Using the Markov chain to predict the outcome of an election in a given situation

Yajie Wan
The Village School, Houston, TX 77077, America
Corresponding author’s e-mail: wyj0283@126.com

Abstract. This paper used Markov Chain, a mathematical model that is often utilized in predicting, to forecast the result of the election process of a person in a specific circumstance. In the way of modeling, starting from the tree diagram to the model of Markov Chain and related transition matrix, this paper aimed at the election result, which relates the object of study, a specific candidate. The research data are collected from the prompt of the research question and the calculation of coding based on the transition matrix. The result is, when it is the 64th year of that specific candidate running for the reelection campaign, he will be likely to retire from politics.

1. Introduction
Markov Chains is a commonly-used statistical model that is always used for modeling specific mathematical problems in numerous special situations, and it also can simplify complicated mathematical models into some terse expressions so people will be likely to pay more attention to them. The election is such an important incident happening periodically, and during every election, every candidate will have to start ahead of time to prepare for the crucial election, such as canvassing, make public speeches, do public benefit activities in order to acquire the masses’ great praises and support during elections.

The result of every election is also significant for not only the candidates but also the citizens and non-citizens, and every time there is an election, there are definitely websites online doing predictions of whom is going to win the election, which came to my notice. Through this essay, I would like to make a prediction of a specific condition of the election using the interesting method, Markov Chain, and its related transition matrix. By means of the calculations, I would like to draw a conclusion of a specific number of years a candidate will retire from politics after he wins his first election. The significance of predicting the result of the election will give the candidates a basic comprehension of how their process of election is going to look like based on a specific circumstance.

2. Analysis

2.1. Tree diagram
A tree diagram is a way to plan multiple stratified tasks. It basically starts from a point and outstretches branches, and the following ones will relatively abide by this pattern[1].

2.2. Markov Chain
A Markov chain is a stochastic model with Markov property. It describes a series of possible events, and the probability of each event only depends on the state that was reached in the previous event[2-3].
2.2.1. **Markov property.** The Markov property is a concept in probability theory named after the Russian mathematician Andrey Markov. A brief definition of it is that when a random process is given the state at present and all other past states, the conditional probability distribution of its next state depends only on this current state. Or this property can be also said that when the present state is given, it is completely and conditionally independent from all the past states. Any process with Markov property is called the Markov process.

2.2.2. **Stochastic process.** A stochastic process is a collection of random variables that are indexed by some mathematical set, and each random variable of the stochastic process is only related to an element in the set [4].

2.2.3. **Transition probability.** Transition probabilities are the probabilities related to diversified state changes. If one change has transition probability, then this change is the state change between two states but not more than two [5].

2.2.4. **Transition matrix.** The transition matrix of the Markov Chain is a stochastic matrix. It gives the probability that a point moves from the ith state to the jth state at the entry (i, j)[6].

2.2.5. **Absorbing state.** An absorbing state in a transition matrix means that once we get to that state, there are no probabilities that allow us to go anywhere else, which means that we will stay at that state forever and never come out. An absorbing state will have a probability 1 shown in the corresponding transition matrix.

3. **Research question**
A popular politician named AP is running for the US Congress.
- If AP has never been elected before, then the probability that they will be likely elected is 1/2, and if they lose the ballot, they are able to campaign in the following election which happens biennially.
- If AP has already been elected and is currently staying in office, then the probability of being reelected is 9/10.
- But if they lose the ballot when campaigning for the reelection, he retires from politics.
Here we will utilize Markov Chain and the transition matrix to solve the following question:
Assume that AP is running for the first time. In how many years should they expect to retire?

4. **Modeling**
Let the status “Never been elected” be symbolized by letter N, letter E with script n, En, which stands for the nth election, and letter R to represent the status “Retire from politics”
Figure 1. Election outcomes-tree diagram

Figure 1 complicatedly elaborates on the tree diagram which symbolizes the overall election outcomes for AP. According to the question stem, it is clear that at the beginning, the probabilities of AP being elected or failing the election are the same, therefore, each of them is 1/2. Once AP is elected, then when he campaigns for the next election, the probability of him being reelected is 9/10, therefore, the probability of AP failing the campaign and retiring from politics is 1/10. Despite AP losing his first campaign, he could go campaign again in the next election which will take place two years later. But if AP gets reelected again, then the probabilities of him still being reelected or getting retired at the next election will keep following the last pattern until he loses the election.

Since the tree diagram modeled seems too complicated, a Markov Chain would be better to be utilized for improvement. Let letter N, E1, and R still respectively symbolize the three general status. In Figure 2, a curved arrow is firstly used so that the branches extended from the N states will be dramatically simplified. As stated before, this initial Markov Chain(MC) stands for the event of transition from the state “never been elected” to the state “get elected”, with probability 1/2 and 1/2 respectively. Since AP once gets elected, it is impossible to go back to the N state, hence no backward arrow is necessary to be drawn.

From Figure 3, state E2 is superimposed to show the possible event in which AP gets reelected while in office (already been elected). The probability of AP gets reelected is 9/10. Since there are no possible ways to go back to state E1 from E2, there is no backward arrow displayed.
Based on Figure 3, the branch of R state is added, which is shown as a transition from state E1 to state R in Figure 4.

The probability of AP failing the reelection and getting retired from politics is $1 - \frac{9}{10} = \frac{1}{10}$. On account of the fact that once loses the reelection, he will go to the R state which is get retired, thus making no rational reason for a backward arrow to appear here pointing from state R to state E1, thus the probability that AP will stay at state R is 1, with a curved arrow shown around the circle of state R. In Figure 4 there are no brand-new things added to the MC, but to keep repeating step (2) and (3).

According to the Markov chain model created above, we could then form a transition matrix[7]:

$$
\begin{bmatrix}
N & E & R \\
1/2 & 1/2 & 0 \\
0 & 9/10 & 1/10 \\
0 & 0 & 1
\end{bmatrix}
$$

Figure 6 obviously reveals all the entries of different transitions: 1/2 for the probability of both events “state N to N” and “to E”, and 0 for the probability of the event “state N to R” because there is no way that AP could skip the state E and go all the way to state R; 0 for the probability of the event “state E to N” because AP cannot go back to the status “never been elected” after being elected, 9/10 for the probability of the event “state E to E”, and 1/10 for the probability of the event “state E to R”; 0 for the probability of both the events “state R to N” and “to E” because it’s not possible for AP to go back to neither state N nor state E after getting retired from the politics(state N and state R are completely two diverse states), and finally, 1 as the probability of the event “state R to R”.

5. Solving process

From the initial tree diagram to Markov chains, and to the transition matrix, we can see that once it goes to the “retire” state, there are no possible ways to transit to other states, therefore, it is an absorbing state.

Since the purpose is to find the expected years until AP’s retirement, given from the matrix that AP can only get retired from state E which stands for “being elected”, we will then raise the transition matrix P to its nth power, and n will be the result of how many years should AP expects to retire from politics. Therefore, we will have a raw formula:
which $x$ stands for the initial array/probability (which always looks like $[x \ x \ x]$), $P$ symbolizes the transition matrix, and $n$ represents the number of years that AP expects to retire from politics.

Since we don’t have the initial array in the research question, we will then do a math justification to prove that

$$xP^n = P^n \quad (2)$$

Here we have a formula:

$$\sum_{i=1}^{\infty} x(i) \times P(j \mid i) - \sum_{i=1}^{\infty} x(j) \times P(i \mid j) \quad (3)$$

and because $x(j)$ is a constant, it will be better for us to move it to the front:

$$= x(j) \sum_{i=1}^{\infty} P(i \mid j) \quad (4)$$

Since for every single row on the matrix, the sum of each probability on the specific row is always equal to 1. Therefore, $\sum_{i=1}^{\infty} x(i) \times P(j \mid i)$ equals 1, which leads to the following:

$$\sum_{i=1}^{\infty} x(i) \times P(j \mid i) = \sum_{i=1}^{\infty} x(j) \times P(i \mid j)$$

$$= x(j)$$

$$= x(j) \sum_{i=1}^{\infty} P(i \mid j) \quad (5)$$

Based on calculations of keep multiplying the transition matrix $P$ by itself, we will find that when $n$ equals 64, the transition matrix will start to look like this:

$$p^{64} =$$

```
1.35525272e-20 1.19375619e-03 9.98806244e-01
0.00000000e+00 9.55004951e-04 9.99044995e-01
0.00000000e+00 0.00000000e+00 1.00000000e+00
```

(6)

$$p^{65} =$$

```
6.77626358e-21 1.07438057e-03 9.98925619e-01
0.00000000e+00 8.59504456e-04 9.99140496e-01
0.00000000e+00 0.00000000e+00 1.00000000e+00
```

(7)

$$p^{66} =$$

```
3.38813179e-21 9.66942513e-04 9.99033057e-01
0.00000000e+00 7.73554010e-04 9.99226446e-01
0.00000000e+00 0.00000000e+00 1.00000000e+00
```

(8)
From $P^{64}$, we can obviously see that the second row, which we pay the main attention to, started to have its third entry (counting from left to the right) staying at 0.999 since then. As we all know from the previous description, this entry represents the state that AP loses his reelection when being in the office, which is “state E to R”.

Therefore, we can tell that when it’s the 64th year of AP staying in the office, he will likely retire from politics.

6. Conclusion
In general, through the detailed modeling and specific solving process, we came to the conclusion that 64 years after the time when AP firstly got elected, he will retire from politics. Throughout the paper, from the tree diagram to the Markov Chain model and the transition matrix, we can be pretty sure that Markov Chain definitely looks neater and easier to read because it does not seem as intricate as the tree diagram, and it simplified some cumbersome steps of the tree diagram. There are still things that I can improve in this study, such as adding more interesting affairs to the research question itself so that the Markov Chain and the relative transition matrix will be more amusing for me to do interesting calculations.

Acknowledgments
In the first place, I would like to thank all the people who have helped me with my essay, especially the professor for providing the vitally important theory of Markov Chain used among the thesis writing, Alex Zong (from Hong Kong University of Science and Technology) for assisting guidance and specialized knowledge throughout this thesis, Steven Hou for supporting me a lot with both the overall background knowledge and the solving process, and Eva Chen for helping me modifying my outline and every draft regarding structure, diction. I also want to give big thanks to my parents, my friends, and my teachers for always supporting me. I may not be able to complete this paper without their help in both theoretical and spiritual ways.

References
[1] Cook, Stephen John, et al. "Object model tree diagram." U.S. Patent No. 7,814,427. 12 Oct. 2010.
[2] Powell, V. (2014). Markov Chains explained visually. Retrieved August 28, 2020, from https://setosa.io/ev/markov-chains/
[3] Tierney, Luke. "Introduction to general state-space Markov chain theory." Markov chain Monte Carlo in practice (1996): 59-74.
[4] Clark, Allan, et al. "Stochastic process algebras." International School on Formal Methods for the Design of Computer, Communication and Software Systems. Springer, Berlin, Heidelberg, 2007.
[5] Uhlmann, Armin. "The “transition probability” in the state space of a*-algebra." Reports on Mathematical Physics 9.2 (1976): 273-279.
[6] Allignol, Arthur, Martin Schumacher, and Jan Beyersmann. "Empirical transition matrix of multi-state models: the etm package." Journal of Statistical Software 38.4 (2011): 1-15.
[7] Franklin, Joel N. Matrix theory. Courier Corporation, 2012.