CS109: Contest

Influences on Brazilian Representatives' Votes

gdasilva

June 2017

1 Summary

In this paper, I present a way to calculate the different influences the Brazilian legislators may be subject to while in Congress. I also present the relevance that has for my country and for a project I've been working on to fix the political misrepresentation in Brazilian politics. Though all the technical ideas are of low complexity, I hope you enjoy seeing how simple concepts can generate some quite interesting results. I for sure did!

2 Motivation

2.1 The purpose

The students in the Brazilian Students Association at Stanford have been working on a mobile app to fix the misrepresentation often present in our country's politics. The app basically allows users to give their opinions (”Agree” or ”Disagree”) on proposals and on their legislators' votes, and notifies users whenever their legislators vote on a new bill. We've always been sure that we can do cool things with the data we gather through and for this project, but we haven't thought much about what those cool things can be. Hence, the CS 109 Contest was an opportunity for me to brainstorm about what these cool things can be [1].

2.2 The problem

Talking about our idea to a GSB student who had worked with Brazilian government, we heard that politicians' votes in Congress don't usually prioritize population's opinions, but rather other interests such as the interests of the parties they belong to and the interests of certain companies (that may sponsor or bribe legislators).
The work I present here sets up a framework to explore this relationship more closely. Depending on the results this work yields, we'll be able to:

- point out what each legislator's behavior depends more heavily on, which itself has significant relevance in Brazil;
- create a self-reinforcing tool for the app we've developed. If we are able to show users that if more of then engage with the app, then their opinion has a heavier impact on their legislators' vote, this can be a powerful tool to engage them even more;
- create a cascade of studies on how various factors may affect representative's votes.

3 Methodology

3.1 Data

It's important to make it clear that most of the data is made up. For example, the tables used with companies', parties', populations' and representatives' votes was created by me, with the purpose

[1] Although the work on the app has been done by a team for a few months, all the work on probability presented in this write-up was done by myself only after the contest was out

of testing the algorithms I present. Real data would demand research and time that I unfortunately didn't have during weeks 9 and 10.

3.2 Algorithms

I used Bayes' Theorem for all calculations. The idea is simple. In order to explain that, let's name a few variables:

$$R_i = \text{The ith representative's vote on a proposal}$$
$$P_i = \text{The orientation that the ith representative's party takes on a proposal}$$
$$C_j = \text{The orientation that the jth company takes on a certain issue}$$
$$U = \text{The orientation that the users of the app take on a proposal}$$

For each random variable X, we calculate $P(R_i = 1|X = x)$, that is, the probability that representative R_i will vote "Yes" given the value of X. It's important to note that we don't differentiate proposals. We consider $P(R_i = 1)$ to be the same for all proposals. Now we explain how we calculate $P(R_i = 1|X = x)$ in more detail given each X.

3.2.1 Parties' influence

I count the number of "Yes"'s (P_y) and the number of "No"'s (P_n) that a given party gave on a set of proposals. I also count the number of times that a representative voted "Yes" AND his/her party had orientation "Yes" (N_y), and the number of times that a representative voted "Yes" AND his/her party had orientation "No" (N_n). Hence,

$$P(R_i = 1|P_i = 1) = \frac{N_y}{P_y}$$

$$P(R_i = 1|P_i = 0) = \frac{N_n}{P_n}$$

I didn't explicitly calculate $P(R_i = 0|P_i = p)$ because that's just $1 - P(R_i = 1|P_i = p)$.

3.2.2 Companies' influence

Just as an explanations of how companies' influence may come into play: imagine a bill that forbids cutting trees commercially in a certain area has come to Congress. Companies that produce paper from those trees would be completely against that bill. Legislators whose campaign was sponsored by those companies or who received bribes could potentially vote "No" to defend the interests of these companies.

The way I calculate companies' influence is very similar to parties' influence except for two major differences:

- Companies don't necessarily have to have an orientation on all proposals (or if may not be possible to infer their orientation with much certainty). Hence, we discard the proposals on which a company doesn't have a particular orientation.
- There's no direct "belonging" relationship between a legislator and a company. Hence, I do the same calculation I did for parties for **all** pairs of legislators and companies. This way, I can find values that that may point to potential associations between specific legislators and companies.

3.2.3 App users' influence

Currently, there is no way through which citizens can communicate their opinions on a certain proposal to influence legislators' decisions, which is one of the roles the app our Association developed plays. I thought that it would be interesting to track the influence the users of our app may have on legislators' votes. Again, the way I calculated the probability of a representative's vote given users' opinions is very similar to what I did for parties' influence expect for one major difference:

- I felt that it would be important to take into consideration the **number** of people who gave a certain opinion, which can't be accounted for with an indicator variable (which is binary). I hence distributed users' opinion in 7 different degrees that vary from "Strongly Disagree" to "Strongly Agree." The bucket in which each case fits can be easily obtained through the ratio of people who were "In favor" of a certain proposal over the total number of opinions on a certain proposal (i.e. # "In favor" + # "Against"). Then, we can consider a ratio from 0.0 to 0.1429 as "Strongly disagree", one from 0.1429 to 0.2858 as "Moderately Disagree", and so forth, up to 0.8571 to 1.0, which would be "Strongly Agree". I omitted the algorithm for calculating that "bucketing" from my code.

I find this influence specially powerful. While the other influences are relevant for pointing out which legislators may have votes biased by private interests, users' influence can track how much our app impacts politics through time. If the results we get for these probabilities start to increase over time, that would imply that our users' engagement have caused legislators to start keeping their constituents opinions in mind while voting. If we show such results to our users, that can be a really powerful tool to engage them even more (it's like saying "Your opinion actually matters! Check how much it has influenced your representative's votes. Engage with the community to foster this even more"), thus reinforcing the process .

4 Conclusion and Next steps

The work presented here may already greatly enhance the relevance of what's presented in the app the Brazilians at Stanford have developed, and it will hopefully impact Brazilian politics as a whole. The first step from here is to gather real data with which we can get real results. We've already built some scrapers to gather representatives' votes and parties' orientation, so that's already a great place to start. We should now also gather some companies' orientation on different proposals to support our second type of influence.
I believe some of the possible next steps are:

- Finding other types of influence. The influences I came up with (parties', companies' ad users') are a good place to start, but we should think of others. Examples of what we may consider are: the political moment (number of protests, for example), industries' orientation (which may or may not be more relevant than a single company's orientation), and so forth.

- Getting more users. Our data for the third type of influence (users') will only be relevant once we have a large userbase, so we should put a lot of effort in making more people download and engage with the app.

- Making predictions. With the data we've gathered and with some of the knowledge we've learned in 109, we can already start making predictions about representatives' votes, and that can be really interesting.

Thanks for the opportunity of coming up with all this through the contest, it was really valuable for me.

5 Acknowledgements

Just wanted to say I'm thankful for: Michela, the TA who helped me to come up with the idea for different degrees of agreement for the users and who told me about FiveThirtyEight (https://fivethirtyeight.com/, which I haven't looked closely at yet, but will); the Brazilians with whom I've been working on the app; my roomate, Karan Sighal, for discussing some ideas that never actually came to life in this paper but that were crucial to get momentum.