

Problem Set 4

Quantitative Political Methodology (L32 PS 363)

Instructions

1. Print out and write your WUSTL ID at the top of **each** page, and complete each exercise in the space allotted. You may attach extra pages if the space provided is not sufficient, but please indicate that you have done so below.

Pages attached: _____

2. Please **show your work** if possible. You may lose points by simply writing in the answer. If the problem requires you to use R, please include the code you used to get your answers. If you are not sure if work needs to be shown for a particular problem, please ask a TA or post a question on Facebook.
3. The various pages of your homework should be **stapled together** (no paper clips please). If pages are lost because of a lack of a staple, no credit will be granted for that portion of the homework.
4. This problem set is **due at the beginning of class on Wednesday, October 25, 2017**. No late assignments will be accepted.
5. Total available points for this homework: 40.

Question 1 (Total: 8 points)

Field experiments have become an important tool to understand various political phenomena. For example, “Getting Out the Vote in Local Elections: Results from Six Door-to-Door Canvassing Experiments” by Green, Gerbern, and Nickerson (2003) is an early field experimental work on political behavior. The paper is available at <http://onlinelibrary.wiley.com/doi/10.1111/1468-2508.t01-1-00126/full>. Read the abstract (and introduction if you wish) of the paper and answer the following questions.

- (a) (2 points) As succinctly as possible, what is the causal claim being made by the authors?

- (b) (2 points) What is the “treatment” (or predictor) variable?

- (c) (2 points) What is the outcome variable?

- (d) (2 points) What allows the authors to claim that their findings are causal?

Question 2 (Total: 16 points)

For the 2006 GSS, a comparison of males and females on the number of hours a day that the subject watched TV gave:

Group	N	Mean	St.Dev	SE Mean
Females	1117	2.99	2.34	0.070
Males	870	2.86	2.22	0.075

- (a) (10 points) Conduct all parts of a significance test to analyze whether the population means differ for females and males. Interpret the p-value, and report the conclusion for α -level = 0.05.
- (b) (2 points) If you were to construct a 95% confidence interval comparing the means, would it contain 0? (You can answer based on the result in (a), without finding the interval.)
- (c) (4 points) Do you think that the distribution of TV watching is approximately normal? Why or why not? Does this affect the validity of your inferences? Explain your answer.

Question 3 (Total: 10 points)

Imagine that the data above is changed as below (note the changed sample size). A comparison of males and females on the number of hours a day that the subject watched TV gave:

Group	N	Mean	St.Dev	SE Mean
Females	11	2.99	2.34	0.070
Males	16	2.86	2.22	0.075

Conduct all parts of a significance test to analyze whether the population means differ for females and males. Interpret the p-value, and report the conclusion for α -level = 0.05.

Question 4 (Total: 6 points)

Using data on the 2008 New Hampshire Democratic Party Primary, visualize the relationship between the proportion of voters for Howard Dean in the 2004 Democratic primary and the proportion of voters for Barack Obama in the 2008 Democratic primary. To get the dataset, type:

```
install.packages("faraway")
library("faraway")
data("newhamp")
help("newhamp")
```

In addition to the relationship between the support for Dean in 2004 and the support for Obama in 2008, we are also interested in whether two different voting systems — hand-counted and machine-counted ballots — matter. At a minimum, you have to do the following things in a single plot:

- Properly label titles and axes
- Set the ranges of the axes appropriately
- Use different colors and symbols to indicate the two ballot systems
- Include a legend that explains colors and symbols