## Calculation of Average Network Size

The iterative fashion employed in calculating network size can be frustrating. This is because at each step we only exclude one ineligible reference group and update the ratios. The situation is exacerbated if you wish to change the tolerable range of ratio or to compare traditional versus MoS estimates. We provide below some code, which can run in R software. You need an intermediate level of experience with R to understand how the codes work. But even with basic R knowledge, you can run the code and get results. Here we explain how the code works.

###### Data preparation and reading in R

First prepare your data set so that replies to each reference groups are entered in one column. The data set should only contain the replies. Number of rows and columns are equal to the number of respondents and reference groups respectively. In the first step, you should read your data set in R. The easiest way is to use the ‘Rcommander’ package. Another way is to save your data in txt format and run the following line:

data<-read.table("Path/ file name.txt", header=T)

Now we determine the inputs. We should define the total population size (shown by ***t***), the actual size of reference groups (in a vector shown by ***ps0***), the minimum and maximum acceptable ratios (shown by ***rmin*** and ***rmax***), the ratio range (1 for 0.5 to 1.5 or 0.5 to 2, and 2 for abs(log(ratio)<1), and the estimation method (1 for traditional and 2 for MoS). Once these decisions are made, you only need to run the following code in R. It automatically performs the iteration steps and provides the final ***C***. In addition to that, for groups remaining in the last step that contribute to calculations, back calculated sizes and ratios are printed.

t #Total population size

# Real size of each reference group in order to columns of dataset (the vector of dim (1\*number of groups))

ps0

rmin #Min acceptable ratio

rmax #Max acceptable ratio

ratio\_range # 1 for rmin<=ratio<=rmax and 2 for abs(Log ratio)

method # 1 for Traditional method and 2 for MOS

Network\_Size<- function(t,ps0,rmin,rmax,ratio\_range,method){

# Initialize empty vectors and matrices for saving result

remove\_group<-0

data<-as.matrix(data)

sample<-nrow(data)

nu\_groups<-ncol(data)

removing<-rep(NA,nu\_groups)

mm<-rep(NA,nu\_groups)

maximum<-rep(NA,nu\_groups)

minimum<-rep(NA,nu\_groups)

ratio<-matrix(nrow=nu\_groups, ncol=nu\_groups);ratio1<-matrix(nrow=nu\_groups, ncol=nu\_groups)

estimate<-matrix(nrow=nu\_groups, ncol=nu\_groups)

c<- rep(NA,sample) #Network size vector

#Calculation of network size

for (l in 1:nu\_groups){

data[,remove\_group]=NA

ps0[remove\_group]=NA

if(method==1){

for(j in 1:sample) {

c[j]<-t\*sum(data[j,],na.rm=T)/sum(ps0,na.rm=T)

if (c[j]==0){c[j]=1}

}

}

if(method==2){

for(j in 1:sample) {

sps0=0

for(jj in 1:length(ps0)){

if(is.na(data[j,jj])==F){

sps0=sps0+(data[j,jj])/(ps0[jj])

}}

sps0=sps0/(nu\_groups-l+1)

c[j]<-t\*sps0

if (c[j]==0){c[j]=1}

}

}

# Babck calculation

for (i in 1:nu\_groups){

mm[i]<-sum(data[,i])

estimate[i,l]<-t\*mm[i]/sum(c,na.rm=T)

ratio[i,l]<- estimate[i,l]/ps0[i]

if(ratio\_range==1){

ratio1[i,l]<- abs((1-ratio[i,l])\*100)

} else{

ratio1[i,l]<- abs(log(ratio[i,l],2))

}

}

maxratio1<- max(ratio1[,l],na.rm=T)

if(ratio\_range==1){

maximum[l]<- max(ratio[,l],na.rm=T)

minimum[l]<- min(ratio[,l],na.rm=T)

} else{

maximum[l]<- max(ratio1[,l],na.rm=T)

minimum[l]<- min(ratio1[,l],na.rm=T)

}

for (k in 1:nu\_groups){

if(is.na(ratio1[k,l])==F & ratio1[k,l]== maxratio1){

remove\_group=k}

removing[l]=remove\_group}

if (ratio\_range==1 & maximum[l]<=rmax & minimum[l]>=rmin){break}

if(ratio\_range==2 & maximum[l]<=1){break}

}

removing\_group<- removing[-(length(removing[!is.na(removing)]))]

options(digits=3)

# output results

cat('\n',"----------------------------------",'\n')

cat('\n',"Method:",ifelse(ratio\_range==1,"Traditional", "MOS"),'\n')

cat('\n',ifelse(ratio\_range==1,"rmin<=Ratio<=rmax", "Absolote(log(Ratio))<1"),'\n')

if(ratio\_range==1){

cat('\n',"Stop Criteria for traditional ratio\_range:", rmin,"<=Ratio<=",rmax,'\n')

}

cat('\n',"Mean of Network size Estimations:\n\n")

print(mean(c))

cat('\n')

cat('\n',"Size of refrence group estimation:",'\n')

print(estimate[,1:l])

if(ratio\_range==1){

cat('\n',"Ratio:",'\n')

print(ratio[,1:l])

} else {

cat('\n',"Absolote Logarithm Ratio:",'\n')

print(ratio1[,1:l])

}

cat('\n',"Removing Group:",'\n')

print((removing\_group[1:l-1]))

cat('\n',"Sum of Network size Estimations:",'\n')

print(sum(c))

cat('\n',"----------------------------------",'\n')

}

Network\_Size(t,ps0,rmin,rmax,ratio\_range,method)