Mathematics of Gerrymandering

Weifan Jiang, Namyoung Kim, Alex Robkin, Leo Segovia, Faculty Mentor: Chris Hoffman, Graduate Mentor: Tejas Devanur
Washington Experimental Mathematics Lab

Undergraduate Research Symposium 2018

Motivation

Gerrymandering refers to the manipulation of congressional district boundaries to favor a particular political party. Our goal is to determine how likely it is that a given state's map is gerrymandered. We began with the state of Iowa. To do this, we approximate and analyze the distribution of all possible redistrictings and determine how similar the current district map is to samples from this distribution.

We consider:

- How to quantify state and federal laws for congressional redistricting.
- How to sample from such a large distribution (Iowa alone has $\approx 4^{99}$ potential redistrictings).

Redistricting Requirements and Energies

We use the The Metropolis Hastings (M.H.) algorithm to produce a random walk that has as its unique stationary distribution, the ideal probability distribution p, of all possible districting plans for a given U.S state.

For Iowa, our model incorporates the redistricting requirements as "energies", measuring the compactness of districts and their population variation. The total energy is a linear combination of these energies.

Population energy:

$$\sum_{\text{Districts}} \left(\text{District Pop.} - \frac{\text{State pop.}}{\text{Number of districts}} \right)^2$$

Compactness energy:

$$\sum_{\text{Districts}} \frac{(\text{District perimeter})^2}{\text{District area}}$$

We begin the random walk with an arbitrary sample from our distribution. We then choose a similar sample randomly and accept with a certain probability. If we choose not to accept, we stay at the current sample. This process repeats for several iterations.

Accept candidate with probability:

$$\min\left(1, \frac{\exp(\text{weighted sum of current energies})}{\exp(\text{weighted sum of candidate energies})}\right)$$

Simulation Statistics

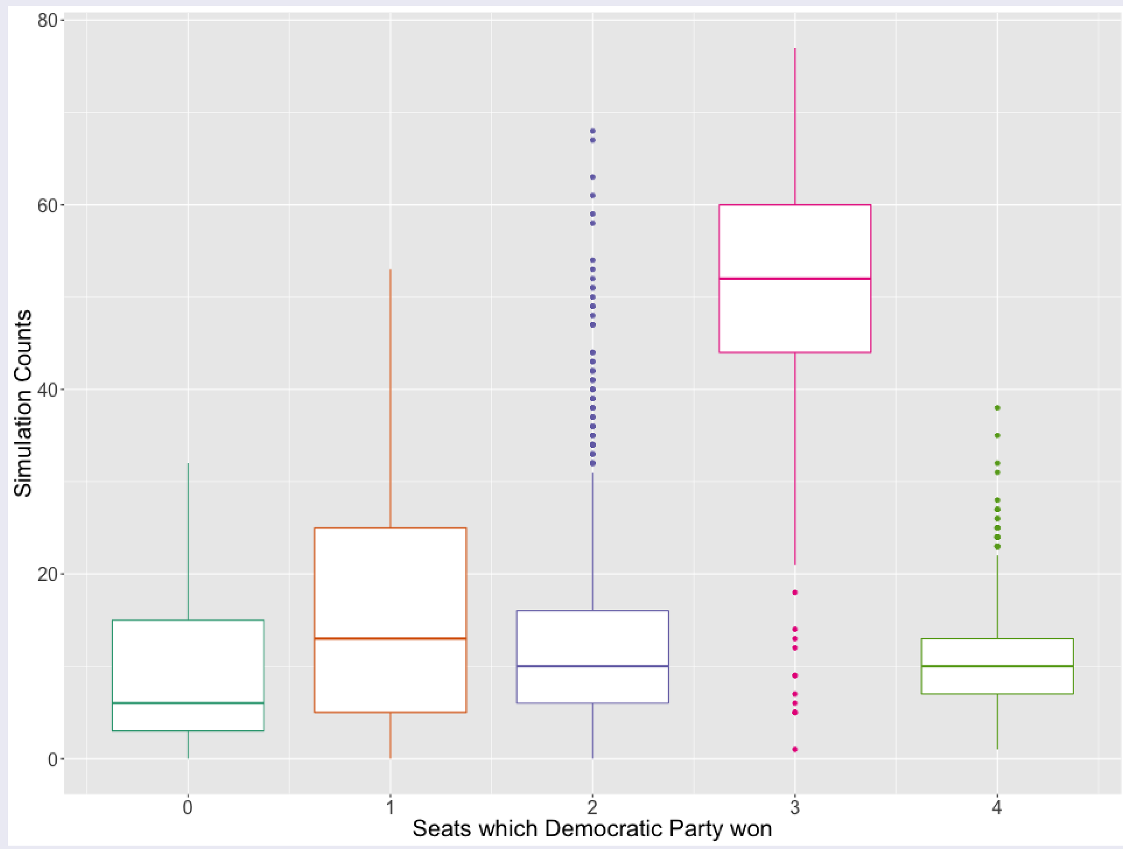

Figure: Box plot of 100 election samples for 1000 redistricted maps each. For seats 0, 1, 2, 3, 4, averages are 9.105, 15.867, 12.726, 51.493, 10.809 respectively and standard deviations are 7.57954, 12.34146, 10.216, 11.71167, 4.94459.

Actual Iowa District Map with Election Simulation

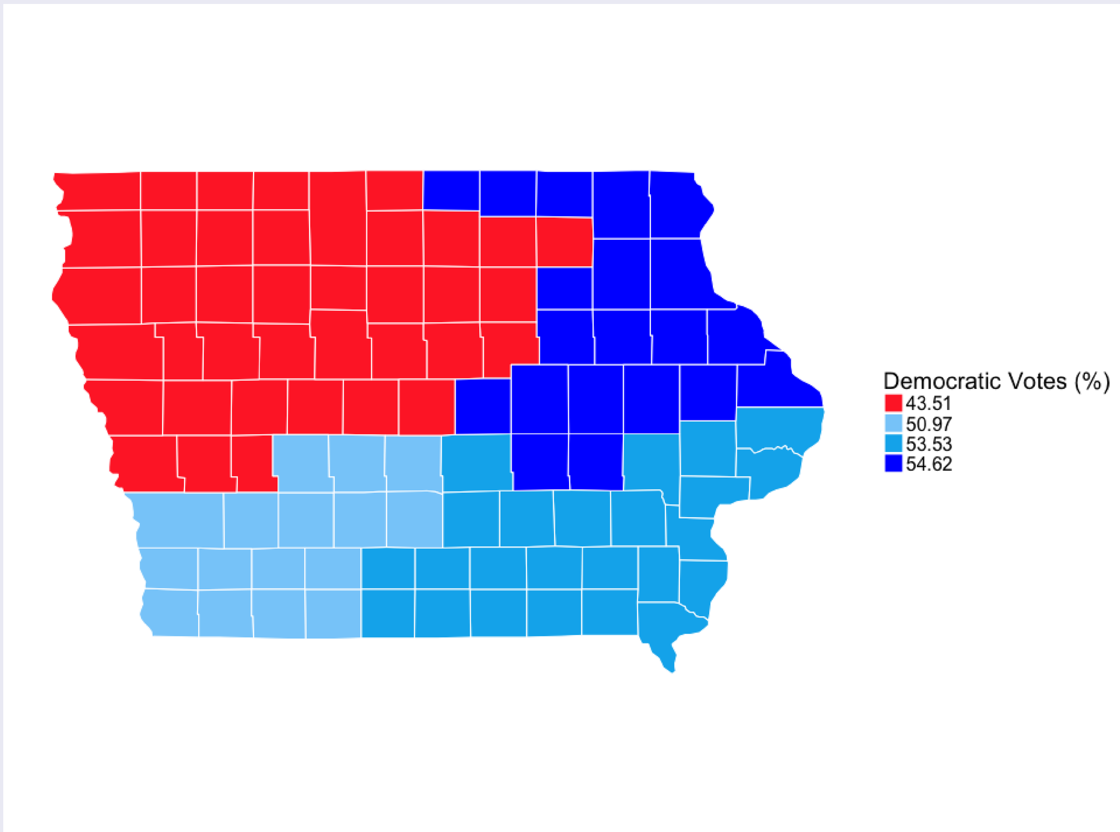

Figure: Iowa's Congressional districts since 2013. For 1000 election simulation, Democratic party got 0 seat 194 times, 1 seat 43 times, 2 seats 99, 3 seats 475, 4 seats 189. This map is one example of 3 seats case.

Redistricted Maps and Election Simulation

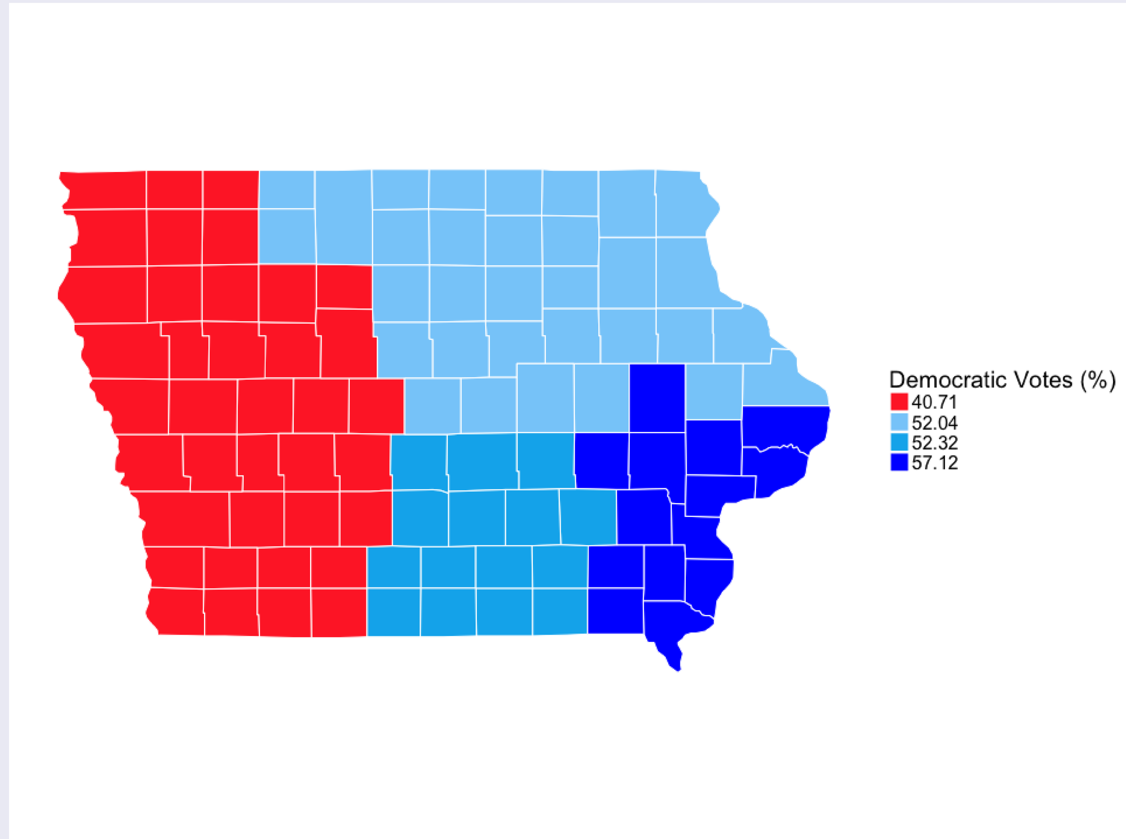

Figure: For 1000 election simulation, Democratic party got 0 seat 54 times, 1 seat 171 times, 2 seats 138, 3 seats 550, 4 seats 87. This map is one example of 3 seats case.

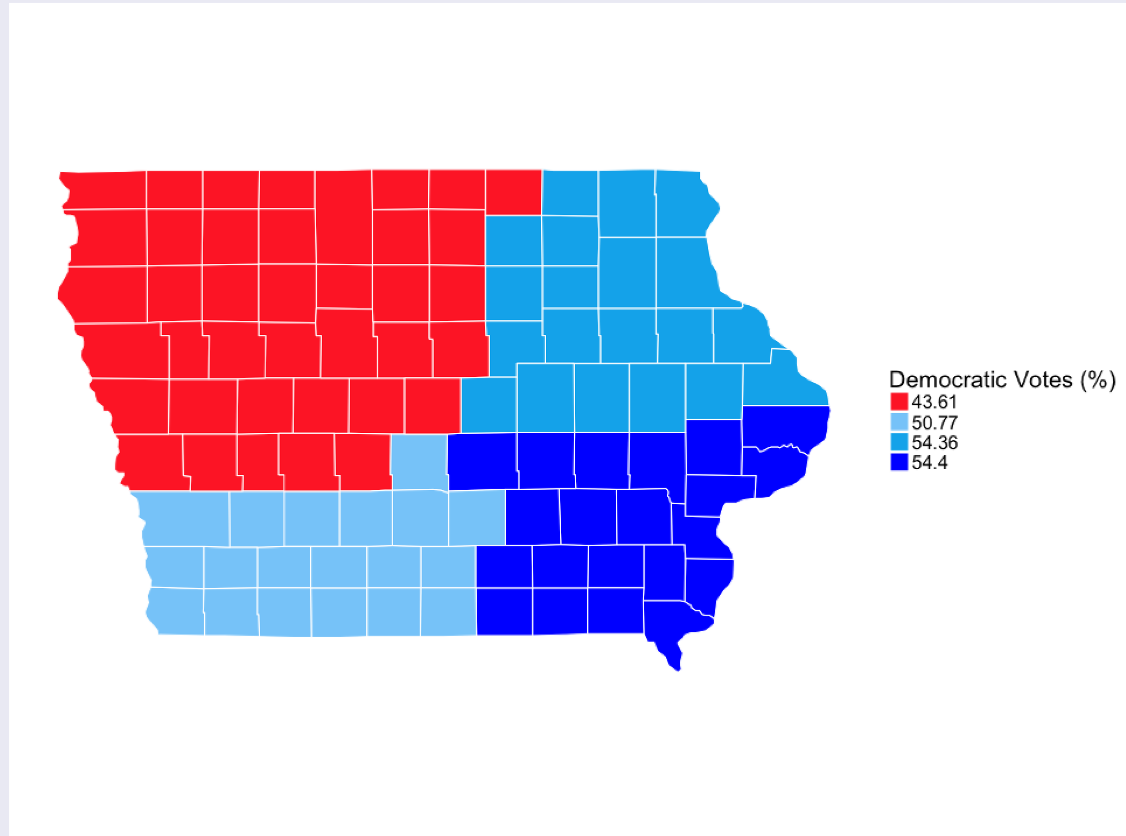

Figure: For 1000 election simulation, Democratic party got 0 seat 231 times, 1 seat 28 times, 2 seats 161, 3 seats 432, 4 seats 148. This map is one example of 3 seats case.

Evaluation of samples and Optimization

After we generated samples of redistricting plans, we wanted to make sure that we sampled correctly. Which means that:

- The final shape of districts must stay reasonably-compact (no S-shaped, no strips)
- The sum of the difference between each district's actual population and average district population must be within 4% of Iowa's total population

We want the samples to be independent from the initial map. We consider the following events as indicators of independence:

- Whether two counties belong to the same district (we considered Story county and Tama county).
- Whether there is a land-locked district (not touching state boundaries)

The independence tests are quantified by calculating "if such event happens in initial map, is it likely to happen in the final sample?"

We decided to use simulated-annealing to preserve both independence and quality. The first half of the random walk has low weights of energies to move away from the initial sample. During the second half, parameters increase to move towards samples with increasing quality.

Election Simulation

We generated a multivariate normal distribution based on the past three presidential election results for Iowa counties. We used these results to simulate congressional elections on 1000 maps generated by the M.H, algorithm, recording the average number of seats won by Democrats in these elections.

References

Bangia, S., Dou, B., Guo, S., Mattingly, J., & Vaughn, C. (n.d.). Quantifying Gerrymandering.

This poster is made with support from the Washington Expiremental Mathematics Lab at the University of Washington at Seattle.