Computer Assignment 5B Costello

Rebecca Costello

9/30/2019

#Rebecca Costello, Ties van der Veen and Niels van Opstal

```
theUrl_ca5b <- "https://surfdrive.surf.nl/files/index.php/s/76yJI8Orr6J9eYN/download"
crime <- read.dta (file = theUrl_ca5b)
```

1. Mechanism

(a) Define the two potential outcomes for a particular city, Y(0,i) and Y(1,i). The amount of sexual abuse and rape stays the same The amount of sexual abuse and rape changes

(b) Provide one or more reasons why tippel zones may affect crime.

Regulating prostitution allows for more control and safety measures. This protects the sex workers there, reducing crime and any opportunity for shady business. Everything happens within the law, which also means the people working there will be less afraid of police/law enforces coming around to check stuff out.

```
crime$citytype <- 3

crime$citytype[crime$city1 == 1] <- 1
crime$citytype[crime$city1 == 2] <- 1
crime$citytype[crime$city1 == 3] <- 1

crime$citytype[crime$city1 == 4] <-2
crime$citytype[crime$city1 == 21] <-2
crime$citytype[crime$city1 == 12] <-2
crime$citytype[crime$city1 == 14] <-2
crime$citytype[crime$city1 == 10] <-2
crime$citytype[crime$city1 == 6] <-2
```

mean and standard distribution of the outcome variables: sexual abuse and rape (rapesexapcN) #mean and standard deviation of the outcome variable

```
crime %>%
  filter(crime$citytype==1) %>%
  summarise(mean=mean(rapesexapcN), sd=sd(rapesexapcN))
```

mean	sd
1.81	0.461

```
crime %>%
  filter(crime$citytype==2) %>%
  summarise(mean=mean(rapesexapcN), sd=sd(rapesexapcN))
```

mean	sd
1.47	0.593

```
crime %>%
  filter(crime$citytype==3) %>%
  summarise(mean=mean(rapesexapcN), sd=sd(rapesexapcN))
```

mean	sd
1.36	0.727

sexual abuse (sexassaultpcN)

```
crime %>%
  filter(crime$citytype==1) %>%
  summarise(mean=mean(sexassaultpcN), sd=sd(sexassaultpcN))
```

mean	sd
0.775	0.242

```
crime %>%
  filter(crime$citytype==2) %>%
  summarise(mean=mean(sexassaultpcN), sd=sd(sexassaultpcN))
```

mean	sd
0.626	0.33

```
crime %>%
  filter(crime$citytype==3) %>%
  summarise(mean=mean(sexassaultpcN), sd=sd(sexassaultpcN))
```

rape (rapepcN)

```
crime %>%
  filter(crime$citytype==1) %>%
  summarise(mean=mean(rapepcN), sd=sd(rapepcN))
```

```
crime %>%
  filter(crime$citytype==2) %>%
  summarise(mean=mean(rapepcN), sd=sd(rapepcN))
```

```
crime %>%
  filter(crime$citytype==3) %>%
  summarise(mean=mean(rapepcN), sd=sd(rapepcN))
```

Average crime rates over the three groups, with group 1 that includes the large cities show on average a higher amount of rape and sexual abuse. These figures are comparable to those from the paper

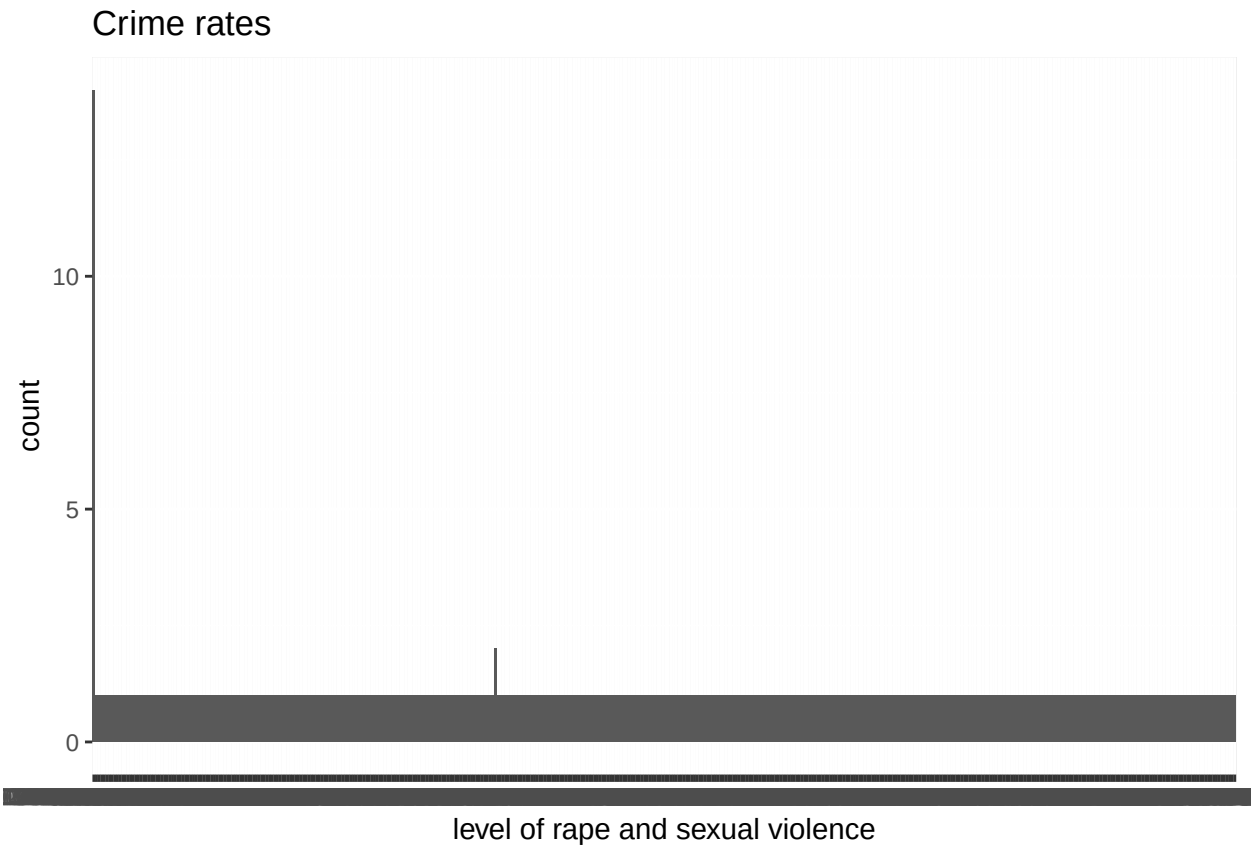
mean	sd
0.664	0.464

mean	sd
1.03	0.342

(b) Is the outcome variable rapesexapcN normally distributed? Just look at the shape of the distribution, do not worry about a statistical test for normality.

```
ggplot(crime, aes(x=as.factor(rapesexapcN)))+
 geom_histogram(stat="count")+
 labs(x='level of rape and sexual violence', y='count', title='Crime rates')
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



3. Statistical test of a treatment effect

(a)

$$Y_{i,s,t} = \alpha + \delta(TREAT_s x POST_t) + \beta TREAT_s + \gamma POST_t + \varepsilon_{i,s,t}$$

(b)

```
didrapesexass <- lm(lnrapesexaN ~ opening + logpopmale1565 + logpopdens + inkhh + educhpc + nondutchpc +
mayorCDA + mayorCU + mayorD66 + mayorVVD + factor(crime$city) + factor(crime$year), data=crime)
summ(didrapesexass)
```

mean	sd
0.846	0.4

mean	sd
0.691	0.432

```
## MODEL INFO:
## Observations: 450
## Dependent Variable: lnrapesexaN
## Type: OLS linear regression
##
## MODEL FIT:
## F(52,397) = 41.65, p = 0.00
## R² = 0.85
## Adj. R² = 0.82
##
## Standard errors: OLS
## ------------------------------------------------------------------------
##                                        Est.   S.E.   t val.      p
## ------------------------------------- ------- ------ -------- ------
## (Intercept)                            8.86   5.25     1.69   0.09
## opening                               -0.14   0.07    -1.96   0.05
## logpopmale1565                        -0.88   0.48    -1.83   0.07
## logpopdens                             0.05   0.52     0.10   0.92
## inkhh                                  0.06   0.14     0.41   0.68
## educhpc                                3.16   1.09     2.90   0.00
## nondutchpc                             0.70   2.77     0.25   0.80
## insurWWAO                              0.00   0.00     0.07   0.95
## mayorCDA                              -0.11   0.10    -1.06   0.29
## mayorCU                                0.16   0.20     0.83   0.41
## mayorD66                               0.01   0.14     0.08   0.94
## mayorVVD                              -0.02   0.06    -0.34   0.73
## factor(crime$city)Amsterdam            3.46   0.96     3.59   0.00
## factor(crime$city)Arnhem               1.07   0.35     3.11   0.00
## factor(crime$city)Breda                0.51   0.68     0.74   0.46
## factor(crime$city)Den Haag             2.91   0.70     4.18   0.00
## factor(crime$city)Deventer            -0.39   0.43    -0.91   0.36
## factor(crime$city)Dordrecht            0.22   0.31     0.70   0.49
## factor(crime$city)Eindhoven            1.32   0.43     3.05   0.00
## factor(crime$city)Enschede             1.18   0.45     2.60   0.01
## factor(crime$city)Groningen            0.91   0.52     1.76   0.08
## factor(crime$city)Haarlem              0.13   0.70     0.18   0.85
## factor(crime$city)Heerlen              0.07   0.48     0.14   0.89
## factor(crime$city)Helmond             -0.19   0.31    -0.62   0.54
## factor(crime$city)Hengelo             -0.45   0.22    -2.02   0.04
## factor(crime$city)Leeuwarden           0.01   0.32     0.02   0.99
## factor(crime$city)Leiden              -0.37   0.80    -0.47   0.64
## factor(crime$city)Maastricht          -0.17   0.30    -0.59   0.56
## factor(crime$city)Nijmegen             0.54   0.43     1.25   0.21
## factor(crime$city)Rotterdam            4.00   0.82     4.90   0.00
## factor(crime$city)Schiedam            -0.11   0.71    -0.16   0.88
## factor(crime$city)sHertogenbosch       0.30   0.60     0.51   0.61
```

```
## factor(crime$city)Tilburg                1.17   0.42     2.80   0.01
## factor(crime$city)Utrecht                1.50   0.66     2.29   0.02
## factor(crime$city)Venlo                 -0.42   0.19    -2.22   0.03
## factor(crime$city)Zwolle                -0.29   0.46    -0.62   0.54
## factor(crime$year)1995                   0.12   0.10     1.24   0.22
## factor(crime$year)1996                   0.15   0.10     1.52   0.13
## factor(crime$year)1997                   0.04   0.10     0.42   0.68
## factor(crime$year)1998                  -0.02   0.11    -0.19   0.85
## factor(crime$year)1999                   0.03   0.11     0.23   0.82
## factor(crime$year)2000                   0.02   0.11     0.17   0.87
## factor(crime$year)2001                   0.20   0.12     1.70   0.09
## factor(crime$year)2002                   0.04   0.12     0.36   0.72
## factor(crime$year)2003                  -0.02   0.12    -0.20   0.84
## factor(crime$year)2004                   0.16   0.12     1.27   0.20
## factor(crime$year)2005                   0.13   0.13     0.97   0.33
## factor(crime$year)2006                   0.19   0.40     0.48   0.63
## factor(crime$year)2007                   0.00   0.23     0.01   1.00
## factor(crime$year)2008                  -0.29   0.17    -1.69   0.09
## factor(crime$year)2009                  -0.46   0.24    -1.92   0.06
## factor(crime$year)2010                  -0.56   0.28    -2.00   0.05
## factor(crime$year)2011                  -0.65   0.28    -2.37   0.02
## ------------------------------------------------------------------------
```

With the opening of the tippelzones there is a 14% decrease in rape and sexual assault crimes. This is relatively a large effect.

(c) For the regression your ran under (b), exclude all the covariates. Do this for the 'sexual abuse and rape' variable only. What happens to the size of the estimated treatment effect? How are we to explain this change in the estimated effect?

```
rapesexassreg1 <- lm(lnrapesexaN ~ opening, data=crime)
summ(rapesexassreg1)
```

```
## MODEL INFO:
## Observations: 450
## Dependent Variable: lnrapesexaN
## Type: OLS linear regression
##
## MODEL FIT:
## F(1,448) = 79.06, p = 0.00
## R² = 0.15
## Adj. R² = 0.15
##
## Standard errors: OLS
## ------------------------------------------------
##                      Est.   S.E.   t val.      p
## ----------------- ------ ------ -------- ------
## (Intercept)          2.79   0.04    67.03   0.00
## opening              0.76   0.09     8.89   0.00
## ------------------------------------------------
```

the estimated effect is now large and positive. This is due to not taking into account different cities with different populations and also different crime rates.

(d)

Falsification is used to test for an effect that should not be there. So here we would be able to see that the

treatment variable and outcome are not related.

```
theft <- lm(lnsimtheftN ~ opening + logpopmale1565 + logpopdens + inkhh + educhpc + nondutchpc + insurW
mayorCDA + mayorCU + mayorD66 + mayorVVD, data=crime)
summ(theft)
```

```
## MODEL INFO:
## Observations: 450
## Dependent Variable: lnsimtheftN
## Type: OLS linear regression
##
## MODEL FIT:
## F(11,438) = 272.18, p = 0.00
## R² = 0.87
## Adj. R² = 0.87
##
## Standard errors: OLS
## ------------------------------------------------------
##                         Est.   S.E.   t val.      p
## -------------------- ------- ------ -------- ------
## (Intercept)            -5.89   0.74    -7.93   0.00
## opening                 0.17   0.04     3.90   0.00
## logpopmale1565          1.23   0.07    17.16   0.00
## logpopdens             -0.08   0.04    -2.00   0.05
## inkhh                  -0.03   0.01    -2.87   0.00
## educhpc                -0.66   0.28    -2.34   0.02
## nondutchpc              1.57   1.13     1.39   0.16
## insurWWAO               0.00   0.00     0.36   0.72
## mayorCDA               -0.17   0.04    -4.07   0.00
## mayorCU                -0.36   0.15    -2.33   0.02
## mayorD66                0.21   0.11     1.96   0.05
## mayorVVD               -0.00   0.04    -0.06   0.95
## ------------------------------------------------------
```

Effect is small which is good, but this regression is incorrect as it needs the correct one from a.