Probability of Election Result

Dylan Vener

Introduction

In 2007, Ross Valley, an area in California, voted on the issue of a flood tax. Of the 15,010 ballots issued, 8,059 ballots were returned. Disregarding 36 ballots which were disqualified for various reasons, 3,938 people voted for the resolution, and 4,085 people voted against the tax, what would be a narrow loss. Importantly though, the ballot had an unusual feature, to be counted, it had to be signed, something which was not required in the past.

The introduction of this caused 1,672 ballots to be disqualified, 730 votes in favor, and 942 against, tipping the scales to 3,208 pro, and 3,143 con, giving the election to those in favor of the election.

The question of note in this analysis is the likelyhood of such a disparity in the number of disqualified votes, given that one would expect, and hope, that the additional requirement of signing the ballot would not have undue harm on a particular faction.

Analysis

```
#Setting the seed to ensure consistent results
set.seed(19120823)
#Number of signed votes in favor
s.pro <- 3208
#Number of signed votes against
s.con <- 3143
#Number of unsigned votes for
u.pro <- 730
#Number of unsigned votes against
u.con <- 942
```

After loading the values into R, I went on and did my analysis, which ultimately was surprisingly simple.

```
#Total Probability of voting in favor of the tax, disregarding signing
t.prob <- (s.pro+u.pro)/(s.pro+s.con+u.pro+u.con)
```

Here I calculate the probability that any given person who returned a ballot returned one in favor of the tax, making no distinction between those who signed and those who did not.

```
#Simulation of 1672 peolple not signing their ballot, run 10,000,000 times, with t.prob of voting in fa
u.sim <- rbinom(10000000,u.pro+u.con,t.prob)
```

And here I simulate 10,000,000 elections, keeping constant the number of people who do not sign their ballots, using the null hypothesis that either party is equally likely to sign.

Conclusion

```
#Percent of times in the simulation that a result as extreme or more was observed
(sum(u.sim<=u.pro)/10000000)*100
```

```
## [1] 0.00045
```

This shows the percentage of times in our simulation that a result as, or more, extreme then the actual result that was observed. As this shows, it is *extremely* unlikely for this to have occured given the assumption provided, that both sides of the vote were equally likely to not sign the ballot, and disqualifying their vote. With such a low P value, the only reasonable course of action is to reject that nully hypothesis.

Thus, we must presume that there is some factor related to which side of the vote one was, which affects the probability of signing the ballot, and thus legitimizing their vote. ## Code

```
#Because of the use of only default functions, this code should work universally
#Setting the seed to ensure consistent results
set.seed(19120823)
#Number of signed votes in favor
s.pro <- 3208
#Number of signed votes against
s.con <- 3143
#Number of unsigned votes for
u.pro <- 730
#Number of unsigned votes against
u.con <- 942

#Total Probability of voting in favor of the tax, disregarding signing
t.prob <- (s.pro+u.pro)/(s.pro+s.con+u.pro+u.con)

#Simulation of 1672 peolple not signing their ballot, run 10,000,000 times, with t.prob of voting in fa
u.sim <- rbinom(10000000,u.pro+u.con,t.prob)

#Percent of times in the simulation that a result as extreme or more was observed
(sum(u.sim<=u.pro)/10000000)*100
```