5-2015

Boundary Problems for One and Two Dimensional Random Walks

Miky Wright
Western Kentucky University, miky.wright768@topper.wku.edu

BOUNDARY PROBLEMS FOR ONE AND
TWO DIMENSIONAL RANDOM WALKS

A Thesis
Presented to
The Faculty of the Department of Mathematics
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By
Miky Wright

May 2015

BOUNDARY PROBLEMS FOR ONE AND
TWO DIMENSIONAL RANDOM WALKS

Date Recommended April 15, 2015

David Neal
Dr. David Neal, Director of Thesis

Ngoc Nguyen
Dr. Ngoc Nguyen

Lan Nguyen
Dr. Lan Nguyen

Cal Abro 4-20-15
Dean, Graduate Studies and Research Date

I would like to dedicate my thesis to my beloved husband Tracy Lee Wright, my parents who are living in China, and my in-laws Damon and Rita Wright for their consistent support all these years while I have been seeking my degrees. Also I dedicate my study and work to my dear daughter Julia Zhong Wright who allows me to practice long-distance parenting and motivates me to be a better woman and mother.

ACKNOWLEDGEMENTS

I thank, first of all, my thesis director Dr. David Neal, for his willingness to spend his summer and winter breaks to work with me on this thesis. Without Dr. Neal's committed directing, provision of his *Mathematica* code, as well as his inspiring suggestions, I would not have accomplished my thesis to this stage. Secondly, I want to thank my committee members Dr. Ngoc Nguyen, who serves as the second reader of my thesis, and Dr. Lan Nguyen, for their valuable time spent on reading my paper and making much criticism that greatly improved my work. I would also express my appreciation to Chase Cain for his help and encouragement when I finally decided to convert the whole typed thesis from Word version to Latex.

CONTENTS

List of Tables

List of Figures

Boundary Problems for One and Two Dimensional Random Walks

Miky Wright May 2015 108 Pages

Directed by: Dr. David Neal, Dr. Ngoc Nguyen, and Dr. Lan Nguyen

Department of Mathematics Western Kentucky University

This thesis provides a study of various boundary problems for one and two dimensional random walks. We first consider a one-dimensional random walk that starts at integer-valued height $k \geq 0$, with a lower boundary being the x-axis, and on each step moving downward with probability q being greater than or equal to the probability of going upward p. We derive the variance and the standard deviation of the number of steps T needed for the height to reach 0 from k, by first deriving the moment generating function of T.

We then study two types of two-dimensional random walks with four boundaries. A Type I walk starts at integer-valued coordinates (h, k), where $0 \leq h \leq m$ and $0 \leq k \leq n$. On each step, the process moves one unit either up, down, left, or right with positive probabilities p_u, p_d, p_l, p_r, respectively, where $p_u + p_d + p_l + p_r = 1$. The process stops when it hits a boundary. A Type II walk is similar to a Type I walk except that on each step, the walk moves diagonally, either left and upward, left and downward, right and downward, or right and upward with positive probabilities p_{lu}, p_{ld}, p_{rd}, p_{ru}, respectively. We mainly answer two questions on these two types of two-dimensional random walks: (1) What is the probability of hitting one boundary before the others from an initial

starting point? (2) What is the average number of steps needed to hit a boundary? To do so, we introduce a Markov Chains method and a System of Equations method.

We then apply the obtained results to a boundary problem involving two independent one-dimensional random walks and answer various questions that arise.

Finally, we develop a conjecture to calculate the probability of a two-sided downward-drifting Type II walk with even-valued starting coordinates hitting the x-axis before the y-axis, and we test the result with *Mathematica* simulations.

Chapter 1

Introduction

A random walk is the process by which randomly-moving objects wander away from the initial starting places. It is a mathematical formalization of a path that consists of a succession of random steps. As early as in 1905, Karl Pearson [6] first introduced the term random walk. Since then, random walks have been used in various fields. For example, modeling a fluctuating stock price in economics, tracing the path of a particular molecule in physics, or simply playing a card game are all related to some type of random walk. In this thesis, we will study some boundary problems for one-dimensional random walks with one or two boundaries and two types of two-dimensional random walks with two or four boundaries.

We first provide the background on one-dimensional boundary problems. In Section 2.1, we describe the process of a one-dimensional random walk with two boundaries, and give the formulas for the probability of either reaching the top boundary before the bottom boundary or the probability of reaching the bottom boundary before the top boundary. In Section 2.2, we analyze the single boundary problem of one-dimensional random walk. Providing that the probability for the walk moving toward the boundary is greater than the probability of moving toward the opposite direction, we give the formula for computing the average number of steps needed to hit the boundary. In Section 2.3, we derive a formula for the moment generating function (mgf) $M_T(t)$, where $T \equiv {}_kT_0$ is the number of steps for a one-dimensional random walk to reach its single boundary height 0 when starting at height $k > 0$. In this case, we assume $q > p$ (the probability of moving downward is greater than the probability of moving upward). Using the properties of the mgf, we are able to derive a formula for variance and standard deviation of T.

In Chapter 3, we discuss the four-sided boundary problem for a Type I two-dimensional random walk that begins at integer-valued coordinates (h, k), with boundaries $x = 0$ (the y-axis), $y = 0$ (the x-axis), $x = m$, and $y = n$. On each step, the random walk moves one unit either up, down, left, or right with positive probabilities p_u, p_d, p_l, p_r, respectively, where $p_u + p_d + p_l + p_r = 1$. The process stops when it hits a boundary. We show five different examples of various possible scenarios. For the walks that start at a given point, we elaborate a Markov Chains method to find the probability of each boundary being hit first, and we then introduce a System of Equations method to find the probability of a single boundary being hit first from any possible starting point. In the appendix we provide the *Mathematica* code to reach the solutions to each example we give in the text. In Section 3.5, we modify the System of Equations method to find the average number of steps needed to hit a boundary and provide the solutions to some of the previous examples. At the end of this chapter, we analyze the number of steps needed for a Type I walk to hit a boundary from a particular starting point (h, k) . In this case, the Markov Chains method allows us to obtain a probability mass function (pmf) value $P(s_{h,k} = x)$, the probability of hitting a boundary after taking exactly

x steps, by taking the difference of cumulative distribution function (cdf) values: $P(s_{h,k} \leq x) - P(s_{h,k} \leq x - 1)$.

In Chapter 4, we study the Type II two-dimensional random walk. Basically, based on the Type I random walk with four boundaries described in Chapter 3, we change the four moving directions from up, down, left, or right to diagonal moving. The main difference we encounter with this process is that the four corners are able to be hit from an interior starting point. Similar to Chapter 3, we use the Markov Chains method to find the probability of each boundary or each corner being hit before the other three boundaries or corners from a specified starting point, and with the System of Equations method we can simultaneously obtain the probabilities of hitting a boundary or a corner from all possible starting points. But in this case we separate the corner points from the boundaries and manage each corner individually. In the last section of this chapter, we modify the System of Equations method to find numerical solution for the average number of steps needed for a Type II walk to hit a boundary or a corner from a starting point.

In Chapter 5, we consider a problem with two one-dimensional random walks, Process X and Process Y, starting at different heights with different boundaries, and we ask some questions that are of interest:

(i) What is the probability that Process X stops before Process Y?

(ii) What is the probability that they stop at the same time?

(iii) Given that Process X stops first, what is the probability that X has hit height m, which is the upper boundary of Process X ?

(iv) Given that they stop at the same time, what is the probability that they have both hit 0?

(v) What is the average number of steps needed for a process to stop? What is the average number of steps needed for *both* processes to stop?

Applying the results we have obtained previously, we are able to solve these problems by converting two processes of one-dimensional random walks into one process of a Type II two-dimensional random walk, assuming that the two processes move simultaneously. Using a particular example, we show step by step how to solve each question listed above.

In Chapter 6, we describe a boundary problem for two-sided downward-drifting Type II random walk. In this case, we consider even-valued starting coordinates and state a conjecture to estimate the probability of hitting the x-axis before hitting the y-axis. Using simulations, we test the accuracy of this conjecture with several examples.

We conclude the thesis with descriptions of other related problems that can be further studied.

Chapter 2

Boundary Problems for One-Dimensional Random Walk

In this chapter, we will first introduce a one-dimensional random walk having two boundaries and examine two known formulas. One formula is for the probability of the process hitting the upper boundary n before the lower boundary 0, and the other is for the average number of steps needed to hit a boundary. We will then generalize these results for any lower boundary of $m < n$. Secondly, we will extend the problem to a single boundary by letting one end go to infinity and analyze how we should amend these two formulas. Thirdly, we aim to derive the variance of the number of steps T needed for the height of a downward-drifting random walk to reach 0 from starting point k where $k \geq 0$. To reach our goal, we will derive the moment generating function of T, and use it to derive $E[T]$, $Var(T)$, and σ_T.

2.1. Background on One-dimensional Random Walk

Back in 1975, Chung [3] discusses the boundary problem for the one-dimensional random walk that begins at integer height k, where $0 \leq k \leq n$, and on each independent step the process either moves upward one unit, or downward one unit, with probabilities p and $q = 1 - p$, respectively, where $p \neq 0$. The process stops upon reaching height 0 or n. Using difference equations, he shows that the random walk will reach a boundary with probability 1, and that the probability ${}_kP_0^n$ of reaching height n before height 0 is

$$ {}_kP_0^n = \begin{cases} k/n, & \text{if } p = q \\ \dfrac{1-(q/p)^k}{1-(q/p)^n}, & \text{if } p \neq q. \end{cases} \tag{2.1.1} $$

In addition, the average number of steps ${}_kT_0^n$ to reach height n or height 0 is

$$ E\big[{}_kT_0^n\big] = \begin{cases} \dfrac{(n-k)k}{2p}, & \text{if } p = q \\ \dfrac{n}{p-q} \times \left(\dfrac{1-(q/p)^k}{1-(q/p)^n} \right) - \dfrac{k}{p-q}, & \text{if } p \neq q. \end{cases} \tag{2.1.2} $$

Now suppose a random walk begins at height k and ends at boundaries m or n, with $m \leq k \leq n$. By vertical translation, we can subtract m from each height and obtain equivalent results for a random walk that begins at height $k-m$ and stops at height 0 or $n-m$. That is, ${}_kP_m^n = {}_{k-m}P_0^{n-m}$ and $E\big[{}_kT_m^n\big] = E\big[{}_{k-m}T_0^{n-m}\big]$. For instance, if $k = 3$ with $m = -5$ and $n = 12$, then by subtracting $m = -5$ from each height, we shift the process up five units. This process is equivalent to starting at height 8 with boundaries of 0 and 17; i.e., ${}_3P_{-5}^{12} = {}_8P_0^{17}$. Thus we still get probability 1 of eventually hitting height m or height n. Moreover, simply by replacing k with $k-m$ and n with $n-m$ in (2.1.1), we obtain the probability ${}_kP_m^n$ of reaching the top boundary n before the bottom boundary m to be

$$ {}_kP_m^n = \begin{cases} \dfrac{k-m}{n-m}, & \text{if } p = q \\ \dfrac{1-(q/p)^{k-m}}{1-(q/p)^{n-m}}, & \text{if } p \neq q. \end{cases} \tag{2.1.3} $$

Then the proability ${}_kQ_m^n$ of reaching height m before height n is

$$
{}_kQ_m^n = 1 - {}_kP_m^n = \begin{cases} \dfrac{n-k}{n-m}, & \text{if } p = q \\[2ex] \dfrac{(q/p)^{k-m} - (q/p)^{n-m}}{1-(q/p)^{n-m}}, & \text{if } p \neq q. \end{cases} \tag{2.1.4}
$$

Replacing k with $k-m$ and n with $n-m$ in (2.1.2), we find the average of the number of steps ${}_kT_m^n$ to reach height n or height m to be

$$
E[{}_kT_m^n] = \begin{cases} \dfrac{(n-k)(k-m)}{2p}, & \text{if } p = q \\[2ex] \dfrac{n-m}{p-q} \times \left(\dfrac{1-(q/p)^{k-m}}{1-(q/p)^{n-m}} \right) - \dfrac{k-m}{p-q}, & \text{if } p \neq q. \end{cases} \tag{2.1.5}
$$

We note that if $p = 0$, then the process never moves up and will never hit n starting from $k < n$; thus, ${}_kP_m^n = 0$ and ${}_kQ_m^n = 1$. And if $p = 0$, the average number of steps needed to hit a boundary is simply the number of steps from the starting point $k < n$ to the bottom boundary m; i.e., $E[{}_kT_m^n] = k - m$.

We also note that if the process moves up with probability $p > 0$, moves down with probability $q \geq 0$, or remains at the same height with probability $r = 1 - p - q$, then it still reaches height 0 or n with probability 1. Moreover, Equations 2.1.1 to 2.1.5 still hold. In this case, all of Chung's derivations still hold with no changes.

2.2. A Single Boundary Problem

We now assume that a one-dimensional random walk has a fixed bottom boundary m , with no upper boundary. We first seek to find the probability that a one-dimensional random walk with single boundary will ever drop to height m

when starting at height k. To do so, we let ${}_kU_m^n$ denote the set of paths that cause one-dimensional random walks with two boundaries to reach height n before height m when starting at height k where $m \le k \le n$. These sets form a nested, decreasing sequence as n increases. Indeed, if a walk reaches height $n+1$ before height m, then it must have reached height n before height m. Since the probability of the intersection is the limit of the probabilities as $n \to \infty$ we have

$$P\left(\bigcap_{i=n}^{\infty} {}_kU_m^i\right) = \lim_{n\to\infty} P\big({}_kU_m^n\big) = \lim_{n\to\infty} {}_kP_m^n$$
$$= \begin{cases} \lim\limits_{n\to\infty} \dfrac{k-m}{n-m}, & \text{if } p = q \\[2ex] \lim\limits_{n\to\infty} \dfrac{1-(q/p)^{k-m}}{1-(q/p)^{n-m}}, & \text{if } p \ne q \end{cases}$$
$$= \begin{cases} 0, & \text{if } p \le q \\[2ex] 1-(q/p)^{k-m}, & \text{if } p > q. \end{cases}$$

We can interpret this limit as follows: we know that a walk will reach a boundary of n or m with probability 1. So we may let ${}_kW_m^n$ be the set of probability 0 consisting of the paths along which walks do not reach either boundary. Then the countable union $W = \bigcup_{i=n}^{\infty} {}_kW_m^i$ still has probability 0. We then exclude these paths and are left with W', those paths that do reach either n or m, where $P(W') = 1$. Within W', paths never drop to height m if and only if they belong to $\bigcap_{i=n}^{\infty} {}_kU_m^i$, which is the intersection of all the paths that hit n first and stop. Then $W \cup \left(\bigcap_{i=n}^{\infty} {}_kU_m^i\right)$ are all paths that never drop to height m. Because $P(W) = 0$, then $P\left(\bigcap_{i=n}^{\infty} {}_kU_m^i\right)$ by itself gives the probability that walks will not drop to height m, which has the value of the limit we just obtained above. By subtracting this value from 1, we have the probability ${}_kP_m$ that a one-sided one-dimensional random walk beginning at height k will drop to height m, for $m \leq k$, which is given by

$$
{}_kP_m = \begin{cases} 1, & \text{if } p \leq q \\ (q/p)^{k-m}, & \text{if } p > q. \end{cases} \tag{2.2.1}
$$

Secondly, we seek to find the average number of steps needed for a one-dimensional random walk with single boundary to decrease to height m when starting at height k. To do so, we again let ${}_kT_m^n$ be the number of steps needed for a two-sided one-dimensional random walk to reach a boundary of height n or m when starting at height k. Then the times $\{{}_kT_m^n\}_{n=k}^{\infty}$ form an increasing sequence such that ${}_kT_m^k \leq {}_kT_m^{k+1} \leq \cdots \leq {}_kT_m^n \leq \cdots$. Furthermore, we let ${}_kT_m$ be the number of steps needed for a one-sided one-dimensional random walk to decrease to height m. We note that if a walk ever hits height m, then ${}_kT_m$ is finite and by Equation 2.2.1, $P({}_kT_m < \infty) = {}_kP_m = 1$ for $p \leq q$. It is also clear that ${}_kT_m^n \leq {}_kT_m$ for all $n \geq k$, for if a walk reaches m before n, then the number of steps are the same thus the equal sign holds; but if the walk reaches n first, then ${}_kT_m^n < {}_kT_m$. Thus, ${}_kT_m < \infty$ with certainty for $p \leq q$, and ${}_kT_m^n$ increases to ${}_kT_m$ as $n \to \infty$. Applying the *Monotone Convergence Theorem* from analysis, we obtain $E[{}_kT_m]$, the average number of steps needed for a one-dimensional random walk with single boundary beginning at height k to decrease to height m, for $m < k$, which is

$$
\begin{aligned}
E[{}_kT_m] &= \lim_{n\to\infty} E[{}_kT_m^n] \\
&= \begin{cases} \lim\limits_{n\to\infty} \dfrac{(n-k)(k-m)}{2p}, & \text{if } p = q \\ \lim\limits_{n\to\infty} \left[\dfrac{n-m}{p-q} \times \left(\dfrac{1-(q/p)^{k-m}}{1-(q/p)^{n-m}}\right) - \dfrac{k-m}{p-q}\right], & \text{if } p \neq q. \end{cases} \\
&= \begin{cases} \infty, & \text{if } p \geq q \\ \dfrac{k-m}{q-p}, & \text{if } p < q. \end{cases}
\end{aligned}
\tag{2.2.2}
$$

2.3. The Stopping Time for the Single Boundary Problem

We now let $m = 0$ and let $T \equiv {}_kT_0$ be the number of steps needed for the height to reach 0 from the starting point k. For $p < q$, by the result in Equation 2.2.2, the random variable T has finite expectation given by $E[T] = k/(q-p)$, which is generally derived by using difference equations and the monotone convergence theorem as outlined in the previous sections. However this technique is not suitable for deriving the variance of T. According to Takacs in [5], due to DeMoivre's original work from 1711, for $T = k+2i$, where i is the number of the upward steps, we know the probability mass function (pmf) of T is given by

$$P(T = k+2i) = \frac{k}{k+2i}\binom{k+2i}{i} p^i q^{k+i}, \quad \text{for } i \geq 0,$$

Which gives the probability of hitting height 0 for the first time in $k+2i$ steps. To compute $E[T]$ using the pmf, we need to compute the following sum for the average of discrete random variables:

$$\begin{aligned} E[T] &= \sum_{i=0}^{\infty}(k+2i)\cdot P(T=k+2i) \\ &= \sum_{i=0}^{\infty} k\binom{k+2i}{i}p^i q^{k+i}. \end{aligned}$$

However, in advance, there is no known way to simplify the sum. That is why the difference equations technique used to derive $E[T]$ is so ingenious. Therefore we will use another technique to achieve our goal. We will first derive the moment generating function (mgf) of T, and then use a property of mgf to re-derive $E[T]$, which we expect to have the same result as the one we have in Equation 2.2.2. Also we can find a formula for $E[T^2]$ by using another property of the mgf. Having formulas for $E[T]$ and $E[T^2]$, we then can derive $Var(T)$ and σ_T.

We begin with finding the mgf of $T \equiv {}_kT_0$, the number of steps for a one-sided one-dimensional random walk to drop to 0 from height k. Then it takes at least k steps to drop to 0. If the walk goes up one unit, it needs to take another step to come down. So if we let i be the number of upward steps, then the total number of steps to drop down to 0 becomes $k+2i$ for some $i \geq 0$. For $q > p$, the paths drift downward and drop to height 0 almost surely. Hence the probabilities $P(T = k+2i)$ will sum to 1 over $i \geq 0$. That is,

$$1 = \sum_{i=0}^{\infty} \frac{k}{k+2i}\binom{k+2i}{i}p^i q^{k+i}. \tag{2.3.1}$$

Factoring out q^k from the summation and dividing it on both sides, we obtain

$$\frac{1}{q^k} = \sum_{i=0}^{\infty} \frac{k}{k+2i}\binom{k+2i}{i}(pq)^i. \tag{2.3.2}$$

Now we let $x = pq = (1-q)q$. Because $0 \le p < q \le 1$, we have $0 < x = (-q^2+q) < 1/4$. We also have $q^2 - q + x = 0$. Using the quadratic formula to solve for q we obtain $q = \frac{1 \pm \sqrt{1-4x}}{2}$. Considering q to be strictly greater than p, we must have

$$q = \frac{1+\sqrt{1-4x}}{2}, \tag{2.3.3}$$

and

$$p = 1 - q = 1 - \frac{1+\sqrt{1-4x}}{2} = \frac{1-\sqrt{1-4x}}{2}.$$

Then we can express

$$q - p = \frac{1+\sqrt{1-4x}}{2} - \frac{1-\sqrt{1-4x}}{2} = \sqrt{1-4x} = \sqrt{1-4pq}. \tag{2.3.4}$$

So Equation 2.3.2 can be written in terms of x, using Equation 2.3.3, as follows:

$$\frac{2^k}{(1+\sqrt{1-4x})^k} = \sum_{i=0}^{\infty} \frac{k}{k+2i}\binom{k+2i}{i}x^i. \tag{2.3.5}$$

We note that with the replacement of $x = pq$ and combining Equations 2.3.3 & 2.3.5, we obtain another equation, which will be convenient for us to derive the mgf later:

$$\frac{2^k}{(1+\sqrt{1-4pq})^k} = \sum_{i=0}^{\infty} \frac{k}{k+2i}\binom{k+2i}{i}(pq)^i = \frac{1}{q^k}. \tag{2.3.6}$$

Also, using $x = pqe^{2t}$ with $0 < pqe^{2t} < 1/4$, we have

$$\sum_{i=0}^{\infty} \frac{k}{k+2i}\binom{k+2i}{i}(pqe^{2t})^i = \frac{2^k}{(1+\sqrt{1-4pqe^{2t}})^k}. \tag{2.3.7}$$

When $t = 0$ in Equation 2.3.7, we get the result in Equation 2.3.6, which is simply $1/q^k$. The same technique can be done with the expected value:

$$\frac{k}{q-p} = E[T] = \sum_{i=0}^{\infty}(k+2i)P(T = k+2i) = \sum_{i=0}^{\infty} k\binom{k+2i}{i}(pq)^i q^k.$$

Dividing by q^k on both sides, we have

$$\frac{k}{(q-p)q^k} = \sum_{i=0}^{\infty} k\binom{k+2i}{i}(pq)^i. \tag{2.3.8}$$

We again let $x = pq$, then $q = \frac{1+\sqrt{1-4x}}{2}$ and $q - p = \sqrt{1-4x} = \sqrt{1-4pq}$. By substitution, we can rewrite Equation 2.3.8 in terms of x:

$$\frac{k2^k}{(\sqrt{1-4x})(1+\sqrt{1-4x})^k} = \sum_{i=0}^{\infty} k\binom{k+2i}{i}x^i.$$

We can also use $x = pqe^{2t}$ for $0 < pqe^{2t} < 1/4$ and obtain

$$\frac{k2^k}{(\sqrt{1-4pqe^{2t}})(1+\sqrt{1-4pqe^{2t}})^k} = \sum_{i=0}^{\infty} k\binom{k+2i}{i}(pqe^{2t})^i. \tag{2.3.9}$$

When $t = 0$ and $\sqrt{1-4pq} = q - p$, then Equation 2.3.9 reduces to Equation 2.3.8. We can now derive the mgf of T.

Theorem 2.3.1. Let $0 < p < q < 1$. The moment generating function of T, the number of steps for a one-dimensional random walk to drop to its single boundary

0 from height k, is given by

$$M_T(t) = \frac{e^{kt} q^k 2^k}{(1 + \sqrt{1 - 4pqe^{2t}})^k},$$

for $0 < pqe^{2t} < 1/4$.

Proof. The mgf $M_T(t)$ is given by

$$\begin{aligned} M_T(t) &= E[e^{Tt}] \\ &= \sum_{i=0}^{\infty} e^{(k+2i)t} P(T = k + 2i) \\ &= \sum_{i=0}^{\infty} e^{kt} e^{(2t)i} \left(\frac{k}{k+2i} \binom{k+2i}{i} p^i q^{k+i} \right) \\ &= e^{kt} q^k \sum_{i=0}^{\infty} \frac{k}{k+2i} \binom{k+2i}{i} (pqe^{2t})^i. \end{aligned}$$

By the result in Equation 2.3.7, we have, for $0 < pqe^{2t} < 1/4$,

$$\begin{aligned} M_T(t) &= e^{kt} q^k \sum_{i=0}^{\infty} \frac{k}{k+2i} \binom{k+2i}{i} (pqe^{2t})^i \\ &= \frac{e^{kt} q^k 2^k}{(1 + \sqrt{1 - 4pqe^{2t}})^k}. \end{aligned}$$

□

We also note that when $t = 0$, by Equation 2.3.1 we have the result

$$\begin{aligned} M_T(0) &= q^k \sum_{i=0}^{\infty} \frac{k}{k+2i} \binom{k+2i}{i} (pq)^i \\ &= \frac{q^k 2^k}{(1 + \sqrt{1 - 4pq})^k}. \\ &= 1. \end{aligned}$$

Corollary 2.3.1. The first and the second derivatives of $M_T(t)$ are:

$$M_T'(t) = M_T(t) \cdot \frac{k}{\sqrt{1-4pqe^{2t}}}$$

and

$$M_T''(t) = M_T(t) \cdot \frac{k^2\sqrt{1-4pqe^{2t}} + k(4pqe^{2t})}{(1-4pqe^{2t})^{3/2}}$$

respectively.

Proof. Taking the first derivative of $M_T(t)$ with respect to t, we have:

$$\begin{aligned}
M_T'(t) &= \frac{(1+\sqrt{1-4pqe^{2t}})^k(e^{kt}q^k2^k)' - e^{kt}q^k2^k\left((1+\sqrt{1-4pqe^{2t}})^k\right)'}{(1+\sqrt{1-4pqe^{2t}})^{2k}} \\
&= \frac{e^{kt}q^k2^k}{(1+\sqrt{1-4pqe^{2t}})^k} \cdot k\left(1 + \frac{4pqe^{2t}}{(1+\sqrt{1-4pqe^{2t}})\sqrt{1-4pqe^{2t}}}\right) \\
&= M_T(t) \cdot k\left(\frac{\sqrt{1-4pqe^{2t}}+1}{\sqrt{1-4pqe^{2t}}+(1-4pqe^{2t})}\right) \\
&= M_T(t) \cdot k\left(\frac{\sqrt{1-4pqe^{2t}}+1}{\sqrt{1-4pqe^{2t}}+(1-4pqe^{2t})}\right) \cdot \frac{\sqrt{1-4pqe^{2t}}-(1-4pqe^{2t})}{\sqrt{1-4pqe^{2t}}-(1-4pqe^{2t})} \\
&= M_T(t) \cdot k\left(\frac{\sqrt{1-4pqe^{2t}}(4pqe^{2t})}{(4pqe^{2t})(1-4pqe^{2t})}\right) \\
&= M_T(t) \cdot \frac{k}{\sqrt{1-4pqe^{2t}}}.
\end{aligned}$$

So we have the first derivative of $M_T(t)$ as

$$M_T'(t) = M_T(t) \cdot \frac{k}{\sqrt{1-4pqe^{2t}}}.$$

Taking the derivative of $M'_T(t)$ using the product rule we have:

$$\begin{aligned}
M''_T(t) &= M'_T(t) \cdot \frac{k}{\sqrt{1-4pqe^{2t}}} + M_T(t) \cdot \left(\frac{k}{\sqrt{1-4pqe^{2t}}} \right)' \\
&= M_T(t) \cdot \frac{k^2}{1-4pqe^{2t}} + M_T(t) \cdot \frac{k(4pqe^{2t})}{(1-4pqe^{2t})^{3/2}} \\
&= M_T(t) \cdot \frac{k^2\sqrt{1-4pqe^{2t}} + k(4pqe^{2t})}{(1-4pqe^{2t})^{3/2}}.
\end{aligned}$$

□

Theorem 2.3.2. Let $0 < p < q < 1$, the variance of $T \equiv {}_kT_0$, the number of steps needed for a one-dimensional random walk to drop to its single boundary 0 from height k is given by

$$Var(T) = \frac{4kpq}{(q-p)^3}.$$

Proof. Properties of moment generating functions state that $M'_T(0) = E[T]$, and $M_T(0) = 1$. Now using $q - p = \sqrt{1-4pq}$, we re-obtain the formula for $E[T]$, which is the same as we have expected:

$$\begin{aligned}
E[T] &= M'_T(0) \\
&= M_T(0) \cdot \frac{k}{\sqrt{1-4pq}} \\
&= \frac{k}{q-p}.
\end{aligned}$$

Another property of moment generating functions states that $M_T''(0) = E[T^2]$. Again with the fact that $M_T(0) = 1$ and $q - p = \sqrt{1-4pq}$, we have

$$\begin{aligned} E[T^2] &= M_T''(0) \\ &= M_T(0) \cdot \frac{k^2\sqrt{1-4pq} + k(4pq)}{(1-4pq)^{3/2}} \\ &= \frac{k^2(q-p) + 4kqp}{(q-p)^3}. \end{aligned}$$

Now we derive the formula for $Var(T)$ by

$$\begin{aligned} Var(T) &= E[T^2] - (E[T])^2 \\ &= \frac{k^2(q-p) + 4kpq}{(q-p)^3} - \left(\frac{k}{q-p}\right)^2 \\ &= \frac{4kpq}{(q-p)^3}. \end{aligned}$$

$\square$

Taking the square root, we have

Corollary 2.3.2. The standard deviation of $T \equiv {}_kT_0$, the number of steps needed for a one-dimensional random walk drop to its single boundary 0 from height k is given by

$$\sigma_T = 2\sqrt{\frac{kpq}{(q-p)^3}}.$$

Chapter 3

Boundary Problems for Type I Two-Dimensional Random Walk

3.1. Introduction

We now study a two-dimensional random walk that begins at integer-valued coordinates (h,k), where $0 \le h \le m$ and $0 \le k \le n$. On each step, the random walk moves one unit either up, down, left, or right with probabilities $p_u \neq 0$, $p_d \neq 0$, $p_l \neq 0$, $p_r \neq 0$, respectively, where $p_u + p_d + p_l + p_r = 1$. We call this process a random walk of Type I. The four boundaries are the lines $x = 0$ (the y-axis), $y = 0$ (the x-axis), $x = m$ and $y = n$. The process stops when it hits a boundary.

As an example, we let $m = 5$, $n = 5$, $(h,k) = (2,3)$, with $p_u = 0.30$, $p_d = 0.25$, $p_l = 0.35$ and $p_r = 0.10$. A possible specific path can be: starting at point $(2,3)$, go down 1 unit, go right 1 unit, go down another 1 unit, go right another 1 unit, finally go down 1 more unit so that it hits the lower boundary and stops. This path is shown below.

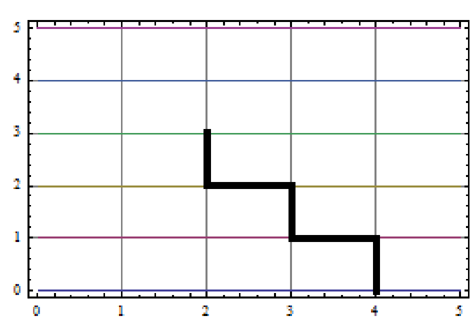

Figure 3.1. A Type I Path

The probability of this path is

$$p_d \cdot p_r \cdot p_d \cdot p_r \cdot p_d = 0.25 \times 0.10 \times 0.25 \times 0.10 \times 0.25 = 1.5625 \times 10^{-4}.$$

Another possible specific path can be starting at point $(2, 3)$, move up 1 unit, move right 1 unit, move down 2 units, move left 1 unit, move down 1 unit, move right 2 units, finally move up 4 units so that it hits the upper boundary and stops. This path is shown in Figure 3.2. The probability of this path is

$$0.30 \times 0.10 \times 0.25^2 \times 0.35 \times 0.25 \times 0.10^2 \times 0.30^4 = 1.3289 \times 10^{-8}.$$

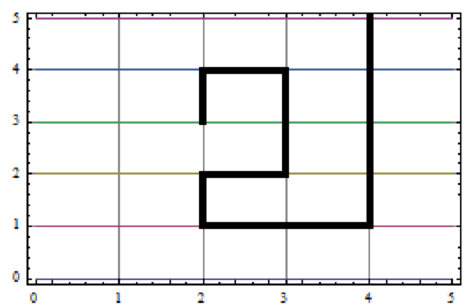

FIGURE 3.2. Another Type I Path

We note that unless we start at one of these corners: $(0, 0)$, $(0, n)$, (m, n), and $(m, 0)$, it is impossible to hit the four corners since we do not allow diagonal movements in this type of walk.

In general, there are infinitely many possible paths, and each single path has a distinct probability. Our goal is to determine the overall probabilities of hitting each boundary first from the initial starting point. In [2], Neal shows how to use a matrix method to solve a boundary problem for one-dimensional random walk and simulates the results with *Mathematica*. In [1], Neal discusses a two-sided boundary problem for Type I two-dimensional random walk in the case where $p_d > p_u$ and $p_l > p_r$. Furthermore, he derives a lengthy formula to compute the probability of such Type I two-dimensional random walk hitting the x-axis before hitting the y-axis. In this chapter, we shall use Neal's technique similar to those in [2] to find the numerical solution for the probability of a four-sided Type I walk hitting one boundary before hitting the other three from a given starting point. Also we will introduce the System of Equations method to find the numerical solution for the probability of a given boundary being hit first and the average number of steps needed to hit a boundary from any possible starting point.

3.2. Background on Random Walk

We recall that Chung [3] shows a one-dimensional random walk with two boundaries 0 and n will reach height 0 or n with probability 1, and that the probability of reaching height n before height 0 is shown in 2.1.1. If we only consider the upward and downward movements of a Type I two-dimensional walk, then a one-dimensional walk begins at the vertical height k and moves up with probability $p = p_u$, down with probability $q = p_d$, or stays at the same height with probability $r = p_l + p_r$. So there is probability 1 that the Type I two-dimensional walk hit a lower or upper boundary if we do not stop when hitting either side boundary. In this case, the probability of hitting the upper boundary first is

$$
{}_kP_0^n = \begin{cases} k/n, & \text{if } p_u = p_d \\ \dfrac{1-(p_d/p_u)^k}{1-(p_d/p_u)^n}, & \text{if } p_u \neq p_d. \end{cases}
$$

Similarly, if we only stop upon hitting the side boundary, then there is probability 1 of hitting a side boundary, and the probability of hitting the right side first is

$$
{}_hP_0^m = \begin{cases} h/m, & \text{if } p_r = p_l \\ \dfrac{1-(p_l/p_r)^h}{1-(p_l/p_r)^m}, & \text{if } p_r \neq p_l. \end{cases}
$$

We note that there are many paths that never hit a boundary. Provided that $1 < k < n-1$ and $1 < h < m-1$, we can always create paths that stay bounded within $h+1, h-1, k+1$, and $k-1$, and which never hit any boundary. However, these paths have probability 0 because, as discussed above, there is probability 1 of eventually hitting a boundary.

In fact, the set of paths that never hit a boundary is uncountable. To see this, we can look at paths that only move up and down (i.e., never sideways) but stay bounded between $k+1$ and $k-1$. So they must move up/down, or down/up on two consecutive steps. A specific possible path can be moving up/down or down/up continuously. But a slight change will make a different path. For example, a path that keeps moving down/up except the second step moving up/down is different from the path that keeps moving down/up except the third step that moves up/down. If we let up/down be 0 and down/up be 1, then all sequences of

0's and 1's are created such as

$$1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...$$

$$1, 0, 0, 1, 0, 1, 1, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, ...$$

The collection of all such sequences of 0's and 1's form the binary version of the interval $[0, 1]$. Any $x \in [0, 1]$ can be written as $x = \frac{a_1}{2^1} + \frac{a_2}{2^2} + \frac{a_3}{2^3} + \cdots$, where all a_i are 0 or 1. For example,

$$0 = (0, 0, 0, 0, 0, \cdots)_2$$

$$1/2 = (1, 0, 0, 0, 0, \cdots)_2$$

$$1 = (1, 1, 1, 1, 1, 1, \cdots)_2$$

$$\vdots$$

Because the interval $[0, 1]$ is uncountable, the collection of these paths that never move sideways and stay bounded between $k + 1$ and $k - 1$ is uncountable. As we mentioned previously, the collection of paths that will hit a boundary has probability 1. Thus the set of paths that never hit a boundary is 0. And we can directly show that the probability of any such path is 0. In particular, the probability of moving up/down or down/up is

$$p_u p_d + p_d p_u = 2 p_u p_d.$$

Then the probability of this event happening infinitely many times is

$$\prod_{i=1}^{\infty} 2p_u p_d = 2p_u p_d \times 2p_u p_d \times 2p_u p_d \times \cdots \tag{3.2.1}$$

We know that $0 < p_u + p_d < 1$, which implies $p_u < 1 - p_d$. Then we have

$$p_u p_d < (1 - p_d) \cdot p_d = p_d - p_d^2,$$

which has a maximum value of $1/4$. Hence, multiplying $2p_u p_d$ by itself ad infinitum in (3.2.1) will yield 0. Similarly, the collection of paths that only move left and right, but stay bounded between $h - 1$ and $h + 1$, are uncountable and have probability 0.

3.3. The Markov Chains Solution

We want the probability of a Type I two-dimensional random walk hitting each boundary first from an initial starting point. Some cases are obvious, but most of them are not. Here are some examples.

Example 3.3.1. If $m = 6$, $n = 6$, $(h, k) = (3, 3)$, and $p_l = p_r = p_u = p_d = 0.25$, then by symmetry, there are equal chances of hitting one boundary before the other three, with each probability being 0.25. However, not each point on one boundary has the same probability of being hit first. To see this, we can look at the five points on y-axis: $(0, 1), (0, 2), (0, 3), (0, 4), (0, 5)$. We would expect that $(0, 3)$ is the most likely to be hit first. And by symmetry, $(0, 2)$ and $(0, 4)$ should have equal probabilities of being hit first, so should $(0, 1)$ and $(0, 5)$.

Example 3.3.2. With the same boundaries and the same starting point as in Example 3.3.1, but with $p_l = 0.4 = p_r$ and $p_u = 0.10 = p_d$, then it will be more likely to hit the side boundaries first than the upper or the lower boundaries. Also, hitting the left side and the right side first must have equal probabilities by symmetry, and so must the upper and the lower boundaries. But it is obvious that not every point on one boundary will have the same probability of being hit first.

Example 3.3.3. If $(h,k) = (5,5)$, with $m = 6 = n$ and equal probabilities of moving in the four directions, then a path will most likely hit the upper boundary or the right side first because the starting point is closer to these two boundaries.

Example 3.3.4. With the same boundaries and the same starting point as in Example 3.3.3, but with $p_l = 0.45$, $p_r = 0.02$, $p_u = 0.03$, $p_d = 0.50$, then we cannot tell which boundary is most likely to be hit first.

Example 3.3.5. If $(h,k) = (2,3)$, $m = 6$, $n = 8$, $p_l = 0.21$, $p_r = 0.23$, $p_u = 0.29$, $p_d = 0.27$, then there is no symmetry at all to help determine which boundary is most likely to be hit first.

General cases like in Example 3.3.4 and Example 3.3.5 are not obvious, and require a method of solution. To solve these boundary problems, we first let $A = (a_{ij,kl})$, for $0 \le i \le m$, $0 \le j \le n$, $0 \le k \le m$, $0 \le l \le n$ be a quadruple-indexed matrix of transition probabilities having dimension $(m+1)(n+1) \times (m+1)(n+1)$. The term $a_{ij,kl}$ gives the probability of a Type I walk moving from coordinates (i,j) to coordinates (k,l) on each step, where $a_{ij,ij} = 1$ if (i,j) is on a boundary. For instance, if $m = 3$ and $n = 4$, then $A = (a_{ij,kl})$ is a 20×20 matrix. If $(i,j) = (0,0)$,

$(0,1)$, $(0,2)$, $(0,3)$, $(0,4)$, $(1,0)$, $(1,4)$, $(2,0)$, $(2,4)$, $(3,0)$, $(3,1)$, $(3,2)$, $(3,3)$, or $(3,4)$, which are the points on the four boundaries, then $a_{ij,ij} = 1$; i.e., if the initial position is on a boundary point, then it stays on that point with probability 1. It goes nowhere. So the probabilities $a_{ij,kl}$ are 0 for other coordinates (k,l). But if we do not start on a boundary, then we have positive probabilities of moving to four other points. For example, if $(i,j) = (1,1)$, then the probability of moving to $(0,1)$ is $a_{11,01} = p_l$. Likewise $a_{11,10} = p_d$, $a_{11,12} = p_u$, and $a_{11,21} = p_r$. Below is the complete 20×20 transition matrix A for boundaries $m = 3$ and $n = 4$.

	00	01	02	03	04	10	11	12	13	14	20	21	22	23	24	30	31	32	33	34
00	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
01	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
02	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
03	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
04	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	p_l	0	0	0	p_d	0	p_u	0	0	0	p_r	0	0	0	0	0	0	0	0
12	0	0	p_l	0	0	0	p_d	0	p_u	0	0	0	p_r	0	0	0	0	0	0	0
13	0	0	0	p_l	0	0	0	p_d	0	p_u	0	0	0	p_r	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	p_l	0	0	0	p_d	0	p_u	0	0	0	p_r	0	0	0
22	0	0	0	0	0	0	0	p_l	0	0	0	p_d	0	p_u	0	0	0	p_r	0	0
23	0	0	0	0	0	0	0	0	p_l	0	0	0	p_d	0	p_u	0	0	0	p_r	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

TABLE 3.1. 20×20 Transition Matrix A for $m = 3$ and $n = 4$

The terms on the top and the left-side are simply place holders that tell the possible coordinates. The place holders on the left represent the previous state, and the place holders on the top represent the possible coordinates after another step is taken.

Secondly, we let $B = (b_{1,ij})$, for $0 \le i \le m$, $0 \le j \le n$ be the $1 \times (m+1)(n+1)$ initial state matrix that designates the initial position of a Type I walk. Then $b_{1,hk} = 1$, and $b_{1,ij} = 0$ when $i \ne h$ or $j \ne k$. In our example with $m = 3$ and $n = 4$, if $(h,k) = (2,3)$, then $B = (b_{1,ij})$ is a 1×20 matrix and $b_{1,ij} = 0$ for all i, j except $b_{1,23} = 1$. Matrix B is shown below, where the top line are place holders.

00	01	02	03	04	10	11	12	13	14	20	21	22	23	24	30	31	32	33	34
0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

TABLE 3.2. 1×20 Initial State Matrix B for $m = 3$ and $n = 4$

To find the probabilities of having all possible positions after x steps, we multiply $B \times A^x$. Letting x be"large" such as $x = 600$, we obtain the final probability states. We let $C_x = B \times A^x = (c_{1,ij,x})$, for $0 \le i \le m$, $0 \le j \le n$. Then C_x is a $1 \times (m+1)(n+1)$ matrix, where $c_{1,ij,x}$ gives probabilities of being at (i,j) after x steps.

Definition 3.3.1. Let X be a Type I two-dimensional random walk. The event of being on the left boundary (i.e., the y-axis) after x steps will be denoted by L_x. Likewise, the events of being on the bottom boundary (the x-axis), the upper boundary ($y = n$), and the right boundary ($x = m$) after x steps will be denoted by D_x, U_x, and R_x, respectively.

Finally, we let $P(L_x)$, $P(D_x)$, $P(U_x)$, and $P(R_x)$ be the probabilities of hitting the left boundary, the lower boundary, the upper boundary, and the right boundary after x steps, respectively.

We obtain the probability of a Type I two-dimensional random walk being at the left boundary after x steps by taking the sum of $c_{1,0j,x}$, where j is from 0 to n. Here $c_{1,0j,x}$ represents the probabilities of hitting $(0, j)$ first after x steps. Likewise, taking the sum of $c_{1,i0,x}$ we have the probability of a Type I walk being at the bottom boundary after x steps, where $1 \leq i \leq m-1$; the sum of $c_{1,in,x}$ where $1 \leq i \leq m-1$ is the probability of being at the upper boundary after x steps; and the sum of $c_{1,mj,x}$ where $0 \leq j \leq n$ is the probability of being at the right boundary after x steps. We include $(0,0)$, $(0,n)$ to the left boundary, and (m,n), $(m,0)$ to the right boundary. Although from an interior starting point, a Type I walk will never hit the four corners, it is possible to start at the corners. Now we can state the theorem as follows.

Theorem 3.3.1. The probabilities of a Type I two-dimensional random walk being at each boundary after x steps are given by

$$(a) \qquad P(L_x) = \sum_{j=0}^{n} c_{1,0j,x} \qquad\qquad (b) \qquad P(D_x) = \sum_{i=1}^{m-1} c_{1,i0,x}$$

$$(c) \qquad P(U_x) = \sum_{i=1}^{m-1} c_{1,in,x} \qquad\qquad (d) \qquad P(R_x) = \sum_{j=0}^{n} c_{1,mj,x}.$$

By taking the limit on the sum of the corresponding $c_{1,ij,x}$ when $x \to \infty$, we can obtain our desired probabilities of a Type-I walk hitting each boundary first. Thus, we can state:

Theorem 3.3.2. The probabilities of a Type-I two-dimensional random walk hitting each boundary first are given by

$$(a) \qquad P(L_\infty) = \lim_{x\to\infty} \sum_{j=0}^{n} c_{1,0j,x} \qquad\qquad (b) \qquad P(D_\infty) = \lim_{x\to\infty} \sum_{i=1}^{m-1} c_{1,i0,x}$$

$$(c) \qquad P(U_\infty) = \lim_{x\to\infty} \sum_{i=1}^{m-1} c_{1,in,x} \qquad\qquad (d) \qquad P(R_\infty) = \lim_{x\to\infty} \sum_{j=0}^{n} c_{1,mj,x}.$$

We call this method we use to achieve the probabilities of each boundary being hit first from a given starting point a Markov Chains method, for the systems we describe above follow a chain of linked events where what happens next depends only on the current state of the system, which is Markov property. Usually, if m, n, and x are large, we are not able to make the matrices and compute the various probabilities by hand. But we can use *Mathematica* for this computation. Now we can quickly obtain the solutions to the previous examples using *Mathematica* (see Appendix A for the code).

Example 3.3.1. The Markov Chains Solution. We have $m = n = 6$, $(h, k) = (3, 3)$, and $p_l = p_r = p_u = p_d = 0.25$. Using $x \geq 600$ steps, we have the results:

(i,j)	$(0,0)$	$(0,1)$	$(0,2)$	$(0,3)$	$(0,4)$	$(0,5)$	$(0,6)$	$P(L_x)$
$c_{1,ij,x}$	0	0.0288	0.0577	0.0770	0.0577	0.0288	0	0.25

This is what we have expected:

(i) $P(L_x) = 0.25$;

(ii) the symmetric points $(0,1)$ & $(0,5)$, $(0,2)$ & $(0,4)$ have the same probabilities of being hit first;

(iii) the corner points $(0,0)$ & $(0,6)$ have probability 0 of being hit first since the process cannot reach these points from $(3,3)$;

(iv) point $(0,3)$ has the greatest probability of being hit first among the left boundary points because it has the shortest distance from $(3,3)$.

The results are the same for $P(D_x)$, $P(U_x)$, $P(R_x)$.

Example 3.3.2. The Markov Chains Solution. With $m = n = 6$ and starting at $(3,3)$, but with $p_l = 0.4 = p_r$, and $p_u = 0.10 = p_d$, we have the following results for hitting the left side first using a maximum of $x \geq 105$ steps.

(i,j)	$(0,0)$	$(0,1)$	$(0,2)$	$(0,3)$	$(0,4)$	$(0,5)$	$(0,6)$	$P(L_x)$
$c_{1,ij,x}$	0	0.0325	0.0922	0.1880	0.0922	0.0325	0	0.4374

Similar to Example 3.3.1, the results in this example show the characteristics (ii), (iii), and (iv), except that hitting the left side boundary first $P(L_x)$ has a higher probability than that of hitting the right side $P(R_x)$. Below are the results for $P(U_x)$, which are the same as for $P(D_x)$:

(i,j)	$(0,6)$	$(1,6)$	$(2,6)$	$(3,6)$	$(4,6)$	$(5,6)$	$(6,6)$	$P(U_x)$
$c_{1,ij,x}$	0	0.0081	0.0145	0.0174	0.0145	0.0081	0	0.0626

Example 3.3.3. The Markov Chains Solution. With boundaries $m = n = 6$ and the probabilities $p_l = p_r = p_u = p_d = 0.25$, but $(h,k) = (5,5)$, we have the following results for hitting the left side first using a maximum of $x \geq 95$ steps.

(i,j)	$(0,6)$	$(1,6)$	$(2,6)$	$(3,6)$	$(4,6)$	$(5,6)$	$(6,6)$	$P(L_x)$
$c_{1,ij,x}$	0	0.0027	0.0055	0.0079	0.0088	0.0064	0	0.0313

We note that there is no symmetry characteristic on this boundary. Nevertheless, in this example, since the distance from the starting point to the left side boundary is the same as the distance to the lower boundary, and we have even probabilities to move toward either direction, we should have $P(L_x) = P(D_x)$. Likewise, we have $P(U_x) = P(R_x)$, with the results as following:

(i,j)	$(0,6)$	$(1,6)$	$(2,6)$	$(3,6)$	$(4,6)$	$(5,6)$	$(6,6)$	$P(U_x)$
$c_{1,ij,x}$	0	0.0064	0.0169	0.0402	0.1035	0.3017	0	0.4687

Example 3.3.4. The Markov Chains Solution. With $m = 6$, $n = 6$, and $(h,k) = (5,5)$, but $p_l = 0.45$, $p_r = 0.02$, $p_u = 0.03$, $p_d = 0.50$, the results for hitting the left, right, upper and lower boundary first respectively are

(i,j)	$(0,0)$	$(0,1)$	$(0,2)$	$(0,3)$	$(0,4)$	$(0,5)$	$(0,6)$	$P(L_\infty)$
$c_{1,ij,x}$	0	0.1095	0.1111	0.0929	0.0608	0.0243	0	0.3986

(i,j)	$(6,0)$	$(6,1)$	$(6,2)$	$(6,3)$	$(6,4)$	$(6,5)$	$(6,6)$	$P(R_\infty)$
$c_{1,ij,x}$	0	0.0015	0.0029	0.0057	0.0106	0.0205	0	0.0412

(i,j)	$(0,6)$	$(1,6)$	$(2,6)$	$(3,6)$	$(4,6)$	$(5,6)$	$(6,6)$	$P(U_\infty)$
$c_{1,ij,x}$	0	0.0016	0.0033	0.0069	0.0144	0.0308	0	0.0570

(i,j)	$(0,0)$	$(1,0)$	$(2,0)$	$(3,0)$	$(4,0)$	$(5,0)$	$(6,0)$	$P(D_\infty)$
$c_{1,ij,x}$	0	0.1216	0.1332	0.1222	0.0878	0.0384	0	0.5032

In this example, though the process starts at a point which is much closer to the upper boundary than the lower one, and is closer to the right side than the left side, the greater probability of moving down than moving up causes $P(D_\infty)$ to be greater than $P(U_\infty)$. Likewise, $P(L_\infty) > P(R_\infty)$.

Example 3.3.5. The Markov Chains Solution. With $(h,k) = (2,3)$, $m = 6$, $n = 8$, and $p_l = 0.21$, $p_r = 0.23$, $p_u = 0.29$, $p_d = 0.27$, using $x \geq 1000$ steps, we have the following results for hitting the left side boundary first:

(i,j)	$(0,0)$	$(0,1)$	$(0,2)$	$(0,3)$	$(0,4)$	$(0,5)$	$(0,6)$	$(0,7)$	$(0,8)$	$P(L_x)$
$c_{1,ij,x}$	0	0.0335	0.0764	0.01237	0.0888	0.0546	0.0307	0.0139	0	0.4216

To apply the Markov Chains method, we need to know exactly which coordinate the walk starts, while the other method that we are going to introduce next does not have this requirement.

3.4. The System of Equations Solution

In [2], Neal uses a system of linear equations to simultaneously solve for the probabilities of a one-dimensional random walk reaching one boundary before the other from all possible starting heights. Using the same technique, we can simultaneously obtain the probabilities of a Type I two-dimensional random walk hitting one boundary before the other three from all possible starting points. We will call this method a System of Equations method. First, we look at one boundary, say the left boundary. We are going to solve for the probabilities of hitting the left boundary first, $P(L_\infty)$, starting at all possible points within the boundaries by setting up a system of equations.

We let $x_{i,j}$ be the probability of hitting the left boundary first when starting at (i,j), for $0 \leq i \leq m$, $0 \leq j \leq n$. Then we know that $x_{0,j} = 1$ for all j (because if we start at the left boundary then we already have hit the left boundary). Also $x_{m,j} = 0$ for all j, and $x_{i,0} = 0 = x_{i,n}$ for $i \geq 1$ (because if we start at the right, the lower, or the upper boundary then we are not going to move and will never hit the left boundary.) Otherwise, by the Law of Total Probability, we have

$$x_{i,j} = p_l \cdot x_{i-1,j} + p_d \cdot x_{i,j-1} + p_u \cdot x_{i,j+1} + p_r \cdot x_{i+1,j},$$

which can be re-written as

$$p_l \cdot x_{i-1,j} + p_d \cdot x_{i,j-1} - x_{i,j} + p_u \cdot x_{i,j+1} + p_r \cdot x_{i+1,j} = 0.$$

For instance, if $m = 2$ and $n = 3$, then $x_{0,j} = 1$ and $x_{2,j} = 0$ for $0 \leq j \leq 3$, and $x_{i,0} = 0 = x_{i,3}$ for $i = 1, 2$. Otherwise, we have

$$p_l \cdot x_{0,1} + p_d \cdot x_{1,0} - x_{1,1} + p_u \cdot x_{1,2} + p_r \cdot x_{2,1} = 0$$

and

$$p_l \cdot x_{0,2} + p_d \cdot x_{1,1} - x_{1,2} + p_u \cdot x_{1,3} + p_r \cdot x_{2,2} = 0.$$

Then we have a 12×12 matrix of coefficients, namely G, a 12×1 matrix of constants, namely H, and a system of equations $GX = H$. We simply solve for X by $X = G^{-1}H$. The augmented matrix of the system of equations is shown below, where the top row are the indices of the unknowns $x_{i,j}$:

00	01	02	03	10	11	12	13	20	21	22	23	
1	0	0	0	0	0	0	0	0	0	0	0	1
0	1	0	0	0	0	0	0	0	0	0	0	1
0	0	1	0	0	0	0	0	0	0	0	0	1
0	0	0	1	0	0	0	0	0	0	0	0	1
0	0	0	0	1	0	0	0	0	0	0	0	0
0	p_l	0	0	p_d	-1	p_u	0	0	p_r	0	0	0
0	0	p_l	0	0	p_d	-1	p_u	0	0	p_r	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	1	0

TABLE 3.3. Augmented Matrix for Boundaries $m = 2$ and $n = 3$

We note that this matrix of coefficients is similar to the matrix of transition probabilities used in the Markov Chains method, with one difference: the entries between p_d and p_u on each row that has them are -1 instead of 0. Applying the System of Equations method to solve $GX = H$, we can obtain the probabilities of hitting one boundary first from all possible starting points simultaneously; however, we will not obtain the individual probabilities of hitting each specific boundary point first.

Now we are going to use *Mathematica* to redo Examples 3.3.1, 3.3.2 and 3.3.5 with the System of Equations method, while considering hitting the left boundary first from all possible starting points. See Appendix B for the code. In each example, for all the starting points on the left boundary, $x_{i,j} = 1$, and for all the starting points on the lower boundary, upper boundary and the right boundary, $x_{i,j} = 0$. Therefore, we are interested in determining $x_{i,j}$ when starting at the interior points.

Example 3.3.1. The System of Equations Solution. With $m = n = 6$ and $p_l = p_r = p_u = p_d = 0.25$, we have the following probabilities of hitting the left boundary first starting at each interior point (i, j):

$$
\begin{array}{lllll}
x_{1,1} = 0.4687 & x_{1,2} = 0.6292 & x_{1,3} = 0.6694 & x_{1,4} = 0.6292 & x_{1,5} = 0.4687 \\
x_{2,1} = 0.2455 & x_{2,2} = 0.3788 & x_{2,3} = 0.4193 & x_{2,4} = 0.3788 & x_{2,5} = 0.2455 \\
x_{3,1} = 0.1346 & x_{3,2} = 0.2212 & x_{3,3} = 0.2500 & x_{3,4} = 0.2212 & x_{3,5} = 0.1346 \\
x_{4,1} = 0.0718 & x_{4,2} = 0.1212 & x_{4,3} = 0.1384 & x_{4,4} = 0.1212 & x_{4,5} = 0.0718 \\
x_{5,1} = 0.0313 & x_{5,2} = 0.0535 & x_{5,3} = 0.0613 & x_{5,4} = 0.0535 & x_{5,5} = 0.0313
\end{array}
$$

We see that $x_{3,3} = 0.25$, which is what we expect due to symmetry. Also we note the equal results for the symmetric starting points, such as $x_{1,2} = x_{1,4}$.

Example 3.3.2. The System of Equations Solution. With $m = n = 6$, but with $p_l = 0.40 = p_r$ and $p_u = 0.10 = p_d$, we have the following results for hitting the left side first starting at each interior point:

$$\begin{array}{lllll}
x_{1,1} = 0.6493 & x_{1,2} = 0.7753 & x_{1,3} = 0.7989 & x_{1,4} = 0.7753 & x_{1,5} = 0.6493 \\
x_{2,1} = 0.4294 & x_{2,2} = 0.5760 & x_{2,3} = 0.6097 & x_{2,4} = 0.5760 & x_{2,5} = 0.4294 \\
x_{3,1} = 0.2802 & x_{3,2} = 0.4049 & x_{3,3} = 0.4374 & x_{3,4} = 0.4049 & x_{3,5} = 0.2802 \\
x_{4,1} = 0.1700 & x_{4,2} = 0.2569 & x_{4,3} = 0.2813 & x_{4,4} = 0.2569 & x_{4,5} = 0.1700 \\
x_{5,1} = 0.0804 & x_{5,2} = 0.1246 & x_{5,3} = 0.1374 & x_{5,4} = 0.1246 & x_{5,5} = 0.0804
\end{array}$$

These results also have symmetric characteristics. For initial point (i, j) with a fixed i, the values of $x_{i,j}$ are symmetric; and for a fixed j, $x_{i,j}$ decreases as i increases.

Example 3.3.5. The System of Equations Solution. With $m = 6$, $n = 8$, $p_l = 0.21$, $p_r = 0.23$, $p_u = 0.29$, and $p_d = 0.27$, we have the following results for hitting the left boundary first from each interior point:

$$\begin{array}{lllllll}
x_{1,1} = 0.4363 & x_{1,2} = 0.6035 & x_{1,3} = 0.6675 & x_{1,4} = 0.6814 & x_{1,5} = 0.6579 & x_{1,6} = 0.5851 & x_{1,7} = 0.4145 \\
x_{2,1} = 0.2232 & x_{2,2} = 0.3569 & x_{2,3} = 0.4215 & x_{2,4} = 0.4365 & x_{2,5} = 0.4095 & x_{2,6} = 0.3360 & x_{2,7} = 0.2023 \\
x_{3,1} = 0.1120 & x_{3,2} = 0.2075 & x_{3,3} = 0.2536 & x_{3,4} = 0.2647 & x_{3,5} = 0.2437 & x_{3,6} = 0.1909 & x_{3,7} = 0.1068 \\
x_{4,1} = 0.0648 & x_{4,2} = 0.1132 & x_{4,3} = 0.1407 & x_{4,4} = 0.1473 & x_{4,5} = 0.1341 & x_{4,6} = 0.1026 & x_{4,7} = 0.0555 \\
x_{5,1} = 0.0279 & x_{5,2} = 0.0491 & x_{5,3} = 0.0615 & x_{5,4} = 0.0645 & x_{5,5} = 0.0584 & x_{5,6} = 0.0441 & x_{5,7} = 0.0236
\end{array}$$

We find no symmetric feature or predictable results in this example.

3.5. The Average Number of Steps to Hit a Boundary

In the previous section, we used the System of Equations method to solve for the probabilities of hitting one boundary before the other three starting from all possible starting points. In this section we are going to apply the same method to solve for the average number of steps needed for a four-sided Type I two-dimensional random walk to hit a boundary. We let $s_{i,j}$ be the number of steps needed to hit a boundary when starting at (i,j), for $0 \le i \le m$ and $0 \le j \le n$, and let $y_{i,j} = E[s_{i,j}]$ be the average number of steps needed to hit a boundary. Then we know that $y_{0,j} = 0 = y_{m,j}$ for all j, and $y_{i,0} = 0 = y_{i,n}$ for all i (because if we start on a boundary then no steps are needed). Otherwise, by the Law of Total Average

$$y_{i,j} = 1 + p_l \cdot y_{i-1,j} + p_r \cdot y_{i+1,j} + p_u \cdot y_{i,j+1} + p_d \cdot y_{i,j-1}. \tag{3.5.1}$$

(Because we must take one step, then start anew from one of the four different coordinates with the respective probabilities.) Applying the System of Equations method, first we can rewrite Equation 3.5.1 as

$$p_l \cdot y_{i-1,j} + p_r \cdot y_{i+1,j} - y_{i,j} + p_u \cdot y_{i,j+1} + p_d \cdot y_{i,j-1} = -1.$$

For instance, if $m = 2$ and $n = 3$, because $(i,j) = (0,j)$, $(2,j)$, $(i,0)$, $(i,3)$ are boundary points for all j and for all i, then $y_{0,j} = 0 = y_{2,j}$ for all j, and $y_{i,0} = 0 = y_{i,3}$ for all i. Otherwise, for the only two interior points $(i,j) = (1,1)$ and $(1,2)$ in this example, which has small boundaries, we have

$$p_l \cdot y_{0,1} + p_r \cdot y_{2,1} - y_{1,1} + p_u \cdot y_{1,2} + p_d \cdot y_{1,0} = -1$$

and

$$p_l \cdot y_{0,2} + p_r \cdot y_{2,2} - y_{1,2} + p_u \cdot y_{1,3} + p_d \cdot y_{1,1} = -1.$$

Then we have a 12×12 matrix of coefficients, namely S, a 12×1 matrix of constants, namely T, and a system of equations $SY = T$. We solve for Y by $Y = S^{-1}T$. The augmented matrix of the system is shown below, where the top row are the indices of the unknowns $y_{i,j}$:

00	01	02	03	10	11	12	13	20	21	22	23	
1	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0
0	p_l	0	0	p_d	-1	p_u	0	0	p_r	0	0	-1
0	0	p_l	0	0	p_d	-1	p_u	0	0	p_r	0	-1
0	0	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	1	0

TABLE 3.4. Augmented Matrix for Boundaries $m = 2$ and $n = 3$

We note that this augmented matrix is very similar to Matrix 3.3, except for all the entries on the boundaries are 0 and the entries on the right hand side of each row that has p_l, p_d, p_u and p_r are -1 instead of 0. Applying the System of Equations method to solve the equation $SY = T$ for Y, we obtain the average number of steps $y_{i,j}$ needed to hit a boundary from all possible starting points simultaneously.

Now we are going to use *Mathematica* to solve for the average number of steps needed to hit a boundary in Examples 3.3.1, 3.3.2 and 3.3.5. Also see Appendix B for the code. In each example, for all the starting points on the boundaries, $y_{i,j} = 0$. Therefore, we are interested in finding $y_{i,j}$ when starting at the interior points.

Example 3.3.1. The Average Number of Steps to Hit a Boundary Solution. With $m = n = 6$, and $p_l = p_r = p_u = p_d = 0.25$, the average number of steps needed to hit a boundary starting at each interior point is:

$y_{1,1} = 3.8077$	$y_{1,2} = 5.6154$	$y_{1,3} = 6.1539$	$y_{1,4} = 5.6154$	$y_{1,5} = 3.8077$
$y_{2,1} = 5.6154$	$y_{2,2} = 8.5000$	$y_{2,3} = 9.3846$	$y_{2,4} = 8.5000$	$y_{2,5} = 5.6154$
$y_{3,1} = 6.1539$	$y_{3,2} = 9.3846$	$y_{3,3} = 10.3846$	$y_{3,4} = 9.3846$	$y_{3,5} = 6.1539$
$y_{4,1} = 5.6154$	$y_{4,2} = 8.5000$	$y_{4,3} = 9.3846$	$y_{4,4} = 8.5000$	$y_{4,5} = 5.6154$
$y_{5,1} = 3.8077$	$y_{5,2} = 5.6154$	$y_{5,3} = 6.1539$	$y_{5,4} = 5.6154$	$y_{5,5} = 3.8077$

The table shows that $y_{1,j} = y_{5,j}$ for all j, and $y_{2,j} = y_{4,j}$ for all j due to symmetry. We can also see that a walk starting in the middle needs more steps on average to hit a boundary since the probabilities are evenly distributed on each step when moving towards four different directions.

Example 3.3.2. The Average Number of Steps to Hit a Boundary Solution. With $m = n = 6$, but with $p_l = 0.4 = p_r$, and $p_u = 0.10 = p_d$, then we have the following results for the average number of steps needed to hit a boundary starting at each interior point.

$$
\begin{array}{lllll}
y_{1,1} = 4.0233 & y_{1,2} = 5.3439 & y_{1,3} = 5.6620 & y_{1,4} = 5.3439 & y_{1,5} = 4.0233 \\
y_{2,1} = 6.2223 & y_{2,2} = 8.4384 & y_{2,3} = 8.9831 & y_{2,4} = 8.4384 & y_{2,5} = 6.2223 \\
y_{3,1} = 6.9229 & y_{3,2} = 9.4506 & y_{3,3} = 10.0766 & y_{3,4} = 9.4506 & y_{3,5} = 6.9229 \\
y_{4,1} = 6.2223 & y_{4,2} = 8.4384 & y_{4,3} = 8.9831 & y_{4,4} = 8.4384 & y_{4,5} = 6.2223 \\
y_{5,1} = 4.0233 & y_{5,2} = 5.3439 & y_{5,3} = 5.6620 & y_{5,4} = 5.3439 & y_{5,5} = 4.0233
\end{array}
$$

Comparing these results to those of Example 3.3.1, we can see that the probabilities affect the results significantly since that is the only difference between them. In particular, for the same starting point $(1,1)$, the average number of steps to hit a boundary in Example 3.3.1 is 3.8077, while it is 4.0233 in this example. However the results still show symmetry.

Example 3.3.5. The Average Number of Steps to Hit a Boundary Solution. With $m = 6$, $n = 8$, $p_l = 0.21$, $p_r = 0.23$, $p_u = 0.29$, and $p_d = 0.27$, we have the following results for the average number of steps needed to hit a boundary starting at each interior point:

$y_{1,1} = 4.5350$	$y_{1,2} = 6.9650$	$y_{1,3} = 8.1271$	$y_{1,4} = 8.3966$	$y_{1,5} = 7.8993$	$y_{1,6} = 6.5736$	$y_{1,7} = 4.1485$
$y_{2,1} = 6.5876$	$y_{2,2} = 10.364$	$y_{2,3} = 12.224$	$y_{2,4} = 12.6588$	$y_{2,5} = 11.8515$	$y_{2,6} = 9.7291$	$y_{2,7} = 5.9725$
$y_{3,1} = 7.0857$	$y_{3,2} = 11.2074$	$y_{3,3} = 13.2519$	$y_{3,4} = 13.7309$	$y_{3,5} = 12.8406$	$y_{3,6} = 10.5075$	$y_{3,7} = 6.4105$
$y_{4,1} = 6.3136$	$y_{4,2} = 9.8904$	$y_{4,3} = 11.6388$	$y_{4,4} = 12.0469$	$y_{4,5} = 11.2922$	$y_{4,6} = 9.2975$	$y_{4,7} = 5.7358$
$y_{5,1} = 4.1628$	$y_{5,2} = 6.3342$	$y_{5,3} = 7.3562$	$y_{5,4} = 7.5923$	$y_{5,5} = 7.1596$	$y_{5,6} = 5.9942$	$y_{5,7} = 3.8230$

Though the given conditions of this example imply unpredictable results, we still get the greatest average value from the centered starting point $(3,4)$. This is because we do not specify which boundary has to be hit and the centered point always has the longest distance to a boundary than any of the other interior points.

We note that the System of Equations method allows us to compute $y_{h,k}$, the average number of steps needed to hit a boundary starting at (h,k), without having the pmf of $s_{h,k}$. Since a closed-form formula for the pmf of $s_{h,k}$ is unknown, we cannot compute $P(s_{h,k} = x)$ for a specific number of steps x. But using the Markov Chains method described in Section 3.3, we will be able to find the pmf value for $s_{h,k}$ after taking exactly x steps.

In Section 3.3, to find the probabilities of having all possible positions after x steps, we multiply the matrices $B \times A^x$. And we obtain the final probabilities when letting x be large such as $x = 600$. We were interested in the final states when all the four boundaries have been hit first and their probabilities of being hit first sum to 1. Now we are interested in the states after x steps have been taken. So for $C_x = B \times A^x = (c_{1,ij,x})$, for $0 \le i \le m$, $0 \le j \le n$ and $x \ge \min\{h, k, m-h, n-k\}$

(because the minimum number of steps to hit a boundary is the shortest distance to a boundary), we can state

Theorem 3.3.1. Let $s_{h,k}$ be the number of steps needed for a Type I two-dimensional random walk to hit a boundary from starting point (h,k), where $0 \le h \le m$ and $0 \le k \le n$. For $x > \min\{h,k,m-h,n-k\}$,

(a) The cdf of $s_{h,k}$ is given by

$$P(s_{h,k} \le x) = P(R_x) + P(L_x) + P(U_x) + P(D_x).$$

(b) The pmf of $s_{h,k}$ is given by

$$\begin{aligned} P(s_{h,k} = x) &= P(s_{h,k} \le x) - P(s_{h,k} \le x-1) \\ &= [P(R_x) + P(L_x) + P(U_x) + P(D_x)] \\ &\quad - [P(R_{x-1}) + P(L_{x-1}) + P(U_{x-1}) + P(D_{x-1})]. \end{aligned}$$

Example 3.5.1. Starting at $(h,k) = (3,3)$, with boundaries $m = n = 6$, and evenly distributed probabilities $p_l = p_d = p_r = p_u = 0.25$, it requires at least 3 steps to hit a boundary. When $x = 10$ steps, using *Mathematica* (the code in Appendix A), we obtain the results for $P(L_{10})$, $P(R_{10})$, $P(U_{10})$, and $P(D_{10})$:

$P(L_{10})$	$P(R_{10})$	$P(U_{10})$	$P(D_{10})$	Sum
0.157766	0.157766	0.157766	0.157766	0.631055

When $x = 9$, we have

$P(L_9)$	$P(R_9)$	$P(U_9)$	$P(D_9)$	Sum
0.144634	0.144634	0.144634	0.144634	0.578537

Applying Theorem 3.3.1, we obtain the probability of hitting a boundary in exactly 10 steps as

$$\begin{aligned} P(s_{3,3} = 10) &= P(s_{3,3} \le 10) - P(s_{3,3} \le 9) \\ &= 0.631055 - 0.578537 \\ &= 0.052518. \end{aligned}$$

In this case, there is 5.2518% chance of hitting a boundary in exactly 10 steps. With the same process, we find the probability of hitting a boundary in exactly 3 steps to be $P(s_{3,3} = 3) = P(s_{3,3} \le 3) - P(s_{3,3} \le 2) = 0.0625 - 0 = 0.0625$.

Example 3.5.2. Starting at $(h, k) = (13, 14)$, with $m = 19$, $n = 25$, $p_l = 0.10$, $p_r = 0.31$, $p_u = 0.17$, and $p_d = 0.42$. The minimum number of steps needed to hit a boundary is 6. When $x = 6$ steps, the results for $P(L_x)$, $P(R_x)$, $P(U_x)$, and $P(D_x)$ are

$P(L_6)$	$P(R_6)$	$P(U_6)$	$P(D_6)$	Sum
0	0.001	0	0	0.001

Therefore $P(s_{13,14} = 6) = P(s_{13,14} \le 6) = 0.001$. Though the probability of hitting a boundary is small, there is still 0.1% chance to hit a boundary when $x = 6$ steps. Now if we increase to $x = 56$ steps, we have

$P(L_{56})$	$P(R_{56})$	$P(U_{56})$	$P(D_{56})$	Sum
0	0.000012	0.8498	0.1307	0.9805

$P(L_{55})$	$P(R_{55})$	$P(U_{55})$	$P(D_{55})$	Sum
0	0.000012	0.8483	0.1307	0.9782

Then $P(s_{13,14} = 56) = 0.9805 - 0.9782 = 0.0023$. Thus for this example, there is 0.23% probability of hitting a boundary in exactly 56 steps.

Chapter 4

Boundary Problems for Type II Two-Dimensional Random Walk

4.1. Introduction

A diagonally-moving two-dimensional random walk begins at integer-valued coordinates (h,k), where $0 \le h \le m$ and $0 \le k \le n$. On each step, the random walk moves one unit towards four different directions, either left and downward, left and upward, right and upward, or right and downward with probabilities $p_{ld} \neq 0$, $p_{lu} \neq 0$, $p_{ru} \neq 0$, $p_{rd} \neq 0$, respectively, where $p_{ld} + p_{lu} + p_{ru} + p_{rd} = 1$. The boundaries are the lines $x = 0$ (the y-axis), $y = 0$ (the x-axis), $x = m$ and $y = n$. The process stops when it hits a boundary or one of the four corner points $(0,0)$, $(0,n)$, $(m,0)$, (m,n). This random walk process will be called a Type II two-dimensional random walk.

As an example, we let $m = 5$, $n = 5$, $(h,k) = (2,3)$, with $p_{ld} = 0.30$, $p_{lu} = 0.25$, $p_{ru} = 0.35$, and $p_{rd} = 0.10$. A possible specific path can be: starting at point $(2,3)$, go right and downward 1 unit, go right and upward 1 unit, go left and upward another 1 unit, finally go right and upward 1 more unit, it hits the upper boundary and stops. The probability of this path is

$$p_{rd} \cdot p_{ru} \cdot p_{lu} \cdot p_{ru} = 0.10 \times 0.35 \times 0.25 \times 0.35 = 0.0030625.$$

This path is shown as:

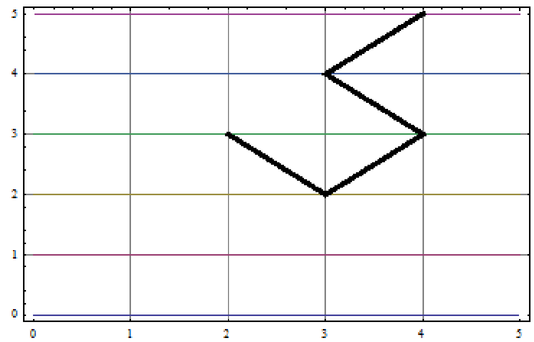

FIGURE 4.1. A Type II Path

Another possible specific path can be: starting at point $(2, 3)$, moving left and downward 1 unit, moving right and downward 1 unit, moving right and upward 1 unit, and finally moving right and downward twice, so that it hits the bottom right corner $(5, 0)$ and stops. This path is shown as:

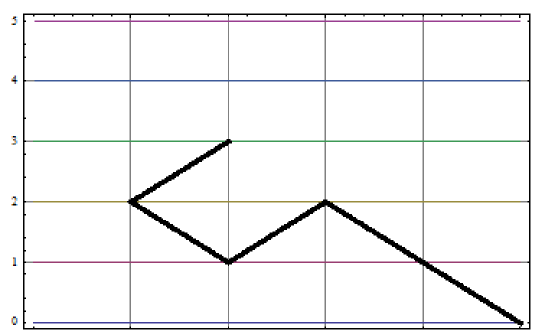

FIGURE 4.2. Another Type II Path

The probability of this path is

$$p_{ld} \cdot p_{rd} \cdot p_{ru} \cdot {p_{rd}}^2 = 0.30 \times 0.10 \times 0.35 \times 0.10^2 = 0.000105.$$

We note that, different from a Type I random walk described in Chapter 3, it is possible for Type II random walk to hit the four corners. However, for the paths that start at a specific point, not every boundary point or every corner point will be hit. For instance, $(0, 2)$ and $(0, 4)$ on the left boundary will never be hit starting from $(2, 3)$. The starting point determines which boundary points or corner points are never going to be hit. Since at each movement, both coordinates will increase or decrease by one unit. If the starting point (h, k) has even or odd numbers at both h and k, then the end points must have either both even or both odd coordinates. Similarly, if the starting point (h, k) has one even and one odd at h and k, then the end points must have one even and one odd coordinate. In this example, if we change the starting point to $(2, 2)$, then it will be $(0, 1)$ and $(0, 3)$ that will never be hit instead since these two end points have an even x-coordinate and an odd y-coordinate, while the starting point has even numbers at both coordinates.

In general, each single path has a distinct probability and there are infinitely many different paths. Our goal is to determine the overall probabilities of hitting each boundary or each corner before hitting the others. We have shown in Section 3.2 that the Type I random walk will hit a boundary with probability 1. Using the same argument, we can show the Type II random walk will hit a boundary or a corner with probability 1. In Chapter 3, we introduced the Markov Chains method to solve for the probabilities of each boundary being hit first from a specified starting point, and the System of Equations method to solve for the probabilities of a specified boundary to be hit first from all possible starting points. In this chapter, we will modify these two methods to solve similar boundary problems for a Type II two-dimensional random walk.

4.2. The Markov Chains Solution

We want the probabilities of hitting each boundary or each corner first from a starting point, and we are interested in the cases that require a complex calculation. We are able to tell some characteristics for symmetric cases, but for the precise numerical results, we need a method to compute. We shall first look at some examples.

Example 4.2.1. Let $m = 6,\ n = 6,\ (h, k) = (3, 3)$, and $p_{ld} = p_{rd} = p_{lu} = p_{ru} = 0.25$. Then by symmetry, there are equal chances of hitting one boundary or one corner before the other three. However, we are not able to tell what probability there is for each boundary or for each corner to be hit first. Also, not every point on one boundary has the same probability of being hit first. To see this, we can look at the five points on y-axis: $(0, 1),\ (0, 2),\ (0, 3),\ (0, 4),\ (0, 5)$, and the two corners on the left side: $(0, 0),\ (0, 6)$. We would expect $(0, 2)$ and $(0, 4)$ to have a bigger chance of being hit first than $(0, 0)$ and $(0, 6)$. And by symmetry, $(0, 2)$ and $(0, 4)$ should be hit first with the same probabilities, as should $(0, 0)$ and $(0, 6)$. Meanwhile, no path starting at $(3, 3)$ will ever hit $(0, 1)$, $(0, 3)$ and $(0, 5)$. Thus these three boundary points must have probability 0 of being hit first.

Example 4.2.2. If $m = 16,\ n = 17,\ (h, k) = (6, 8)$, $p_{ld} = 0.21$, $p_{rd} = 0.27$, $p_{lu} = 0.23$, and $p_{ru} = 0.29$, then there is no symmetry to help determine which boundary is most likely to be hit first, and the boundaries are too big to allow computing by hand.

General cases like this require a method and a computing device to find the solutions. We are going to adjust the Markov Chains method introduced in Section 3.3 and use *Mathematica* to solve these boundary problems for Type II two-dimensional random walk.

We first let $A = (a_{ij,kl})$, for $0 \le i \le m$, $0 \le j \le n$, $0 \le k \le m$, $0 \le l \le n$ be a quadruple-indexed matrix of transition probabilities, which has dimension $(m+1)(n+1) \times (m+1)(n+1)$. The term $a_{ij,kl}$ gives the probability of the Type II walk moving from coordinates (i,j) to coordinates (k,l) on each step, where $a_{ij,ij} = 1$ if (i,j) is on a boundary or a corner. For example, with $m = 3$, $n = 4$, $A = (a_{ij,kl})$ is a 20×20 matrix. If $(i,j) = (0,0)$, $(0,1)$, $(0,2)$, $(0,3)$, $(0,4)$, $(1,0)$, $(1,4)$, $(2,0)$, $(2,4)$, $(3,0)$, $(3,1)$, $(3,2)$,$(3,3)$, or $(3,4)$, which are the points on the four boundaries or the four corners, then $a_{ij,ij} = 1$; i.e., if the initial position is on a boundary or corner point, then it stays on that point with probability 1. It goes nowhere. Then the probabilities $a_{ij,kl} = 0$ for the other coordinates (k,l). But if we do not start at a boundary point or a corner, then we have positive probabilities of moving toward four other coordinates. For example, if $(i,j) = (1,1)$, then the probability of moving to $(0,0)$ is $a_{11,00} = p_{ld}$. Likewise, $a_{11,02} = p_{lu}$, $a_{11,22} = p_{ru}$, and $a_{11,20} = p_{rd}$.

Below is the complete transition matrix A for our example with $m = 3$ and $n = 4$. The terms on the top and the left side are simply place holders that tell the possible coordinates. The place holders on the left represent the previous state, and the place holders on the top represent the possible coordinates after another step is taken.

	00	01	02	03	04	10	11	12	13	14	20	21	22	23	24	30	31	32	33	34
00	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
01	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
02	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
03	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
04	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	p_{ld}	0	p_{lu}	0	0	0	0	0	0	0	p_{rd}	0	p_{ru}	0	0	0	0	0	0	0
12	0	p_{ld}	0	p_{lu}	0	0	0	0	0	0	0	p_{rd}	0	p_{ru}	0	0	0	0	0	0
13	0	0	p_{ld}	0	p_{lu}	0	0	0	0	0	0	0	p_{rd}	0	p_{ru}	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	p_{ld}	0	p_{lu}	0	0	0	0	0	0	0	p_{rd}	0	p_{ru}	0	0
22	0	0	0	0	0	0	p_{ld}	0	p_{lu}	0	0	0	0	0	0	0	p_{rd}	0	p_{ru}	0
23	0	0	0	0	0	0	0	p_{ld}	0	p_{lu}	0	0	0	0	0	0	0	p_{rd}	0	p_{ru}
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

TABLE 4.1. 20×20 Transition Matrix A for $m = 3$ and $n = 4$

Secondly, we let $B = (b_{1,ij})$, for $0 \leq i \leq m$, $0 \leq j \leq n$ be the $1 \times (m+1)(n+1)$ initial state matrix that designates the initial position of the Type II walk. Then we know $b_{1,hk} = 1$, and $b_{1,ij} = 0$ when $i \neq h$ or $j \neq k$. In our example with $m = 3$ and $n = 4$, if $(h,k) = (2,3)$, then the initial state matrix $B = (b_{1,ij})$ is a 1×20 matrix and $b_{1,ij} = 0$ for all $i,\ j$ except $b_{1,23} = 1$. Matrix B is shown below, where the terms on the top line are place holders.

00	01	02	03	04	10	11	12	13	14	20	21	22	23	24	30	31	32	33	34
0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

TABLE 4.2. 1×20 Initial State Matrix B for $m = 3$ and $n = 4$

To find the probabilities of having all possible positions after x steps, we multiply $B \times A^x$. Letting x be large enough so that the probabilities of each boundary or each corner to be hit first sum to 1, then we obtain the final states of each boundary or each corner being hit first. Otherwise, we obtain the cdf value for each boundary or each corner being hit first when up to x steps are taken. We then can compute the pmf value for a boundary or a corner being hit first in exactly x steps without knowing a closed-form formula for the pmf.

We let $C_x = B \times A^x = (c_{1,ij,x})$, for $0 \le i \le m,$, $0 \le j \le n$. Here C_x is a $1 \times (m+1)(n+1)$ matrix, where $c_{1,ij,x}$ gives probabilities of being at (i,j) after x steps. We let $P(L_x)$, $P(D_x)$, $P(U_x)$, and $P(R_x)$ be the probabilities of hitting the left boundary, the lower boundary, the upper boundary, and the right boundary in x steps, respectively. Thus, we can state:

Theorem 4.2.1. The probabilities of a four-sided Type II two-dimensional random walk being at each boundary, excluding corner points, after x steps are given by

$$(a) \qquad P(L_x) = \sum_{j=1}^{n-1} c_{1,0j,x} \qquad\qquad (b) \qquad P(D_x) = \sum_{i=1}^{m-1} c_{1,i0,x}$$

$$(c) \qquad P(U_x) = \sum_{i=1}^{m-1} c_{1,in,x} \qquad\qquad (d) \qquad P(R_x) = \sum_{j=1}^{n-1} c_{1,mj,x}.$$

By taking the limit when $x \to \infty$ we reach the final states. Hence, we have

Theorem 4.2.2. Excluding corner points, the probabilities of a four-sided Type II two-dimensional random walk hitting each boundary first from its initial starting point are given by

$$(a) \qquad P(L_\infty) = \lim_{x\to\infty} \sum_{j=1}^{n-1} c_{1,0j,x} \qquad\qquad (b) \qquad P(D_\infty) = \lim_{x\to\infty} \sum_{i=1}^{m-1} c_{1,i0,x}$$

$$(c) \qquad P(U_\infty) = \lim_{x\to\infty} \sum_{i=1}^{m-1} c_{1,in,x} \qquad\qquad (d) \qquad P(R_\infty) = \lim_{x\to\infty} \sum_{j=1}^{n-1} c_{1,mj,x}.$$

We let $P(B_x^L)$, $P(T_x^L)$, $P(T_x^R)$, and $P(B_x^R)$ be the probabilities of hitting the bottom left corner, the top left corner, the top right corner, and the bottom right corner first after x steps, respectively. We then have:

Theorem 4.2.3. The probabilities of a four-sided Type II two-dimensional random walk being at each corner point after x steps are given by

$$(a) \qquad P(B_x^L) = c_{1,00,x} \qquad\qquad (b) \qquad P(T_x^L) = c_{1,0n,x}$$

$$(c) \qquad P(T_x^R) = c_{1,mn,x} \qquad\qquad (d) \qquad P(B_x^R) = c_{1,m0,x}.$$

Taking the limit as $x \to \infty$, we obtain

Theorem 4.2.4. The probabilities of a four-sided Type II two-dimensional random walk hitting each corner point first from its initial starting point are given by

$$(a) \quad P(B_\infty^L) = \lim_{x\to\infty} c_{1,00,x} \qquad (b) \quad P(T_\infty^L) = \lim_{x\to\infty} c_{1,0n,x}$$

$$(c) \quad P(T_\infty^R) = \lim_{x\to\infty} c_{1,mn,x} \qquad (d) \quad P(B_\infty^R) = \lim_{x\to\infty} c_{1,m0,x}.$$

We shall use *Mathematica* for these computations. See Appendix C for the code. Doing so, we can quickly obtain the solutions to Examples 4.2.1 and 4.2.2.

Example 4.2.1. The Markov Chains Solution. For $m = 6$, $n = 6$, $(h,k) = (3,3)$, and $p_{ld} = p_{rd} = p_{lu} = p_{ru} = 0.25$. When taking $x = 10$ steps, the probabilities of being on each left boundary point are:

(i,j)	$(0,1)$	$(0,2)$	$(0,3)$	$(0,4)$	$(0,5)$	$P(L_{10})$
$c_{1,ij,10}$	0	0.0964	0	0.0964	0	0.1928

These results for the other three boundaries are similar. The probability of each corner point being hit first is approximately 0.0321. In summary, we have

$P(L_{10})$	$P(U_{10})$	$P(R_{10})$	$P(D_{10})$	$P(B_{10}^L)$	$P(T_{10}^L)$	$P(T_{10}^R)$	$P(B_{10}^R)$	Sum
0.1928	0.1928	0.1928	0.1928	0.0321	0.0321	0.0321	0.0321	0.8999

The total sum of the probabilities that boundaries or corners being hit has not yet reached 1 after taking 10 steps. Now we increase the number of steps to 500, and we obtain the following final probabilities:

$P(L_{500})$	$P(U_{500})$	$P(R_{500})$	$P(D_{500})$	$P(B^L_{500})$	$P(T^L_{500})$	$P(T^R_{500})$	$P(B^R_{500})$	Sum
0.2143	0.2143	0.2143	0.2143	0.0357	0.0357	0.0357	0.0357	1

Specifically, the values $c_{1,0j,500}$ for each left boundary point including the two corner points on the left side (round to the fourth decimal place) are

(i,j)	$(0,0)$	$(0,1)$	$(0,2)$	$(0,3)$	$(0,4)$	$(0,5)$	$(0,6)$
$c_{1,ij,500}$	0.0357	0	0.1071	0	0.1071	0	0.0357

The results are as we have expected:

(i) Points on the boundaries that are symmetric about the initial point have the same probabilities of being hit first;

(ii) the boundary points $(0,2)$ and $(0,4)$ have greater probabilities of being hit first than the corner points $(0,0)$ and $(0,6)$.

(iii) $(0,1)$, $(0,3)$ and $(0,5)$ are never hit thus have probability 0.

Example 4.2.2. The Markov Chains Solution. With $(h,k) = (6,8)$, $m = 16$, $n = 17$, $p_{ld} = 0.21$, $p_{lu} = 0.23$, $p_{ru} = 0.29$, and $p_{rd} = 0.27$, the transition matrix A has dimensions $(16+1)(17+1)\times(16+1)(17+1) = 306\times306$, and matrix B has dimensions 1×306. It is almost impossible to compute this problem by hand. Thankfully, we can obtain the solutions quickly and accurately using *Mathematica*. The final state for the probability of hitting each boundary or each corner is obtained when $x \geq 342$ steps. The minimum number of steps to obtain the final state is done by test and trial on the *Mathematica* code with different inputs on MaximumNumberOfSteps. The results are as follows:

$P(L_x)$	$P(U_x)$	$P(R_x)$	$P(D_x)$	$P(B_x^L)$	$P(T_{10}^L)$	$P(T_x^R)$	$P(B_x^R)$	Sum
0.1547	0.2574	0.4212	0.1600	0.0015	0	0	0.0052	1

We notice that the top left corner $(0,17)$ and the top right corner $(16,17)$ both have probability 0 of being hit first. No path will reach them starting at $(6,8)$ since both coordinates of the starting point are even numbers while these two corner points have one even coordinate and one odd.

4.3. The System of Equations Solution

We now know the Markov Chains method can give us simultaneously the probabilities of the four boundaries or the four corners being hit first if we know where the walk starts. But what if a Type II walk can possibly start at any point within the boundaries, and we want the probability of a certain boundary or a certain corner being hit first? We use the System of Equations method to fulfill this goal for the Type I random walk. We now will amend this method to achieve our goal for the Type II walk.

We let $x_{i,j}$ be the probability of hitting the left boundary first when starting at (i,j), for $0 \leq i \leq m$, $0 \leq j \leq n$. Then we know that $x_{0,j} = 1$ for all $1 \leq j \leq n-1$ (because if we start at left boundary then we stay there and already hit the left boundary). Also $x_{m,j} = 0$ for all j, and $x_{i,0} = 0 = x_{i,n}$ for all i (because if we start at the corners, the right boundary, the lower boundary, or the upper boundary, then we are not going to move and will never hit the left boundary). Otherwise, by the

Law of Total Probability,

$$x_{i,j} = p_{ld} \cdot x_{i-1,j-1} + p_{lu} \cdot x_{i-1,j+1} + p_{ru} \cdot x_{i+1,j+1} + p_{rd} \cdot x_{i+1,j-1},$$

which can be re-written as

$$p_{ld} \cdot x_{i-1,j-1} + p_{lu} \cdot x_{i-1,j+1} - x_{i,j} + p_{ru} \cdot x_{i+1,j+1} + p_{rd} \cdot x_{i+1,j-1} = 0.$$

Similarly, we let $c_{i,j}$ be the probability of hitting the bottom left corner $(0,0)$ first when starting at (i,j), for $0 \le i \le m$, $0 \le j \le n$. Then we know $c_{0,0} = 1$. Also, $c_{0,j} = 0$ for $j > 0$, $c_{i,0} = 0$ for $i > 0$, $c_{i,n} = 0$ for all i, and $c_{m,j} = 0$ for all j (because if we start at the other three corners or any boundary then we will not move and will never hit the bottom left corner). Otherwise,

$$c_{i,j} = p_{ld} \cdot c_{i-1,j-1} + p_{lu} \cdot c_{i-1,j+1} + p_{ru} \cdot c_{i+1,j+1} + p_{rd} \cdot c_{i+1,j-1},$$

which can be re-written as

$$p_{ld} \cdot c_{i-1,j-1} + p_{lu} \cdot c_{i-1,j+1} - c_{i,j} + p_{ru} \cdot c_{i+1,j+1} + p_{rd} \cdot c_{i+1,j-1} = 0.$$

For instance, if $m = 2$ and $n = 3$, then $x_{0,j} = 1$, $c_{0,0} = 1$ and $x_{2,j} = 0$, $c_{2,j} = 0$ for $j = 0,\ 1,\ 2,\ 3$, and $x_{i,0} = 0 = x_{i,3}$ for $i = 0,\ 1,\ 2$. $c_{0,j} = 0$ for $j = 1,\ 2,\ 3$, $c_{i,0} = 0$ for $i = 1,\ 2$ and $c_{i,3}$ for $i = 0,\ 1,\ 2$. Otherwise, for starting at $(1,1)$, we have

$$p_{ld} \cdot x_{0,0} + p_{lu} \cdot x_{0,2} - x_{1,1} + p_{ru} \cdot x_{2,2} + p_{rd} \cdot x_{2,0} = 0.$$

$$p_{ld} \cdot c_{0,0} + p_{lu} \cdot c_{0,2} - c_{1,1} + p_{ru} \cdot c_{2,2} + p_{rd} \cdot c_{2,0} = 0.$$

and for starting at $(1,2)$, we have

$$p_{ld} \cdot x_{0,1} + p_{lu} \cdot x_{0,3} - x_{1,2} + p_{ru} \cdot x_{2,3} + p_{rd} \cdot x_{2,1} = 0.$$

$$p_{ld} \cdot c_{0,1} + p_{lu} \cdot c_{0,3} - c_{1,2} + p_{ru} \cdot c_{2,3} + p_{rd} \cdot c_{2,1} = 0.$$

Then we have a 12×12 matrix of coefficients, namely G, a 12×1 matrix of constants, namely H, and a system of equations $GX = H$. We solve for X by $X = G^{-1}H$. The augmented matrix shown below is for the system $GX = H$, where the top row are the indices of the unknowns $x_{i,j}$:

00	01	02	03	10	11	12	13	20	21	22	23	
1	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	1
0	0	1	0	0	0	0	0	0	0	0	0	1
0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0
p_{ld}	0	p_{lu}	0	0	-1	0	0	p_{rd}	0	p_{ru}	0	0
0	p_{ld}	0	p_{lu}	0	0	-1	0	0	p_{rd}	0	p_{ru}	0
0	0	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	1	0

TABLE 4.3. Augmented Matrix for Boundaries $m = 2$ and $n = 3$

The following augmented matrix is for the system $GC = H$, where the top row are the indices of the unknowns $c_{i,j}$:

00	01	02	03	10	11	12	13	20	21	22	23	
1	0	0	0	0	0	0	0	0	0	0	0	1
0	1	0	0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0
p_{ld}	0	p_{lu}	0	0	-1	0	0	p_{rd}	0	p_{ru}	0	0
0	p_{ld}	0	p_{lu}	0	0	-1	0	0	p_{rd}	0	p_{ru}	0
0	0	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	1	0

TABLE 4.4. Augmented Matrix for Boundaries $m = 2$ and $n = 3$

We now use *Mathematica* to redo Examples 4.2.1 and 4.2.2, only considering hitting the left boundary first from all possible starting points. See Appendix D for the code. In each example, for all the starting points on the left boundary, $x_{i,j} = 1$, and for all the starting points at the corners or on the lower boundary, upper boundary and the right boundary, $x_{i,j} = 0$. Therefore, we are interested in showing $x_{i,j}$ when starting at the interior points.

Example 4.2.1. The System of Equations Solution. With $m = n = 6$ and $p_l = p_r = p_u = p_d = 0.25$, the probabilities (i, j) of hitting the left boundary first starting at all interior points are given by:

$x_{1,1} = 0.3324$	$x_{1,2} = 0.6648$	$x_{1,3} = 0.6648$	$x_{1,4} = 0.6648$	$x_{1,5} = 0.2198$
$x_{2,1} = 0.2198$	$x_{2,2} = 0.3297$	$x_{2,3} = 0.4396$	$x_{2,4} = 0.3297$	$x_{2,5} = 0.2198$
$x_{3,1} = 0.1071$	$x_{3,2} = 0.2143$	$x_{3,3} = 0.2143$	$x_{3,4} = 0.2143$	$x_{3,5} = 0.1071$
$x_{4,1} = 0.0695$	$x_{4,2} = 0.0989$	$x_{4,3} = 0.1319$	$x_{4,4} = 0.0989$	$x_{4,5} = 0.0659$
$x_{5,1} = 0.0247$	$x_{5,2} = 0.0495$	$x_{5,3} = 0.0495$	$x_{5,4} = 0.0495$	$x_{5,5} = 0.0247$

We note that when the process starts at the center $(3, 3)$, then $x_{3,3} = 0.2143$. This is the same result as we obtained using the Markov Chains method. The symmetric starting points such as $(1, 1)$ and $(1, 5)$ yield the same value for $x_{i,j}$. Starting at $(1, 2)$, $(1, 3)$ and $(1, 4)$ give an equal chance to hit the left boundary first, but they yield a greater chance than starting at $(1, 1)$ and $(1, 5)$ since we don't include the two corner points on the boundary, and that reduces the probability of hitting the left boundary from $(1, 1)$ and $(1, 5)$.

Example 4.2.2. The System of Equations Solution. With $m = 16$, $n = 17$, $p_ld = 0.21$, $p_lu = 0.23$, $p_ru = 0.29$, and $p_rd = 0.27$, when the starting point is at $(6, 8)$, we have $x_{6,8} = 0.1547$. This is the value we obtained when using the Markov Chains method for the final state of $P(L_x)$. Within the boundaries $m = 16$, $n = 17$, there are $(16 - 2) \times (17 - 2) = 210$ interior points. We are not going to list all the solutions at the other interior starting points in this paper.

Remark. With some minor alterations to the code, finding the solutions to $c_{i,j}$, the probabilities of hitting a certain corner first from all possible interior points, can also be done through *Mathematica*.

4.4. The Average Number of Steps To Hit a Boundary or a Corner

We again let $s_{i,j}$ be the number of steps needed to hit a boundary (including the corner points) when a Type II two-dimensional random walk starts at (i,j), for $0 \le i \le m$ and $0 \le j \le n$. And we let $e_{i,j} = E[s_{i,j}]$ be the average number of steps needed for a Type II random walk to hit a boundary starting from (i,j). Then we know that $e_{0,j} = 0 = e_{m,j}$ for all $0 \le j \le n$, and $e_{i,0} = 0 = e_{i,n}$ for all $0 \le i \le m$ (because if we start on a boundary or a corner then no steps are needed). Otherwise, by the Law of Total Average,

$$e_{i,j} = 1 + p_{ld} \cdot e_{i-1,j-1} + p_{lu} \cdot e_{i-1,j+1} + p_{ru} \cdot e_{i+1,j+1} + p_{rd} \cdot e_{i+1,j-1} \tag{4.4.1}$$

Because we must take one step, then start anew from one of the four different coordinates with the respective probabilities. To apply the System of Equations method, first we rewrite Equation 4.4.1 as

$$p_{ld} \cdot e_{i-1,j-1} + p_{lu} \cdot e_{i-1,j+1} - e_{i,j} + p_{ru} \cdot e_{i+1,j+1} + p_{rd} \cdot e_{i+1,j-1} = -1.$$

For instance, if $m = 2$ and $n = 3$, then $e_{0,j} = 0 = e_{2,j}$ for $0 \le j \le 3$, and $e_{1,0} = 0 = e_{1,3}$. Otherwise, from the interior starting points $(1,1)$ and $(1,2)$, we have the equations

$$p_{ld} \cdot e_{0,0} + p_{lu} \cdot e_{0,2} - e_{1,1} + p_{ru} \cdot e_{2,2} + p_{rd} \cdot e_{2,0} = -1.$$

and

$$p_{ld} \cdot e_{0,1} + p_{lu} \cdot e_{0,3} - e_{1,2} + p_{ru} \cdot e_{2,3} + p_{rd} \cdot e_{2,1} = -1.$$

Then we have a 12×12 matrix of coefficients, namely P, a 12×1 matrix of constants, namely Q, and a system of equations $PE = Q$. We solve for the system by $E = P^{-1}Q$. The augmented matrix shown below is for the system of equations $PE = Q$, where the top row are the indices of the unknowns $e_{i,j}$:

00	01	02	03	10	11	12	13	20	21	22	23	
1	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0
p_{ld}	0	p_{lu}	0	0	-1	0	0	p_{rd}	0	p_{ru}	0	-1
0	p_{ld}	0	p_{lu}	0	0	-1	0	0	p_{rd}	0	p_{ru}	-1
0	0	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	1	0

TABLE 4.5. Augmented Matrix for Boundaries $m = 2$ and $n = 3$

We note that this matrix is similar to Table 4.3 except the constant terms on the right side.

Now we are going to use *Mathematica* to solve for the average number of steps needed to hit a boundary in Example 4.2.1. Also see Appendix D for the code. In this example, for all the starting points on the boundaries, $e_{i,j} = 0$. Therefore, we are interested in showing $e_{i,j}$ when starting at the interior points.

Example 4.4.1. The Average Number of Steps to Hit a Boundary or a corner Solution. With $m = n = 6$ and $p_l d = p_r d = p_l u = p_r u = 0.25$, the average number of steps needed to hit a boundary starting at each interior point is:

$$\begin{array}{ccccc}
e_{1,1} = 2.1429 & e_{1,2} = 3.0 & e_{1,3} = 3.2857 & e_{1,4} = 3.0 & e_{1,5} = 2.1429 \\
e_{2,1} = 3.0 & e_{2,2} = 4.5714 & e_{2,3} = 5.0 & e_{2,4} = 4.5714 & e_{2,5} = 3.0 \\
e_{3,1} = 3.2857 & e_{3,2} = 5.0 & e_{3,3} = 5.5714 & e_{3,4} = 5.0 & e_{3,5} = 3.2857 \\
e_{4,1} = 3.0 & e_{4,2} = 4.5714 & e_{4,3} = 5.0 & e_{4,4} = 4.5714 & e_{4,5} = 3.0 \\
e_{5,1} = 2.1429 & e_{5,2} = 3.0 & e_{5,3} = 3.2857 & e_{5,4} = 3.0 & e_{5,5} = 2.1429
\end{array}$$

We note that the boundaries form a 6×6 square, and the walk starting at the center needs the most steps on average to hit a boundary since the probabilities are evenly distributed and the center has the longest distance to the boundaries. The symmetric points, such as $(1, 2)$ and $(1, 4)$, have the same distance to the left or the right boundary, and the distance from $(1, 2)$ to the upper boundary is the same as the distance from $(1, 4)$ to the lower boundary, etc. Thus the results show the first row and the fifth row are the same, so are the second row and the fourth row.

Chapter 5

Applications to Two One-Dimensional Random Walks

In this chapter, we will apply the results we obtained from Chapter 4 to various boundary problems for two processes of one-dimensional random walks. The initial problem is to analyze two one-dimensional random walks as follows:

(1) Process X begins at positive integer height h and on each step moves either upward or downward one unit at a time with probabilities p_x and $q_x = 1 - p_x$, respectively. This process stops upon reaching boundaries of 0 or m;

(2) Process Y begins at positive integer height k and on each step moves either upward or downward one unit at a time with probabilities p_y and $q_y = 1 - p_y$, respectively. This process stops upon reaching boundaries of 0 or n;

We want to answer the following questions:

(i) What is the probability that Process X stops before Process Y? And the probability that Process Y stops before Process X?

(ii) What is the probability that they stop at the same time?

(iii) Given that Process X stops first, what is the probability that X has hit height m?

(iv) Given that they stop at the same time, what is the probability that they have both hit 0?

(v) What is the average number of steps needed for a process to stop? What is the average number of steps needed for both processes to stop?

For instance, suppose Process X starts at positive integer height $h = 6$, with $p_x = 0.60$ and $q_x = 1 - p_x = 0.40$. This process stops upon reaching boundaries of 0 or $m = 10$. There are uncountably many different paths. One example of Process X is demonstrated in Figure 5.1 (the horizontal axis represents the number of steps that are taken).

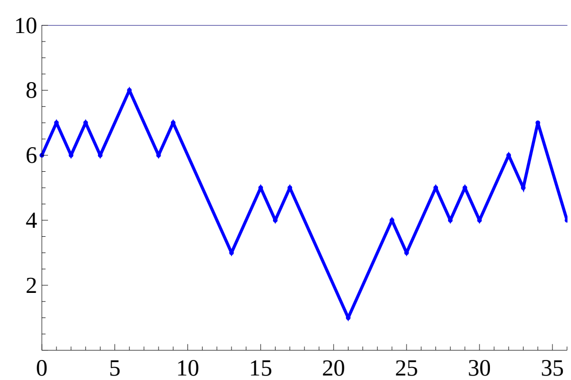

FIGURE 5.1. An example of Process X

Process Y begins at positive integer height $k = 10$, with $p_y = 0.20$ and $q_y = 1 - p_y = 0.80$. This process stops upon reaching boundaries of 0 or $n = 14$. An example of Process Y is demonstrated in Figure 5.2:

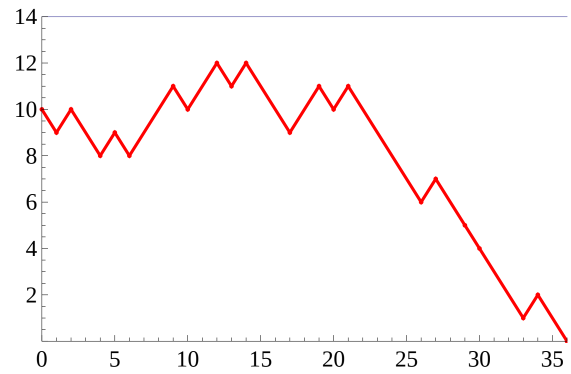

FIGURE 5.2. An example of Process Y

In these figures, we can see Process X never stops within 36 steps, and on the 36^{th} step Process Y hits 0 and stops. So for this specific path, we can say Y stops before X. Then one of the questions we are interested in is the overall probability of Y stopping before X.

Although we may not be able to find closed-form solutions, we can use the techniques of a Type II two-dimensional random walk to find numerical solutions. By converting two one-dimensional random walks to one Type II two-dimensional random walk, we are able to apply the results of Chapter 4 to determine the probabilities of various boundaries being hit first from the initial starting point, and the average number of steps needed to hit a boundary.

We begin with converting the example above to a Type II two-dimensional random walk with $m = 10$, $n = 14$, starting at $(h, k) = (6, 10)$. Hence, Y represents the vertical movements between 0 and 14, and X represents the horizontal movements between 0 and 10. Now on each step, the Type II two-dimensional random walk either moves left and downward with probability $p_{ld} = q_x \cdot q_y = 0.40 \times 0.80 = 0.32$, or moves left and upward with probability $p_{lu} = q_x \cdot p_y = 0.40 \times 0.20 = 0.08$, or moves right and upward with probability $p_{ru} = p_x \cdot p_y = 0.60 \times 0.20 = 0.12$, or moves right and downward with probability $p_{rd} = p_x \cdot q_y = 0.60 \times 0.80 = 0.48$. Now for Process X to stop upon reaching boundaries of 0 or m is to say the Type II random walk hits the left or the right boundaries; for Process Y to stop upon reaching boundaries of 0 or n is to say the Type II random walk hits the upper or the lower boundaries. Hitting the corners is when the two processes stop at the same time. Applying the results from Chapter 4, we are able to answer the previous questions.

Question(i). What is the probability that Process X stops before Process Y? And the probability that Process Y stops before Process X?

Solution. We want the probability that Process X stops at 0 or $m = 10$ before Process Y reaches 0 or $n = 14$. We can interpret this question as asking for the total probability of a Type II random walk hitting the points of $(0, j)$ or $(10, j)$ first, where $0 < j < 14$; i.e., the probability of hitting the left or the right boundaries before hitting the upper or the lower boundaries from the starting point $(h, k) = (6, 10)$. Using *Mathematica* (changing the parameters in Appendix C, we can easily obtain the Markov Chains solutions of the final state for each boundary and each corner being hit first. By trial and error, when $x \geq 58$, the total probability of each boundary and each corner being hit first sums to 1, and the results are:

$P(L_x)$	$P(U_x)$	$P(R_x)$	$P(D_x)$	Sum
0.0283	0.0028	0.5613	0.3500	0.9424

Because the events of the left boundary being hit first and the right boundary being hit first are mutually exclusive, the probability of hitting the left or the right boundaries before hitting the upper or the lower boundaries from the starting point $(h, k) = (6, 10)$ is simply the sum of $P(L_x) + P(R_x) = 0.5896$. This implies the probability that Process X stops before Process Y is 0.5896 (or 58.96%). Similarly, we can obtain the probability of Process Y stopping before Process X as the sum of $P(U_x) + P(D_x) = 0.3528$ (or 35.28%).

Question(ii). What is the probability that they stop at the same time?

Solution. To find the probability that Process X stops at 0 or $m = 10$ and Process Y stops at 0 or $n = 14$ at the same time, we only need to find the total probability of a Type II two-dimensional random walk hitting the four corners first, because there are only four combinations for the two processes to stop at the same time; they are either $(0,0), (0,14), (10,14)$ or $(10,0)$. With the same procedure using *Mathematica*, we have the following results:

$P(B_x^L)$	$P(T_x^L)$	$P(T_x^R)$	$P(B_x^R)$	Sum
0.0051	0.00001	0.0004	0.0521	0.0576

Thus, the probability of Process X and Process Y stopping at the same time is the total probability of the four corners being hit first, which is 0.0576 (or 5.76%).

Remark. The event of Process X and Process Y stopping at the same time is the complement of the event of either Process X stopping before Process Y or Process Y stopping before Process X. Thus we can also obtain the solution to this question by 1 – (0.9424) = 0.0576.

Question(iii). Given that Process X stops first, what is the probability that X has hit height 10?

Solution. We simply apply the rule for conditional probability and obtain

$$\begin{aligned} P(X \text{ has hit height } 10 \mid X \text{ stops first}) &= \frac{P(X \text{ has hit height 10 and stops first})}{P(X \text{ stops first})} \\ &= \frac{P(R_x)}{P(L_x) + P(R_x)} \\ &= \frac{0.5613}{0.5896} = 0.9520. \end{aligned}$$

Thus, the probability that X has hit height 10 given that Process X stops first is 0.9520 (or 95.20%).

Question(iv). Given that X and Y stop at the same time, what is the probability that they have both hit 0?

Solution. The probability of X and Y stop at the same time is the total probability when the four-sided Type II walk hitting the four corners first. Then by the rule of conditional probability again, we have

$$
\begin{aligned}
&P(X \text{ \& } Y \text{ have hit } 0 \mid X \text{ \& } Y \text{ stop at the same time}) \\
&= \frac{P(X \text{ \& } Y \text{ have hit } 0 \text{ and } X \text{ \& } Y \text{ stop at the same time})}{P(X \text{ and } Y \text{ stop at the same time})} \\
&= \frac{P(B_x^L)}{P(B_x^L) + P(T_x^L) + P(B_x^R) + P(T_x^R)} \\
&= \frac{0.0051}{0.0576} = 0.0885.
\end{aligned}
$$

Given that X and Y stop at the same time, the probability that they have both hit 0 is 0.0885 (or 8.85%).

Question(v). What is the average number of steps needed for a process to stop? What is the average number of steps needed for both processes to stop?

Solution. The average number of steps for Process X to stop from its starting point 6 is the average number of steps needed for a Type II two-dimensional random walk to hit the left boundary or the right boundary from a starting point $(6, 10)$. Likewise, the average number of steps for Process Y to stop from a starting point 10 is the average number of steps that a Type II two-dimensional random walk needs to hit the upper boundary or the lower boundary from $(6, 10)$. Thus the average number of steps needed for a process (either X or Y) to stop from their initial starting points can be understood as the average number of steps for a Type II two-dimensional random walk needed to hit any boundary from $(6, 10)$, which is $e_{6,10}$ in Section 4.4. Thus we can use the System of Equations method to find our desired solution. Entering the value of each corresponding probability and the boundaries $m = 10$ and $n = 14$ into *Mathematica* code (Appendix D, the output shows $e_{6,10} = 11.1526$, which is the average number of steps needed for a process (either X or Y) to stop from its starting point, where Process X starts at 6 and Process Y starts at 10.

To solve for the average number of steps needed for both processes to stop from their starting points, we first let s_X be the number of steps needed for Process X to stop from its starting point 6, and s_Y be the number of steps needed for Process Y to stop from 10. Then $\max\{s_X, s_Y\}$ is the number of steps which guarantees that *both* processes have hit a boundary and have stopped. Thus, we want the solution to $E[\max\{s_X, s_Y\}]$. We note that

$$s_X + s_Y = \min\{s_X, s_Y\} + \max\{s_X, s_Y\}.$$

By taking the average on both sides, and using the fact that the average of a sum is the sum of the averages, we have

$$E[s_X] + E[s_Y] = E[\min\{s_X, s_Y\}] + E[\max\{s_X, s_Y\}].$$

From Section 2.1, both $E[s_X]$ and $E[s_Y]$ are known by Equation 2.1.2. And we have $\min\{s_X, s_Y\} = s_{6,10}$, which is the number of step needed for a Type II two-dimensional random walk to hit a boundary; thus,

$$E[\min\{s_X, s_Y\}] = E[s_{6,10}] = e_{6,10}.$$

We now can obtain our desired solution by

$$
\begin{aligned}
E[\max\{s_X, s_Y\}] &= E[s_X] + E[s_Y] - e_{6,10} \\
&= E[{}_6T_0^{10}] + E[{}_{10}T_0^{14}] - 11.1526 \\
&= \left[\frac{10}{0.6-0.4} \times \left(\frac{1-(0.4/0.6)^6}{1-(0.4/0.6)^{10}}\right) - \frac{6}{0.6-0.4}\right] \\
&\quad + \left[\frac{14}{0.2-0.8} \times \left(\frac{1-(0.8/0.2)^{10}}{1-(0.8/0.2)^{14}}\right) - \frac{10}{0.2-0.8}\right] - 11.1526 \\
&= 21.9383
\end{aligned}
$$

Hence, the average number of steps needed for *both* processes to stop from their starting points is 21.9383.

Chapter 6

Downward-drifting Type II Walk

From Chapters 3 and 4, we can find the numerical solutions for four-sided boundary problems of Type I and Type II two-dimensional random walks. Also for downward-drifting Type I two-dimensional random walk, with the x-axis and the y-axis as the two boundaries, Neal [1] derived a lengthy formula for the probability of hitting one axis before the other. In this chapter, we aimed to adopt Neal's technique to derive a closed-form formula for the same question applying to two-sided downward-drifting Type II walk. However, we encountered unexpected difficulties in this case, so we will develop a conjecture for a special case when both coordinates of a starting point are even, then perform a statistical hypothesis test on the conjecture using *Mathematica* simulation data. Eventually, we wish to prove this conjecture and adjust it to apply for the cases when both coordinates are odd or one is odd and one is even.

6.1. Introduction

A two-sided Type II two-dimensional random walk starts at coordinates (h, k) in the first quadrant, with the x-axis and the y-axis as the two boundaries, and moves diagonally on each independent step one unit at a time in one of four directions with probabilities p_{ld}, p_{lu}, p_{rd}, and p_{ru}, respectively, shown as the figure below:

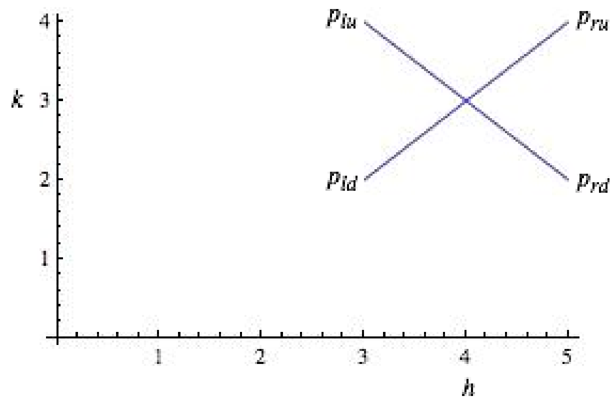

FIGURE 6.1. A Two-Sided Type II Walk

If we consider just the upward and downward movements, a one-dimensional random walk Y is created that begins at height k and moves upward or downward one unit at a time with probabilities $p_y = p_{lu} + p_{ru}$ and $q_y = p_{ld} + p_{rd}$, respectively. In order to drop to the x-axis with probability 1, we need $q_y \geq p_y$, and in order to do so with a finite expected number of steps, we need $q_y > p_y$. Thus we shall assume that $p_{ld} + p_{rd} > p_{lu} + p_{ru}$, so that the Type II two-dimensional random walk will almost surely reach the x-axis with a finite expected number of steps.

Likewise, if we consider just the leftward and rightward movements, a one-dimensional walk X is created that begins at height h and gains one unit or loses one unit at a time with probabilities $p_x = p_{rd} + p_{ru}$ and $q_x = p_{ld} + p_{lu}$, respectively. In order to drop to height 0 with probability 1 and with a finite expected number of steps, we need $q_x > p_x$. Thus we shall assume that $p_{ld} + p_{lu} > p_{rd} + p_{ru}$, so that the Type II two-dimensional random walk will almost surely reach the y-axis with a finite expected number of steps.

6.2. A Conjecture on Downward-Drifting Type II Walk

We are going to develop a conjecture for the probability of a certain downward-drifting Type II walk hitting the x-axis before ever hitting the y-axis. To do so, we will assume that both h and k are even. In this case, the only coordinates that can be hit on the x-axis, without hitting the y-axis are the points $(2,0)$, $(4,0)$, $(6,0),\cdots$. So we want to determine, or at least estimate, the probability of hitting any of these points before ever touching the y-axis.

For Y to hit the x-axis for the first time in exactly $k+2i$ steps, there must be i upward movements and $k+i$ downward movements for some $i \geq 0$. The coefficient ${}_kC_i = \frac{k}{k+2i}\binom{k+2i}{i}$ gives the number of ways for Y to move upward and downward $k+2i$ times while hitting the x-axis for the first time on the $(k+2i)^{th}$ step. So as stated in Section 2.3, the pmf for the number of steps T_Y needed for Y to drop to the x-axis is

$$\begin{aligned} P(T_Y = k+2i) &= \frac{k}{k+2i}\binom{k+2i}{i}(p_{lu} + p_{ru})^i (p_{ld} + p_{rd})^{k+i} \qquad (6.2.1) \\ &= {}_kC_i (p_y)^i (q_y)^{k+i}, \end{aligned}$$

for $i \geq 0$. We note that $\sum_{i=0}^{\infty} P(T_Y = k+2i) = 1$. Thus for large N we have $\sum_{i=0}^{N} P(T_Y = k+2i) \approx 1$. Using trial and error technique, we can find how large this N must be using *Mathematica*.

For each path that drops to the x-axis, we now consider the possible horizontal movements described by the process X. Using the *Reflection Principle* described by Feller [5], the number of ways for X to move from value h to value $2z$ in exactly $k+2i$ steps without ever reaching value 0 is given by

$$D(h,k,i,z) = \binom{k+2i}{\frac{k+2i-h}{2}+z} - \binom{k+2i}{\frac{k+2i+h}{2}+z},$$

for $1 \leq z \leq (h+k+2i)/2$. In this case, a two-sided Type II two-dimensional random walk will end at coordinates $(2z, 0)$, without ever touching the y-axis, and there must be $(k+2i+h)/2 - z$ leftward movements and $(k+2i-h)/2 + z$ rightward movements.

Thus, considering just the horizontal movements, the probability of the x-coordinate ending at value $2z$ in exactly $k+2i$ steps without ever touching the y-axis is

$$\begin{aligned} &\left(\binom{k+2i}{\frac{k+2i-h}{2}+z} - \binom{k+2i}{\frac{k+2i+h}{2}+z}\right)(p_{ld}+p_{lu})^{\frac{k+2i+h}{2}-z}(p_{rd}+p_{ru})^{\frac{k+2i-h}{2}+z} \\ &= D(h,k,i,z)(q_x)^{\frac{k+2i+h}{2}-z}(p_x)^{\frac{k+2i-h}{2}+z}, \end{aligned} \tag{6.2.2}$$

for $1 \leq z \leq (h+k+2i)/2$.

As an example, we consider the easiest case with $h = k = 2$. To do so, we consider the following chart that lists the numbers of possible paths:

i (U)	$2+i$ (D)	${}_2C_i$	$1 \leq z \leq \frac{2+2i+2}{2}$	$D(2,2,i,z)$	$\frac{2+2i+2}{2} - z$ (L)
0	2	1	$1 \leq z \leq 2$	2,1	1,0
1	3	2	$1 \leq z \leq 3$	5,4,1	2,1,0
2	4	5	$1 \leq z \leq 4$	14,14,6,1	3,2,1,0
3	5	14	$1 \leq z \leq 5$	42,48,27,8,1	4,3,2,1,0
4	6	42	$1 \leq z \leq 6$	132,165,110,44,10,1	5,4,3,2,1,0

We now shall analyze the third row when $i = 2$: Starting at $(h,k) = (2,2)$, the walk hits the x-axis in exactly $k + 2i = 6$ steps with 2 upward movements and 4 downward movements. There are ${}_2C_2 = 5$ ways for this to happen:

U,U,D,D,D,D U,D,U,D,D,D U,D,D,U,D,D

D,U,U,D,D,D D,U,D,U,D,D

It takes a few thoughts to list all the possible ways. First, the last two steps must be downward. Second, the first two steps cannot be both downward, otherwise the walk already reaches x-axis and no upward step is needed.

Now since $1 \leq z \leq 4$ and we want the path to end at some coordinate $(2z, 0)$ without hitting the y-axis, the path must end at one of the coordinates $(2,0)$, $(4,0)$, $(6,0)$, or $(8,0)$ in exactly 6 steps. Considering just the horizontal movements, there are 14 ways to end at $(2,0)$ and there must be 3 left movements and 3 right movements. There are 14 ways to end at $(4,0)$ and there must be 2 left movements and 4 right movements. There are 6 ways to end at $(6,0)$ and there must be 1 left movement and 5 right movements. There is 1 way to end at $(8,0)$ and there must be 0 left movements and 6 right movements.

Let us now consider the 14 ways to end at $(2,0)$:

L, R, L, R, L, R   L, R, R, L, R, L   L, R, R, R, L, L   L, R, R, L, L, R

L, R, L, R, R, L   R, L, R, L, R, L   R, L, L, R, L, R   R, L, R, R, L, L

R, L, R, L, L, R   R, L, L, R, R, L   R, R, R, L, L, L   R, R, L, L, R, L

R, R, L, L, L, R   R, R, L, R, L, L.

Any of these 14 L/R permutations can be matched with any of the 5 possible U/D permutations giving a total of 70 ways for a Type-II random walk to move from initial coordinates $(2,2)$ to final coordinates $(2,0)$ in exactly 6 steps without ever hitting the y-axis. Matching the first two permutations of each, we have LU, RU, LD, RD, LD, RD, which makes the process moves from $(2,2)$ to $(1,3)$ to $(2,4)$ to $(1,3)$ to $(2,2)$ to $(1,1)$ and to $(2,0)$. This particular path occurs with probability $p_{lu} \cdot p_{ru} \cdot p_{ld}^2 \cdot p_{rd}^2$. This path is shown as in Figure 6.2:

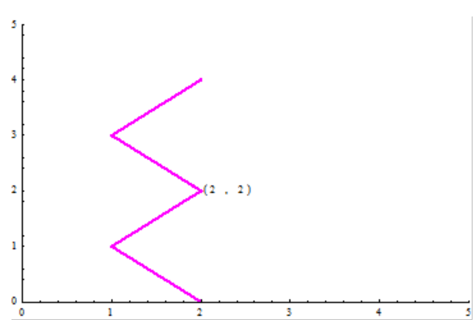

FIGURE 6.2. A Path From $(2,2)$ to $(2,0)$

We note that not all of these 70 matchings are equally likely because we do not always have the same number of RU steps. For instance, the matching of L, R, L, R, L, R with D, U, D, U, D, D gives LD, RU, LD, RU, LD, RD which occurs with probability $p_{ld}^3 \cdot p_{ru}^2 \cdot p_{rd}$.

To obtain the exact probability of hitting the x-axis first, we would have to enumerate further among the number of ${}_kC_i \cdot D(h,k,i,z)$ paths for $i \geq 0$ and $1 \leq z \leq (h+k+2i)/2$, to determine how many different ways there are to have various amounts of RU paths while still ending at $(2z,0)$, without hitting the y-axis, in exactly $k+2i$ steps. Instead, we estimate this desired probability with a formula which combines Equations 6.2.1 and 6.2.2.

Conjecture 6.2.1. Assume h and k are both even, with $h > 0$ and $k > 0$, and assume that $p_{ld} + p_{rd} > p_{lu} + p_{ru}$ and $p_{ld} + p_{lu} > p_{rd} + p_{ru}$. The probability of a two-sided Type II random walk hitting the x-axis before hitting the y-axis can be estimated by the formula

$$\sum_{i=0}^{\infty}\left(\sum_{z=1}^{\frac{h+k+2i}{2}} D(h,k,i,z)(q_x)^{\frac{k+2i+h}{2}-z}(p_x)^{\frac{k+2i-h}{2}+z}\right)P(T_Y = k+2i)$$
$$= \sum_{i=0}^{\infty}\left(\sum_{z=1}^{\frac{h+k+2i}{2}} D(h,k,i,z)(q_x)^{\frac{k+2i+h}{2}-z}(p_x)^{\frac{k+2i-h}{2}+z}\right){}_kC_i(p_y)^i(q_y)^{k+i},$$

Where $p_x = p_{rd} + p_{ru}$, $q_x = p_{ld} + p_{lu}$, $p_y = p_{lu} + p_{ru}$, and $q_y = p_{ld} + p_{rd}$.

6.3. Testing the Conjecture

At this point, rather than giving a formal proof of this conjecture, we aim to test its accuracy by using simulations and an appropriate hypothesis test for proportions. We are looking for the probability of a downward-drifting Type II walk starting at even-valued coordinates hitting the x-axis before ever hitting the y-axis. This event includes hitting any of the coordinates $(2,0)$, $(4,0)$, etc. We include $(0,0)$ on the y-axis. This formula appears to be verified with *Mathematica* simulations. However, even if it is true, when computing numerically, the initial infinite sum on i must be limited to a finite number, for instance, i from 0 to 100. Therefore, if we wish to have a numerical approximation of this probability, it is much easier to run a simulation. For even with several thousand trials, the simulation concludes quickly and still gives a close approximation of the desired probability. We now perform the statistical hypothesis test through some examples.

Example 6.3.1. Let $h = 4$, $k = 10$, $p_{ld} = 0.35$, $p_{rd} = 0.40$, $p_{lu} = 0.20$, and $p_{ru} = 0.05$, where the downward-drifting condition $p_{ld} + p_{rd} > p_{lu} + p_{ru}$ is satisfied. Now using trial and error, with $N \geq 50$ trials for the first summation in *Mathematica* (see Appendix E), the output for $\sum_{i=0}^{N} P(T_Y = k+2i)$ reaches 1, implying the distribution of Y completes. For the particular $i = 4$, the number of ways for Y to move upward and downward $k + 2i = 18$ times while hitting the x-axis for the first time on the 18^{th} step is ${}_{10}C_4 = 1700$. The coordinates on the x-axis that are possible to be hit are $(2,0)$, $(4,0)$, $\cdots$, $(22,0)$. The number of ways $D(h,k,i,z)$ for X to move from value $h = 4$ to value $2z$ where $1 \leq z \leq 11$ in exactly 18 steps without ever reaching value 0 are given by:

$$\{2, 25194\}, \{4, 40052\}, \{6, 40698\}, \{8, 31008\}, \{10, 18411\},$$

$$\{12, 8550\}, \{14, 3059\}, \{16, 816\}, \{18, 153\}, \{20, 18\}, \{22, 1\}.$$

For example, there are 25194 possible ways for the horizontal movements X to move from $h = 4$ to coordinate $(2, 0)$ in exactly 18 steps without ever reaching value $x = 0$, and there is only 1 way to reach $(22, 0)$ in such manner.

Below we show two example paths simulated by *Mathematica*. One path hits the y-axis first and stops, while the other hits the x-axis first and stops.

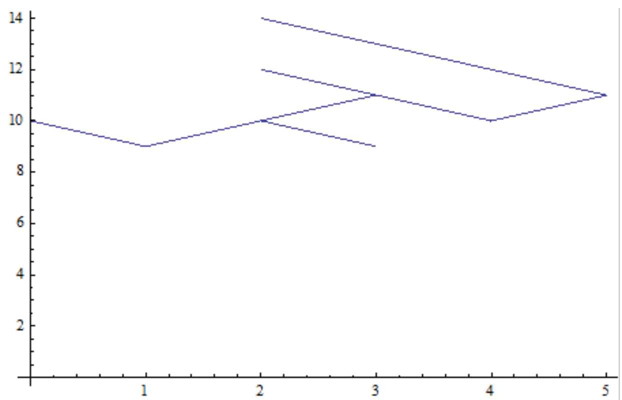

FIGURE 6.3. **Example** 6.3.1 Hitting the y-axis first From $(4, 10)$

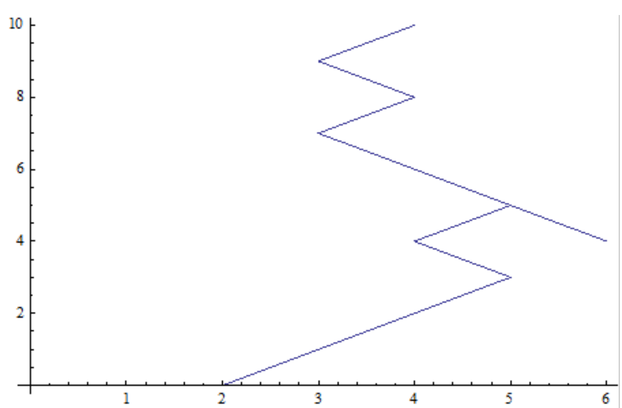

FIGURE 6.4. **Example** 6.3.1 Hitting the x-axis first From $(4, 10)$

We cannot tell by looking at the graph how many steps it actually takes to move from one coordinate to the other because there may be some overlapping steps in between.

The result given by the conjecture for the probability of hitting the x-axis first is Conj = 0.479502. Using 9200 trials to run the simulation several times, we will have different outputs for the probability of hitting the x-axis first. We randomly pick one sample which has output Sim = 0.480217 for our test. Now we claim that for this example the probability of hitting the x-axis before hitting the y-axis is $P(A_{x,y}) = 0.479502$. So we set up the hypothesis as follows:

$$\text{Null hypothesis} \quad H_0 : P(A_{x,y}) = 0.479502;$$

$$\text{Alternative hypothesis} \quad H_a : P(A_{x,y}) \neq 0.479502.$$

We choose the general-purpose significance level $\alpha = 0.05$ to test the accuracy of our conjecture. Since we use large enough random samples, $n = 9200$ trials, we can conduct two-tailed Z-test. The test statistic is computed as

$$z_0 = \frac{\text{Sim} - \text{Conj}}{\sqrt{\dfrac{\text{Conj} \cdot (1 - \text{Conj})}{n}}} = 0.137383.$$

As we use a two-tailed test, we have

$$\text{P-Value} = 2 \times P(z > 0.137383) = 0.890728.$$

Since P-Value > 0.05, there is insignificant evidence to show that the alternative hypothesis is true, thus we fail to reject the null hypothesis. However, we cannot state that the conjecture is proved to be correct, only that it gives a good approximation in this example. To see whether our conjecture works for other cases, we are going to test two more examples with different entries.

Example 6.3.2. Let $h = 6$, $k = 4$, $p_{ld} = 0.35$, $p_{rd} = 0.28$, $p_{lu} = 0.32$, and $p_{ru} = 0.05$. Again the downward-drifting condition is satisfied. Using $N \geq 150$ for the first summation in *Mathematica* code (see Appendix F), we note that the distribution of Y completes as the output for $\sum_{i=0}^{N} P(T_Y = k + 2i)$ reaches 1. For the particular $i = 5$, we have ${}_4C_5 = 572$ possible ways for the vertical movements Y to move $k + 2i = 14$ times while hitting the x-axis for the first time on the 14^{th} step. As $1 \leq z \leq 10$, only the following coordinates on the x-axis can be hit in exactly 14 steps: $(2,0)$, $(4,0)$, $(6,0)$, $(8,0)$, $(10,0)$, $(12,0)$, $(14,0)$, $(16,0)$, $(18,0)$, $(20,0)$. The number of ways $D(h,k,i,z)$ for the horizontal movements X to move from value $h = 6$ to value $2z$ in exactly 14 steps without ever reaching value 0 are given by:

$$\{2,1638\},\{4,2912\},\{6,3418\},\{8,3002\},\{10,2002\},$$

$$\{12,1001\},\{14,364\},\{16,91\},\{18,14\},\{20,1\}.$$

From the data above, we notice that from $h = 6$ to coordinate $(6,0)$, there are the most ways for the possible horizontal movements X to move in exactly 14 steps without ever reaching value 0. Using $n = 10000$ trials, the probability of hitting the x-axis first given by running the simulation, choosing a random sample Sim, comparing to the result given by the conjecture, Conj, are as follows:

$$\text{Conj} : 0.588113$$

$$\text{Sim} : 0.5906.$$

We claim that for this example the probability of hitting the x-axis before ever hitting the y-axis is $P(A_{x,y}) = 0.588113$. Now we set up the hypothesis to test the accuracy of the conjecture:

$$\text{Null hypothesis} \quad H_0 : P(A_{x,y}) = 0.588113;$$

$$\text{Alternative hypothesis} \quad H_a : P(A_{x,y}) \neq 0.588113.$$

Again we use the significance level $\alpha = 0.05$. For random samples $n = 10000$ trials, the test statistics is computed by

$$z_0 = \frac{\text{Sim} - \text{Conj}}{\sqrt{\frac{\text{Conj}\cdot(1-\text{Conj})}{n}}} = 0.505394.$$

For a two-tailed test, we have

$$\text{P-Value} = 2 \times P(z > 0.505394) = 0.613282.$$

The P-Value is again larger than the significance level $\alpha = 0.05$, we fail to reject the null hypothesis, although the conjecture seems to be a little less accurate than the previous example; thus, the conjecture is still good. Let us look at one more example with the starting point much further away from the origin.

Example 6.3.3. Let $h = 18$, $k = 16$, $p_{ld} = 0.45$, $p_{rd} = 0.30$, $p_{lu} = 0.15$, and $p_{ru} = 0.10$. In this case $\sum_{i=0}^{N} P(T_Y = k + 2i)$ reaches 1 when $N \geq 60$ (see Appendix G for the code). For $i = 6$, we have ${}_{16}C_6 = 215280$ ways for the vertical movements Y to move $k + 2i = 28$ times while hitting the x-axis for the first time on the 28^{th} step. As $1 \leq z \leq 23$, there are 23 coordinates on the x-axis that can be hit: $(2,0)$, $(4,0)$,$\cdots$, $(46,0)$. The number of ways $D(h,k,i,z)$ for the horizontal movements X to move from value $h = 18$ to value $2z$ in exactly 28 steps without ever reaching value 0 are given by:

$$\{2,356265\},\ \{4,1180764\},\ \{6,3107727\},\ \{8,6906872\},\ \{10,13123109\},$$
$$\{12,21474180\},\ \{14,30421755\},\ \{16,37442160\},\ \{18,40116600\},\ \{20,37442160\},$$
$$\{22,30421755\},\ \{24,21474180\},\ \{26,13123110\},\ \{28,6906900\},\ \{30,3108105\},$$
$$\{32,1184040\},\ \{34,376740\},\ \{36,98280\},\ \{38,20475\},\ \{40,3276\},\ \{42,378\},$$
$$\{44,28\},\ \{46,1\}.$$

We notice that there is a symmetric feature between coordinates $(16,0)$ and $(20,0)$, $(14,0)$ and $(20,0)$, etc. Also, there is only one way for the possible horizontal movements X to move from $h = 18$ to the furthest coordinate $(46,0)$ that is possible to reach in exactly 28 steps without ever reaching value 0.

Using $n = 12000$ trials, we ran two sets of trial simultaneously, each stopped when an axis was hit. The two sets of output for the probability of hitting the

x-axis first are:

$$\mathrm{Conj} : 0.952916 \qquad\qquad \mathrm{Conj} : 0.952916$$

$$\mathrm{Sim} : 0.95325; \qquad\qquad \mathrm{Sim} : 0.953083$$

The result on the right set shows that the conjecture is more accurate comparing to the result given by the simulation. We are going to test the conjecture using the left set. So we claim that for this example the probability of hitting the x-axis before ever hitting the y-axis is $P(A_{x,y}) = 0.952916$. Then the hypotheses are:

$$\text{Null hypothesis } \; H_0 : P(A_{x,y}) = 0.952916;$$

$$\text{Alternative hypothesis } \; H_a : P(A_{x,y}) \neq 0.952916.$$

Again we use the significance level $\alpha = 0.05$. For random sample $n = 12000$ trials, the test statistics is 0.172857, and the P-Value is 0.862764, which is much greater than the significance level. We again fail to reject the null hypothesis in this case.

We have applied a hypothesis test on three examples, respectively. On each example we failed to reject the null hypothesis with a significance level $\alpha = 0.05$. This does not prove our conjecture, yet it does show our conjecture gives a nice approximation to our desired probability. For the next step as future work, we can seek a more rigorous proof for our conjecture and eventually find a closed-form formula to apply any case of starting coordinates.

Chapter 7

Conclusion and Future Work

We have been working on the boundary problems of one-dimensional and two-dimensional random walks in this thesis. In the case of single boundary problems for a one-dimensional random walk, we derived the moment generating function of the stopping time and the formulas for its variance and standard deviation. For four-sided boundary problems of a two-dimensional random walk, either the Type I walk or the Type II walk, we found the numerical solutions of one boundary being hit before the other three, and the average number of steps needed for the walk to hit one boundary from the initial starting point. With the results we found in Type II random walk, we were able to apply them to answer various boundary problems for two one-dimensional random walks. At the end of this thesis, we deveopped a conjecture to approximate the probability of a two-sided downward-drifting Type II random walk starting at even-valued coordinate hitting one axis before the other, and used the simulations to test its accuracy.

In the future, besides giving an elegant proof of the conjecture that we stated in Chapter 6, there are at least three different directions that we are interested in. Firstly, we can combine Type I and Type II two-dimensional random walks together; that is, a walk starts at an initial point, and on each step, the process moves independently toward eight different directions one unit at a time. Can we still answer the similar problems for this type of random walk with the Markov Chains method and the System of Equations method?

Secondly, suppose the process of a four-sided two-dimensional random walk, either Type I or Type II, does not stop when it hits the boundaries. What is the probability for the walk to ever come back to its initial starting point, and what is the average number of steps for that goal to fulfill? Thirdly, we may consider a random walk with a fixed initial starting point and a fixed number of steps to the boundaries, but varying probabilities of moving toward a direction on each step. So far we have been working on the random walks with a fixed probability toward the same direction on each step, various initial points, and various number of steps to the boundaries.

As a motivating example, a game is played where two players bet on one end of a number line, which has zero in the middle, and five steps to each end. The game starts by drawing a card randomly from a deck, with replacement, and placing it at zero. Now a player draws one card randomly. If the number on the card is greater than the previous card, the card is replaced and is moved one step toward the side this player bets on. On the other hand, if the number is smaller, then the card is replaced and is moved one step toward the opposite side, and the previous card remains when the new drawn card has the same value. The game ends when one end is reached. The player wins when the end he or she bets on is reached first. This is an example of a symmetric two-sided one-dimensional random walk with different probabilities on each step, because on each step the probability of moving towards one end or the other, or remaining at the same place, depends on the previous card. However, the starting point is fixed, and the steps to each end are symmetric and fixed. The common questions on the boundary problems of two-dimensional random walks can also be asked in this case. What will the probability of one end being reached before the other be? What is the average number of steps to reach one end? This problem may be studied in the future.

Appendix A

The Markov Chains Solution for Type I Walk

(Applied to Example 3.3.1; change parameters to use for other examples)

Enter Probabilities and Right Boundary m and Upper Boundary n:

```
pl = 0.25; pr = 0.25; pd = 0.25; pu = 0.25;
m = 6; n = 6;
```

Enter Initial Position and Create Initial State Matrix:

```
h = 3; k = 3;
Do[b[i, j] = 0, {i, 1, 1}, {j, 1, (m + 1)(n + 1)}]; b[1, h(n + 1) + k + 1] = 1;
Do[d[i, j] = 0, {i, 1, 2}, {j, 1, (m + 1)(n + 1)}];
Do[Do[d[1, i(n + 1) + j + 1] = {i, j}, {j, 0, n}], {i, 0, m}]
d[2, h(n + 1) + k + 1] = 1;
MatrixForm[Table[d[i, j], {i, 1, 2}, {j, 1, (m + 1)(n + 1)}]]
```

Create Transition Matrix:

```
Do[a[i, j] = 0, {i, 0, (m + 1)(n + 1)}, {j, 0, (m + 1)(n + 1)}];
Do[a[i, i] = 1, {i, 1, n + 1}];
Do[a[i, i] = 1, {i, m(n + 1) + 1, (m + 1)(n + 1)}];
Do[a[k * (n + 1) + 1, k * (n + 1) + 1] = 1, {k, 1, m − 1}];
Do[a[k * (n + 1), k * (n + 1)] = 1, {k, 2, m}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j] = pd, {k, 1, m − 1}, {j, 1, n − 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + 2 + j] = pu, {k, 1, m − 1}, {j, 1, n − 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j − n] = pl, {k, 1, m − 1}, {j, 1, n − 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j + n + 2] = pr, {k, 1, m − 1}, {j, 1, n − 1}];
MatrixForm[Table[a[i, j], {i, 1, (m + 1)(n + 1)}, {j, 1, (m + 1)(n + 1)}]]
```

Enter Maximum Number of Steps.
Output Gives Final State After the Maximum Number of Steps:

```
MaximumNumberOfSteps = max = 600;
A = Table[a[i, j], {i, 1, (m + 1)(n + 1)}, {j, 1, (m + 1)(n + 1)}];
B = Table[b[i, j], {i, 1, 1}, {j, 1, (m + 1)(n + 1)}];
Z = B.MatrixPower[A, max];
Do[e[i, j] = 0, {i, 1, 2}, {j, 1, (m + 1)(n + 1)}];
Do[Do[e[1, i(n + 1) + j + 1] = {i, j}, {j, 0, n}], {i, 0, m}]
Do[e[2, i] = Z[[1]][[i]], {i, 1, (m + 1)(n + 1)}];
MatrixForm[Table[e[i, j], {i, 1, 2}, {j, 1, (m + 1)(n + 1)}]]
```

Probability of Hitting Boundaries:

```
Lft = Sum[Z[[1]][[i]], {i, 1, (n + 1)}];
Rght = Sum[Z[[1]][[i]], {i, m(n + 1) + 1, (m + 1)(n + 1)}];
Lower = Sum[Z[[1]][[i(n + 1) + 1]], {i, 1, m − 1}];
Upper = Sum[Z[[1]][[i(n + 1) + n + 1]], {i, 1, m − 1}];
S = Lft + Rght + Upper + Lower;
MatrixForm[{{"Left", "Upper", "Right", "Lower", "Sum"},
{Lft, Upper, Rght, Lower, S}}]
```

Simulation

```
NumberOfIterations = num = 5000;
Do[x[s, 0] = h, {s, 1, num}]; Do[y[s, 0] = k, {s, 1, num}];
p[1] = pl; p[2] = pr; p[3] = pd; p[4] = pu;
t[0] = 0; Do[t[i] = t[i − 1] + p[i], {i, 1, 4}];
```

Stop When Hit Boundary or Make the Maximum Number of Steps

```
Do[i = 1;
While[(0 < x[s, i - 1] < m)&&(0 < y[s, i - 1] < n)&&i ≤ max, z = Random[];
x[s, i] = If[z ≤ t[1], x[s, i - 1] - 1, If[z ≤ t[2], x[s, i - 1] + 1, x[s, i - 1]]];
y[s, i] = If[z > t[3], y[s, i - 1] + 1, If[z > t[2], y[s, i - 1] - 1, y[s, i - 1]]];
r[s] = i; i = i + 1], {s, 1, num}]

For[s = 1, s ≤ 10, s++, Print[ListLinePlot[Table[{x[s, j], y[s, j]},
{j, 0, r[s]}], AxesOrigin → {0, 0}]]]
```

Sample Average Number of Steps Needed to Hit a Boundary:

```
N[Mean[Table[r[s], {s, 1, num}]]]
```

Compare Simulation with Theoretical:

```
Do[w[s] = If[x[s, r[s]] == 0, 1, 0], {s, 1, num}];
We = N[Sum[w[s], {s, 1, num}]/num];
Do[b[s] = If[y[s, r[s]] == 0, 1, 0], {s, 1, num}];
So = N[Sum[b[s], {s, 1, num}]/num];
Do[e[s] = If[x[s, r[s]] == m, 1, 0], {s, 1, num}];
Ea = N[Sum[e[s], {s, 1, num}]/num];
Do[t[s] = If[y[s, r[s]] == n, 1, 0], {s, 1, num}];
No = N[Sum[t[s], {s, 1, num}]/num];
Si = We + No + Ea + So;
MatrixForm[{{" ", "Left", "Upper", "Right", "Lower", "Sum"},
{"Theory", Lft, Upper, Rght, Lower, S}, {"Sim", We, No, Ea, So, Si}}]
```

Appendix B

The System of Equations Solution for Type I Walk

(Applied to Example 3.3.1; change parameters to use for other examples)

To derive the Probabilities of Hitting the Left Boundary First:
Enter Probabilities and Right Boundary m and Upper Boundary n:

```
pl = 0.25; pr = 0.25; pd = 0.25; pu = 0.25; m = 6; n = 6;
```

Create Matrix of Coefficients:

```
Do[a[i, j] = 0, {i, 0, (m + 1)(n + 1)}, {j, 0, (m + 1)(n + 1)}];
Do[a[i, i] = 1, {i, 1, n + 1}];
Do[a[i, i] = 1, {i, m(n + 1) + 1, (m + 1)(n + 1)}];
Do[a[k * (n + 1) + 1, k * (n + 1) + 1] = 1, {k, 1, m - 1}];
Do[a[k * (n + 1), k * (n + 1)] = 1, {k, 2, m}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j] = pd, {k, 1, m - 1}, {j, 1, n - 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + 2 + j] = pu, {k, 1, m - 1}, {j, 1, n - 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + 1 + j] = -1, {k, 1, m - 1}, {j, 1, n - 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j - n] = pl, {k, 1, m - 1}, {j, 1, n - 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j + n + 2] = pr, {k, 1, m - 1}, {j, 1, n - 1}];
A = Table[a[i, j], {i, 1, (m + 1)(n + 1)}, {j, 1, (m + 1)(n + 1)}];
MatrixForm[A]
```

Create Matrix of Constants:

```
Do[f[i, j] = 0, {i, 1, (m + 1)(n + 1)}, {j, 1, 1}];
Do[f[i, j] = 1, {i, 1, (n + 1)}, {j, 1, 1}];
F = Table[f[i, j], {i, 1, (m + 1)(n + 1)}, {j, 1, 1}];
```

Solve the System:

```
Z = Inverse[A].F;
Do[e[i, j] = 0, {i, 1, 2}, {j, 1, (m + 1)(n + 1)}];
Do[Do[e[1, i(n + 1) + j + 1] = {i, j}, {j, 0, n}], {i, 0, m}]
Do[e[2, i] = Z[[i]][[1]], {i, 1, (m + 1)(n + 1)}];
MatrixForm[Table[e[i, j], {i, 1, 2}, {j, 1, (m + 1)(n + 1)}]]
```

Find the Average Time to Hit a Boundary.
Create the Augmented Column:

```
Do[f[i, j] = 0, {i, 1, (m + 1)(n + 1)}, {j, 1, 1}];
Do[Do[f[k * (n + 1) + 1 + j, 1] = -1, {k, 1, m - 1}], {j, 1, n - 1}];
F = Table[f[i, j], {i, 1, (m + 1)(n + 1)}, {j, 1, 1}]
```

```
Z = Inverse[A].F;
Do[e[i, j] = 0, {i, 1, 2}, {j, 1, (m + 1)(n + 1)}];
Do[Do[e[1, i(n + 1) + j + 1] = {i, j}, {j, 0, n}], {i, 0, m}]
Do[e[2, i] = Z[[i]][[1]], {i, 1, (m + 1)(n + 1)}];
MatrixForm[Table[e[i, j], {i, 1, 2}, {j, 1, (m + 1)(n + 1)}]]
```

Appendix C

The Markov Chains Solution for Type II Walk

(Applied to Example 4.2.1; change parameters to use for other examples)

Enter Probabilities and Right Boundary m and Upper Boundary n:

```
pld = 0.25; plu = 0.25; pru = 0.25; prd = 0.25;
m = 6; n = 6;
```

Enter Initial Position and Create Initial State Matrix:

```
h = 3; k = 3;
Do[b[i, j] = 0, {i, 1, 1}, {j, 1, (m + 1)(n + 1)}];
b[1, h(n + 1) + k + 1] = 1;
Do[d[i, j] = 0, {i, 1, 2}, {j, 1, (m + 1)(n + 1)}];
Do[Do[d[1, i(n + 1) + j + 1] = {i, j}, {j, 0, n}], {i, 0, m}]
d[2, h(n + 1) + k + 1] = 1;
MatrixForm[Table[d[i, j], {i, 1, 2}, {j, 1, (m + 1)(n + 1)}]]
```

Create Transition Matrix:

```
Do[a[i, j] = 0, {i, 0, (m + 1)(n + 1)}, {j, 0, (m + 1)(n + 1)}];
Do[a[i, i] = 1, {i, 1, n + 1}];
Do[a[i, i] = 1, {i, m(n + 1) + 1, (m + 1)(n + 1)}];
Do[a[k * (n + 1) + 1, k * (n + 1) + 1] = 1, {k, 1, m - 1}];
Do[a[k * (n + 1), k * (n + 1)] = 1, {k, 2, m}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j - n - 1] = pld, {k, 1, m - 1}, {j, 1, n - 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j + n + 1] = prd, {k, 1, m - 1}, {j, 1, n - 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j - n + 1] = plu, {k, 1, m - 1}, {j, 1, n - 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j + n + 3] = pru, {k, 1, m - 1}, {j, 1, n - 1}];
MatrixForm[Table[a[i, j], {i, 1, (m + 1)(n + 1)}, {j, 1, (m + 1)(n + 1)}]];
```

Enter Maximum Number of Steps.
Output Gives Final State After the Maximum Number of Steps:

```
MaximumNumberOfSteps = max = 10;
A = Table[a[i, j], {i, 1, (m + 1)(n + 1)}, {j, 1, (m + 1)(n + 1)}];
B = Table[b[i, j], {i, 1, 1}, {j, 1, (m + 1)(n + 1)}];
Z = B.MatrixPower[A, max];
Do[e[i, j] = 0, {i, 1, 2}, {j, 1, (m + 1)(n + 1)}];
Do[Do[e[1, i(n + 1) + j + 1] = {i, j}, {j, 0, n}], {i, 0, m}]
Do[e[2, i] = Z[[1]][[i]], {i, 1, (m + 1)(n + 1)}];
MatrixForm[Table[e[i, j], {i, 1, 2}, {j, 1, (m + 1)(n + 1)}]]
```

Probability of Hitting Boundaries:

```
Lft = Sum[Z[[1]][[i]], {i, 2, n}];
Rght = Sum[Z[[1]][[i]], {i, m(n + 1) + 2, (m + 1)(n + 1) - 1}];
Lower = Sum[Z[[1]][[i(n + 1) + 1]], {i, 1, m - 1}];
Upper = Sum[Z[[1]][[i(n + 1) + n + 1]], {i, 1, m - 1}];
OO = Z[[1]][[1]]; ON = Z[[1]][[n + 1]];
mn = Z[[1]][[(m + 1)(n + 1)]]; mO = Z[[1]][[m(n + 1) + 1]];
S = Lft + Rght + Upper + Lower + OO + ON + mn + mO;
MatrixForm[{{"Left", "Upper", "Right", "Lower",
"00", "0n", "mn", "m0", "Sum"},
{Lft, Upper, Rght, Lower, OO, ON, mn, mO, S}}]
```

Simulation

```
NumberOfIterations = num = 5000;
Do[x[s, 0] = h, {s, 1, num}]; Do[y[s, 0] = k, {s, 1, num}];
p[1] = pld; p[2] = plu; p[3] = pru; p[4] = prd;
t[0] = 0; Do[t[i] = t[i - 1] + p[i], {i, 1, 4}];
```

Stop When Hit Boundary or Make the Maximum Number of Steps

```
Do[i = 1; While[(0 < x[s, i - 1] < m)
&&(0 < y[s, i - 1] < n)&&i ≤ max,
z = Random[]; x[s, i] = If[z ≤ t[2], x[s, i - 1] - 1, x[s, i - 1] + 1];
y[s, i] = If[t[1] ≤ z < t[3], y[s, i - 1] + 1, y[s, i - 1] - 1];
r[s] = i; i = i + 1], {s, 1, num}]

For[s = 1, s ≤ 10, s++,
Print[ListLinePlot[Table[{x[s, j], y[s, j]}, {j, 0, r[s]}],
AxesOrigin → {0, 0}]]]
```

Sample Average Number of Steps Needed to Hit a Boundary:

```
N[Mean[Table[r[s], {s, 1, num}]]]
```

Compare Simulation with Theoretical:

```
Do[w[s] = If[x[s, r[s]] == 0&&0 < y[s, r[s]] < n, 1, 0], {s, 1, num}];
We = N[Sum[w[s], {s, 1, num}]/num];
Do[b[s] = If[y[s, r[s]] == 0&&0 < x[s, r[s]] < m, 1, 0], {s, 1, num}];
So = N[Sum[b[s], {s, 1, num}]/num];
Do[e[s] = If[x[s, r[s]] == m&&0 < y[s, r[s]] < n, 1, 0], {s, 1, num}];
Ea = N[Sum[e[s], {s, 1, num}]/num];
Do[t[s] = If[y[s, r[s]] == n&&0 < x[s, r[s]] < m, 1, 0], {s, 1, num}];
No = N[Sum[t[s], {s, 1, num}]/num];
Do[w[s] = If[x[s, r[s]] == 0&&y[s, r[s]]==0, 1, 0], {s, 1, num}];
dd = N[Sum[w[s], {s, 1, num}]/num];
Do[b[s] = If[x[s, r[s]] == 0&&y[s, r[s]]==n, 1, 0], {s, 1, num}];
du = N[Sum[b[s], {s, 1, num}]/num];
Do[e[s] = If[x[s, r[s]] == m&&y[s, r[s]]==n, 1, 0], {s, 1, num}];
```

```
uu = N[Sum[e[s], {s, 1, num}]/num];
Do[t[s] = If[x[s, r[s]] == m&&y[s, r[s]]==0, 1, 0], {s, 1, num}];
ud = N[Sum[t[s], {s, 1, num}]/num];
Si = We + No + Ea + So + dd + du + uu + ud;
MatrixForm[{{" ", "Left", "Upper", "Right", "Lower",
"00", "0n", "mn", "m0", "Sum"},
{"Theory", Lft, Upper, Rght, Lower, OO, ON, mn, mO, S},
{"Sim", We, No, Ea, So, dd, du, uu, ud, Si}}]
```

Appendix D

The System of Equations Solution for Type II Walk

(Applied to Example 4.2.1; change parameters to use for other examples)

To derive the Probabilities of Hitting the Left Boundary First
Enter Probabilities and Right Boundary m and Upper Boundary n:

```
pld = 0.25; plu = 0.25; pru = 0.25; prd = 0.25;
m = 6; n = 6;
```

Create Matrix of Coefficients:

```
Do[a[i, j] = 0, {i, 0, (m + 1)(n + 1)}, {j, 0, (m + 1)(n + 1)}];
Do[a[i, i] = 1, {i, 1, n + 1}];
Do[a[i, i] = 1, {i, m(n + 1) + 1, (m + 1)(n + 1)}];
Do[a[k * (n + 1) + 1, k * (n + 1) + 1] = 1, {k, 1, m - 1}];
Do[a[k * (n + 1), k * (n + 1)] = 1, {k, 2, m}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j - n - 1] = pld, {k, 1, m - 1}, {j, 1, n - 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j + n + 1] = prd, {k, 1, m - 1}, {j, 1, n - 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j - n + 1] = plu, {k, 1, m - 1}, {j, 1, n - 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + j + n + 3] = pru, {k, 1, m - 1}, {j, 1, n - 1}];
Do[a[k * (n + 1) + 1 + j, k * (n + 1) + 1 + j] = -1, {k, 1, m - 1}, {j, 1, n - 1}];
A = Table[a[i, j], {i, 1, (m + 1)(n + 1)}, {j, 1, (m + 1)(n + 1)}];
MatrixForm[A]
```

Create Matrix of Constants:

```
Do[f[i, j] = 0, {i, 1, (m + 1)(n + 1)}, {j, 1, 1}]; Do[f[i, j] = 1, {i, 2, n}, {j, 1, 1}];
F = Table[f[i, j], {i, 1, (m + 1)(n + 1)}, {j, 1, 1}]
```

Solve the System:

$Z = \text{Inverse}[A].F;$

$\text{Do}[e[i,j] = 0, \{i,1,2\}, \{j,1,(m+1)(n+1)\}];$

$\text{Do}[\text{Do}[e[1, i(n+1)+j+1] = \{i,j\}, \{j,0,n\}], \{i,0,m\}]$

$\text{Do}[e[2,i] = Z[[i]][[1]], \{i,1,(m+1)(n+1)\}];$

$\text{MatrixForm}[\text{Table}[e[i,j], \{i,1,2\}, \{j,1,(m+1)(n+1)\}]]$

Find the Average Time to Hit a Boundary

$\text{Do}[f[i,j] = 0, \{i,1,(m+1)(n+1)\}, \{j,1,1\}];$

$\text{Do}[\text{Do}[f[k * (n+1)+1+j, 1] = -1, \{k,1,m-1\}], \{j,1,n-1\}];$

$F = \text{Table}[f[i,j], \{i,1,(m+1)(n+1)\}, \{j,1,1\}];$

$Z = \text{Inverse}[A].F;$

$\text{Do}[e[i,j] = 0, \{i,1,2\}, \{j,1,(m+1)(n+1)\}];$

$\text{Do}[\text{Do}[e[1, i(n+1)+j+1] = \{i,j\}, \{j,0,n\}], \{i,0,m\}]$

$\text{Do}[e[2,i] = Z[[i]][[1]], \{i,1,(m+1)(n+1)\}];$

$\text{MatrixForm}[\text{Table}[e[i,j], \{i,1,2\}, \{j,1,(m+1)(n+1)\}]]$

Appendix E

Testing the Conjecture in Example 6.3.1

(Applied to Example 6.3.1; change parameters to use for other examples)

Enter Parameters:

$h = 4; k = 10;$

$\text{pld} = .35; \text{prd} = 0.40; \text{plu} = 0.20; \text{pru} = 0.05;$

Use Trial and Error to Determine
when the distribution of Y is essentially complete:

$k * \sum_{i=0}^{50} \text{Binomial}[k + 2i, i]/(k + 2i) * (\text{prd} + \text{pld})^{k+i} * (\text{plu} + \text{pru})^{i}$

For a particular i, display kCi and all D(h, k, i, z):

$i = 4;$

$k * \text{Binomial}[k + 2i, i]/(k + 2i)$

$\text{Table}[\{h + (k + 2i - h)/2 + z - ((k + 2i + h)/2 - z),$

$\text{Binomial}[k + 2i, (k + 2i - h)/2 + z] - \text{Binomial}[k + 2i, (k + 2i + h)/2 + z]\},$

$\{z, 1, (h + k + 2i)/2\}]$

$\text{Conj} =$

$k*$

$\sum_{i=0}^{50} \text{Binomial}[k + 2i, i]/(k + 2i) * (\text{prd} + \text{pld})^{k+i} * (\text{plu} + \text{pru})^{i}$

$* \sum_{z=1}^{\frac{k+2i+h}{2}} (\text{Binomial}[k + 2i, (k + 2i - h)/2 + z] - \text{Binomial}[k + 2i, (k + 2i + h)/2 + z])$

$*(\text{plu} + \text{pld})^{(k+2i+h)/2-z} * (\text{pru} + \text{prd})^{(k+2i-h)/2+z}$

Enter Number of Trials for the Simulation:

$\text{NumberOfTrials} = \text{num} = 9200;$

$\text{Do}[x[s, 0] = h, \{s, 1, \text{num}\}]; \text{Do}[y[s, 0] = k, \{s, 1, \text{num}\}];$

```
Do[x1[s,0] = h, {s,1,num}]; Do[y1[s,0] = k, {s,1,num}];
p[1] = pld; p[2] = prd; p[3] = plu; p[4] = pru;
t[0] = 0;
Do[t[i] = t[i - 1] + p[i], {i,1,4}]
```

Run Two Sets of Trials Simultaneously. Each Stops When Hit an Axis

```
Do[i = 1; While[(0 < x[s,i - 1])&&(0 < y[s,i - 1]), z = Random[];
x[s,i] = If[z ≤ t[1]||t[2] ≤ z < t[3], x[s,i - 1] - 1, x[s,i - 1] + 1];
y[s,i] = If[t[2] ≤ z, y[s,i - 1] + 1, y[s,i - 1] - 1];
r[s] = i; i = i + 1]; j = 1;
While[(0 < x1[s,j - 1])&&(0 < y1[s,j - 1]), z1 = Random[];
x1[s,j] = If[z1 ≤ t[1]||t[2] ≤ z1 < t[3], x1[s,j - 1] - 1, x1[s,j - 1] + 1];
y1[s,j] = If[t[2] ≤ z1, y1[s,j - 1] + 1, y1[s,j - 1] - 1];
r1[s] = j; j = j + 1], {s,1,num}]
```

Display Results of Simulations:

```
Do[b[s] = If[y[s,r[s]] == 0&&0 < x[s,r[s]], 1, 0], {s,1,num}];
sim = N[Sum[b[s], {s,1,num}]/num];

Do[b[s] = If[y1[s,r1[s]] == 0&&0 < x1[s,r1[s]], 1, 0], {s,1,num}];
sim1 = N[Sum[b[s], {s,1,num}]/num];

MatrixForm[{{" ", "Bottom"}, {"Conj", Conj}, {"Sim", sim}}]

MatrixForm[{{" ", "Bottom"}, {"Conj", Conj}, {"Sim1", sim1}}]
```

Hypothesis Test:

```
p0 = Conj; Z = NormalDistribution[0,1]; F[x_] = CDF[Z,x];
```

$$\text{TestStat} = \text{z0} = \frac{(\text{sim}-\text{p0})}{\sqrt{\frac{\text{p0}*(1-\text{p0})}{\text{num}}}}$$

P-Value for Two-Sided Alternative:

```
1 - Abs[F[z0] - F[-z0]]
```

Display Some Graphs:

```
For[s = 1, s ≤ 5, s++,
Print[ListLinePlot[Table[{x[s, j], y[s, j]}, {j, 0, r[s]}],
AxesOrigin → {0, 0}]]];
```

Appendix F

Testing the Conjecture in Example 6.3.2

(Applied to Example 6.3.2; change parameters to use for other examples)

Enter Parameters:

```
h = 6; k = 4;

pld = .35; prd = 0.28; plu = 0.32; pru = 0.05;
```

Use Trial and Error to Determine
when the distribution of Y is essentially complete:

$k * \sum_{i=0}^{150} \text{Binomial}[k + 2i, i]/(k + 2i) * (\text{prd} + \text{pld})^{k+i} * (\text{plu} + \text{pru})^i$

For a particular i, display kCi and all D(h, k, i, z):

$i = 5;$

$k * \text{Binomial}[k + 2i, i]/(k + 2i)$

$\text{Table}[\{h + (k + 2i - h)/2 + z - ((k + 2i + h)/2 - z),$

$\text{Binomial}[k + 2i, (k + 2i - h)/2 + z] - \text{Binomial}[k + 2i, (k + 2i + h)/2 + z]\},$

$\{z, 1, (h + k + 2i)/2\}]$

$\text{Conj} =$

$k*$

$\sum_{i=0}^{150} \text{Binomial}[k + 2i, i]/(k + 2i) * (\text{prd} + \text{pld})^{k+i} * (\text{plu} + \text{pru})^i$

$* \sum_{z=1}^{\frac{k+2i+h}{2}} (\text{Binomial}[k + 2i, (k + 2i - h)/2 + z] - \text{Binomial}[k + 2i, (k + 2i + h)/2 + z])$

$*(\text{plu} + \text{pld})^{(k+2i+h)/2-z} * (\text{pru} + \text{prd})^{(k+2i-h)/2+z}$

Enter Number of Trials for the Simulation:

```
NumberOfTrials = num = 10000;

Do[x[s, 0] = h, {s, 1, num}]; Do[y[s, 0] = k, {s, 1, num}];
```

```
Do[x1[s, 0] = h, {s, 1, num}]; Do[y1[s, 0] = k, {s, 1, num}];
p[1] = pld; p[2] = prd; p[3] = plu; p[4] = pru;
t[0] = 0;
Do[t[i] = t[i − 1] + p[i], {i, 1, 4}]
```

Run Two Sets of Trials Simultaneously. Each Stops When Hit an Axis

```
Do[i = 1; While[(0 < x[s, i − 1])&&(0 < y[s, i − 1]), z = Random[];
x[s, i] = If[z ≤ t[1]||t[2] ≤ z < t[3], x[s, i − 1] − 1, x[s, i − 1] + 1];
y[s, i] = If[t[2] ≤ z, y[s, i − 1] + 1, y[s, i − 1] − 1];
r[s] = i; i = i + 1]; j = 1;
While[(0 < x1[s, j − 1])&&(0 < y1[s, j − 1]), z1 = Random[];
x1[s, j] = If[z1 ≤ t[1]||t[2] ≤ z1 < t[3], x1[s, j − 1] − 1, x1[s, j − 1] + 1];
y1[s, j] = If[t[2] ≤ z1, y1[s, j − 1] + 1, y1[s, j − 1] − 1];
r1[s] = j; j = j + 1], {s, 1, num}]
```

Display Results of Simulations:

```
Do[b[s] = If[y[s, r[s]] == 0&&0 < x[s, r[s]], 1, 0], {s, 1, num}];
sim = N[Sum[b[s], {s, 1, num}]/num];

Do[b[s] = If[y1[s, r1[s]] == 0&&0 < x1[s, r1[s]], 1, 0], {s, 1, num}];
sim1 = N[Sum[b[s], {s, 1, num}]/num];

MatrixForm[{{" ", "Bottom"}, {"Conj", Conj}, {"Sim", sim}}]

MatrixForm[{{" ", "Bottom"}, {"Conj", Conj}, {"Sim1", sim1}}]
```

Hypothesis Test:

```
p0 = Conj; Z = NormalDistribution[0, 1]; F[x_] = CDF[Z, x];
```

$$\text{TestStat} = \text{z0} = \frac{(\text{sim}-\text{p0})}{\sqrt{\frac{\text{p0}*(1-\text{p0})}{\text{num}}}}$$

P-Value for Two-Sided Alternative:

```
1 – Abs[F[z0] – F[–z0]]
```

Display Some Graphs:

```
For[s = 1, s ≤ 5, s++,
Print[ListLinePlot[Table[{x[s, j], y[s, j]}, {j, 0, r[s]}],
AxesOrigin → {0, 0}]]];
```

Appendix G

Testing the Conjecture in Example 6.3.3

(Applied to Example 6.3.3; change parameters to use for other examples)

Enter Parameters:

$h = 18; k = 16;$

$\text{pld} = .45; \text{prd} = 0.30; \text{plu} = 0.15; \text{pru} = 0.10;$

Use Trial and Error to Determine
when the distribution of Y is essentially complete:

$k * \sum_{i=0}^{60} \text{Binomial}[k+2i, i]/(k+2i) * (\text{prd}+\text{pld})^{k+i} * (\text{plu}+\text{pru})^{i}$

For a particular i, display kCi and all D(h, k, i, z):

$i = 6;$

$k * \text{Binomial}[k+2i, i]/(k+2i)$

$\text{Table}[\{h + (k+2i-h)/2 + z - ((k+2i+h)/2 - z),$

$\text{Binomial}[k+2i, (k+2i-h)/2 + z] - \text{Binomial}[k+2i, (k+2i+h)/2 + z]\},$

$\{z, 1, (h+k+2i)/2\}]$

$\text{Conj} =$

$k*$

$\sum_{i=0}^{60} \text{Binomial}[k+2i, i]/(k+2i) * (\text{prd}+\text{pld})^{k+i} * (\text{plu}+\text{pru})^{i}$

$* \sum_{z=1}^{\frac{k+2i+h}{2}} (\text{Binomial}[k+2i, (k+2i-h)/2+z] - \text{Binomial}[k+2i, (k+2i+h)/2+z])$

$*(\text{plu}+\text{pld})^{(k+2i+h)/2-z} * (\text{pru}+\text{prd})^{(k+2i-h)/2+z}$

Enter Number of Trials for the Simulation:

$\text{NumberOfTrials} = \text{num} = 12000;$

$\text{Do}[x[s,0] = h, \{s, 1, \text{num}\}]; \text{Do}[y[s,0] = k, \{s, 1, \text{num}\}];$

$\mathrm{Do}[\mathrm{x1}[s,0]=h,\{s,1,\mathrm{num}\}];\mathrm{Do}[\mathrm{y1}[s,0]=k,\{s,1,\mathrm{num}\}];$

$p[1]=\mathrm{pld};p[2]=\mathrm{prd};p[3]=\mathrm{plu};p[4]=\mathrm{pru};$

$t[0]=0;$

$\mathrm{Do}[t[i]=t[i-1]+p[i],\{i,1,4\}]$

Run Two Sets of Trials Simultaneously. Each Stops When Hit an Axis

$\mathrm{Do}[i=1;\mathrm{While}[(0<x[s,i-1])\&\&(0<y[s,i-1]),z=\mathrm{Random}[];$

$x[s,i]=\mathrm{If}[z\le t[1]\|t[2]\le z<t[3],x[s,i-1]-1,x[s,i-1]+1];$

$y[s,i]=\mathrm{If}[t[2]\le z,y[s,i-1]+1,y[s,i-1]-1];$

$r[s]=i;i=i+1];j=1;$

$\mathrm{While}[(0<\mathrm{x1}[s,j-1])\&\&(0<\mathrm{y1}[s,j-1]),\mathrm{z1}=\mathrm{Random}[];$

$\mathrm{x1}[s,j]=\mathrm{If}[\mathrm{z1}\le t[1]\|t[2]\le \mathrm{z1}<t[3],\mathrm{x1}[s,j-1]-1,\mathrm{x1}[s,j-1]+1];$

$\mathrm{y1}[s,j]=\mathrm{If}[t[2]\le \mathrm{z1},\mathrm{y1}[s,j-1]+1,\mathrm{y1}[s,j-1]-1];$

$\mathrm{r1}[s]=j;j=j+1],\{s,1,\mathrm{num}\}]$

Display Results of Simulations:

$\mathrm{Do}[b[s]=\mathrm{If}[y[s,r[s]]==0\&\&0<x[s,r[s]],1,0],\{s,1,\mathrm{num}\}];$

$\mathrm{sim}=N[\mathrm{Sum}[b[s],\{s,1,\mathrm{num}\}]/\mathrm{num}];$

$\mathrm{Do}[b[s]=\mathrm{If}[\mathrm{y1}[s,\mathrm{r1}[s]]==0\&\&0<\mathrm{x1}[s,\mathrm{r1}[s]],1,0],\{s,1,\mathrm{num}\}];$

$\mathrm{sim1}=N[\mathrm{Sum}[b[s],\{s,1,\mathrm{num}\}]/\mathrm{num}];$

MatrixForm[{{“ ”, “Bottom”}, {“Conj”, Conj}, {“Sim”, sim}}]

MatrixForm[{{“ ”, “Bottom”}, {“Conj”, Conj}, {“Sim1”, sim1}}]

Hypothesis Test:

$\mathrm{p0}=\mathrm{Conj};Z=\mathrm{NormalDistribution}[0,1];F[\mathrm{x\_}]=\mathrm{CDF}[Z,x];$

$\mathrm{TestStat}=\mathrm{z0}=\frac{(\mathrm{sim}-\mathrm{p0})}{\sqrt{\frac{\mathrm{p0}*(1-\mathrm{p0})}{\mathrm{num}}}}$

P-Value for Two-Sided Alternative:

```
1 – Abs[F[z0] – F[–z0]]
```

Display Some Graphs:

```
For[s = 1, s ≤ 5, s++,
Print[ListLinePlot[Table[{x[s, j], y[s, j]}, {j, 0, r[s]}],
AxesOrigin → {0, 0}]]];
```

Bibliography

[1] D. Neal, “A Two-Sided Boundary Problem for Two-Dimensional Random Walks”, *International Electronic Journal of Pure and Applied Mathematics*, 6, No. 2 (2013), 63-75.

[2] D. Neal, “Generalized Boundary Problem For One-Dimensional Random Walks”, *Mathematica in Education*, 3, No. 4 (1994), 11-16.

[3] K. L. Chung, “Elementary Probability Theory with Stochastic Processes”. *Springer-Verlag*, New York, 1975.

[4] K. Pearson, “The Problem of the Random Walk”. *Nature*. 72. 294, 1905.

[5] L. Takacs, “On the Classical Ruin Problems”, *Journal of the American Statistical Association*, Vol. 64, No. 327 (1969), pp. 889 - 906.

[6] W. Feller, “An Introduction to Probability Theory and Its Applications”, Vol. I, 3rd Edition, *John Wiley and Sons*, 1968.