

Selected Problems in Discrete Probability

Mohamed Syaheer Altaf

September 23, 2025

In this note, we will explore a selection of questions in (discrete) probability and how to solve them. We begin with simple sequential experiments, continue with problems involving expectations, and finish with several classical topics—such as the *problem of points*, the *coupon collector's problem*, and *gambler's ruin*, to name a few—that often serve as abstractions for many discrete probability systems.

1 Simple Sequential Experiments

When tackling problems of this kind, one typically **constructs** suitable **events** (or defines appropriate **random variables**). It is also good practice to identify—or at least acknowledge—the experiment's **sample space**, often via the **counting principle** (e.g., enumerate all ordered pairs $(x, y) \in X \times Y$ with $x \in X$ and $y \in Y$).

In many cases, it is easier to compute **conditional probabilities** first and then use them with i) the **multiplication rule** to find **joint probabilities**, and ii) the **law of total probability** to obtain unconditional probabilities.

Q 1. *We roll two fair six-sided die. Compute the probability that the second die shows 5 given that the sum of the two die is odd.*

Answer. This is a straightforward exercise in building the appropriate sample space under the stated *condition*. Define $\Omega' = \{(x, y) \mid x, y \in \{1, \dots, 6\} \text{ and } x+y \text{ is odd}\}$. One checks quickly that $\Pr(\Omega') = \frac{18}{36}$.

Next, consider the intersection of the events (i) the second die shows 5 and (ii) the sum is odd. Fixing the outcome $(x, 5)$ and noting that 5 is odd, we need x to be even; thus $x \in \{2, 4, 6\}$.

Hence, by the definition of conditional probability,

$$\Pr(\text{second die is 5} \mid \text{sum is odd}) = \frac{\frac{3}{36}}{\frac{18}{36}} = \frac{1}{6}.$$

□

Q 2. *There are three special dice with the following faces:*

1. **die A:** 1, 2, 2, 2, 3, 3
2. **die B:** 1, 1, 2, 2, 3, 3
3. **die C:** 1, 1, 2, 2, 2, 3

All three dice are fair, each face appearing with probability $\frac{1}{6}$. Suppose you throw the three dice in an unknown order and observe two 2's and one 3. What is the probability that the die showing 3 is die C?

Answer. Without loss of generality, we construct an ordered tuple (a, b, c) that contains exactly two 2's and one 3; this will be the *conditional* sample space Ω' , and it is easy to see that $|\Omega'| = \frac{3!}{2!} = 3$.

We proceed to find the intersection. Notice that if we fix $c = 3$, there is only one such sequence—namely $(2, 2, 3)$. It remains to compute $\Pr(\Omega')$ via the multiplication rule:

$$\Pr(\Omega') = \Pr(2, 2, 3) + \Pr(2, 3, 2) + \Pr(3, 2, 2).$$

For example,

$$\Pr(2, 2, 3) = \Pr(A = 2) \Pr(B = 2 \mid A = 2) \Pr(C = 3 \mid A = 2 \cap B = 2) = \frac{1}{36} \quad (\text{by independence}).$$

Thus $\Pr(\Omega') = \frac{1}{6}$.

Finally, by the definition of conditional probability,

$$\Pr(\text{die C shows 3} \mid \text{two 2's and one 3}) = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6}.$$

□

Q 3. A bag contains three colours of balls: 2 red, 3 yellow, and 4 green. Three balls are drawn **without replacement**. Let the event \mathcal{E} = “the first ball is yellow” and the event \mathcal{F} = “the third ball is not red”. Find $\Pr(\mathcal{E} \mid \mathcal{F})$.

Answer. We have

$$\Pr(\mathcal{E} \mid \mathcal{F}) = \frac{\Pr(\mathcal{E}) \Pr(\mathcal{F} \mid \mathcal{E})}{\Pr(\mathcal{F})},$$

and, by the law of total probability,

$$\Pr(\mathcal{F}) = \Pr(\mathcal{E}) \Pr(\mathcal{F} \mid \mathcal{E}) + \Pr(\bar{\mathcal{E}}) \Pr(\mathcal{F} \mid \bar{\mathcal{E}}).$$

We compute each term. Clearly

$$\Pr(\mathcal{E}) = \frac{3}{9}, \quad \Pr(\bar{\mathcal{E}}) = \frac{6}{9}.$$

Conditional probabilities. Sampling is without replacement, so the draws are not independent. Let \mathcal{G} (taking values r, y, g) denote the colour of the *second* ball. For convenience we also let \mathcal{E} and \mathcal{F} stand for the colours of the first and third balls, respectively.

$$\begin{aligned} \Pr(\mathcal{F} \neq r \mid \mathcal{E} = y) &= \Pr(\mathcal{G} = r \mid \mathcal{E} = y) \Pr(\mathcal{F} \neq r \mid \mathcal{E} = y \cap \mathcal{G} = r) \\ &\quad + \Pr(\mathcal{G} = y \mid \mathcal{E} = y) \Pr(\mathcal{F} \neq r \mid \mathcal{E} = y \cap \mathcal{G} = y) \\ &\quad + \Pr(\mathcal{G} = g \mid \mathcal{E} = y) \Pr(\mathcal{F} \neq r \mid \mathcal{E} = y \cap \mathcal{G} = g) \\ &= \frac{2}{8} \cdot \frac{6}{7} + \frac{2}{8} \cdot \frac{5}{7} + \frac{4}{8} \cdot \frac{5}{7} = \frac{3}{4}. \end{aligned}$$

$$\begin{aligned}
\Pr(\mathcal{F} \neq r \mid \mathcal{E} = r) &= \Pr(\mathcal{G} = r \mid \mathcal{E} = r) \Pr(\mathcal{F} \neq r \mid \mathcal{E} = r \cap \mathcal{G} = r) \\
&\quad + \Pr(\mathcal{G} = y \mid \mathcal{E} = r) \Pr(\mathcal{F} \neq r \mid \mathcal{E} = r \cap \mathcal{G} = y) \\
&\quad + \Pr(\mathcal{G} = g \mid \mathcal{E} = r) \Pr(\mathcal{F} \neq r \mid \mathcal{E} = r \cap \mathcal{G} = g) \\
&= \frac{1}{8} \cdot \frac{7}{7} + \frac{3}{8} \cdot \frac{6}{7} + \frac{4}{8} \cdot \frac{6}{7} = \frac{7}{8}.
\end{aligned}$$

$$\begin{aligned}
\Pr(\mathcal{F} \neq r \mid \mathcal{E} = g) &= \Pr(\mathcal{G} = r \mid \mathcal{E} = g) \Pr(\mathcal{F} \neq r \mid \mathcal{E} = g \cap \mathcal{G} = r) \\
&\quad + \Pr(\mathcal{G} = y \mid \mathcal{E} = g) \Pr(\mathcal{F} \neq r \mid \mathcal{E} = g \cap \mathcal{G} = y) \\
&\quad + \Pr(\mathcal{G} = g \mid \mathcal{E} = g) \Pr(\mathcal{F} \neq r \mid \mathcal{E} = g \cap \mathcal{G} = g) \\
&= \frac{2}{8} \cdot \frac{6}{7} + \frac{3}{8} \cdot \frac{5}{7} + \frac{3}{8} \cdot \frac{5}{7} = \frac{3}{4}.
\end{aligned}$$

Putting the pieces together. We must weight the probability appropriately (note that we only know the first ball is not yellow—so what is the probability that it is red?). Given that the first ball is not yellow,

$$\Pr(\mathcal{E} = r \mid \bar{\mathcal{E}}) = \frac{2}{6}, \quad \Pr(\mathcal{E} = g \mid \bar{\mathcal{E}}) = \frac{4}{6}.$$

Hence

$$\Pr(\mathcal{F} \mid \bar{\mathcal{E}}) = \frac{2}{6} \cdot \frac{7}{8} + \frac{4}{6} \cdot \frac{3}{4} = \frac{19}{24}.$$

Finally,

$$\Pr(\mathcal{E} \mid \mathcal{F}) = \frac{\frac{3}{9} \cdot \frac{3}{4}}{\frac{3}{9} \cdot \frac{3}{4} + \frac{6}{9} \cdot \frac{19}{24}} = \frac{9}{28}.$$

□

Q 4 (Finding Probability Mass Function). *A fair four-sided die is rolled twice, with the two rolls independent. Let X and Y denote the results of the first and second rolls, respectively. Compute the conditional probability $\Pr(\mathcal{E} \mid \mathcal{F})$ where $\mathcal{E} = \{\max(X, Y) = m\}$ and $\mathcal{F} = \{\min(X, Y) = 2\}$ for $m \in \{1, 2, 3, 4\}$.*

Answer. First, note the entire sample space of the experiment: $\Omega = \{(x, y) \mid x, y \in \{1, 2, 3, 4\}\}$, so $|\Omega| = 4 \times 4 = 16$. Computing $\Pr(\mathcal{F})$ is straightforward: fix either x or y at 2 and let the other value be ≥ 2 ; subtract one to avoid double-counting (2, 2):

$$\Pr(\mathcal{F}) = \frac{2 \cdot 3 - 1}{16} = \frac{5}{16}.$$

Using the same strategy, we find $\Pr(\mathcal{E} \cap \mathcal{F})$ by fixing one value at m and the other at 2; thus $m \geq 2$ to satisfy \mathcal{E} . Enumerating all values of m and applying the definition of conditional probability, we obtain the following (conditional) probability mass function:

$$p_{\mathcal{E} \mid \mathcal{F}}(m) = \begin{cases} 0 & \text{if } m = 1, \\ \frac{1}{16} = \frac{1}{5} & \text{if } m = 2, \\ \frac{16}{16} = \frac{2}{5} & \text{if } m = 3, \\ \frac{16}{16} = \frac{2}{5} & \text{if } m = 4. \end{cases}$$

□

Q 5 (Binomial Random Variable). *Two players are engaged in a match that ends as soon as one of them wins 10 games. Each game is independent, and each player wins a game with probability $\frac{1}{2}$ (a game cannot be drawn). What is the probability that exactly 11 games have been played when the match ends?*

Answer. First, note that any two-outcome experiment in which we count the number of successes in a fixed number of trials is naturally modeled by the *binomial distribution*. (If, instead, we were interested in the number of trials required to achieve a fixed number of successes, we would use the *negative-binomial distribution*.)

If the match ends on the 11th game, then after 10 games one of the players must have accumulated 9 wins. Without loss of generality assume it is player *A*; we will later multiply by two for the symmetric case of player *B*.

Let A_n denote the number of wins by player *A* after n games. Then $A_n \sim \text{Binomial}(n, \frac{1}{2})$, so

$$\Pr(A_{10} = 9) = \binom{10}{9} \left(\frac{1}{2}\right)^9 \left(1 - \frac{1}{2}\right)^{10-9} = \frac{5}{512}.$$

Now apply the law of total probability:

$$\begin{aligned} \Pr(A \text{ wins on game } 11) &= \Pr(A_{10} = 9) \Pr(A \text{ wins game } 11 | A_{10} = 9) \\ &\quad + \Pr(A_{10} < 9) \Pr(A \text{ wins game } 11 | A_{10} < 9). \end{aligned}$$

Here $\Pr(A \text{ wins game } 11 | A_{10} = 9) = \frac{1}{2}$ and $\Pr(A \text{ wins game } 11 | A_{10} < 9) = 0$, so

$$\Pr(A \text{ wins on game } 11) = \frac{5}{512} \cdot \frac{1}{2} = \frac{5}{1024}.$$

Finally, multiply by two to account for player *B* in the symmetric scenario:

$$\Pr(\text{match ends in exactly 11 games}) = 2 \cdot \frac{5}{1024} = \frac{5}{512}.$$

□

Q 6 (Geometric Random Variable). *The probability mass functions for players *A* and *B* in a dart game are given below. Let X be the Bernoulli random variable for player *A* that takes the value 1 if *A* hits the bull's-eye (and 0 otherwise); define Y analogously for player *B*.*

$$p_A(x) = \begin{cases} \frac{1}{5} & \text{if } x = 1, \\ \frac{4}{5} & \text{if } x = 0, \end{cases} \quad p_B(y) = \begin{cases} \frac{1}{6} & \text{if } y = 1, \\ \frac{5}{6} & \text{if } y = 0. \end{cases}$$

*The players alternate turns, and the winner is the first to hit the bull's-eye. If player *A* throws first, find the probability that *A* wins.*

Answer. This problem is not strictly about the *geometric distribution*, but the analysis is identical: we have a two-outcome experiment repeated until the first success.

The probability that *A* wins on the first throw is $\frac{1}{5}$. The probability that *A* wins on the second throw is $\frac{4}{5} \cdot \frac{5}{6} \cdot \frac{1}{5}$ by the multiplication rule. In general, the probability that *A* wins on the n^{th} throw is

$$\Pr(X_1 = 0, Y_1 = 0, \dots, X_{n-1} = 0, Y_{n-1} = 0, X_n = 1) = \frac{1}{5} \left(\frac{4}{5}\right)^{n-1} \left(\frac{5}{6}\right)^{n-1} = \frac{1}{5} \left(\frac{2}{3}\right)^{n-1}.$$

Therefore, the total probability that A wins must be the sum as $n \rightarrow \infty$:

$$\Pr(A \text{ wins}) = \frac{1}{5} \sum_{n=1}^{\infty} \left(\frac{2}{3}\right)^{n-1} = \frac{1}{5} \cdot \frac{1}{1 - \frac{2}{3}} = \frac{3}{5}.$$

□

Q 7 (Hypergeometric Random Variable). *Three cards were drawn at random **without replacement** from an ordinary deck of 52 cards. What is the probability that none of the three cards is a heart?*

Answer. This is a classic problem pertaining to the *hypergeometric distribution*, as each draw is a binary outcome (i.e., the card is either a heart or not) from a population without replacement; thus, in this particular case, we have the random variable $X \sim H(N = 52, K = 13, n = 3)$, which pertains to the number of successes. Hence,

$$\Pr(X = 0) = \frac{\binom{13}{0} \binom{39}{3-0}}{\binom{52}{3}} = \frac{703}{1700}.$$

□

Q 8 (Partitioning). *A class has 4 male and 12 female students. They are randomly divided into 4 groups of 4 students each. What is the probability that each group has a female student?*

Answer. When an experiment involves *partitioning* discrete objects, we employ the *multinomial* method together with the *counting principle*. In this case, we have $|\Omega| = \binom{16}{4,4,4,4} = 63,063,000$ ways to partition the students into 4 groups of 4.

We are interested in partitions in which there is *at least* one female student in every group—there is only one way this can happen, namely by distributing 3 female students to each group (because otherwise one group would contain all four female students, making it impossible to include a male student). There are $\binom{12}{3,3,3,3} = 369,600$ ways to accomplish this, and the remaining four male students are distributed one to each group in $\binom{4}{1,1,1,1} = 24$ ways.

Therefore, by the counting principle, the number of favorable partitions is $369,600 \times 24 = 8,870,400$. Hence the probability is

$$\frac{8,870,400}{63,063,000} = \frac{64}{455}.$$

□

To conclude this section, we will look at an interesting problem that involves the law of total probability and recurrence.

Q 9. *Suppose there is a machine that can give either a correct or an incorrect answer at each trial. The machine gives a correct answer with probability 0.8 on the first trial. If the machine gives a correct answer in the previous trial, it will give a correct answer in the current trial with probability 0.6. If the machine gives an incorrect answer in the previous trial, it will give a correct answer in the current trial with probability 0.4. What is the probability that the machine will give a correct answer on the third trial?*

Answer. This problem is interesting because we can set up a recurrence relation via the law of total probability. Suppose we are on the i th trial and define an indicator random variable X that takes the value 1 if the machine outputs the correct answer. By the law of total probability, we have

$$\begin{aligned}\Pr(X_i = 1) &= \Pr(X_{i-1} = 0) \Pr(X_i = 1 \mid X_{i-1} = 0) + \Pr(X_{i-1} = 1) \Pr(X_i = 1 \mid X_{i-1} = 1) \\ &= 0.4 \Pr(X_{i-1} = 0) + 0.6 \Pr(X_{i-1} = 1).\end{aligned}$$

with the initial conditions $\Pr(X_1 = 0) = 0.2$ and $\Pr(X_1 = 1) = 0.8$. This is a general solution; therefore, one can easily compute for $i = 3$, which yields $\Pr(X_3 = 1) = 0.512$ by recursively computing the previous trials. \square

2 Problems with Expectation

An expectation, by definition, can be computed via $E[X] = \sum_x x \Pr(X = x)$. However, finding the unconditional probabilities can sometimes be elusive. In that scenario, we often cleverly define an auxiliary random variable and exploit the **linearity of expectation**. Other times, we employ **conditional expectation** to find its unconditional counterpart, because oftentimes computing conditional probabilities is much easier. We will look at some problems pertaining to these ideas.

Q 10. *A deck of n playing cards, which contains five jokers, is well shuffled. The cards are turned up one by one from the top until the third joker appears. What is the expected number of cards to be turned up?*

Answer. The key idea here is that these five jokers divide the deck into six piles: Pile 1 → Joker 1 → Pile 2 → ⋯ → Pile 5 → Joker 5 → Pile 6. A pile can consist of zero cards. This means that the third joker will always be in the middle and therefore the probability of finding the third joker from the top equals the probability of finding it from the bottom. If X is the number of cards that need to be turned up either from the top or the bottom until we get the third joker, then $\Pr(X = k) = \Pr(X = n - k + 1)$. Then, by the definition of expectation, we get

$$\begin{aligned}2E[X] &= \sum_k k \cdot \Pr(X = k) + \sum_k k \cdot \Pr(X = n - k + 1) \\ &= \sum_k (k + n - k + 1) \cdot \Pr(X = k) && (\text{let the second summand } k \leftarrow n - k + 1) \\ &= (n + 1) \sum_k \Pr(X = k) \\ &= n + 1 && (\text{by the normalization axiom of probability}).\end{aligned}$$

Thus, the expected value is $\frac{n+1}{2}$. \square

Q 11. *We have a complete graph with n vertices. On each edge of the graph, there is a light bulb that is initially in the OFF state. On each vertex, there is a switch. The switch is connected to all the light bulbs on the incident edges, and pressing that switch changes the status of these light bulbs (OFF → ON or vice versa). Suppose each switch is pressed once, independently, with probability $\frac{1}{2}$; what is the expected number of light bulbs in the ON state after the process is completed?*

Answer. Since the graph is complete, we can label each edge from 1 to $\binom{n}{2}$. Let Z be the number of light bulbs in the ON state after the process; then we can define an indicator random variable Z_i that takes the value 1 if the i th light bulb is ON—so, $Z = \sum_i Z_i$. Thus, $E[Z_i] = \Pr(Z_i = 1)$. Since each edge is between two vertices, say v_1 and v_2 , the bulb can be ON if and only if one of those vertices is pressed but **not both**. Hence,

$$\begin{aligned}\Pr(Z_i = 1) &= \Pr(v_1 \text{ is pressed}) \Pr(v_2 \text{ is not pressed}) + \Pr(v_1 \text{ is not pressed}) \Pr(v_2 \text{ is pressed}) \\ &= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \\ &= \frac{1}{2}.\end{aligned}$$

Thus, by linearity of expectation, the number of light bulbs turned ON after the process is $E[Z] = \frac{1}{2} \cdot \binom{n}{2} = \frac{n(n-1)}{4}$. \square

Q 12. 20 couples are invited to a party. They are asked to be seated at a long table with 20 seats on each side, with husbands sitting on one side and wives on the other. If the seating is done at random, what is the expected number of married couples that are seated face-to-face?

Answer. Let X be the random variable representing the number of couples sitting face-to-face. We define an indicator random variable X_i that takes the value 1 if the i th sitting position seats a married couple sitting face-to-face. Then we know $X = \sum_i X_i$ and $E[X_i] = \Pr(X_i = 1)$. To find the probability, suppose we fix an arbitrary husband at position i ; then the probability of choosing his wife is $\frac{1}{20}$. By linearity of expectation, we expect to see $20 \times \frac{1}{20} = 1$ married couple sitting face-to-face. \square

Q 13. 14 boys and 6 girls are arranged to sit randomly at a circular table. Suppose that all possible arrangements are equally likely to occur. Let N denote the number of times a girl is sitting next to two boys (i.e., one boy on her left and one boy on her right). Compute $E[N]$.

Answer. Again, we employ linearity of expectation by first defining an indicator random variable N_i denoting the i th girl sitting between two boys. To find $\Pr(N_i = 1)$, we fix the girl in the middle and randomly choose one boy to her left (the probability is $\frac{14}{14+6-1} = \frac{14}{19}$) and the other on her right (the probability is $\frac{13}{14+6-2} = \frac{13}{18}$). Thus, multiplying the two yields $\Pr(N_i = 1) = \frac{182}{342}$. Therefore, $E[N] = 6 \times \frac{182}{342} = \frac{182}{57}$, which is roughly 3 girls. \square

Q 14. There are 60 men, 40 women, and 50 tables seating two. If they are seated at random, what is the expected number of tables with opposite genders?

Proof. Again, we define an indicator random variable X_i that equals 1 if the i th table has one man and one woman. Note that X_i can be interpreted as a *hypergeometric random variable* obtained by drawing exactly one man in two samples without replacement from a population of 60 men and 40 women; thus, $\Pr(X_i = 1) = \frac{\binom{60}{1} \cdot \binom{40}{1}}{\binom{100}{2}} = \frac{16}{33}$. Then, by linearity of expectation, we get the expected number of tables with different genders to be $\frac{800}{33}$, which is roughly 24. \square

Q 15. A permutation $\pi : [1, n] \rightarrow [1, n]$ can be represented as a set of **cycles** as follows. Let there be one vertex for each i , $i = 1, 2, \dots, n$. If the permutation maps the number i to the number $\pi(i)$, then a **directed arc** is drawn from vertex i to vertex $\pi(i)$ (you

may use **Cauchy notation** to aid the process of determining cycles). This leads to a graph that is a set of **disjoint cycles**. Notice that the cycles could be **self-loops**, i.e., $\pi(i) = i$. What is the expected number of cycles in a random permutation of n numbers?

Solution 1. Let X be the random variable representing the number of cycles. This is difficult to compute head-on, but one can cleverly define auxiliary random variables and exploit the linearity of expectation. Notice that each vertex v_i can belong to a cycle of **length** k , where k is between 1 (i.e., a self-loop) and n (i.e., one big cycle). Define the random variable $X_{v_i} = \frac{1}{k}$; one can immediately confirm that $\sum_i X_{v_i} = X$. It remains to determine $\Pr(X_{v_i} = \frac{1}{k})$, which is equivalent to $\Pr(v_i \text{ is in a cycle of length } k)$. Thus, by the definition of expectation,

$$E[X_{v_i}] = \sum_k \frac{1}{k} \Pr(v_i \text{ is in a cycle of length } k).$$

From here we continue with a simple analysis of how the arc is drawn to form a cycle: i) for $k = 1$ we require $\pi(i) = i$, so $\Pr(v_i \text{ is in a cycle of length } 1) = \frac{1}{n}$; and ii) for $k > 1$ we fix an arbitrary i and ensure $\pi_1 \neq i$ (with probability $\frac{n-1}{n}$), $\pi_2 \neq i$ (with probability $\frac{n-2}{n-1}$), \dots , $\pi_{k-1} \neq i$ (with probability $\frac{n-k+1}{n-k+2}$), and finally $\pi_k = i$ (with probability $\frac{1}{n-k+1}$); thus $\Pr(v_i \text{ is in a cycle of length } k \text{ for } k > 1) = \frac{1}{n}$. Therefore,

$$E[X] = n \cdot \frac{1}{n} \left(\frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n} \right) = H(n),$$

where $H(n)$ is the n th harmonic number. \square

Solution 2. Alternatively, we can employ the *law of total expectation*, but we must partition the sample space cleverly. We define the random variable

$$Y_n = \# \text{ cycles for an } n\text{-permutation.}$$

For the base case, only one permutation is possible; thus,

$$E[Y_1] = 1.$$

For an arbitrary vertex i , there are two possibilities after the permutation: i) $\pi(i) = i$ or ii) $\pi(i) \neq i$ —this partitions the sample space nicely. Thus, we have

$$E[Y_n] = E[Y_n \mid \pi(i) = i] \cdot \Pr(\pi(i) = i) + E[Y_n \mid \pi(i) \neq i] \cdot \Pr(\pi(i) \neq i).$$

As seen in the first solution, $\Pr(\pi(i) = i) = \frac{1}{n}$ and $\Pr(\pi(i) \neq i) = 1 - \frac{1}{n} = \frac{n-1}{n}$. When $\pi(i) = i$, we know there is one cycle followed by the permutation of $n - 1$ numbers; so $E[Y_n \mid \pi(i) = i] = 1 + E[Y_{n-1}]$. On the other hand, we have fixed a value $\pi(i)$ such that $\pi(i) \neq i$, so we are left with the remaining permutation of $n - 1$ numbers; thus, $E[Y_n \mid \pi(i) \neq i] = E[Y_{n-1}]$. Putting it all together,

$$\begin{aligned} E[Y_n] &= (1 + E[Y_{n-1}]) \cdot \frac{1}{n} + E[Y_{n-1}] \cdot \frac{n-1}{n} \\ &= E[Y_{n-1}] + \frac{1}{n}. \end{aligned}$$

From the above recursion, it is obvious by *mathematical induction* that

$$E[Y_n] = H(n).$$

\square

Q 16. A lost tourist arrives at a point with 4 roads. There are no signs on the roads. The first road brings them back to the same point after two hours' walk. The second road gets them to the city after three hours' walk. The third road leads them back to the same point after four hours' walk. Lastly, the fourth road gets them to the city after five hours' walk. Assuming the tourist chooses a road with equal probability at all times (that is, a road may be chosen again and again), what is the mean time until the tourist arrives at the city?

Answer. This is the kind of problem for which it is immediately obvious to utilize the law of total expectation. Let T represent the time it takes for the tourist to arrive at the city. Then,

$$\begin{aligned} E[T] &= E[T \mid \text{first road was chosen}] \cdot \Pr(\text{first road was chosen}) \\ &\quad + E[T \mid \text{second road was chosen}] \cdot \Pr(\text{second road was chosen}) \\ &\quad + E[T \mid \text{third road was chosen}] \cdot \Pr(\text{third road was chosen}) \\ &\quad + E[T \mid \text{fourth road was chosen}] \cdot \Pr(\text{fourth road was chosen}). \end{aligned}$$

Since choosing a road is equally likely at all times, the probability of choosing any road is $\frac{1}{4}$. It follows that $E[T \mid \text{first road was chosen}] = E[T] + 2$ (as two hours are wasted getting back to the same point), $E[T \mid \text{second road was chosen}] = 3$, $E[T \mid \text{third road was chosen}] = E[T] + 4$, and $E[T \mid \text{fourth road was chosen}] = 5$. Therefore,

$$\begin{aligned} E[T] &= \frac{1}{4}(E[T] + 2 + 3 + E[T] + 4 + 5) \\ \Rightarrow E[T] &= 7 \text{ hours.} \end{aligned}$$

□

Q 17. Alice and Bob are betting on the patterns observed in the output sequence produced by rolling a fair six-sided die 1000 times. For each time the pattern "66" occurs (i.e., a 6 is followed immediately by another 6), Alice wins \$1 from Bob. For each time the pattern "65" occurs (i.e., a 6 is followed immediately by a 5), Alice loses \$1 to Bob. Let W denote Alice's winnings. What is $E[W]$?

Remark. For an illustration, suppose we roll the die 15 times instead, and the output sequence is 142566621656243. Then Alice wins \$1 twice and loses \$1 once, so that $W = 1$.

Answer. Let I_i^{66} be the indicator random variable that equals 1 exactly when rolls i and $i+1$ show the pattern "66" (that is, $(X_i, X_{i+1}) = (6, 6)$) and equals 0 otherwise. Likewise, let I_i^{65} be the indicator that equals 1 when rolls i and $i+1$ show the pattern "65" (so $(X_i, X_{i+1}) = (6, 5)$) and 0 otherwise. Alice's net winning can then be written as

$$W = \sum_{i=1}^{999} (I_i^{66} - I_i^{65}).$$

Because each ordered pair of consecutive rolls is independent and uniformly distributed over the 36 ordered pairs of faces,

$$\Pr((X_i, X_{i+1}) = (6, 6)) = \Pr((X_i, X_{i+1}) = (6, 5)) = \frac{1}{36}, \quad i = 1, \dots, 999.$$

Hence, by linearity of expectation,

$$E[W] = \sum_{i=1}^{999} (E[I_i^{66}] - E[I_i^{65}]) = 999 \left(\frac{1}{36} - \frac{1}{36} \right) = 0.$$

□

Q 18. Suppose Alice and Bob play a different game with a fair six-sided die: they each take turns and count the number of rolls needed to observe their desired pattern—Alice chooses “66” (i.e., a 6 immediately followed by another 6), and Bob chooses “65” (i.e., a 6 immediately followed by a 5). Whoever has the larger number of rolls loses the game. Who has an expected advantage?

Answer. Let X denote the number of rolls needed to observe the sequence “66,” and let Y denote the number of rolls needed to observe the sequence “65.” Define three events:

$$\mathcal{A} = \text{the first roll is a 6}, \quad \mathcal{B} = \text{the second roll is a 6}, \quad \mathcal{C} = \text{the second roll is a 5}.$$

Alice. We first analyze $E[X]$:

$$E[X] = E[X | \mathcal{A}] \Pr(\mathcal{A}) + E[X | \overline{\mathcal{A}}] \Pr(\overline{\mathcal{A}}).$$

Because $E[X | \overline{\mathcal{A}}] = E[X] + 1$,

$$E[X | \mathcal{A}] = E[X | \mathcal{A} \cap \mathcal{B}] \Pr(\mathcal{B} | \mathcal{A}) + E[X | \mathcal{A} \cap \overline{\mathcal{B}}] \Pr(\overline{\mathcal{B}} | \mathcal{A}),$$

with $E[X | \mathcal{A} \cap \mathcal{B}] = 2$ and $E[X | \mathcal{A} \cap \overline{\mathcal{B}}] = E[X] + 2$. Hence

$$E[X] = \frac{1}{6} \left[\frac{1}{6}(2) + \frac{5}{6}(E[X] + 2) \right] + \frac{5}{6}(E[X] + 1),$$

which yields $E[X] = 42$ rolls.

Bob. Next we analyze $E[Y]$:

$$E[Y] = E[Y | \mathcal{A}] \Pr(\mathcal{A}) + E[Y | \overline{\mathcal{A}}] \Pr(\overline{\mathcal{A}}).$$

Again $E[Y | \overline{\mathcal{A}}] = E[Y] + 1$. For the first term,

$$\begin{aligned} E[Y | \mathcal{A}] &= E[Y | \mathcal{A} \cap \mathcal{C}] \Pr(\mathcal{C} | \mathcal{A}) \\ &\quad + E[Y | \mathcal{A} \cap \mathcal{B}] \Pr(\mathcal{B} | \mathcal{A}) \\ &\quad + E[Y | \mathcal{A} \cap \overline{\mathcal{B}} \cap \overline{\mathcal{C}}] \Pr(\overline{\mathcal{B}} \cap \overline{\mathcal{C}} | \mathcal{A}), \end{aligned}$$

with $E[Y | \mathcal{A} \cap \mathcal{C}] = 2$, $E[Y | \mathcal{A} \cap \mathcal{B}] = E[Y | \mathcal{A}] + 1$, and $E[Y | \mathcal{A} \cap \overline{\mathcal{B}} \cap \overline{\mathcal{C}}] = E[Y] + 2$. Therefore,

$$\begin{aligned} E[Y | \mathcal{A}] &= \frac{1}{6}(2) + \frac{1}{6}(E[Y | \mathcal{A}] + 1) + \frac{4}{6}(E[Y] + 2) \\ &\Rightarrow \\ E[Y | \mathcal{A}] &= \frac{4}{5}E[Y] + \frac{11}{5}. \end{aligned}$$

Substituting these conditional expectations gives

$$E[Y] = \frac{1}{6} \left(\frac{4}{5} E[Y] + \frac{11}{5} \right) + \frac{5}{6} \left(E[Y] + 1 \right),$$

and solving yields $E[Y] = 36$ rolls.

Conclusion. Since $E[X] > E[Y]$, Bob's expected number of rolls is smaller, so Bob has the expected advantage. The reason for this discrepancy is how they both wait for their respective pattern, as evident from the computation above. \square

We complete this section with two problems relating to variance.

Q 19. A **fixed point** of a permutation $\pi : [1, n] \rightarrow [1, n]$ is a value i such that $\pi(i) = i$. Find the variance of the number of fixed points of a permutation chosen uniformly at random from all permutations of $[1, n]$.

Answer. Define an indicator random variable X_i that takes the value 1 if $\pi(i) = i$ and 0 otherwise. For any i , $\Pr(X_i = 1) = p = \frac{1}{n}$. Hence, for an n -permutation, the expected number of fixed points is

$$E[X] = \sum_i E[X_i] = \sum_i p = n \cdot \frac{1}{n} = 1.$$

Recall the identity

$$\text{Var}\left[\sum_i X_i\right] = \sum_i \text{Var}[X_i] + 2 \sum_{j < i} \text{Cov}(X_i, X_j),$$

where $\text{Cov}(X_i, X_j) = E[X_i X_j] - E[X_i] E[X_j]$ and the second sum ranges over all **unordered pairs** (i, j) .

Since each X_i is Bernoulli,

$$\text{Var}[X_i] = p(1-p) = \frac{n-1}{n^2} \implies \sum_i \text{Var}[X_i] = \frac{n-1}{n}.$$

Next, from the definition of expectation we have

$$\begin{aligned} E[X_i X_j] &= (0)(0) \Pr(X_i = 0 \cap X_j = 0) + (0)(1) \Pr(X_i = 0 \cap X_j = 1) \\ &\quad + (1)(0) \Pr(X_i = 1 \cap X_j = 0) + (1)(1) \Pr(X_i = 1 \cap X_j = 1) \\ &= \Pr(X_i = 1 \cap X_j = 1) \\ &= \Pr(X_i = 1) \Pr(X_j = 1 \mid X_i = 1) \\ &= \frac{1}{n} \cdot \frac{1}{n-1} = \frac{1}{n(n-1)}. \end{aligned}$$

while $E[X_i]E[X_j] = \frac{1}{n^2}$. There are $\binom{n}{2} = \frac{n(n-1)}{2}$ unordered pairs, so

$$\begin{aligned} \text{Var}[X] &= \frac{n-1}{n} + 2 \frac{n(n-1)}{2} \left(\frac{1}{n(n-1)} - \frac{1}{n^2} \right) \\ &= \frac{n-1}{n} + 1 - \frac{n-1}{n} \\ &= 1. \end{aligned}$$

\square

Q 20. Suppose that n gentlemen deposit their hats in a coatroom, and later the hats are returned to them randomly. Let X be the number of gentlemen who get their own hats back. Compute $\text{Var}[X]$.

Answer. This problem is equivalent to finding the number of fixed points for a uniformly random permutation, as shown above (i.e., the fixed-point abstraction applies to counting the gentlemen who receive their own hats). Using the same analysis, we obtain $\text{Var}[X] = 1$. \square

3 Classical Problems in Discrete Probability

This section covers important problems in discrete probability that are often used as abstractions for many real-world applications.

Problem 1 (The Problem of Points). *Two players, A and B , agree to play a series of points in a game, each contributing equally to the pot. Player A wins each point with probability p , whereas player B wins each point with probability $1 - p$. The first player to win N points wins the game. Suppose the game is interrupted when player A has won $a < N$ points and player B has won $b < N$ points. How should they divide the stakes fairly if they no longer wish to resume playing?*

Solution. It is reasonable to assume that the fairest division of the stakes is proportional to the probability that each player would ultimately win if play were continued from the current score. Player A still needs $n = N - a$ points to win, whereas player B needs $m = N - b$ points. Consequently, **at most** $n + m - 1$ additional points must be played before a winner is certain. (The match could end earlier, but we conceptually extend play to the $(n+m-1)$ -st point; that is, we imagine the players continue even after the winner is determined, until a total of $n + m - 1$ points have been played.) There are 2^{n+m-1} possible sequences of outcomes for these points. The key is to count the sequences in which A wins—equivalently, to compute the probability that A wins **at least** n of the next $n + m - 1$ points. This probability is given by the *binomial distribution*:

$$p_A = \sum_{k=n}^{n+m-1} \binom{n+m-1}{k} p^k (1-p)^{n+m-1-k}.$$

Similarly,

$$p_B = \sum_{k=m}^{n+m-1} \binom{n+m-1}{k} p^k (1-p)^{n+m-1-k}.$$

If the total stake in the pot is sum , then player A receives $p_A \text{sum}$, while player B receives $p_B \text{sum}$. \square

Problem 2 (The Branching Process Problem). *Suppose we have an infected machine \mathcal{M} . Because \mathcal{M} is infected, it sends out a virus program to n new machines, each of which becomes infected independently with probability p . If a machine is infected, it in turn sends out a virus program to n new machines, and so on—i.e., at generation 0 we have 1 infected machine (namely \mathcal{M}); at generation 1 there can be at most n infected machines, and this probabilistic growth continues in later generations. If Y is the random variable denoting the total number of infected machines far into the future (irrespective of any particular generation), compute $E[Y]$.*

Solution. Let $Y_j = \#$ infected machines at generation j ; clearly $Y_0 = 1$. The number of infected machines at generation j depends on the number at generation $j - 1$, with each infected machine spreading according to a *binomial distribution*. Assume that at generation $j - 1$ there are y_{j-1} infected machines. Denote by $Z_{j|k}$ the number of machines infected at generation j by the k th infected machine of generation $j - 1$. Then

$$\begin{aligned} E[Y_j | Y_{j-1} = y_{j-1}] &= E\left[\sum_{k=1}^{y_{j-1}} Z_{j|k}\right] \\ &= \sum_{k=1}^{y_{j-1}} E[Z_{j|k}] && (\text{linearity of expectation}) \\ &= \sum_{k=1}^{y_{j-1}} np && (\text{each is } \text{Bin}(n, p)) \\ &= y_{j-1} np. \end{aligned}$$

By the law of total expectation,

$$\begin{aligned} E[Y_j] &= E[E[Y_j | Y_{j-1}]] \\ &= E[Y_{j-1} np] = np E[Y_{j-1}] \\ &= (np)^j && (\text{by induction}). \end{aligned}$$

Therefore,

$$E[Y] = \sum_{j=0}^{\infty} E[Y_j] = \sum_{j=0}^{\infty} (np)^j.$$

Hence, if $np \geq 1$ the expectation diverges, whereas for $np < 1$ it is

$$E[Y] = \frac{1}{1 - np}.$$

□

Problem 3 (The Coupon Collector's Problem). *Suppose each box of cereal contains one of n different types of coupons (assume the coupon contained in each box is chosen independently and uniformly at random from the n possibilities). A collector wishes to obtain all of them; how many cereal boxes are they expected to buy to collect at least one of every type of coupon?*

Solution. We denote by X the random variable representing the number of boxes to buy to complete the collection. Let

$$X_i = \# \text{ boxes to buy to obtain the } i\text{th new coupon.}$$

Then $X = \sum_i X_i$, and each X_i is a *geometric random variable* with parameter p_i . Therefore $E[X] = \sum_i \frac{1}{p_i}$. The probability of getting a first (arbitrarily new) type is $p_1 = 1$; the next new type appears with probability $1 - \frac{1}{n} = \frac{n-1}{n}$; the third with probability $1 - \frac{2}{n} = \frac{n-2}{n}$; hence, in general,

$$p_i = \frac{n - i + 1}{n}.$$

Consequently,

$$E[X] = \sum_{i=1}^n \frac{n}{n-i+1} = nH(n).$$

With a more sophisticated analysis, we can show that the random variable is indeed **closely concentrated around its mean** ($nH(n) \approx n \ln n$). To do this, we invoke the *balls-and-bins* analogy, viewing coupons as bins and cereal boxes as balls.

We consider the experiment in which each bin receives a Poisson number of balls with mean

$$\mu = \ln n + c,$$

so the expected total number of balls is

$$m = n \ln n + nc$$

(by linearity of expectation). This setup is known as the *Poisson approximation*, where the occupancies of the bins are assumed independent. Define

$$\mathcal{E} = \text{no bin is empty}, \quad X = \text{number of balls thrown}, \quad r = \sqrt{2m \ln m}.$$

The probability that a specific bin is empty is

$$\Pr(\text{bin empty}) = \frac{\mu^0 e^{-\mu}}{0!} = e^{-(\ln n + c)} = \frac{e^{-c}}{n}.$$

Under the Poisson approximation the bins are independent, and therefore

$$\lim_{n \rightarrow \infty} \Pr(\mathcal{E}) = \left(1 - \frac{e^{-c}}{n}\right)^n = e^{-e^{-c}}.$$

Thus, by the law of total probability,

$$\begin{aligned} \lim_{n \rightarrow \infty} \Pr(\mathcal{E}) &= \Pr(\mathcal{E} \mid |X - m| \leq r) \Pr(|X - m| \leq r) \\ &\quad + \Pr(\mathcal{E} \mid |X - m| > r) \Pr(|X - m| > r) \\ &= e^{-e^{-c}} \quad (\text{as shown above}). \end{aligned}$$

Step 1. We aim to show that $\Pr(|X - m| > r) = o(1)$.

We employ the *Chernoff bound* for a Poisson random variable Z :

$$\Pr(Z \geq z) \leq \frac{e^{-\mu}(e\mu)^z}{z^z}.$$

Right tail:

$$\begin{aligned} \Pr(X > m + r) &\leq e^r m^{m+r} / (m + r)^{m+r} \\ &= e^{r-(m+r) \ln(1+r/m)} \\ &\leq e^{r-(m+r)(r/m - r^2/(2m^2))} \quad (\text{using } \ln(1+z) \geq z - \frac{z^2}{2}, z \geq 0) \\ &= e^{\ln m - 2 \ln m + o(\ln m)} = o(1) \quad (\text{substituting } r). \end{aligned}$$

Left tail:

$$\begin{aligned}
\Pr(X < m - r) &\leq e^{-r} m^{m-r} / (m-r)^{m-r} \\
&= e^{-r-(m-r)\ln(1-r/m)} \\
&\leq e^{-r-(m-r)\left(-r/m-\frac{3}{2}r^2/(2m^2)\right)} \quad (\text{using } \ln(1+z) \geq z - \frac{3}{4}z^2, z \geq -0.47) \\
&= e^{1.5\ln m - 2\ln m - o(\ln m)} = o(1) \quad (\text{substituting } r).
\end{aligned}$$

Step 2. We aim to prove that

$$|\Pr(\mathcal{E} \mid |X - m| \leq r) - \Pr(\mathcal{E} \mid X = m)| = o(1).$$

Because $\Pr(\mathcal{E} \mid X = k)$ is **monotonically increasing** in k (intuitively, the more balls that are thrown, the more likely it is that all bins are non-empty),

$$\Pr(\mathcal{E} \mid X = m - r) \leq \Pr(\mathcal{E} \mid |X - m| \leq r) \leq \Pr(\mathcal{E} \mid X = m + r).$$

Hence

$$|\Pr(\mathcal{E} \mid |X - m| \leq r) - \Pr(\mathcal{E} \mid X = m)| \leq \Pr(\mathcal{E} \mid X = m + r) - \Pr(\mathcal{E} \mid X = m - r).$$

The term on the *right-hand side* equals the difference between the probability that all bins receive at least one ball when $m + r$ balls are thrown and the probability that this happens when $m - r$ balls are thrown. The difference can be interpreted as the following experiment: we first throw $m - r$ balls, after which **at least** one bin is still empty; we then throw an additional $2r$ balls (so that a total of $m + r$ balls have been thrown), and now every bin is non-empty. The probability that these extra $2r$ balls succeed in covering all previously empty bins is **at most** $2r/n = o(1)$ by the *union bound*. Hence we obtain the desired result.

Combining Steps 1 and 2. Inserting these results into the law of total probability,

$$\begin{aligned}
\Pr(\mathcal{E}) &= \Pr(\mathcal{E} \mid |X - m| \leq r)(1 - o(1)) + o(1) \\
&= \Pr(\mathcal{E} \mid X = m)(1 - o(1)) + o(1).
\end{aligned}$$

Therefore,

$$\lim_{n \rightarrow \infty} \Pr(\mathcal{E}) = \lim_{n \rightarrow \infty} \Pr(\mathcal{E} \mid X = m).$$

This shows that, for sufficiently large n , there is effectively no difference between observing “no empty bins” under the Poisson approximation and when exactly m balls are thrown at random (as in the original coupon-collector model). Consequently,

$$\Pr(|X - n \ln n| > cn) = e^{-e^{-c}}.$$

For example, 98% of the time the number of cereal boxes needed to collect all n coupon types lies between $n \ln n - 4n$ and $n \ln n + 4n$. \square

Problem 4 (Birthday Problem). *In a distant galaxy, suppose there is an alien civilization that counts n days in a year. If there is a group of m aliens, what is the probability that **at least** two aliens share the same birthday?*

Note. We assume $m \leq n$; otherwise, the probability is obviously 1 by the *pigeonhole principle*.

Solution. We use the *balls-and-bins* abstraction: throw m balls into n bins (with $m \leq n$) and ask for the probability that at least one bin receives more than one ball. Clearly,

$$\Pr(\text{at least one bin with more than one ball}) = 1 - \Pr(\text{each bin has at most one ball}).$$

Suppose we throw the first ball; it can land in any arbitrary bin. The second ball must avoid that bin, so it lands in a different bin with probability $1 - \frac{1}{n} = \frac{n-1}{n}$. The third ball must avoid the first two bins, and so on. Hence

$$\begin{aligned}\Pr(\text{each bin has at most one ball}) &= \prod_{k=1}^m \left(1 - \frac{k-1}{n}\right) \\ &\approx \prod_{k=1}^m e^{-(k-1)/n} = e^{-\frac{(m-1)m}{2n}} \\ &\approx e^{-\frac{m^2}{2n}}.\end{aligned}$$

Therefore,

$$\Pr(\text{at least one bin with more than one ball}) \approx 1 - e^{-\frac{m^2}{2n}}.$$

When $n = 365$ and $m \leq n$, the number of people m required so that the probability of at least two sharing a birthday is one-half satisfies

$$1 - e^{-\frac{m^2}{2 \cdot 365}} = \frac{1}{2} \implies m \approx 23 \text{ people.}$$

This seemingly counter-intuitive result is known as the **birthday paradox** (though it is simply a consequence of the underlying probability calculation). \square

Problem 5 (Gambler's Ruin Problem). *A determined gambler starts with a wealth of i dollars. They then make a series of independent bets, each with a winning probability q (and a losing probability $p = 1 - q$). In each bet, a win increases their wealth by \$1, while a loss decreases it by \$1. The gambler's goal is to reach a total of $\$N$ before going bankrupt (i.e., their wealth drops to \$0—the gambler is said to be **ruined**).*

Solution. This problem can be viewed and studied under a mathematical process called *random walk*. We denote q_i the probability the gambler wins starting with i dollars. Hence,

- $q_0 = 0$ is the state where the gambler is ruined, no chance of winning.
- In the event the gambler wins a bet starting with i dollars with probability q , they then *promoted* to the next state with $i + 1$ dollars.
- In the event the gambler loses a bet starting with i dollars with probability $p = 1 - q$, they then *demoted* to the previous state with $i - 1$ dollars.
- The gambler wins the independent series of bet at the N -th state.

From these definitions, we have the following *recurrence relation*

$$q_i = (q)q_{i+1} + (p)q_{i-1} \tag{1}$$

and since $p + q = 1$, we also have another condition

$$q_i = (p)q_i + (q)q_i \quad (2)$$

From (??) and (??) yield

$$q_{i+1} - q_i = \frac{p}{q}(q_i - q_{i-1}) \quad (3)$$

We can use mathematical induction to show that this expression is equivalent to

$$q_{i+1} - q_i = \left(\frac{p}{q}\right)^i q_1 \quad (4)$$

We start with the base case of $i = 0$, clearly $q_1 - (0) = \left(\frac{p}{q}\right)^0 q_1$ (because $q_0 = 0$). Suppose it is true for all k integers. We compute for $k + 1$: by definition we know that $q_k = (q)q_{k+1} + (p)q_{k-1}$ and we know that $q_k - q_{k-1} = \left(\frac{p}{q}\right)^k q_1$. Thus $q_k = (q)q_{k+1} + p(q_k - \left(\frac{p}{q}\right)^k q_1) \Rightarrow p\left(\frac{p}{q}\right)^k q_1 = (q)q_{k+1} + (p)q_k - q_k \Rightarrow \frac{p}{q}\left(\frac{p}{q}\right)^k q_1 = q_{k+1} + \left(\frac{p-1}{q}\right)q_k \Rightarrow \left(\frac{p}{q}\right)^{k+1} q_1 = q_{k+1} - q_k$ as expected.

We use this identity, together with the telescoping series to get the following

$$q_{i+1} - q_1 = \sum_{k=1}^i q_{k+1} - q_k = q_1 \sum_{k=1}^i \left(\frac{p}{q}\right)^k \quad (5)$$

and to put it succinctly

$$q_{i+1} = q_1 \sum_{k=0}^i \left(\frac{p}{q}\right)^k \quad (6)$$

Employing the sum of a finite geometric series $\sum_{n=0}^i a^n = \frac{1-a^{i+1}}{1-a}$ for any a ($\neq 1$) and integer $i \geq 1$, equation (??) can be written as

$$q_{i+1} = \begin{cases} q_1 \frac{1-\left(\frac{p}{q}\right)^{i+1}}{1-\frac{p}{q}} & \text{if } p \neq q \\ q_1(i+1) & \text{if } p = q = 0.5 \end{cases} \quad (7)$$

The other end of the boundary is when the gambler achieves their goal, $q_N = 1$, and together with (??) we can solve for q_1

$$q_N = \begin{cases} q_1 \frac{1-\left(\frac{p}{q}\right)^N}{1-\frac{p}{q}} = 1 & \text{if } p \neq q \\ q_1(N) = 1 & \text{if } p = q = 0.5 \end{cases} \quad (8)$$

$$q_1 = \begin{cases} \frac{1-\frac{p}{q}}{1-\left(\frac{p}{q}\right)^N} & \text{if } p \neq q \\ \frac{1}{N} & \text{if } p = q = 0.5 \end{cases} \quad (9)$$

Substitute the value of q_1 above appropriately to equation (??) concludes the following:

$$q_i = \begin{cases} \frac{1-\left(\frac{p}{q}\right)^i}{1-\left(\frac{p}{q}\right)^N} & \text{if } p \neq q \\ \frac{i}{N} & \text{if } p = q = 0.5 \end{cases}$$

□

To give a concrete example, suppose two players, A and B , play a series of bets determined by flips of a coin that shows HEAD with probability $\frac{2}{3}$ and TAIL with probability $\frac{1}{3}$. Player A begins with 100 tokens, while player B begins with 3 tokens. The players agree that after each flip the winner receives one token from the loser, and the game ends when one player's tokens are completely depleted. If player A always bets on TAIL, what is the probability that A wins the game?

This scenario fits the classic *Gambler's Ruin* model: a victory occurs when either party reaches the absorbing state $N = 100 + 3 = 103$ tokens (reducing the other to zero). We wish to compute the winning probability for player A , who starts with 100 tokens.

At first glance, one might think that winning 3 tokens is considerably easier than winning 100 tokens; however, as the calculation shows, the per-bet bias dominates the initial-capital advantage. Using the standard Gambler's Ruin formula, player A 's winning probability is

$$q_{100} = \frac{1 - \left(\frac{2/3}{1/3}\right)^{100}}{1 - \left(\frac{2/3}{1/3}\right)^{103}} = \frac{1 - 2^{100}}{1 - 2^{103}} \approx 0.125.$$

Consequently, player B wins with probability $q_3 \approx 1 - 0.125 = 0.875$. The discrepancy is significant in this case, and B enjoys the victory in most runs of the game.