Ranked Choice Modelling

Matthew Q. Knipfer

MTH-4910

Preliminary Progress Report

Introduction

This semester, I have been studying the current state of modelling ranked choice election data and building a model of my own that attempts to overcome some of the various deficiencies. Ultimately, this resulted in trade-offs for predictive power and geometric feasibility, but nevertheless the axiom system I created has explanatory power for the ways voters behave functionally.

Ranked Choice Voting

This method of voting is less popular in the US, only being used in California (and recently tested in Maine this past midterm for a House seat), but has the potential to disrupt for the better our polar political system. Plurality, our current model, trends strongly towards a duopoly, whereas ranked choice voting better aids the development of third parties. Instead of having one single non-transferrable vote, voters rank the candidates from most to least favorable. The ballots are sorted by first choice, and the candidate in last place is eliminated, with their votes being redistributed to the next options on the ballot. This continues until one candidate has >50% of the remaining votes.

Because of the redistribution, this means that voters are less concerned about picking who they think has the best chance at winning, and can be honest about whom they want without fear of having cast a wasted vote. Take for example the candidate Bernie Sanders. Although he was not in the final national race, there were Democrat voters who preferred him, but voted for Hillary in the primaries because they didn’t think he had a chance. This mismatches the will of the people with the vote they cast. By having the losers’ ballots redistributed, people can vote for someone and feel like their will is preserved in the backup options.

Two problems

Dr. Graham-Squire first proposed I try to find anomaly situations through modelling the voter data, which could lead to some visual or mathematical indicator that something weird would happen. As I was reading the literature, a paper by Tideman and Matje[[1]](#footnote-1) caught my attention in how it was trying to describe the voters and candidates with an axis-free space that used bisectors to separate the candidates. Voters were ideological points on this space and would rank the candidates based on their distance. Voters who were between two candidates almost exactly would be indifferent or nearly so between them. I disagreed with this model in some ways, although I thought it was insightful, so I decided to follow it instead of the anomaly situations this semester.

Axiom system

I believed that where the Tideman-Matje model failed was creating an axiom system to describe non-voters and partial voters (people who ranked some, but not all candidates). Instead of establishing a bisector system, I redefined the region to incorporate all features including ideology and enthusiasm, and I defined a boundary region around each candidate within which a voter that fell there would add them to their ballot. They might not be listed first, but they would be put on the ballot. This took care of the non and partial voters.

I next dealt with the issue of knowing where exactly these voters are. We don’t have the capability to survey them all to know where their point is exactly, and even if we could survey them, how would we know we asked the objective questions to make this space perfectly defined. Instead of following the route of points, I decided to treat each voters as a particle wave in which they affect the area the candidates have control over, but they are not marked as points in the end. So long as a ballot contains a candidate’s name in it, it contributes to the size of control that candidate has, increasing the boundary region defined as the area of voters they have been listed by.

Thus, the axiom system I characterize the Tideman-Matje model to use is:

* Candidates and voters exist as points in an axis-free space.
* Candidates are placed in an equilateral triangle.
* Voters are placed in the space based on their closeness of ideology to each candidate.
* Voters rank the candidates by shortest to greatest distance.
  + The crossover line between two candidates is defined as the bisector of the line between two candidates.
* Along the bisector’s width, voters are indifferent between the candidate on either side.

And, my system is as such:

* Candidates and voters exist as points in an axis-free space.
* Distance in this space between a voter and a candidate is defined as a composite of favorability in ideology, enthusiasm, and other all other factors.
* Around each candidate exists a circular boundary of radius rn within which a voter will add that candidate to their ballot.
* Voters rank candidates that appear on their ballot in the order of least to greatest distance from them to each candidate.
* Voters exist as a point in this space, but are only observed as a contribution to area.

Calculations

To create my model, take three candidates, A,B,C. Place A at the origin, B along the x-axis, and C above and between them. The axes are only there to help graph this space. The axes do not actually exist in the space. We will be defining this region by the distances between the center points of circles A, B, and C.

We add up all of the ballots that contain each candidate’s name and that comprises the area of each circle. Then, we calculate the radius. Using the radius of the circles and the overlap between two circles (for example the AB + BA + ABC + BAC+ BCA + ABC + ACB votes), we can calculate the distance between the center points. *This function will displayed during the presentation*.

By repeating this method three times, once for each relationship of A to B, B to C, A to C, we can get the distances between all three center points. We then have all we need to construct the three overlapping circles.

Lastly, by using the law of Sines, we can find the angle between the x-axis and A to C, projecting that down and across to get the cartesian coordinates we will represent C to be at. Then, we plot the circles at their respective points.

Results

The regions lined up surprisingly well. The circles were not missing a triple overlap contact for any of the data sets. However, the percent error of the triple overlap seemed to have suffered, varying in two cases from 50% to 24%. This could be either a calculation error of the overlap or a genuine issue in the model.

Future Work

This model, in my opinion, lacks predictive power, but can help to describe the radicalism or conformity of candidates in this kind of voting method. Maybe it could be used with polling, although I think polling statistics will take a hit in accuracy under this less stable and more diverse system. Nevertheless, I believe the circular boundary description and wave particle interaction are insightful for how we view voting outcomes from a pragmatic level due to the lack of knowledge on where people stand objectively.

In the future, I think this model could take some more axioms like cutting off the third listed candidate since ABC is congruent to AB. Having C on there functionally does nothing, but it does affect my model. I also wish to try out creating a new round cutoff for ranked choice voting that doesn’t just eliminate candidates on first place votes, but possibly the square of the candidates left in the field. That way, more moderate candidates gain an edge. In the end, I know that there is no perfectly just system, and it is all based on a series of wants and desires I have been conditioned to have for my democracy, but having another model that passes fairness criteria in another way is an addition to the toolbelt of democracy.

1. Tideman, T. N., & Matje, T. (n.d.). *A Statistical Characterization of Rankings from Political Survey Data* [Scholarly project]. Retrieved 2018. [↑](#footnote-ref-1)