HW7 Patrick Neyland

Patrick Neyland

2022-10-18

## Question 1

1. and (iii) can cause the usual OLS t statistic to be invalid.

## Question 2

### Part i

### Part ii All positive

### Part iii

10 percent increase in pop will reult in .66 percent increase in rent.

### Part iv

t = ,0056 - 0 / 0.0017 = 56/17 = 3.294

pt(3.294, 60, lower.tail = FALSE)

[1] 0.0008300033

Easily reject the null

0.99917

or

calculate the critical value. |t| > |c| ?

## Question 3

var(betahat1) + var(9betahat2) - 6cov(beathat1, beathat2)

## Question 5

### Part i

df <- filter(k401ksubs, fsize == 1)  
length(df$fsize)

[1] 2017

The data set contains 2017 single person households.

### Part ii

model5\_2 <- lm(nettfa ~ inc + age, data = df)  
stargazer(model5\_2, type = "text")

===============================================  
 Dependent variable:   
 ---------------------------  
 nettfa   
-----------------------------------------------  
inc 0.799\*\*\*   
 (0.060)   
   
age 0.843\*\*\*   
 (0.092)   
   
Constant -43.040\*\*\*   
 (4.080)   
   
-----------------------------------------------  
Observations 2,017   
R2 0.119   
Adjusted R2 0.118   
Residual Std. Error 44.683 (df = 2014)   
F Statistic 136.465\*\*\* (df = 2; 2014)   
===============================================  
Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

No, there are no surprises in the the slope estimates.

### Part iii

I think it shows that a lot of people start out with a lot of debt early in their careers—student loans, mortgage, etc.

### Part iv

summary(model5\_2)

Call:  
lm(formula = nettfa ~ inc + age, data = df)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-179.95 -14.16 -3.42 6.03 1113.94   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) -43.03981 4.08039 -10.548 <2e-16 \*\*\*  
inc 0.79932 0.05973 13.382 <2e-16 \*\*\*  
age 0.84266 0.09202 9.158 <2e-16 \*\*\*  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 44.68 on 2014 degrees of freedom  
Multiple R-squared: 0.1193, Adjusted R-squared: 0.1185   
F-statistic: 136.5 on 2 and 2014 DF, p-value: < 2.2e-16

### Part v

model5\_5 <- lm(nettfa ~ inc, data = df)  
stargazer(model5\_5, type = "text")

===============================================  
 Dependent variable:   
 ---------------------------  
 nettfa   
-----------------------------------------------  
inc 0.821\*\*\*   
 (0.061)   
   
Constant -10.571\*\*\*   
 (2.061)   
   
-----------------------------------------------  
Observations 2,017   
R2 0.083   
Adjusted R2 0.082   
Residual Std. Error 45.592 (df = 2015)   
F Statistic 181.599\*\*\* (df = 1; 2015)   
===============================================  
Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

The estimated coefficient on is not much different from the estimate in part ii.

## Question 6

### Part i

model6\_1 <- lm(log(psoda) ~ prpblck + log(income) + prppov, data = discrim)  
stargazer(model6\_1, type = "text")

===============================================  
 Dependent variable:   
 ---------------------------  
 log(psoda)   
-----------------------------------------------  
prpblck 0.073\*\*   
 (0.031)   
   
log(income) 0.137\*\*\*   
 (0.027)   
   
prppov 0.380\*\*\*   
 (0.133)   
   
Constant -1.463\*\*\*   
 (0.294)   
   
-----------------------------------------------  
Observations 401   
R2 0.087   
Adjusted R2 0.080   
Residual Std. Error 0.081 (df = 397)   
F Statistic 12.604\*\*\* (df = 3; 397)   
===============================================  
Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Yes, is statistically different from zero at a 5% level in a two-sided test. However, it is not significant at the 1% level.

### Part ii

df1 <- na.omit(discrim)  
cor(df1$lincome, df1$prppov)

[1] -0.8468178

cor(log(df1$income), df1$prppov)

[1] -0.8468178

summary(model6\_1)

Call:  
lm(formula = log(psoda) ~ prpblck + log(income) + prppov, data = discrim)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-0.32218 -0.04648 0.00651 0.04272 0.35622   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) -1.46333 0.29371 -4.982 9.4e-07 \*\*\*  
prpblck 0.07281 0.03068 2.373 0.0181 \*   
log(income) 0.13696 0.02676 5.119 4.8e-07 \*\*\*  
prppov 0.38036 0.13279 2.864 0.0044 \*\*   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 0.08137 on 397 degrees of freedom  
 (9 observations deleted due to missingness)  
Multiple R-squared: 0.08696, Adjusted R-squared: 0.08006   
F-statistic: 12.6 on 3 and 397 DF, p-value: 6.917e-08

The correlation between and is -0.847 Yes, each of the variables is statistically significant.

| Variable | p-value |  |  |  |
| --- | --- | --- | --- | --- |
| prpblck | 0.181 |  |  |  |
| log(income) | 0.000 |  |  |  |
| prppov | 0.004 |  |  |  |