POLI 171: Problem Set 3: Experiments

Due Thursday, February 8, at the beginning of class

For this assignment, you may collaborate with one other student. You may also complete this assignment on your own, if you prefer. If you collaborate, you and your co-author will turn in a single document, and you will each receive the same grade. You and your co-author are expected to work together on *all* parts of this assignment; do not simply divide the questions between yourselves.

Please type your answers and submit them as a hardcopy. Please show your work by including your `R` code and output in your answers.

The Oregon Health Plan and Access to Health Care

In 2008, the state of Oregon resumed enrollment in its Medicaid expansion program, Oregon Health Plan Standard (OHP), after a four-year freeze on new enrollments. Because the state lacked the resources to offer the OHP to everyone who wanted it, the state government decided to determine who would have access to the program through a random lottery.

You can find the data for this assignment on the TritonEd page for our course (Problem Sets $\rightarrow$ Problem Set 3). In this dataset, the variable `Selected` measures whether an individual was selected by the lottery to have an opportunity to enroll in the OHP; this is the treatment-assignment variable. The variable `OHP` measures whether an individual actually enrolled in the program; only some of the lottery winners ended up enrolling. In this homework, the Intention to Treat (ITT) effect is the effect of being selected by the lottery, while the Average Treatment Effect on the Treated (ATT) is the effect of participating in the OHP among the individuals who actually enrolled. An individual "complies" with treatment assignment if

- They participate in the OHP if they are selected by the lottery
- They do not participate in the OHP if they are not selected by the lottery

For this homework, use the `lm()` function to run regressions, and use the `coeftest()` function to run significance tests on your regression results:

```
mod <- lm(OUTCOME ~ TREATMENT, data=data)
coeftest(mod, vcov=vcovHC(mod, "HC2"))
```

where `OUTCOME` is your outcome variable, and `TREATMENT` is your treatment variable. You will need the `sandwich` and `lmtest` libraries. A coefficient is significant if its p-value is less than 0.05.

Variable	Description
person_id	Individual's ID number
OHP	Indicates whether the individual was enrolled in the OHP in 2008
Selected	Equals 1 if the individual was selected by the lottery to have an opportunity to enroll in the OHP, and 0 otherwise
Female	Equals 1 if the individual is female, and 0 otherwise
Age	The individual's age in 2008
White	Equals 1 if the individual identifies as White, and 0 otherwise
HospitalVisits08	The number of times that the individual visited a hospital during the last 6 months, as measured in the 2008 pre-treatment survey
Income07	The individual's household income in 2007, rounded to the nearest $\$2,500$
NeedUnmetMedicalCare09	Equals 1 if the individual was unable to get all of the medical care that they needed during the last six months, as measured in the 2009 post-treatment survey
PrimaryCareVisits	The number of primary care visits the individual had during the last six months, as measured in the 2009 post-treatment survey
Responded	Equals 1 if the individual responded to the 2009 post-treatment survey

Question 1: Compliance (3 points)

The variable `Selected` measures whether an individual was offered treatment, and the variable `OHP` measures whether an individual took treatment (i.e., enrolled in the Oregon Health Plan).

a) Use these two variables to calculate the compliance rate in the treatment group. This is the proportion of individuals assigned to treatment (`Selected == 1`) who ended up participating in the program (`OHP == "Enrolled"`).

b) Is the noncompliance in this experiment one-sided or two-sided? Justify your answer by referring to the compliance rates for both the treatment group and the control group.

Question 2: Balance on Pre-Treatment Variables (3 points)

The variables `Female`, `Age`, `White`, `HospitalVisits08`, and `Income07` were assigned and measured prior to treatment. If the treatment and control groups are truly equivalent groups, we should not expect to see statistically-significant differences between the two groups on these variables.

a) Test for balance on these pre-treatment variables. In a neatly-formatted table, report the treatment group means, the control group means, the differences between the treatment and control group means, and the p-values for each variable. Are there any significant imbalances between the two groups on these variables?

Option 1: Regress each variable in turn on the treatment-assignment variable (`Selected`). The regression model for the `Female` variable would be

$$Female = \alpha + \delta Selected$$

where δ gives you the difference in that variable between the treatment and the control group. In `R`, the code for this model would be `lm(Female ~ Selected, data=data)`.
Option 2: Use the `t.test()` function: `t.test(Female ~ Selected, data=data)`.

Question 3: The Effect of Treatment on Access to Health Care (6 points)

Now you are ready to analyze how the treatment affects an individual's access to health care.

a) The first outcome variable of interest is `NeedUnmetMedicalCare09`, which equals 1 if the individual reported in the 2009 follow-up survey that they had been unable to get all of the medical care that they needed during the last six months. Calculate the Intention to Treat (ITT) effect on the likelihood of having unmet medical needs by regressing this variable on `Selected`. Interpret this effect in words. Is the ITT effect statistically significant?

b) The second outcome variable of interest is `PrimaryCareVisits`, which is the number of times during the last six months that the individual went to a health care provider for primary care (day-to-day, non-specialized medical care). Calculate the ITT effect on the number of primary care visits, and interpret the effect in words. Is it statistically significant?

c) Use the estimated ITT on `PrimaryCareVisits` that you calculated in part (b) and the compliance rate in the treatment group that you calculated in question 1(a) to calculate the Average Treatment Effect on the Treated (ATT) of *enrolling* in the OHP on the number of primary care visits. You do not need to worry about significance for this subquestion; just estimate the effect by using the equation that we saw in lecture on Tuesday, January 30. Interpret the effect in words.

d) Based on these results, was the OHP program effective at increasing access to health care? Your answer should reference the results that you found in parts (a) (b), and (c) of this question, and it should discuss the statistical significance of the effects calculated in parts (a) and (b).

Question 4: Attrition (2 points)

The outcome variables used in Question 3 were measured with a post-treatment survey. However, only a fraction of the sample responded to this survey. There is a risk that non-response to the survey may be correlated with both treatment status and pre-treatment variables, which would imply that the treatment and control groups are no longer equivalent groups for the purposes of the analysis in Question 3.

a) The variable `Responded` is equal to 1 if the individual answered the post-treatment survey, and 0 otherwise. Test for whether the response rate was statistically different between the treatment and control groups by regressing `Responded` on `Selected`, or by using the `t.test()` function. Which group was more likely to respond to the survey? Is the difference statistically significant?