If You're Not Informed, Don't Vote

Malik Magdon-Ismail
Department of Computer Science,
Rensselaer Polytechnic Institute,
Troy, NY 12180.
Email: magdon@cs.rpi.edu

January 24, 2005

Abstract

We present a simple model which indicates that if you are not informed, don't vote. Alternatively, inform yourself.

1 Introduction

This little article arose from the unfortunate but generally accurate introspection that I am not politically savy. I am not well informed, so should I cast my vote? In order to make some form of formal discussion of it, let's build a simple model. Albeit simple, the model conveys the intuition.

Suppose there are N potential voters, and two candidates c_1, c_2. Assume that by some global judgement, candidate c_1 is the best candidate. Assume that the population (N) is composed of N_S savy voters N_I ignorant voters ($N = N_S + N_I$). More formally, an ignorant voter picks a candidate randomly, specifically, an ignorant voter picks the best candidate c_1, with probability $\frac{1}{2}$. A savy voter picks the best candidate with some probability $p > \frac{1}{2}$. Due to randomness in sampling, we denote the election as a wash (no outcome) if the winning margin is less than some desired threshold; equivalently, the percentage of votes required to win is strictly greater than $\frac{1}{2}$. As an example, consider a vote by the house of representatives to determine whether a particular bill will be passed. This can be viewed as an election between two "candidate" outcomes, the bill is passed, and the bill is not passed. An amendment to the constitution requires a $\frac{2}{3}$ majority of 290 votes in the house to be passed. The question we ask is:

> "What is the probability that the best candidate c_1 gets elected if $M \leq N_I$ of the ignorant voters vote, assuming that all savy voters vote?".

A "no-outcome" vote occurs if neither candidate obtains the required number of votes, and this is considered a loss for c_1. Let n_{c_1} be the number of votes c_1 obtains, and n_{c_2} the number of votes that c_2 obtains, $n_{c_1} + n_{c_2} = N_S + M$. For illustration, let the fraction of votes required to win be $\frac{1}{2} + \epsilon$, where $\epsilon > 0$, i.e., c_1 wins if

$$n_{c_1} \geq (\tfrac{1}{2} + \epsilon)(n_{c_1} + n_{c_2}).$$

The intuition behind the title of this paper is that each random voter brings down the average probability of selecting the best candidate closer to $\frac{1}{2}$. Since the threshold for winning is strictly greater than $\frac{1}{2}$, if there are enough random voters, the average probability of voting in the best candidate will become close to $\frac{1}{2}$, and eventually drop below the required threshold of $\frac{1}{2}+\epsilon$, resulting in the best candidate not winning. A randomly selected voter will vote for c_1 with probability

$$\frac{\frac{1}{2}M + pN_S}{M + N_S} = p - \frac{M}{N_S + M}(p - \tfrac{1}{2}),$$

which is monotonically decreasing in M. This probability is maximized at $M = 0$, and so we expect that c_1 will be elected with highest probability when all ignorant voters do not vote.

The conclusion is that the pervasive mantra "Vote; your vote counts." should be changed. Certainly your vote counts, however it is not positively biasing the outcome unless it is an informed vote: "Vote; your informed vote counts!".

The Probability of Electing the Best Candidate

Winning by majority vote can be defined in one of two ways:

i. [**Entire Population Majority**] By the raw fixed number of votes required out of the entire population, for example the $\frac{2}{3}$ house of representatives majority requirement of 290 votes for passing amendments to the constitution.

ii. [**Voting Population Majority**] By the number of votes required from the actual voters, as in the presidential election. The number of votes for the best candidate must exceed the number of votes against by some threshold τ. τ could be a fixed number, for example 1 in a simple majority election, or a fixed fraction of the voting population, for example $\frac{1}{2} + \epsilon$.

Formal Model Description

Assume that all voters are independent, and that there are N_S savy voters and N_I ignorant voters. Assume that all savy voters vote, and the probability that n_s of the savy voters vote for c_1 is given by the binomial distribution,

$$P_S[n_s] = \binom{N_s}{n_s} p^{n_s}(1-p)^{N_s - n_s}$$

Similarily, if $M \leq N_I$ ignorant voters vote, the probabiltiy that n_i of them vote for c_1 is

$$P_I[n_i] = \binom{M}{n_i} \frac{1}{2^M}$$

We seek the probability that c_1 wins within this model, where M of the ignorant voters vote.

Entire Population Majority

If some fixed number of votes are required, the it is always better to have more voters. To be specific, let $S_1 \subseteq S_2$ be two sets of voters. If some subset of S_1 gives a win for c_1, that same subset (with the same probability) can occur in S_2 giving a win to c_1. However, there may be other different subsets in S_2 that can give a win to c_1 which are not available in S_1. The conclusion is that it is always better to vote.

Lets resolve an apparent contradiction. If more voters vote, then it is better for c_1. Certainly the same argument applies to c_2, thus if more voters vote, it is better for both c_1 and c_2, which appears to be a contradiction. The contradiction is resolved by realizing that the probability of a no-outcome vote must drop.

To be specific, consider an example in which $N_s = 5$ and 5 votes are required to win the election. If no ignorant voters vote, then the probability that c_1 wins is p^5 and the probability that c_2 wins is $(1-p)^5$. The probability of a no-outcome vote is $1 - p^5 - (1-p)^5$. Now suppose we add one one ignorant voter. The probability that c_1 wins is $p^5 + \frac{1}{2}\binom{5}{4} p^4(1-p) > p^5$, and the probability that c_2 wins is $(1-p)^5 + \frac{1}{2}\binom{5}{4} p(1-p)^4 > (1-p)^5$. The probability of a no-outcome vote has dropped by $\frac{1}{2}\binom{5}{4} p(1-p)(p^3 + (1-p)^3)$.

Voting Population Majority

Lets first consider the case where c_1 wins if she obtains strictly more votes than c_2, by some fixed number, i.e., $n_{c_1} - n_{c_2} \geq \kappa$. For concreteness, take $\kappa = 1$, i.e., simple majority voting. Let's first consider the case N_S is odd.

Suppose that $M \geq 0$ of the ignorant voters vote, and you are the $M+1^{\text{th}}$ ignorant voter trying to determine whether to vote or not. If $N_s + M$ is odd, then if you do not vote, the majority is given by $A = \frac{1}{2}(N_s + M + 1)$. Let P_{win} denote the probability that c_1 wins, $P_{win} = P[n_s + n_i \geq A]$. We find that

$$P_{win}(M) = \sum_{\substack{n_s+n_i \geq A \\ n_s \leq N_s, n_i \leq M}} \binom{N_s}{n_s} \binom{M}{n_i} \frac{p^{n_s}(1-p)^{N_s-n_s}}{2^M} \tag{1}$$

In the event that you do vote, then the majority is given by $\frac{1}{2}(N_s + M + 1) + 1 = A + 1$, so

$$P_{win}(M+1) = \sum_{\substack{n_s+n_i \geq A+1 \\ n_s \leq N_s, n_i \leq M+1}} \binom{N_s}{n_s} \binom{M+1}{n_i} \frac{p^{n_s}(1-p)^{N_s-n_s}}{2^{M+1}}$$

We will need the identity $\binom{M+1}{n_i} = \binom{M}{n_i} + \binom{M}{n_i-1}$, and we will use the conventions that $\binom{N}{n} = 0$ if $N < 0$, $n < 0$ or $N < n$. Using the identity, we get

$$P_{win}(M+1) = \sum_{\substack{n_s+n_i \geq A+1 \\ n_s \leq N_s, n_i \leq M+1}} \binom{N_s}{n_s} \left(\binom{M}{n_i} + \binom{M}{n_i - 1} \right) \frac{p^{n_s}(1-p)^{N_s-n_s}}{2^{M+1}}$$

If we make a change of variables in this summation to $\hat{n}_i = n_i - 1$, we obtain

$$P_{win}(M+1) = \sum_{\substack{n_s+\hat{n}_i \geq A \\ n_s \leq N_s, \hat{n}_i \leq M}} \binom{N_s}{n_s} \left(\binom{M}{\hat{n}_i + 1} + \binom{M}{\hat{n}_i} \right) \frac{p^{n_s}(1-p)^{N_s-n_s}}{2^{M+1}}. \tag{2}$$

Subtracting (2) from (1), we obtain the increase in the probability that c_1 wins on account of the fact that you (an ignorant voter) voted,

$$P_{win}(M+1) - P_{win}(M) = \sum_{\substack{n_s+n_i \geq A \\ n_s \leq N_s, n_i \leq M}} \binom{N_s}{n_s} \left(\binom{M}{n_i+1} - \binom{M}{n_i} \right) \frac{p^{n_s}(1-p)^{N_s-n_s}}{2^{M+1}}.$$

Changing variables in the first term of the summation to $\hat{n}_i = n_i + 1$ and noting that $\binom{M}{n_i+1}$ is non-zero only if $n_i + 1 \leq M$, the first term is a summation similar to the second term, except that the condition $n_s + n_i \geq A$ is replaced by $n_s + n_i \geq A + 1$. The conclusion is that

$$P_{win}(M+1) - P_{win}(M) = - \sum_{\substack{n_s+n_i = A \\ n_s \leq N_s, n_i \leq M}} \binom{N_s}{n_s} \binom{M}{n_i} \frac{p^{n_s}(1-p)^{N_s-n_s}}{2^{M+1}} < 0.$$

Thus it is disadvantagous to c_1 if you (an ignorant voter) votes in this situation. However, the conclusion is not as clear cut as it seems. To illustrate, suppose, instead, that you were a savy voter. A similar argument (assuming that $M = 0$) obtains that

$$P_{win}(N_S+1) - P_{win}(N_S) = - \binom{N_S}{A} p^A (1-p)^{N_S-A+1} < 0,$$

where $A = \frac{1}{2}(N_S + 1)$. Thus, it is *still* detrimental to c_1 for you (a savy voter) to cast a vote. Even more confusing, the situation reverses if $N_S + M$ is even. It is now beneficial to c_1 for an ignorant voter to vote. The reason is that when $N_S + M$ is even, the majority is given by $\frac{1}{2}(N_S + M) + 1$. If you add one more voter, the majority does not change, and as we argued in the case of a fixed number of votes needed, it always pays to add more voters. Thus, it appears that the value of your vote to c_1 peculiarly depends on whether the current size of the voting population is odd or even, and apparently does not depend on how savy you are. We thus need to appeal to a more general type of argument.

There are two possible avenues. One is to average, i.e., compute the expected advantage to c_1 of your (ignorant) vote. The alternative is to use a game theoretic view, namely, if it is advantageous for one ignorant voter to vote, all of them will vote. If not, then none of them will vote. Both avenues lead to the same computation and imply that we need to compare $P_{win}(N_S + N_I) - P_{win}(N_S)$. For illustration, assume that N_S is even, and $N_S + N_I$ is odd (the other three cases can be handled in exactly the same way.). This is the worst possible case, since we already know that it is good for one ignorant voter to vote, so we ask what happens if all the ignorant voters follow this logic. For simplicity, assume that $N_I \leq N_S$. The case $N_I \geq N_S$ can be handled in a similar fashion. Let $A = \frac{N_S}{2} + \frac{N_I+1}{2}$ be the majority requirement when all voters vote, and let $B = \frac{N_I-1}{2}$.

$$\begin{aligned} P_{win}(N_S) &= \sum_{n_s=\frac{N_S}{2}+1}^{N_S} P_S[n_s] \\ P_{win}(N_S+N_I) &= \sum_{n_s=\frac{N_S}{2}-B}^{N_S} P_S[n_s] \sum_{n_i=A-n_s}^{N_I} P_I[n_i] \end{aligned}$$

The second summation in $P_{win}(N_S+N_I)$ equals 1 when $n_s \geq \frac{N_S}{2}+\frac{N_I+1}{2}$, so

$$P_{win}(N_S+N_I) = \sum_{n_s=\frac{N_S}{2}-B}^{\frac{N_S}{2}+B} P_S[n_s] \sum_{n_i=A-n_s}^{N_I} P_I[n_i] + \sum_{n_s=\frac{N_S}{2}+B+1}^{N_S} P_S[n_s].$$

Using the fact that $P_I[n_i]=P_I[N_I-n_i]$, we can rewrite $P_{win}(N_S+N_I)$ as

$$\begin{aligned} P_{win}(N_S+N_I) &= \frac{1}{2}P_S[\tfrac{N_S}{2}] + \sum_{\frac{N_S}{2}+1}^{\frac{N_S}{2}+B}\left(P[n_s]\sum_{n_i=A-n_s}^{N_I}P_I[n_i] + P[N_S-n_s]\sum_{n_i=0}^{A-n_s-1}P_I[n_i]\right) \\ &= \frac{1}{2}P_S[\tfrac{N_S}{2}] + \sum_{\frac{N_S}{2}+1}^{\frac{N_S}{2}+B}P[n_s] - \sum_{\frac{N_S}{2}+1}^{\frac{N_S}{2}+B}\left(P[n_s]-P[N_S-n_s]\right)\sum_{n_i=0}^{A-n_s-1}P_I[n_i] \end{aligned}$$

and so we find that

$$P_{win}(N_S+N_I) - P_{win}(N_S) = \frac{1}{2}P_S[\tfrac{N_S}{2}] - \sum_{\frac{N_S}{2}+1}^{\frac{N_S}{2}+B}\left(P[n_s]-P[N_S-n_s]\right)\sum_{n_i=0}^{A-n_s-1}P_I[n_i]$$

Every term in the summation is positive, so the summation is strictly positive. Let's consider a single term in the summation, corresponding to $n_s^* = \frac{N_S}{2}+O(\sqrt{N_I})$. By the central limit theorem, the second summation converges to some constant, call it $\kappa < \frac{1}{2}$. Since $P_S[\frac{N_S}{2}] > P_S[N_S-n_s^*]$, and $\kappa < \frac{1}{2}$,

$$P_{win}(N_S+N_I) - P_{win}(N_S) < P_S[\tfrac{N_S}{2}] - \kappa P_S[n_S^*]$$

Using Stirling's approximation in the asymptotic regime, when $N_S, N_I \to \infty$, $P[\frac{N_S}{2}]/P[n_s^*] \to \exp\left(-\frac{(p-\frac{1}{2})\sqrt{N_I}}{p(1-p)} + \frac{N_I}{2N_S p(1-p)}\right) \to 0$, so $P_{win}(N_S+N_I) - P_{win}(N_S) < 0$ for large enough N_S, N_I. The conclusion is that none of the ignorant voters should vote in a large enough voting population.

If N_S is even and N_S+N_I is also even, then from the discussion on adding one ignorant voter, we know that $P_{win}(N_S+N_I+1) > P_{win}(N_S+N_I)$. From the previous result we therefore have that asymptotically, $P_{win}(N_S) > P_{win}(N_S+N_I+1) > P_{win}(N_S+N_I)$, which settles the case N_S even, N_S+N_I even.

The remaining two cases when N_S is odd can be handled in exactly the same way. Let N_S+N_I also be odd. The majority is $A = \frac{N_S}{2}+\frac{N_I}{2}+\frac{1}{2}$ and let $B = \frac{N_I}{2}-\frac{1}{2}$. Then,

$$P_{win}(N_S+N_I) - P_{win}(N_S) = -\sum_{\frac{N_S}{2}+\frac{1}{2}}^{\frac{N_S}{2}+B}\left(P[n_s]-P[N_S-n_s]\right)\sum_{n_i=0}^{A-n_s-1}P_I[n_i] < 0$$

Once again, adding one ignorant voter gives $P_{win}(N_S+N_I) > P_{win}(N_S+N_I+1)$, which settles the case N_S odd, N_S+N_I even. To recapitulate, if N_S is even, then it may be the case that for small odd N_I, it is advantageous to c_1 for the ignorant voters to vote. However, as N_S, N_I becomes large, it is disadvantageous (to c_1) for the ignorant voters to vote. When N_S is odd, it is bad for *any* number of ignorant voters to vote.

Let's now briefly consider the case when the winning margin has to be some fraction of the voting population, i.e., $n_s - n_i \geq (N_S + M)(\frac{1}{2} + \epsilon)$. Let $A(M) = \lceil (N_S + M)(\frac{1}{2} + \epsilon) \rceil$. If $A(M+1) > A(M)$, then the preceeding arguments indicate that is better for a single additional voter not to vote. If $A(M+1) = (M)$, then it is better for an additional voter to vote. Considering the case where all ignorant voters vote versus none of them voting, we can perform a similar analysis to the case of a fixed majority to obtain that asymptotically as $N_S, N_I \to \infty$, none of the ignorant voters should vote. In fact, asymptotically as $N_I \to \infty$, the probability that c_1 wins approaches 0. To see this, we can use a Chebyshev bounding argument. $n_{c_1} = n_s + n_i$. $E[n_s + n_i] = E[n_s] + E[n_i] = pN_S + \frac{1}{2}N_I$, and since n_s and n_i are independent, $Var[n_s + n_i] = Var[n_s] + Var[n_i] = p(1-p)N_S + \frac{1}{4}N_I$. Thus, for fixed N_S,

$$
\begin{aligned}
P[n_{c_1} \geq (N_S + N_I)(\frac{1}{2} + \epsilon)] &= P[n_{c_1} - E[n_{c_1}] \geq N_S(\tfrac{1}{2} + \epsilon - p) + N_I\epsilon] \\
&\leq P[|n_{c_1} - E[n_{c_1}]| \geq N_S(\tfrac{1}{2} + \epsilon - p) + N_I\epsilon] \\
&\leq \frac{Var(n_{c_1})}{(N_S(\frac{1}{2} + \epsilon - p) + N_I\epsilon)^2} \\
&= \frac{p(1-p)N_S + \frac{1}{4}N_I}{(N_S(\frac{1}{2} + \epsilon - p) + N_I\epsilon)^2} \to 0.
\end{aligned}
$$

Since n_{c_1} is the sum of Bernouilli random variables, tighter bounds could be obtained using a Chernoff bounding arguments.

Food for Thought. Interesting conversation might ensue in considering the case where the voters are not independent. More constructive conversation might result in pondering how one should inform oneself. More specifically, how do annaeling type strategies behave. For example, every voter polls her "neighborhood" of voters to determine their majority opinion, and makes that her vote. Does this help c_1, and if the process is repeated, does it converge to an equilibrium in which the best candidate gets elected with high probability if there are enough savy voters?