

MATCHING ANALYSIS OF THE NHANES DATA

The variables are:

- BPSysAve (systolic blood pressure, response),
- SmokeNow (smoking status: 0–No, 1–Yes),
- Gender, Age, Race3, Education, MaritalStatus, and Poverty

where the first two are the outcome and exposure of interest and the remaining are potential confounders. We restrict our attention to 1377 adults (> 17 years old) in the second wave of the survey.

```
suppressPackageStartupMessages({
library(NHANES); library(tableone); library(Matching)      #Must be pre-loaded in R
library(survey); library(Hmisc); library(htmlwidgets)      #Must be pre-loaded in R
})
NHANES$SmokeNow <- as.numeric(NHANES$SmokeNow)-1
small.nhanes <- na.omit(NHANES[NHANES$SurveyYr=="2011_12"
                             & NHANES$Age > 17,c(3,4,8:11,13,25,61)])
vars <- c("Gender", "Age", "Race3", "Education", "MaritalStatus", "Poverty")
```

```
tabUnmatched<-CreateTableOne(vars=vars,strata="SmokeNow",data=small.nhanes,test=FALSE)
print(tabUnmatched, smd = TRUE)
```

Stratified by SmokeNow				
	0	1		SMD
n	782	595		
Gender = male (%)	432 (55.2)	369 (62.0)		0.138
Age (mean (SD))	54.33 (16.52)	44.96 (15.11)		0.592
Race3 (%)				0.315
Asian	25 (3.2)	15 (2.5)		
Black	43 (5.5)	64 (10.8)		
Hispanic	26 (3.3)	38 (6.4)		
Mexican	45 (5.8)	35 (5.9)		
White	630 (80.6)	416 (69.9)		
Other	13 (1.7)	27 (4.5)		
Education (%)				0.512
8th Grade	59 (7.5)	33 (5.5)		
9 - 11th Grade	71 (9.1)	120 (20.2)		
High School	152 (19.4)	151 (25.4)		
Some College	256 (32.7)	210 (35.3)		
College Grad	244 (31.2)	81 (13.6)		
MaritalStatus (%)				0.488
Divorced	85 (10.9)	77 (12.9)		
LivePartner	61 (7.8)	96 (16.1)		
Married	453 (57.9)	240 (40.3)		
NeverMarried	108 (13.8)	142 (23.9)		
Separated	6 (0.8)	14 (2.4)		
Widowed	69 (8.8)	26 (4.4)		
Poverty (mean (SD))	3.11 (1.65)	2.38 (1.58)		0.453

```
Smoke<-small.nhanes$SmokeNow
```

Propensity score matching:

```
#Build the PS model
fmods<-formula(SmokeNow~Gender+Age+Race3+Education+MaritalStatus+HHIncome+Poverty)
ps.mod <- glm(fmods,data=small.nhanes,family="binomial")
ps.lr <- predict(ps.mod,type="response")
small.nhanes$ps.lr<-ps.lr

#Matching for the ATE
set.seed(1234)
ps.lr.match <- Match(Tr=small.nhanes$SmokeNow,X=small.nhanes$ps.lr,
                    estimand="ATE",ties=FALSE)
summary(ps.lr.match)

+
+ Estimate... 0
+ SE..... 0
+ T-stat..... NaN
+ p.val..... NA
+
+ Original number of observations..... 1377
+ Original number of treated obs..... 595
+ Matched number of observations..... 1377
+ Matched number of observations (unweighted). 1377

match.details<-cbind(ps.lr.match$index.control,
                    small.nhanes$SmokeNow[ps.lr.match$index.control],
                    ps.lr.match$index.treated,small.nhanes$SmokeNow[ps.lr.match$index.treated],
                    small.nhanes$ps.lr[ps.lr.match$index.control],
                    small.nhanes$ps.lr[ps.lr.match$index.treated])
cols<-c('Control ID','Control Z','Treated ID','Treated Z','Control PS','Treated PS')
colnames(match.details)<-cols

match.details[1:18,]

+      Control ID Control Z Treated ID Treated Z Control PS Treated PS
+ 1310      1310      0      1      1 0.5777708 0.5778596
+ 2      2      0      165      1 0.2830976 0.2863810
+ 3      3      0      165      1 0.2830976 0.2863810
+ 4      4      0      165      1 0.2830976 0.2863810
+ 5      5      0      165      1 0.2830976 0.2863810
+ 6      6      0      88      1 0.3224138 0.3227360
+ 345      345      0      7      1 0.3817483 0.3811699
+ 8      8      0      766      1 0.6064829 0.6075296
+ 9      9      0      676      1 0.2068104 0.2022726
+ 10     10      0      675      1 0.2068104 0.2022726
+ 11     11      0      398      1 0.1900244 0.1847449
+ 12     12      0      118      1 0.5645810 0.5638680
+ 13     13      0      1286      1 0.4613472 0.4615596
+ 14     14      0      602      1 0.5189636 0.5179700
+ 15     15      0      602      1 0.5189636 0.5179700
+ 946     946      0      16      1 0.5704613 0.5708199
+ 946     946      0      17      1 0.5704613 0.5708199
+ 946     946      0      18      1 0.5704613 0.5708199
```

```

matched.samp <- small.nhanes[c(ps.lr.match$index.control,ps.lr.match$index.treated),]

table(table(c(ps.lr.match$index.control, ps.lr.match$index.treated)))

+
+   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  23  29  38
+ 846 252 128  60  27  20   9   6   1   4   5   3   4   4   4   1   1   1   1

tabMatched<-CreateTableOne(vars = vars, strata = "SmokeNow",data = matched.samp,
                           test = FALSE)

print(tabMatched)

+
+               Stratified by SmokeNow
+               0               1
+   n               1377               1377
+   Gender = male (%)   760 (55.2)   753 (54.7)
+   Age (mean (SD))    50.25 (16.85) 50.31 (16.16)
+   Race3 (%)
+     Asian              35 ( 2.5)   44 ( 3.2)
+     Black              114 ( 8.3)  119 ( 8.6)
+     Hispanic            81 ( 5.9)   68 ( 4.9)
+     Mexican             90 ( 6.5)  125 ( 9.1)
+     White              1016 (73.8)  979 (71.1)
+     Other               41 ( 3.0)   42 ( 3.1)
+   Education (%)
+     8th Grade           88 ( 6.4)   123 ( 8.9)
+     9 - 11th Grade      198 (14.4)  194 (14.1)
+     High School         296 (21.5)  316 (22.9)
+     Some College        458 (33.3)  472 (34.3)
+     College Grad       337 (24.5)  272 (19.8)
+   MaritalStatus (%)
+     Divorced            180 (13.1)  168 (12.2)
+     LivePartner         144 (10.5)  143 (10.4)
+     Married             702 (51.0)  690 (50.1)
+     NeverMarried        225 (16.3)  228 (16.6)
+     Separated           33 ( 2.4)   19 ( 1.4)
+     Widowed             93 ( 6.8)   129 ( 9.4)
+   Poverty (mean (SD))  2.74 (1.63)  2.65 (1.62)

#####
#Matching for the ATT
set.seed(1234)
ps.lr.match <- Match(Tr=small.nhanes$SmokeNow,X=small.nhanes$ps.lr,
                    estimand="ATT",ties=FALSE)

summary(ps.lr.match)

+
+ Estimate... 0
+ SE..... 0
+ T-stat..... NaN
+ p.val..... NA
+
+ Original number of observations..... 1377
+ Original number of treated obs..... 595
+ Matched number of observations..... 595
+ Matched number of observations (unweighted). 595

```

```

match.details<-cbind(ps.lr.match$index.control,
  small.nhanes$SmokeNow[ps.lr.match$index.control],
  ps.lr.match$index.treated,small.nhanes$SmokeNow[ps.lr.match$index.treated],
  small.nhanes$ps.lr[ps.lr.match$index.control],
  small.nhanes$ps.lr[ps.lr.match$index.treated])
colnames(match.details)<-cols

match.details[1:18,]

+      Control ID Control Z Treated ID Treated Z Control PS Treated PS
+ 1310      1310      0      1      1 0.5777708 0.5778596
+ 304      304      0      7      1 0.3812127 0.3811699
+ 946      946      0     16      1 0.5704613 0.5708199
+ 946      946      0     17      1 0.5704613 0.5708199
+ 946      946      0     18      1 0.5704613 0.5708199
+ 413      413      0     22      1 0.6185314 0.6183710
+ 841      841      0     28      1 0.1783055 0.1786832
+ 841      841      0     29      1 0.1783055 0.1786832
+ 373      373      0     30      1 0.1785356 0.1786832
+ 843      843      0     31      1 0.1783055 0.1786832
+ 573      573      0     36      1 0.3156519 0.3147168
+ 1097     1097      0     39      1 0.1402311 0.1404715
+ 1098     1098      0     40      1 0.1402311 0.1404715
+ 305      305      0     41      1 0.6029375 0.6024898
+ 139      139      0     42      1 0.1419909 0.1413732
+ 683      683      0     43      1 0.2717209 0.2725138
+ 683      683      0     44      1 0.2717209 0.2725138
+ 683      683      0     45      1 0.2717209 0.2725138

matched.samp <- small.nhanes[c(ps.lr.match$index.control,ps.lr.match$index.treated),]

table(table(c(ps.lr.match$index.control, ps.lr.match$index.treated)))

+
+ 1 2 3 4 5 6 7 8 9 11 12 13 14 15
+ 735 50 28 17 11 7 1 2 2 1 1 1 1 1

tabMatched<-CreateTableOne(vars = vars, strata = "SmokeNow",data = matched.samp,
  test = FALSE)
print(tabMatched)

+
+      Stratified by SmokeNow
+      0      1
+ n      595      595
+ Gender = male (%)      330 (55.5)      369 (62.0)
+ Age (mean (SD))      45.52 (15.97) 44.96 (15.11)
+ Race3 (%)
+ Asian      7 ( 1.2)      15 ( 2.5)
+ Black      75 (12.6)      64 (10.8)
+ Hispanic      55 ( 9.2)      38 ( 6.4)
+ Mexican      40 ( 6.7)      35 ( 5.9)
+ White      388 (65.2)      416 (69.9)
+ Other      30 ( 5.0)      27 ( 4.5)
+ Education (%)
+ 8th Grade      28 ( 4.7)      33 ( 5.5)

```

+	9 - 11th Grade	130 (21.8)	120 (20.2)
+	High School	143 (24.0)	151 (25.4)
+	Some College	206 (34.6)	210 (35.3)
+	College Grad	88 (14.8)	81 (13.6)
+	MaritalStatus (%)		
+	Divorced	97 (16.3)	77 (12.9)
+	LivePartner	74 (12.4)	96 (16.1)
+	Married	262 (44.0)	240 (40.3)
+	NeverMarried	115 (19.3)	142 (23.9)
+	Separated	26 (4.4)	14 (2.4)
+	Widowed	21 (3.5)	26 (4.4)
+	Poverty (mean (SD))	2.24 (1.48)	2.38 (1.58)

Estimation for the ATE:

```
ps.lr.match.ate <- Match(Y=small.nhanes$BPSysAve,Tr=small.nhanes$SmokeNow,
                        X=small.nhanes$ps.lr,estimand="ATE",ties=FALSE)
summary(ps.lr.match.ate)

+
+ Estimate... -0.93319
+ SE..... 0.62896
+ T-stat..... -1.4837
+ p.val..... 0.13789
+
+ Original number of observations..... 1377
+ Original number of treated obs..... 595
+ Matched number of observations..... 1377
+ Matched number of observations (unweighted). 1377
```

Estimation for the ATT:

```
ps.lr.match.att <- Match(Y=small.nhanes$BPSysAve,Tr=small.nhanes$SmokeNow,
                        X=small.nhanes$ps.lr,estimand="ATT",ties=FALSE)
summary(ps.lr.match.att)

+
+ Estimate... 0.25714
+ SE..... 0.9274
+ T-stat..... 0.27727
+ p.val..... 0.78157
+
+ Original number of observations..... 1377
+ Original number of treated obs..... 595
+ Matched number of observations..... 595
+ Matched number of observations (unweighted). 595
```