

WU #8 - Permuting

Math 154 - Jo Hardin

Thursday, September 30, 2021

Name: _____

Consider the MacNell experimental data.

Skeptic: the differences in averages are due to random variability (null hypothesis)

Advocate: the difference in averages are due to the identity of the TA (alternative hypothesis)¹

The null hypothesis that the gender **identity** of the instructor is the same with respect to the probability distribution of the course evaluations. Below are some possible ways to permute the data, which is correct?

1. Permute the identity variable
2. Permute the gender variable
3. Permute the gender variable after grouping by the identity variable
4. Permute the identity variable after grouping by the gender variable

Provide one possible permutation under your suggested permutation strategy (for the data provided). (That is, permute the variable(s) that should be permuted.)

```
## # A tibble: 20 x 3
## # Groups:   tagender, taidgender [4]
##   tagender taidgender overall
##   <int>      <int>    <dbl>
## 1         0         0        5
## 2         0         0        1
## 3         0         0        1
## 4         0         0        4
## 5         0         0        4
## 6         0         1        3
## 7         0         1        4
## 8         0         1        4
## 9         0         1        4
## 10        0         1        4
## 11        1         0        4
## 12        1         0        4
## 13        1         0        4
## 14        1         0        4
## 15        1         0        4
## 16        1         1        4
## 17        1         1        3
## 18        1         1        4
```

¹Note that there should generally be no causal claim / conclusion in the alternative. Usually, the conclusion is that of an association. **However**, here the data were collected under experimental conditions, so there is a possible causal claim if warranted by the data.

## 19	1	1	5
## 20	1	1	4

Solution

The structure of the permutation test will be to permute the identity variable after grouping by the gender variable. One possible permutation is:

```
## # A tibble: 20 x 4
## # Groups:   tagender [2]
##   tagender taidgender permTAID overall
##   <int>      <int>      <int>    <dbl>
## 1         0         0         0        5
## 2         0         0         0        1
## 3         0         0         1        1
## 4         0         0         0        4
## 5         0         0         1        4
## 6         0         1         1        3
## 7         0         1         0        4
## 8         0         1         1        4
## 9         0         1         1        4
## 10        0         1         0        4
## 11        1         0         0        4
## 12        1         0         0        4
## 13        1         0         1        4
## 14        1         0         1        4
## 15        1         0         1        4
## 16        1         1         1        4
## 17        1         1         0        3
## 18        1         1         0        4
## 19        1         1         0        5
## 20        1         1         1        4
```