PROBABILISTIC COUNTING-OUT GAME ON A LINE

by
Tingting Ou

A thesis submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Science in Engineering

Baltimore, Maryland

May 2020

Abstract

This thesis attempts to solve a novel problem originally posted on a question-and-answer website. In a counting-out game, 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 people 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 two questions: which initial position has the greatest chance to win and which has the longest expected time to stay in the line.

To answer the two questions above, we use a recursive approach to solve for exact values of the winning probabilities and the expected survival time, prove the exact formula for the winning probabilities and derive the asymptotic behaviors of the expected survival time for some locations. We have also considered a variation 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. Recursive equations are constructed for this generalized case, and a proof of the exact formula for the winning probabilities is provided as well. Finally, we present other possible extensions and discuss future research directions concerning this problem.

Primary Reader and Advisor: John Wierman

Acknowledgments

I am very grateful to my advisor, Professor John Wierman for his guidance along the road of my research. I also want to thank Michelle Shu, who introduced this problem to me and collaborated with me in the early stage of my research.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This thesis 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 thesis, we will answer the question from two different perspectives. The solution can be based on the probability to survive all rounds of elimination and win the game eventually, which we refer to as *winning probability*, or 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 largest winning probability or the longest expected survival time would be the most favorable compared to other positions.

The problem studied in this thesis resembles the famous Josephus Problem [1]. In the Josephus problem, some players stand in a circle and one is chosen 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, who 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 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 because the elimination process is probabilistic rather than deterministic.

As mentioned previously, this problem was originally 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 more than 3,000 users on the website, the question has drawn a great deal of attention during the past two years. There are more than a hundred 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. The problem is also posted on Quora [4], receiving answers based on either simulations or recursive calculations.

In this thesis, 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. In addition to the original problem where we select a person at an odd position k such that $k \equiv 1 \pmod 2$ each time, we will also study a generalized version of the problem, which we will refer to as the $\bmod m$ case: a person at a position k such that $k \equiv 1 \pmod m$ is selected in each round, where $m \in \mathbb{N},\ m \geq 2$. Note that by definition, the original problem is the $\bmod 2$ case. Unless noted otherwise, the thesis will be addressing the $\bmod 2$ case. Results of the $\bmod m$

general case are presented in Chapter 7.

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 thesis is structured as follows:

In Chapter 2, recursive equations are constructed, then the exact values of the winning probabilities and the expected survival time for each person are determined recursively for n up to 9.

In Chapter 3, the exact formula of the winning probability for each person is provided and proved by induction.

Chapter 4 gives the exact expected survival time formula of Person 1.

Lower bounds and upper bounds for the expected survival time of Person 2 are provided in Chapter 5. An asymptotic approach is introduced in this chapter as well.

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

Chapter 7 gives the exact expected survival time formula of Person n (the last person). The formula is proved by induction.

Chapter 8 is dedicated to the discussion of the generalized $\bmod m$ case. The formula of the winning probabilities in the generalized case is proved by induction.

Chapter 9 is a brief summary of our results. We conclude the thesis with Chapter 10, which mentions future directions of this research and some open questions.

Chapter 2

Recursive Methods and Calculations

We will first introduce some notation used in this chapter. Let $p_n(k)$ denote the probability that Person k wins the game and $E_n(k)$ denote the expected survival time of Person k with n people in the line initially. In this thesis, 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 $n - 1$.

2.1 Winning Probability

We will use a bottom-up recursion to solve for exact values of the winning probabilities numerically. To start with, we first consider the base cases when n is small.

If $n = 1$, Person 1 wins the game in the first round automatically. Therefore, $p_1(1) = 1$.

If $n = 2$, Person 1 will be eliminated in the first round because s/he is the only odd-indexed person, so $p_2(1) = 0$. Person 2 will be the only person in the second round and win the game. Thus, $p_2(2) = 1$.

If $n = 3$, let us consider Person 1 first. Since 1 is the smallest index, Person 1's position remains the same until s/he is selected and eliminated. Thus, Person 1 will never win the game. Therefore, we have $p_n(1) = 0,$ for all n. This implies that $p_3(1) = 0.$

Then let us consider Person 2. In the first round, this person is even-positioned so s/he will not get eliminated. Consider the result of the first round: P(Person 1 selected) = P(Person 3 selected) = $\frac{1}{2}$, i.e. Person 2 has equal probabilities to shift to the first position (if Person 1 is eliminated at first) or remain at the second position (if Person 3 is eliminated at first). If Person 2 is shifted to the first position, then s/he has probability $p_2(1) = 0$ to win. If Person 2 keeps the second position, then s/he has probability

$p_2(2) = 1$ to win. Thus, we use a first-step decomposition and condition on the result of the first round:

$$
\begin{aligned}
p_3(2) &= P(\text{shifted to position } 1) \cdot P(\text{win} \mid \text{shifted}) \\
&\quad + P(\text{remain at position } 2) \cdot P(\text{win} \mid \text{remain}) \\
&= \frac{1}{2} \cdot p_2(1) + \frac{1}{2} \cdot p_2(2) \\
&= \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1 \\
&= \frac{1}{2}.
\end{aligned}
$$

Finally, we consider Person 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 3 is either eliminated in the first round or shifted to the second position. If the person is shifted to the second position , s/he has probability $p_2(2)$ to win. Thus we have:

$$
\begin{aligned}
p_3(3) &= P(\text{shifted to position } 2) \cdot P(\text{win} \mid \text{shifted}) \\
&= \frac{1}{2} \cdot p_2(2) \\
&= \frac{1}{2}.
\end{aligned}
$$

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 use this bottom-up recursive approach to calculate the exact value for any winning probability $p_n(k)$.

$$\begin{aligned}
p_n(k) = {} & P(\text{shifted to position } (k-1)) \cdot P(\text{win} \mid \text{shifted to position } (k-1)) \\
& + P(\text{remain at position } (k-1)) \cdot P(\text{win} \mid \text{remain at position } (k-1)) \\
= {} & P(\text{shifted to position } (k-1)) \cdot p_{n-1}(k-1) \\
& + P(\text{remain at position } (k-1)) \cdot p_{n-1}(k).
\end{aligned}$$

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 for the winning probabilities becomes:

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

Since we have computed the winning probabilities for $n = 1, 2, 3$, we can compute the winning probabilities for $n \geq 4$ using the above recursion. Up to $n = 9$, using the recursion method, we have obtained the following results in Table 2.1. Note that Person k is abbreviated as Pk in the table, and we have used a common denominator for all probabilities within the same row.

n	$P1$	$P2$	$P3$	$P4$	$P5$	$P6$	$P7$	$P8$	$P9$
1	1								
2	0	1							
3	0	1/2	1/2						
4	0	1/4	1/4	2/4					
5	0	1/6	1/6	2/6	2/6				
6	0	1/9	1/9	2/9	2/9	3/9			
7	0	1/12	1/12	2/12	2/12	3/12	3/12		
8	0	1/16	1/16	2/16	2/16	3/16	3/16	4/16	
9	0	1/20	1/20	2/20	2/20	3/20	3/20	4/20	4/20

Table 2.1: Winning probabilities of each person for $n = 1, ..., 9$.

From Table 2.1, we see the winning probability increases approximately linearly as index k increases. Also, we can see a clear pattern in the winning probabilities within each level (row). We will prove the exact formula of winning probabilities in Chapter 3 and the generalized formula in Chapter 7.

2.2 Expected Survival Time

The idea of recursion for the expected survival time is very similar to that for the winning probabilities. Again, we first 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, $E_2(2) = 1$.

If $n = 3$, we can clearly see that 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:

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

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:

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

With the base cases solved above, we can use recursion to calculate more results. 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. Using this recursive idea, we can construct the following recursive equation based on the first-step decomposition:

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

Recall that we have obtained these two formulae in the previous section:

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

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

Using the above formulae of probabilities, the general recursive relation for the expected survival time becomes:

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

$$P(\text{remain at position } k) \cdot (1 + E_{n-1}(k))$$

$$= \frac{\lceil \frac{k-1}{2} \rceil}{\lceil \frac{n}{2} \rceil}(1 + E_{n-1}(k-1)) + \left(1 - \frac{\lceil \frac{k}{2} \rceil}{\lceil \frac{n}{2} \rceil}\right)(1 + E_{n-1}(k)).$$

Up to $n = 9$, using the recursion method, we have obtained the following table of results. Person k is abbreviated as Pk, 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 2.2: 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 2.1 shows a plot when n varies from 2 (the bottom line) to 20 (the top line). XieZhuoFan [3] on Zhihu claims that

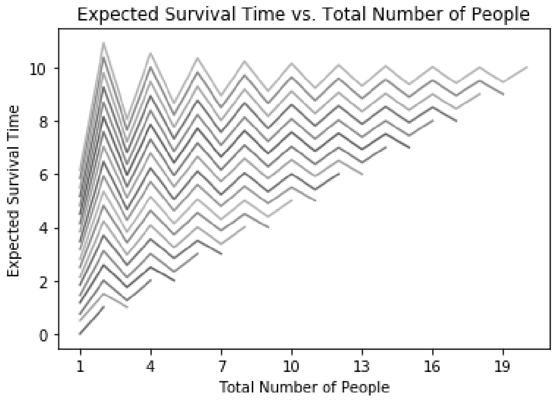

Figure 2.1: Expected survival time of each person for $n = 2, ..., 20$. For odd-indexed positions, the expected survival time increases. For even-indexed positions, the expected survival time decreases. The peak occurs at position 2.

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 Chapter 4 and Chapter 5. Based on our recursive calculations and plots, XieZhuoFan's first claim also appears to be true. In Chapter 6, 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.

Chapter 3

Winning Probabilities

We can express the winning probability of each person as a function of k and n, where k is the initial position index and n is the total number of people in the game. From observation of the values in Table 2.1, we claim that the winning probability grows linearly with respect to the initial position index k and that the formula is $\frac{\lfloor \frac{k}{2} \rfloor}{\lfloor \frac{n}{2} \rfloor \lfloor \frac{n+1}{2} \rfloor}$. A natural idea to prove the formula is to use induction and break into four cases based on the parity of k and n. The initial proof was excessively long, and it could not be adapted to prove the generalized mod m case since there are too many cases (m^2 cases) to consider. We provide a refined proof that does not rely on cases and could be generalized to prove the mod m case.

Theorem: The winning probability for the person at initial position k

(i.e. Person k) among n people initially is:

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

Proof: We will use induction to prove the result. Note that we have shown Person 1 always loses the game unless s/he is the only player, i.e. $n = 1$.

Base cases:

Let $n = 1$. This formula is trivially true.

Let $n = 2$. Person 1 is selected in the first round. Therefore, Person 1 never wins, and Person 2 wins with probability one. When $k = 1$, we have $p_2(1) = \frac{\lfloor \frac{1}{2} \rfloor}{\lfloor \frac{2}{2} \rfloor \lfloor \frac{2+1}{2} \rfloor} = \frac{0}{1 \cdot 1} = 0$. When $k = 2$, we have $p_2(2) = \frac{\lfloor \frac{2}{2} \rfloor}{\lfloor \frac{2}{2} \rfloor \lfloor \frac{2+1}{2} \rfloor} = \frac{1}{1 \cdot 1} = 1$.

Let $n = 3$. If Person 1 gets picked (with probability $\frac{1}{2}$) in the first round, then Person 2 becomes the new Person 1 and can never win, so Person 3 is the winner. If Person 3 gets picked (with probability $\frac{1}{2}$), then Person 2 is the winner because the Person 1 would be eliminated. We then verify the formula for all possible k. When $k = 1$, $p_3(1) = \frac{\lfloor \frac{1}{2} \rfloor}{\lfloor \frac{3}{2} \rfloor \lfloor \frac{3+1}{2} \rfloor} = \frac{0}{1 \cdot 2} = 0$. When $k = 2$, $p_3(2) = \frac{\lfloor \frac{2}{2} \rfloor}{\lfloor \frac{3}{2} \rfloor \lfloor \frac{3+1}{2} \rfloor} = \frac{1}{1 \cdot 2} = \frac{1}{2}$. When $k = 3$, $p_3(3) = \frac{\lfloor \frac{3}{2} \rfloor}{\lfloor \frac{3}{2} \rfloor \lfloor \frac{3+1}{2} \rfloor} = \frac{1}{1 \cdot 2} = \frac{1}{2}$.

Induction Hypothesis: We assume our formula is correct for all possible k, $k = 1, 2, 3, ..., n$ in level n, i.e. the game where there are n people initially.

Inductive Step: We prove the correctness of our formula for level $n+1$.

Let S_k be the event that Person k gets shifted to the previous position $(k-1)$ in the next round when there are $(n+1)$ people, R_k be the event that Person k remains at the same position k when there are $(n+1)$ people.

If Person k is shifted, we must have an odd-indexed person selected standing in front of him/her. There are $\lfloor \frac{k}{2} \rfloor$ odd-indexed people among Person $1, 2, ..., k-1$, and there are a total of $\lfloor \frac{n+2}{2} \rfloor$ odd-indexed people among $(n+1)$ people. Thus, we have $P(S_k) = \frac{\lfloor \frac{k}{2} \rfloor}{\lfloor \frac{n+2}{2} \rfloor}$.

If Person k remains at the same position, we must have an odd-indexed person selected standing behind him/her. That is to say, Person $1, 2, ..., k$ must NOT be chosen in that particular round. There are $\lfloor \frac{k+1}{2} \rfloor$ odd-indexed people among Person $1, 2, ..., k$, and there are a total of $\lfloor \frac{n+2}{2} \rfloor$ odd-indexed people among $(n+1)$ people. Thus, we have $P(R_k) = 1 - \frac{\lfloor \frac{k+1}{2} \rfloor}{\lfloor \frac{n+2}{2} \rfloor}$.

Now let's consider the recursion. Since we know $p_n(k-1)$ and $p_n(k)$ by the induction hypothesis, we can plug these values into the recursive relation and obtain $p_{n+1}(k)$. In the following calculations, in the step from (3.3) to (3.4), we multiply the middle term by $\frac{\lfloor \frac{n+2}{2} \rfloor}{\lfloor \frac{n+2}{2} \rfloor}$ so that all the terms share a common denominator. In the step from (3.5) to (3.6), notice that:

$\lfloor\frac{k-1}{2}\rfloor - \lfloor\frac{k+1}{2}\rfloor = \lfloor\frac{k-1}{2}\rfloor - \left(\lfloor\frac{k-1}{2}\rfloor + 1\right) = -1$. We have:

$$
\begin{aligned}
p_{n+1}(k) &= P(S_k)p_n(k-1) + P(R_k)p_n(k) && (3.1)\\
&= \left(\frac{\lfloor\frac{k}{2}\rfloor}{\lfloor\frac{n+2}{2}\rfloor}\right)\left(\frac{\lfloor\frac{k-1}{2}\rfloor}{\lfloor\frac{n}{2}\rfloor\lfloor\frac{n+1}{2}\rfloor}\right) + \left(1 - \frac{\lfloor\frac{k+1}{2}\rfloor}{\lfloor\frac{n+2}{2}\rfloor}\right)\left(\frac{\lfloor\frac{k}{2}\rfloor}{\lfloor\frac{n}{2}\rfloor\lfloor\frac{n+1}{2}\rfloor}\right) && (3.2)\\
&= \frac{\lfloor\frac{k-1}{2}\rfloor\lfloor\frac{k}{2}\rfloor}{\lfloor\frac{n}{2}\rfloor\lfloor\frac{n+1}{2}\rfloor\lfloor\frac{n+2}{2}\rfloor} + \frac{\lfloor\frac{k}{2}\rfloor}{\lfloor\frac{n}{2}\rfloor\lfloor\frac{n+1}{2}\rfloor} - \frac{\lfloor\frac{k}{2}\rfloor\lfloor\frac{k+1}{2}\rfloor}{\lfloor\frac{n}{2}\rfloor\lfloor\frac{n+1}{2}\rfloor\lfloor\frac{n+2}{2}\rfloor} && (3.3)\\
&= \frac{\lfloor\frac{k-1}{2}\rfloor\lfloor\frac{k}{2}\rfloor + \lfloor\frac{k}{2}\rfloor\lfloor\frac{n+2}{2}\rfloor - \lfloor\frac{k}{2}\rfloor\lfloor\frac{k+1}{2}\rfloor}{\lfloor\frac{n}{2}\rfloor\lfloor\frac{n+1}{2}\rfloor\lfloor\frac{n+2}{2}\rfloor} && (3.4)\\
&= \frac{\lfloor\frac{k}{2}\rfloor\lfloor\frac{n+2}{2}\rfloor + \lfloor\frac{k}{2}\rfloor\left(\lfloor\frac{k-1}{2}\rfloor - \lfloor\frac{k+1}{2}\rfloor\right)}{\lfloor\frac{n}{2}\rfloor\lfloor\frac{n+1}{2}\rfloor\lfloor\frac{n+2}{2}\rfloor} && (3.5)\\
&= \frac{\lfloor\frac{k}{2}\rfloor\lfloor\frac{n+2}{2}\rfloor + \lfloor\frac{k}{2}\rfloor\{-1\}}{\lfloor\frac{n}{2}\rfloor\lfloor\frac{n+1}{2}\rfloor\lfloor\frac{n+2}{2}\rfloor} && (3.6)\\
&= \frac{\lfloor\frac{k}{2}\rfloor\left\{\lfloor\frac{n+2}{2}\rfloor - 1\right\}}{\lfloor\frac{n}{2}\rfloor\lfloor\frac{n+1}{2}\rfloor\lfloor\frac{n+2}{2}\rfloor} && (3.7)\\
&= \frac{\lfloor\frac{k}{2}\rfloor\left\{\lfloor\frac{n}{2}\rfloor\right\}}{\lfloor\frac{n}{2}\rfloor\lfloor\frac{n+1}{2}\rfloor\lfloor\frac{n+2}{2}\rfloor} && (3.8)\\
&= \frac{\lfloor\frac{k}{2}\rfloor}{\lfloor\frac{n+1}{2}\rfloor\lfloor\frac{n+2}{2}\rfloor}. && (3.9)
\end{aligned}
$$

Thus the theorem is correct by mathematical induction.

Chapter 4

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. Our goal is to calculate $E[T]$, which is the same as $E_n(1)$ defined in Chapter 2. To obtain the expected survival time, 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:

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

$$P(T=2) = P(\text{selected in the 3nd round, survive the first 2 rounds})$$

$$= \frac{1}{\lceil \frac{n-2}{2} \rceil} \left(1 - \frac{1}{\lceil \frac{n}{2} \rceil}\right) \left(1 - \frac{1}{\lceil \frac{n-1}{2} \rceil}\right)$$

$$= \frac{1}{m-1} \left(\frac{m-1}{m}\right) \left(\frac{m-2}{m-1}\right)$$

$$= \frac{1}{m} \left(\frac{m-2}{m-1}\right).$$

$$P(T=3) = P(\text{selected in the 4th round, survive the first 3 rounds})$$

$$= \frac{1}{\lceil \frac{n-3}{2} \rceil} \left(1 - \frac{1}{\lceil \frac{n}{2} \rceil}\right) \left(1 - \frac{1}{\lceil \frac{n-1}{2} \rceil}\right) \left(1 - \frac{1}{\lceil \frac{n-2}{2} \rceil}\right)$$

$$= \frac{1}{m-2} \left(\frac{m-1}{m}\right) \left(\frac{m-2}{m-1}\right) \left(\frac{m-2}{m-1}\right)$$

$$= \frac{1}{m} \left(\frac{m-2}{m-1}\right).$$

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 ($T = 2i + 1$).

For even values,

$$P(T = 2i) = P(\text{selected in the } (2i+1)\text{th round, not the first } (2i) \text{ rounds})$$

$$= \frac{1}{\lceil \frac{n-2i}{2} \rceil} \left(1 - \frac{1}{\lceil \frac{n}{2} \rceil}\right) \left(1 - \frac{1}{\lceil \frac{n-1}{2} \rceil}\right) \cdots \left(1 - \frac{1}{\lceil \frac{n-(2i-1)}{2} \rceil}\right)$$

$$= \frac{1}{m-i} \left(\frac{m-1}{m}\right) \left(\frac{m-2}{m-1}\right) \left(\frac{m-2}{m-1}\right) \cdots$$

$$\left(\frac{m-i}{m-(i-1)}\right) \left(\frac{m-i}{m-(i-1)}\right) \left(\frac{m-(i+1)}{m-i}\right)$$

$$= \frac{(m-1)(m-i)^2[m-(i+1)]}{(m-i)(m)(m-1)^2(m-i)}$$

$$= \frac{1}{m} \left(\frac{m-1-i}{m-1}\right).$$

For odd values,

$$P(T = 2i+1) = P(\text{selected in } (2i+2)\text{th round, not in first } (2i+1) \text{ rounds})$$

$$= \frac{1}{\lceil \frac{n-(2i+1)}{2} \rceil} \left(1 - \frac{1}{\lceil \frac{n}{2} \rceil}\right) \cdots \left(1 - \frac{1}{\lceil \frac{n-2i}{2} \rceil}\right)$$

$$= \frac{1}{m-i-1} \left(\frac{m-1}{m}\right) \left(\frac{m-2}{m-1}\right) \left(\frac{m-2}{m-1}\right) \cdots$$

$$\left(\frac{m-(i+1)}{m-i}\right) \left(\frac{m-(i+1)}{m-i}\right)$$

$$= \frac{1}{m} \left(\frac{m-1-i}{m-1}\right).$$

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 Chapter 2 as well as the formula of the odd case to solve for the expected survival time:

$$\begin{aligned} E_n(1) &= P(\text{remain at position } 1) \cdot (E_{n-1}(1) + 1) \\ &= \left(1 - \frac{1}{n/2}\right)\left(\frac{1}{3}(n-1) - \frac{1}{2} + 1\right) \\ &= \frac{1}{3}(n-1) + \frac{1}{2} - \frac{2}{n}\left(\frac{1}{3}(n-1) + \frac{1}{2}\right) \\ &= \frac{1}{3}n + \frac{1}{6} - \left(\frac{2}{3} - \frac{2}{3n} + \frac{1}{n}\right) \\ &= \boxed{\frac{1}{3}n - \frac{1}{2} - \frac{1}{3n}}, \text{ if } n \text{ is even.} \end{aligned}$$

When n is sufficiently large, the term $\frac{1}{3n}$ is negligible. What's more, the two lower-order terms $\frac{1}{2}$ and $\frac{1}{3n}$ together are $O(1)$, while the leading term is $O(n)$. Therefore, the expected survival time is approximately $\frac{1}{3}n$ if n is sufficiently large.

Chapter 5

Expected Survival Time of Person 2

5.1 Lower and Upper Bounds

In Chapter 4, we calculated the expected survival time of Person 1. For odd n, $E_n(1) = \frac{1}{3}n - \frac{1}{2}$, and for even n, $E_n(1) = \frac{1}{3}n - \frac{1}{2} - \frac{1}{3n}$. 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 for $E_n(2)$. 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

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

Adding all three terms, we get:

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

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 for $E_n(2)$.

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

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

5.2 Asymptotic Approach

If n is sufficiently large, then the exact value of the lower-order terms in $E_n(1)$ is unimportant. The lower-order terms are $O(1)$, so Person 1 is expected to survive for time $E[T] = \frac{1}{3}n + O(1)$.

CHAPTER 5. EXPECTED SURVIVAL TIME OF PERSON 2

Let the survival time of Person 2 be denoted by T_2 and T denote the survival time of Person 1 again. We apply the law of total expectation, conditioning on the time when Person 2 gets shifted (denoted by T_{shift}) to compute $E[T_2]$. Shifting Person 2 is equivalent to selecting Person 1, so we have $T_{shift} = T$ and $E[T_{shift}] = E[T] = \frac{1}{3}n + O(1)$.

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 $\frac{1}{3}(n - T_{shift}) + O(1)$. Thus,

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

Therefore, the expected survival time of Person 2 is approximately $\frac{5}{9}n$ from an asymptotic perspective.

Chapter 6

Expected Survival Time of Person 3 and Person 4

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

Person 3

We consider the third person first. 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:

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

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 approximately $\frac{5}{9}(n-T_{ss})$ 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 Chapter 4. In order to recognize the pattern, we perform more calculations, but for the sake of brevity, only the first few example calculations are included:

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

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

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

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

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

After simplification, the generalized formulae of the probability mass functions of even T ($T = 2i$) and odd T ($T = 2i+1$) for $i = 0, 1, 2, 3...$ are

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

$$P(T_{ss} = 2i+1) = P(\text{Person 1 or 3 eliminated in the (2i+2)th round,}$$

$$\text{both survive the first (2i+1) rounds})$$

$$= \frac{2(m-(i+2))(m-(i+2))(m-(i+1))}{(m-2)(m-1)(m-1)m}$$

$$= \frac{2(m-i-2)(m-i-2)(m-i-1)}{(m-2)(m-1)(m-1)m}$$

$$= \frac{2(m-i-2)^2(m-i-1)}{(m-2)(m-1)^2m}.$$

Using the probability mass function above, we can compute the expected time until Person 1 or 3 gets selected. Note that $m = \frac{n+1}{2}$. We used Wolfram Alpha to compute the sum of powers, i.e. $\sum_{i=0}^{m-1} 4i \cdot (m-i-2)(m-i-1)^2$ and $\sum_{i=0}^{m-1} 4i \cdot (m-i-2)^2(m-i-1)$.

$$E[T_{ss}] = \sum_{t=0}^{n} t \cdot P(T_{ss} = t)$$

$$= \sum_{i=0}^{m-1} 2i \cdot P(T_{ss} = 2i) + \sum_{i=0}^{m-1} (2i+1) \cdot P(T_{ss} = 2i+1)$$

$$= \sum_{i=0}^{m-1} 2i \cdot \left(\frac{2(m-i-2)(m-i-1)^2}{(m-2)(m-1)(m-1)m} \right)$$

$$+ \sum_{i=0}^{m-1} (2i+1) \cdot \left(\frac{2(m-i-2)^2(m-i-1)}{(m-2)(m-1)(m-1)m} \right)$$

$$= \frac{\sum_{i=0}^{m-1} 4i \cdot (m-i-2)(m-i-1)^2}{(m-2)(m-1)^2m} + \frac{\sum_{i=0}^{m-1} 4i \cdot (m-i-2)^2(m-i-1)}{(m-2)(m-1)^2m}$$

$$= \frac{1}{(m-2)(m-1)^2 m} \cdot \left(\frac{1}{5}m^5 - \frac{4}{3}m^4 + 3m^3 - \frac{8}{3}m^2 + \frac{4}{5}m\right)$$

$$+ \frac{1}{(m-2)(m-1)^2 m} \cdot \left(\frac{1}{5}m^5 - \frac{7}{6}m^4 + \frac{8}{3}m^3 - \frac{17}{6}m^2 + \frac{17}{15}m\right)$$

$$= \frac{1}{(m-2)(m-1)^2 m} \left(\frac{2}{5}m^5 - \frac{5}{2}m^4 + \frac{17}{3}m^3 - \frac{11}{2}m^2 + \frac{29}{15}m\right).$$

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 asymptotic 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 chapter and Chapter 4, 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 $\frac{1}{5}n$. Additionally, we have made a conjecture that the expected time until Person k is shifted before selected is approximately $\frac{1}{3}n, \frac{1}{5}n, \frac{1}{7}n, \frac{1}{9}n$ when $k = 1, 3, 5, 7, ...$

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.

Although we have only computed the expected survival time for Person 1, 2, 3, 4 (respectively $\frac{1}{3}n$, $\frac{5}{9}n$, $\frac{19}{45}n$ and $\frac{121}{225}n$), we can still observe a zig-zag pattern in the numbers: $\frac{5}{9} > \frac{19}{45} > \frac{1}{3}$, and $\frac{19}{45} < \frac{121}{225} < \frac{5}{9}$. If we use the method mentioned above and compute the expected survival time of more people (Persons 5, 6, 7,...), we may be able to observe the zig-zag pattern more clearly, or even prove it.

Chapter 7

Expected Survival Time of Person n

From Table 2.2 in Chapter 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 2.2 in Chapter 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 $1, 2, ..., n$.

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 Chapter 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

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

$$= E_n(n) + 1$$

$$= \lfloor \frac{n}{2} \rfloor + 1$$

$$= \frac{n-1}{2} + 1$$

$$= \lfloor \frac{n+1}{2} \rfloor.$$

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

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

$$= \left(1 - \frac{1}{\lceil \frac{n+1}{2} \rceil}\right)\left(\lfloor \frac{n}{2} \rfloor + 1\right) \quad \text{by hypothesis}$$

$$= \left(1 - \frac{2}{n+2}\right) \frac{n+2}{2}$$

$$= \frac{n}{2}$$

$$= \lfloor \frac{n+1}{2} \rfloor.$$

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. $\square$

Chapter 8

Generalization: the $\bmod m$ case

Now we discuss a generalized version of the problem. In the original problem, we select one person who stands at position k such that $k \equiv 1 \pmod 2$ to leave the line in each round. In the generalized $\bmod m$ problem, we still eliminate one person per round, but now people standing at position k such that $k \equiv 1 \pmod m$ for some constant $m \geq 2$ are vulnerable. We have derived the formula of the winning probabilities in this generalized case.

Theorem: In the mod m case (where $m \geq 2$), the winning probability for the person at position k among n people initially is:

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

Proof:

Base cases: Consider the base case where $n = 3$. We will show that our formula is correct for all possible $k = 1, 2, 3$.

When $n \leq m$, this game is simply eliminating whoever stands in the first position in each round. Therefore, the last person, Person n always wins, and all other people will lose. Therefore, we have $p_n(k) = 1$ when $k = n$ and $p_n(k) = 0$ when $k < n$.

When $n = 3 > m$, the only feasible value for m is 2. Then this is our $\mod 2$ case which has been proved before. Thus the formula is correct when $n = 3$ for all possible k.

Induction Hypothesis: We assume our formula is correct for all possible k, $k = 1, 2, 3, ..., n$ in level n.

Inductive Step: We prove the correctness of our formula for level $n + 1$.

First, among a total of x people, there are $\lfloor\frac{x+m-1}{m}\rfloor$ people who are vulnerable. This can be verified by considering all m possible remainders when n is divided by m.

Let S be the event that Person k gets shifted to the previous position $(k-1)$ in the next round when there are $(n+1)$ people, and R be the event that Person k remains at the same position k when there are $(n+1)$ people.

If Person k is shifted, a person standing in front of him/her must be selected. There are $\lfloor\frac{k-1+m-1}{m}\rfloor = \lfloor\frac{k+m-2}{m}\rfloor$ vulnerable people among Persons $1, 2, ..., k-1$, and there are a total of $\lfloor\frac{n+1+m-1}{m}\rfloor = \lfloor\frac{n+m}{m}\rfloor$ vulnerable people among $(n+1)$ people. Thus, we have $P(S) = \frac{\lfloor\frac{k+m-2}{m}\rfloor}{\lfloor\frac{n+m}{m}\rfloor}$.

If Person k remains at the same position, we must have a person selected standing behind him/her. That is to say, Persons $1, 2, ..., k$ must not be chosen in that particular round. There are $\lfloor\frac{k+m-1}{m}\rfloor$ vulnerable people among Persons $1, 2, ..., k$, and there are a total of $\lfloor\frac{n+m}{m}\rfloor$ vulnerable people among $(n+1)$ people. Thus, we have $P(R) = 1 - \frac{\lfloor\frac{k+m-1}{m}\rfloor}{\lfloor\frac{n+m}{m}\rfloor}$.

Since we know the value of $p_n(k-1)$ and $p_n(k)$ by the induction hypothesis, we have the following equation hold for every $k = 1, 2, ..., n+1$. In the step from (8.2) to (8.3) we simply expand the equation. In the step from (8.3) to (8.4), we turn the product of $\prod_{j=0}^{m-1}\lfloor\frac{n+j}{m}\rfloor$ and $\lfloor\frac{n+m}{m}\rfloor$ in the denominators into a common factor, $\prod_{j=0}^{m}\lfloor\frac{n+j}{m}\rfloor$. In the step from (8.8) to (8.9), we cancel the $j = 0$ factor in the denominator. We re-index and change the range of the index in the final step.

$$
\begin{aligned}
p_{n+1}(k) &= P(S)p_n(k-1) + P(R)p_n(k) && (8.1)\\
&= \left(\frac{\lfloor\frac{k+m-2}{m}\rfloor}{\lfloor\frac{n+m}{m}\rfloor}\right)\left(\frac{\prod_{i=0}^{m-2}\lfloor\frac{k-1+i}{m}\rfloor}{\prod_{j=0}^{m-1}\lfloor\frac{n+j}{m}\rfloor}\right) && (8.2)\\
&+ \left(1 - \frac{\lfloor\frac{k+m-1}{m}\rfloor}{\lfloor\frac{n+m}{m}\rfloor}\right)\left(\frac{\prod_{i=0}^{m-2}\lfloor\frac{k+i}{m}\rfloor}{\prod_{j=0}^{m-1}\lfloor\frac{n+j}{m}\rfloor}\right) && (8.3)\\
&= \left(\frac{\lfloor\frac{k-1}{m}\rfloor}{\lfloor\frac{n+m}{m}\rfloor}\right)\left(\frac{\prod_{i=0}^{m-2}\lfloor\frac{k+i}{m}\rfloor}{\prod_{j=0}^{m-1}\lfloor\frac{n+j}{m}\rfloor}\right) + \left(\frac{\prod_{i=0}^{m-2}\lfloor\frac{k+i}{m}\rfloor}{\prod_{j=0}^{m-1}\lfloor\frac{n+j}{m}\rfloor}\right)\left(\frac{\lfloor\frac{n+m}{m}\rfloor}{\lfloor\frac{n+m}{m}\rfloor}\right) && (8.4)\\
&- \left(\frac{\prod_{i=0}^{m-2}\lfloor\frac{k+i}{m}\rfloor}{\prod_{j=0}^{m-1}\lfloor\frac{n+j}{m}\rfloor}\right)\left(\frac{\lfloor\frac{k+m-1}{m}\rfloor}{\lfloor\frac{n+m}{m}\rfloor}\right) && (8.5)\\
&= \left(\frac{\prod_{i=0}^{m-2}\lfloor\frac{k+i}{m}\rfloor}{\prod_{j=0}^{m}\lfloor\frac{n+j}{m}\rfloor}\right)\left(\left\lfloor\frac{k-1}{m}\right\rfloor + \left\lfloor\frac{n+m}{m}\right\rfloor - \left\lfloor\frac{k+m-1}{m}\right\rfloor\right) && (8.6)\\
&= \left(\frac{\prod_{i=0}^{m-2}\lfloor\frac{k+i}{m}\rfloor}{\prod_{j=0}^{m}\lfloor\frac{n+j}{m}\rfloor}\right)\left(\left\lfloor\frac{k-1}{m}\right\rfloor + \left\lfloor\frac{n+m}{m}\right\rfloor - \left\lfloor\frac{k-1}{m}\right\rfloor - 1\right) && (8.7)\\
&= \left(\frac{\prod_{i=0}^{m-2}\lfloor\frac{k+i}{m}\rfloor}{\prod_{j=0}^{m}\lfloor\frac{n+j}{m}\rfloor}\right)\left\lfloor\frac{n}{m}\right\rfloor && (8.8)\\
&= \frac{\prod_{i=0}^{m-2}\lfloor\frac{k+i}{m}\rfloor}{\prod_{j=1}^{m}\lfloor\frac{n+j}{m}\rfloor} && (8.9)\\
&= \frac{\prod_{i=0}^{m-2}\lfloor\frac{k+i}{m}\rfloor}{\prod_{j=0}^{m-1}\lfloor\frac{n+1+j}{m}\rfloor} && (8.10)
\end{aligned}
$$

Chapter 9

Conclusion

In summary, this thesis has introduced a recursive approach to solve for exact values of the winning probabilities and the expected survival time, and proved the exact formula of the winning probabilities and the expected survival time of Person 1 and Person n. Additionally, we have shown 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 Chapter 6 may be adapted to calculate the expected survival time of the other persons as well. Finally, we have also generalized the original problem and proved the exact formula of the winning probabilities in the generalized case.

Chapter 10

Future Research

As mentioned in Chapter 6, 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 chapters, 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 made the 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.

Another future direction is to work on the expected survival time of the generalized case. Although we have already proved the formula of the winning probabilities, we have not done any work on the expected survival time yet.

Finally, note that the generalized pattern of elimination in the mod m case is still very restricted. In the $\mathrm{mod}\ m$ case, in every group of m people, only the first person is vulnerable. It would be an interesting open problem if we consider the even more generalized case, where more people are vulnerable in addition to the first one in each group of m people. For example, when $m = 5$, in every group of 5 people, we can assume that the first and the third are vulnerable ("XOXOO"), or that the first and the fourth are vulnerable ("XOOXO").

Bibliography

[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

Vita

Tingting Ou is from Shenzhen, China. She is a student in the combined Bachelor's Master's program in Applied Mathematics and Statistics at Johns Hopkins University, under the supervision of Professor John Wierman. She also majors in Computer Science and (Pure) Mathematics. She is interested in applied probability, statistics and machine learning. In the future, Tingting will pursue a Ph.D. degree in Operations Research.