# Project_Adults

Larissa Xu

2024-11-14

```
library(tableone)
library(Matching)
```

```
## Loading required package: MASS
```

```
## ##
## ##  Matching (Version 4.10-15, Build Date: 2024-10-14)
## ##  See https://www.jsekhon.com for additional documentation.
## ##  Please cite software as:
## ##   Jasjeet S. Sekhon. 2011. ``Multivariate and Propensity Score Matching
## ##   Software with Automated Balance Optimization: The Matching package for R.''
## ##   Journal of Statistical Software, 42(7): 1-52.
## ##
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:MASS':
##
##     select
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(exact2x2)
```

```
## Loading required package: exactci
```

```
## Loading required package: ssanv
```

```
## Loading required package: testthat
```

```
##
## Attaching package: 'testthat'
```

```
## The following object is masked from 'package:dplyr':
##
##     matches
```

```
library(haven)
df<-read.csv("/Users/xuchuning/Desktop/Experimental Design/Project/Merged/OLD(Wave1and4only)/addhealth_merged.csv")
```

```
# rename variables
# H1HS3: have you received psycholgoical or emotional counseling
# H1SU2: During the past 12 months, how many times did you actually attempt suicide?
df <- df %>%
  rename(counselling = h4hs9,
         suicide_attempts = h4se2)
```

```
# Age
# IYEAR4 is the year filled in the questionnaire
# H4OD1Y is the year of birth
# calculate the age of the participants
# Create the age variable in the data frame
df$age <- df$iyear4 - df$h4od1y
```

```
# data management for all categorical covariates
df %>%
  mutate(race_white=ifelse(H1GI6A=="1",1,0),
         race_black_african=ifelse(H1GI6B=="1",1,0),
         race_indian_native=ifelse(H1GI6C=="1",1,0),
         race_asian_pacificislander=ifelse(H1GI6D=="1",1,0),
         race_other=ifelse(H1GI6E=="1",1,0),
         depression=ifelse(h4id5h =="1",1,0), # mental disorders
         anxiety=ifelse(h4id5j=="1",1,0),
         ptsd=ifelse(h4id5i=="1",1,0),
         no_insure=ifelse(h4hs1=="1",1,0), # insurance
         SA_1=ifelse(H1SU2=="1",1,0), # if attempt suicide at Wave 1 or not
         counselling=ifelse(counselling=="1",1,0), # explanatory
         suicide_attempts=ifelse(suicide_attempts %in% c("1","2","3","4"),1,0)) -> df
```

```
# create a new dataset include explanatory, response, covariate
df  %>%
  dplyr::select(race_white, race_black_african, race_indian_native,
                race_asian_pacificislander, race_other,
                depression, anxiety, ptsd, SA_1,
                no_insure, counselling, suicide_attempts)-> mydata
```

```
# All covariates we will be using
xvars<-c("race_white", "race_black_african", "race_indian_native",
         "race_asian_pacificislander", "race_other",
         "depression", "anxiety", "ptsd", "SA_1",
         "no_insure")
```

```
#look at a table 1
table1<- CreateTableOne(vars=xvars,strata="suicide_attempts", data=mydata, test=FALSE)
## include standardized mean difference (SMD)
print(table1,smd=TRUE)
```

```
##                                       Stratified by suicide_attempts
##                                        0           1           SMD
##   n                                    6439          65
##   race_white (mean (SD))               0.66 (0.47) 0.58 (0.50)  0.157
##   race_black_african (mean (SD))       0.25 (0.43) 0.32 (0.47)  0.166
##   race_indian_native (mean (SD))       0.04 (0.19) 0.06 (0.24)  0.118
##   race_asian_pacificislander (mean (SD)) 0.04 (0.20) 0.02 (0.12)  0.159
##   race_other (mean (SD))               0.07 (0.25) 0.08 (0.27)  0.045
##   depression (mean (SD))               0.16 (0.36) 0.48 (0.50)  0.727
##   anxiety (mean (SD))                  0.12 (0.33) 0.37 (0.49)  0.597
##   ptsd (mean (SD))                     0.03 (0.17) 0.17 (0.38)  0.476
##   SA_1 (mean (SD))                     0.02 (0.14) 0.09 (0.29)  0.319
##   no_insure (mean (SD))                0.21 (0.41) 0.42 (0.50)  0.454
```

```
# how diff treatment and ctonrtol are
```

```
# get rid of missing values
mydata <- na.omit(mydata)
```

```
set.seed(123)
# run this before the matching process
# this allows for:
# if I run the matching for 10 times
# people review this doc will see exactly what I see.
# and I report the first set (results are sensitive to matching process)
greedymatch<-Match(Tr=mydata$counselling,M=1,X=mydata[xvars],replace=FALSE)
matched<-mydata[unlist(greedymatch[c("index.treated","index.control")]), ]

#get table 1 for matched data with standardized differences
matchedtab1<-CreateTableOne(vars=xvars, strata ="counselling",
                            data=matched, test = FALSE)
print(matchedtab1, smd = TRUE)
```

```
##                                        Stratified by counselling
##                                         0           1           SMD
##   n                                      526         526
##   race_white (mean (SD))                0.76 (0.42) 0.77 (0.42)  0.009
##   race_black_african (mean (SD))        0.17 (0.37) 0.17 (0.38)  0.020
##   race_indian_native (mean (SD))        0.02 (0.16) 0.02 (0.16) <0.001
##   race_asian_pacificislander (mean (SD)) 0.02 (0.16) 0.03 (0.16)  0.012
##   race_other (mean (SD))                0.06 (0.24) 0.06 (0.25)  0.008
##   depression (mean (SD))                0.50 (0.50) 0.54 (0.50)  0.088
##   anxiety (mean (SD))                   0.40 (0.49) 0.43 (0.50)  0.058
##   ptsd (mean (SD))                      0.14 (0.35) 0.14 (0.35) <0.001
##   SA_1 (mean (SD))                      0.03 (0.17) 0.03 (0.17) <0.001
##   no_insure (mean (SD))                 0.16 (0.37) 0.16 (0.37)  0.005
```

```
# disregard people who have not received counselling in the past year that
# are very different than the people who did receive counselling.
# till this point, the main diff is "if a person receive counselling" or not
# between two groups
# we of course exclude the suicide attempts (response var)

# Outcome analysis
y_trt<-matched$suicide_attempts[matched$counselling==1]
# pull out response variables for everyone in treatment
y_con<-matched$suicide_attempts[matched$counselling==0]
# pull out the response for everyone in control
# the first person in treatment is matched with the first person in control

# Mcnemar test
mcnemar.exact(table(y_con,y_trt))
```

```
##
##  Exact McNemar test (with central confidence intervals)
##
## data:  table(y_con, y_trt)
## b = 16, c = 9, p-value = 0.2295
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.7397548 4.5643094
## sample estimates:
## odds ratio
##   1.777778
```

The Mcnemar test reveals that receiving psychological or emotional counseling in the past 12 months does not have a significant causal relationship with whether someone attempt suicide over the past year (p = .2295).