Problem Set #4 - DATASCI W241

Greg Ceccarelli

November 15, 2015

FE exercise 5.2.

```
##clear global env
rm( list = ls() )

library(dplyr)
library(stargazer)
library(memisc)
library(data.table)

set.seed(123)

## Hypothetical Schedule
e_52_hyp_schedule <- data.table(D_Z_1=c(1,0,1,0,1,0),
                                Y_D_1=c(8,10,8,10,10,8),
                                Y_D_0=c(9,8,9,7,10,2))

## ATE - Positive in my test example
ATE <- e_52_hyp_schedule %>%
  summarise(t=sum(Y_D_1-Y_D_0)/n())
ATE #print positive ATE
```

```
##      t
## 1: 1.5
```

```
##CACE - Negative in my test example
CACE<- e_52_hyp_schedule %>%
  group_by(D_Z_1) %>%
  summarise(t=sum(Y_D_1-Y_D_0)/n())

as.data.frame(CACE)[1,2] #print negative CACE
```

```
## [1] -0.6666667
```

In what ways would the estimated CACE be informative or misleading?

[To Answer]

Which population is more relevant to study for future decisionmaking: the set of Compliers, or the set of Compliers plus Never-Takers? Why?

It really depends on what you're trying to estimate, but since we have negative CACE and we're probably trying to understand the effect of an intervention (that we presumably hypothesize has a negative effect), if we used ATE, we'd underestimate the effect (think it was positive when really its negative) if we considered Compliers + Never takers. Thus it probably makes sense to focus on Compliers (those who receive the treatment), in this example

FE exercise 5.6.

```
##clear global env
##rm( list = ls() )

## ITT_D What proportion of those ASSIGNED were actually treated?

Assigned = 1000
Treatment_Lie = 500
Treatment_Real = 250
Treatment_Voted = 400
Control = 2000
Control_Voted = 700
```

a. If you believe that 500 subjects were actually contacted, what would your esimate of CACE be?

```
#What proportion of those ASSIGNED were actually treated?
ITT_d = Treatment_Lie/Assigned
#Intent to Treat Effect
ITT = (Treatment_Voted/Assigned) - (Control_Voted/Control)
CACE = ITT/ITT_d
CACE
```

```
## [1] 0.1
```

b. Suppose you learned that only 250 subjects were actually treated. What would your estimate of CACE be?

```
#What proportion of those ASSIGNED were actually treated?
ITT_d = Treatment_Real/Assigned
#Intent to Treat Effect
ITT = (Treatment_Voted/Assigned) - (Control_Voted/Control)
CACE = ITT/ITT_d
CACE
```

```
## [1] 0.2
```

c. Do the canvassers' exaggerated reports make their efforts seem more or less effective? In terms of Complier's average causal effect (CACE), the exaggerated report of 500 contacts actually makes the CACE smaller. This is because the CACE ratio increases as ITT_d decreases.

FE exercise 5.10.

```
#setwd
library(foreign)
#setwd("/Users/ceccarelli/MIDS/DATASCI_W241/Assignments/Problem Set 4/")
d5.10 <- read.dta("Guan_Green_CPS_2006.dta.dta")

summary(d5.10)
```

```
##     turnout          contact          dormid             treat2
##  Min.   :0.0000   Min.   :0.0000   Min.   : 1010101   Min.   :0.0000
##  1st Qu.:1.0000   1st Qu.:0.0000   1st Qu.: 6010146   1st Qu.:0.0000
##  Median :1.0000   Median :1.0000   Median : 9020212   Median :1.0000
##  Mean   :0.7568   Mean   :0.5922   Mean   :10226152   Mean   :0.6685
##  3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:14010136   3rd Qu.:1.0000
##  Max.   :1.0000   Max.   :1.0000   Max.   :24033068   Max.   :1.0000
##  NA's   :2
```

a. Using the dataset, estimate ITT

```
itt <- as.vector(d5.10 %>%
                 na.omit %>%
  group_by(treat2) %>%
  summarise(mean(turnout)))

#ITT
as.data.frame(itt)[2,2] - as.data.frame(itt)[1,2]
```

```
## [1] 0.1319296
```

b. Use regression to test the sharp null that ITT is zero for all observations, taking into account the fact that random assignment was clustered by dorm room. Interpret results.

```
#initialize clustered standard errors function
cl <- function(fm, cluster){
  require(sandwich, quietly = TRUE)
  require(lmtest, quietly = TRUE)
    M <- length(unique(cluster))
    N <- length(cluster)
    K <- fm$rank
    dfc <- (M/(M-1))*((N-1)/(N-K))
    uj <- apply(estfun(fm),2, function(x) tapply(x, cluster, sum));
    vcovCL <- dfc*sandwich(fm, meat=crossprod(uj)/N)
    coeftest(fm, vcovCL)
}

#make sure to set cluster variable as a factor
d5.10$cluster <- factor(as.character(d5.10$dormid))

#remove NAs
d5.10.nona <- d5.10 %>%
                 na.omit
#run standard errors
c.se <- cl(lm(turnout ~ treat2,d5.10.nona), d5.10.nona$cluster)

c.se.ate <- c.se[2,1]
c.se.se <- c.se[2,2]

c.se.ci <- c(c.se.ate -1.96 *c.se.se, c.se.ate+1.96*c.se.se)
```

The p-value is highly statistically significant and I'd be quite confident rejecting the sharpe null that the difference between each observation is zero. Also, a look at the 95% confidence interval (that doens't cross zero), would indicate that there is a good chance of an intent to treat effect.

c. Assume that the leaflet had no effect on turnout. Estimate the CACE.

```
library("AER")
cace_fit <- ivreg(turnout ~ treat2,~contact, data = d5.10)
mod <- summary(cace_fit)

#CACE
cat(" CACE = ", mod$coefficients[2,1])
```

```
##  CACE =  0.1430509
```

d. Assume the leaflet raised the probability of voting by 1 percentage point among the compliers and never takers.

```
Treatment = 2380
Assigned = 2688
Treatment_Voted = 2152
Control_Voted = 892
Control = 1334

ITT_d = Treatment/Assigned
#Intent to Treat Effect
ITT = ((Treatment_Voted/Assigned)+.01) - ((Control_Voted/Control))

CACE = ITT/ITT_d
CACE
```

```
## [1] 0.1602969
```

E. Given this assumption, estimate CACE of canvassing

```
CACE = ITT/ITT_d
CACE
```

```
## [1] 0.1602969
```

f. Raised turnout among never takers by 3%, given this assumption, estimate CACE of canvassing

Exercise 5.11

a. Estimate the proportion of Compliers by using the data on the Treatment group. Then compute a second estimate of the proportion of Compliers by using the data on the Placebo group. Are these sample proportions statistically significantly different from each other? Explain why you would not expect them to be different, given the experimental design.

b. Do the data suggest that Never-Takers in the treatment and placebo groups havet the same rate of turnout?

c. Estimate CACE of receiving the placebo. Is this estimate consistent with the assumption that the placebo has no effect on turnout?

d. Estimate the CACE of receiving treatments using two different methods.

Question 5

a. In the advertising example of Lewis and Reiley (2014), assume some treatment-group members are friends with control-group members.

b. Consider the police displacement example from the bulleted list in the introduction to FE 8, where we are estimating the effects of enforcement on crime.

In the case where an untreated location was adjacent to a treated neighbor, there could be an "artifical" surge in crime resulting in an over-estimate of the ate between the two locations of the police intervention effect.

c. Suppose employees work harder when you experimentally give them compensation that is more generous than they expected, that people feel resentful (and therefore work less hard) when they learn that their compensation is less than others, and that some treatment-group members talk to control group members.

In this communication interference example, the causal effect of the direct treatment to the employees receiving the compensation increase may be substantial, but given the interference may be "canceled" out by resentful employees working less hard. Thus the effect would be underestimated

d. When Olken (2007) randomly audits local Indonesian governments for evidence of corruption, suppose control-group governments learn that treatment-group governments are being randomly audited and assume they are likely to get audited too.

In this deterrence/communication interference example, if treatment and control are compared, the effect of the audit may be underestimated because the control groups would see an equal decrease in corruption with the expectation that they will be audited

Exercise 8.2

National Surveys indicate that college roommates tend to have correlated weights. The more one roommate weighs at the end of the freshman year, the more the other freshman roomate weights. On the other hand, researchers studying housing arrangements in which roommates are randomly paired together find no correlation between two roommates' weights at the end of their freshman year. Explain how these two facts can be reconciled.

Exercise 8.10

a. Suppose you were seeking to estimate the ATE of running on Tetris. Explain the assumptions needed.

b. Estimate the effect of running on Tetris score. Use randomization inference to test the sharp null hypothesis

```r
library(foreign)

d8.10 <- read.dta("Hough_WorkingPaper_2010.dta.dta")

summary(d8.10)
```

```
##       day             run              weight          tetris
##  Min.   : 1.00   Min.   :0.0000   Min.   :18.00   Min.   : 3939
##  1st Qu.: 7.25   1st Qu.:0.0000   1st Qu.:19.25   1st Qu.:11700
##  Median :13.50   Median :1.0000   Median :20.00   Median :16982
##  Mean   :13.50   Mean   :0.5385   Mean   :20.00   Mean   :20747
##  3rd Qu.:19.75   3rd Qu.:1.0000   3rd Qu.:21.00   3rd Qu.:24008
##  Max.   :26.00   Max.   :1.0000   Max.   :22.00   Max.   :56047
##                                                   NA's   :2
##       mood           energy         appetite          gre
##  Min.   :1.000   Min.   :1.000   Min.   :0.000   Min.   :0.00
##  1st Qu.:3.000   1st Qu.:2.000   1st Qu.:0.000   1st Qu.:0.00
##  Median :3.000   Median :3.000   Median :0.000   Median :1.00
##  Mean   :3.167   Mean   :3.042   Mean   :0.625   Mean   :0.72
##  3rd Qu.:4.000   3rd Qu.:4.000   3rd Qu.:1.000   3rd Qu.:1.00
##  Max.   :4.000   Max.   :5.000   Max.   :3.000   Max.   :1.00
##  NA's   :2       NA's   :2       NA's   :2       NA's   :1
```

```r
#remove NAs
d8.10.nona <- d8.10 %>%
                 na.omit

d8.10.nona %>%
  group_by(run) %>%
  summarise(mean(tetris))
```

```
## Source: local data frame [2 x 2]
##
##     run mean(tetris)
##   (int)        (dbl)
## 1     0      12806.4
## 2     1      26419.5
```

```r
mod <- summary(lm(tetris ~ run,d8.10.nona))
mod
```

```
##
## Call:
## lm(formula = tetris ~ run, data = d8.10.nona)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -19294  -6707  -1154   4890  29628
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    12806       3708   3.453  0.00226 **
## run            13613       4856   2.804  0.01035 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11730 on 22 degrees of freedom
## Multiple R-squared:  0.2632, Adjusted R-squared:  0.2297
## F-statistic:  7.86 on 1 and 22 DF,  p-value: 0.01035
```

```r
ate <- mod$coefficients[2,1]


##  grab the standard error
se <- mod$coefficients[2,2]

ci <- c(ate-1.96 *se, ate+1.96*se)

cat(" CACE = ", mod$coefficients[2,1])
```

```
##  CACE =  13613.1
```

c. One way to lend credibility to within-subjects results is to verify the no-anticipation assumption. Use the variable Run to predict the tetris score on the preceding day.

```r
d8.10.nona_lag <- d8.10.nona

#created lagged variable
d8.10.nona_lag$tetris_lag <- lag(d8.10.nona$tetris)

#remove nas from dataset
d8.10.nona_lag<-d8.10.nona_lag %>%
  na.omit

#compare means
d8.10.nona %>%
  group_by(run) %>%
  summarise(mean(tetris))
```

```
## Source: local data frame [2 x 2]
##
##     run mean(tetris)
##   (int)        (dbl)
## 1     0      12806.4
## 2     1      26419.5
```

```r
d8.10.nona_lag %>%
  na.omit %>%
  group_by(run) %>%
  summarise(mean(tetris_lag))
```

```
## Source: local data frame [2 x 2]
##
##     run mean(tetris_lag)
##   (int)            (dbl)
## 1     0         18957.80
## 2     1         19408.62
```

```r
#predict effect
mod <- summary(lm(tetris_lag ~ run,d8.10.nona_lag))
mod
```

```
##
## Call:
## lm(formula = tetris_lag ~ run, data = d8.10.nona_lag)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -15470  -7478  -2842   2242  29360
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  18957.8     3655.0   5.187 3.86e-05 ***
## run            450.8     4861.6   0.093    0.927
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11560 on 21 degrees of freedom
## Multiple R-squared:  0.0004093,  Adjusted R-squared:  -0.04719
## F-statistic: 0.008599 on 1 and 21 DF,  p-value: 0.927
```

The effect is very small and seems to confirm the no-anticipation assumption

d. If Tetris responds to exercise, one might suppose that energy levels and GRE scores would as well. Are these hypothesis borne out by the data?

e. Given that, would you expect randomization inference to give you a better answer than the regression answer you just obtained in (b)? Which number(s) do you expect to be different in regression than in randomization inference? What is the direction of the bias?