Psych 205
Assignment 5
Due March 2nd at 9:00am

In this problem set, we will be using data from Evelyn Hooker's paper, "The Adjustment of the Male Overt Homosexual" from 1957. This paper was important historically in changing the DSM to remove homosexuality as a psychological disorder.

First, you should read the paper. It is in the assignment folder.

We will be analyzing data from Table IV, which contains expert ratings on a psychological adjustment, rated while blind to group (homosexual/heterosexual). We won't be using the same statistics as Hooker – instead, you'll be using a *permutation test* to look for differences between the groups. In this test, you will form the null hypothesis by permuting (randomizing) the group labels on individual responses. Remember that in class, we formed the null hypothesis sometimes through simulation. For this homework, the simulation is one where you *scramble labels* (homosexual/heterosexual) on data points.

Q1. [2pts, HELP] Enter the data from Table IV into a data frame. To do this, you should ensure that the data is in standard ggplot format, meaning one response per row, and a column for group. Note that this is <u>not</u> the format that the table is in, since for example there are 7 ratings of "2" in the heterosexual group, which should give 7 lines in your dataframe. DO NOT make your data frame in excel or enter these 7 responses by hand. Instead, use the *rep* function in R and format your R code in a way that makes it easy to detect errors.

Q2. [1pt, SOLO]. Print a summary of your data table and check against a friend's.

Q3. [2pts, HELP]. Using your data and R's *table* function, replicate Hooker's Table IV. (Note R's *table* function may defaultly leave off the "1" ratings since there are none, and that's fine)

Q4. [3pts, SOLO] Make a "publication-quality" plot of the mean ratings in each group. For this plot, put bars for the means, and individual scatter points for each rating using geom_dotplot (you may want to read the help files). Fiddle with the parameters of geom_dotplot to make it look nice (you may want to set binwidth, stackdir, alpha, and colors). Save your graph as a pdf in an aspect ratio that makes it look nice.

Our main task in this assignment is to implement a "permutation test" on the data. To do this, we will be creating a null hypothesis by random permuting (shuffling) labels (rows of the group column) on each rating. For each time you permute, you will compute a test statistic consisting of the difference between the simulated "groups" divided by their "pooled" standard error (below), where x and y are the samples of the two groups and *sem* is a function (we wrote as *f* in class) to compute a standard error.

```
sem <- function(x) { sd(x) / (sqrt(length(x))) }

test.statistic <- function(x,y) {
	(mean(x)-mean(y)) / sqrt(sem(x)^2 + sem(y)^2)
}
```

The test statistic here is really the same as the t-test we did in class – it's a mean (here, a mean difference) divided by a standard error of the difference (pooled since there are two groups).

Q5. [6pts, HELP] One reason to like a permutation test is that it leave the overall statistics of responses (mean, distribution, etc.) the same. Explain why. Does it leave the distribution of responses within each group the same or not? Explain why. Does it matter if we form the null distribution by shuffling the group labels or shuffling the responses? Explain why or why not.

Q6. [10pts, HELP] Write 3-4 sentences explaining at an undergraduate level why shuffling labels corresponds to a null hypothesis of "no difference" between groups. Think about the groups would look like if there was no difference, and what shuffling does.

Q7. [15pts, HELP] Write a loop that permutes the group labels, computes the test statistics on the permuted samples, and stores them in a vector called *permuted.stats*. Store 1000 of these permutations. Plot a histogram of these permuted differences with ggplot; if the default binning of the histogram doesn't look great, read about how to change it and change it to something reasonable. Place a vertical red line at the true (un-permuted) test statistic.

Q8. [2pts, SOLO] Why does the count on the histogram go so high, when there are so few subjects? What determines the count and is the count relevant to the inferences we'll make, why or why not?

Q9. [2pts, SOLO] What should the mean of the permuted test statistics be? Explain why. Does yours match?

Q10. [5pts, HELP] From *permuted.stats*, print a one-tailed p-value, testing the hypothesis that the homosexual group has higher ratings than the heterosexual.

Q11. [5pts, HELP] From *permuted.stats*, print a two-tailed p-value, testing the hypothesis that the groups differ.

Q12. [5pts, SOLO] Explain what the p-value for Q11 means at the level a college undergraduate could understand.

Q13. [10pts, HELP] You've probably heard that $p < 0.05$ is a threshold for getting published. Your p-values should not have been < 0.05 in the previous questions. Should this work have been published? Write a few sentences as you might in a review or discussion, explaining why or why not.