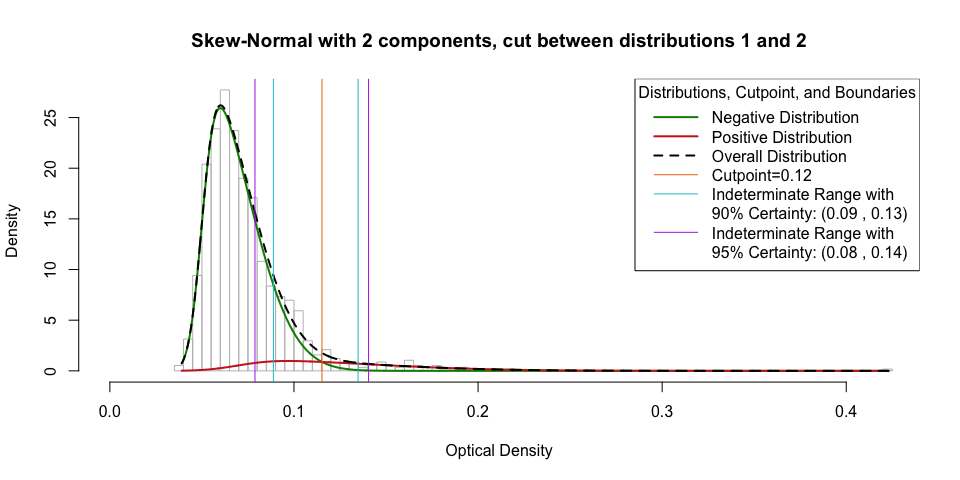
rawhistcuts(uncertobj=uncertobject1,multcutobj=multcutobject1)



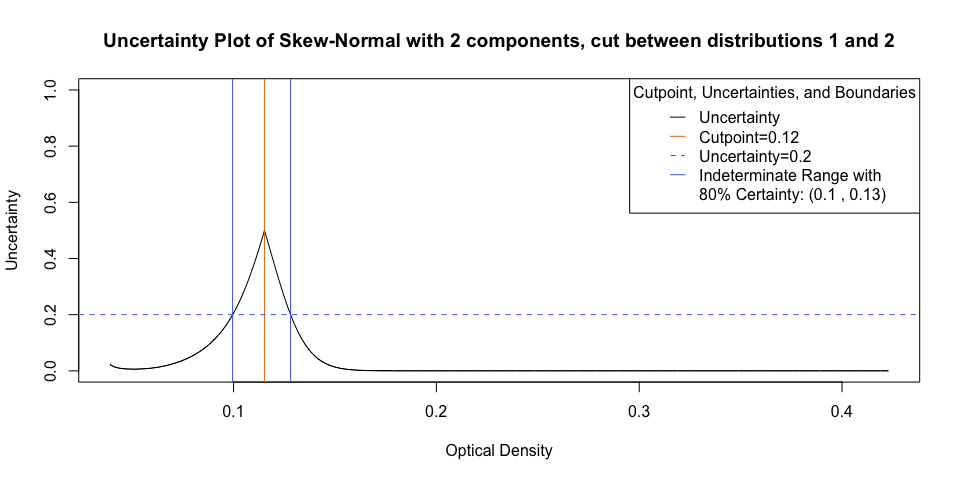
standindetobject1<-standindet(uncertobj=uncertobject1,multcutobj=multcutobject1,cutcomp=1)  
specindetobject1<-specindet(uncertobj=uncertobject1,multcutobj=multcutobject1,cutcomp=1,certlevel=0.8)  
  
cutuncertgraph(cutobj=standindetobject1)



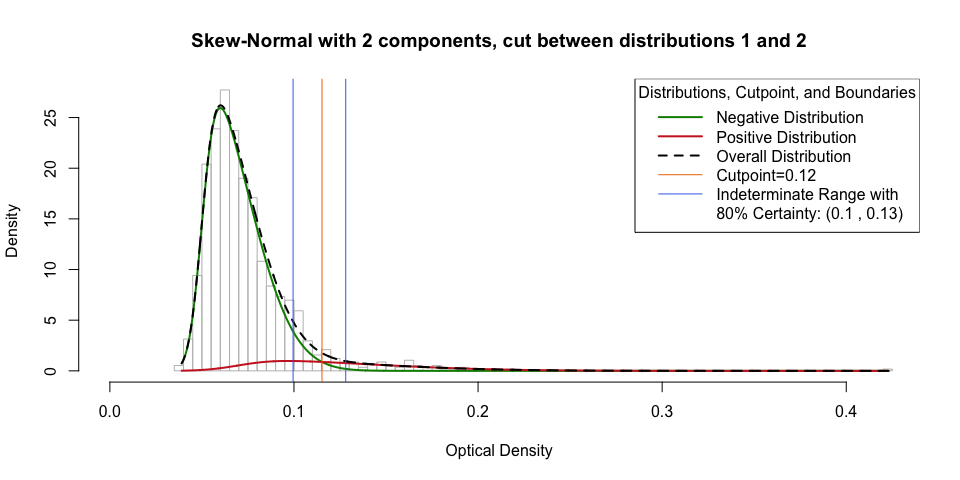
cutdistgraph(cutobj=standindetobject1, pickobj=fitpickobject1)



cutuncertgraph(cutobj=specindetobject1)



cutdistgraph(cutobj=specindetobject1, pickobj=fitpickobject1)

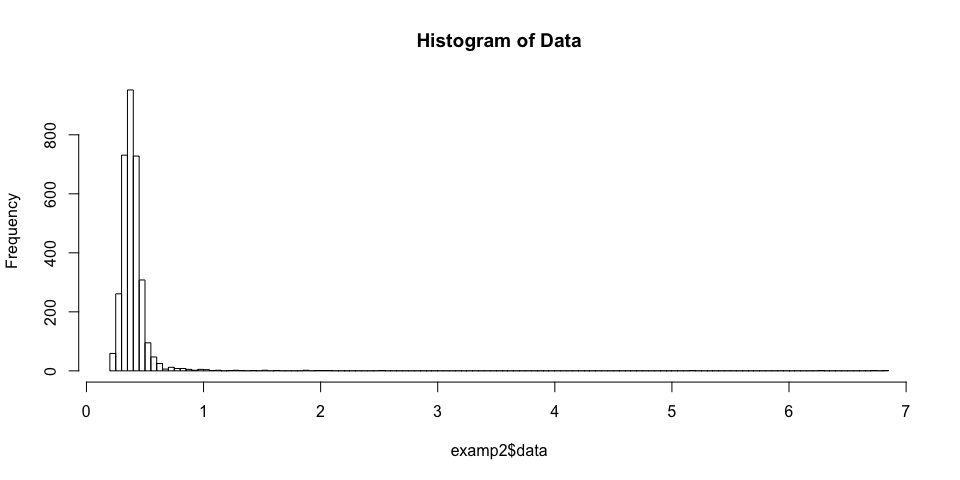


outdataobject1<-outputdata(standindet=standindetobject1,specindet=specindetobject1)  
summaryobject1<-summarytable(outdataobj=outdataobject1, standindetobj=standindetobject1, specindetobj=specindetobject1)  
outtable1<-summaryobject1$outtab  
outtable1

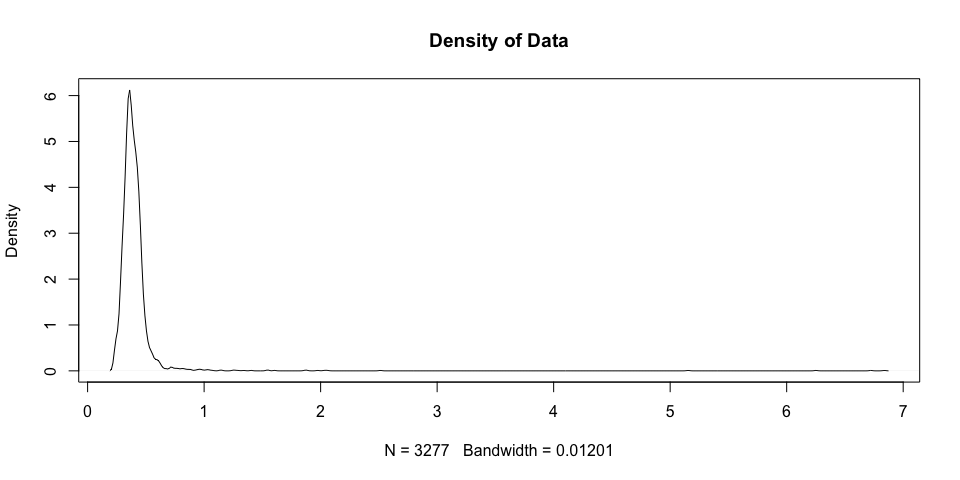
## No. Neg No. Indet No. Pos % Neg % Indet % Pos  
## Raw Cutoff 1086 0 63 94.52% 0% 5.48%  
## Cutoff with 80% Certainty 1025 81 43 89.21% 7.05% 3.74%  
## Cutoff with 90% Certainty 936 178 35 81.46% 15.49% 3.05%  
## Cutoff with 95% Certainty 812 305 32 70.67% 26.54% 2.79%

# Example 2

examp2<-read.csv(file="/Users/sarahsullivan/Documents/SullivanDocs/Rwd/cutoff/ex2.csv")  
  
hist(examp2$data, breaks=100, main="Histogram of Data")



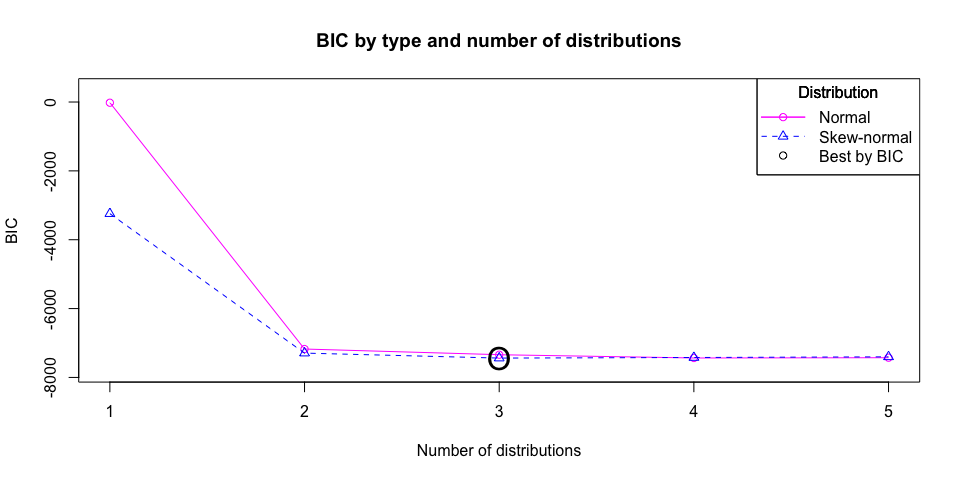
dens2<-density(examp2$data)  
plot(dens2, main="Density of Data")



fitloopsobject2<-fitloops(datawithids=examp2)  
bestfitsobject2<-bestfits(fitres=fitloopsobject2)

## [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): -7340.68612 -7437.852  
## 4 Component(s): -7435.41094 -7421.051  
## 5 Component(s): -7426.65105 -7401.471  
## [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 -7435.411  
## Two Skew-normal Skew-normal with 2 components -7293.374  
## Two Normal Normal with 2 components -7174.868

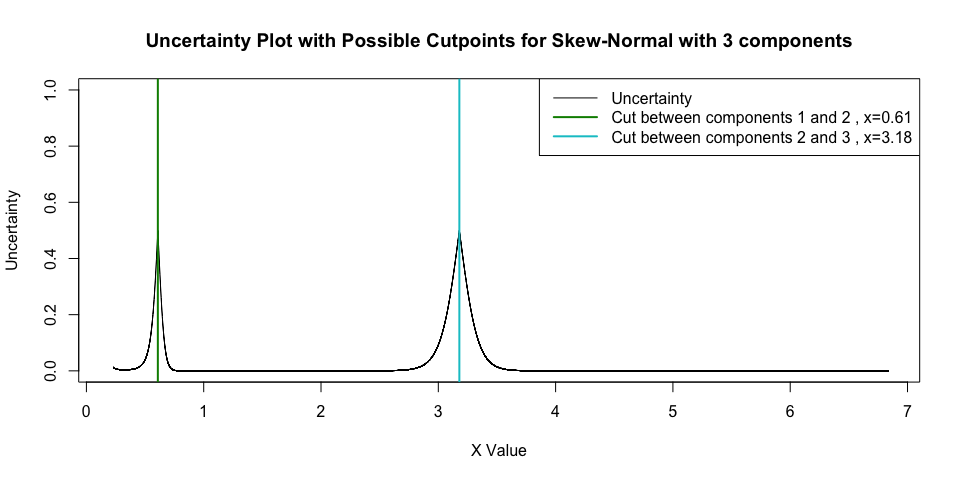
bicgraph(model=bestfitsobject2)



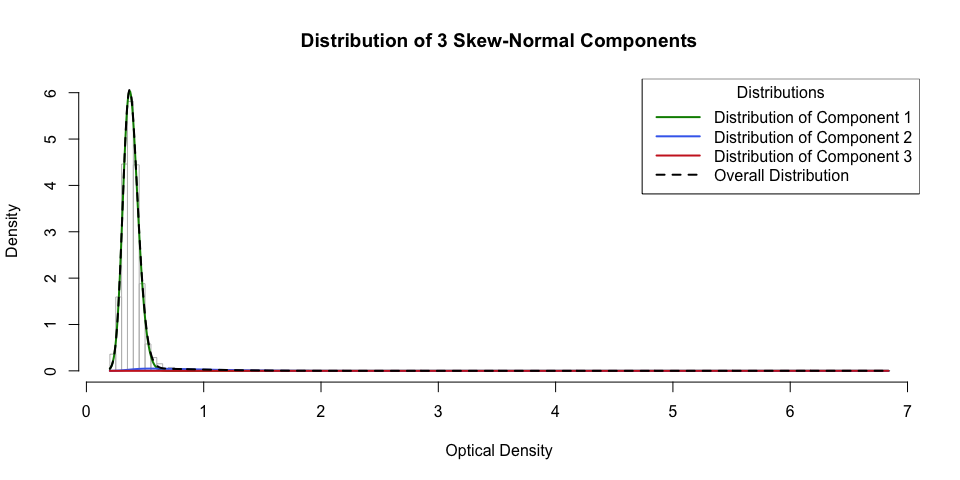
fitpickobject2<-fitpick(fitres=fitloopsobject2,bestfits=bestfitsobject2)  
  
uncertobject2<-uncert(fitpickobj=fitpickobject2)  
multcutobject2<-multcut(uncertobj=uncertobject2)

## 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%

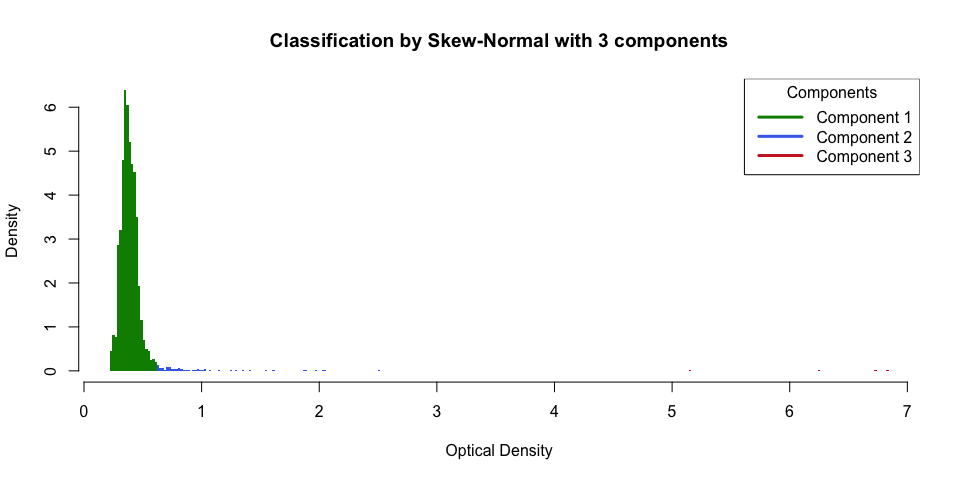
rawuncertgraph(uncertobj=uncertobject2,multcutobj=multcutobject2)



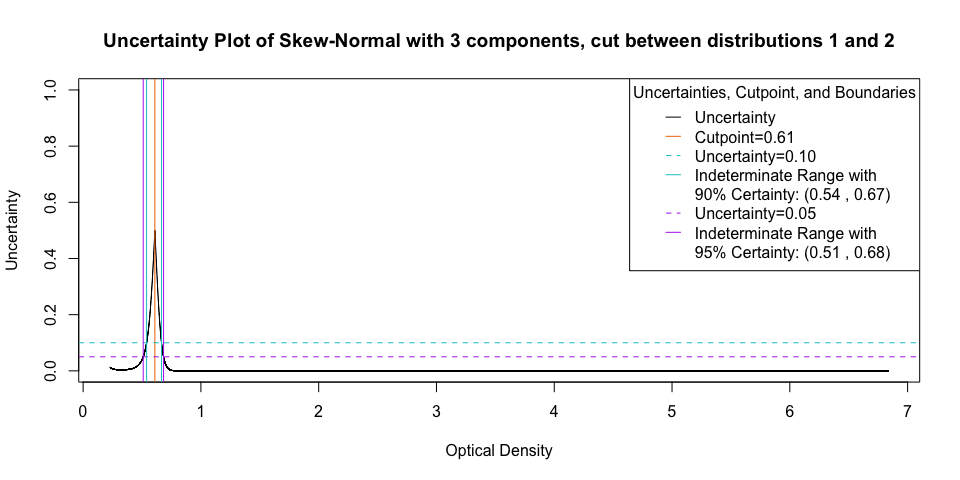
rawdistgraph(pickobj=fitpickobject2)



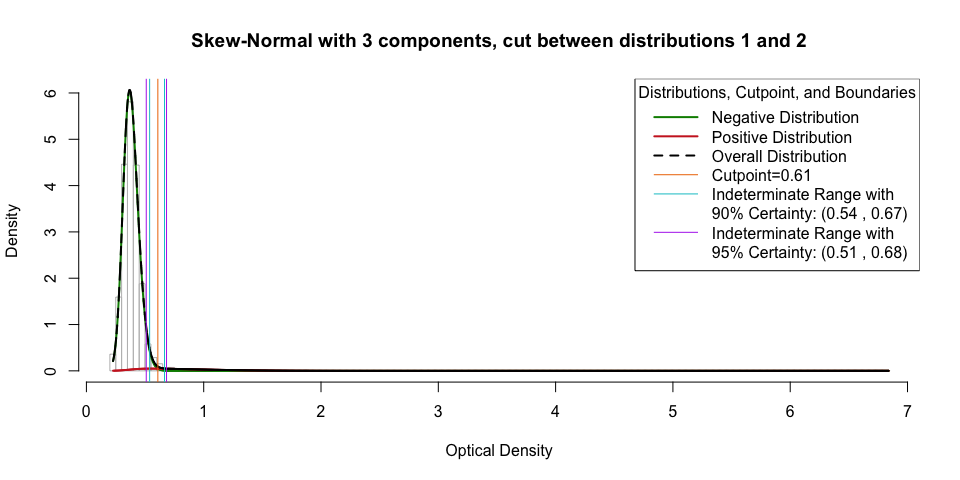
rawhistcuts(uncertobj=uncertobject2,multcutobj=multcutobject2)



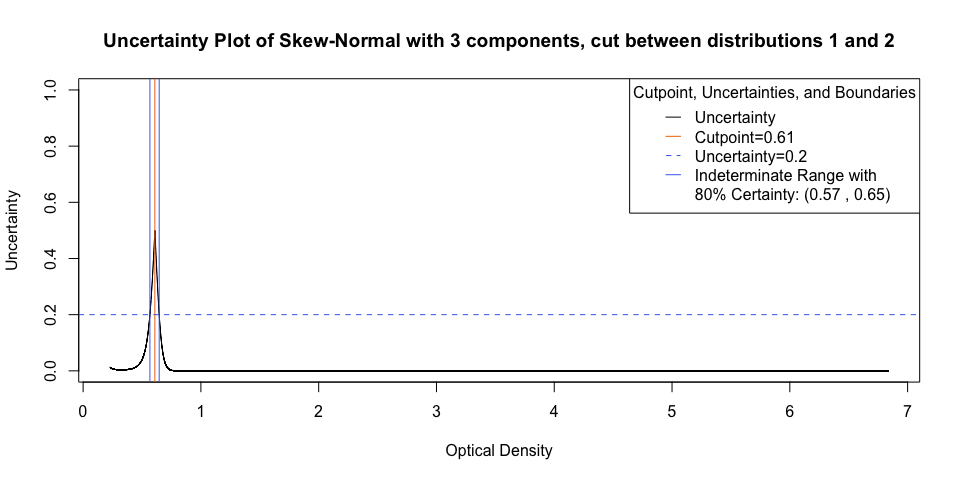
standindetobject2<-standindet(uncertobj=uncertobject2,multcutobj=multcutobject2,cutcomp=1)  
specindetobject2<-specindet(uncertobj=uncertobject2,multcutobj=multcutobject2,cutcomp=1,certlevel=0.8)  
  
cutuncertgraph(cutobj=standindetobject2)



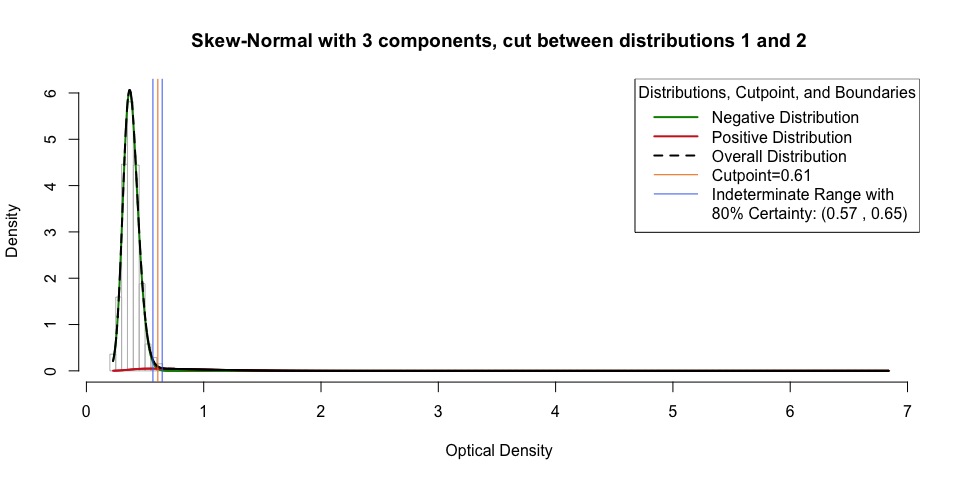
cutdistgraph(cutobj=standindetobject2, pickobj=fitpickobject2)



cutuncertgraph(cutobj=specindetobject2)



cutdistgraph(cutobj=specindetobject2, pickobj=fitpickobject2)



outdataobject2<-outputdata(standindet=standindetobject2,specindet=specindetobject2)  
summaryobject2<-summarytable(outdataobj=outdataobject2, standindetobj=standindetobject2, specindetobj=specindetobject2)  
outtable2<-summaryobject2$outtab  
outtable2

## No. Neg No. Indet No. Pos % Neg % Indet % Pos  
## Raw Cutoff 3868 0 101 97.46% 0% 2.54%  
## Cutoff with 80% Certainty 3835 52 82 96.62% 1.31% 2.07%  
## Cutoff with 90% Certainty 3791 98 80 95.52% 2.47% 2.02%  
## Cutoff with 95% Certainty 3736 158 75 94.13% 3.98% 1.89%