

' 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?'

Da Qi Ren, Rui Li, Dicky Woo

Contents

1	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?	2
1.1	Description of Data	2
1.2	Importance and Context	2
1.3	Most appropriate test	3
1.4	Test Limitations	5

1 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?

1.1 Description of Data

It is often the governors who make a bigger impact on people's everyday lives. The coronavirus pandemic has shined a spotlight on the role of these leaders, who have largely been responsible for leading their states' responses to the public health crisis. With little guidance from the federal government, state governors have decided the rules on the shutdown of nonessential businesses and the timelines for reopening. They've been responsible for getting personal protective gear into the hands of medical workers and first responders, as well as expanding coronavirus testing in their states. And they're the ones people look to, to find out exactly how their state is doing and what they can expect during this time of uncertainty.

To 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? We will address this question using data from the 2018 American National Election Studies (ANES). This is an observational dataset, based on a sample of respondents drawn from the YouGov platform. I use data items

- V201624 group of people "Covid-19 infected" and "Not infected"
- V201145 approval or disapproval
- V201146 strong or normal (approval or disapproval)

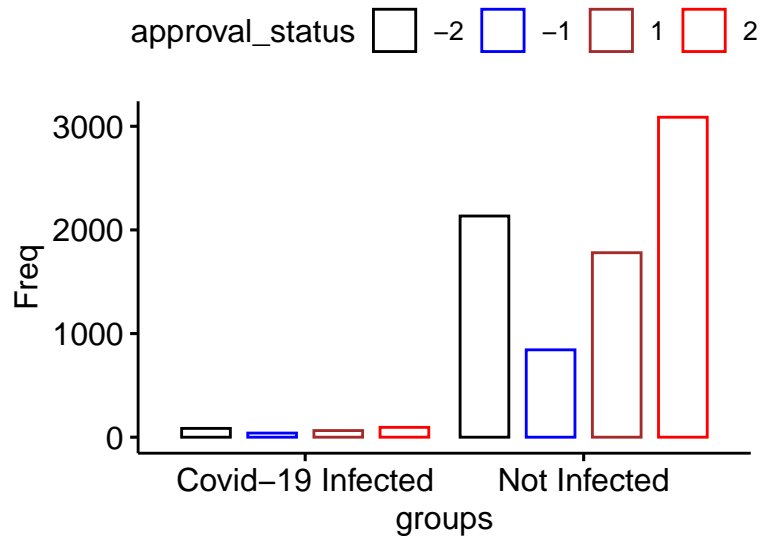
The first thing to note in this data there are differences in the rank ordering of strong disapproval and disapproval, approval and strong approval, or normal approval/disapproval. The next figure plots the relative prevalence of ANES respondent's feelings of approval and disapproval. I combine V201145/V201146 as "approval_status", with rank of -2, -1, 1, 2 to represent strong disapproval, disapproval, approval and strong approval, respectively.

1.2 Importance and Context

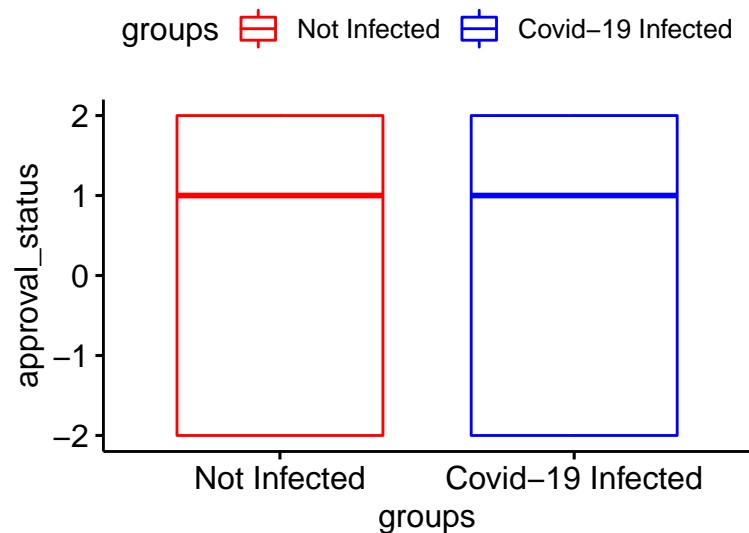
COVID-19 has affected every state, but no two governors have handled it in exactly the same way. Some, like Gov. Andrew Cuomo in New York and Gov. Gavin Newsom in California, have instituted long-term stay-at-home orders and mandatory shutdowns of nonessential businesses. Others, like Gov. Kristi Noem of South Dakota and Gov. Asa Hutchinson of Arkansas, have resisted placing sweeping restrictions on travel and business operations. In February, The approval rating of California Governor Gavin Newsom plummeted, and this situation may exacerbate the possibility of his removal. A poll conducted by the Berkeley Institute of Government on more than 10,000 registered voters in California found that Newsom's work performance was only approved by 46%, a sharp drop from his 64% approval rate in September last year. In addition, Newsom has been widely criticized for the previous strict stay-at-home order. Report pointed out that 36% of California voters currently support the removal of Newsom, and most of them support the Republican Party. They are dissatisfied with Newsom's performance in handling the epidemic. In addition, less than one-third of the interviewees said that the governor had "excelled" in responding to the epidemic. In recent weeks, the number of new confirmed cases of new coronary pneumonia in a single day in California has declined. According to statistics from the state, only 12,000 new confirmed cases were reported on the 2nd, while in December last year, the highest number of new confirmed cases in a single day exceeded 50,000.

While governors' responses to the public health emergency have varied, one thing is true across all 50 states: Constituents are keeping a careful eye on how authorities are addressing COVID-19, and judging them accordingly.

```
##               approval_status
## groups          -2    -1     1     2
## Covid-19 Infected   85   40   64   95
## Not Infected       2134  843 1780 3087
```



```
## # A tibble: 2 x 4
##   groups      count  mean    sd
## * <chr>      <int> <dbl> <dbl>
## 1 Covid-19 Infected   284  0.155  1.70
## 2 Not Infected      7844  0.362  1.69
```



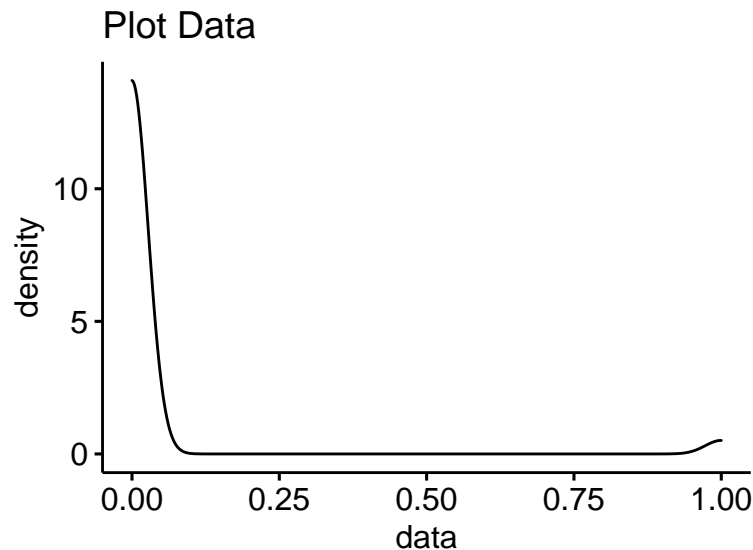
1.3 Most appropriate test

An underlying assumption for appropriate use of the tests described was that the continuous outcome was approximately normally distributed or that the samples were sufficiently large (usually $n_1 > 30$ and $n_2 > 30$) to justify their use based on the Central Limit Theorem.

In detail, to find the most appropriate test, I did the check test as the follows:

- Hypothesis 1: Are the two samples independent? Yes, because the samples from Covid-19 infected and not infected have nothing to do.
- Hypothesis 2: Does the data of each of the two groups obey a normal distribution? I use the function of `shapiro.test()` to calculate the Shapiro-Wilk test for each set of samples. In the Shapiro-Wilk normality test result, $p\text{-value} < 2.2e-16$, which is less than the significance level of 0.05, indicating that the distribution of the two sets of data is significantly different from the normal distribution.

```
##
## Shapiro-Wilk normality test
##
## data: approval_status[groups == "Covid-19 Infected"]
## W = 0.78207, p-value < 2.2e-16
```



- Hypothesis 3: Do these two populations conform to the homogeneity of variance? I use F test to test the homogeneity of variance with the `var.test()` function. The F test results $p = 0.9084$. It is greater than the significance level $\alpha = 0.05$. Therefore, there is no significant difference between the variances of the two sets of data. Therefore, we believe that the variance of the two groups of men and women is equal (homogeneity of variance).

```
##
## F test to compare two variances
##
## data: approval_status by groups
## F = 1.0076, num df = 283, denom df = 7843, p-value = 0.9084
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.8574517 1.1998611
## sample estimates:
## ratio of variances
##          1.007593
```

Due to the combination of the above three assumptions, also because the data does not conform to a normal distribution, we cannot use the student-t test. A nonparametric test is appropriate, particularly two independent samples Wilcoxon rank sum test.

I then test whether 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. The modules on hypothesis testing two independent samples, i.e. a group has family members infected, and another group without infected. A nonparametric test to compare outcomes between two independent groups is the Wilcoxon Rank Sum Test. The parametric test compares the means between independent groups. In contrast, the null and two-sided research hypotheses for the nonparametric test are stated as follows:

- H0: Someone in their home infected by COVID-19 more likely to disapprove the governor.
- H1: Someone in their home infected by COVID-19 has equal probability to disapprove the governor.

A one-sided research hypothesis is used and, thus, the research hypothesis H0 indicates that Someone in their home infected by COVID-19 more likely to disapprove of the governor. The procedure for the test involves pooling the observations from the two samples into one combined sample, keeping track of which sample each observation comes from, and then ranking lowest to highest from, respectively.

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: approval_status by groups
## W = 1038950, p-value = 0.9786
## alternative hypothesis: true location shift is greater than 0
```

- Interpretation of results. In the above result, the p value is the significance level of the Wilcoxon test (p-value = 0.9786). The p-value for the test is 0.9786, which is greater than the significance level $\alpha = 0.05$. We can conclude that who have had someone in their home infected by COVID-19 more likely to disapprove of the way their governor is handling the pandemic.
- Practical Significance. I use Pearson's r as a useful descriptor of the degree of linear association between two variables. In the result below, t is the t-test statistic value ($t = -2.0291$), df is the degrees of freedom ($df = 8126$), p-value is the significance level of the t-test (p-value = 0.04248). sample estimates is the correlation coefficient (Cor.coeff = -0.023). The p-value of the test is 0.042, which is less than the significance level $\alpha = 0.05$. We can conclude that the Covid-19 infection and governor approval are medium correlated with a correlation coefficient of -0.023 and p-value of 0.042.

```
##
## Pearson's product-moment correlation
##
## data: df0$groups and df0$approval_status
## t = -2.0291, df = 8126, p-value = 0.04248
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.044222374 -0.000763539
## sample estimates:
## cor
## -0.02250359
```

1.4 Test Limitations

We have conducted this test based on the data available in the ANES. However, even after an appropriate analysis of the data on hand, it shows only a medium correlation to support the conclusion that who have had someone in their home infected by COVID-19 more likely to disapprove of the way their governor is handling the pandemic. People disapprove their governor could because of many different reasons. If this were to be the case, then our study design couldn't find all the causes. Use more data items to analysis could be helpful to give a more accurate result about the Covid-19 impections.