Survival Probability of a Probabilistic Counting-Out Game on a Line

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

Abstract

Our research focuses on the winning probability of 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 i, where i is odd, to leave the line, and shift each person at a position j such that $j > i$ to position $j - 1$. 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 greatest chance to survive, that is, the highest probability to be the last one remaining. Specifically, we have derived recursions to solve for exact values and the formula of the winning probabilities. We have also considered variations of the problem, where people are grouped into triples, quadruples, etc., and the first person in each group is at the risk of being selected. We will also present various sequences we have discovered while solving for the winning probabilities of the different variations, as well as other possible extensions and related findings concerning this problem.

1 Introduction

When the original version of this probabilistic counting-out game, where the total number of people was limited to 600, appeared on the Chinese question-and-answer website Zhihu[2], it received vast attention from the Zhihu community. Within less than a year, it gained over more than 130 answers, 3000 followers and 600,000 reviews.

The original post is rather simple: it asks when 600 people stand in a line, which position is the "safest". What grasps people's attention is that the definition of "safest" is unclear. Moreover, after redefining "safest" in two different ways: the initial position with highest probability to survive or the initial position with the longest expected time to stay in the line, people realize that according to simulations, the two ways produce different answers. While the initial position that has the highest probability to survive seems to be the last position, the initial position that has the longest expected survival time appears to be the second position.

CONGRESSUS NUMERANTIUM 234 (2019), pp.213-224

At first sight, both answers make intuitive sense. Standing at the second initial position, the person stays safe for a long period of time until the person with the first initial position gets selected (we also recognize that the first position is the most dangerous position). On the other hand, the person at the last position is constantly switching from unsafe positions to safe positions, so overall this person has more chances than any other person to be switched to safe positions.

Of course, to study this mismatch between longest expected survival time and highest survival probability, we need more than just intuition. In this paper, we focus particularly on one of the answers described above – the survival probabilities. We formally define the problem as the following: there are n people in a line at positions $1, 2, \ldots, n$. For each round, we randomly select a person at position i, where i is odd, to leave the line, and shift the people at positions j such that $j > i$ to positions $j-1$. We continue to select people until there is only one person left, who then becomes the winner. The survival probability $p_n(i)$ is defined as the probability of the person with initial position i to be the only person left after $n-1$ turns.

By applying a first-step analysis and recursion, we are able to solve for exact values for survival probabilities and prove a formula. We have also designed variations of this problem, and found related sequences and formulas while solving these variations. In the following sections, we will present the above results as follows:

In Section 2, we will introduce a first-step analysis method and its recursion equation, which help us solve for the exact values for survival probabilities.

In Section 3, we provide the formula for the survival probabilities and the proof for this formula.

In Section 4, we present variations of this problem and the exact survival probabilities values for these variations. Some speculations and discoveries on the relationship between the original problem and the variations are provided.

Last but not least, section 5 briefly mentions some open questions related to the problem and some of our future research directions.

2 First Step Analysis

To introduce the idea of first step analysis, let us first start with some base case examples. Let $p_n(i)$ denote the probability that the person with initial position i wins the game with n people in the line initially ($1 \leq i \leq n$). For convenience, we also denote A_i to be the person with initial position i, and s_i be the his/her position. We first consider the base cases when n is small:

If $n = 1$, person A_1 is the only person in the line, so s/he automatically becomes the winner. Thus we have $\boxed{p_1(1) = 1}$.

If $n = 2$, person A_1 will be eliminated in the first round because s/he is the only odd-indexed person. Person A_2, who is the only one remaining in the next round, will be the winner. Thus, $\boxed{p_2(1) = 0}$ and $\boxed{p_2(2) = 1}$.

If $n = 3$, let us consider person A_1 first. Since 1 is the smallest index, person A_1's position remains the same until s/he is selected and eliminated. Thus, if a person in the line has initial position of s_1, s/he will eventually lose the game. Thus, $p_n(1) = 0 \implies \boxed{p_3(1) = 0}$.

Then let us consider person A_2. In the first round, the person is even-positioned so s/he will not get eliminated. Consider the result of the first round: $P(\text{person } A_1 \text{ selected}) = P(\text{person } A_3 \text{ selected}) = \frac{1}{2}$, i.e. the person A_2 has equal probabilities to shift to s_1 (if A_1 is eliminated at first) or remain at s_2 (if A_3 is eliminated at first). If person A_2 is shifted to s_1, then s/he has probability $p_2(1) = 0$ to win. If person A_2 keeps the position s_2, then s/he has probability $p_2(2) = \frac{1}{2}$ to win. Thus, we use a first-step decomposition and condition on the result of the first round:

$$\begin{aligned}
P(\text{person } A_2 \text{ wins}) &= P(\text{shifted to } s_1) \cdot P(\text{win} \mid \text{shifted to } s_1) \\
&\quad + P(\text{remain at } s_2) \cdot P(\text{win} \mid \text{remain at } s_2) \\
&= \frac{1}{2} \cdot p_2(1) + \frac{1}{2} \cdot p_2(2) \\
&= \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1 \\
&= \frac{1}{2}.
\end{aligned}$$

Hence $\boxed{p_3(2) = \frac{1}{2}}$.

Finally, we consider person A_3. Although we can simply compute $1 - p_3(1) - p_3(2)$, we will still do a first-step decomposition here, since we want to generalize the formula for the winning probability using the decomposition later. Person A_3 is either eliminated in the first round or shifted to s_2. If the person is shifted to s_2, s/he has probability $p_2(2)$ to win. Thus we have:

$$\begin{aligned}
P(\text{person } A_3 \text{ wins}) &= P(\text{shifted to } s_2) \cdot P(\text{win} \mid \text{shifted to } s_2) \\
&= \frac{1}{2} \cdot p_2(2) \\
&= \frac{1}{2}.
\end{aligned}$$

Therefore, $\boxed{p_3(3) = \frac{1}{2}}$.

With the base cases shown above, we see that each calculation depends only on its very first step. After the first step is taken, the cases break into situations we have already calculated. We call this bottom-up idea the first-step analysis, and use it to generalize the recursion equation to calculate the exact value for any winning probability $p_n(i)$:

$$\begin{aligned}
p_n(i) &= P(\text{shifted to } s_{(i-1)}) \cdot P(\text{win} \mid \text{shifted to } s_{(i-1)}) \\
&\quad + P(\text{remain at } s_i) \cdot P(\text{win} \mid \text{remain at } s_i) \\
&= P(\text{shifted to } s_{(i-1)}) \cdot p_{n-1}(i-1) + P(\text{remain at } s_i) \cdot p_{n-1}(i)
\end{aligned}$$

This process is a non-homogenous Markov chain. The probability to shift (or not to shift) only depends on the current state, the position of the person and the total number of people at present.

3 Survival Probability Formula

From the first step analysis above, since we have the winning probabilities for $n = 1, 2, 3$, we can compute the winning probabilities for $n \geq 4$.

n	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$	$p(6)$	$p(7)$	$p(8)$	$p(9)$
1	1								
2	0	1							
3	0	$\frac{1}{2}$	$\frac{1}{2}$						
4	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{2}{4}$					
5	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{2}{6}$				
6	0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{3}{9}$			
7	0	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{3}{12}$	$\frac{3}{12}$		
8	0	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	
9	0	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{2}{20}$	$\frac{3}{20}$	$\frac{3}{20}$	$\frac{4}{20}$	$\frac{4}{20}$

Table 1: The winning probabilities for each person for $n = 1, 2, 3, ..., 9$.

From Table 1, we see the winning probability increases approximately linearly as index i increases. Also, we can see a clear pattern in the winning probabilities within each level (row).

The following theorem provides an exact formula for $p_n(i)$ for all n and $i < n$:

Theorem: the winning probability for the person at position i among n people initially is:

$$p_n(i) = \begin{cases} 1, & \text{if } n = 1 \\ \frac{\lfloor \frac{i}{2} \rfloor}{\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil}, & \text{otherwise} \end{cases}$$

Proof: We use induction to prove the result. Note that we have shown that the person A_1 always loses the game unless s/he is the only person.

Base cases:

The $n = 1$ case is trivial.

If $n = 2$, then person A_1 gets picked in the first round and A_2 wins with probability one.

When $i = 1$, we have $p_2(1) = 0$ and $\frac{\lfloor \frac{1}{2} \rfloor}{\lfloor \frac{2}{2} \rfloor \lceil \frac{2}{2} \rceil} = \frac{0}{1 \cdot 1} = 0$.

When $i = 2$, we have $p_2(2) = 1$ and $\frac{\lfloor \frac{2}{2} \rfloor}{\lfloor \frac{2}{2} \rfloor \lceil \frac{2}{2} \rceil} = \frac{1}{1 \cdot 1} = 1$. Both are in agreement with the formula

Let $n = 3$. If A_1 gets picked (with probability $\frac{1}{2}$) in the first round, then A_2

becomes the person at the first position and thereby cannot win, so A_3 automatically becomes winner. If A_3 gets picked (with probability $\frac{1}{2}$), then A_2 is the winner because the person A_1 will be eliminated. We verify agreement with the formula:
When $i = 1$, $p_3(1) = 0$ and $\frac{\lfloor \frac{1}{2} \rfloor}{\lfloor \frac{3}{2} \rfloor \lceil \frac{3}{2} \rceil} = \frac{0}{1 \cdot 1} = 0$.
When $i = 2$, $p_3(2) = \frac{1}{2}$ and $\frac{\lfloor \frac{2}{2} \rfloor}{\lfloor \frac{3}{2} \rfloor \lceil \frac{3}{2} \rceil} = \frac{1}{1 \cdot 2} = \frac{1}{2}$.
When $i = 3$, $p_3(3) = \frac{1}{2}$ and $\frac{\lfloor \frac{3}{2} \rfloor}{\lfloor \frac{3}{2} \rfloor \lceil \frac{3}{2} \rceil} = \frac{1}{1 \cdot 2} = \frac{1}{2}$.
Thus the base cases are correct.

Induction Hypothesis: We assume our formula is correct for all possible i, $i = 1, 2, 3, ..., n$ in level n (the n-th row).

Inductive Step: We prove the correctness of our formula by doing two inductive steps: (1) if our formula is correct for level n (even), then it is correct for $n + 1$ (odd), and (2) if our formula is correct for level n (odd), then it is correct for $n + 1$ (even).

Notation: Let S be the event that the person A_i gets shifted to position s_{i-1} in the next round when there are $n + 1$ people, R be the event that the person A_i remains at the same position s_i, and E be the event that the person A_i gets eliminated. We then have $P(S) + P(R) + P(E) = 1$, where $P(E) = 0$ if i is even.

Case (1): Let n be even. We want to show that the formula is correct in the case $n + 1$ for any $i = 1, 2, ..., n + 1$.

Sub-case (1.1): Let i be odd ($i = 1, 3, 5, ..., n - 1$) and n be even. By the first step decomposition,

$$p_{n+1}(i) = P(S)p_n(i - 1) + P(R)p_n(i).$$

In this case,

$$p_n(i - 1) = \frac{\lfloor \frac{i-1}{2} \rfloor}{\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil} = \frac{\frac{i-1}{2}}{\frac{n}{2}\frac{n}{2}}$$

and

$$p_n(i) = \frac{\lfloor \frac{i}{2} \rfloor}{\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil} = \frac{\frac{i-1}{2}}{\frac{n}{2}\frac{n}{2}}.$$

Since i is odd and n is even, when we increase to $n + 1$ people,

$$P(S) = \frac{(i - 1)/2}{(n + 2)/2}$$

and

$$P(R) = 1 - P(E) - P(S) = 1 - \frac{1}{(n + 2)/2} - \frac{(i - 1)/2}{(n + 2)/2}.$$

Thus,

$$\begin{aligned}
p_{n+1}(i) &= P(S)p_n(i-1) + P(R)p_n(i) \\
&= \frac{(i-1)/2}{(n+2)/2}\left(\frac{\frac{i-1}{2}}{\frac{n}{2}\frac{n}{2}}\right) + \left(1 - \frac{1}{(n+2)/2} - \frac{(i-1)/2}{(n+2)/2}\right)\left(\frac{\frac{i-1}{2}}{\frac{n}{2}\frac{n}{2}}\right) \\
&= \frac{\frac{i-1}{2}}{\frac{n+2}{2}\frac{n}{2}} \\
&= \frac{\lfloor\frac{i}{2}\rfloor}{\lceil\frac{n+1}{2}\rceil\lfloor\frac{n+1}{2}\rfloor}.
\end{aligned}$$

Thus the formula is correct.

Sub-case (1.2): Let i be even ($i = 2, 4, 6, ..., n$) and n be even. By the first step decomposition,

$$p_{n+1}(i) = P(S)p_n(i-1) + P(R)p_n(i)$$

In this case, $p_n(i-1) = \frac{\lfloor\frac{i-1}{2}\rfloor}{\lfloor\frac{n}{2}\rfloor\lceil\frac{n}{2}\rceil} = \frac{i/2-1}{\frac{n}{2}\frac{n}{2}}$, and $p_n(i) = \frac{\lfloor\frac{i}{2}\rfloor}{\lfloor\frac{n}{2}\rfloor\lceil\frac{n}{2}\rceil} = \frac{\frac{i}{2}}{\frac{n}{2}\frac{n}{2}}$
Since i is even and n is even, when we increase to $n+1$ people, $P(S) = \frac{i/2}{(n+2)/2}$ and $P(R) = 1 - P(S) = 1 - \frac{i/2}{(n+2)/2}$.
Thus,

$$\begin{aligned}
p_{n+1}(i) &= P(S)p_n(i-1) + P(R)p_n(i) \\
&= \left(\frac{i/2}{(n+2)/2}\right)\left(\frac{i/2-1}{\frac{n}{2}\frac{n}{2}}\right) + \left(1 - \frac{i/2}{(n+2)/2}\right)\left(\frac{\frac{i}{2}}{\frac{n}{2}\frac{n}{2}}\right) \\
&= \frac{\frac{i}{2}}{\frac{n+2}{2}\frac{n}{2}} \\
&= \frac{\lfloor\frac{i}{2}\rfloor}{\lceil\frac{n+1}{2}\rceil\lfloor\frac{n+1}{2}\rfloor}.
\end{aligned}$$

Thus the formula is correct.

Case (2): Let n be odd. We want to show that the formula is correct in the case $n+1$ for any $i = 1, 2, ..., n+1$.

Sub-case (2.1): Let i be odd and n be odd. By the first step decomposition,

$$p_{n+1}(i) = P(S)p_n(i-1) + P(R)p_n(i).$$

In this case, $p_n(i-1) = \frac{\lfloor\frac{i-1}{2}\rfloor}{\lfloor\frac{n}{2}\rfloor\lceil\frac{n}{2}\rceil} = \frac{\frac{i-1}{2}}{\frac{n-1}{2}\frac{n+1}{2}}$, and $p_n(i) = \frac{\lfloor\frac{i}{2}\rfloor}{\lfloor\frac{n}{2}\rfloor\lceil\frac{n}{2}\rceil} = \frac{\frac{i-1}{2}}{\frac{n-1}{2}\frac{n+1}{2}}$
Since i is odd and n is odd, when we increase to $n+1$ people, $P(S) = \frac{(i-1)/2}{(n+1)/2}$ and $P(R) = 1 - P(E) - P(S) = 1 - \frac{1}{(n+1)/2} - \frac{(i-1)/2}{(n+1)/2}$.

Thus,

$$\begin{aligned}
p_{n+1}(i) &= P(S)p_n(i-1) + P(R)p_n(i) \\
&= \frac{(i-1)/2}{(n+1)/2}\frac{\frac{i-1}{2}}{\frac{n-1}{2}\frac{n+1}{2}} \\
&= \frac{\frac{i-1}{2}}{\frac{n+1}{2}\frac{n+1}{2}} \\
&= \frac{\lfloor\frac{i}{2}\rfloor}{\lceil\frac{n+1}{2}\rceil\lfloor\frac{n+1}{2}\rfloor}.
\end{aligned}$$

Thus the formula is correct.

Sub-case (2.2): Let i be even and n be odd. By the first step decomposition,

$$p_{n+1}(i) = P(S)p_n(i-1) + P(R)p_n(i).$$

In this case, $p_n(i-1) = \frac{\lfloor\frac{i-1}{2}\rfloor}{\lfloor\frac{n}{2}\rfloor\lceil\frac{n}{2}\rceil} = \frac{i/2-1}{\frac{n-1}{2}\frac{n+1}{2}}$, and $p_n(i) = \frac{\lfloor\frac{i}{2}\rfloor}{\lfloor\frac{n}{2}\rfloor\lceil\frac{n}{2}\rceil} = \frac{i/2}{\frac{n-1}{2}\frac{n+1}{2}}$. Since i is even and n is odd, when we increase to $n+1$ people, $P(S) = \frac{i/2}{(n+1)/2}$ and $P(R) = 1 - P(S) = 1 - \frac{i/2}{(n+1)/2}$.
Thus,

$$\begin{aligned}
p_{n+1}(i) &= P(S)p_n(i-1) + P(R)p_n(i) \\
&= \frac{i/2}{(n+1)/2}\frac{i/2-1}{\frac{n-1}{2}\frac{n+1}{2}} + \left(1 - \frac{i/2}{(n+1)/2}\right)\frac{i/2}{\frac{n-1}{2}\frac{n+1}{2}} \\
&= \frac{\frac{i}{2}}{\frac{n+1}{2}\frac{n+1}{2}} \\
&= \frac{\lfloor\frac{i}{2}\rfloor}{\lceil\frac{n+1}{2}\rceil\lfloor\frac{n+1}{2}\rfloor}.
\end{aligned}$$

Thus the formula is correct.

Remark: A concern is that if $i = n+1$ (the last person in the next level), then we may be using $p_n(i) = p_n(n+1)$ which does not exist. However, if we consider the last person in a queue of $(n+1)$ people, then the probability that he remains at the same position $s_{(n+1)}$ is 0 because the person either gets eliminated or gets shifted. Thus the second term involving $p_n(n+1)$ would disappear: $P(R)\cdot p_n(n+1) = 0\cdot p_n(n+1) = 0$. For the first term, $P(S)\cdot p_n(n-1) = P(S)\cdot p_n(n)$, and the term $p_n(n)$ can still be found in the previous level. Hence, we do not have "index out of range" issue.

4 Winning Probability: mod k variations

In this section, we will present variations of this problem. We call these variations the mod k variations. First, let us reconsider the original problem. Instead of recognizing the selection process as randomly selecting a player at odd position in the line, we can think of it as grouping people in pairs, and the first one in each group is at the risk of being selected. We call it the mod 2 case (problem operating on a mod 2 basis).

After obtaining the specific formula for the winning probabilities of the mod 2 case, we want to examine those cases where we group players in triplets, quadruplets, etc. Similarly, in each round, the first person in the group, i.e. the players with position number $\equiv 1 \pmod{k}$, are at risk of being selected.

We re-use our notation from the mod 2 case to state the problem:

There are n people in a line at positions $1, 2, \ldots, n$. For each round, we randomly select a person at position i, where where $i \equiv 1 \pmod{k}$, to leave the line, and shift the people at position i such that $j > i$ to position $j - 1$. We continue to select people until there is only one person left, who then becomes the winner. The survival probability $p_n(i)$ is defined as the probability of the the person with initial position i to be the only person left after $n - 1$ turns.

Again, in these cases, we can re-use the recursion equation to compute the probabilities directly:

$$\begin{aligned} p_n(i) &= P(\text{shifted to } s_{(i-1)}) \cdot P(\text{win} \mid \text{shifted to } s_{(i-1)}) \\ &\quad + P(\text{remain at } s_i) \cdot P(\text{win} \mid \text{remain at } s_i) \\ &= P(\text{shifted to } s_{(i-1)}) \cdot p_{n-1}(i-1) + P(\text{remain at } s_i) \cdot p_{n-1}(i) \end{aligned}$$

From Section 3, we see that the winning probability grows linearly in pairs. Therefore, a natural conjecture would be that for a problem operating on mod i basis, the probability grows linearly in groups of i with perhaps a different slope. However, surprisingly, we will see that it is not the case. More interestingly, the winning probability grows in a polynomial fashion.

n	$p(1)$	$p(2)$	$p(3)$	$p(4)$	$p(5)$	$p(6)$	$p(7)$	$p(8)$	$p(9)$
3	0	0	1						
4	0	0	$\frac{1}{2}$	$\frac{1}{2}$					
5	0	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{2}{4}$				
6	0	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{4}{8}$			
7	0	0	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{4}{12}$	$\frac{4}{12}$		
8	0	0	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{2}{18}$	$\frac{4}{18}$	$\frac{4}{18}$	$\frac{6}{18}$	
9	0	0	$\frac{1}{27}$	$\frac{1}{27}$	$\frac{2}{27}$	$\frac{4}{27}$	$\frac{4}{27}$	$\frac{6}{27}$	$\frac{9}{27}$

Table 2: Some values of the winning probabilities for the mod 3 case

We see from Table 2 that the numerator does not follow a linear trend. After adjusting the values to share a common denominator, the numerators form such a sequence: 0, 0, 1, 1, 2, 4, 4, 6, 9, 9, 12, 16, 16, 20, 25, 25, 30, 36, 36, 42, 49, 49, ...

With the help of the On-line Encyclopedia of Integer Sequences (OEIS)[6], we have identified the sequence of numerators to be $A008133$ with formula $\lfloor \frac{i}{3} \rfloor \lfloor \frac{i+1}{3} \rfloor$ where i is the initial position index, while the denominator is very similar to $AE006501$. After making slight adjustments in the formula for $AE006501$, we obtain the formula for the denominator, which is $\lfloor \frac{n+1}{3} \rfloor \lfloor \frac{n+2}{3} \rfloor \lfloor \frac{n+3}{3} \rfloor$ for the mod 3 case, where n is the total number of people.

Therefore, we obtain the general formula for mod 3 case:

Formula: In the mod 3 case, the winning probability for the person at position i among n people initially is:

$$p_n(i) = \begin{cases} 1, & \text{if } n = 1 \\ \frac{\lfloor \frac{i}{3} \rfloor \lfloor \frac{i+1}{3} \rfloor}{\lfloor \frac{n+1}{3} \rfloor \lfloor \frac{n+2}{3} \rfloor \lfloor \frac{n+3}{3} \rfloor}, & \text{otherwise.} \end{cases}$$

Notice that another way to look at this sequence is to group the numbers into triplets. Then the third number (last number) in each group would exhibit quadratic growth: $1, 4, 9, 16, 25...$

We have also obtained adjusted numerators for mod 4 and 5:

- mod 4: 0, 0, 0, 1, 1, 2, 4, 8, 8, 12, 18, 27, 27, 36, 48, 64, ...
  Numerator OEIS Sequence $A008218$, formula $\lfloor \frac{i}{4} \rfloor \lfloor \frac{i+1}{4} \rfloor \lfloor \frac{i+2}{4} \rfloor$.
  Denominator OEIS Sequence $A008233$ (slight adjustments made), formula $\lfloor \frac{n+1}{4} \rfloor \lfloor \frac{n+2}{4} \rfloor \lfloor \frac{n+3}{4} \rfloor \lfloor \frac{n+4}{4} \rfloor$.
- mod 5: 0, 0, 0, 0, 1, 1, 2, 4, 8, 16, 16, 24, 36, 54, 81, 81,...
  Numerator OEIS Sequence $A008381$, formula $\lfloor \frac{i}{5} \rfloor \lfloor \frac{i+1}{5} \rfloor \lfloor \frac{i+2}{5} \rfloor \frac{i+3}{5} \rfloor$.
  Denominator OEIS Sequence $A008382$ (slight adjustments made), formula $\lfloor \frac{n+1}{5} \rfloor \lfloor \frac{n+2}{5} \rfloor \lfloor \frac{n+3}{5} \rfloor \lfloor \frac{n+4}{5} \rfloor \lfloor \frac{n+5}{5} \rfloor$.

Using similar ways to look at these sequences, we observe that if the game mode is mod k, then for each person A_i where $i \equiv 0 \pmod{k}$, i.e. i is a multiple of k, the numerator of survival probability (with a common denominator being the least common multiple) grows polynomially exponent of $k - 1$.

Additionally, with the numerator and denominator formulae, we are able to make a conjecture for general formulae for mod 4 and mod 5 as well:

Formula: In the mod 4 case, the winning probability for the person at position i among n people initially is:

$$p_n(i) = \begin{cases} 1, & \text{if } n = 1 \\ \frac{\lfloor \frac{i}{4} \rfloor \lfloor \frac{i+1}{4} \rfloor \lfloor \frac{i+2}{4} \rfloor}{\lfloor \frac{n+1}{4} \rfloor \lfloor \frac{n+2}{4} \rfloor \lfloor \frac{n+3}{4} \rfloor \lfloor \frac{n+4}{4} \rfloor}, & \text{otherwise.} \end{cases}$$

Formula: In the mod 5 case, the winning probability for the person at position i among n people initially is:

$$p_n(i) = \begin{cases} 1, & \text{if } n = 1 \\ \frac{\lfloor \frac{i}{5} \rfloor \lfloor \frac{i+1}{5} \rfloor \lfloor \frac{i+2}{5} \rfloor \frac{i+3}{5} \rfloor}{\lfloor \frac{n+1}{5} \rfloor \lfloor \frac{n+2}{5} \rfloor \lfloor \frac{n+3}{5} \rfloor \lfloor \frac{n+4}{5} \rfloor \lfloor \frac{n+5}{5} \rfloor}, & \text{otherwise.} \end{cases}$$

There are some patterns we could observe from our proved mod 2 formula and our conjectured formulae for the mod 3, mod 4 and mod 5 variations. When $n = 1$, the survival probability is 1 since there is only one person to begin, so this person has automatically survived the game. When $n > 1$ and in mod k, the numerator has $k-1$ floor function terms and the dominator has k floor function terms. Inside the floor functions, the numerators of the $k-1$ terms for the numerator of the formula are just $i+1, i+2, ..., i+(k-2)$ and the denominator $n+1, n+2, ..., n+k$, while all the denominators are k. By these observations. we conjecture the generalized formula for any mod k case:

Formula: In the mod k case, the winning probability for the person at position i among n people initially is:

$$p_n(i) = \begin{cases} 1, & \text{if } n = 1 \\ \frac{\prod_{j=0}^{k-2} \lfloor \frac{i+j}{k} \rfloor}{\prod_{q=1}^{k} \lfloor \frac{n+q}{k} \rfloor}, & \text{otherwise.} \end{cases}$$

5 Open Questions and Future Research

In this paper, we have presented a novel type of probabilistic counting-out game. By using first-step analysis, we have solved for exact survival probabilities values for any person in this game, and proved an exact formula. We have also presented different variations of the game, the mod k modes, where instead of randomly selecting the first person in every pair of persons, we randomly select every first person in groups of 2, 3, ..., k.

With the help of $OEIS$[6], we have identified various sequences in the numerators and denominators of mod k survival probabilities, and made a conjecture on the general formula of survival probabilities on any mod k game.

As shown earlier, we are able to prove the original problem, the mod 2 case formula, by induction and first-step analysis. While we agree that it is possible to apply induction and first-step analysis to any mod k game, one of the key challenges we are currently facing is the subcase explosion. In Section 3, due to the constant odd-even switch resulting in slight variations, we are forced to split the problem into four different subcases, and use first-step analysis to deal with each of them separately. In mod 3, the problem will further expand to six different subcases, because for any person in the line, s/he can start as the first, second or last person in the group, and after taking the first step, the person can end up in two different positions (shift up or stay). Likewise, mod 4 induction proof will require discussions on eight subcases, and mod 5 on ten subcases...

Although by tedious repetitions, we are able to use first-step analysis and induction to prove any mod k formula, in order to verify the general mod k formula, we need to find other more efficient strategies.

Another interesting aspect we are planning to explore is the correlation between this problem and other preexisting questions. It is easy to see that this counting-out game bear some resemblances to the Josephus problem [1]. After identifying the numerator and denominator sequences on OEIS, we are excited about the potentials of drawing similarities between our problem and other topics.

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): 4752. doi:10.2307/3608532. JSTOR 3608532, 1960.

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

[3] LingJian: Answer to ZhiHu "Probabilistic Count-Out Problem", `https://www.zhihu.com/question/55445739/answer/144612718`

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

[5] 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`

[6] The On-Line Encyclopedia of Integer Sequences `https://oeis.org/`