Lab 1 Question 3: Are survey respondents who have had someone in their home infected by COVID-19 more likely to disapprove of the way their governor is handling the pandemic?

Yao Chen, Jenny Conde, Satheesh Joseph, Paco Valdez, Yi Zhang

Contents

Table 1: Cross Tab of Governor Approval and Positive COVID-19 Tests

	No Positive Test	Positive Test
Approve	0.599	0.020
Disapprove	0.366	0.015

1 Importance and Context

The COVID-19 has changed everything, and one of the most important aspects is how the campaigns for the US 2020 elections were conducted. The pandemic forced to cancel conventions and prompted many states to change how people get and submit their ballots, with all the uncertainty it created. And to make things worse, all of this happened in the middle of social distancing restrictions not seen since almost a century. It's essential to understand how elected officials' approval changed when people got in close contact with COVID-19.

2 Description of Data

The ANES data set contains information from 8,280 pre-election interviews with U.S. citizens of voting age. Two variables are particularly relevant for us to answer this question:

- `V201145: APPROVE OR DISAPPROVE R'S GOVERNOR HANDLING COVID-19`
- `V201624: ANYONE IN HOUSEHOLD TESTED POS FOR COVID-19`

```
summary(df_clean)
```

```
##       gov            covid
##  Min.   :1.000   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:2.000
##  Median :1.000   Median :2.000
##  Mean   :1.381   Mean   :1.965
##  3rd Qu.:2.000   3rd Qu.:2.000
##  Max.   :2.000   Max.   :2.000
```

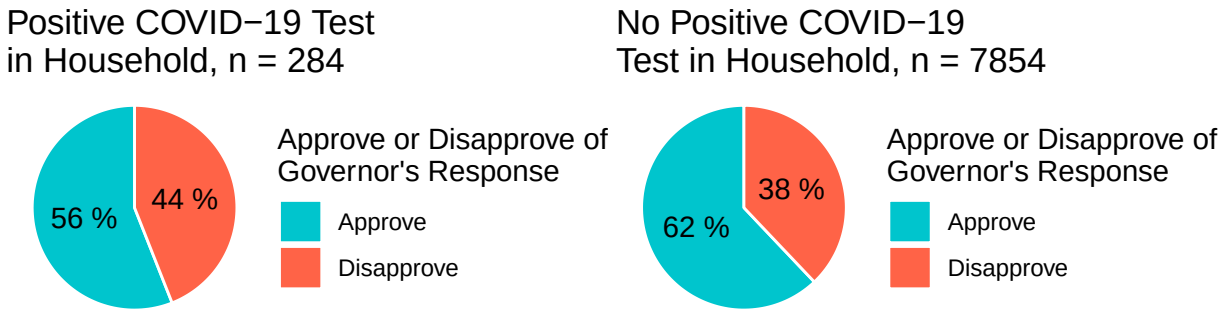

Figure 1: Distribution of Voters' Approval of Governor based on Positive COVID-19 Test in Household

3 Most appropriate test

First, we notice that the governor approval variable is a Boolean variable. There are only two valid answers: Approve, or Disapprove. As a result, a parametric test based on some underlying distribution that resembles Normal would not be appropriate.

At the same time, the groups (people who had a positive test in their household v.s. people who did not) are distinct people, and they don't have a natural pairing.

Furthermore, given the sampling frame based on a cross-section of registered addresses across 50 states and the District of Columbia, we feel the data are sufficiently close to be i.i.d.

Based on the above diagnose, the Wilcoxon Rank Sum test is the most appropriate in this case.

4 Test, results and interpretation

We establish the *null hypothesis* to be that the average support for the respondent's governor is the same among people that had a member of the household test positive for COVID-19 and those who did not. Given we have no strong initial inclination in either direction, this should be a two tailed test. We use the standard 5% significance level.

First take a look at the table in each of the 4 cases for context, then run the test.

```
table(df_clean$approval, df_clean$covid_tested)
```

```
##
##              Not positive Positive
##   Approve            4876      159
##   Disapprove         2978      125
```

```
wilcox.test(df_clean$gov ~ df_clean$covid)
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  df_clean$gov by df_clean$covid
## W = 1183267, p-value = 0.0377
## alternative hypothesis: true location shift is not equal to 0
```

From the test we can see that the p-value is 0.0377, which is less than the significance level $\alpha = 0.05$, meaning that we reject the null hypothesis in favor of the alternative that people with a positive COVID-19 test in their household have a different opinion of their Governor than people without.

Practically, we can calculate the correlation between the two variables.

As presented in Fig. 2, having a positive test in the household does linearly correlate to less likely to approve the Governor's handling of the pandemic. However given the magnitude of the the correlation coefficient. The linear relationship isn't very strong.

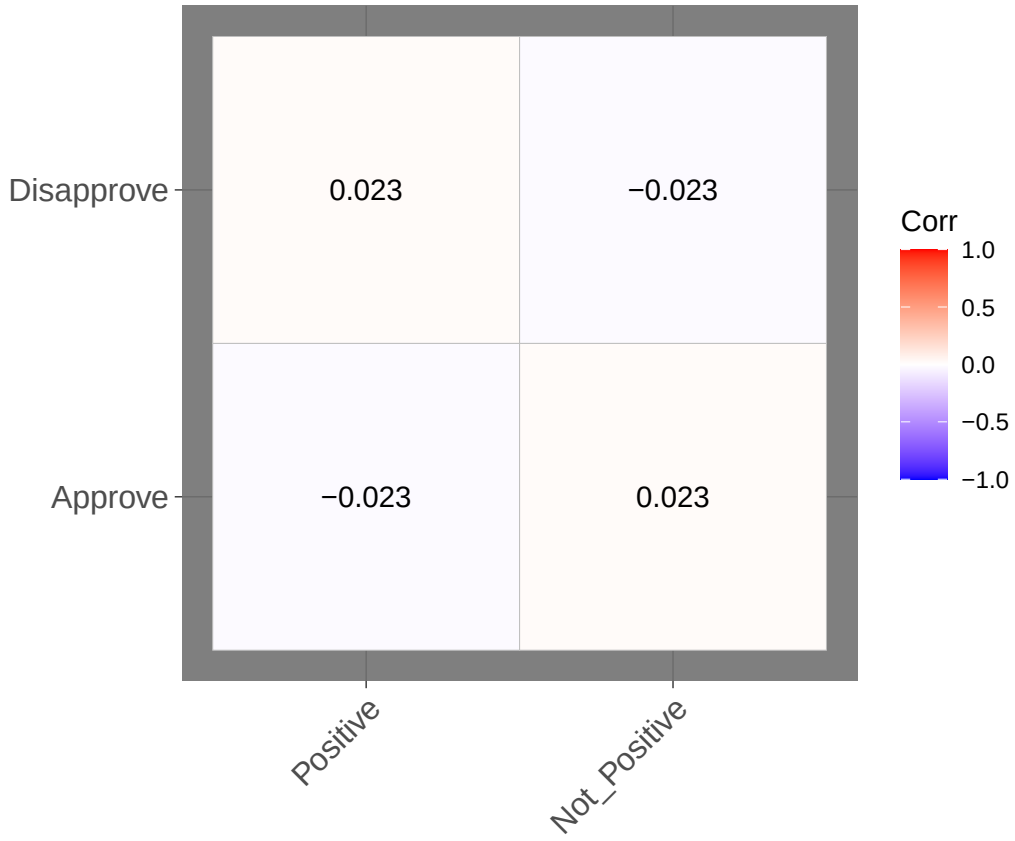

Figure 2: Correlation Table for Voters' Approval of Governor and Household COVID-19 Tests