Determining the number of ways there can be a tie in a USA Presidential election

Karl Sigman *

April 2013

Abstract

The winner of a USA presidential election is determined by which candidate wins a majority of the 538 Electoral College (EC) votes, not the national popular vote. We introduce the breakdown of the Electoral College (EC) votes for each state and study the problem of determining the total number of ways there could be a tie in a presidential election. We use stochastic simulation (random number generation), to obtain an approximation to the answer. It turns out that an exact (non-random recursive) method is possible to get the exact answer, but here we focus on stochastic simulation as a powerful approximation method.

*Department of Industrial Engineering and Operations Research, Columbia University, New York, NY 10027 (ks20@columbia.edu).

1 Introduction

The USA Presidential election, every four (4) years, is not decided by the nation's popular vote, but instead by what is called the *Electoral College (EC)* vote. In short, there are 538 EC votes distributed among the 50 states of the USA (plus Washington D.C.) for a total of 51 "states" that each decide how to allocate their allotment of EC votes. The number of EC votes that each state is granted is based on population, and is updated every 10 years based on the "Census". The way the election works: The presidential election takes place on the Tuesday after the first Monday in November, every 4 years. (The last such election (as of the writing of this paper) was November 6, 2012: Obama versus Romney; Obama won his second term.) Whichever candidate wins the majority of the popular vote in a given state, wins all the EC votes in that state, a *winner-takes-all* method. For example, in the State of New York (which currently has $n = 29$ EC votes) if candidate A wins the popular vote, even by as little as 51 percent, then candidate A is awarded all 29 EC votes.[1] The sum for each candidate is totaled, a tie being 269 each; the winner is the one with at least 270 EC votes. (Yes, a tie is theoretically possible!) The number of EC votes varies from as little as $n = 3$, for small population states such as Alaska, and D.C., to as large as $n = 55$ for California (the state with the largest population). We will study the problem of determining the total number of ways there can be a tie. We will assume that there are two (2) candidates A and B.

2 The EC votes

Let n_i denote the number of EC votes granted to state i, $i = 1, 2, \ldots, 51$. Each state has a number of members in the *USA House of Representatives*, based on its population (each member represents about $750,000$ constituents in their district.) Let H_i denote the number of House Representatives for state i. Each state also has 2 Senators serving in the USA Senate. Then $n_i = H_i + 2$ by definition. Washington D.C. is not a state, so it has no Representatives and no Senators, but it is awarded 3 EC votes because its population is the same as having 1 Representative (about $750,000$), and 2 (for Senators) is thrown in to be fair. This is done because citizens of D.C. are USA Citizens and thus have the right to vote (even though they have no representatives in Congress).

A list of the allocation of the EC (Electors) over the last several elections is given here up to and including the 2010 Census (hence including the 2012 election):

http://www.thegreenpapers.com/Census10/HouseAndElectors.phtml

The numbers that were used in 2012 are under the heading (2012, 2016,2020). The numbers that were used in 2008 are different because of the 2010 Census. This change will change the answer to our question, but probably not by much. (You will be able to compare!)

[1]Maine and Nebraska use the Congressional District Method instead of the winner-takes-all method, but we will not take this into account in our analysis given that they are small states, and given that in past elections their method essentially ends up being winner-takes-all anyhow (for example, Maine has never split its votes). We do note, however, that in the 2008 election, Nebraska's $n = 5$ EC votes were split for the very first time, 4 for John McCaine and 1 for Barack Obama. On the other hand, the Nebraska's state legislature has put forth a bill (still pending) to revert back to a winner-takes-all method.

3 The problem statement

Given two (2) candidates, A and B, let $I_i(A) = 1$ if candidate A wins state i, and 0 if not. Let $I_i(B) = 1$ if candidate B wins state i, and 0 if not. Here $i \in \{1, 2, \ldots, 51\}$. Letting C_A and C_B denote the total number of EC votes won, we have

$$C_A = \sum_{i=1}^{51} n_i I_i(A), \quad (1)$$

$$C_B = \sum_{i=1}^{51} n_i I_i(B), \quad (2)$$

and $C_A + C_B = 538$.

A tie occurs if and only if $C_A = C_B = 269$.

In any case, there are $N = 2^{51} = 2,251,799,813,685,248$ ways that the EC votes can be allocated among the two candidates. This is a very large number!! Equivalently, there are $N = 2^{51}$ ways that the USA's 51 states can be partitioned into two groups (one for candidate A and one for candidate B).

Let N_T denote the number of ways there could be a tie. We want to determine N_T.

Given how large N is, this at first seems like an onerous task. But we will indeed solve for N_T using two methods to be introduced next.

4 Simulation method for estimation of N_T

Note that if we used a fair coin flip to determine each states outcome ($Heads = A$, $Tails = B$), flipping the coin 51 times, then each of the $N = 2^{51}$ possibilities would be *equally likely*. Moreover, the ratio $P_T = N_T/N$ would then yield *the probability that there is a tie.*

The probability there is a tie is equal to the number of ways there can be a tie divided by the total number of ways possible.

Since $N_T = P_T N = P_T 2^{51}$: If we can estimate P_T, then we get an estimate for N_T.

We will use simulation to obtain our estimate of P_T. We will ask our computer to flip a fair coin for us 51 times. We will set $I_i(A) = 1$ if it lands Heads, 0 otherwise, and set $I_i(B) = 1$ if it lands Tails, 0 otherwise. (If U_i denotes the i^{th} independent uniform number over the interval $(0, 1)$ generated by our computer, then we can set $I_i(A) = 1$ if $U_i \leq 0.5$; $I_i(A) = 0$ if $U_i > 0.5$.)

Then we can use the formula in (1) to get a copy C_A, C_B of the outcome. We will denote this (first) outcome as $C_A(1), C_B(1)$. We next do this again independently, to obtain $C_A(2), C_B(2)$, and so on, a very large number m of times yielding m such C_A copies $C_A(1), C_A(2), \ldots, C_A(m)$. (Since $C_B = 538 - C_A$, we do not need to separately compute the C_B copies.) We now let $T_i = 1$ if $C_A(i) = 269$ and 0 otherwise. For then

$$\sum_{j=1}^{m} T_i$$

is the total number out of the m copies that yielded a tie.

The following approximation results from the famous *strong law of large numbers (SLLN)* in probability theory:

$$P_T \approx \hat{P}_T(m) = \frac{1}{m}\sum_{j=1}^{m} T_i.$$

The SLLN asserts that P_T is exactly (with probability one) the long-run *proportion of copies out of* m *that yield a tie*; $P_T = \lim_{m\to\infty} \hat{P}_T(m)$. For example, it asserts that the proportion of fair coin flips that land Heads approaches exactly a half as the number of flips increases to ∞.

How big should m be? Well, we will use $m = 10^6$ to be sure it is large enough!

5 Exact algorithm method for determining N_T: a recursive approach

In general, trying to count through a huge number of possibilities (such as $N = 2^{51}$) to compute exactly how many produce a tie, can be a daunting and time consuming task. Mathematically, one can show that such problems are too hard to solve exactly in general (complexity theory). But it turns out that in our case, because the values of the n_i are not large, we indeed can do the counting in a fast way (15 seconds on your laptop)! We will learn how to do this, and we will see that our exact answer is very similar to our approximation thus showing that the simulation method was a good one for approximation.

As we will see , the method of solving exactly for N_T involves the use of a mathematical *recursion.* As a simple example of a recursion consider

$$X_{n+1} = 2X_n + 1,\ \ n \geq 0,$$

where the values of n are the non-negative integers. If we start with a given initial value for X_0, then the recursion sequentially yields all future values of X_n. For example, if we choose $X_0 = 1$, then $X_1 = 2X_0 + 1 = 2 \times 1 + 1 = 3$, and $X_2 = 2X_1 + 1 = 2 \times 3 + 1 = 7$, and so on. Computationally, this means that we can produce the values of X_n sequentially one at a time and could have a computer do so up to any number (such as $n = 500$.) Sometimes a recursion even allows us to explicitly solve for each value of X_n in terms of only the initial value X_0. For example, in our recursion here, we can do so: $X_2 = 4X_0 + 3$, and $X_3 = 8X_0 + 7$ and so on: Can you derive the fact that

$$X_n = 2^n X_0 + 2^n - 1,\ \ n \geq 1,$$

?

5.1 The recursion

We have 51 integers $n_1, \ldots n_{51}$. For any $1 \leq j \leq 51$, and any non-negative integer I, let $N(j, I)$ denote the number of ways that a subset of the first j, $n_1, \ldots n_j$, of all the 51 integers can be used to sum up to I. (What we want is $I = 269$, and $j = 51$.) Then the following recursion holds:

$$N(j + 1, I) = N(j, I) + N(j, I - n_{j+1}). \tag{3}$$

To motivate the recursion (3) consider a small case of only 6 states with n_i values

$\{n_1, n_2, n_3, n_4, n_5, n_6\} = \{3, 4, 4, 5, 6, 8\}$. A tie is then 15. How many ways can this be achieved? The answer is easily seen as 4: Here are the 4 ways: $\{6, 5, 4\}$, $\{6, 5, 4\}$, $\{3, 4, 8\}$, $\{3, 4, 8\}$. So, $N(6, 15) = 4$.

Now let us use the recursion (3) to compute this answer: If we use $I = 15$, and we take $j = 5$, then $j + 1 = 6$. For $j = 5$ we use the first 5 numbers $\{n_1, n_2, n_3, n_4, n_5\} = \{3, 4, 4, 5, 6\}$, and we have $n_6 = 8$. Note that $I - n_6 = 15 - 8 = 7$. The recursion then says that

$$N(6, 15) = N(5, 15) + N(5, 7).$$

To compute $N(5, 15)$ we see that there are two ways (subsets) to get the sum 15 from the 5 elements $\{3, 4, 4, 5, 6\}$: $\{n_5, n_4, n_3\} = \{6, 5, 4\}$ and $\{n_5, n_4, n_2\} = \{6, 5, 4\}$. Thus $N(5, 15) = 2$.

Similarly, there are two ways (subsets) to get the sum 7 from the 5 elements: $\{n_1, n_2\} = \{3, 4\}$ and $\{n_1, n_3\} = \{3, 4\}$. Thus $N(5, 7) = 2$.

Thus indeed $N(6, 15) = N(5, 15) + N(5, 7) = 2 + 2 = 4$.

Expanding further another step we could use the recursion to compute each of the two pieces $N(5, 15)$ and $N(5, 7)$:)

$$N(5, 15) = N(4, 15) + N(4, 9),$$

and

$$N(5, 7) = N(4, 7) + N(4, 1).$$

It is easy to see that $N(4, 9) = 2$ (subsets $\{n_2, n_4\} = \{4, 5\}$ and $\{n_3, n_4\} = \{4, 5\}$), and that $N(4, 7) = 2$ (same subsets $\{n_1, n_2\} = \{3, 4\}$ and $\{n_1, n_3\} = \{3, 4\}$), and $N(4, 15) = 0 = N(4, 1)$. Thus $N(5, 15) = N(5, 7) = 2$.

If we go a step further to compute

$$N(4, 7) = N(3, 7) + N(3, 2),$$

and

$$N(4, 9) = N(3, 9) + N(3, 4),$$

it is easy to see that $N(3, 7) = 2$ (same subsets $\{n_1, n_2\} = \{3, 4\}$ and $\{n_1, n_3\} = \{3, 4\}$), and $N(3, 2) = 0$, and $N(3, 9) = 0$ and $N(3, 4) = 2$ (subsets $\{n_2\} = \{4\}$ and $\{n_3\} = \{4\}$).

If we go a step further to compute

$$N(3, 7) = N(2, 7) + N(2, 3),$$

$$N(3, 4) = N(2, 4) + N(2, 0),$$

it is easy to see that $N(2, 7) = 1$ (subset $\{n_1, n_2\} = \{3, 4\}$), and $N(2, 3) = 1$ (subset $\{n_1\} = \{3\}$). Also $N(2, 4) = 1$ ($\{n_2\} = \{4\}$, and we do have $N(2, 0) = 1$ since we can use the so-called *empty subset* $\emptyset$ consisting of no elements (so indeed they sum up to 0). Continuing:

$$N(2, 7) = N(1, 7) + N(1, 3),$$

$$N(2, 3) = N(1, 3) + N(1, -1)$$

$$N(2, 4) = N(1, 4) + N(1, 0),$$

where we immediately see that $N(1,7) = 0$ and $N(1,3) = 1$ (subset $\{n_1\} = \{3\}$), and of course $N(1,-1) = 0$, and $N(1,4) = 0$ and $N(1,0) = 1$ (empty set again).

To actually use the recusion the way it is intended, we would start at the bottom initializing with the immediate obvious pieces $N(1,3) = 1$ and $N(1,7) = 0$, and $N(1,4) = 0$ then going up one step to get $N(2,7) = N(1,7) + N(1,3) = 1$ and $N(2,3) = N(1,3) + N(1,-1) = 1$, $N(2,4) = N(1,4) + N(1,0) = 1$. Then moving up another step to get $N(3,7) = 2$, $N(3,4) = 2$ and so on moving step by step all the way to the top to get the final answer $N(6,15) = N(5,15) + N(5,7) = 2 + 2 = 4$. All this can be done by your computer!

5.2 Further applications of the simulation approach: A particular election

Putting aside the problem of determining the number of ways N_T there can be a tie, note that the simulation approach can yield *the probability P_T there is a tie* regardless of what the individual state probabilities are. We used "equally likely" $p = 0.5$ for each state so that we could assert the simple relation

$$N_T = P_T 2^{51}; \tag{4}$$

our objective was to determine the value of N_T.

But when the values of p differ from state to state (as they would in a real election), we no longer have the relation in (4). But it would be useful to compute P_T as well as certain expected values and probabilities under such a scenario such as $E(C_A)$, the expected number of EC votes that candidate A will get, or $P(C_A \geq 270)$, the probability that candidate A will be the winner. In a real election, then, one can in principle make predictions of who will win. The idea is to obtain good estimates of $p_i \stackrel{\text{def}}{=}$ *the probability that candidate A will win state i*, for each $1 \leq i \leq 51$. One way of estimating p_i is through extensive state-by-state polling data. Note how this method completely bypasses using the overall national poll estimate: That is not what determines who will win a USA Presidential election. Of course, if one candidate is leading the other by a wide margin in the national polls, then it becomes increasingly unlikely that the other will win; a simulation study is not likely to shed any light on the matter. It is only when the polls are relatively close that a simulation study can really add new help to predict what will happen.