Econometrics_Problem_Set_2.R

Bernardo

Mon Sep 16 21:06:14 2019

```
library(readstata13)
```

```
## Warning: package 'readstata13' was built under R version 3.4.4
```

```
library(fastDummies)
```

```
## Warning: package 'fastDummies' was built under R version 3.4.4
```

```
## create dataset
cpsdata <- read.dta13("cps09mar.dta")

## compute hourly wages
cpsdata$hourly_wage <- cpsdata$earnings / cpsdata$hours / cpsdata$week

## convert  earnings to log earnings
logwage <- cpsdata
logwage$earnings <- log(cpsdata$earnings)
logwage$hourly_wage <- log(cpsdata$hourly_wage)

## regress hourly wage on female

femalelm <- lm(hourly_wage ~ female, data = logwage)
summary(femalelm)
```

```
##
## Call:
## lm(formula = hourly_wage ~ female, data = logwage)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.9092  -0.3771   0.0100   0.3940   2.6019
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.045938   0.003901  780.84   <2e-16 ***
## female      -0.234314   0.005979  -39.19   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6659 on 50740 degrees of freedom
## Multiple R-squared:  0.02938,    Adjusted R-squared:  0.02936
## F-statistic:  1536 on 1 and 50740 DF,  p-value: < 2.2e-16
```

```
## create gender subsets and calculate mean conditional on marital status
fsubset <- subset(logwage, female == 1)
msubset <- subset(logwage, female == 0)

flogwagemarital <- by(fsubset$hourly_wage, fsubset$marital, mean )
mlogwagemarital <- by(msubset$hourly_wage, msubset$marital, mean )
```

```
## create one dataframe with 14 values for expected wage, corresponding to the 14 combinations of
## gender and marital status
cmean <- cbind(flogwagemarital, mlogwagemarital)
conditionalmean <- data.frame(y = c(cmean[, "flogwagemarital"], cmean[, "mlogwagemarital"]))

conditionalmean <- cbind(conditionalmean, female = c(1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0))
conditionalmean <- cbind(conditionalmean, marital1 = c(1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0))
conditionalmean <- cbind(conditionalmean, marital2 = c(0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0))
conditionalmean <- cbind(conditionalmean, marital3 = c(0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0))
conditionalmean <- cbind(conditionalmean, marital4 = c(0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0))
conditionalmean <- cbind(conditionalmean, marital5 = c(0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0))
conditionalmean <- cbind(conditionalmean, marital6 = c(0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0))
conditionalmean <- cbind(conditionalmean, marital7 = c(0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1))

conditionalmean
```

```
##           y female marital1 marital2 marital3 marital4 marital5 marital6
## 1  2.871203      1        1        0        0        0        0        0
## 2  2.768660      1        0        1        0        0        0        0
## 3  2.692750      1        0        0        1        0        0        0
## 4  2.738468      1        0        0        0        1        0        0
## 5  2.821688      1        0        0        0        0        1        0
## 6  2.622708      1        0        0        0        0        0        1
## 7  2.682811      1        0        0        0        0        0        0
## 8  3.150610      0        1        0        0        0        0        0
## 9  3.154428      0        0        1        0        0        0        0
## 10 2.814120      0        0        0        1        0        0        0
## 11 2.974115      0        0        0        0        1        0        0
## 12 2.966238      0        0        0        0        0        1        0
## 13 2.865955      0        0        0        0        0        0        1
## 14 2.709468      0        0        0        0        0        0        0
##    marital7
## 1         0
## 2         0
## 3         0
## 4         0
## 5         0
## 6         0
## 7         1
## 8         0
## 9         0
## 10        0
## 11        0
## 12        0
## 13        0
## 14        1
```

```
## regress expected wage on dummy variables female & marital(1-7) along with interaction terms

conditionalmeanlm <- lm(y ~ female + marital1 + female*marital1 + marital2 + female*marital2 +
                          marital3 + female*marital3 + marital4 + female*marital4 + marital5 +
                          female*marital5 + marital6 + female*marital6 + marital7 + female*marital7,
                        data= conditionalmean)
conditionalmeanlm
```

```
##
## Call:
## lm(formula = y ~ female + marital1 + female * marital1 + marital2 +
##     female * marital2 + marital3 + female * marital3 + marital4 +
##     female * marital4 + marital5 + female * marital5 + marital6 +
##     female * marital6 + marital7 + female * marital7, data = conditionalmean)
##
## Coefficients:
##     (Intercept)           female         marital1         marital2
##         2.70947         -0.02666          0.44114          0.44496
##        marital3         marital4         marital5         marital6
##         0.10465          0.26465          0.25677          0.15649
##        marital7  female:marital1  female:marital2  female:marital3
##              NA         -0.25275         -0.35911         -0.09471
## female:marital4  female:marital5  female:marital6  female:marital7
##        -0.20899         -0.11789         -0.21659               NA
```