Econometrics-HW5

Wei Ye

4/18/2022

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.8
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1
## Warning: package 'tidyr' was built under R version 4.0.5
## Warning: package 'readr' was built under R version 4.0.5
## Warning: package 'dplyr' was built under R version 4.0.5
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
apple <- read_csv("apple.csv")
```

```
## Rows: 660 Columns: 17
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (1): state
## dbl (16): id, educ, date, regprc, ecoprc, inseason, hhsize, male, faminc, ag...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
head(apple)
```

```
## # A tibble: 6 x 17
##      id  educ   date state regprc ecoprc inseason hhsize  male faminc   age
##   <dbl> <dbl>  <dbl> <chr>  <dbl>  <dbl>    <dbl>  <dbl> <dbl>  <dbl> <dbl>
## 1 10002    16 111597 SD      1.19   1.19        1      4     0     45    43
## 2 10004    16 121897 KS      0.59   0.79        0      1     0     65    37
## 3 10034    18 111097 MI      0.59   0.99        1      3     0     65    44
## 4 10035    12 111597 TN      0.89   1.09        1      2     1     55    55
## 5 10039    15 122997 NY      0.89   1.09        0      1     1     25    22
## 6 10041    12 112297 WV      0.59   0.79        1      4     0     15    34
## # ... with 6 more variables: reglbs <dbl>, ecolbs <dbl>, numlt5 <dbl>,
## #   num5_17 <dbl>, num18_64 <dbl>, numgt64 <dbl>
```

```
#create dummy variables
apple$ecobuy <- ifelse(apple$ecolbs>0,1,0)
```

```
attach(apple)
```

a. Derive the log-likelihood function.

$\max l_i = p_1^{regprc_i} p_2^{ecoprc_i} (1-p_1-p_2)^{age_i}$ Take log on the above eqn: $\max \sum_i^N \log l_i = \sum_i^N regprc_i \log p_i + ecoprc_i \log p_2 + age_i \log(1-p_1-p_2)$

b. Estimate the parameters of interest via M-estimation

```
source('~/Dropbox/My Mac (Wei's MacBook Air)/Downloads/PhD-Coursework/22Spring/Econometrics/My_solution_
 source('~/Dropbox/My Mac (Wei's MacBook Air)/Downloads/PhD-Coursework/22Spring/Econometrics/My_solutio
 source('~/Dropbox/My Mac (Wei's MacBook Air)/Downloads/PhD-Coursework/22Spring/Econometrics/My_solutio
```

```
X=cbind(regprc,ecoprc,age)
thetahat=cbind(0.2,0.6,0.2)
out <- qderivfun(ecobuy,X,t(thetahat))
out1 <- qderivfun2(ecobuy,X,t(thetahat))
out1
```

```
## $H
##              [,1]       [,2]         [,3]
## [1,] -2207533.70  -11158.73   -1304947.8
## [2,]   -11158.73 -301384.07    -533228.8
## [3,] -1304947.80 -533228.77 -194413608.1
##
## $AOhat
##              [,1]       [,2]         [,3]
## [1,] -3344.74802  -16.90717   -1977.1936
## [2,]   -16.90717 -456.64253    -807.9224
## [3,] -1977.19364 -807.92238 -294566.0729
```

c. Make t-test:

```
ava_theta <- solve(out1$AOhat)%*% out$BOhat %*% solve(out1$AOhat)/length(ecobuy)
ava_theta
```

```
##               [,1]         [,2]         [,3]
## [1,] 2.860107e-05 8.278539e-05 2.523338e-05
## [2,] 8.278539e-05 2.547686e-04 7.605291e-05
## [3,] 2.523338e-05 7.605291e-05 3.348403e-05
```

```
se_theta <- diag(ava_theta)^(1/2)
se_theta
```

```
## [1] 0.005347996 0.015961471 0.005786539
```

Now, make some t-tests on theta:

```
t_1 <- log(thetahat)/se_theta
t_2 <- thetahat/se_theta
t_1
```

```
##           [,1]      [,2]      [,3]
## [1,] -300.9422 -32.00367 -278.1348
```

```
t_2
```

```
##          [,1]     [,2]     [,3]
## [1,] 37.39718 37.59052 34.56298
```

No matther what, abs of t_test values are way larger than 1.96. So we can react the zero null for all variables.

d. Use probit command in R:

```
probitout <- glm(ecobuy~regprc+ecoprc+age,family=binomial(link="probit"))
probitout
```

```
##
## Call:  glm(formula = ecobuy ~ regprc + ecoprc + age, family = binomial(link = "probit"))
##
## Coefficients:
## (Intercept)       regprc       ecoprc          age
##     1.26121      2.03914     -2.33504     -0.00432
##
## Degrees of Freedom: 659 Total (i.e. Null);  656 Residual
## Null Deviance:       873.8
## Residual Deviance: 813.5     AIC: 821.5
```

The conclusions will not be changed.

e. Compute the likelihood ratio test and justify whether reject the null hypothesis.

```
probitout_red <- glm(ecobuy~ecoprc+age,family=binomial(link="probit"))
probitout_red
```

```
##
## Call:  glm(formula = ecobuy ~ ecoprc + age, family = binomial(link = "probit"))
##
## Coefficients:
## (Intercept)       ecoprc          age
##    1.491366    -0.911912    -0.003958
##
## Degrees of Freedom: 659 Total (i.e. Null);  657 Residual
## Null Deviance:       873.8
## Residual Deviance: 843.3     AIC: 849.3
```

```
#compute likelihood ratio
library(epiDisplay)
```

```
## Loading required package: foreign
```

```
## Warning: package 'foreign' was built under R version 4.0.5
```

```
## Loading required package: survival
```

```
## Loading required package: MASS
```

```
## Warning: package 'MASS' was built under R version 4.0.5
```

```
##
## Attaching package: 'MASS'
```

```
## The following object is masked from 'package:dplyr':
##
##     select
```

```
## Loading required package: nnet
```

```
## Warning: package 'nnet' was built under R version 4.0.5
```

```
##
## Attaching package: 'epiDisplay'
```

```
## The following object is masked from 'package:ggplot2':
##
##     alpha
```

```
lrtest(probitout,probitout_red)
```

```
## Likelihood ratio test for MLE method
## Chi-squared 1 d.f. =  29.71658 , P value =  5.000524e-08
```

Actually from the command we can derive Chi-squared number from lrtest command without replying on dchisq. It's chi-squared value is high, and P-value is way lower than 0.05, therefore, we reject the null hypothesis.