

# Evaluating causal effects in network-based studies

TingFang Lee

7/27/2021

Load packages and function script...

```
library(lme4)
library(plyr)
library(dplyr)
library(igraph)
library(numDeriv)
library(gtools)
rm(list = ls())

source("https://github.com/uri-ncipher/Nearest-Neighbor-estimators/blob/main/functions.R?raw=TRUE")
```

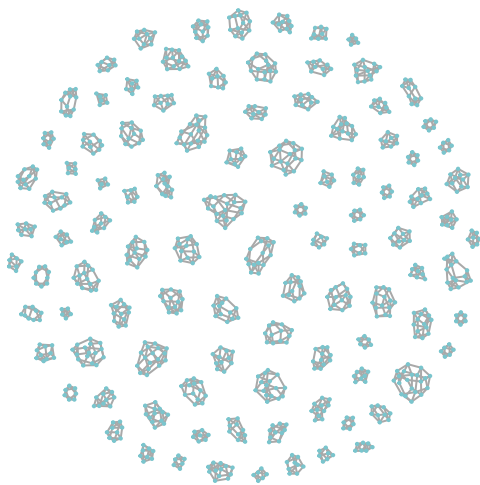
## Data Preparation

Read in the synthetic nodes and edges file for creating the simulated network.

```
nodes=read.csv("https://github.com/uri-ncipher/Nearest-Neighbor-estimators/blob/main/nodes.csv?raw=TRUE")
edges=read.csv("https://github.com/uri-ncipher/Nearest-Neighbor-estimators/blob/main/edges.csv?raw=TRUE")
net0=graph_from_data_frame(d=edges, vertices = nodes, directed=F)
```

Network visuliation

```
plot(net0, vertex.size=1, vertex.label=NA, vertex.color="cadetblue3", vertex.frame.color="cadetblue3")
```



Make the data for modeling

```
n=length(V(net0))
m=components(net0)$no
```

```

data=data.frame(id=1:n, na=unlist(lapply(1:n, num_neighbors, net=net0)),
               component=components(net0)$membership)

#assign treatment, outcome, and baseline covariates to data
data$treatment=nodes$treatment
data$outcome=nodes$outcome
data$var1=nodes$var1
data$var2=nodes$var2
data$na_a= unlist(lapply(1:n, trt_neighbors, net=net0))
data$notna_a=data$na-data$na_a

base_covariate=c("var1", "var2")
#averaging baseline covariates
data$avg_var1=unlist(lapply(1:n, avg_neighbors, net=net0, variable="var1"))
data$avg_var2=unlist(lapply(1:n, avg_neighbors, net=net0, variable="var2"))

avg_covariate=c("avg_var1", "avg_var2")

```

## Modeling

### IPW 1

Using IPW1 to evaluate the average potential outcome and causal effects under allocation strategies  $\alpha$ . The model will output the point estimation and the estimated variance of average potential outcomes  $\hat{Y}(1, \alpha), \hat{Y}(0, \alpha), \hat{Y}(\alpha)$ .

```

alpha=c(0.25, 0.5, 0.75)
IPW_1_model(data, base_covariate, alpha)

## [[1]]
##           a=1          a=0        margin alpha          type
## 1 0.1852175678 0.188372738 0.1875839458 0.25 point estimate
## 2 0.1451336794 0.276759930 0.2109468049 0.50 point estimate
## 3 0.1352103897 0.261785879 0.1668542621 0.75 point estimate
## 4 0.0014465618 0.000919702 0.0005100593 0.25      variance
## 5 0.0003415270 0.001480020 0.0004108478 0.50      variance
## 6 0.0004228426 0.001566447 0.0003414386 0.75      variance
##
## [[2]]
##           estimation alpha0 alpha1      type
## 1 -0.0031551706 0.25 0.25 Direct
## 2 -0.1316262511 0.50 0.50 Direct
## 3 -0.1265754895 0.75 0.75 Direct
## 4  0.0028872409 0.25 0.25 Var DE
## 5  0.0019997039 0.50 0.50 Var DE
## 6  0.0019589605 0.75 0.75 Var DE
## 7 -0.0883871920 0.25 0.50 Indirect
## 8 -0.0734131408 0.25 0.75 Indirect
## 9  0.0149740512 0.50 0.75 Indirect
## 10 0.0004993204 0.25 0.50 Var IE
## 11 0.0018240734 0.25 0.75 Var IE
## 12 0.0012634087 0.50 0.75 Var IE
## 13 -0.0915423626 0.25 0.50 Total

```

```
## 14 -0.0765683114 0.25 0.75 Total
## 15 -0.1166521998 0.50 0.75 Total
## 16 0.0033180932 0.25 0.50 Var IE
## 17 0.0024878018 0.25 0.75 Var IE
## 18 0.0017330800 0.50 0.75 Var IE
## 19 -0.0262624477 0.25 0.50 Overall
## 20 -0.0551034003 0.25 0.75 Overall
## 21 -0.0288409526 0.50 0.75 Overall
## 22 0.0001937051 0.25 0.50 Var IE
## 23 0.0006692608 0.25 0.75 Var IE
## 24 0.0003768236 0.50 0.75 Var IE
```

## IPW 2

Using IPW2 to evaluate the average potential outcome and causal effects under allocation strategies  $\alpha$ . The model will output the point estimation and the estimated variance of average potential outcomes  $\hat{Y}(1, \alpha), \hat{Y}(0, \alpha), \hat{Y}(\alpha)$ .

```
formula_1=paste("cbind(na_a, notna_a)", "~", "treatment", "+",
                paste(base_covariate, collapse = "+"), "+",
                paste(avg_covariate, collapse = "+"))
formula_2=paste("treatment", "~", paste(base_covariate, collapse = "+"))
M1=glm(formula_1, family = binomial(link = "logit"), data=data)
M2=glm(formula_2, family = binomial(link = "logit"), data=data)

IPW_2_model(data, M1, M2, alpha)
```

```
## [[1]]
##           a=1           a=0      margin alpha      type
## 1 0.1796249651 0.1454643863 0.1540045310 0.25 point estimate
## 2 0.1597258863 0.2266322694 0.1931790778 0.50 point estimate
## 3 0.1470394052 0.2526265019 0.1734361794 0.75 point estimate
## 4 0.0011119881 0.0003719574 0.0002913227 0.25      variance
## 5 0.0003355471 0.0004879431 0.0002299846 0.50      variance
## 6 0.0004107754 0.0016837676 0.0003741267 0.75      variance
##
## [[2]]
##      estimation alpha0 alpha1      type
## 1 0.0341605788 0.25 0.25 Direct
## 2 -0.0669063830 0.50 0.50 Direct
## 3 -0.1055870966 0.75 0.75 Direct
## 4 0.0014167592 0.25 0.25 Var DE
## 5 0.0007270418 0.50 0.50 Var DE
## 6 0.0018927829 0.75 0.75 Var DE
## 7 -0.0811678831 0.25 0.50 Indirect
## 8 -0.1071621155 0.25 0.75 Indirect
## 9 -0.0259942325 0.50 0.75 Indirect
## 10 0.0002295839 0.25 0.50 Var IE
## 11 0.0017476652 0.25 0.75 Var IE
## 12 0.0010428090 0.50 0.75 Var IE
## 13 -0.0470073043 0.25 0.50 Total
## 14 -0.0730015368 0.25 0.75 Total
## 15 -0.0929006155 0.50 0.75 Total
## 16 0.0014390786 0.25 0.50 Var IE
```

## 17	0.0020005766	0.25	0.75	Var IE
## 18	0.0015989317	0.50	0.75	Var IE
## 19	-0.0349567019	0.25	0.50	Overall
## 20	-0.0653733126	0.25	0.75	Overall
## 21	-0.0304166108	0.50	0.75	Overall
## 22	0.0001526009	0.25	0.50	Var IE
## 23	0.0005000172	0.25	0.75	Var IE
## 24	0.0002177706	0.50	0.75	Var IE