Problem Set 2

Tate Mason

2024-02-12

Problem 1

Hansen 2.2

If $\mathbb{E}[Y|X] = a + bX$, we can do the following:

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]|X] = \mathbb{E}[X(a + bX)] \to \mathbb{E}[YX] = a\mathbb{E}[X] + b\mathbb{E}[X]^2$$

Problem 2

Hansen 2.5

Part A

$$\mathbb{E}[(Y - h(X))^2] = \mathbb{E}[((Y - m(X)) + (m(X) - h(X)))^2] =$$
$$\mathbb{E}[(e + (m(X) - h(X)))^2] = \mathbb{E}[e^2 + 2e(m(X) - h(X)) + (m(X) - h(X))^2] = \mathbb{E}[e^2]$$

Part B

Predicting e^2 is to estimate the amount of variation in the model in expectation. That is, it estimates σ^2 which is present due to the potential omission of omni-present factors.

Part C

Problem 3

Hansen 2.6

Problem 4

Hansen 2.10

This is true. We can use the law of iterated expectations to prove it. $\mathbb{E}[X^2e] = \mathbb{E}[\mathbb{E}[X^2e|X]$ We can then do some factoring to yield $X^2\mathbb{E}[e|X]$. Because of the fact that $\mathbb{E}[e|X] = 0$, we can say that $\mathbb{E}[X^2e] = \mathbb{E}[0] = 0$

Problem 5

Hansen 2.11

This is false. Mainly, this is due to the fact that we are not given information of the error term's correlation to X. If we were to make X a binary variable, with equal possibility of equaling negative one or one, $\mathbb{E}[Xe] = \mathbb{E}[X^2] = 1$. This is a contradiction as, in this case, the equality is not zero.

Problem 6

Hansen 2.12

This is also false. The condition of $\mathbb{E}[e|X] = 0$ implies zero-mean expectation, but not full independence.

Problem 7

Hansen 2.13

False. The condition of $\mathbb{E}[Xe] = 0$ implies only expectationally zero correlation between X and the error term e. In reality, this is no guarantee of the zero mean assumption holding.

Problem 8

Hansen 2.14

Finally, this is also false. From the given conditions, we can assume that the error is homoskedastic, but nothing is given regarding independence.

Problem 9

Hansen 2.21

Part A

$\gamma_1 = \beta_1$ if and only if the quadratic term's $\beta_2 = 0$. This would imply that in the short term regression, there is no omitted variable bias, thus meaning the two parameters would be equivalent.

Part B

It is a similar case here in which the cubic term's $\theta_2 = 0$ or in the case that $\mathbb{E}[X^3 X] = 0$, or that X^3 is uncorrelated with X.

Problem 10

Part A

Part B

Part C

Part D

Problem 11

Part A

```r
library(Ecdat)
```

```
Loading required package: Ecfun


Attaching package: 'Ecfun'

The following object is masked from 'package:base':

    sign


Attaching package: 'Ecdat'

The following object is masked from 'package:datasets':

    Orange
```

```r
library(dplyr)
```

```

Attaching package: 'dplyr'
```

```
The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```r
data(Star,package="Ecdat")
str(Star)
```

```
'data.frame':	5748 obs. of  8 variables:
 $ tmathssk: int  473 536 463 559 489 454 423 500 439 528 ...
 $ treadssk: int  447 450 439 448 447 431 395 451 478 455 ...
 $ classk  : Factor w/ 3 levels "regular","small.class",..: 2 2 3 1 2 1 3 1 2 2 ...
 $ totexpk : int  7 21 0 16 5 8 17 3 11 10 ...
 $ sex     : Factor w/ 2 levels "girl","boy": 1 1 2 2 2 2 1 1 1 1 ...
 $ freelunk: Factor w/ 2 levels "no","yes": 1 1 2 1 2 2 2 1 1 1 ...
 $ race    : Factor w/ 3 levels "white","black",..: 1 2 2 1 1 1 2 1 2 1 ...
 $ schidkn : int  63 20 19 69 79 5 16 56 11 66 ...
 - attr(*, "na.action")= 'omit' Named int [1:5850] 1 4 6 7 8 9 10 15 16 17 ...
  ..- attr(*, "names")= chr [1:5850] "1" "4" "6" "7" ...
```

This loads the required package and associated data into the program. Next, we will subset the data and perform the necessary actions to find the ATT

```r
boys <- Star %>%
  filter(sex == "boy", classk %in% c("small.class","regular")) %>%
  mutate(small_class = ifelse(classk == "small.class", 1,0))
mean_score <- boys %>%
  group_by(small_class) %>%
  summarise(mean_score = mean(tmathssk, na.rm=TRUE), .groups="drop")
mean_small <- mean_score %>% filter(small_class == 1) %>% pull(mean_score)
mean_not <- mean_score %>% filter(small_class == 0) %>% pull(mean_score)
ATT_mean <- mean_small - mean_not
ATT_mean
```

```
[1] 13.67522
```

Part B

```
Y <- as.matrix(boys$tmathssk)
X <- as.matrix(cbind(1,boys$small_class))
beta_hat <- solve(t(X) %*% X) %*% t(X) %*% Y
beta_hat
```

```
          [,1]
[1,] 476.69066
[2,]  13.67522
```

The effect of being in a small class is the same as in part (A). It also matches with the built in regression function.

Part C

```
X_c <- as.matrix(cbind(1, boys$small_class, boys$totexpk, boys$freelunk))
beta_hat_c <- solve(t(X_c) %*% X_c) %*% t(X_c) %*% Y
beta_hat_c
```

```
            [,1]
[1,] 505.2189955
[2,]  13.4233260
[3,]   0.5397799
[4,] -22.4616584
```

In this case, the effect of a small class is slightly smaller. This makes sense as with more factors being taken into account, one would expect less "oomph" from a singular variable. With that said, parameter coefficients match with both matrix algebra and `lm`.