

A Naive Bayes Framework for Estimating the Competence of an Institution Applied to the United States Congress

November 25, 2019

1 Introduction

In this note I introduce a Naive Bayes framework for estimating the competence of an institution in representing the interest of the public, for the simple case of an institution which makes binary decisions exclusively informed by independently voting representatives of the public. Making the assumptions 1) that the competence of a representative in representing the public interest can be approximated by a function of the public approval and disapproval ratings of the representative and 2) that public approval and disapproval ratings of a representative can be approximated by the public approval and disapproval ratings of the political party of the representative, I then use Gallup's Party Images polling and CSPAN voting records to estimate the competence of the chambers of the United States Congress in representing the public interest over successive six-month periods from 1993 to the present.

2 Naive Bayes Framework for Estimating the Competence of an Institution

An institution can be defined as acting in the public interest when making a decision if it makes the decision most likely to be in the public interest given the available evidence. Under that definition, the competence of an institution in representing the public interest over some time period can be approximated by the agreement between the decisions actually made by the institution during that time period and the decisions which were most likely to be in the public interest given the available evidence. Calling the decision most likely to be in the public interest given the available evidence the "epistemic decision" as shorthand, under this approximation an institution which always made the epistemic decision would have a competence of 1 for the time period and an institution which never made the epistemic decision would have a competence of 0 for the time period.

If the evidence available to inform each institutional decision is an observed voting profile of a body of representatives of the public, then the competence of an institution over a time period is the agreement between the decisions actually made by the institution during that time period and the decisions most likely to be in the public interest given the observed representative voting profiles. Therefore, if I can define a method for determining the epistemic decision given an observed voting profile of a body of representatives of the public as evidence, then, by comparing the actual institutional decisions to the epistemic decisions determined by the method, I can estimate the competence of an institution in representing the public interest over a time period.

In order to define a method for determining the epistemic decision given an observed voting profile of a body of representatives of the public as evidence, and therefore be able to estimate the competence of institutions which make binary decisions informed by voting profiles of bodies of representatives of the public, this section first reviews how Bayes theorem can be applied in a binary setting to estimate the probability a decision is correct given a set of evidence, then reviews how an observed voting profile can be used as evidence.

2.1 Bayes Theorem for Binary Decisions

Assume a decision must be made between two options - call them option 0 and option 1. Assume one of the options is "correct" - call it option X . Assume the decision is to be informed exclusively by a set of evidence v , which implies both options are equally probable prior to considering v . The respective posterior probabilities of option 0 and option 1 being correct having considered evidence v are then the respective probabilities of observing the evidence v given option 0 and option 1 being correct relative to the total probability of observing the evidence v .

$$P(X = 0|v) = \frac{P(v|X = 0)}{P(v|X = 0) + P(v|X = 1)}$$
$$P(X = 1|v) = \frac{P(v|X = 1)}{P(v|X = 0) + P(v|X = 1)}$$

Where $P(v|X = 0)$ is the probability of observing the evidence v given the correct option being 0 and $P(v|X = 1)$ is similarly the probability of observing the evidence v given the correct option being 1. In general, the probability option x is the correct option given observed evidence v is

$$P(X = x|v) = \frac{P(v|X = x)}{\sum_x P(v|X = x)}$$

Where the sum is over all possibilities for the correct option. The value of x for which $P(X = x|v)$ is maximized is the option most likely to be correct given the available evidence v .

2.2 Voting as Producing Evidence

Assume a body of N voters is to produce the evidence meant to inform a decision between two options - call them option 0 and option 1. Assume one of the options is correct - call it option X . By Bayes theorem, The probability that the x th option is correct given an observed voting profile $\vec{v} = [v_1, v_2, \dots, v_N]$ as evidence is

$$P(X = x|\vec{v}) = \frac{P(\vec{v}|X = x)}{\sum_x P(\vec{v}|X = x)}$$

Where $P(\vec{v}|X = 0)$ is the probability of observing the voting profile \vec{v} given the correct decision being option 0 and $P(\vec{v}|X = 1)$ is the probability of observing the voting profile \vec{v} given the correct decision being option 1.

If the voters vote independently conditional on the correct decision such that $P(v_i = u|v_j = w, X = x) = P(v_i = u|X = x)P(v_j = w|X = x) \forall (i, j, x, u, v) i \neq j$ and the i th voter is assumed have competence p_i such that $P(X = 1|v_i = 1) = P(X = 0|v_i = 0) = p_i$ and $P(X = 0|v_i = 1) = P(X = 1|v_i = 0) = 1 - p_i$, then the probability of observing the voting profile \vec{v} given the correct choice being the x th option is

$$P(\vec{v}|X = x) = \prod_{i=1}^N (x - p_i - v_i)(-1)^{v_i}$$

And the probability that the x th option is correct given the voting profile \vec{v} as evidence simplifies to

$$P(X = x|\vec{v}) = \frac{1}{1 + \prod_{i=1}^N \left(\frac{1 - p_i - v_i}{p_i - v_i} \right)^{-x}}$$

For a given a set of voter competence parameters \vec{p} , the value of x which maximizes this equation is the option most likely to be correct given an observed voting profile \vec{v} as evidence.

3 Application to the United States Congress

The chambers of the United States Congress make institutional decisions by majority rule. Generally, the decisions are between taking and not taking an institutional action. For each action considered by a chamber, each representative of the chamber votes Yea or Nay; if a majority (or supermajority) of the representatives vote Yea, then the proposed action is taken; if not, then the proposed action is not taken.

Congressional voting profiles and the resulting institutional decisions are available on CSPAN's website [1] [2] for each roll-call vote since 1989. If I assume vectors of representative competencies for each vote, then, by applying the method defined in the previous section, I can use the congressional voting profiles of a time period to estimate the corresponding epistemic decisions, and I can compare those epistemic decisions to the institutional decisions actually made in order to estimate the competencies of the chambers of Congress in representing the public interest during that time period.

It is, of course, not possible to know the actual values of the parameters; there is no way of knowing exactly how competent a representative is in representing the public interest. In lieu of having the resources necessary to determine reasonable estimates empirically, the best I can do is make heuristically reasonable assumptions as to what the parameters might be approximated by, then interpret the resulting estimations as approximations particular to the assumptions.

This section first goes over the competence parameter assumptions I made and how they might be improved upon, then uses the assumed parameters to estimate the competencies of the chambers of the U.S. Congress over successive six-month periods from 1993 to the present.

3.1 Representative Competence Assumption

I assume the competence of a representative in representing the public interest can be approximated by a function of the approval and disapproval ratings of the representative. For simplicity, I take the function which maps ratings to competence to be the logistic of the difference between public approval and disapproval. This mapping assures that a representative who is more approved than disapproved by the public will be more likely than not to vote in the public interest and that representative competence increases diminishingly as the difference between representative approval and disapproval increases.

Unfortunately, approval polling of individual representatives is sparse; generally it only occurs when they are running for office. Consequently, I'm forced to assume that the public approval and disapproval ratings of a representative can be approximated by those of the political party of the representative. This assumption allows me to use the long-running Gallup Party Images poll [3] to estimate the competencies of Democratic and Republican representatives at the time of each congressional vote from 1993 - the first year the poll ran - to the present. Because the public frequently elects representatives based on their party affiliation instead of their individual characteristics, and because representatives frequently vote along party lines instead of according to individual convictions, I feel that this assumption is heuristically reasonable.

Applying these two assumptions, I take the competencies of Democratic and Republican representatives for past votes to be

$$p_D = \frac{1}{1 + e^{-(\text{approval}_{D,t} - \text{disapproval}_{D,t})}}$$

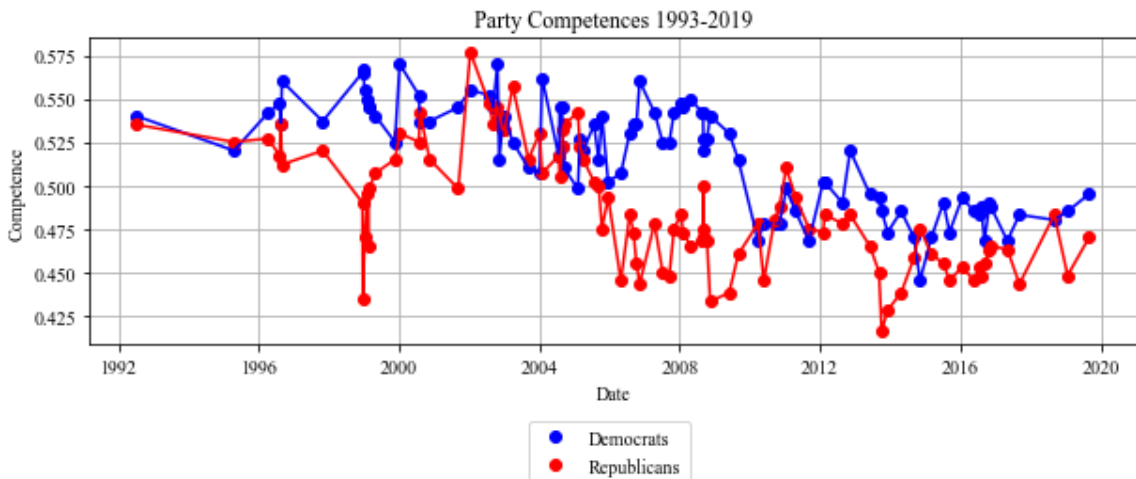
$$p_R = \frac{1}{1 + e^{-(\text{approval}_{R,t} - \text{disapproval}_{R,t})}}$$

Where $\text{approval}_{D,t}$ is the decimal-percent approval rating of the Democratic party as reported on date t , the closest date to the vote a poll result was reported; $\text{disapproval}_{D,t}$ is the decimal-percent disapproval rating of the Democratic party as reported on the same date; and $\text{approval}_{R,t}$ and $\text{disapproval}_{R,t}$ are similarly the decimal-percent approval and disapproval ratings for the Republican party. For example, polls concluding on September 15th 2019 reported the Democratic party as having decimal-percent approval and disapproval ratings of .48 and .50, respectively, and the Republican party as having decimal-percent approval and disapproval ratings of .43 and .55. Consequently, for dates of votes closer to September 15th 2019 than to any other date polls concluded, I take the competencies of Democratic and Republican representatives, p_D and p_R , to be

$$p_D = \frac{1}{1 + e^{-(.48 - .50)}} = .495$$

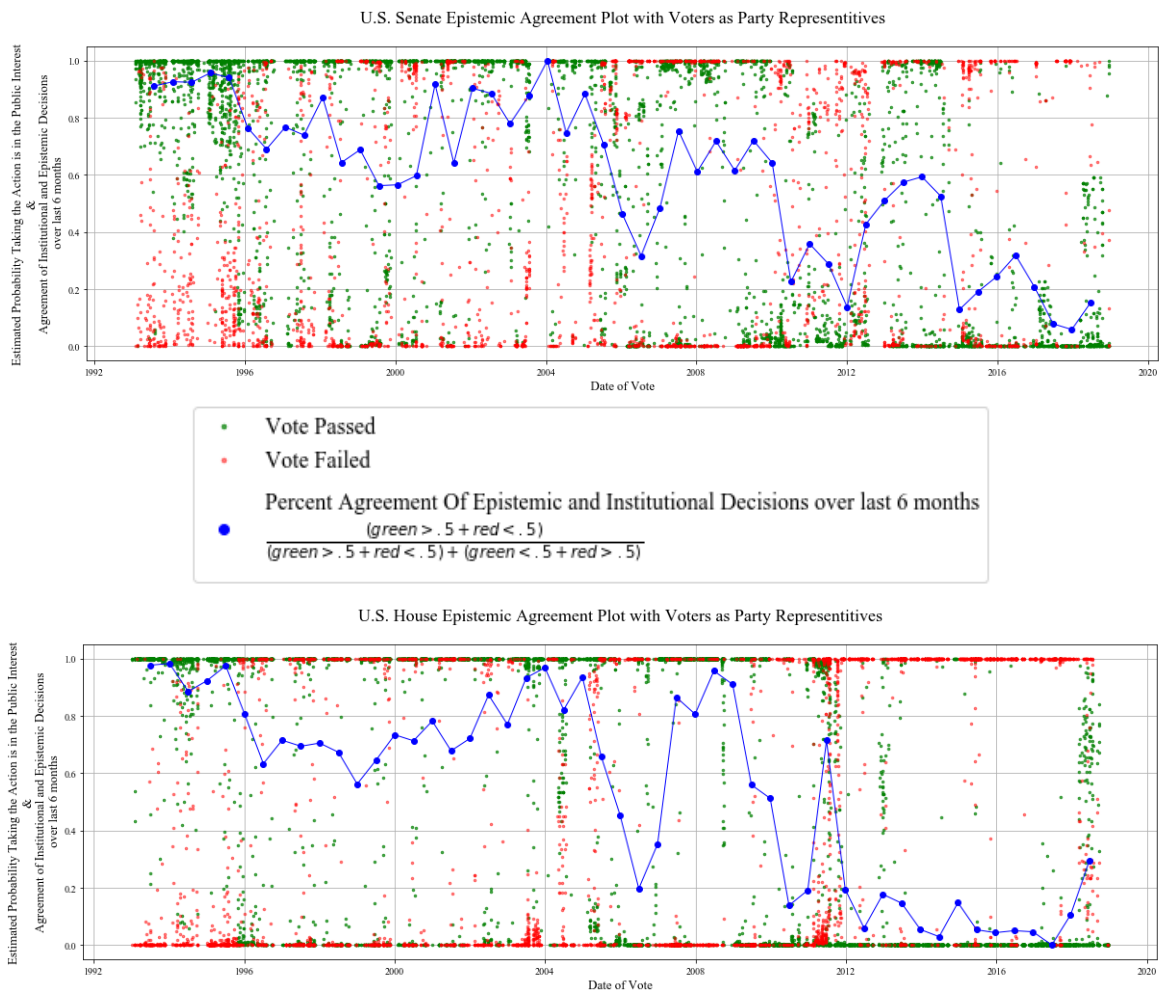
$$p_R = \frac{1}{1 + e^{-(.43 - .55)}} = .470$$

The figure below displays the competencies of Democratic and Republican representatives from 1993 to the present, calculated using the above method. These are the values which I take as representative competencies while producing the institution-epistemic agreement plots displayed in the next section.



This competence parameter assumption could be improved by fitting a trend line to the poll results then taking the competencies from the exact date on the fit line instead of from the date of the nearest reported poll result. The assumptions could be further improved by conducting and aggregating a number of national polls asking how much each individual representative is trusted to act in the public interest. Preferably the polls would be taken at regular intervals and according to a regular standard. This would allow representatives to be treated as individuals instead of as homogeneous party members, and for the opinions of the public to directly factor into a standardized measure of the competence of institutions in representing the public.

3.2 Institutional-Epistemic Decision Agreement



The vertical position of each blue dot represents the estimated competence of a chamber of Congress in representing the public interest over the preceding six months. Each green/red dot represents an action voted on by a chamber of congress; its color indicates whether or not the action was taken, and its vertical position is the estimated probability taking it was in the public interest. By my assumed definition of institutional competence, a perfectly competent institution would have all green dots above .5 vertically and all red dots below .5 vertically, indicating that it always made the decision most likely to be in the public interest given the available evidence.

The competencies of the House and the Senate have decayed dramatically over the last generation. The decay reflects the transition of our political parties from trustworthy to untrustworthy in the eyes of the public; once the parties crossed the competence threshold of .5, they became more likely to vote against the public interest than for it. Across the threshold, individual representatives are less reliable than a coin flip; a coin with Yea on one side and Nay on the other is more likely to vote in the public interest than a representative of the public. In fact, across the threshold, each vote by a representative for an option becomes evidence against it.

Because both political parties presently have a competence below .5, unanimous majority decisions are estimated as having an approximately 0% chance of being in the public interest. For party line votes, because the Republican party is presently less competent than the Democratic party, the Democratic choice is seen as being more likely than the Republican choice to be in the public interest; the Republican votes provide more evidence for the Democratic choice than the Democratic votes do for the Republican choice.

Because congressional voting data only started to be digitized in 1989 and Gallup only started to poll party images in 1993, there is no convenient way for me to compare the trajectories observed here to earlier trends.

References

- [1] <https://www.c-span.org/congress/votes/?chamber=house>
- [2] <https://www.c-span.org/congress/votes/?chamber=senate>
- [3] news.gallup.com/poll/24655/party-images.aspx