EDS 241: Assignment 2

Alexandra Yousefivand

01/29/2022

```
# load packages
packages=c("stargazer", "here", "tidyr", "dplyr","stringr", "janitor",
           "cowplot", "ggplot2", "tinytex", "datasets", "tibble", "estimatr")

for (i in packages) {
  if (require(i,character.only=TRUE)==FALSE) {
    install.packages(i,repos='http://cran.us.r-project.org')
  }
  else {
    require(i,character.only=TRUE)
  }
}

#devtools::install_github('rstudio/rmarkdown')
options(scipen=999) # not scientific notation
```

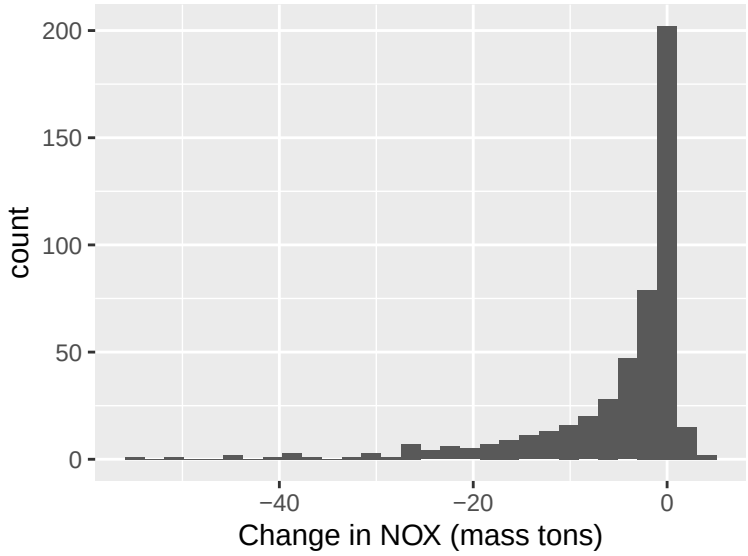
```
# load data
data_raw <- read.csv("NBP.csv")

data <- data_raw %>% clean_names()
```

Question (a) Histogram

```
histogram <- ggplot(data = data, mapping = aes(x = dnox_masstons)) +
                    geom_histogram() +
                    labs(x = "Change in NOX (mass tons)")

histogram
```



Question (b) Indicator = 1 if pct_black > sample median, 0 otherwise

```
# create indicator variable column
data <- mutate(.data = data,
                  D = case_when(pct_black >  median(data$pct_black) ~ 1,
                                pct_black <= median(data$pct_black) ~ 0))
```

```
data_1 <- filter(.data = data, D == 1)

avg_pct_black <- mean(data_1$pct_black)
print(paste0(round(avg_pct_black, 2), "%"))
```

```
## [1] "19.91%"
```

Question (c) Regression

```
model <- lm_robust(formula = dnox_masstons ~ nbp, data = data)
summary(model)
```

```
##
## Call:
## lm_robust(formula = dnox_masstons ~ nbp, data = data)
##
## Standard error type:  HC2
##
## Coefficients:
##             Estimate Std. Error t value             Pr(>|t|) CI Lower CI Upper
```

```
## (Intercept)   -3.615     0.4207  -8.593 0.0000000000000001181   -4.442   -2.789
## nbp           -3.908     0.7958  -4.911 0.0000012411634147570   -5.472   -2.345
##               DF
## (Intercept) 483
## nbp         483
##
## Multiple R-squared:  0.05152 ,   Adjusted R-squared:  0.04956
## F-statistic: 24.12 on 1 and 483 DF,  p-value: 0.000001241
```

Intercept = -3.615 tons

The change in annual NOx emissions from all power plants in a county between 2000 and 2008, if the county was NOT regulated under the NOx Budget Program (NBP = 0).

Slope coefficient = -3.908 tons

The difference in the change in annual NOx emissions from all power plants between 2000 and 2008 in a county regulated under the NOx Budget Program and another county NOT regulated under the NOx Budget Program (i.e. regulation / no regulation difference in annual NOx emissions).

Question (d) Interaction between NBP andd D

```
model_interaction <- lm_robust(formula = dnox_masstons ~ nbp + D + nbp:D,
                               data = data)
summary(model_interaction)
```

```
##
## Call:
## lm_robust(formula = dnox_masstons ~ nbp + D + nbp:D, data = data)
##
## Standard error type:  HC2
##
## Coefficients:
##             Estimate Std. Error t value      Pr(>|t|) CI Lower CI Upper  DF
## (Intercept)   -2.601      0.469  -5.547 0.00000004815   -3.523  -1.6798 481
## nbp           -6.333      1.216  -5.208 0.00000028281   -8.722  -3.9436 481
## D             -2.215      0.859  -2.578 0.01022558126   -3.902  -0.5269 481
## nbp:D          5.035      1.592   3.163 0.00165763434    1.908   8.1630 481
##
## Multiple R-squared:  0.07284 ,   Adjusted R-squared:  0.06706
## F-statistic: 12.27 on 3 and 481 DF,  p-value: 0.00000009556
```

Intercept: -3.615 tons

The change in annual NOx emissions from all power plants in a county between 2000 and 2008, if the county was NOT regulated under the NOx Budget Program (NBP = 0) and the pct_black is below the sample median.

NBP coefficient = -6.333 tons

The difference in the change in annual NOx emissions from all power plants between 2000 and 2008 in a county regulated under the NOx Budget Program compared to another county NOT regulated under the NOx Budget Program (i.e. regulation / no regulation difference in annual NOx emissions), with all else constant.

D coefficient = -2.215 tons

The difference in the change in annual NOx emissions from all power plants between 2000 and 2008 in a county with pct_black above the sample median compared to a county with pct_black below the sample median, with all else constant.

Interaction coefficient = 5.035 tons

The difference in the change in annual NOx emissions from all power plants between 2000 and 2008 in a county regulated with the NOx Budget Program and with pct_black above the sample median compared to another county with the NOx Budget Program and below the sample median.

Question (e) Prediction and 95% confidence interval

```
new_data <- data.frame(nbp = 0, D = 1)

ci <- predict(object = model_interaction,
              newdata = new_data,
              se.fit = TRUE,
              interval = "confidence")
ci
```

```
## $fit
##         fit       lwr       upr
## [1,] -4.816 -6.229991 -3.402009
##
## $se.fit
##         1
## 0.7196221
```