# POLS 5003: Problem Set # 2

Team Members: Christopher Eubanks, Dongkyu Kim, and Desmond Wallace

Date: March 3, 2014

The dataset `Obama.dta` is a subset of the 2008 American National Election Survey. We will use it to examine attitudes toward Barack Obama, using the feeling thermometer `obama`.

```
> # Setup
> require(foreign)
> obama <- read.dta("Obama.dta")
> var.labels <- attr(obama,"var.labels")
> data.key <- data.frame(var.name=names(obama),var.labels)
> data.key
```

```
  var.name                      var.labels
1    obama       Obama feeling thermometer
2      age                   Years of age
3   income         Household income, $000s
4     educ              Years of education
5   female                          Female
6    black      R self-identifies as black
7      dem   R self-identifies as Democrat
8      rep R self-identifies as Republican
```

1. Suppose we hypothesize that a respondent's income affects her or his attitudes toward Obama, that those with higher incomes will express cooler feelings toward him. Controlling for age, education, gender, race, and partisanship, is this hypothesis supported? How do you know?

```
> m1 <- lm(obama ~ income + age + educ + female + black + dem + rep, data=obama)
> summary(m1)

Call:
lm(formula = obama ~ income + age + educ + female + black + dem +
    rep, data = obama)

Residuals:
    Min      1Q  Median      3Q     Max
-75.815 -11.761   3.395  12.594  66.320

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  60.20277    3.24800  18.535  < 2e-16 ***
income       -0.03332    0.01043  -3.193  0.00143 **
```

```
age          -0.03495    0.03013  -1.160  0.24629
educ          0.04891    0.21070   0.232  0.81647
female        4.48527    0.99574   4.504 7.07e-06 ***
black        16.76626    1.22609  13.675  < 2e-16 ***
dem          13.76778    1.14550  12.019  < 2e-16 ***
rep         -16.71796    1.40899 -11.865  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 21.03 on 1850 degrees of freedom
  (465 observations deleted due to missingness)
Multiple R-squared:  0.3779,        Adjusted R-squared:  0.3756
F-statistic: 160.6 on 7 and 1850 DF,  p-value: < 2.2e-16
```

According to the results above, the coefficient for `income`, representing household income in thousands of dollars, is -0.0333. This means that for every $1,000 increase in household income, respondents' feelings toward Obama will decrease by 0.0333 points. This coefficient is statistically significant at $\alpha = 0.01$. Based on these results, there is sufficient support for the hypothesis that a respondent's income affects their attitudes towards President Obama. Not only does household income affects respondents' feelings towards Obama, household income has a statistically significant *negative* effect.

2. Suppose we think Democrats' feelings toward Obama will be less influenced by their incomes than others' feelings are. Is there support for this conditional hypothesis? How do you know?

```
> m2 <- lm(obama ~ income + age + educ + female + black
+          + dem + rep + dem:income, data=obama)
> summary(m2)

Call:
lm(formula = obama ~ income + age + educ + female + black + dem +
    rep + dem:income, data = obama)

Residuals:
   Min     1Q Median     3Q    Max
-76.67 -11.64   3.05  12.73  69.79

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  61.666783   3.271568  18.849  < 2e-16 ***
income       -0.053484   0.012147  -4.403 1.13e-05 ***
```

```
age           -0.030398   0.030092  -1.010  0.31255
educ          -0.004112   0.210809  -0.020  0.98444
female         4.433373   0.993360   4.463 8.57e-06 ***
black         17.070766   1.226655  13.917  < 2e-16 ***
dem           10.504224   1.527455   6.877 8.34e-12 ***
rep          -16.009862   1.422543 -11.254  < 2e-16 ***
income:dem     0.067813   0.021063   3.219  0.00131 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 20.97 on 1849 degrees of freedom
  (465 observations deleted due to missingness)
Multiple R-squared:  0.3814,        Adjusted R-squared:  0.3787
F-statistic: 142.5 on 8 and 1849 DF,  p-value: < 2.2e-16
```

According to the results above, the coefficient for `income`, representing household income in thousands of dollars, is -0.0535. This means that for every $1,000 increase in household income, there is an approximately 0.0535 point decrease in Obama's feeling thermometer for respondents who self-identify as non-Democrats, holding all other factors constant. For respondents who self-identify as Democrats, there is a 0.0143 point increase for every $1,000 increase, again holding all other factors constant. Not only doess `income` remain statistically significant in this model, the included interaction term is significant when $\alpha = 0.01$. Although these two values have different signs, it appears that income has a larger effect on respondents who self-identify as non-Democrats. Based on this information, there is sufficient evidence to conlcude there is support for the conditional hypothesis. In other words, Democrats' feelings toward Obama are less influenced by their incomes than others' feelings are.

3. Does income have a statistically significant effect on the feelings toward Obama of those who aren't Democrats? On the feelings of Democrats? Report the estimated effect and $p$-value for each.

```
> obama$nondem <- 1-obama$dem
> m3 <- lm(obama ~ income + age + educ + female +
+            black + nondem + rep + nondem:income, data=obama)
> summary(m3)

Call:
lm(formula = obama ~ income + age + educ + female + black + nondem +
    rep + nondem:income, data = obama)

Residuals:
```

```
   Min     1Q Median     3Q    Max
-76.67 -11.64   3.05  12.73  69.79

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)    72.171007   3.438792  20.987  < 2e-16 ***
income          0.014329   0.018093   0.792  0.42847
age            -0.030398   0.030092  -1.010  0.31255
educ           -0.004112   0.210809  -0.020  0.98444
female          4.433373   0.993360   4.463 8.57e-06 ***
black          17.070766   1.226655  13.917  < 2e-16 ***
nondem        -10.504224   1.527455  -6.877 8.34e-12 ***
rep           -16.009862   1.422543 -11.254  < 2e-16 ***
income:nondem  -0.067813   0.021063  -3.219  0.00131 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 20.97 on 1849 degrees of freedom
  (465 observations deleted due to missingness)
Multiple R-squared:  0.3814,        Adjusted R-squared:  0.3787
F-statistic: 142.5 on 8 and 1849 DF,  p-value: < 2.2e-16
```

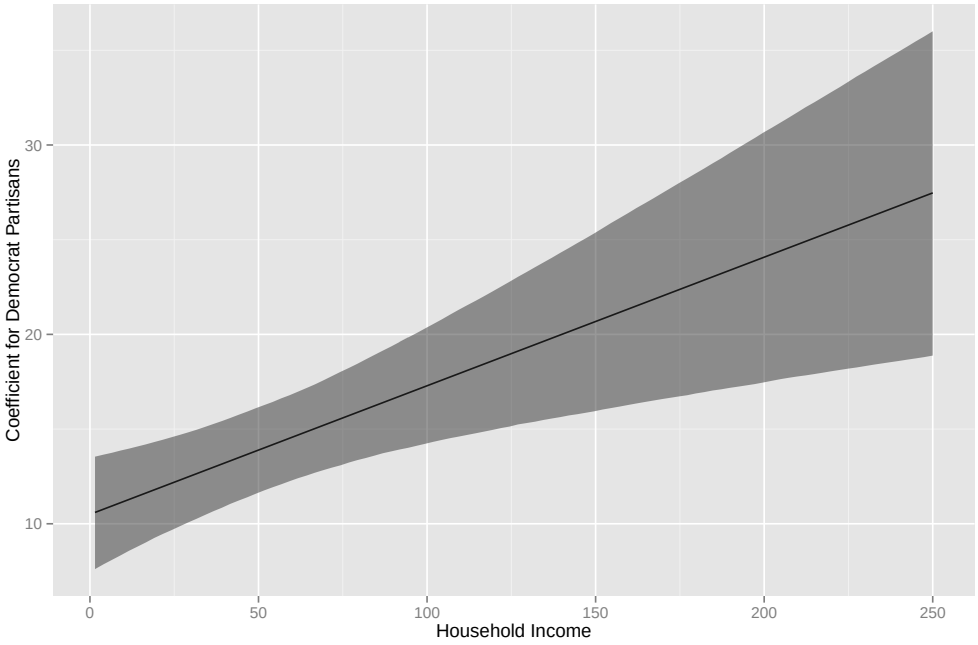
According to the results from question #2, for respondents who self-identify as non-Democrats, household income does have a statistically significant effect on their feelings towards Obama, with a coefficient value of -0.0333 and a $p$-value less than $2x10^{-16}$, well below $\alpha = 0.05$. This result implies that all the other factors remain constant. In order to answer the second part of the question, we had to change the reference level of the model, so now we can show the effect income has on respondents self-identified as Democrats. According to this new model, household income does *not* have a statistically significant effect on Democrats' feelings towards Obama, with a coefficient value of 0.0143 and a $p$-value of 0.4285, well above $\alpha = 0.05$, holding all other factors constant.

4. Suppose we were really more interested in how being a Democrat affects feelings towards Obama. What effect does income have on this effect? Graph your answer and insert the graph in your LaTeX file.

```
[1]  61.650052314  -0.053464096  -0.029886126  -0.003878492   4.422827025
[6]  17.066358972  10.496633740 -16.031033308   0.067896264
```

According to the figure below, it appears that household income has a positive effect on the coefficient `dem`. In other words, as income increases, the effect being a self-

Figure 1: Effect of Income on the Effect Being A Democrat Has on Respondents' Feelings Towards Obama



identified Democrat has on respondents' feelings towards President Obama increases as well, holding other factors constant.

Table 1: Linear Regression Results (Q1 and Q2)

| | *Dependent variable:* | |
|---|---|---|
| | Obama Feeling Thermometer | |
| | (1) | (2) |
| Household Income (Thousands) | $-0.033^{***}$ | $-0.053^{***}$ |
| | (0.010) | (0.012) |
| Age | $-0.035$ | $-0.030$ |
| | (0.030) | (0.030) |
| Education | 0.049 | $-0.004$ |
| | (0.211) | (0.211) |
| Gender | $4.485^{***}$ | $4.433^{***}$ |
| | (0.996) | (0.993) |
| Race | $16.766^{***}$ | $17.071^{***}$ |
| | (1.226) | (1.227) |
| Democrat | $13.768^{***}$ | $10.504^{***}$ |
| | (1.145) | (1.527) |
| Republican | $-16.718^{***}$ | $-16.010^{***}$ |
| | (1.409) | (1.423) |
| Democrat:Income | | $0.068^{***}$ |
| | | (0.021) |
| Constant | $60.203^{***}$ | $61.667^{***}$ |
| | (3.248) | (3.272) |
| Observations | 1,858 | 1,858 |
| $R^2$ | 0.378 | 0.381 |
| *Note:* | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 | |

Table 2: Linear Regression Results (Q3)

| | *Dependent variable:* |
|---|---|
| | Obama Feeling Thermometer |
| Household Income (Thousands) | 0.014 |
| | (0.018) |
| Age | −0.030 |
| | (0.030) |
| Education | −0.004 |
| | (0.211) |
| Gender | 4.433*** |
| | (0.993) |
| Race | 17.071*** |
| | (1.227) |
| Non-Democrat | −10.504*** |
| | (1.527) |
| Republican | −16.010*** |
| | (1.423) |
| Non-Democrat:Income | −0.068*** |
| | (0.021) |
| Constant | 72.171*** |
| | (3.439) |
| Observations | 1,858 |
| $R^2$ | 0.381 |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 |