STATISTICS IN TRANSITION new series, June 2018
Vol. 19, No. 2, pp. 351–357, DOI 10.21307/stattrans-2018-020

ON A SURPRISING RESULT OF TWO-CANDIDATE ELECTION FORECAST BASED ON THE FIRST LEADERSHIP TIME

Czeslaw Stępniak[1]

ABSTRACT

This is a simple but provocative note. Consider an election with two candidates and suppose that candidate A was the leader until counting n votes. How to use this information in predicting the final results of the election? According to the common belief the final number of votes for the leader should be a strictly increasing function of n. Assuming the votes are counted in random order we derive the Maximum Likelihood predictor of the final number of votes for the future winner and loser based on the first leadership time. It appears that this time has little effect on the predicting.

Key words: two-candidate election, winner, leader, leadership time, predicting number of votes for winner, Maximum Likelihood.

1 Introduction

Two-candidate election such as the last round of presidential election always attracts a great attention. Suppose that candidate A was the leader until counting n votes. We write $T = n$ for the first leadership time T. The problem is how to use this information in predicting the final results of the election. According to the common belief the final number of votes for the leader should be a strictly increasing function of n.

Assume the votes are counted in random order. Combinatorial tools for the process of counting of votes in this situation may be found in many books and articles (Brémaud (1994), Feller (1968), Goulden and Serrano (2003), Lengyel (2011), Renault (2007) and Takacs (1997), among others) under the name of the ballot problem. The results are usually formulated in probabilistic terms.

Statistical inference is often based on the notion of likelihood (cf. Azzalini (1996)) and the Maximum Likelihood principle plays the fundamental role in the process. In the present note we derive the Maximum Likelihood predictor of the final number of votes for the future winner.

Presentation of this note is accessible not only for specialists.

[1] Faculty of Mathematics and Natural Sciences, University of Rzeszów.
E-mail: stepniak@umcs.lublin.pl.

2 Initial formalization and predicting the future winner

Consider election with two candidates A and B. In this note the number of all votes is known and is denoted by N. So that all potential results of the election were conclusive we assume that N is odd and each vote indicates exactly one candidate. The votes are counted in random order, that is all permutations are equally probable.

The results of the election are usually announced by the winner (W) and by the final number of votes for W. For some technical reasons, instead of this number, it will be convenient to handle the final number of votes for loser. Denote the last number by M. In this situation the pair (W,M), where $W \in \{A,B\}$ and $M \in \{0,1,...,\frac{N-1}{2}\}$ plays the role of unknown parameter.

One can consider two problems: predicting W under assumption that M is the nuisance parameter, and predicting M. The both predictors are based on the observation (L,T), where $L \in \{A,B\}$ is the first leader and T is the first leadership time.

As regards the first problem we may choose between two predictors: $W = L$ and $W \neq L$. Intuitively, the first one is better. We shall formally confirm this intuition. To this aim we only need to observe that a candidate will be the first leader if and only if the first vote is for him.

In consequence

$$P_M(W=L) = P_M(L=W) = \frac{\text{final number of votes for winner}}{\text{number of all votes}} = \frac{N-M}{N} > \frac{1}{2}$$

while

$$P_M(W \neq L) = P_M(L \neq W) = \frac{\text{final number of votes for loser}}{\text{number of all votes}} = \frac{M}{N} < \frac{1}{2}$$

for all $M \leq \frac{N-1}{2}$. Therefore, the predictor $W = L$ is better.

For predicting the final number of votes for winner and loser we shall use distribution $P_M(T=n)$ of the first leadership time T.

3 Towards distribution of the first leadership time

Some results on this distribution were derived in Stępniak (2015) under silent assumption that $M > 0$. We shall prove the following

Theorem 1 *For all $M = 0,1,...,\frac{N-1}{2}$ distribution $P_M(T=n)$ of the first leadership time T is given by*

$$P_M(T=n)=\begin{cases} \frac{2}{n}\binom{n}{\frac{n+1}{2}}\binom{N-n-1}{M-\frac{n+1}{2}}\binom{N}{M}^{-1}, & \text{if } n \text{ is a positive odd integer such that } \frac{n+1}{2}\le M,\\ \frac{N-2M}{N}, & \text{if } n=N,\\ 0, & \text{otherwise.}\end{cases} \quad (1)$$

Proof. Let us consider three cases:

I. $M=0$ (n arbitrary),

II. $n=N$ (M arbitrary),

III. $M>0$ and $n<N$.

We mention that the classical ballot problem refers only to the probability $P_M(T=N,L=W)$. In our notation the well-known Ballot Theorem (see Brémaud(1994), Feller (1968), Goulden and Serrano (2003), Lengyel (2011), Renault (2007) and Takacs (1997)) may be expressed in the form

$$P_M(T=N,L=W)=\frac{(N-M)-M}{N}=\frac{N-2M}{N} \quad \text{for all } M. \quad (2)$$

In the case II $P_M(T=N,L\neq W)=0$ and, therefore,

$$P_M(T=N)=\frac{N-2M}{N} \text{ for all } M=0,1,...,\frac{N-1}{2}.$$

The case I is trivial and it leads to

$$P_0(T=n)=\begin{cases} 1, & \text{if } n=N,\\ 0, & \text{otherwise.}\end{cases} \quad (3)$$

In this situation the set of all positive integers n such that $\frac{n+1}{2}\le M$ is empty and hence the formula (3) coincides with (1).

Now let us consider the case III.

Any record of the counting of votes may be represented as a lattice path from the origin to $(N,N-2M)$ with steps of type $(1,1)$ and $(1,-1)$ indicating that a successive voice is for the future winner or loser. In particular, the first leader is the future winner with the leadership time n, if and only if the path is touching the x-axis for $x=n+1$ and lying above the axis for all positive integers $x\le n$. Similarly, the first leader is the future loser with the leadership time n, if the corresponding segment of the path is lying below the x-axis. In consequence, $P_M(T=n,L=W)$ is positive only for n odd and less than $2M$. Moreover

$$P_M(T=n,L\neq W)=P_M(T=n,L=W). \quad (4)$$

This fact is known as the Reflection Principle (see, Brémaud(1994) or Feller (1968), for instance). Thus it remains to find $P_M(T=n,L\neq W)$ for $n=1,3,...,2M-1$. The desired probability may be expressed in the form

$$P_M(T=n, L \neq W) = P(A)P(B/A)P(C/A \cap B),$$

where

• A is the event that $\frac{n+1}{2}$ among the first n votes will be for loser (and $\frac{n-1}{2}$ for winner),

• B is the event that $(n+1)$-th vote will be for winner,

• C is the event that during counting the first n votes the loser will be always the leader.

By the well-known formula for the hypergeometric distribution we get

$$P(A) = \frac{\binom{M}{\frac{n+1}{2}}\binom{N-M}{\frac{n-1}{2}}}{\binom{N}{n}}.$$

Moreover

$$P(B/A) = \frac{N-M-\frac{n-1}{2}}{N-n}.$$

On the other hand, by the Ballot Theorem (2) for $N=n$ and $M=\frac{n-1}{2}$

$$P(C/A \cap B) = P(C/A) = \frac{n-(n-1)}{n} = \frac{1}{n}.$$

In consequence for $n = 1, 3, ..., 2M-1$

$$P_M(T=n, L \neq W) = \frac{1}{n}\frac{N-M-\frac{n-1}{2}}{N-n}\frac{\binom{M}{\frac{n+1}{2}}\binom{N-M}{\frac{n-1}{2}}}{\binom{N}{n}}.$$

By some elementary operations on the binomial coefficients the last one reduces to

$$P_M(T=n, L \neq W) = \frac{1}{n}\binom{n}{\frac{n+1}{2}}\binom{N-n-1}{M-\frac{n+1}{2}}\binom{N}{M}^{-1} \tag{5}$$

for all $M > 0$ and for all $n = 1, 3, ..., 2M-1$.

Finally, by collecting the formulae (3), (4) and (5) we get the desired result (1). ■

In the next section we will predict the final number M of votes for loser by the Maximum Likelihood.

4 Predicting the final number of votes for winner and loser

Let us recall that under given leadership time n the Likelihood Function is a function of the unknown parameter M defined by the formula

$$\mathscr{L}_n(M) = P_M(T = n)$$

and the Maximum Likelihood predictor $\widehat{M}(n)$ of M is defined by the argument of $\mathscr{L}_n$ realizing its maximum. This may be expressed more precisely in the form

$$\widehat{M}(n) = \arg \max_{M \in \{0,1,\ldots,\frac{N-1}{2}\}} \mathscr{L}_n(M).$$

We shall prove

Theorem 2 *The Maximum Likelihood predictor $\widehat{M}(n)$ of the final number of votes for loser based on the leadership time n is given by*

$$\widehat{M}(n) = \begin{cases} 0, & \text{if } n = N, \\ \frac{N-1}{2}, & \text{if } n < N \end{cases}$$

while the ML predictor of the final number of votes for winner is given by

$$\widehat{N-M}(n) = \begin{cases} N, & \text{if } n = N, \\ \frac{N+1}{2}, & \text{if } n < N. \end{cases}$$

Proof. For $n = N$ the probability P_M attains its maximum for $M = 0$ and hence $\widehat{M}(N) = 0$.

For all odd positive integers $n < N$ the Likelihood Function is defined by the formula

$$\mathscr{L}_n(M) = \begin{cases} \frac{2}{n} \binom{n}{\frac{n+1}{2}} \binom{N-n-1}{M-\frac{n+1}{2}} \binom{N}{M}^{-1}, & \text{for } M \in \{\frac{n+1}{2},\ldots,\frac{N-1}{2}\}, \\ 0, & \text{otherwise} \end{cases}$$

and the ML predictor $\widehat{M}(n)$ of M is given by

$$\widehat{M}(n) = \arg \max_{M \in \left\{\frac{n+1}{2},\ldots,\frac{N-1}{2}\right\}} \mathscr{L}_n(M).$$

We will show that in this case

$$\arg \max_{M \in \left\{\frac{n+1}{2},\ldots,\frac{N-1}{2}\right\}} \mathscr{L}_n(M) = \frac{N-1}{2}.$$

To this aim we only need to verify that

$$\frac{\mathfrak{L}_n(M+1)}{\mathfrak{L}_n(M)} > 1 \tag{6}$$

for all integers M belonging to the interval $\left[\frac{n+1}{2}, \frac{N-1}{2}\right)$.

Indeed, for such M

$$\frac{\mathfrak{L}_n(M+1)}{\mathfrak{L}_n(M)} = \frac{(M+1)(N-M-k)}{(M-k+1)(N-M)}, \text{ where } k = \frac{n+1}{2}.$$

Thus the condition

$$\frac{\mathfrak{L}_n(M+1)}{\mathfrak{L}_n(M)} > 1$$

may be presented in the form

$$(M+1)(N-M-k) > (M-k+1)(N-M).$$

Since the last inequality holds for all integers $M \in \left[\frac{n+1}{2}, \frac{N-1}{2}\right)$, the desired condition (6) was verified.

Reassuming, the ML predictor of the final number of votes for loser based on the leadership time n is given by

$$\widehat{M}(n) = \begin{cases} 0, & \text{if } n = N, \\ \frac{N-1}{2}, & \text{if } n < N. \end{cases}$$

Finally, by the well-known fact that the results of the ML estimation do not depend on the parametrization (see, for instance, Schervish (1995, Th. 5.28, p. 308)) we get the predictor of the final number of votes for winner in the form

$$\widehat{N-M}(n) = \begin{cases} N, & \text{if } n = N, \\ \frac{N+1}{2}, & \text{if } n < N. \end{cases}$$

■

Therefore, the ML predictor of the final number of votes for winner does not depend on the first leadership time n unless $n = N$. This leads to the following conclusion.

5 Conclusion

The first leadership time is informative for the final results of the election only in the trivial case.

Acknowledgements

The author thanks the reviewers of the manuscript for their valuable suggestions.

REFERENCES

AZZALINI, A., (1996). Statistical Inference based on the Likelihood, Chapman & Hall, London, UK.

BRÉMAUD, P., (1994). An Introduction to Probabilistic Modelling, 2nd ed., Springer-Verlag, New York.

FELLER, W., (1968). An Introduction to Probability Theory and its Applications, Vol. 1, 3rd ed., Wiley, New York.

GOULDEN, I. P., SERRANO, L. G., (2003). Maintaining the spirit of the reflection principle when the boundary has arbitrary integer slope, J. Comb. Theory Ser. A, 104, pp. 317–326,

LENGYEL, T., (2011). Direct consequences of the basic Ballot Theorem, Statist. Probab. Lett., 81, pp. 1476–1481.

RENAULT, M., (2007). Four proofs of the Ballot Theorem, Math. Mag., 80, pp. 345–352.

SCHERVISH, M. J., (1995). Theory of Statistics, Springer-Verlag, New York.

STĘPNIAK, C., (2015). On distribution of leadership time in counting votes and predicting winners, Statist. Probab. Lett., 106, pp. 109–112.

TAKACS, L., (1997). On the ballot theorem, In: Advances in Combinatorial Methods and Applications to Probability and Statistics, Birkhauser, pp. 97–114.