Statistics and Probability Letters 101 (2015) 40–48

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

The truncated geometric election algorithm: Duration of the election

Guy Louchard [a], Mark Daniel Ward [b,*]

[a] Département d'Informatique, Université Libre de Bruxelles, CP 212, Boulevard du Triomphe, B-1050, Bruxelles, Belgium

[b] Department of Statistics, Purdue University, West Lafayette, IN 47907, USA

ARTICLE INFO

Article history:
Received 12 January 2015
Received in revised form 18 February 2015
Accepted 18 February 2015
Available online 26 February 2015

MSC:
60C05
68W40

Keywords:
Analysis of algorithms
Asymptotic analysis
Leader election
Generating function
Recurrence

ABSTRACT

The present paper makes three distinct improvements over an earlier investigation of Kalpathy and Ward. We analyze the length of the entire election process (not just one participant's duration), for a randomized election algorithm, with a truncated geometric number of survivors in each round. We not only analyze the mean and variance; we analyze the asymptotic distribution of the entire election process. We also introduce a new variant of the election that guarantees a unique winner will be chosen; this methodology should be more useful in practice than the previous methodology. The method of analysis includes a precise analytic (complex-valued) approach, relying on singularity analysis of probability generating functions.

1. Motivation

This is a sequel paper to Kalpathy and Ward (2014), which also appears in Statistics and Probability Letters. The earlier paper introduces a randomized leader election algorithm, in which a truncated geometric number of participants survive from round to round. (Louchard and Prodinger (2009), constitutes a good starting point for readers who are unfamiliar with randomized leader election algorithms; they also have extensive references, to provide a broader context.)

Kalpathy and Ward (2014) precisely analyze the number of rounds that a *particular* contestant survives in the election. That analysis, however, had three key shortcomings. (1) It also focused on only the duration of a particular contestant; it did not analyze the length of the entire election (which should be more interesting and more useful in practice). (2) It only analyzed the mean and variation of the duration of a particular contestant (who was not necessarily the winner), but it did not discuss the asymptotic distribution of the duration (the analysis contained in the present investigation is more informative and useful). (3) The method of election in the earlier paper did not guarantee a unique winner.

The present paper addresses all three of these issues with the earlier paper. We focus on the duration of the entire election, not just of one participant. We go beyond the mean and variance, and analyze also the asymptotic distribution of the entire election. We also analyze two variants of the election, namely, the version from the original paper, and also a new style of election in which a unique winner is guaranteed to appear at the end of the election process.

* Corresponding author.
E-mail addresses: louchard@ulb.ac.be (G. Louchard), mdw@purdue.edu (M.D. Ward).

http://dx.doi.org/10.1016/j.spl.2015.02.018

As another point of motivation for this sequel paper, the authors discover (in the proofs of Theorems 3.1 and 3.4) that the total length of each of these styles of election, when starting with n participants, can be decomposed into a sum of $n-1$ independent random variables that do not have identical distributions. This surprising point about the decomposition of the length of the whole election is enlightened in the analysis by using moment generating functions. More precisely (in the proof of Theorem 3.1), from the representation of $\phi_n(t) := \mathbb{E}(e^{tX_n})$ for $n \geq 2$, shown in Eq. (1), it is surprising to see a decomposition of X_n in a sum of $n-1$ independent random variables. Indeed, we see that X_n has the same distribution as $Z_2 + Z_3 + \cdots + Z_n$, where the Z_j's are independent, non-negative random variables, and Z_2 has moment generating function $\phi_2(t) = \frac{e^t p(1+q)}{1-q^2(q+e^t p)}$, while Z_j has moment generating function $\frac{1-q^j}{1-q^j(q+e^t p)}$ (for $3 \leq j \leq n$). An analogous decomposition holds in the proof of Theorem 3.4, because as we see in Eq. (5), we have a decomposition of Y_n into a sum of $n-1$ independent, non-negative random variables that do not have identical distributions.

2. Definitions

We consider elections for which, if n contestants are present in a round, then K_n contestants proceed to the next round, where K_n is a *truncated geometric random variable* with parameters p and $q := 1-p$. So the mass of K_n is

$$\mathbf{P}(K_n = \ell) = \frac{pq^\ell}{1-q^{n+1}}, \quad \text{for } \ell = 0, 1, \ldots, n.$$

We study the number of rounds needed for the election is two distinct situations. The difference occurs in how the one of the base conditions is handled, namely, what happens when $K_n = 0$.

1. **Setup #1.** As in Kalpathy and Ward (2014), if $K_n = 0$ in one of the rounds, the election stops, and the remaining n participants can all be treated as winners, or (alternatively) all considered as losers, but the main point is that no additional rounds of the election take place. (It is easy to show that the Kalpathy–Ward election, starting with n participants, will end without a unique winner with probability $\frac{1}{1+q}\prod_{j=3}^{n}\frac{1-q^j}{(1-pq^j-q^{j+1})}$.)
2. **Setup #2.** In the literature, it is very common that, when all of the current participants fail to advance to the next round, all of them are given another chance to participate in one more (renewal) round.

We define X_n and Y_n as the number of rounds, when starting with n participants, for the elections given in setups #1 and #2, respectively.

For example, consider an election with 20 initial participants. If 8 survive in round 1 (and 12 are eliminated), and 5 survive in round 2 (and 3 are eliminated), and 0 survive in round 3 (all 5 are eliminated), then $X_{20} = 3$ because 3 rounds were needed for the election. On the other hand, in such a situation, setup #2 mandates that the 5 participants from round 3 are all resurrected and get to continue for subsequent rounds. Suppose that these 5 participants are resurrected, and exactly 1 of the 5 survives in round 4. Then $Y_{20} = 4$ because 4 rounds were needed for the election.

3. Main results

First we give some results about the Kalpathy–Ward model from Kalpathy and Ward (2014) (Setup #1).

Theorem 3.1. *The expected number of rounds* $\mathbf{E}(X_n)$ *in an election (in setup #1), for* $n \geq 2$*, is*

$$\mathbf{E}(X_n) = p + \sum_{\ell=1}^{\infty}\frac{pq^\ell}{1-q^\ell} - \sum_{k=0}^{\infty}\frac{pq^{n+k+1}}{1-q^{n+k+1}}.$$

The variance $\mathbf{Var}(X_n)$ *is*

$$\mathbf{Var}(X_n) = \sum_{\ell=1}^{\infty}\frac{pq^\ell(1-q^{\ell+1})}{(1-q^\ell)^2} - q(q+1) - \sum_{k=0}^{\infty}\frac{pq^{n+k+1}(1-q^{n+k+2})}{(1-q^{n+k+1})^2}.$$

Remark 3.1. Since $C_1 := p + \sum_{\ell=1}^{\infty}\frac{pq^\ell}{1-q^\ell}$ is a constant (depending only on p an q, but not on n), we have

$$\mathbf{E}(X_n) = C_1 - q^{n+1} + \Theta(q^{2n}).$$

Similarly, because $C_2 := \sum_{\ell=1}^{\infty}\frac{pq^\ell(1-q^{\ell+1})}{(1-q^\ell)^2} - q(q+1)$ is a constant, then we have

$$\mathbf{Var}(X_n) = C_2 - q^{n+1} + \Theta(q^{2n}).$$

Using Maple (with some human guidance), we can derive more terms in the asymptotic expansion, if desired. We can also analyze higher moments of X_n with the methods found in the proofs section.

Theorem 3.2. *The probability that the number of rounds X_n in an election (in setup #1) is equal to k is exponentially decreasing as $k \to \infty$. Namely, for $n \geq 2$, we have*

$$P(X_n = k) \sim \frac{p(q+1)z_2}{(1-q^3)} \left(\prod_{j=3}^{n} \frac{1-q^j}{(1-q^{j+1})(1-z_2/z_j)} \right) (z_2)^{-k} \quad \text{as } k \to \infty,$$

where $z_j = \frac{1-q^{j+1}}{pq^j}$.

For an example of the rapid decrease in $\lim_{n\to\infty} P(X_n = k)$, when $p = 3/10$ and $q = 7/10$, we have $z_2 = 4.469387755\ldots$.

Theorem 3.3. *The limiting probability that the number of rounds in an election (in setup #1) is equal to k is exponentially decreasing as $k \to \infty$. We have*

$$\lim_{n\to\infty} P(X_n = k) \sim (r)(z_2)^{-k+1} \quad \text{as } k \to \infty,$$

where

$$r := \sum_{\ell=2}^{\infty} pq^{\ell} \frac{p(q+1)z_2}{(1-q^3)} \left(\prod_{j=3}^{\ell} \frac{1-q^j}{(1-q^{j+1})(1-z_2/z_j)} \right),$$

with $z_j = \frac{1-q^{j+1}}{pq^j}$.

Now we give some results about a modification to the model of election, so that the election is guaranteed to end with a unique winner (Setup #2).

Theorem 3.4. *The expected number of rounds $\mathbf{E}(Y_n)$ in an election (in setup #2), for $n \geq 2$, is*

$$\mathbf{E}(Y_n) = \frac{q+1}{q} + \sum_{\ell=1}^{\infty} \frac{pq^{\ell}}{1-q^{\ell}} - \sum_{k=0}^{\infty} \frac{pq^{n+k}}{1-q^{n+k}}.$$

The variance $\mathbf{Var}(Y_n)$ is

$$\mathbf{Var}(Y_n) = \frac{q+1}{q^2} + \sum_{\ell=1}^{\infty} \frac{pq^{\ell-1}(1+p-q^{\ell+2})}{(1-q^{\ell})^2} - \sum_{k=0}^{\infty} \frac{pq^{n+k-1}(1+p-q^{n+k+2})}{(1-q^{n+k})^2}.$$

Remark 3.2. Since $C_3 := \frac{q+1}{q} + \sum_{\ell=1}^{\infty} \frac{pq^{\ell}}{1-q^{\ell}}$ is a constant, it follows that

$$\mathbf{E}(Y_n) = C_3 - q^n + \Theta(q^{2n}).$$

Similarly, because $C_4 := \frac{q+1}{q^2} + \sum_{\ell=1}^{\infty} \frac{pq^{\ell-1}(1+p-q^{\ell+2})}{(1-q^{\ell})^2}$ is a constant, then we have

$$\mathbf{Var}(Y_n) = C_4 - (1+p)q^{n-1} + \Theta(q^{2n}).$$

Again, using Maple, more asymptotic terms and larger moments can both be achieved by the same methods discussed in the proofs.

Theorem 3.5. *The probability that the number of rounds Y_n in an election (in setup #2) is equal to k is exponentially decreasing as $k \to \infty$. Namely, for $n \geq 2$, we have*

$$P(Y_n = k) \sim \frac{pqz_2^*}{(1-q^3)} \left(\prod_{j=3}^{n} \frac{1-pz_2^*-q^j}{(1-q^{j+1})(1-z_2^*/z_j^*)} \right) (z_2^*)^{-k} \quad \text{as } k \to \infty,$$

where $z_j^* = \frac{1-q^{j+1}}{p(q^j+1)}$.

For an example of the rapid decrease in $\lim_{n\to\infty} P(Y_n = k)$, when $p = 3/10$ and $q = 7/10$, we have $z_2^* = 1.469798658\ldots$.

Theorem 3.6. *The limiting probability that the number of rounds in an election (in setup #2) is equal to k is exponentially decreasing as $k \to \infty$. We have*

$$\lim_{n\to\infty} P(Y_n = k) \sim \frac{Rz_2^*}{1-pz_2^*} (z_2^*)^{-k} \quad \text{as } k \to \infty,$$

where

$$R := \sum_{\ell=2}^{\infty} pq^{\ell} \frac{pqz_2^*}{(1-q^3)} \left(\prod_{j=3}^{\ell} \frac{1-pz_2^*-q^j}{(1-q^{j+1})(1-z_2^*/z_j^*)} \right),$$

with $z_j^* = \frac{1-q^{j+1}}{p(q^j+1)}$.

4. Proofs

The random variables X_n, Y_n, and K_n were all defined in Section 2.

Proof of Theorem 3.1. If $n = 0$ or $n = 1$, no election is needed, and thus $X_n = 0$ for $n = 0$ and for $n = 1$. We use the notation $\phi_n(t) = \mathbf{E}(e^{tX_n})$ to denote the moment generating function of X_n.

In the cases $n = 0$ and $n = 1$, we have $\phi_0(t) = \phi_1(t) = 1$.

For $n \geq 2$, we use conditional expectation to get the recursion

$$\phi_n(t) = \mathbf{E}(e^{tX_n}) = \sum_{\ell=0}^{n} \mathbf{E}(e^{tX_n} \mid K_n = \ell)\, \mathbf{P}(K_n = \ell),$$

and $\mathbf{E}(e^{tX_n} \mid K_n = \ell) = \mathbf{E}(e^{t(1+X_\ell)}) = e^t\phi_\ell(t)$, and $\mathbf{P}(K_n = \ell) = \frac{pq^\ell}{1-q^{n+1}}$. So we obtain

$$\phi_n(t) = e^t \sum_{\ell=0}^{n} \phi_\ell(t) \frac{pq^\ell}{1-q^{n+1}} \quad \text{for } n \geq 2.$$

Multiplying on both sides by $1 - q^{n+1}$, and subtracting the $\ell = n$ term, it follows that

$$(1 - q^{n+1} - e^t pq^n)\phi_n(t) = e^t \sum_{\ell=0}^{n-1} \phi_\ell(t) pq^\ell.$$

Shifting the n by 1, we obtain, for $n \geq 3$, $(1 - q^n - e^t pq^{n-1})\phi_{n-1}(t) = e^t \sum_{\ell=0}^{n-2} \phi_\ell(t)pq^\ell$. Using differencing, it follows that

$$(1 - q^{n+1} - e^t pq^n)\phi_n(t) - (1 - q^n - e^t pq^{n-1})\phi_{n-1}(t) = e^t\phi_{n-1}(t)pq^{n-1} \quad \text{for } n \geq 3.$$

Thus, we obtain

$$\phi_n(t) = \phi_{n-1}(t) \frac{1-q^n}{1-q^n(q+e^t p)} \quad \text{for } n \geq 3.$$

Telescoping, this yields,

$$\phi_n(t) = \phi_2(t) \prod_{j=3}^{n} \frac{1-q^j}{1-q^j(q+e^t p)};$$

this representation is seen to hold not only for $n \geq 3$, but for $n = 2$ as well. We manually compute $\phi_2(t) = e^t p(1+q)/(1 - q^2(q + e^t p))$. So we conclude, for $n \geq 2$, that

$$\phi_n(t) = \frac{e^t p(1+q)}{1-q^2(q+e^t p)} \prod_{j=3}^{n} \frac{1-q^j}{1-q^j(q+e^t p)}. \tag{1}$$

From the representation of $\phi_n(t)$ in Eq. (1), it is surprising to see a decomposition of X_n in a sum of random variables. Indeed, we see that X_n has the same distribution as $Z_2 + Z_3 + \cdots + Z_n$, where the Z_j's are independent, and Z_2 has moment generating function $\phi_2(t) = \frac{e^t p(1+q)}{1-q^2(q+e^t p)}$, and Z_j (for $3 \leq j \leq n$) has moment generating function $\frac{1-q^j}{1-q^j(q+e^t p)}$. Computing the expected values of each of the Z_j's, we obtain

$$\mathbf{E}(Z_2) = \lim_{t\to 0} \frac{d}{dt} \frac{e^t p(1+q)}{1-q^2(q+e^t p)} = \frac{q^2+q+1}{q+1},$$

and, for $3 \leq j \leq n$,

$$\mathbf{E}(Z_j) = \lim_{t\to 0} \frac{d}{dt} \frac{1-q^j}{1-q^j(q+e^t p)} = \frac{pq^j}{1-q^j}.$$

Therefore, we conclude

$$\mathbf{E}(X_n) = \frac{q^2+q+1}{q+1} + \sum_{j=3}^{n} \frac{pq^j}{1-q^j} = p + \sum_{\ell=1}^{\infty} \frac{pq^\ell}{1-q^\ell} - \sum_{k=0}^{\infty} \frac{pq^{n+k+1}}{1-q^{n+k+1}}.$$

This establishes the first half of Theorem 3.1.

Computing the second moments and then the variances of each of the Z_j's, we obtain

$$\mathbf{E}(Z_2^2) = \lim_{t\to 0} \frac{d^2}{dt^2} \frac{e^t p(1+q)}{1-q^2(q+e^t p)} = \frac{p(1+q)(1-q^3)(1-q^2+2pq^2)}{(1-q^2)^3},$$

and

$$\mathbf{Var}(Z_2) = \mathbf{E}(Z_2^2) - (\mathbf{E}(Z_2))^2 = \frac{p(1+q)(1-q^3)(1-q^2+2pq^2)}{(1-q^2)^3} - \left(\frac{q^2+q+1}{q+1}\right)^2 = \frac{(q^2+q+1)q^2}{(q+1)^2},$$

and, for $3 \le j \le n$,

$$\mathbf{E}(Z_j^2) = \lim_{t\to 0} \frac{d^2}{dt^2} \frac{1-q^j}{1-q^j(q+e^t p)} = \frac{pq^j(1+q^j(1-2q))}{(1-q^j)^2},$$

and

$$\mathbf{Var}(Z_j) = \mathbf{E}(Z_j^2) - (\mathbf{E}(Z_j))^2 = \frac{pq^j(1+q^j(1-2q))}{(1-q^j)^2} - \left(\frac{pq^j}{1-q^j}\right)^2 = \frac{pq^j(1-q^{j+1})}{(1-q^j)^2}.$$

Thus, we have

$$\mathbf{Var}(X_n) = \frac{(q^2+q+1)q^2}{(q+1)^2} + \sum_{j=3}^{n} \frac{pq^j(1-q^{j+1})}{(1-q^j)^2} = \sum_{\ell=1}^{\infty} \frac{pq^\ell(1-q^{\ell+1})}{(1-q^\ell)^2} - q(q+1) - \sum_{k=0}^{\infty} \frac{pq^{n+k+1}(1-q^{n+k+2})}{(1-q^{n+k+1})^2}.$$

This establishes the second half of Theorem 3.1.

Proof of Theorem 3.2. Now we analyze the asymptotic distribution of X_n. We proceed as in Lavault and Louchard (2006, Section 3). For $n \ge 2$, we have

$$P(X_n = 1) = \frac{p}{1-q^{n+1}} + \frac{pq}{1-q^{n+1}} = \frac{p(q+1)}{1-q^{n+1}},$$

and therefore $\lim_{n\to\infty} P(X_n = 1) = p(q+1)$. For $n \ge 2$ and $j \ge 2$, we have the recurrence

$$P(X_n = j) = \sum_{\ell=2}^{n} \frac{pq^\ell}{1-q^{n+1}} P(X_\ell = j-1). \tag{2}$$

It follows that

$$\lim_{n\to\infty} P(X_n = j) = \sum_{\ell=2}^{\infty} pq^\ell P(X_\ell = j-1). \tag{3}$$

We define an ordinary generating function

$$h(z) := \sum_{j=1}^{\infty} \lim_{n\to\infty} P(X_n = j) z^j.$$

Now we define the probability generating function of X_n as

$$f_n(z) = \sum_{j=1}^{\infty} P(X_n = j) z^j.$$

From Eq. (2), it follows that, for $n \ge 2$,

$$f_n(z) = \frac{p(q+1)z}{1-q^{n+1}} + \sum_{\ell=2}^{n} \frac{pq^\ell}{1-q^{n+1}} z f_\ell(z). \tag{4}$$

When $n = 2$, it follows that

$$f_2(z) = \frac{p(q+1)z}{(1-q^3)(1-z/z_2)},$$

where $z_2 = \frac{1-q^3}{pq^2}$. Returning to Eq. (4), we multiply by $1 - q^{n+1}$, subtract the $F_n(z)$ terms on each side, shift the n by 1, and use differencing. It follows that, for $n \geq 3$, we have

$$f_n(z) = \frac{1 - q^n}{1 - pq^n z - q^{n+1}} f_{n-1}(z).$$

Telescoping, this yields, for $n \geq 2$, an explicit expression for $f_n(z)$:

$$f_n(z) = \frac{p(q+1)z}{(1-q^3)(1-z/z_2)} \prod_{j=3}^{n} \frac{1-q^j}{1-pq^j z - q^{j+1}}.$$

Now we define $z_j := \frac{1-q^{j+1}}{pq^j}$, so we can rewrite the expression for $F_n(z)$ as:

$$f_n(z) = \frac{p(q+1)z}{(1-q^3)(1-z/z_2)} \prod_{j=3}^{n} \frac{1-q^j}{(1-q^{j+1})(1-z/z_j)}.$$

Since $1 < z_2 < z_3 < z_4 < \cdots$, it follows that z_2 is a simple pole (i.e., a pole of order 1) of $f_n(z)$. Thus, we can use singularity analysis (Flajolet and Sedgewick, 2009, Chapter 6) to extract the first order asymptotic growth of $P(X_n = k)$ as $k \to \infty$, namely, for $n \geq 2$,

$$P(X_n = k) = [z^k]f_n(z) \sim \frac{p(q+1)z_2}{(1-q^3)} \left(\prod_{j=3}^{n} \frac{1-q^j}{(1-q^{j+1})(1-z_2/z_j)} \right) (z_2)^{-k} \quad \text{as } k \to \infty.$$

Proof of Theorem 3.3. Now we define another ordinary generating function $g(z)$ as

$$g(z) := \sum_{j=1}^{\infty} \sum_{\ell=2}^{\infty} pq^\ell P(X_\ell = j) z^j.$$

It follows from (3) that

$$h(z) := p(q+1)z + zg(z).$$

We have $g(z) = \sum_{\ell=2}^{\infty} pq^\ell f_\ell(z)$. Since each $f_\ell(z)$ has a simple pole at z_2, and this pole is the closest singularity to the origin, it follows that $g(z)$'s closest singularity to the origin is also a simple pole at z_2. Thus, if we define

$$r := \sum_{\ell=2}^{\infty} pq^\ell \frac{p(q+1)z_2}{(1-q^3)} \left(\prod_{j=3}^{\ell} \frac{1-q^j}{(1-q^{j+1})(1-z_2/z_j)} \right),$$

it follows that $g(z) \asymp r/(1 - z/z_2)$ as $z \to z_2$. Since $g(z)$'s closest singularity to the origin is a pole at z_2, it follows that $h(z) \asymp p(q+1)z_2 + z_2 r/(1 - z/z_2)$ as $z \to z_2$, and moreover,

$$\lim_{n\to\infty} P(X_n = k) = [z^k]h(z) \sim rz_2(z_2)^{-k} \quad \text{as } k \to \infty.$$

Proof of Theorem 3.4. This proof will mirror the proof of Theorem 3.1, so we omit some of the details. We mostly highlight the differences from the earlier proof, to keep the argument succinct.

If $n = 0$ or $n = 1$, no election is needed, and thus $Y_n = 0$ for $n = 0$ and for $n = 1$. We use $\psi_n(t) = \mathbf{E}(e^{tY_n})$ to denote the moment generating function of Y_n. As before, for $n = 0$ and $n = 1$, we have $\psi_0(t) = \psi_1(t) = 1$.

For $n \geq 2$, we again use conditional expectation to get a recursion, namely

$$\psi_n(t) = \mathbf{E}(e^{tY_n}) = \sum_{\ell=0}^{n} \mathbf{E}(e^{tY_n} \mid K_n = \ell)\, \mathbf{P}(K_n = \ell),$$

and $\mathbf{E}(e^{tY_n} \mid K_n = \ell) = \mathbf{E}(e^{t(1+Y_\ell)}) = e^t \psi_\ell(t)$ for $\ell \geq 1$, but the $\ell = 0$ case differs from setup #1, so we have

$$\mathbf{E}(e^{tY_n} \mid K_n = 0) = \mathbf{E}(e^{t(1+Y_n)}) = e^t \psi_n(t),$$

and as before, $\mathbf{P}(K_n = \ell) = \frac{pq^\ell}{1-q^{n+1}}$. So we obtain

$$\psi_n(t) = e^t \psi_n(t) \frac{p}{1-q^{n+1}} + e^t \sum_{\ell=1}^{n} \psi_\ell(t) \frac{pq^\ell}{1-q^{n+1}} \quad \text{for } n \geq 2.$$

We again multiply by $1-q^{n+1}$; subtract both ψ_n terms; shift the n by 1; and use differencing and telescoping; it follows that

$$\psi_n(t) = \frac{e^t pq}{1-q^3-e^tp-e^tpq^2}\prod_{j=3}^{n}\frac{1-q^j-e^tp}{1-e^tp-q^j(q+e^tp)} \quad \text{for } n \geq 2. \tag{5}$$

Again, it is surprising to see that Y_n can be decomposed as a sum of $n-1$ independent random variables with moment generating functions corresponding to the factors of the displayed equation for $\psi_n(t)$ written above. Thus, differentiating each MGF and taking $t \to 0$ in each factor, we obtain

$$\mathbf{E}(Y_n) = \frac{q^2+q+1}{q} + \sum_{j=3}^{n}\frac{pq^{j-1}}{1-q^{j-1}} = \frac{q+1}{q} + \sum_{\ell=1}^{\infty}\frac{pq^\ell}{1-q^\ell} - \sum_{k=0}^{\infty}\frac{pq^{n+k}}{1-q^{n+k}}.$$

This establishes the first half of Theorem 3.4.

Computing the second moments and then the variances of each of the terms in the decomposition of Y_n into $n-1$ independent random variables, we obtain

$$\begin{aligned}\mathbf{Var}(Y_n) &= \frac{(1-q^3)(1-q^3+p+pq^2)}{(pq)^2} - \left(\frac{q^2+q+1}{q}\right)^2 \\ &\quad + \sum_{j=3}^{n}\left(\frac{pq^{j+1}(1-q^{j-1})(1+pq^j-q^{j+1}+p)}{(1-pq^j-q^{j+1}-p)^3} - \left(\frac{pq^{j-1}}{1-q^{j-1}}\right)^2\right)\end{aligned}$$

which simplifies to

$$\begin{aligned}\mathbf{Var}(Y_n) &= \frac{(q^2+1)(q^2+q+1)}{q^2} + \sum_{j=3}^{n}\frac{pq^{j-2}(1+p-q^{j+1})}{(1-q^{j-1})^2} \\ &= \frac{q+1}{q^2} + \sum_{\ell=1}^{\infty}\frac{pq^{\ell-1}(1+p-q^{\ell+2})}{(1-q^\ell)^2} - \sum_{k=0}^{\infty}\frac{pq^{n+k-1}(1+p-q^{n+k+2})}{(1-q^{n+k})^2}.\end{aligned}$$

This establishes the second half of Theorem 3.4.

Proof of Theorem 3.5. Now we analyze the asymptotic distribution of Y_n. We proceed as in Lavault and Louchard (2006, Section 3). For $n \geq 2$, we have

$$P(Y_n = 1) = \frac{pq}{1-q^{n+1}},$$

and therefore $\lim_{n\to\infty} P(Y_n = 1) = pq$. For $n \geq 2$ and $j \geq 2$, we have the recurrence

$$P(Y_n = j) = \frac{p}{1-q^{n+1}}P(Y_n = j-1) + \sum_{\ell=2}^{n}\frac{pq^\ell}{1-q^{n+1}}P(Y_\ell = j-1). \tag{6}$$

It follows that

$$\lim_{n\to\infty} P(Y_n = j) = p\lim_{n\to\infty} P(Y_n = j-1) + \sum_{\ell=2}^{\infty} pq^\ell P(Y_\ell = j-1). \tag{7}$$

Unlike in the analogous proof for X_n, we now need to iterate the previous equation $j-1$ times. This yields, for $j \geq 2$,

$$\lim_{n\to\infty} P(Y_n = j) = p^{j-1}\lim_{n\to\infty} P(Y_n = 1) + \sum_{k=1}^{j-1} p^{k-1}\sum_{\ell=2}^{\infty} pq^\ell P(Y_\ell = j-k).$$

We define an ordinary generating function

$$H(z) := \sum_{j=1}^{\infty}\lim_{n\to\infty} P(Y_n = j)z^j.$$

Unlike in the analogous case for X_n, the $H(z)$ has a singularity of its own (because of the denominator $1-pz$), namely, at $z_1^* := 1/p$. This is the closest singularity of $H(z)$ to the origin.

Now we define the probability generating function of Y_n as

$$F_n(z) = \sum_{j=1}^{\infty} P(Y_n = j)z^j.$$

From Eq. (6), it follows that, for $n \geq 2$,

$$F_n(z) = \frac{pqz}{1-q^{n+1}} + \frac{pz}{1-q^{n+1}}F_n(z) + \sum_{\ell=2}^{n} \frac{pq^\ell}{1-q^{n+1}} zF_\ell(z). \tag{8}$$

When $n = 2$, it follows that

$$F_2(z) = \frac{pqz}{(1-q^3)(1-z/z_2^*)},$$

where $z_2^* = \frac{1-q^3}{p(q^2+1)}$. Returning to Eq. (8), we multiply by $1-q^{n+1}$, subtract the $F_n(z)$ terms on each side, shift the n by 1, and use differencing. It follows that, for $n \geq 3$, we have

$$F_n(z) = \frac{1-pz-q^n}{1-pz-pq^nz-q^{n+1}}F_{n-1}(z).$$

Telescoping, this yields, for $n \geq 2$, an explicit expression for $F_n(z)$:

$$F_n(z) = \frac{pqz}{(1-q^3)(1-z/z_2^*)} \prod_{j=3}^{n} \frac{1-pz-q^j}{1-pz-pq^jz-q^{j+1}}.$$

Now we define $z_j^* := \frac{1-q^{j+1}}{p(q^j+1)}$, so we can rewrite the expression for $F_n(z)$ as:

$$F_n(z) = \frac{pqz}{(1-q^3)(1-z/z_2^*)} \prod_{j=3}^{n} \frac{1-pz-q^j}{(1-q^{j+1})(1-z/z_j^*)}.$$

As before, since $1 < z_2^* < z_3^* < z_4^* < \cdots$, it follows that z_2^* is a simple pole (i.e., of order 1) of $F_n(z)$. Thus, by singularity analysis, we extract the first order asymptotic growth of $P(Y_n = k)$ as $k \to \infty$, for $n \geq 2$:

$$P(Y_n = k) = [z^k]F_n(z) \sim \frac{pqz_2^*}{(1-q^3)} \left(\prod_{j=3}^{n} \frac{1-pz_2^*-q^j}{(1-q^{j+1})(1-z_2^*/z_j^*)} \right) (z_2^*)^{-k} \quad \text{as } k \to \infty.$$

Proof of Theorem 3.6. Now we define another ordinary generating function $G(z)$ as

$$G(z) := \sum_{j=1}^{\infty} \sum_{\ell=2}^{\infty} pq^\ell P(Y_\ell = j) z^j.$$

It follows from (7) that

$$H(z) = pqz + pzH(z) + zG(z),$$

or equivalently,

$$H(z) = \frac{pqz + zG(z)}{1-pz}.$$

We have $G(z) = \sum_{\ell=2}^{\infty} pq^\ell F_\ell(z)$. Since each $F_\ell(z)$ has a simple pole at z_2^*, and this pole is the closest singularity to the origin, it follows that $G(z)$'s closest singularity to the origin is also a simple pole at z_2^*. Thus, if we define

$$R := \sum_{\ell=2}^{\infty} pq^\ell \frac{pqz_2^*}{(1-q^3)} \left(\prod_{j=3}^{\ell} \frac{1-pz_2^*-q^j}{(1-q^{j+1})(1-z_2^*/z_j^*)} \right),$$

it follows that $G(z) \asymp R/(1-z/z_2^*)$ as $z \to z_2^*$. Since $H(z)$'s closest singularity to the origin is a pole at z_2^*, it follows that $H(z) \asymp \frac{pqz_2^* + z_2^* R/(1-z/z_2^*)}{1-pz_2^*}$ as $z \to z_2^*$, and moreover,

$$\lim_{n\to\infty} P(Y_n = k) = [z^k]H(z) \sim \frac{Rz_2^*}{1-pz_2^*}(z_2^*)^{-k} \quad \text{as } k \to \infty.$$

Final remarks

The number of survivors at the beginning and ending stages of the election can also be analyzed by similar techniques. The authors already conducted this analysis too. Interested readers are welcome to contact us, to obtain these further details.

Acknowledgments

M. D. Ward's work is supported by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-0939370, and by NSF DMS-1246818. He thanks R. Kalpathy and H. Mahmoud for their kind hospitality during a sabbatical visit to The George Washington University in 2013–14.

Both authors thank R. Kalpathy, who was a co-author with M. D. Ward on the earlier paper Kalpathy and Ward (2014), on which this paper builds.

References

Flajolet, P., Sedgewick, R., 2009. Analytic Combinatorics. Cambridge University Press.
Kalpathy, R., Ward, M.D., 2014. On a leader election algorithm: truncated geometric case study. Statist. Probab. Lett. 87, 40–47.
Lavault, C., Louchard, G., 2006. Asymptotic analysis of a leader election algorithm. Theoret. Comput. Sci. 359, 239–254.
Louchard, G., Prodinger, H., 2009. The asymmetric leader election algorithm. Ann. Comb. 12, 449–478.