

Chapter 7

Thalles Quinaglia Liduares

2022-07-12

Exercise 7.6

```
library(wooldridge)
library(lmreg)
library(car)
```

Use the data in SLEEP75.RAW for this exercise. The equation of interest is

```
data<-wooldridge::sleep75

attach(data)
```

```
## The following object is masked from package:datasets:
##
##      sleep
```

Use the data in SLEEP75.RAW for this exercise. The equation of interest is

$$\text{sleep} = \beta_0 + \beta_1 \text{totwork} + \beta_2 \text{educ} + \beta_3 \text{age} + \beta_4 \text{age}^2 + \beta_5 \text{yngkid} + u$$

(i) Estimate this equation separately for men and women and report the results in the usual form. Are there notable differences in the two estimated equations?

```
summary(lm_men<-lm(sleep~totwrk+educ+age+agesq+yngkid, data=subset(data, male==1)))
```

```
##
## Call:
## lm(formula = sleep ~ totwrk + educ + age + agesq + yngkid, data = subset(data,
##   male == 1))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1793.96  -216.05    7.93   244.57  1141.21
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3648.20826   310.03933   11.767 < 2e-16 ***
## totwrk      -0.18212     0.02449   -7.438 6.45e-13 ***
## educ       -13.05238     7.41422   -1.760  0.0791 .
## age         7.15659    14.32037    0.500  0.6175
## agesq      -0.04477     0.16841   -0.266  0.7905
## yngkid      60.38021    59.02278    1.023  0.3069
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 402.3 on 394 degrees of freedom
## Multiple R-squared:  0.1562, Adjusted R-squared:  0.1455
## F-statistic: 14.59 on 5 and 394 DF,  p-value: 3.952e-13
```

The estimated coefficients for men equation is given by

$$\widehat{sleep}_{men} = 3648.20 - 0.182totwork - 13.05educ + 7.15age - 0.04age^2 + 60.38yngkid$$

Now, for the woman case:

```
summary(lm_woman<-lm(sleep~totwrk+educ+age+agesq+yngkid, data=subset(data, male==0)))
```

```
##
## Call:
## lm(formula = sleep ~ totwrk + educ + age + agesq + yngkid, data = subset(data,
##   male == 0))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2485.02  -244.18    7.24   270.64  1376.91
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4238.72933   384.89226   11.013 < 2e-16 ***
## totwrk       -0.13995    0.02766   -5.060 7.33e-07 ***
## educ        -10.20514    9.58885   -1.064  0.288
## age         -30.35657   18.53091   -1.638  0.102
## agesq         0.36794    0.22334    1.647  0.101
## yngkid      -118.28256   93.18757   -1.269  0.205
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 437 on 300 degrees of freedom
## Multiple R-squared:  0.09768,    Adjusted R-squared:  0.08264
## F-statistic: 6.495 on 5 and 300 DF,  p-value: 9.42e-06
```

For woman's case, the estimated coefficients are given by

$$\widehat{sleep}_{woman} = 4238.72 - 0.139totwork - 10.205educ - 30.35age + 0.367age^2 - 118.28yngkid$$

The coefficients differ very significantly between these two categories.

(ii) Compute the Chow test for equality of the parameters in the sleep equation for men and women. Use the form of the test that adds male and the interaction terms male?totwrk, ..., male?yngkid and uses the full set of observations. What are the relevant df for the test? Should you reject the null at the 5% level?

```
summary(lm_chow<-lm(sleep~totwrk+educ+age+agesq+yngkid+male+male*totwrk+male*educ
+male*age+male*agesq+male*yngkid))
```

```
##
## Call:
## lm(formula = sleep ~ totwrk + educ + age + agesq + yngkid + male +
##      male * totwrk + male * educ + male * age + male * agesq +
##      male * yngkid)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2485.02  -226.74    7.93   257.77  1376.91
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4238.72933   367.85193   11.523 < 2e-16 ***
## totwrk        -0.13995    0.02643   -5.294 1.61e-07 ***
## educ         -10.20514    9.16432   -1.114  0.2658
## age          -30.35657   17.71049   -1.714  0.0870 .
## agesq         0.36794    0.21345    1.724  0.0852 .
## yngkid       -118.28256   89.06187   -1.328  0.1846
## male         -590.52107  488.79159   -1.208  0.2274
## totwrk:male   -0.04217    0.03667   -1.150  0.2506
## educ:male     -2.84724   11.96795   -0.238  0.8120
## age:male      37.51316   23.12332    1.622  0.1052
## agesq:male    -0.41271    0.27591   -1.496  0.1352
## yngkid:male  178.66277  108.10510    1.653  0.0988 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 417.6 on 694 degrees of freedom
## Multiple R-squared:  0.1306, Adjusted R-squared:  0.1168
## F-statistic: 9.479 on 11 and 694 DF,  p-value: 4.947e-16
```

```
linearHypothesis(lm_chow, c("male=0", "totwrk:male=0", "educ:male=0", "age:male=0", "agesq:male=0", "yngkid:male =0"))
```

```
## Linear hypothesis test
##
## Hypothesis:
## male = 0
## totwrk: male = 0
## educ: male = 0
## age: male = 0
## agesq: male = 0
## yngkid: male = 0
##
## Model 1: restricted model
## Model 2: sleep ~ totwrk + educ + age + agesq + yngkid + male + male *
##      totwrk + male * educ + male * age + male * agesq + male *
##      yngkid
##
##      Res.Df      RSS Df Sum of Sq      F Pr(>F)
## 1      700 123267451
## 2      694 121052555   6    2214896 2.1164 0.04949 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Based on the p-value of the performed hypothesis test we can reject the null that the coefficients between men and woman are equal at 5% level.