Expected Survival Time of a Probabilistic Counting-Out Game on a Line

Tingting Ou, Michelle Shu and John C. Wierman
Department of Applied Mathematics and Statistics
Johns Hopkins University

Abstract

Our research studies the expected survival time in a novel problem posted on a question-and-answer website. There are n people in a line at positions $1, 2, \ldots, n$. For each round, we randomly select a person at position k, where k is odd, to leave the line, and shift the person at each position i such that $i > k$ to position $i-1$. We continue to select people until there is only one person left, who then becomes the winner. We are interested in which initial position has the largest expected number of turns to stay in the line before being selected, which we refer to as "expected survival time." In this paper, we use a recursive approach to solve for exact values of the expected survival time. We have proved the exact formula of the expected survival time of the first and the last position as well. We will also present our work on the expected survival time of the other positions, 2, 3, 4, ... from an asymptotic perspective.

1 Introduction

This article addresses a problem originally posted by ChengYiYi [2] on the Chinese Question-and-Answer website Zhihu. There are n people in a line at positions $1, 2, \ldots, n$. For each round, we randomly select a person at position k, where k is odd, to leave the line, and shift the person at each position i such that $i > k$ to position $i - 1$. We continue to select people until there is only one person left, who then becomes the winner. The question is, which initial position is the most favorable? In this paper, we will answer the question based on the expected number of turns to stay in the line before being selected and forced to quit the game, which we refer to as *expected survival time.* Of course, the initial position that has the longest expected survival time would be the most favorable. A companion paper by Shu, Ou and Wierman studies the same problem, but will give a different answer based on the probability to survive all rounds of elimination and win the game eventually.

The problem studied in this paper resembles the famous Josephus Problem [1]. In the Josephus problem, some players stand in a circle and one is chosen

CONGRESSUS NUMERANTIUM 234 (2019), pp.195-212

randomly to be the starting point. For each round, in a specified direction, we skip a certain number of people and execute the next. The procedure is repeated until only one person remains, and is then freed. Mathematicians and computer scientists studying the Josephus problem are interested in which position in the initial circle can avoid execution, given the total number of players, the direction, the starting point and the number of people to skip in each round. Both the Josephus problem and our problem are variations of the counting-out games, in which players are eliminated one-by-one until there is only one person left. However, our problem is different in that people stand in a line instead of a circle, and the elimination process is probabilistic rather than deterministic.

As mentioned previously, this problem was posted on the Chinese Question-and-Answer website Zhihu. The original problem was only concerned with the special case where the total number of people in the line initially is 600. Viewed more than 600,000 times and followed by 3,184 users on the website, the question has drawn a great deal of attention during the past two years. There are 123 answers to the question, but most of them are completely based on computer simulations and include work only about the winning probability rather than the expected survival time of each person. XieZhuoFan [3] has given the exact values of the winning probability and the expected survival time when $n = 600$ using recursive calculation methods, but no analytical solution is provided. In this article, we will assume the total number of people in the line initially is n, an arbitrary positive integer, and present our results obtained numerically as well as analytically. The generalized problem is also posted on Quora [4].

For notational convenience, the person initially standing at the k-th position is referred to as "Person k," $1 \leq k \leq n$. If Person k gets shifted, we still refer to this person as Person k, but at the $(k-1)$-th position.

The paper is structured as follows:

In section 2, recursive equations are constructed, then the exact values of the expected survival time for each person are determined recursively.

In section 3, the expected survival time of Person 1 is calculated exactly as a function of n.

Lower bounds and upper bounds for the expected survival time of Person 2 are provided in section 4. An asymptotic approach is also introduced in this section.

The expected survival time of Persons 3 and 4 are calculated asymptotically in section 5. The method explained in this section may be used to calculate the asymptotic survival time for other persons (5, 6 and so on) as well.

Section 6 gives an exact formula of the expected survival time of Person n (the last person), which is proved by induction.

Section 7 is a brief summary of our results. We conclude the paper with section 8, which mentions the future direction of this research and some open questions.

2 Recursive Methods and Calculations

Let $E_n(k)$ denote the expected survival time of Person k in a queue of n people. In this paper, we define "survival time" to be the number of rounds a person survives in the game, excluding the first round. That is to say, if Person k gets eliminated in the first round, then the survival time is 0. If s/he becomes the winner, then the survival time is $n-1$.

To start with, we solve the base cases when n is small.

If $n = 1$, we only have one person in the line initially, so s/he wins the game in the first round. Therefore, $E_1(1) = 0$.

If $n = 2$, Person 1 will be eliminated in the first round because s/he is the only odd-indexed person, so $E_2(1) = 0$. Person 2 will be the only person in the second round, and the game terminates. Thus, $E_2(2) = 1$.

We will now introduce the idea of the recursion. Assume that a person survives the first round of the game, then in the next round this person either gets shifted to the previous position (if someone in front of this person gets selected) or remains at the same position (if someone behind this person gets selected). If the person is selected in the first round, then the survival time is 0 by our definition, and will not contribute to the calculation of the expected survival time. In the next round, the game of n people reduces to a smaller game with $n-1$ people initially.

Take $n = 3$ as an example. We will calculate the expected survival time of Person 1 first. Person 1 will never be shifted. Therefore, if Person 1 survives the first round of selection, then it must be the case that Person 3 is selected in the first round with probability $\frac{1}{2}$ and that Person 1 remains at the first position in the next round. Starting from the second round, Person 1 will be a new "Person 1" in the 2-people game, and is expected to survive for $E_2(1) = 0$ rounds in the new game. Thus, the expected survival time of Person 1 when $n = 3$ can be calculated as follows:

$$\begin{aligned} E_3(1) &= P(\text{remain at position 1}) \cdot (1 + E_2(1)) \\ &= \frac{1}{2} \cdot (1+0) \\ &= \frac{1}{2}. \end{aligned}$$

We now consider Person 2 when $n = 3$. If Person 1 is selected in the first round with probability $\frac{1}{2}$, then Person 2 will be shifted and become "Person 1" in the new game of 2 people. If Person 3 is selected in the first round with probability $\frac{1}{2}$, then Person 2 will remain at the same position and still be "Person 2" in the new game of 2 people. Thus, $E_3(2)$ can be calculated as follows:

$$\begin{aligned} E_3(2) = &\ P(\text{shifted to position 1}) \cdot (1 + E_2(1)) \\ &+ P(\text{remain at position 2}) \cdot (1 + E_2(2)) \\ = &\ \frac{1}{2} \cdot (1+0) + \frac{1}{2} \cdot (1+1) \\ = &\ \frac{3}{2}. \end{aligned}$$

If not selected, Person 3 will be shifted (with probability $\frac{1}{2}$) and become "Person 2" in the reduced 2-people game because s/he is already the last person in the line. S/he will never remain at the same position. Then, we have the following:

$$\begin{aligned} E_3(3) &= P(\text{shifted to position } 2) \cdot (1 + E_2(2)) \\ &= \frac{1}{2} \cdot (1+1) \\ &= 1. \end{aligned}$$

Using the idea above, we can construct the following recursive equation based on the first-step decomposition:

$$\begin{aligned} E_n(k) &= P(\text{shifted to position } (k-1)) \cdot E(\text{survival time} \mid \text{shifted}) \\ &\quad + P(\text{remain at position } k) \cdot E(\text{survival time} \mid \text{remain}) \\ &= P(\text{shifted to position } (k-1)) \cdot (1 + E_{n-1}(k-1)) \\ &\quad + P(\text{remain at position } k) \cdot (1 + E_{n-1}(k)). \end{aligned}$$

The two probabilities in the recursion can be expressed as a function of n and k. Given that there are n people in the line, $\lceil \frac{n}{2} \rceil$ people are odd-indexed. Also, $\lceil \frac{k-1}{2} \rceil$ odd-indexed people stand in front of Person k. If Person k is shifted, then an odd-indexed person in front must be selected, thus:

$$P(\text{shifted to position } (k-1)) = \frac{\lceil \frac{k-1}{2} \rceil}{\lceil \frac{n}{2} \rceil}.$$

Similarly, if Person k remains at the same position in the next round, then an odd-indexed person standing behind Person k must be selected in this round. There are $\lceil \frac{k}{2} \rceil$ odd-indexed people among the first k people. Therefore, there are $\left(\lceil \frac{n}{2} \rceil - \lceil \frac{k}{2} \rceil\right)$ odd-indexed people standing behind Person k. Thus, we have:

$$P(\text{remain at position } k) = \frac{\lceil \frac{n}{2} \rceil - \lceil \frac{k}{2} \rceil}{\lceil \frac{n}{2} \rceil} = 1 - \frac{\lceil \frac{k}{2} \rceil}{\lceil \frac{n}{2} \rceil}.$$

Using the above formulae of probabilities, the general recursive relation becomes:

$$\begin{aligned} E_n(k) &= P(\text{shifted to position } (k-1)) \cdot (1 + E_{n-1}(k-1)) + \\ &\quad P(\text{remain at position } k) \cdot (1 + E_{n-1}(k)) \\ &= \frac{\lceil \frac{k-1}{2} \rceil}{\lceil \frac{n}{2} \rceil}(1 + E_{n-1}(k-1)) + \left(1 - \frac{\lceil \frac{k}{2} \rceil}{\lceil \frac{n}{2} \rceil}\right)(1 + E_{n-1}(k)). \end{aligned}$$

Since we have already solved the base cases, we may use the above recursion to calculate the expectations bottom-up and solve for $E_n(k)$ for any n and k. Up to $n = 9$, using the recursion method, we have obtained the following table of results. Person 1 is abbreviated as P1, and the numbers are rounded to two decimal places.

n	P1	P2	P3	P4	P5	P6	P7	P8	P9
1	0								
2	0	1.00							
3	0.50	1.50	1.00						
4	0.75	2.00	1.25	2.00					
5	1.17	2.58	1.75	2.50	2.00				
6	1.44	3.11	2.11	3.00	2.33	3.00			
7	1.83	3.69	2.58	3.56	2.83	3.50	3.00		
8	2.13	4.23	2.97	4.07	3.24	4.00	3.38	4.00	
9	2.50	4.81	3.43	4.63	3.72	4.54	3.88	4.50	4.00

Table 1: Expected survival time of each person for $n = 1, ..., 9$.

It is instructive to plot the expected survival time versus the initial position index for different n. Figure 1 shows a plot when n varies from 2 (the bottom line) to 20 (the top line).

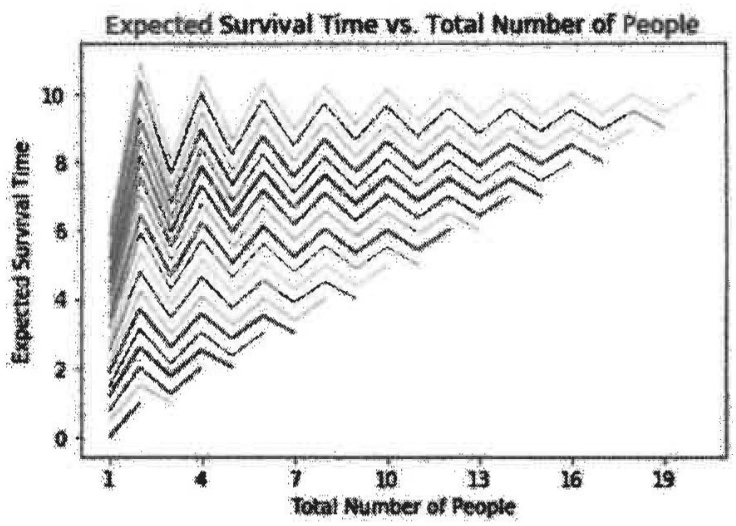

Figure 1: Expected survival time of each person for $n = 2, ..., 20$. For odd-indexed positions, the expected survival time increases. For even-indexed positions, the expected survival time decreases. The peak occurs at position 2.

XieZhuoFan [3] on Zhihu claims that the expected survival time is increasing with respect to the initial position index for odd-indexed people and decreasing with respect to the initial position index for even-indexed people. He also believes that Person 1 has the shortest expected survival time, approximately $\frac{1}{3}n$ and that Person 2 has the longest expected survival time, approximately $\frac{5}{9}n$, although no proof is provided.

We will prove that $E_n(1) = \frac{1}{3}n$ and $E_n(2) = \frac{5}{9}n$ asymptotically when n is sufficiently large in section 3 and 4. Based on our recursive calculations and plots, XieZhuoFan's first claim also appears to be true. In section 5, we will introduce a method to calculate the asymptotic expected survival time of any person, which may be used to prove XieZhuoFan's first claim.

3 The Expected Survival Time of Person 1

Although we can determine the expected survival time numerically from the bottom-up recursion, it would be useful to obtain a solution analytically without calculating all the expectations for each Person k and for each n.

If n is odd, then the expected survival time can be calculated exactly for Person 1. Let n be the initial number of people in the line and m be the initial number of odd-indexed people, so we have $m = \lceil \frac{n}{2} \rceil = \frac{n+1}{2}$ and $n = 2m - 1$. Let T be the survival time of Person 1. We will calculate the probability mass function of T.

If there are n people, $\lceil \frac{n}{2} \rceil$ of them are at odd-indexed positions. If $T = 0$, Person 1 should be selected out of the $\lceil \frac{n}{2} \rceil$ people in the first round. Then,

$$P(T=0) = P(\text{Person 1 selected in the 1st round}) = \frac{1}{\lceil \frac{n}{2} \rceil} = \frac{1}{m}.$$

If the survival time is t where $t > 0$, then Person 1 should have survived the first t rounds but get selected in the $(t+1)$-th round. The probabilities of T being $1, 2$ and 3 are calculated respectively as follows:

$$\begin{aligned} P(T=1) &= P(\text{selected in the 2nd round, survive the 1st round}) \\ &= \frac{1}{\lceil \frac{n-1}{2} \rceil}\left(1 - \frac{1}{\lceil \frac{n}{2} \rceil}\right) \\ &= \frac{1}{m-1}\left(\frac{m-1}{m}\right) \\ &= \frac{1}{m}. \end{aligned}$$

$$\begin{aligned} P(T=2) &= P(\text{selected in the 3nd round, survive the first 2 rounds}) \\ &= \frac{1}{\lceil \frac{n-2}{2} \rceil}\left(1 - \frac{1}{\lceil \frac{n}{2} \rceil}\right)\left(1 - \frac{1}{\lceil \frac{n-1}{2} \rceil}\right) \\ &= \frac{1}{m-1}\left(\frac{m-1}{m}\right)\left(\frac{m-2}{m-1}\right) \\ &= \frac{1}{m}\left(\frac{m-2}{m-1}\right). \end{aligned}$$

$$\begin{aligned}
P(T=3) &= P(\text{selected in the 4th round, survive the first 3 rounds}) \\
&= \frac{1}{\lceil \frac{n-3}{2} \rceil}\left(1-\frac{1}{\lceil \frac{n}{2} \rceil}\right)\left(1-\frac{1}{\lceil \frac{n-1}{2} \rceil}\right)\left(1-\frac{1}{\lceil \frac{n-2}{2} \rceil}\right) \\
&= \frac{1}{m-2}\left(\frac{m-1}{m}\right)\left(\frac{m-2}{m-1}\right)\left(\frac{m-2}{m-1}\right) \\
&= \frac{1}{m}\left(\frac{m-2}{m-1}\right).
\end{aligned}$$

Having recognized the pattern of cancellation of factors in the calculations, we can generalize the formula of the probability mass function of even T ($T = 2i$) and odd T ($T = 2i + 1$).

When T is even,

$$\begin{aligned}
P(T=2i) &= P(\text{selected in the } (2i+1)\text{th round, not the first } (2i) \text{ rounds}) \\
&= \frac{1}{\lceil \frac{n-2i}{2} \rceil}\left(1-\frac{1}{\lceil \frac{n}{2} \rceil}\right)\left(1-\frac{1}{\lceil \frac{n-1}{2} \rceil}\right)\cdots\left(1-\frac{1}{\lceil \frac{n-(2i-1)}{2} \rceil}\right) \\
&= \frac{1}{m-i}\left(\frac{m-1}{m}\right)\left(\frac{m-2}{m-1}\right)\left(\frac{m-2}{m-1}\right)\cdots \\
&\quad \left(\frac{m-i}{m-(i-1)}\right)\left(\frac{m-i}{m-(i-1)}\right)\left(\frac{m-(i+1)}{m-i}\right) \\
&= \frac{(m-1)(m-i)^2[m-(i+1)]}{(m-i)(m)(m-1)^2(m-i)} \\
&= \frac{(m-i-1)}{m(m-1)} \\
&= \frac{1}{m}\left(\frac{m-1-i}{m-1}\right).
\end{aligned}$$

When T is odd,

$$\begin{aligned}
P(T=2i+1) &= P(\text{selected in } (2i+2)\text{th round, not in first } (2i+1) \text{ rounds}) \\
&= \frac{1}{\lceil \frac{n-(2i+1)}{2} \rceil}\left(1-\frac{1}{\lceil \frac{n}{2} \rceil}\right)\cdots\left(1-\frac{1}{\lceil \frac{n-2i}{2} \rceil}\right) \\
&= \frac{1}{m-i-1}\left(\frac{m-1}{m}\right)\left(\frac{m-2}{m-1}\right)\left(\frac{m-2}{m-1}\right)\cdots \\
&\quad \left(\frac{m-(i+1)}{m-i}\right)\left(\frac{m-(i+1)}{m-i}\right) \\
&= \frac{(m-1)[m-(i+1)]^2}{(m-i-1)(m)(m-1)^2} \\
&= \frac{m-i-1}{m(m-1)} \\
&= \frac{1}{m}\left(\frac{m-1-i}{m-1}\right).
\end{aligned}$$

Using the probability mass function, we can compute $E[T]$, the expected survival time of Person 1:

$$
\begin{aligned}
E[T] &= \sum_{t=0}^{n} t \cdot P(T=t) \\
&= \sum_{i=0}^{m-1} 2i \cdot P(T=2i) + \sum_{i=0}^{m-1} (2i+1) \cdot P(T=2i+1) \\
&= \sum_{i=0}^{m-1} 2i \cdot \left(\frac{1}{m} \cdot \frac{m-1-i}{m-1}\right) + \sum_{i=0}^{m-1} (2i+1) \cdot \left(\frac{1}{m} \cdot \frac{m-1-i}{m-1}\right) \\
&= \frac{1}{m} \sum_{i=0}^{m-1} (2i+2i+1) \frac{m-1-i}{m-1} \\
&= \frac{1}{m} \sum_{i=0}^{m-1} (4i+1) \left(1 - \frac{i}{m-1}\right) \\
&= \frac{1}{m} \sum_{i=0}^{m-1} \left(4i + 1 - \frac{4i^2}{m-1} - \frac{i}{m-1}\right) \\
&= \frac{1}{m} \left[4 \sum_{i=0}^{m-1} i + m - \frac{4}{m-1} \sum_{i=0}^{m-1} i^2 - \frac{1}{m-1} \sum_{i=0}^{m-1} i\right] \\
&= \frac{1}{m} \left[2m(m-1) + m - \frac{4}{m-1} \frac{(m-1)(m)(2m-1)}{6} - \frac{m}{2}\right] \\
&= 2(m-1) + 1 - \frac{2}{3}(2m-1) - \frac{1}{2} \\
&= \frac{2}{3}m - \frac{5}{6}.
\end{aligned}
$$

We then replace m by $\frac{n+1}{2}$, because $n = 2m - 1$ by assumption, to obtain

$$
E_n(1) = E[T] = \frac{2}{3}\left(\frac{n+1}{2}\right) - \frac{5}{6} = \boxed{\frac{1}{3}n - \frac{1}{2}}, \text{ if } n \text{ is odd.}
$$

We do not need to re-do all the calculations above if we want to compute the expected survival time when n is even. Instead, we use the recursive equation we constructed in section 2 as well as the formula of the odd case to solve for the expected survival time:

$$
\begin{aligned}
E_n(1) &= P(\text{remain at position } 1) \cdot (E_{n-1}(1) + 1) \\
&= \left(1 - \frac{1}{n/2}\right)\left(\frac{1}{3}(n-1) - \frac{1}{2} + 1\right) \\
&= \boxed{\frac{1}{3}n - \frac{1}{2} - \frac{1}{3n}}, \text{ if } n \text{ is even.}
\end{aligned}
$$

When n is sufficiently large, the constant $\frac{1}{2}$ and the term $\frac{1}{3n}$ are negligible. Therefore, the expected survival time is approximately $\frac{1}{3}n$ if n is sufficiently large.

4 The Expected Survival Time of Person 2

Lower and Upper Bounds

In section 3, we have computed the expected survival time of Person 1: $E_n(1) = \frac{1}{3}n - \frac{1}{2}$ for odd n, and $E_n(1) = \frac{1}{3}n - \frac{1}{2} - \frac{1}{3n}$ for even n. Then, we have the inequality that $\frac{1}{3}n - 1 \leq E_n(1) \leq \frac{1}{3}n$ for all n. We will use this inequality, which gives an upper bound and a lower bound for $E_n(1)$, to calculate the bounds for the expected survival time of Person 2.

We first calculate the upper bound of $E_n(2)$. We break into cases based on the round in which Person 2 gets shifted. The probability that Person 2 gets shifted in round k is the same as the probability that Person 1 gets eliminated in round k. This is also equal to the probability that Person 1 survives for time $t = k - 1$.

Because Person 2 is standing at an even position, s/he will not be eliminated in the game until shifting to the first position. If Person 2 gets shifted to the first position in round k, then s/he will become the new "Person 1" in the reduced game of $n - k$ people. Thus, we have:

$$\begin{aligned}
E_n(2) &= \sum_{k=1}^{n-1} P(\text{Person 2 shifted in round } k) \cdot (E_{n-k}(1) + k) \\
&= \sum_{k=1}^{n-1} P(\text{Person 1 eliminated in round } k) \cdot (E_{n-k}(1) + k) \\
&= \sum_{t=0}^{n-2} P(T = t) \cdot (E_{n-(t+1)}(1) + (t+1)) \\
&= \sum_{t=0}^{n-2} P(T = t) \cdot E_{n-t-1}(1) + \sum_{t=0}^{n-2} P(T = t) \cdot t + \sum_{t=0}^{n-2} P(T = t).
\end{aligned}$$

The expression for $E_n(2)$ includes three terms. The second term is the expected survival time of Person 1. If we apply the upper bound, then the term $\sum_{t=0}^{n-2} P(T = t) \cdot t = E_n(1) \leq \frac{1}{3}n$. The third term is the summation of the probabilities for all of the possible T values in the support, so the term equals 1. Finally, in order to compute the first term, we replace $E_{n-t-1}(1)$ by the upper bound, $\frac{1}{3}(n - t - 1)$.

In the calculations below, we have $-\frac{1}{3}\sum_{t=0}^{n-2} P(T = t) \cdot t = -\frac{1}{3}E_n(1)$, and because of the minus sign, we will need to apply a lower bound instead of the upper bound for $E_n(1)$, Therefore, we have

$$\begin{aligned}
&\sum_{t=0}^{n-2} P(T=t)\cdot E_{n-t-1}(1)\\
&\le \sum_{t=0}^{n-2} P(T=t)\cdot \frac{1}{3}(n-t-1)\\
&= \frac{1}{3}n\sum_{t=0}^{n-2} P(T=t) - \frac{1}{3}\sum_{t=0}^{n-2} P(T=t)\cdot t - \frac{1}{3}\sum_{t=0}^{n-2} P(T=t)\\
&\le \frac{1}{3}n - \frac{1}{3}\left(\frac{1}{3}n-1\right) - \frac{1}{3}\\
&= \frac{2}{9}n.
\end{aligned}$$

Adding all three terms, we get:

$$\begin{aligned}
E_n(2) &= \sum_{t=0}^{n-2} P(T=t)\cdot E_{n-t-1}(1) + \sum_{t=0}^{n-2} P(T=t)\cdot t + \sum_{t=0}^{n-2} P(T=t)\\
&\le \frac{2}{9}n + \frac{1}{3}n + 1\\
&= \frac{5}{9}n + 1.
\end{aligned}$$

Thus, we obtain $(\frac{5}{9}n+1)$ as an upper bound for $E_n(2)$.

In a similar fashion, we may calculate a lower bound of $E_n(2)$.

$$\begin{aligned}
E_n(2) &= \sum_{k=1}^{n-1} P(\text{player 1 eliminated in round k})\cdot (E_{n-k}(1)+k)\\
&= \sum_{t=0}^{n-2} P(T=t)\cdot (E_{n-(t+1)}(1)+(t+1))\\
&= \sum_{t=0}^{n-2} P(T=t)\cdot E_{n-t-1}(1) + \sum_{t=0}^{n-2} P(T=t)\cdot t + \sum_{t=0}^{n-2} P(T=t)\\
&\ge \sum_{t=0}^{n-2} P(T=t)\cdot \left(\frac{1}{3}(n-t-1)-1\right) + \left(\frac{1}{3}n-1\right) + 1\\
&= \frac{1}{3}n\sum_{t=0}^{n-2} P(T=t) - \frac{1}{3}\sum_{t=0}^{n-2} P(T=t)\cdot t - \frac{1}{3}\sum_{t=0}^{n-2} P(T=t) - 1 + \frac{1}{3}n\\
&\ge \frac{1}{3}n - \frac{1}{3}\left(\frac{1}{3}n\right) - \frac{1}{3} - 1 + \frac{1}{3}n\\
&= \frac{5}{9}n - \frac{4}{3}.
\end{aligned}$$

Thus, we obtain $(\frac{5}{9}n-\frac{4}{3})$ as a lower bound for $E_n(2)$.

Asymptotic Approach

If n is sufficiently large, then we can ignore the lower-order terms in $E_n(1)$. In this case, Person 1 is expected to survive for time $E[T] \approx \frac{1}{3}n$.

Let the survival time of Person 2 be denoted by T_2. Still, let T denote the survival time of Person 1. We apply the law of total expectation, conditioning on the time when Person 2 gets shifted (denoted by T_{shift}) to compute $E[T_2]$. Person 2 getting shifted is equivalent to Person 1 getting selected, so we have:

$$T_{shift} = T,$$

$$E[T_{shift}] = E[T] \approx \frac{1}{3}n.$$

If T_{shift} is known, then after Person 2 is shifted and becomes "Person 1," the game will be reduced to a smaller game with $(n - T_{shift})$ people. In the reduced game, Person 2, who moves to the first position, will be expected to survive for time $\frac{1}{3}(n - T_{shift})$ approximately. Therefore, we have:

$$\begin{aligned}
E[T_2] &= E[E(T_2|T_{shift})] \\
&\approx E[T_{shift} + \frac{1}{3}(n - T_{shift})] \\
&= E[T_{shift}] + \frac{1}{3}n - \frac{1}{3}E[T_{shift}] \\
&= \frac{2}{3}E[T_{shift}] + \frac{1}{3}n \\
&\approx \frac{2}{3}\left(\frac{1}{3}n\right) + \frac{1}{3}n \\
&= \frac{5}{9}n.
\end{aligned}$$

The expected survival time of Person 2 is $E_n(2) \approx \frac{5}{9}n$ from an asymptotic perspective.

5 The Expected Survival Time of Persons 3 and 4

In this section, we will discuss the method to compute the asymptotic expected survival time of Persons 3 and 4. In principle, the method can be applied to compute the asymptotic expected survival time of an arbitrary person as well.

Person 3

We consider the third person. Instead of conditioning on the shift time as we did for Person 2, we condition on the time until Person 3 is shifted or selected. We denote this time by T_{ss}. Also, we define I to be the indicator such that:

$$I = \begin{cases} 1, & \text{if Person 3 is shifted first,} \\ 0, & \text{if Person 3 is selected first.} \end{cases}$$

We denote the survival time of Person 3 by T_3. If T_{ss} is known, the survival time before Person 3 gets shifted or selected is of course T_{ss}. If Person 3 is shifted, then s/he will become "Person 2" in the new game of $(n - T_{ss})$ people and is expected to survive for time $\frac{5}{9}(n - T_{ss})$ after the shift. If Person 3 is selected, then s/he will leave the queue and the survival time after the selection is 0. Thus, using the indicator notation, Person 3 can survive for time $I \cdot \frac{5}{9}(n - T_{ss})$ on average after the shift or selection. Then the expected survival time of Person 3 is:

$$\begin{aligned} E_n(3) &= E[T_3] \\ &= E[E(T_3|T_{ss})] \\ &\approx E[T_{ss} + I \cdot \frac{5}{9}(n - T_{ss})]. \end{aligned}$$

However, I is independent of T_{ss}. Knowing the time it takes until Person 3 is shifted or selected does not provide us with any information about whether the person is indeed shifted first or selected first. Besides, T_{ss} is just the time until Person 1 or 3 gets selected, because Person 3 will shift if and only if Person 1 gets selected.

In each round, the probability that Person 1 is selected is the same as the probability that Person 3 is selected, both being the reciprocal of the number of odd-indexed people in that round. Therefore, given that Person 3 is either shifted or selected, with probability $\frac{1}{2}$ s/he is shifted first (and Person 1 being selected first), and with probability $\frac{1}{2}$ s/he is selected first. Then, the expected value of I is

$$E[I] = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0 = \frac{1}{2}.$$

Because of the independence,

$$\begin{aligned} E_n(3) &= E[T_{ss} + I \cdot \frac{5}{9}(n - T_{ss})] \\ &= E[T_{ss}] + \frac{5}{9}E[I]E[n - T_{ss}] \\ &= E[T_{ss}] + \frac{5}{18}(n - E[T_{ss}]) \\ &= \frac{5}{18}n + \frac{13}{18}E[T_{ss}]. \end{aligned}$$

We now are interested in $E[T_{ss}]$, the expected time until Person 1 or 3 gets selected. To start with, assume n is odd, and $n = 2m-1$, so that m is the number of odd-indexed people initially. We calculate the probability mass function of T_{ss} using the same approach as in section 3. In order to recognize the pattern, we perform more calculations, but for the sake of brevity, only the first few example calculations are included:

$$\begin{aligned} P(T_{ss} = 0) &= P(\text{either Person 1 or 3 selected in the 1st round}) \\ &= \frac{2}{\lceil \frac{n}{2} \rceil} \\ &= \frac{2}{m}. \end{aligned}$$

$$P(T_{ss}=1) = P(\text{Person 1 or 3 eliminated in the 2nd round, both survive the 1st round})$$
$$= \frac{2}{\lceil \frac{n-1}{2} \rceil}\left(1 - \frac{2}{\lceil \frac{n}{2} \rceil}\right)$$
$$= \frac{2}{m-1}\left(\frac{m-2}{m}\right)$$
$$= \frac{2(m-2)}{(m-1)m}.$$

$$P(T_{ss}=2) = P(\text{Person 1 or 3 eliminated in the 3rd round, both survive the first 2 rounds})$$
$$= \frac{2}{\lceil \frac{n-2}{2} \rceil}\left(1 - \frac{2}{\lceil \frac{n-1}{2} \rceil}\right)\left(1 - \frac{2}{\lceil \frac{n}{2} \rceil}\right)$$
$$= \frac{2}{m-1}\left(\frac{m-3}{m-1}\right)\left(\frac{m-2}{m}\right)$$
$$= \frac{2(m-3)(m-2)}{(m-1)(m-1)m}.$$

$$P(T_{ss}=3) = P(\text{Person 1 or 3 eliminated in the 4th round, both survive the first 3 rounds})$$
$$= \frac{2}{\lceil \frac{n-3}{2} \rceil}\left(1 - \frac{2}{\lceil \frac{n-2}{2} \rceil}\right)\left(1 - \frac{2}{\lceil \frac{n-1}{2} \rceil}\right)\left(1 - \frac{2}{\lceil \frac{n}{2} \rceil}\right)$$
$$= \frac{2}{m-2}\left(\frac{m-3}{m-1}\right)\left(\frac{m-3}{m-1}\right)\left(\frac{m-2}{m}\right)$$
$$= \frac{2(m-3)(m-3)}{(m-1)(m-1)m}.$$

$$P(T_{ss}=4) = P(\text{Person 1 or 3 eliminated in the 5th round, both survive the first 4 rounds})$$
$$= \frac{2}{\lceil \frac{n-4}{2} \rceil}\left(1 - \frac{2}{\lceil \frac{n-3}{2} \rceil}\right)\left(1 - \frac{2}{\lceil \frac{n-2}{2} \rceil}\right)\left(1 - \frac{2}{\lceil \frac{n-1}{2} \rceil}\right)\left(1 - \frac{2}{\lceil \frac{n}{2} \rceil}\right)$$
$$= \frac{2}{m-2}\left(\frac{m-4}{m-2}\right)\left(\frac{m-3}{m-1}\right)\left(\frac{m-3}{m-1}\right)\left(\frac{m-2}{m}\right)$$
$$= \frac{2(m-4)(m-3)(m-3)}{(m-2)(m-1)(m-1)m}.$$

After simplification, the generalized formulae of the probability mass functions of even T ($T = 2i$) and odd T ($T = 2i + 1$) are,

$$
\begin{aligned}
P(T_{ss} = 2i) &= P(\text{Person 1 or 3 eliminated in the (2i+1)th round,} \\
&\qquad \text{both survive the first 2i rounds)} \\
&= \frac{2(m-(i+2))(m-(i+1))(m-(i+1))}{(m-2)(m-1)(m-1)m} \\
&= \frac{2(m-i-2)(m-i-1)(m-i-1)}{(m-2)(m-1)(m-1)m} \\
&= \frac{2(m-i-2)(m-i-1)^2}{(m-2)(m-1)(m-1)m}.
\end{aligned}
$$

$$
\begin{aligned}
P(T_{ss} = 2i+1) &= P(\text{Person 1 or 3 eliminated in the (2i+2)th round,} \\
&\qquad \text{both survive the first (2i+1) rounds)} \\
&= \frac{2(m-(i+2))(m-(i+2))(m-(i+1))}{(m-2)(m-1)(m-1)m} \\
&= \frac{2(m-i-2)(m-i-2)(m-i-1)}{(m-2)(m-1)(m-1)m} \\
&= \frac{2(m-i-2)^2(m-i-1)}{(m-2)(m-1)(m-1)m}.
\end{aligned}
$$

Using the probability mass function above, we can compute the expected time until Person 1 or 3 gets selected:

$$
\begin{aligned}
E[T_{ss}] &= \sum_{t=0}^{n} t \cdot P(T_{ss} = t) \\
&= \sum_{i=0}^{m-1} 2i \cdot P(T_{ss} = 2i) + \sum_{i=0}^{m-1} (2i+1) \cdot P(T_{ss} = 2i+1) \\
&= \sum_{i=0}^{m-1} 2i \cdot \left(\frac{2(m-i-2)(m-i-1)^2}{(m-2)(m-1)(m-1)m} \right) \\
&\quad + \sum_{i=0}^{m-1} (2i+1) \cdot \left(\frac{2(m-i-2)^2(m-i-1)}{(m-2)(m-1)(m-1)m} \right) \\
&= \frac{\sum_{i=0}^{m-1} 4i \cdot (m-i-2)(m-i-1)^2}{(m-2)(m-1)^2 m} + \frac{\sum_{i=0}^{m-1} 4i \cdot (m-i-2)^2(m-i-1)}{(m-2)(m-1)^2 m} \\
&= \frac{1}{(m-2)(m-1)^2 m} \cdot \left(\frac{1}{5}m^5 - \frac{4}{3}m^4 + 3m^3 - \frac{8}{3}m^2 + \frac{4}{5}m \right) \\
&\quad + \frac{1}{(m-2)(m-1)^2 m} \cdot \left(\frac{1}{5}m^5 - \frac{7}{6}m^4 + \frac{8}{3}m^3 - \frac{17}{6}m^2 + \frac{17}{15}m \right) \\
&= \frac{1}{(m-2)(m-1)^2 m} \left(\frac{2}{5}m^5 - \frac{5}{2}m^4 + \frac{17}{3}m^3 - \frac{11}{2}m^2 + \frac{29}{15}m \right).
\end{aligned}
$$

If we ignore the lower order terms, and do not distinguish between $m-1$ and m, then the expected time until Person 1 or 3 gets selected will be asymptotically

$\frac{1}{m^4} \cdot \frac{2}{5} m^5 = \frac{2}{5} m$. Since $m = \frac{n+1}{2} \approx \frac{n}{2}$, the expected time $E[T_{ss}] \approx \frac{1}{5} n$ when n is sufficiently large. Using this result, we can compute the expected survival time of Person 3:

$$\begin{aligned} E_n(3) &= \frac{5}{18} n + \frac{13}{18} E[T_{ss}] \\ &\approx \frac{5}{18} n + \frac{13}{18} \left(\frac{1}{5} n \right) \\ &= \frac{19}{45} n. \end{aligned}$$

Person 4

The case of Person 4 is similar to the case of Person 2. We condition on the shift time T_{shift} of Person 4 again because s/he will not be selected until being shifted to the third position. However, T_{shift} of Person 4 is the time until Person 1 or 3 gets eliminated. Thus,

$$T_{shift} = T_{ss},$$

$$E[T_{shift}] = E[T_{ss}] \approx \frac{1}{5} n.$$

Let T_4 denote the survival time of Person 4, we have

$$\begin{aligned} E_n(4) &= E[T_4] \\ &= E[E(T_4 | T_{shift})] \\ &\approx E[T_{shift} + \frac{19}{45}(n - T_{shift})] \\ &= \frac{19}{45} n + \frac{26}{45} E[T_{shift}] \\ &\approx \frac{19}{45} n + \frac{26}{45} \left(\frac{1}{5} n \right) \\ &= \frac{121}{225} n. \end{aligned}$$

Other Persons

The calculation methods explained above can be adapted to compute the expected survival time of Person 5, 6, 7, ...

For odd-indexed persons, as we have seen in the case of Person 3, the asymptotic expectation of survival time only depends on the expected value of the indicator I and the expected time until the person gets shifted or selected. The expected value of the indicator is the probability that the indicator takes the value of 1, which is the same as the probability that the person is shifted (i.e. any person in front is selected) before selected. For Person k, where k is odd, there are $\frac{k-1}{2}$ odd-indexed people in front of Person k. Therefore, the probability that Person k is shifted before selected is $\frac{\frac{k-1}{2}}{\frac{k-1}{2}+1} = \frac{k-1}{k+1}$. Thus, $E[I] = \frac{k-1}{k+1}$. The expected time until Person k gets shifted or selected may be calculated by computing the probability mass function. In the previous part of this section and section 3, we have used the probability mass function to obtain the result that when $k = 1$, the expected time is approximately $\frac{1}{3}n$, and that when $k = 3$, the expected time is approximately $\frac{1}{5}n$.

For even-index persons, such as Person 2 or 4, we only need to calculate the expected time until the person gets shifted to obtain the expected survival time. The expected time until Person k shifts, where k is even, is the same as the the expected time until Person $(k-1)$ gets shifted or selected. As mentioned above, the expected time until Person $(k-1)$ gets shifted or selected, where $k-1$ is odd, may be calculated using the probability mass function.

6 The Expected Survival Time of Person n

From Table 1 in section 2, we observe that the expected survival time of the last person (Person n) is: 0, 1, 1, 2, 2, 3, 3, etc. Based on this observation, we claim that $E_n(n) = \lfloor \frac{n}{2} \rfloor$. A proof using mathematical induction follows.

Lemma: *The expected survival time of Person n is:*

$$E_n(n) = \lfloor \frac{n}{2} \rfloor.$$

Proof: We prove the formula above by induction.

Base Cases: By Table 1 in section 2, we have $E_1(1) = \lfloor \frac{1}{2} \rfloor = 0$, $E_2(2) = \lfloor \frac{2}{2} \rfloor = 1$, and $E_3(3) = \lfloor \frac{3}{2} \rfloor = 1$. Thus the formula is correct for the base cases.

Induction Hypothesis: We assume that the formula is correct for $1, 2, 3, ..., n$.

Inductive Step: Now we show that the formula is correct for $n+1$. There are two cases: n is odd, $n+1$ is even; or n is even, $n+1$ is odd. From section 2, we know the recursive equation below holds:

$$E_{n+1}(k) = P(\text{shifted to position } k-1) \cdot (E_n(k-1)+1).$$

We will use the above equation when $k = n+1$. That is,

$$E_{n+1}(n+1) = P(\text{shifted to position } n) \cdot (E_n(n)+1).$$

Case (1): If n is odd, $n+1$ is even, then Person $(n+1)$ will shift to position n with probability one because s/he is even-indexed in the first round and thus safe. By hypothesis, $E_n(n) = \lfloor \frac{n}{2} \rfloor$, so

$$\begin{aligned} E_{n+1}(n+1) &= P(\text{shifted to position } n) \cdot (E_n(n)+1) \\ &= E_n(n)+1 \\ &= \lfloor \frac{n}{2} \rfloor + 1 \\ &= \frac{n-1}{2} + 1 \\ &= \lfloor \frac{n+1}{2} \rfloor. \end{aligned}$$

Case (2): If n is even, $n+1$ is odd, so Person $(n+1)$ will be selected in the first round with probability $\frac{1}{\lceil\frac{n+1}{2}\rceil}$, because s/he is one of the $\lceil\frac{n+1}{2}\rceil$ odd-indexed people in the line. Thus, s/he will shift to position n in the next round with probability $\left(1-\frac{1}{\lceil\frac{n+1}{2}\rceil}\right)$. We have

$$\begin{aligned}
E_{n+1}(n+1) &= P(\text{shifted to position } n)\cdot(E_n(n)+1)\\
&= \left(1-\frac{1}{\lceil\frac{n+1}{2}\rceil}\right)\left(\lfloor\frac{n}{2}\rfloor+1\right) \qquad \text{by hypothesis}\\
&= \left(1-\frac{2}{n+2}\right)\frac{n+2}{2}\\
&= \frac{n}{2}\\
&= \lfloor\frac{n+1}{2}\rfloor.
\end{aligned}$$

In either case, the claim holds true for $n+1$.

Therefore, by mathematical induction, the formula $E_n(n) = \lfloor\frac{n}{2}\rfloor$ is correct for any positive integer n. □

7 Conclusion

In summary, this paper introduces a recursive approach to solve for exact values of the expected survival time, proves the exact formula of the expected survival time of Person 1 and Person n, and shows that the asymptotic expected survival time of Person 1, 2, 3 and 4 are $\frac{1}{3}n$, $\frac{5}{9}n$, $\frac{19}{45}n$ and $\frac{121}{225}n$ respectively. The method in section 5 may be adapted to calculate the expected survival time of the other persons as well.

8 Future Research

As mentioned in section 5, the expected survival time can be solved asymptotically if we can compute the expected time until Person k gets shifted or selected for all k. This quantity is just the expected time until Person $1, 3, ..., k$ gets selected if k is odd, and is the expected time until Person $1, 3, ..., (k-1)$ gets selected if k is even. In previous sections, we used the probability mass function to obtain the expected time until Person 1 is selected ($\approx \frac{1}{3}n$) and the expected time until Person 1 or 3 is selected ($\approx \frac{1}{5}n$). However, the computational difficulty greatly increases if we consider the expected time until more people get shifted. By observing the exact values of expectation obtained by recursion, we conjecture that the expected time until selection is $\frac{1}{7}n$ for Person 1, 3 and 5, and $\frac{1}{9}n$ for Person 1, 3, 5 and 7, etc. In the future, we will try to prove this conjecture without the cumbersome calculations.

Acknowledgement: The authors thank the Acheson J. Duncan Fund for the Advancement of Research in Statistics for support for undergraduate research.

References

[1] Robinson, W. J. *"the Josephus Problem"* [*The Mathematical Gazette*]. 44 (347): 47–52. doi:10.2307/3608532. JSTOR 3608532, 1960.

[2] ChengYiYi: Original Question on ZhiHu "Probabilistic Count-Out Problem",
`https://www.zhihu.com/question/55445739/`

[3] XieZhuoFan: Answer to ZhiHu "Probabilistic Count-Out Problem",
`https://www.zhihu.com/question/55445739/answer/144619265`

[4] WangBoChen: "If there are N people in a line, and every hour randomly kill a person with odd index, and the survivors re-indexed, who will be most likely to survive?",
`http://qr.ae/TUTFeX`