

Waiting Time Until Success

Jyotirmoy Sarkar

*Department of Mathematical Sciences
Indiana University Indianapolis, IN 46202, USA*

Received: 30 January 2026; Revised: 04 August 2026; Accepted: 19 August 2026

Dedication

This research is respectfully dedicated to honor my guruji, mamababu, and guide, Professor Bikas Kumar Sinha, whose association, blessings, and collaboration can reduce the waiting time until success for any student of probability and statistics.

Abstract

Waiting time until success is a fundamental concept in probability, with the geometric distribution providing the classic example. We investigate waiting times associated with more elaborate notions of success, including prescribed patterns of outcomes, events in multinomial experiments, features of finite Markov chains, and return times of random walks on graphs. A common theme throughout the paper is the use of renewal-type recursions derived from one-step transition probabilities. They generate the probability mass functions of the waiting times through matrix iterations. The unified methodology produces explicit distributions for a variety of stochastic processes and reveals connections between pattern-occurrence problems, Markov chains, and random walks.

Key words: Probability mass function; Lack of memory; Renewal time; Recursive relation; Matrix inversion; Stationary distribution.

AMS Subject Classifications: 97K60, 60J10

1. Introduction

Suppose that we toss a biased coin with a known $P(\text{Heads}) = p \in (0, 1)$. Then the number X of tosses until the first head appears is said to have the $\text{geometric}(p)$ distribution. Its probability mass function (PMF) is given by

$$p_x = P\{X = x\} = q^{x-1}p, \quad \text{for } x = 1, 2, 3, \dots, \quad (1)$$

where $q = 1 - p = P(\text{Tails})$. The probability masses $p_1 = p, p_2 = qp, p_3 = q^2p, \dots$ of X form a geometric series. Hence, X is called a geometric random variable (RV).

Most textbooks (see, for example, Devore and Berk (2007), Dudewicz and Mishra (1988), Hoel *et al.* (1971), Rohatgi (2003)) develop the properties of a geometric RV starting from the PMF (1). We list a few such properties below.

(G1) Total probability: The sequence of probability masses is geometric, and adds to

$$p + qp + q^2p + \dots = p[1 + q + q^2 + \dots] = p(1 - q)^{-1} = 1.$$

No mass escapes to infinity; the process stops in finite time with probability one.

(G2) Probability generating function (PGF): For $0 < t < q^{-1}$,

$$P_X(t) = E[t^X] = pt + qpt^2 + q^2pt^3 + \dots = pt [1 + (qt) + (qt)^2 + \dots] = pt(1 - qt)^{-1}.$$

(G3) First moment: The expected waiting time until the first head appears is

$$\begin{aligned} E[X] &= p + 2qp + 3q^2p + \dots = p \frac{d}{dq} \{q + q^2 + q^3 + \dots\} \\ &= p \frac{d}{dq} \{1 + q + q^2 + q^3 + \dots\} = p \frac{d}{dq} (1 - q)^{-1} = p(1 - q)^{-2} = p^{-1}. \end{aligned}$$

(G4) Moment generating function (MGF): For $0 < t < -\ln q$,

$$\begin{aligned} M_X(t) = E[e^{tX}] &= pe^t + qpe^{2t} + q^2pe^{3t} + \dots \\ &= pe^t [1 + (qe^t) + (qe^t)^2 + \dots] = pe^t(1 - qe^t)^{-1}. \end{aligned}$$

(G5) Lack of memory property: For $x \geq 1$, we have $P\{X > x\} = q^x p + q^{x+1}p + \dots = q^x$. Consequently, for $x, y \geq 1$, we have

$$P\{X > x + y | X > x\} = \frac{P\{X > x + y\}}{P\{X > x\}} = \frac{q^{x+y}}{q^x} = q^y = P\{X > y\}.$$

In the above, properties G1–G5 are proved by evaluating a geometric series. Below, we derive these properties starting from the **renewal property** given by

$$X = \begin{cases} 1 & \text{with probability } p \\ 1 + \tilde{X} & \text{with probability } q, \end{cases} \quad (2)$$

where $\tilde{X}$ has the same probability distribution as X . In other words, after the first toss, either a success (a head) has occurred and the process stops, or the process has *renewed* itself in the sense that the number of additional tosses (after the first toss that did not yield a success) until a success occurs has the same distribution as the original RV X itself. Equivalently, (2) can be written as

$$p_x = \begin{cases} p & \text{if } x = 1 \\ qp_{x-1} & \text{if } x > 1. \end{cases} \quad (3)$$

Writing all probability masses as a vector $\mathbf{p}_X = (p_1, p_2, \dots)$, (3) implies $\mathbf{p}_X = (p, q \mathbf{p}_X)$. Thereafter, by mathematical induction on x , one can prove that (3) implies (1). Below, we show that properties G1–G5 follow from (3), without requiring either evaluation of a geometric series or a differentiation technique. We prove these relatively easy results as a precursor to the extensions we have in mind.

(G1) Total probability: Writing $T = \sum_{x=1}^{\infty} p_x$ and substituting $y = x - 1$, we note that

$$T = p + \sum_{x=2}^{\infty} q p_{x-1} = p + q \sum_{y=1}^{\infty} p_y = p + q T.$$

Solve this relation to get $T = p/(1 - q) = 1$.

(G2) PGF: Writing $P_X(t) = \sum_{x=1}^{\infty} p_x t^x$ and substituting $y = x - 1$, we have

$$P_X(t) = pt + \sum_{x=2}^{\infty} q p_{x-1} t^x = pt + qt \sum_{y=1}^{\infty} p_y t^y = pt + qt P_X(t),$$

which we solve to get $P_X(t) = pt(1 - qt)^{-1}$, for $0 < t < q^{-1}$.

(G3) First moment (or mean): Taking expectations on the two sides of (2), we have

$$E[X] = p + q(1 + E[\tilde{X}]) = 1 + q E[X],$$

solving which we get $E[X] = p^{-1}$.

(G4) MGF: Writing $M_X(t) = \sum_{x=1}^{\infty} p_x e^{tx}$ and substituting $y = x - 1$, we have

$$M_X(t) = p e^t + \sum_{x=2}^{\infty} q p_{x-1} e^{tx} = p e^t + q e^t \sum_{y=1}^{\infty} p_y e^{ty} = p e^t + q e^t M_X(t).$$

Hence, we get $M_X(t) = p e^t (1 - q e^t)^{-1}$, for $0 < t < -\ln q$.

(G5) Lack of memory property: Every time a tail (T) occurs, the process renews itself. The event $\{X > x\}$ means the first x tosses are all T's. Likewise, the conditional event $\{X > x + y | X > x\}$ means that when the first x tosses are known to be all T's, the next y tosses also turn out to be all T's. However, the second batch of y tosses is independent of the first batch of x tosses. Hence, the desired conditional probability $P\{X > x + y | X > x\}$ equals the unconditional probability $P\{X > y\}$.

Immediate extension: The geometric(p) RV is also called a negative binomial($1, p$) RV. If we change the definition of success from “the appearance of the first head” to “the appearance of the r -th head” (for $r \geq 2$) or until we have a cumulative total of r heads, then the waiting time $Y = X_1 + \dots + X_r$, where $X_1, X_2, \dots, X_r$ are independent geometric(p) RVs, is said to have a negative binomial(r, p) distribution. Its first moment is r/p ; its PGF is $P_Y(t) = (pt)^r (1 - qt)^{-r}$, for $0 < t < q^{-1}$; and its MGF is $M_Y(t) = (p e^t)^r (1 - q e^t)^{-r}$, for $0 < t < -\ln q$.

1.1. Relation to existing literature

Waiting-time problems for patterns in random sequences have a long history in probability theory. Classical approaches include martingale methods, generating functions, finite-state Markov chain embeddings, and renewal-theoretic techniques. For example, Li (1980) studied the occurrence of patterns using martingale arguments, while Ebneshrashoob and Sobel (1990) focused on the sooner and later waiting-time problems in Bernoulli sequences. Fu and Chang (2002) and Uchida (1998) developed generating-function methods for waiting times associated with compound and multistate patterns.

The present paper has a different objective. Rather than seeking the most general theorem for a particular class of patterns, we emphasize a unified recursive framework based on one-step transition probabilities and renewal relations. The same methodology is used throughout the paper to derive exact PMFs for waiting times arising in independent trials, multinomial experiments, finite Markov chains, and random walks on graphs.

The novelty of the present paper is not in the examples considered, but in the use of a common recursive framework based on renewal relations and one-step transition probabilities. Repeated matrix recursion extends the familiar recursive construction of the geometric distribution. This perspective yields explicit PMFs for waiting times in several seemingly different settings and highlights their structural connections. The paper serves both as a study of waiting-time distributions and as an exposition of a unifying probabilistic method.

1.2. Motivation and applications

Waiting-time distributions arise naturally in a wide variety of scientific settings. In quality control, one may be interested in the number of inspections preceding a prescribed sequence of outcomes. In reliability and communications, the occurrence of particular patterns of successes and failures may signal the onset of an important event. In bioinformatics, researchers study waiting times until specified nucleotide patterns appear in DNA sequences. In stochastic processes, return times in Markov chains and random walks are fundamental quantities related to recurrence, mixing behavior, and long-run structure. While expected waiting times are often reported, the full PMF provides substantially richer information by quantifying variability and tail behavior. The examples studied in this paper illustrate how a common recursive methodology can be used to derive exact waiting-time distributions in several such settings. In statistical applications, waiting-time distributions provide benchmark models against which simulation studies and empirical observations can be compared.

The rest of the paper considers other, more complex notions of success. The paper is organized as follows: In Sections 2–4, we study the waiting time until a specified block of outcomes appears in (i) a sequence of independent, binary tosses, (ii) in a sequence of independent, multinomial outcomes, and (iii) in a Markov chain. In Section 5, we study the time until a random walk on the vertices of a graph returns to its starting position. Section 6 concludes the paper with a summary and several open problems.

2. Until a specific block of binary outcomes occurs

Consider a sequence of independent, binary RVs in which the appearance of a specific block of outcomes (for example, HTH or HTT or HTHTH) signifies a success. A popular

brain teaser goes as follows: “Consider tossing a *fair* coin repeatedly. Let a denote the expected number of tosses until the pattern HTH occurs, and b until HTT occurs. Which of the three cases holds true: $a < b$, $a = b$, or $a > b$?”

Many readers initially conjecture that $a = b$. Refer to Donnelly (2005). But actually, $a > b$. Here is a layman’s justification: When you are waiting for either pattern, you are somewhat excited to see the first H; then you are even more excited to see the pattern HT emerge. Now, if you were waiting for the pattern HTH, but instead HTT occurs, all hope is lost, and the process begins afresh for you. However, if you were waiting for HTT, but HTH occurs, not all is lost! You still have some excitement left because you already see the initial H of the desirable pattern. In fact, $a = 10, b = 8$; the difference of 2 is exactly the expected waiting time until the first H occurs. In Lemma 1 below, we give a rigorous proof not just for the fair coin, but for the more general case of a possibly biased coin.

It is worth repeating the key idea: Partial progress toward a target pattern may or may not be retained after a mismatch occurs. For HTT, the appearance of HTH still preserves the initial H needed for a future occurrence of HTT. By contrast, when searching for HTH, the appearance of HTT destroys all progress toward the target pattern. The recursive equations below quantify this phenomenon.

Lemma 1: The expected number of tosses of a biased coin with $P(\text{Heads}) = p$ until the pattern HTH occurs is $a = p^{-2}q^{-1} + p^{-1}$, and that until HTT occurs is $b = p^{-1}q^{-2}$.

Figure 1 shows the schematic diagram of the patterns of outcomes until success.



Figure 1: Transitions until a specific block of outcomes occurs: (a) HTH, and (b) HTT (each arrow represents $1/2$ probability)

Proof of Lemma 1 For any (partial) pattern Z such as ϕ, H, T, HT , etc., let $\mu(Z)$ denote the expected number of additional tosses until the target pattern HTH appears, starting after the appearance of the pattern Z . Clearly, $\mu(\phi) = a$, and the following system of equalities hold:

$$\begin{pmatrix} a \\ \mu(H) \\ \mu(HT) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \begin{bmatrix} q & p & 0 \\ 0 & p & q \\ q & 0 & 0 \end{bmatrix} \begin{pmatrix} a \\ \mu(H) \\ \mu(HT) \end{pmatrix},$$

solving which we get

$$\begin{pmatrix} a \\ \mu(H) \\ \mu(HT) \end{pmatrix} = \begin{bmatrix} p & -p & 0 \\ 0 & q & -q \\ -q & 0 & 1 \end{bmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{p^2q} \begin{pmatrix} 1 + pq \\ 1 \\ (1 + p)q \end{pmatrix}. \quad (4)$$

In particular, $a = (1 + pq)/(p^2q) = p^{-2}q^{-1} + p^{-1}$. Incidentally, from the second and third coordinates in (4), we get $\mu(H) = p^{-2}q^{-1}$ and $\mu(HT) = p^{-2} + p^{-1}$.

Likewise, for any pattern Z , let $\nu(Z)$ denote the expected number of additional tosses until the target pattern HTT appears, after the appearance of Z . Then $\nu(\phi) = b$, and

$$\begin{pmatrix} b \\ \nu(H) \\ \nu(HT) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \begin{bmatrix} q & p & 0 \\ 0 & p & q \\ 0 & p & 0 \end{bmatrix} \begin{pmatrix} b \\ \nu(H) \\ \nu(HT) \end{pmatrix},$$

This system of equations decomposes into: $b = 1 + p \nu(H) + q b = p^{-1} + \nu(H)$ and

$$\begin{pmatrix} \nu(H) \\ \nu(HT) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{bmatrix} p & q \\ p & 0 \end{bmatrix} \begin{pmatrix} \nu(H) \\ \nu(HT) \end{pmatrix},$$

which, when solved, gives us

$$\begin{pmatrix} \nu(H) \\ \nu(HT) \end{pmatrix} = \begin{bmatrix} q & -q \\ -p & 1 \end{bmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{q^2} \begin{pmatrix} 1+q \\ 1 \end{pmatrix}. \quad (5)$$

That is, $\nu(H) = q^{-2} + q^{-1}$ and $\nu(HT) = q^{-2}$. Thereafter, we obtain $b = p^{-1} + \nu(H) = p^{-1} + q^{-2} + q^{-1} = (pq)^{-1} + q^{-2} = p^{-1}q^{-2}$. The proof is now complete. Q.E.D.

Lemma 1 can be generalized to the following proposition. See Ross (1996).

Theorem 1: Suppose that the pattern Z , consisting of i H's and j T's, when the **fewest** number of leading symbols is deleted, becomes a sub-pattern $\tilde{Z}$ that is the same as when that many trailing symbols are deleted from Z . Then the expected number of tosses until Z appears is $p^{-i}q^{-j}$ more than the expected number of tosses until $\tilde{Z}$ appears.

For example, the expected time until pattern HTHTH appears is

$$\mu(HTHTH) = p^{-3}q^{-2} + \mu(HTH) = p^{-3}q^{-2} + p^{-2}q^{-1} + \mu(H) = p^{-3}q^{-2} + p^{-2}q^{-1} + p^{-1}.$$

For more general results regarding: (1) the expected time until a pattern occurs starting from another given pattern, (2) the expected time until one of several patterns occurs, and (3) the probability that one pattern occurs before some other patterns, see Li (1980).

2.1. Waiting time until HTT

So far, in this section, we studied the *expected* waiting time until a pattern appears. What about the PMF of the waiting time? Indeed, we can derive the PMF by suitably utilizing the renewal property. That is, having studied a property analogous to (G3), let us now focus on a property analogous to (G1).

Let X denote the waiting time until the pattern HTT occurs. Also, let $X(H)$ and $X(HT)$ denote the number of additional tosses until the target pattern HTT appears, starting after the appearance of H and HT , respectively. Then in the spirit of (2), we have

$$X = \begin{cases} 1 + X(H) & \text{with probability } p \\ 1 + \tilde{X} & \text{with probability } q, \end{cases} \quad (6)$$

where $\tilde{X}$ has the same distribution as X . Also, $X(H)$ and $X(HT)$ are interrelated as

$$X(H) = \begin{cases} 1 + \tilde{X}(H) & \text{w.p. } p \\ 1 + X(HT) & \text{w.p. } q; \end{cases} \quad \text{and} \quad X(HT) = \begin{cases} 1 + X(H) & \text{w.p. } p \\ 1 & \text{w.p. } q; \end{cases} \quad (7)$$

where $\tilde{X}(H)$ has the same distribution as $X(H)$. The two equations in (7) imply that

$$X(H) = \begin{cases} 1 + \tilde{X}(H) & \text{w.p. } p \\ 2 + \tilde{X}(H) & \text{w.p. } qp \\ 2 & \text{w.p. } q^2. \end{cases} \quad (8)$$

Thus, the distribution of $X(H)$ is related to itself, or the renewal property holds. Let the PMF of $X(H)$ be $\mathbf{p}_{X(H)} = (p_1, p_2, p_3, \dots)$. From (8), we have $p_1 = 0, p_2 = q^2$, and

$$p_{n+2} = p \cdot p_{n+1} + qp \cdot p_n, \text{ for } n \geq 1.$$

Let us write $\begin{pmatrix} p_2 \\ p_1 \end{pmatrix} = \begin{pmatrix} q^2 \\ 0 \end{pmatrix}$, and thereafter, for $k \geq 2$, we have

$$\begin{pmatrix} p_{k+1} \\ p_k \end{pmatrix} = \begin{bmatrix} p & pq \\ 1 & 0 \end{bmatrix} \begin{pmatrix} p_k \\ p_{k-1} \end{pmatrix} = \dots = \begin{bmatrix} p & pq \\ 1 & 0 \end{bmatrix}^{k-1} \begin{pmatrix} q^2 \\ 0 \end{pmatrix}.$$

Thus, we obtain the PMF $\mathbf{p}_{X(H)}$ of RV $X(H)$ by repeated pre-multiplication by a matrix, similar to how in (1) the geometric(p) probabilities were obtained by repeated pre-multiplication by q . We can verify that the total probability is

$$\sum_{k=1}^{\infty} \begin{pmatrix} p_{k+1} \\ p_k \end{pmatrix} = \sum_{k=1}^{\infty} \begin{bmatrix} p & pq \\ 1 & 0 \end{bmatrix}^{k-1} \begin{pmatrix} q^2 \\ 0 \end{pmatrix} = \begin{bmatrix} q & -pq \\ -1 & 1 \end{bmatrix}^{-1} \begin{pmatrix} q^2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Thereafter, from the second equation within (7), we can obtain the PMF of $X(HT)$ as $(q, p \cdot \mathbf{p}_{X(H)})$; and from (6), the PMF of X is $\mathbf{p}_X = (0, p \cdot \mathbf{p}_{X(H)} + q \cdot \mathbf{p}_X)$, or

$$pq^2(0, 0, 1, 1, 1, 1 - pq^2, 1 - 2pq^2, 1 - 3pq^2, \dots).$$

2.2. Waiting time until HTH

Let Y denote the waiting time until the pattern HTH occurs; and let $Y(H)$ and $Y(HT)$ denote the number of additional tosses until HTH appears, after H and HT , respectively. Then writing $\mathbf{q}_k = (P\{Y(H) = k\}, P\{Y = k\}, P\{Y = k - 1\})^T$, a vector of three probability masses, we have $\mathbf{q}_2 = (pq, 0, 0)^T$, and for $k \geq 2$, we have

$$\mathbf{q}_{k+1} = \begin{bmatrix} p & 0 & q^2 \\ p & q & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{q}_k = \dots = \begin{bmatrix} p & 0 & q^2 \\ p & q & 0 \\ 0 & 1 & 0 \end{bmatrix}^{k-1} \mathbf{q}_2.$$

Again, note the similarity between $\mathbf{q}_{k+1}$ and the mass of a geometric(p) RV in (1)! Clearly, $P\{Y = 1\} = 0$, and for $k \geq 2$, $P\{Y = k\}$ is the second element of $\mathbf{q}_k$. Hence, the entire PMF of Y , written as a vector, is

$$p^2q(0, 0, 1, 1, p^2 + q, p^3 + q, p^4 + q, p^4 + [1 + p + 2p^2]q^2, \dots).$$

Again, we can verify that the total probability is

$$\sum_{k=2}^{\infty} \mathbf{q}_k = \sum_{k=2}^{\infty} \begin{bmatrix} p & 0 & q^2 \\ p & q & 0 \\ 0 & 1 & 0 \end{bmatrix}^{k-2} \begin{pmatrix} pq \\ 0 \\ 0 \end{pmatrix} = \begin{bmatrix} q & 0 & -q^2 \\ -p & p & 0 \\ 0 & -1 & 1 \end{bmatrix}^{-1} \begin{pmatrix} pq \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

An extension: If success is redefined as the occurrence of either HTH or HTT, then the waiting time until success is exactly *one more than* the waiting time until HT occurs. As such, the waiting time to success has a PMF given by

$$\begin{aligned} & pq(0, 0, 1, p+q, p^2+q, p^3+q(p^2+q), p^4+q(p^2+q^2), p^5+q(p^4+p^2q+q^3), \dots) \\ &= pq(0, 0, 1, 1, (p^2+q), (p^2+q^2), (p^4+p^2q+q^3), (p^4+p^2q^2+q^4), \dots). \end{aligned}$$

We leave it to the reader to check that the sum of the probabilities is one.

3. Until a block of multinomial outcomes occurs

Each coin toss results in a binary outcome. What if the outcome were multinomial?

Several authors studied the waiting-time distributions for multinomial patterns. Uchida (1998) employed generalized PGFs to study waiting times and counts of pattern occurrences. Fu and Chang (2002) used finite Markov chain embedding techniques to derive PGFs and develop computational procedures for compound patterns. In contrast, the present paper emphasizes direct recursive construction of the probability mass function through one-step transition probabilities. This approach naturally leads to matrix recursions analogous to the recursion defining the geometric distribution and provides a common framework applicable throughout the paper.

As an example, consider the popular game of “Rock-Paper-Scissors.” The main objective of this section is not the game itself, but to illustrate how the recursive framework extends naturally from binary outcomes to multinomial outcomes. The game provides a familiar setting in which several competing patterns may define success.

In Rock-Paper-Scissors, two players simultaneously display with their hands one of these three symbols R-P-S. If both players display the same symbol, the play ends in a tie, and no one scores any points; otherwise, rock loses to paper, paper loses to scissors, and scissors loses to rock, and the winner scores one point. What are the probabilities of the three outcomes—a tie, A wins, and B wins? The probabilities depend on the strategies the players employ. We will consider several models of their strategies.

Scenario 1: Suppose that Player A selects the symbols R-P-S independently at random with probabilities (.3, .4, .3), Player B selects the symbol independently at random with probabilities (.2, .4, .4); and neither player knows the probability distribution of the other. Below, we tabulate the joint probabilities of the two players’ selections, with the rows representing Player A’s choices, the columns Player B’s, and the superscripts telling us who wins

(here T denotes a tie):

$$\begin{array}{ccccc}
 \begin{pmatrix} .06^T & .12^B & .12^A \\ .08^A & .16^T & .16^B \\ .06^B & .12^A & .12^T \end{pmatrix} & \begin{matrix} .3 \\ .4 \\ .3 \end{matrix} & \begin{matrix} R \\ P \\ S \end{matrix} & \text{A's choice} & \\
 \begin{matrix} .2 \\ R \end{matrix} & \begin{matrix} .4 \\ P \end{matrix} & \begin{matrix} .4 \\ S \end{matrix} & \text{B's choice} &
 \end{array}$$

In this case, the play ends in a tie with probability $.06 + .16 + .12 = .34$, A wins the play with probability $.12 + .08 + .12 = .32$, and B wins with probability $.12 + .16 + .06 = .34$. Thus, the game is slightly advantageous to Player B.

Scenario 2: Suppose Player A selects the symbols R-P-S independently at random with probabilities $(.3, .4, .3)$, and Player B knows this strategy. To beat Player A's most likely choice P, Player B selects the symbol S all the time. In this case, the play ends in a tie with probability $.3$, A wins with probability $.3$ and B wins with probability $.4$. Thus, by utilizing her knowledge of A's strategy, B substantially increases her advantage over A.

However, B's advantage is not sustainable in the long run. As soon as Player A would recognize Player B's strategy (of always choosing S), he ought to switch his own strategy to always choosing R; and thereby always winning.

Scenario 3: Suppose that Player A selects the symbols R-P-S independently at random with probabilities $(1/3, 1/3, 1/3)$, and declares it to Player B. Then, no matter which probability distribution Player B uses, the play ends in a tie, or A wins, or B wins with probability $1/3$ each. Not to give Player A any potential advantage by discovering her strategy, Player B should choose R-P-S with equal probabilities also.

In the rest of this subsection, we assume that Scenario 3 holds; that is, we suppose that at least one player chooses R-P-S with equal probability in each play. We address three questions: How many plays are needed until

(Q1) Player A wins three in a row (or the pattern AAA appears)?

(Q2) either Player A or Player B wins three in a row (either AAA or BBB appears)? and

(Q3) the same outcome occurs three times in a row (either AAA or BBB or TTT occurs)?

Figure 2 shows the recursive relations involved in answering the three questions.

Answer to (Q1):

Let N_{A^3} denote the number of plays until Player A wins three in a row. Using Proposition 2.1, $E[N_{A^3}] = 3^3 + 3^2 + 3 = 39$. To find the PMF of N_{A^3} , let us write

$$\mathbf{q}_k = (P\{N_{A^3} = k\}, P\{N_{A^3}(A) = k\}, P\{N_{A^3}(AA) = k\})^T.$$

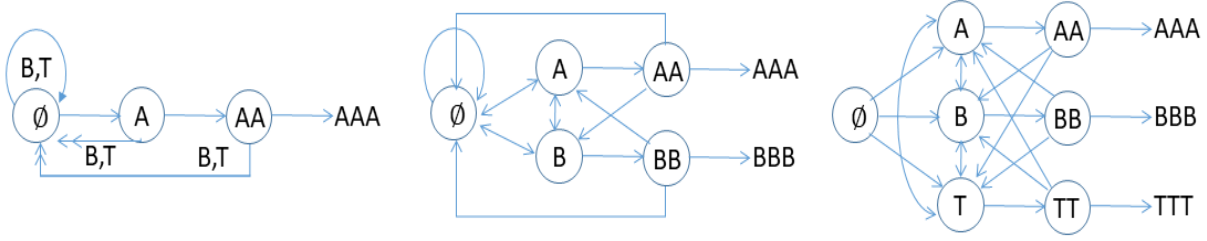


Figure 2: How many plays until (a) A^3 , (b) $\{A^3, B^3\}$, and (c) $\{A^3, B^3, T^3\}$ occurs? (Each arrowhead represents $1/3$ probability of transition in that direction.)

Then, we have

$$\mathbf{q}_1 = \begin{pmatrix} 0 \\ 0 \\ 1/3 \end{pmatrix} \text{ and for } k \geq 1, \mathbf{q}_{k+1} = \begin{bmatrix} 2/3 & 1/3 & 0 \\ 2/3 & 0 & 1/3 \\ 2/3 & 0 & 0 \end{bmatrix} \mathbf{q}_k,$$

solving which, we have

$$p_k = P\{N_{A^3} = k\} = 3^{-k} (1, 0, 0) \begin{bmatrix} 2 & 1 & 0 \\ 2 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix}^{k-1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \text{ for } k \geq 1.$$

Note the similarity to the PMF of the geometric(p) distribution in (1). In particular, N_{A^3} has PMF $(0, 0, 1/3^3, 2/3^4, 6/3^5, 18/3^6, 52/3^7 \dots)^T$; or,

$$p_k = P\{N_{A^3} = k\} = 3^{-k} a_k,$$

where $a_1 = 0, a_2 = 0, a_3 = 1$ and $a_{k+1} = 2(a_k + a_{k-1} + a_{k-2})$ for $k \geq 3$.

Answer to (Q2):

Let N_{A^3, B^3} denote the number of plays until either player wins three in a row. Then $E[N_{A^3, B^3}] = 39/2$. To find the PMF of N_{A^3, B^3} , note that $N_{A^3, B^3}(A)$ and $N_{A^3, B^3}(B)$ have identical distributions, and so do $N_{A^3, B^3}(AA)$ and $N_{A^3, B^3}(BB)$. Writing

$$\mathbf{q}_k = (P\{N_{A^3, B^3} = k\}, P\{N_{A^3, B^3}(A) = k\}, P\{N_{A^3, B^3}(AA) = k\})^T,$$

we have

$$\mathbf{q}_1 = \begin{pmatrix} 0 \\ 0 \\ 1/3 \end{pmatrix} \text{ and for } k \geq 1, \mathbf{q}_{k+1} = \begin{bmatrix} 1/3 & 2/3 & 0 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 \end{bmatrix} \mathbf{q}_k.$$

Therefore, the PMF of N_{A^3, B^3} is given by

$$p_k = P\{N_{A^3, B^3} = k\} = 3^{-k} (1, 0, 0) \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}^{k-1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \text{ for } k \geq 1.$$

Again, note the similarity to the PMF of the geometric(p) distribution in (1). In particular, N_{A^3, B^3} has PMF $(0, 0, 2/3^3, 4/3^4, 12/3^5, 34/3^6, 96/3^7 \dots)^T$; or,

$$p_k = P\{N_{A^3, B^3} = k\} = 3^{-k} b_k,$$

where $b_1 = 0, b_2 = 0, b_3 = 1$, and $b_{k+1} = 2(b_k + b_{k-1}) + b_{k-2}$ for $k \geq 3$.

Answer to (Q3):

Let N denote the number of plays until one of the patterns AAA or BBB or TTT appears. Then $E[N] = 39/3 = 13$. To find the PMF $\mathbf{p}$ of N , note that $N(A), N(B), N(T)$ have identical distributions, and so do $N(AA), N(BB), N(TT)$. Writing

$$\mathbf{q}_k = (P\{N(A) = k\}, P\{N(AA) = k\})^T,$$

we have

$$\mathbf{q}_1 = \begin{pmatrix} 0 \\ 1/3 \end{pmatrix} \text{ and for } k \geq 1, \mathbf{q}_{k+1} = \begin{bmatrix} 2/3 & 1/3 \\ 2/3 & 0 \end{bmatrix} \mathbf{q}_k.$$

Therefore, the PMF $\mathbf{p}$ of N is given by $p_1 = 0$, and

$$p_{k+1} = P\{N = k+1\} = P\{N(A) = k\} = 3^{-k} (1, 0) \begin{bmatrix} 2 & 1 \\ 2 & 0 \end{bmatrix}^{k-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \text{ for } k \geq 1.$$

Again, note the similarity to the PMF of the geometric(p) distribution in (1). Specifically, the PMF of N is $\mathbf{p} = (0, 0, 1/3^2, 2/3^3, 6/3^4, 16/3^5, 44/3^6, \dots)$; or,

$$p_k = P\{N = k\} = 3^{-(k-1)} c_k,$$

where $c_1 = 0, c_2 = 0, c_3 = 1$, and thereafter $c_{k+1} = 2(c_k + c_{k-1})$ for $k \geq 3$.

Another extension: Suppose that the game ends as soon as one player earns 10 points, and that player receives as many dollars as her points exceed the other player's. What is the PMF of the number of plays N_{10} when the game ends? What is the PMF of the winnings W_{10}^A of Player A? Who will win the game?

The number of plays until Player A earns 10 points has a negative binomial(10, 1/3) distribution, which is the sum of 10 independent, geometric(1/3) RVs. However, the distribution of N_{10} , the time until either Player A or Player B earns 10 points (the new definition of success), is not commonly known. Let us call its distribution the *doubly negative binomial*(10, 10; 1/3, 1/3). The attentive reader should check that:

- (i) The PMF of N_{10} is $P\{N_{10} = 10 + k + t\} = 2(9 + k + t)!/[9!k!t!] 3^{-(10+k+t)}$, where k denotes the points earned by the loser (with $0 \leq k \leq 9$) and $t \geq 0$ denotes the number of ties during the game.
- (ii) The PMF of W_{10}^A is symmetric over $\{\pm 1, \pm 2, \dots, \pm 10\}$; and it is given by

$$P\{W_{10}^A = -j\} = P\{W_{10}^A = j\} = \binom{19-j}{9} 2^{-(20-j)} \text{ for } j = 1, 2, \dots, 10.$$

- (iii) By symmetry, it follows that $P\{A \text{ wins}\} = 1/2 = P\{B \text{ wins}\}$.

4. Until a Markov chain exhibits a desired feature

So far we have considered sequences of outcomes that are independent. What if they are not? As the simplest departure from independence, we consider Markov chains. In a Markov chain, the probability distribution of the next outcome depends only on the current outcome and is independent of all previous outcomes.

For example, in Rock-Paper-Scissors, suppose that Player A chooses R-P-S according to a probability distribution that depends on his or her latest choice, but not on any previous choices. Suppose that the conditional probabilities of the next choice, given the current choice, are as follows: If the latest choice was R, then the next choice will be R-P-S with respective probabilities (.1, .5, .4). If the latest choice was P, then the probabilities are (.4, .3, .3). Finally, if the latest choice was S, then the probabilities are (.3, .2, .5). We summarize all such information in a one-step transition probability matrix

$$P = \begin{array}{ccc|c} \begin{pmatrix} .1 & .5 & .4 \\ .4 & .3 & .3 \\ .3 & .2 & .5 \end{pmatrix} & \begin{matrix} R \\ P \\ S \end{matrix} & \text{A's current choice} & \\ \begin{matrix} R \\ P \\ S \end{matrix} & & & \text{A's next choice} \end{array}$$

where P_{ij} denotes the conditional probability that A's next choice is j , given that her current choice is i , and is free of all previous choices.

The above example is a random walk on three states $\{R, P, S\} = \{\text{rock, paper, scissors}\}$. (The graph includes loops at each vertex). The given information suffices to determine what proportion of choices are (R, P, S) in the long run. Here is how: Reasoning backwards, we say that if such a probability distribution π exists and if the latest choice was made with probability distribution π , then the next choice will also have the same distribution π ; that is, $\pi^T P = \pi^T$, $\pi^T \mathbf{1} = 1$, where $\mathbf{1}$ denotes the vector of all entries 1. Therefore, it suffices to find the eigen vector of the matrix $(I - P^T)$ corresponding to the eigenvalue 0, and standardize it to ensure the sum is one. Alternatively, we may compute P^n for a sufficiently large $n \geq 1$ (in fact, to be more efficient, we compute $P^2, P^4, P^8, \dots$) until its rows are (almost) identical. For the above P matrix, the answer is $\pi_A^T = (.2762, .3143, .4095)$.

Suppose that Player B chooses R-P-S according to another Markov chain with transition matrix

$$Q = \begin{array}{ccc|c} \begin{pmatrix} .3 & .4 & .3 \\ .2 & .3 & .5 \\ .4 & .4 & .2 \end{pmatrix} & \begin{matrix} R \\ P \\ S \end{matrix} & \text{B's current choice} & \\ \begin{matrix} R \\ P \\ S \end{matrix} & & & \text{B's next choice} \end{array}$$

Then B chooses R-P-S with marginal stationary probabilities $\pi_B^T = (.2975, .3636, .3388)$.

If each of Players A and B acts according to his or her own transition matrix, but independently of each other's choice, then in the long run the probabilities of a tie, of A winning, and of B winning can be obtained as was done in Scenario 1, with marginal selection probabilities replaced by their respective stationary probabilities. The reader may

verify that B's probabilities of win, loss, and tie are .3288, .3360, and .3352, respectively. Thus, the game is slightly unfavorable to Player B.

However, suppose that after playing for a long time, Player B analyzed Player A's behavior and discovered his transition matrix P . (To do so, simply use the relative frequencies of RR-RP-RS to estimate the first row of P , those of PR-PP-PS to estimate the second row, and those of SR-SP-SS to estimate the third row.) However, Player A remained oblivious to Player B's transition matrix Q . How should Player B take advantage of her knowledge?

Since whatever option A chooses in one play determines his probability distribution in the next play, B should choose that option which will beat A's most likely next choice. In other words, if A chose R/P/S in the current play, then in the next play B should choose respectively S/P/R. With this strategy, in the next play B will win with respective probabilities (.5, .4, .5), and lose with respective probabilities (.1, .3, .2) and the game will end up in a tie with respective probabilities (.4, .4, .3). Overall, B's probabilities of win, loss and tie are .4686, .2038 and .3276 respectively. The game became highly advantageous to Player B who analyzed Player A's past behavior to discover his transition matrix. It pays to analyze the opponent's strategy!

5. Until a random walk returns to origin

The examples in this section demonstrate that the same recursive methodology applies beyond pattern-occurrence problems. Return times in random walks are traditionally studied through expectations and asymptotic methods; here we construct their full distributions by identifying suitable renewal structures.

In general, a random walk is a discrete-time Markov chain; that is, a sequence of states in which the next state is chosen according to a conditional probability distribution that depends on the current state. For a symmetric random walk on the vertices of a graph, the next vertex is equally likely to be any one of the adjacent vertices. Random walks on graphs are related to reversible Markov chains (see Aldous and Fill (2014)) and to electric networks (see Doyle and Snell (1984)). We focus on symmetric random walks on a linear path, a cycle, a wheel-and-spokes graph, and some familiar solids such as a tetrahedron, an octahedron, and a hexahedron.

At time 0, the walk is at a particular vertex, which we call the starting vertex or the origin. We will study the time until the walk returns to the origin. Much of the literature on random walks (see Sarkar (2006); Maiti and Sarkar (2019)) emphasizes expectations, variances, recurrence properties, and asymptotic behavior of return times. Motivated by these studies, we examine the full PMF of the return time for several finite graphs and show how it can be obtained recursively from one-step transition probabilities.

5.1. Random walks on a linear path, a cycle, or a wheel-and-spokes

Here we study a symmetric random walk on some simple graphs. See Figure 3.

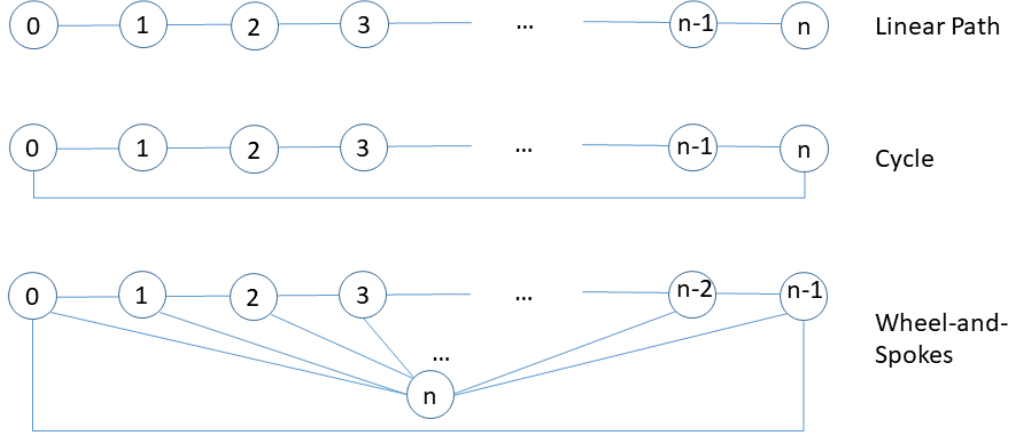


Figure 3: Graphs of (a) a linear path L_n , (b) a cycle C_n , and (c) a wheel-and-spokes W_n . Transition on each edge happens in both directions.

5.1.1. Linear path

A linear path L_n is a graph consisting of vertices $\{0, 1, 2, \dots, n\}$ and edges $(i-1, i)$ for $i = 1, 2, \dots, n$. The walk begins at vertex 0 at time $t = 0$. At time $t = 1$, the walk surely reaches vertex 1. Whenever the walk is at a vertex i (for $i = 1, 2, \dots, n-1$), at the next instance, it moves to either its left neighbor $(i-1)$ or its right neighbor $(i+1)$, with probability $1/2$ each. Finally, from vertex n the walk surely moves to vertex $(n-1)$. Let T_R denote the waiting time until the walk returns to the starting vertex 0. It is well known that $E[T_R] = n$. Here, we derive the PMF of T_R .

Let ${}_i T_j$ denote the number of steps needed for the walk to go from vertex i to vertex j . Then $T_R = 1 + {}_1 T_0$; and the following recursive relations hold:

$$\begin{aligned} {}_1 T_0 &= \begin{cases} 1 & \text{w.p. } 1/2 \\ 1 + {}_2 T_0 & \text{w.p. } 1/2 \end{cases}, \quad {}_2 T_0 = \begin{cases} 1 + {}_1 T_0 & \text{w.p. } 1/2 \\ 1 + {}_3 T_0 & \text{w.p. } 1/2 \end{cases}, \quad \dots, \\ {}_{n-1} T_0 &= \begin{cases} 1 + {}_{n-2} T_0 & \text{w.p. } 1/2 \\ 1 + {}_n T_0 & \text{w.p. } 1/2 \end{cases}, \end{aligned}$$

and ${}_n T_0 = 1 + {}_{n-1} T_0$. Thus, to find the PMF of T_R , we need the PMF of ${}_1 T_0$, for which we need the PMF of ${}_2 T_0$, etc. In fact, we can simultaneously find the PMF of the entire random vector $({}_1 T_0, {}_2 T_0, {}_3 T_0, \dots, {}_n T_0)$. To do so, define the following $n \times 1$ vectors

$$\mathbf{p}_k = (P\{{}_1 T_0 = k\}, P\{{}_2 T_0 = k\}, P\{{}_3 T_0 = k\}, \dots, P\{{}_n T_0 = k\})^T, \text{ for } k \geq 1.$$

Then $\mathbf{p}_1 = 2^{-1}(1, 0, \dots, 0)^T$, and for $k \geq 2$,

$$\mathbf{p}_k = R \mathbf{p}_{k-1} = \dots = R^{k-1} \mathbf{p}_1, \tag{9}$$

where R is an $n \times n$ matrix given by

$$R = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 0 & \dots & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & \dots & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & \dots & 0 \\ \dots & & \dots & & \dots & & \\ \dots & & \dots & & \dots & & \\ 0 & & \dots & 0 & 1/2 & 0 & 1/2 \\ 0 & & \dots & 0 & 0 & 1 & 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 1 & 0 & \dots & 0 \\ \dots & & \dots & & \dots & & \\ \dots & & \dots & & \dots & & \\ 0 & & \dots & 0 & 1 & 0 & 1 \\ 0 & & \dots & 0 & 0 & 2 & 0 \end{bmatrix}.$$

Note that $\mathbf{p}_k$ in (9) is formally similar to a geometric(p) RV given in (1). Using (9), we evaluate $\mathbf{p}_2 = 2^{-2}(0, 1, 0, 0, \dots, 0)^T$, $\mathbf{p}_3 = 2^{-3}(1, 0, 1, 0, \dots, 0)^T$, etc. The attentive reader should verify that $\sum_{k=1}^{\infty} \mathbf{p}_k = (I - R)^{-1} \mathbf{p}_1 = \mathbf{1}$, the $n \times 1$ vector of all entries 1.

5.1.2. Cycle

To a linear path L_n , let us augment an edge joining vertices n and 0 . See Figure 3 (middle panel). The resulting graph C_n is called the cycle graph consisting of vertices $\{0, 1, 2, \dots, n\}$ on a circle numbered clockwise. Suppose that at time $t = 0$, the walk begins at vertex 0 . Whenever the walk is at a vertex i (for $i = 0, 1, 2, \dots, n$), at the next instance it moves to either its counterclockwise neighbor ($i - 1$) or its clockwise neighbor ($i + 1$) with probability $1/2$ each. Here, addition and subtraction are done modulo $(n + 1)$. We are interested in the waiting time T_R until the walk returns to the starting vertex 0 . Again, it is well known that $E[T_R] = n + 1$. Here, we derive the full PMF of T_R .

Without loss of generality, we assume that the walk goes to vertex 1 at time $t = 1$ (for otherwise, we can reverse-number the vertices). Also, splitting vertex 0 into two vertices 0 and $(n + 1)$, we can convert cyclic graph C_n into linear path L_{n+1} . Let us write $\mathbf{0}$ to denote the subset of vertices $\{0, n + 1\}$, and ${}_i T_{\mathbf{0}}$ to denote the number of steps needed for the walk to go from vertex i to either vertex 0 or vertex $n + 1$ on the linear path L_{n+1} . Then T_R on the cycle graph C_n is the same as $1 + {}_1 T_{\mathbf{0}}$ on linear path L_{n+1} ; and the following recursive relations hold:

$$\begin{aligned} {}_1 T_{\mathbf{0}} &= \begin{cases} 1 & \text{w.p. } 1/2 \\ 1 + {}_2 T_{\mathbf{0}} & \text{w.p. } 1/2 \end{cases}, \quad {}_2 T_{\mathbf{0}} = \begin{cases} 1 + {}_1 T_{\mathbf{0}} & \text{w.p. } 1/2 \\ 1 + {}_3 T_{\mathbf{0}} & \text{w.p. } 1/2 \end{cases}, \quad \dots, \\ {}_n T_{\mathbf{0}} &= \begin{cases} 1 + {}_{n-1} T_{\mathbf{0}} & \text{w.p. } 1/2 \\ 1 & \text{w.p. } 1/2. \end{cases} \end{aligned}$$

Furthermore, by reverse numbering, we note that ${}_1 T_{\mathbf{0}}$ and ${}_n T_{\mathbf{0}}$ have the same distribution; so do ${}_2 T_{\mathbf{0}}$ and ${}_{n-1} T_{\mathbf{0}}$; etc. Thus, if n is odd, then T_R on C_n has the same distribution as T_R on $L_{(n+1)/2}$. On the other hand, if n is even, then T_R on C_n has the same distribution as T_R on a modified linear path $L_{\lfloor n/2 \rfloor}$ that has an extra loop at vertex $\lfloor n/2 \rfloor$. In both cases, (9) holds, at least formally. However, for even n , the R matrix changes to an $\lfloor n/2 \rfloor \times \lfloor n/2 \rfloor$

matrix given by

$$R_1 = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 0 & \dots & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & \dots & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & \dots & 0 \\ \dots & & \dots & & & \dots & \\ \dots & & \dots & & & \dots & \\ 0 & \dots & 0 & 1/2 & 0 & 1/2 & \\ 0 & \dots & 0 & 0 & 1/2 & 1/2 & \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 1 & 0 & \dots & 0 \\ \dots & & \dots & & & \dots & \\ \dots & & \dots & & & \dots & \\ 0 & \dots & 0 & 1 & 0 & 1 & \\ 0 & \dots & 0 & 0 & 1 & 1 & \end{bmatrix}.$$

For this R_1 matrix also, the reader can verify that $\sum_{k=1}^{\infty} \mathbf{p}_k = (I - R_1)^{-1} \mathbf{p}_1 = \mathbf{1}$.

5.1.3. Wheel-and-spokes

To a cycle C_{n-1} , let us augment a new vertex n (at the center of the cycle) and join it to all vertices 0 through $(n-1)$. See Figure 3 (lower panel). The resultant graph W_n is called the wheel-and-spokes graph consisting of vertices $\{0, 1, 2, \dots, n-1\}$ on a circle and numbered clockwise, and vertex n at the center. Whenever the walk is at a vertex i (for $i = 0, 1, 2, \dots, n-1$), at the next instance it moves to either its counterclockwise neighbor $(i-1)$ or its clockwise neighbor $(i+1)$, or to the central hub with probability $1/3$ each. [Again, addition and subtraction are done modulo $(n+1)$.] We are interested in the waiting time T_R until the walk returns to the starting vertex.

If the starting vertex is n , without loss of generality, we assume that the walk goes to vertex 0 at time $t = 1$ (for otherwise, we can renumber the vertices). Thereafter, the walk returns to vertex n with probability $1/3$ or goes to another vertex, which we can renumber as vertex 0 . Hence, T_R equals one more than a geometric($1/3$) RV.

On the other hand, if the starting vertex is on the circle, we renumber it as vertex 0 . Then we have the following interrelations:

$$\begin{aligned} T_R &= \begin{cases} 1 + {}_n T_0 & \text{w.p. } 1/3 \\ 1 + {}_1 T_0 & \text{w.p. } 2/3 \end{cases} ; \quad {}_n T_0 = \begin{cases} 1 & \text{w.p. } 1/n \\ 1 + {}_i T_1 & \text{w.p. } 1/n, \text{ for } 1 \leq i \leq n-1 \end{cases} ; \\ {}_1 T_0 &= \begin{cases} 1 & \text{w.p. } 1/3 \\ 1 + {}_2 T_0 & \text{w.p. } 1/3 \\ 1 + {}_n T_0 & \text{w.p. } 1/3 \end{cases} ; \quad {}_{n-1} T_0 = \begin{cases} 1 & \text{w.p. } 1/3 \\ 1 + {}_{n-2} T_0 & \text{w.p. } 1/3 \\ 1 + {}_n T_0 & \text{w.p. } 1/3 \end{cases} ; \\ \text{and for } 2 \leq i \leq n-2, \quad {}_i T_0 &= \begin{cases} 1 + {}_{i-1} T_0 & \text{w.p. } 1/3 \\ 1 + {}_{i+1} T_0 & \text{w.p. } 1/3 \\ 1 + {}_n T_0 & \text{w.p. } 1/3 \end{cases} . \end{aligned}$$

Writing ${}_i p_k = P\{{}_i T_0 = k\}$ for $i = 1, \dots, n$, and stacking them to form a column vector $\mathbf{p}_k = ({}_1 p_k, {}_2 p_k, \dots, {}_n p_k)^T$, we note that $\mathbf{p}_1 = (1/3, 0, 0, \dots, 0, 1/3, 1/n)^T$ and $\mathbf{p}_{k+1} =$

$R_w \mathbf{p}_k = \dots = R_w^k \mathbf{p}_1$, formally agreeing with the geometric(p) distribution, where

$$R_w = \begin{bmatrix} 0 & 1/3 & 0 & 0 & 0 & \dots & 0 & 1/3 \\ 1/3 & 0 & 1/3 & 0 & 0 & \dots & 0 & 1/3 \\ 0 & 1/3 & 0 & 1/3 & 0 & \dots & 0 & 1/3 \\ \dots & & \dots & & & \dots & & \dots \\ \dots & & \dots & & & \dots & & \dots \\ 0 & & \dots & 0 & 0 & 1/3 & 0 & 1/3 \\ 1/n & & \dots & 1/n & 1/n & 1/n & 1/n & 0 \end{bmatrix}.$$

For this R_w matrix also, the reader can verify that $\sum_{k=1}^{\infty} \mathbf{p}_k = (I - R_w)^{-1} \mathbf{p}_1 = \mathbf{1}$. Finally, $P\{T_R = k + 1\} = ({}_1p_k + {}_{n-1}p_k + {}_1p_k)/3 = (2/3) {}_1p_k + (1/3) {}_{n-1}p_k$, for $k \geq 2$.

5.2. Random walks on a tetrahedron, an octahedron, or a hexahedron

Whereas the expected return time to the origin is documented in the literature (see Sarkar (2020)), we study the return time distribution for a random walk on the vertices of three Platonic solids—tetrahedron, octahedron, and hexahedron. See Figure 4. We leave it to the reader to study the remaining two Platonic solids—dodecahedron and icosahedron. They should also ponder why there are exactly five Platonic solids. See Wi (date).

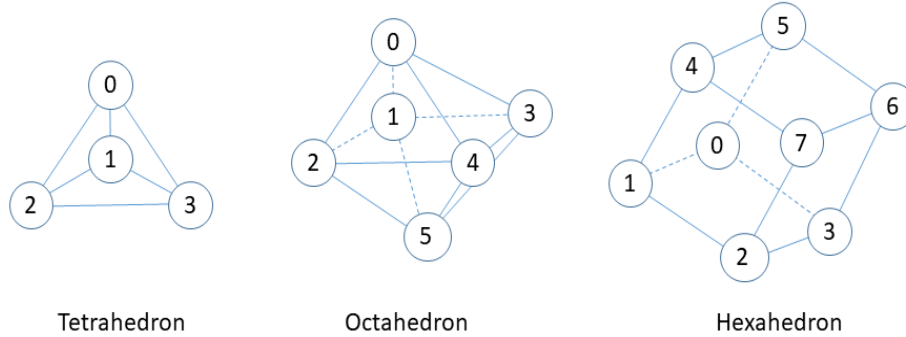


Figure 4: Graphs of (a) a tetrahedron, (b) an octahedron, and (c) a hexahedron

5.2.1. Tetrahedron

The graph of a tetrahedron is the complete graph on four vertices. Starting from the origin, the walk goes to any one of the other three vertices. Thereafter, at each step, the walk can return to the origin with probability $1/3$, or the process renews. Hence, the waiting time until return to origin, T_R , is exactly one more than a geometric($1/3$) RV.

5.2.2. Octahedron

The graph of an octahedron can be thought of as follows: Begin with a square whose four vertices, labeled counterclockwise as 1, 2, 4, 3, are edge-connected to a fifth vertex, labeled 0, above the square and also to a sixth vertex, labeled 5, below the square. The graph is not complete because members within the pairs of diametrically opposite vertices, $(0, 5)$, $(1, 4)$, $(2, 3)$, are not directly edge-connected. At each step, the walk goes to one of its four neighbors with equal probability $1/4$.

Starting from vertex 0, without loss of generality, the walk surely goes to vertex 1 (for otherwise we can rotate the octahedron suitably). Thereafter, at each step it returns to vertex 0 with probability $1/4$, or it goes to vertex 5 with probability $1/4$ (and in the next step it surely returns to one of vertices $1, 2, 3, 4$), or it goes to vertices 2 or 3 with probability $1/2$. In these latter cases, we can renumber the vertices of the square to ensure that the walk is at vertex 1. Thus, we have $T_R = 1 + {}_1T_0$, where ${}_1T_0$ satisfies a slightly more involved renewal property than (2); namely,

$${}_1T_0 = \begin{cases} 1 & \text{w.p. } 1/4 \\ 1 + {}_1\bar{T}_0 & \text{w.p. } 1/2 \\ 2 + {}_1\tilde{T}_0 & \text{w.p. } 1/4 \end{cases} \quad (10)$$

where ${}_1\bar{T}_0$ and ${}_1\tilde{T}_0$ have the same distribution as ${}_1T_0$. Writing $\mathbf{p} = (p_1, p_2, \dots)^T$ as the PMF of ${}_1T_0$, we note that $p_1 = 1/4, p_2 = 1/8$, and thereafter the following recursive relation holds:

$$p_{k+2} = \frac{1}{2} p_{k+1} + \frac{1}{4} p_k, \quad \text{for any } k \geq 1.$$

Consequently, the PMF of T_R is the same as the PMF of $X(H)$, described in (8) of Subsection 2.1, specialized to $p = 1/2$. In particular,

$$P\{T_R = k\} = 2^{-2(k-1)} (1, 0) \begin{bmatrix} 2 & 1 \\ 4 & 0 \end{bmatrix}^{k-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{for } k \geq 1.$$

Indeed, the PMF of T_R is $2^{-2(k-1)} a_k$, where $a_1 = 0, a_2 = 1$, and $a_{k+1} = a_k + a_{k-1}$ for $k \geq 2$, the Fibonacci sequence!

5.2.3. Hexahedron

Let us describe the graph of a hexahedron slightly unconventionally: Consider a hexagon whose six vertices are labeled counterclockwise as $1, 2, 3, 6, 5, 4$. Thereafter, vertices $1, 3, 5$ are edge-connected to a seventh vertex, labeled 0, below the hexagon; and vertices $2, 4, 6$ are edge-connected to an eighth vertex, labeled 7, above the hexagon. At each step, the RW goes to one of three neighbors with probability $1/3$ each.

Starting from vertex 0, without loss of generality, the walk surely goes to vertex 1 (for otherwise we can rotate the hexahedron suitably). Thereafter, at each step it returns to vertex 0 with probability $1/3$, or it goes to one of the vertices $2, 4, 6$ which are two edges away from vertex 0 with probability $2/3$. In the latter case, in the next step it goes to vertex 7 with probability $1/3$ (and then it surely returns to one of the vertices $2, 4, 6$), or it goes to one of the vertices $1, 3, 5$, which are adjacent to vertex 0, with probability $2/3$. In these latter cases, holding vertices 0 and 7 fixed, we can rotate the hexahedron to ensure that the walk is either at vertex 1 or at vertex 2. Thus, we have $T_R = 1 + {}_1T_0$, where

$${}_1T_0 = \begin{cases} 1 & \text{w.p. } 1/3 \\ 1 + {}_2T_0 & \text{w.p. } 2/3 \end{cases}; \quad {}_2T_0 = \begin{cases} 2 + {}_2\bar{T}_0 & \text{w.p. } 1/3 \\ 1 + {}_1T_0 & \text{w.p. } 2/3 \end{cases}$$

where ${}_2\bar{T}_0$ has the same distribution as ${}_2T_0$. Writing $\mathbf{p} = (p_1, p_2, \dots)^T$ as the PMF of ${}_1T_0$, and $\mathbf{q} = (q_1, q_2, \dots)^T$ as the PMF of ${}_2T_0$, we note that $p_1 = 1/3, q_1 = 0$; and defining $q_0 = 0$, the following recursive relations hold:

$$p_{k+1} = \frac{2}{3} q_k \quad \text{and} \quad q_{k+1} = \frac{2}{3} p_k + \frac{1}{3} q_{k-1}, \quad \text{for any } k \geq 1.$$

We rewrite these recursive relations in matrix notation and iterate to obtain

$$\begin{pmatrix} p_{k+1} \\ q_{k+1} \\ q_k \end{pmatrix} = \begin{bmatrix} 0 & 2/3 & 0 \\ 2/3 & 0 & 1/3 \\ 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} p_k \\ q_k \\ q_{k-1} \end{pmatrix} = \dots = \begin{bmatrix} 0 & 2/3 & 0 \\ 2/3 & 0 & 1/3 \\ 0 & 1 & 0 \end{bmatrix}^k \begin{pmatrix} 1/3 \\ 0 \\ 0 \end{pmatrix}.$$

The solution is formally similar to the geometric(p) distribution. To verify the total probability, one can check that

$$\begin{bmatrix} 1 & -2/3 & 0 \\ -2/3 & 1 & -1/3 \\ 0 & -1 & 1 \end{bmatrix}^{-1} \begin{pmatrix} 1/3 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

In particular, the PMF $\mathbf{p}$ of ${}_1T_0$ simplifies to

$$p_k = P\{{}_1T_0 = k\} = 3^{-k} (1, 0, 0) \begin{bmatrix} 0 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 3 & 0 \end{bmatrix}^{k-1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \text{for } k \geq 1.$$

Thus, the PMF $\mathbf{p}$ of ${}_1T_0$ is $3^{-k}a_k$, where $a_{2m} = 0$ for all $m \geq 1$; $a_1 = 1, a_3 = 4$, and $a_{2m+1} = 7a_{2m-1}$ for all $m \geq 2$.

Finally, the PMF of T_R is $(0, \mathbf{p})^T$. That is, T_R is an *even* natural number with $P\{T_R = 2\} = 1/3, P\{T_R = 4\} = 4/27, P\{T_R = 2m + 2\} = (7/9)P\{T_R = 2m\}$ for all $m \geq 2$.

6. Summary and exercises

If each trial results in a success (with probability p) or a failure (with probability $q = 1 - p$), then the waiting time X until success has a geometric(p) distribution, with PMF $P\{X = x\} = q^{x-1}p$ for $x = 1, 2, 3, \dots$. We consider several more complicated notions of success and illustrate the PMF of the waiting time until success. Specifically, the success may involve a block of outcomes—either binomial or multinomial—that are independent, or Markovian, or symmetric random walks on the vertices of a given graph. Diligent researchers may extend the examples to more general situations, which may be useful in applications involving reliability systems, communication protocols, biological sequence analysis, game strategies, and stochastic network models.

To test their comprehension, student readers may solve the following exercises:

- 1 Without evaluating a sequence or differentiating, find $E[X^2]$, where X follows (a) a geometric(p), or (b) a negative binomial(r, p) distribution.
- 2 Find the PMF of the number of tosses of a fair or a biased coin until the pattern (a) HTTHT, or (b) HTHTH occurs, or (c) either HTTHT or HTTT occurs.

- 3 How many rolls of a fair die are necessary until: (a) any one of the six values shows up three times in a row? (b) the sequence (1, 2, 3, 4, 5, 6) shows up exactly in that order? (c) the six values show up successively (but in no particular order)?
- 4 In an R-P-S game, the two players choose their respective options according to Markov chains with transition probability matrices P and Q , respectively.

$$P = \begin{array}{ccc|c} \begin{pmatrix} .2 & .4 & .4 \\ .3 & .3 & .4 \\ .4 & .3 & .3 \end{pmatrix} & \begin{matrix} R \\ P \\ S \end{matrix} & \text{A's choice} & \\ \begin{matrix} R \\ P \\ S \end{matrix} & & & \text{A's next choice} \end{array}, \quad Q = \begin{array}{ccc|c} \begin{pmatrix} .3 & .3 & .4 \\ .4 & .3 & .3 \\ .2 & .4 & .4 \end{pmatrix} & \begin{matrix} R \\ P \\ S \end{matrix} & \text{B's choice} & \\ \begin{matrix} R \\ P \\ S \end{matrix} & & & \text{B's next choice} \end{array}$$

Neither player knows the other player's matrix. What are the probabilities of winning for Player A, Player B, and a tie?

- 5 Find the PMF of the return time to the starting vertex for (a) a dodecahedron and (b) an icosahedron. See these objects and their planar graphs in Wi (no date).

Acknowledgment

Dr. Vahid Andalib gave feedback on an earlier draft, and Dr. Honglang Wang mentioned several references. I thank them both. A learned referee offered many helpful suggestions to highlight the purpose, motivation, and applications of the mathematical principles advocated in this paper and to improve the exposition.

Conflict of interest

The author does not have any financial or non-financial conflict of interest to declare for the research work included in this paper.

References

- Aldous, D. J. and Fill, J. A. (2014). *Reversible Markov Chains and Random Walks on Graphs*. Unpublished monograph.
- Devore, J. L. and Berk, K. N. (2007). *Modern Mathematical Statistics*. Thomson-Brooks/Cole, Belmont, CA.
- Donnelly, P. (2005). How juries are fooled by statistics. TED: Ideas worth spreading. See also: <https://www.ted.com/talks?q=peter>
- Doyle, P. G. and Snell, J. L. (1984). *Random Walks and Electric Networks*, volume 22 of *Carus Mathematical Monographs*. Mathematical Association of America, Washington, DC.
- Dudewicz, E. J. and Mishra, S. N. (1988). *Modern Mathematical Statistics*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York.
- Hoel, P. G., Port, S. C., and Stone, C. J. (1971). *Introduction to Probability Theory*. Houghton Mifflin, Boston, MA.

- Li, S. R. (1980). A martingale approach to the study of occurrence of a sequence pattern in repeated experiments. *The Annals of Probability*, **8**, 1171–1176.
- Maiti, S. I. and Sarkar, J. (2019). Symmetric walks on paths and cycles. *Mathematics Magazine*, **92**, 252–268.
- Rohatgi, V. K. (2003). *Statistical Inference*. Dover, New York, 2 edition.
- Ross, S. M. (1996). *Stochastic Processes*. Wiley Series in Probability and Statistics. John Wiley & Sons, New York, 2 edition.
- Sarkar, J. (2006). Random walk on a polygon. In *Recent Developments in Nonparametric Inference and Probability*, volume 50 of *IMS Lecture Notes*, pages 31–43. Institute of Mathematical Statistics, Beachwood, OH.
- Sarkar, J. (2020). A symmetric random walk on the vertices of a hexahedron. *Mathematics Student*, **89**, 63–85.
- Wi (no date). Platonic solid. Wikipedia: The Free Encyclopedia. Online; accessed 2025-10-29.