SIMPLE RANDOM WALKS: IMPROBABILITY OF PROFITABLE STOPPING

EMILY GENTLES

Abstract. This paper introduces the basics of the simple random walk with a flair for the statistical approach. Applications in biology and game theory will be discussed. The Impossibility of Profitable Stopping will be explored and proven.

Contents

1. Introduction

Imagine you are on an integer number line, standing at zero. You flip a fair coin and if it comes up heads you move one step right to +1 and if it comes up tails you move one step left to -1. You continue this process, moving one step right or left for each coin toss, until you have returned to your starting position or have taken a predetermined number of steps. You have just executed a symmetric simple random walk. This simple random walk spans many fields of study and was first noticed by the Romans in 60BC. However, one of the subject's most notable beginnings occurred in biology when Scottish botanist Robert Brown rediscovered the phenomena in 1827. While studying pollen suspended in water Brown noticed the jittery motion of the pollen's organelles. At the time Brown speculated that the organelles were tiny organisms, alive and moving. Although his hypothesis was later disproved, the movement was coined Brownian motion. Since its start the subject has been fundamental in economics and physics with less significant uses in various other courses of study. In this paper we will explore the basic attributes of simple random walks in Section 2, how they relate to Brownian motion and the various distributions they can be described with in Section 3, and we will touch on their role in game theory in Section 4.

Date: DEADLINES: Draft AUGUST 15 and Final version AUGUST 29, 2016.

2. What it looks like

Definition 2.1. The *configuration space of dimension* N is the collection of binary sequences of length N for a fixed $N \in \mathbb{N}$. We denote this as

$$\Omega_N = \{\omega = (\omega_1, \ldots, \omega_N) \in \{+1, -1\}^N\}.$$

Note that Ω can be thought of as the set of all possible outcomes of a simple random walk.

Definition 2.2. The *step* of the random walk at time k is the direction the walk is moving in at time k, denoted

$$X_k(w) = w_k$$

for $1 \leq k \leq N$ and $w \in \Omega_N$.

Definition 2.3. The *position* of the random walk after n steps is the sum of all n steps:

$$S_n(w) = \sum_{k=1}^{n} X_k(w)$$

for $1 \leq n \leq N$ and $S_0(w) = 0$

For every $w \in \Omega_N$ we obtain a *path* (i.e. a trajectory). That is to say, for each w there is a valid collection of steps from the starting position to another chosen point.

Example 2.1 If $N = 2$ then $\Omega = \{(1,1), (1,-1), (-1,1), (-1,-1)\}$ where the position 0 can be achieved by two different paths. Thus $X_2(w) = -1$ or 1 and $S_2(w) = -2$ or 0 or 2.

We can take the uniform distribution $P_N(A) = |A|2^{-N}$ for $A \subseteq \Omega_N$ as the *probability distribution* on Ω_N. Hence all binary sequences w (all paths) have the same probability.

Definition 2.4. The sequence of random variables $(S_n)_{n=0}^N$ on the finite probability space (Ω_N, P_N) is called a *simple random walk* of length N, starting at 0.

Notice, also, that $P(X_k = x_{k_1}, \ldots, X_{k_n} = x_{k_n}) = 2^{N-n}2^{-N} = 2^{-n}$ as $1 \leq k_1 < \ldots < k_N \leq N$ and $x_{k_i} \in \{+1, -1\}$ for $i = 1, \ldots, n$. For a fair coin the probability of getting heads is the same as getting tails:2^{-1}. The probability of any given step is the product of the probability of all previous steps along with the probability of the particular step. Thus landing on 3 after 3 steps has probability $P(X_3) = 2^{-1}2^{-1}2^{-1} = 2^{-3}$.

We will now quickly introduce the idea and implementation of expected value. Finding the expected value of a function or process means that you are calculating the most likely outcome. Random walks have a binomial distribution (Section 3) and the expected value of such a distribution is simply $E(x) = np$ where n is the total number of trials, steps in our case, and p is the probability of success, a right step in our case. Thus, if we are going to take 10 steps and the probability of taking a right step is .3 then $E(x) = 10 * .3 = 3$; we would expect to take 3 steps to the right out of 10 total steps. In Theorem 2.5 we follow our previously established notation and denote right steps as 1 and left steps as -1.

Let X denote the position of a simple random walk after k steps. L be the total number of steps.

Theorem 2.5.

(1) $E(X_k) = 0$
(2) $E({X_k}^2) = 1$
(3) $E(S_N) = 0$
(4) $E({S_N}^2) = N$

Proof.

(1) $E(X_k) = \frac{1}{2}(-1)\frac{1}{2}(1) = 0$
(2) $E({X_k}^2) = \frac{1}{2}(-1)^2\frac{1}{2}(1)^2 = 1$
(3) $E(S_N) = \sum_{k=1}^{N} E(X_k) = \sum_{k=1}^{N} 0 = 0$
(4) $E({S_N}^2) = \sum_{k=1}^{N} E({X_k}^2) = \sum_{k=1}^{N} 1 = N$

□

Theorem 2.6. *For* $x \in \{-n, -n+2, \ldots, n-2, n\}$ *the probability that a simple random walk is in* x *after* n *steps is* $P(S_n = x) = \binom{n}{\frac{n+x}{2}}2^{-n}$.

Proof. $S_n = x$ if and only if $k = \frac{n+x}{2}$ is the number of steps to the right. Therefore $S_n(w) = k(1) + (n-k)(-1) = 2k - n = x$. Thus $P(S_n = x) = \binom{n}{k}2^{-n}$ (where 2^{-n} is derived from the fact that each step happens with probability 2^{-1}). □

3. Brownian Motion

Brownian Motion was the real start to the study of random walks. However, Brownian motion covers many more cases than just that of the simple symmetric walk. Brownian motion is inspired by the movement of particles living in a three-dimensional space as opposed to the two directions we limited the simple random walk to. Thus the mathematical representation of the various scenarios quickly becomes more complex. For now we will focus our efforts on the simplest form of Brownian Motion.

By definition standard Brownian motion is a random process
$X = \{x_t : t \in [0, \infty)\}$ that operates in $\mathbb{R}$ and has the following conditions:

(1) The starting position is always zero so $X_0 = 0$ with probability 1.
(2) X has *stationary increments* thus $X_t - X_s$ has the same distribution as X_{t-s}. This means that the distribution is a function of 'how long', t, and not of 'when', s. The underlying dynamics don't change over time whether they be jostling particles or coin flips.
(3) X has *independent increments* meaning that the random variables $X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent for $t_1, t_2, \ldots, t_n \in [0, \infty)$ and $t_1 < \ldots < t_n$. In other words, each increment does not depend on the previous or following increments.
(4) X_t is normally distributed with a mean of 0 and a variance of t for $t \in (0, \infty)$.
(5) $t \mapsto X_t$ is always continuous on $[0, \infty)$.

Let's use these requirements to consider the journey of a random walker. What is the probability, $P(m, N)$, that the walker will be at position m after N steps? (We are using $P(m, N)$ here as a general probability statement.) Let p be the probability of a right step and $1 - p = q$ be the probability of a left step. There are many ways to obtain such a path when $m < N$ but each path is independent thus we can simply add up their probabilities. We know from (1) that the walker must start from 0 and we also know that the walker must make $n_1 = m + n_2$ steps to the right and n_2 steps to the left as $n_1 + n_2 = N$. It follows that $n_1 = \frac{1}{2}(N + m)$ and $n_2 = \frac{1}{2}(N - m)$. Now, since each path is independent (a conclusion that can be attributed to (3)), we know that there must be n_1 factors of p and n_2 factors of q. Using this we get

$$p^{n_1} q^{n_2} = p^{\frac{1}{2}(N+m)} q^{\frac{1}{2}(N-m)}$$

which is simply the probability of any one path. We must multiply that by the total number of ways a path of this sort can occur which is

$$\frac{N!}{n_1! n_2!} = \frac{N!}{n_1!(N - n_1)!}$$

Thus we ultimately obtain

$$P(m, N) = \frac{N!}{(\frac{N+m}{2})!(\frac{N-m}{2})!} p^{\frac{1}{2}(N+m)} q^{\frac{1}{2}(N-m)}$$

This probability can be explored through multiple distributions.

First we'll look at the binomial distribution. This distribution is of the form

$$b(n, p, k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

This is used to calculate the number of successes in n fixed trials where p =probability of success and k =number of successes. Viewing our random walk within this distribution we use the fact that $n = \frac{1}{2}(N + m)$ to get

$$P(m, N) = \frac{N!}{n!(N - n)!} p^n q^{N-n} = \binom{N}{n} p^n q^{N-n}$$

Thus our random walk can clearly be expressed by a binomial distribution where we define success as a rightward step.

Next let's look at the Poisson distribution. This distribution has the form

$$P(m, N) = \frac{e^{-\lambda} \lambda^x}{x!}$$

and calculates the number of arrivals in a fixed period of time. The frequency or success per unit of time is denoted by λ while x is the number of successes in a given unit of time. In general, $\lambda = np$ for large n and small p. Thus as N goes to infinity, p goes to 0 and λ remains constant we can see that

$$Pois_N(n) = \frac{N!}{n!(N - n)!} p^n q^{N-n} = \frac{N(N-1)\dots(N-n+1)}{n!} p^n (1 - p)^{N-n}$$

$$= (1 - \frac{1}{N}) \dots (1 - \frac{n-1}{N}) \frac{(Np)^n}{n!} (1 - \frac{Np}{N})^{N-n}$$

This means that $n \approx Np$ and $n << N$. From this we can see that the limit as N goes to infinity of $(\frac{1-\lambda}{N})^N$ is $e^{-\lambda}$. Therefore, the limit as N goes to infinity of Np

is λ and $Pois(n) = \frac{\lambda^n e^{-\lambda}}{n!}$. Hence our random walk can be easily expressed by the Poisson distribution.

4. Game Theory

Now we will look at what part random walks play in game theory. Let's first take note of a few things.
Observations:
1. The distribution of S_n is symmetric around 0.

$$P(S_n = x) = \frac{n!}{\left(\frac{n-x}{2}\right)!\left(\frac{n+x}{2}\right)!} = P(S_n = -x) \tag{4.1}$$

In other words the probability of a right step is equal to the probability of a left step.
2. The maximal weight (mode) is achieved in the middle.

$$P(S_{2n} = 0) = P(S_{2n-1} = 1) = \binom{2n}{2} 2^{-2n} \tag{4.2}$$

3. Sterling's formula, an approximation for factorials, states that $n! \sim n^n e^{-n}\sqrt{2\pi n}$ as $n \to \infty$. From this we can conclude

$$P(S_n = 0) \sim \frac{1}{\sqrt{\pi n}}, n \to \infty \tag{4.3}$$

This means that $P(S_{2n} = 0) * \sqrt{\pi n} = 0$ so $P(S_{2n} = 0)$ must go to 0 and $\sqrt{\pi n}$ must go to infinity.
Furthermore for a finite interval [a,b],

$$\lim_{n\to\infty} P(a \le S_n \le b) = 0. \tag{4.4}$$

This simply means that as n grows very large the probability of the path ending within some finite range is 0.

We shall now view the simple random walk as a game. Suppose that in round k a player wins an amount X_k and S_n is the player's capital after n rounds. Before each round the player chooses to either continue in the game or stop where they are. We saw in Section 2 that $E(S_n) = 0$ so the expected gain after each round is 0. Considering this, is it possible to stop the game in a favorable moment such that it leads to positive expected gain? Let's find out.

Definition 4.5. An event $A \subseteq \Omega$ is *observable until time* n when it can be written as the union of basic events of the form $\{w \in \Omega : w_1 = o_1, \ldots, w_n = o_n\}$ with $o_1, \ldots, o_n \in \{+1, -1\}$.

This really just means that a player cannot see into the future. After round n the player does not know the outcome of round $n+1$. $\mathcal{A}_n$ denotes this class of events A that are observable until time n.
The *indicator function* is a random variable $\mathcal{I}_A$ for an event $A \subseteq \Omega$ where $\mathcal{I}_A(w) = 1$ when $w \in A$ and $\mathcal{I}_A(w) = 0$ when $w \notin A$. This function is useful when determining whether or not a round is within the time frame that the player can observe. Note that $\{\emptyset, \Omega\} = \mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \ldots \subseteq \mathcal{A}_n = \{\text{set of all subsets of } \Omega\}$. This makes sense because if the player knows what happened in round n then the player also knows

what happened in round $n-1$.

Lemma: For $n = 0, 1, \ldots, N-1$ and $A_n \in \mathcal{A}_n$, $P(A_n \bigcap \{X_{n+1} = +1\}) = \frac{1}{2}P(A_n)$ and $E(X_{n+1}\mathcal{I}_{A_n}) = 0$.

This expected value simply means that the player could have a positive capital and the expected value of the next round would be 0, indicating that the player still has a chance to end the game at a favorable moment.

Theorem 4.6. ***Impossibility of Profitable Stopping***: *For any stopping time* $T : \Omega_N \to \{0, \ldots, N\}, E(S_T) = 0$ *where* $S_T = S_{T(w)}(w)$ *is the outcome of the path* w *at stopping time* $T(w)$.

Proof. $\{T \geq k\}^C = \bigcup_{l=0}^{k-1}\{T = l\} \subseteq \mathcal{A}_{k-1}$ for $k = 0, \ldots, N$. This means that $\{w \in \Omega_N | T(w) = l\} \subseteq \mathcal{A}_l \subseteq \mathcal{A}_{k-1}$. Since $S_T = \sum_{k=1}^{N} X_k \mathcal{I}_{\{T \geq k\}}$ it follows that $E(S_T) = \sum_{k=1}^{N} E(X_k \mathcal{I}_{\{T \geq k\}}) = 0$. □

Therefore, it is impossible to strategically choose a stopping time such that the capital will be positive.

Acknowledgments. It is a pleasure to thank my mentor, MurphyKate Montee, for excellence guidance, insight and enthusiasm as well as the University of Chicago for the opportunity to grow in my mathematical journey. I would also like to thank the University of Arkansas Mathematics Department for the additional support and encouragement.

References

[1] Luca Avena, Markus Heydenreich, Frank den Hollander, Evgeny Verbitskiy, Willem van Zuijlen. Random Wlaks (Lecture Notes). http://websites.math.leidenuniv.nl/probability/lecturenotes/RandomWalks.pdf.
[2] Random Walks, Chapter 2. http://physics.gu.se/ frtbm/joomla/media/mydocs/LennartSjogren/kap2.pdf.
[3] Standard Brownian Motion. http://www.math.uah.edu/stat/brown/Standard.html.