Lab 1: Question 2

Jot, Leena, Jessica

Contents

Table 1: Likelyhood to vote in 2020 election

Likelyhood	Freq
Extremely likely	0.79
Inapplicable	0.06
Moderately likely	0.03
Not likely at all	0.01
Slightly likely	0.01
Very likely	0.09

0.1 Importance and Context

Are Democratic voters more enthusiastic about Joe Biden or Kamala Harris?

The 2020 US presidential election will be remembered in history as one of the most highly divisive and consequential presidential elections of the early 21st century. At the forefront of voter's minds was handling of a global pandemic and a social rights movement - both issues driven by current political climate. In the presidential office, was Donald Trump, a Republican who was running for reelection. The opposing Democrats had Joe Biden as their presidential candidate, running with Kamala Harris for vice president. Democratic voters hailed Joe Biden as a leader of the previous Democratic administration - who would potentially bring the nation back to 'normal'. Kamala Harris was seen by Democrats as one to redefine the role of black woman in leadership.

For Democrats, it is important to understand the party's enthusiasm for their candidates in order to drive Democratic voter turnout. Is one of their candidates more effective at bringing Democrats to the polls? The answer to this question will provide guidance for the campaign leading up to the election. If there is less enthusiasm for one candidate, the campaign can focus on improving that person's image. Similarly, the candidate with more enthusiasm can act as the face of the ticket.

0.2 Description of Data

In order to answer this question, we will use data drawn from the 2020 American National Election Studies (ANES). This survey data was collected prior to the 2020 national election. Questions asked of respondents include their political party, their likelihood to vote in the election, their enthusiasm for Kamala Harris, and their enthusiasm for Joe Biden.

We controlled for democratic voters by filtering for those who are actually registered as Democrat. We did not want to include data from Republican, Independent, or Unregistered respondents to align with our research question.

We also identified votes as those who responded to the survey question on likelihood of voting. Examining Table 1, we see that respondent's likelihood of voting range from 'Extremely Likely', with a frequency of 79%, to 'Not likely at all', at only 1%. 6% of respondents indicated that this question was not applicable to them. While someone living in American territories, such as Puerto Rico, have the right to participate in party primaries, they are not eligible to vote in the presidential election. Therefore, we interpreted that if the respondent had any response other than 'Not likely at all' or 'Inapplicable' then they were considered a voter.

The survey data prompts the respondent to score their satisfaction with candidate Joe Biden, then prompts them to score their satisfaction with Kamala Harris. These scores range from 0 to 100. The higher the score, the more satisfied they are in the candidate. Each respondent is given both questions to answer.

Our research question is interested in finding out if there a difference in enthusiasm towards the two candidates. According to the histogram in figure 1, the difference in enthusiasm takes on a distribution around 0. That means that we need to remove respondents who have equal enthusiasm in order to assess only those who showed a difference in opinion.

After subsetting, we are left with a sample size of 988 to conduct our analysis.

Enthusiasm is a subjective measure and there is no methodology to standardize a definition for scoring candidates. Therefore, to find out if there is a difference in enthusiasm towards the two candidates, we need to create a new variable. This variable is binomial, a value of 1 for if the respondent has a higher enthusiasm for Kamala Harris and 0 for Joe Biden. According to the bar chart in figure 1, there is an equal number of respondents who is more enthusiastic about one candidate over the other. We will use this binomial field to determine if we have statistical evidence for a difference in enthusiasm.

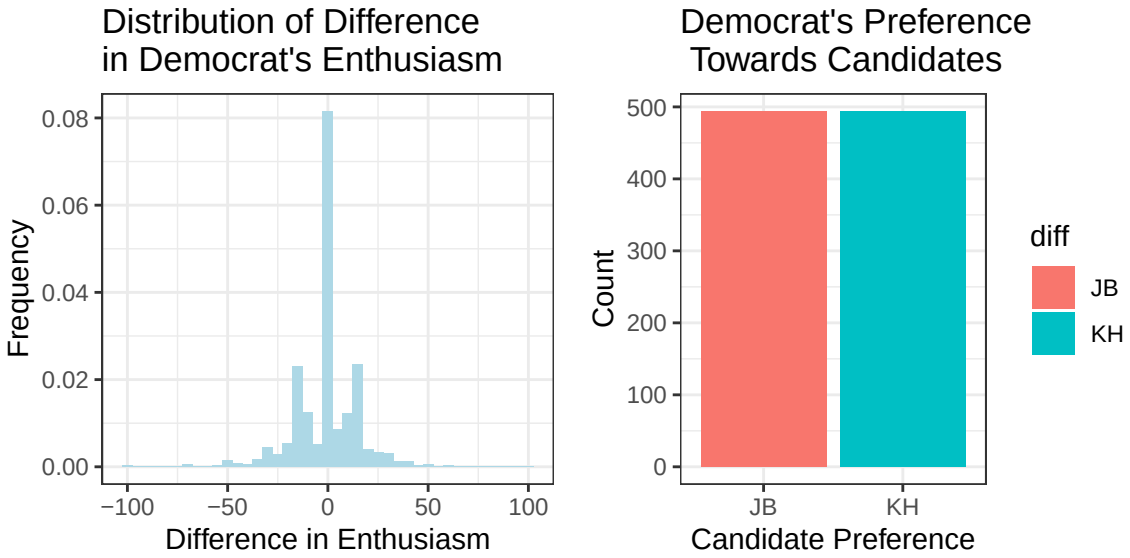

Figure 1: Democratic Voter Enthusiasm for Party Candidates

0.3 Most appropriate test

We chose to use the Binomial test to answer the research question because of two assumptions. Our data is on an ordinal scale since we are comparing average frequency of a subjective enthusiasm for one candidate over the other. The other assumption to be true is that our sample is independently and identically distributed (iid).

We can examine the sample's iid in the sample design and respondent recruitment for the survey. The survey was conducted three ways: self-administered online survey, live video interviews, and telephone interviews. It is identically distributed in that each respondent is drawn from the population of interest, the US citizens aged 18+. Independence is met in that sampling occurred from a random draw from the USPS delivery addresses across the US. The mode of survey was also computer randomized.

The null hypothesis of our test is that there is no difference in enthusiasm for our candidates. We will use an alpha of 0.05, meaning our test must produce a p-value less than 0.05 in order for us to reject the null.

If we reject the null hypothesis, then our data would suggest that there is a difference in candidate enthusiasm. We will conduct a two tailed test. If our confidence interval is closer to 1, then enthusiasm for Kamala Harris is higher than Joe Biden, and if the confidence interval is closer to 0, then enthusiasm is higher for Joe Biden.

```
(Q2_Biom <- binom.test(nrow(Q2[Q2$diff=='KH',]), nrow(Q2),alternative = "two.sided"))
```

```
Exact binomial test
```

data: nrow(Q2[Q2$diff == "KH",]) and nrow(Q2) number of successes = 494, number of trials = 988, p-value = 1 alternative hypothesis: true probability of success is not equal to 0.5 95 percent confidence interval: 0.4683562 0.5316438 sample estimates: probability of success 0.5

0.4 Test, results and interpretation

The binomial test results in a p-value of 1, which is larger than our alpha of 0.05. Therefore, we fail to reject the null hypothesis, the data does not suggest any evidence of difference in voter enthusiasm for Kamala Harris compared to Joe Biden. This finding is supported by our figure 1, where there is an equal number of respondents preferring either candidates.

Although there is no practical significance to examine with a when we fail to reject, we would like to point out that it would have been logical to expect a difference in enthusiasm. In the democratic primaries, Kamala Harris and Joe Biden were campaigning against each other for the party nomination for president. It was publicly known that their policies did not align at the time, so we would expect there to be a difference of enthusiasm by public opinion.

A future research question to better measure this would be to understand the difference in enthusiasm based on the survey respondent's likelihood of voting. If the respondent is already extremely likely to vote, then it can be assumed that they have a similar favorability of both candidates representing their party. Since 80% of the respondents indicated that was the case, as shown in table 1, then their opinion is driving our test results. In understanding the enthusiasm of those less then 'Extremely likely' to vote, we can recommend how to campaign the democratic candidates in order to influence more Democratic people to vote.