HW11 - Patrick Neyland

Patrick Neyland

2022-11-17

## Question 1

### Part 1

The coefficients will all switch signs and the new intercept will be 0.414

### Part 2

Standard errors will not change. ### Part 3 No change in

## Question 2 ### Part 1

df <- mutate(apple, ecobuy = ifelse(ecolbs>0, 1, 0 ))  
sum(df$ecobuy)/length(df$ecobuy)

[1] 0.6242424

The fraction of families who claim they would buy ecolabeled apples is 62.4%

### Part 2

MRM <- lm(ecobuy ~ ecoprc + regprc + faminc + hhsize + educ + age, df)  
stargazer(MRM, digits = 4, type = "text")

===============================================  
 Dependent variable:   
 ---------------------------  
 ecobuy   
-----------------------------------------------  
ecoprc -0.8026\*\*\*   
 (0.1094)   
   
regprc 0.7193\*\*\*   
 (0.1316)   
   
faminc 0.0006   
 (0.0005)   
   
hhsize 0.0238\*   
 (0.0125)   
   
educ 0.0248\*\*\*   
 (0.0084)   
   
age -0.0005   
 (0.0012)   
   
Constant 0.4237\*\*   
 (0.1650)   
   
-----------------------------------------------  
Observations 660   
R2 0.1098   
Adjusted R2 0.1016   
Residual Std. Error 0.4594 (df = 653)   
F Statistic 13.4270\*\*\* (df = 6; 653)   
===============================================  
Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

If ecoprc increases by 1 dollar, on average the likelihood of buying eco-labled apples decreases by 80.3 percent. If the regprc increases by 1 dollar, on average the likelihood of buying eco-labeled apples increases by 71.9 percent.

### Part 3

linearHypothesis(MRM, c("faminc=0", "hhsize=0", "educ = 0", "age = 0"))

Linear hypothesis test  
  
Hypothesis:  
faminc = 0  
hhsize = 0  
educ = 0  
age = 0  
  
Model 1: restricted model  
Model 2: ecobuy ~ ecoprc + regprc + faminc + hhsize + educ + age  
  
 Res.Df RSS Df Sum of Sq F Pr(>F)   
1 657 141.55   
2 653 137.81 4 3.7376 4.4276 0.001544 \*\*  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

With a p-value of 0.0015, the non-price variables are jointly significant. Looking at the model, the variable seems to have the most important effect. This makes sense to me, someone with more education may be more informed on the importance of purchasing and consuming eco-friendly produce. However, also acts as proxy variable for other variables that are also more significant.

### Part 4

MRM\_log <- lm(ecobuy ~ ecoprc + regprc + log(faminc) + hhsize + educ + age, df)  
stargazer(MRM,MRM\_log, digits = 4, type = "text")

===========================================================  
 Dependent variable:   
 ----------------------------  
 ecobuy   
 (1) (2)   
-----------------------------------------------------------  
ecoprc -0.8026\*\*\* -0.8007\*\*\*   
 (0.1094) (0.1093)   
   
regprc 0.7193\*\*\* 0.7214\*\*\*   
 (0.1316) (0.1315)   
   
faminc 0.0006   
 (0.0005)   
   
log(faminc) 0.0445   
 (0.0287)   
   
hhsize 0.0238\* 0.0227\*   
 (0.0125) (0.0125)   
   
educ 0.0248\*\*\* 0.0231\*\*\*   
 (0.0084) (0.0085)   
   
age -0.0005 -0.0004   
 (0.0012) (0.0013)   
   
Constant 0.4237\*\* 0.3038\*   
 (0.1650) (0.1790)   
   
-----------------------------------------------------------  
Observations 660 660   
R2 0.1098 0.1116   
Adjusted R2 0.1016 0.1034   
Residual Std. Error (df = 653) 0.4594 0.4589   
F Statistic (df = 6; 653) 13.4270\*\*\* 13.6730\*\*\*   
===========================================================  
Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

The log model has a slightly higher adjusted however, because the two different variables between the models are not statistically significant, I would say the models are about the same. In the second model, as faminc increases by 1 percent, ecobuy will increase by 0.00045.

### Part 5

sum(predict(MRM\_log)<0)

[1] 0

sum(predict(MRM\_log)>1)

[1] 2

summary(predict(MRM\_log))

Min. 1st Qu. Median Mean 3rd Qu. Max.   
 0.1854 0.4937 0.6241 0.6242 0.7551 1.0507

0 are negative and only 2 are above 1. We do not need to be concerned.

### Part 6

confusion\_matrix <- table(df$ecobuy, predict(MRM\_log) >= 0.5)  
confusion\_matrix

FALSE TRUE  
 0 102 146  
 1 72 340

ecobuy = 0 | 102/(102+146) = .411 or 41.1% ecobuy = 1 | 340/(72+340) = .825 or 82.5% (better) Predicted correctly | (102 + 340)/(102+146+72+340) = 0.669 or 66.9% The ecobuy = 1 model does a better job of predicting the decision to buy eco-labeled apples.