*library(sampling)*

*library(outliers)*

*library(resample)*

*library(sampledata)*

*library(resamplingdata)*

library(vegan)

library(agricolae)

AllOTU=read.table('D:/CAS/Projects/QXN MiSeq 201420152016/16S data/R/OTU\_table QXN v2.csv',header=T, sep=",", row.names = 1)

All.group0=read.table('D:/CAS/Projects/QXN MiSeq 201420152016/16S data/R/Group.csv',header=T, sep=",", row.names = 1)

All.group=All.group0[1:290,]

All0=AllOTU

All0[is.na(All0)]<-0

R=(1: ncol(All0))

TempL<-which(R<=7)

TempD=which(R>7)

dataL=All0[,TempL]

dataD=All0[,TempD]

sort(colSums(dataD))

datacombine<-dataD

**##remove<7000 sequences**

dataLcut=dataL

dataDcut=datacombine

rall<-colSums(dataDcut)

tempall<-which(rall<7000)

length(tempall)

if(length(tempall)!=0) {

dataDcut=dataDcut[,-tempall]

}

ncol(dataDcut)

dataD1=dataDcut

dataD2=dataD1

ncol(dataD1)

##resample number

a1=100

for (i in 1: ncol(dataD1))

{

a2=sum(dataD1[,i])

B1=sample(a2, a1)

B2=sort(B1)

a3=nrow(dataD1)

m1=0

for (j in 1: nrow(dataD1))

{

m2=m1

m1=dataD1[j,i]+m1

B3=which(B2<=m1)

B4=which(B2>m2)

a4=length(B3)+length(B4)-a1

dataD2[j,i]=a4

}

print(i)

}

colSums(dataD2)

datacombine<-dataD1

**##delete ghost species**

dataLcut=dataL

dataDcut=datacombine

rall<-rowSums(dataDcut)

tempall<-which(rall==0)

length(tempall)

if(length(tempall)!=0) {

**dataLcut<-dataLcut[-tempall]**

dataDcut=dataDcut[-tempall,]

}

nrow(dataDcut)

**##delete singleton**

tempall<-which(rall==1)

length(tempall)

if(length(tempall)!=0) {

**dataLcut<-dataLcut[-tempall]**

dataDcut=dataDcut[-tempall,]

}

nrow(dataDcut)

dataDcom1=dataDcut

dataLcom1=dataLcut

datacombine<-dataD2

**##delete ghost species**

dataLcut=dataL

dataDcut=datacombine

rall<-rowSums(dataDcut)

tempall<-which(rall==0)

length(tempall)

if(length(tempall)!=0) {

dataLcut<-dataLcut[-tempall]

dataDcut=dataDcut[-tempall,]

}

nrow(dataDcut)

**##delete singleton**

tempall<-which(rall==1)

length(tempall)

if(length(tempall)!=0) {

dataLcut<-dataLcut[-tempall]

dataDcut=dataDcut[-tempall,]

}

nrow(dataDcut)

dataDcom2=dataDcut

dataLcom2=dataDcut

**##OTU#**

D<-c(rep(0,ncol(dataDcom1)))

for (j in 1:ncol(dataDcom1))

{

C=which(dataDcom1[,j]>0)

D[j]=length(C)

}

D

**##OTU#**

D2<-c(rep(0,ncol(dataDcom2)))

for (j in 1:ncol(dataDcom2))

{

C=which(dataDcom2[,j]>0)

D2[j]=length(C)

}

D2

YLcut1=t(dataLcom1)

Y11= t(dataDcom1)

YLcut2=t(dataLcom2)

Y12= t(dataDcom2)

**##RA**

**Yra=Y11**

B5=rowSums(Yra)

for (i in 1: nrow(Yra))

{

Yra[i,]=(Yra[i,]/B5[i])

}

rowSums(Yra)

**##RA**

**Yra2=Y12**

B6=rowSums(Yra2)

for (i in 1: nrow(Yra2))

{

Yra2[i,]=(Yra2[i,]/B6[i])

}

rowSums(Yra2)

Y2=Y10

Y2=Yra

Y2=Yra2

dcaY2<-decorana(Y2)

summary(dcaY2, display="site")