Proof of Random Walk Theorem

Phillip Compeau

The Random Walk Theorem states that the average distance that a randomly walking particle will find itself from its starting point after taking n steps of unit length is $\sqrt{n}$. Below, we provide a justification for why this is true for interested readers who are familiar with probability.

Let $\mathbf{x_i}$ denote the (random) vector corresponding to the particle's i-th step. The particle's position $\mathbf{x}$ after n steps is the sum of the $\mathbf{x_i}$,

$$\mathbf{x} = \mathbf{x_1} + \mathbf{x_2} + \cdots + \mathbf{x_n} \,.$$

The distance d traveled by the particle is the distance from $\mathbf{x}$ to the origin, which is the square root of the inner product $\langle \mathbf{x}, \mathbf{x} \rangle$. We will show that the *expected value* of d^2 is equal to n. First, note that

$$d^2 = \langle \mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{x_1} + \mathbf{x_2} + \cdots + \mathbf{x_n}, \mathbf{x_1} + \mathbf{x_2} + \cdots + \mathbf{x_n} \rangle \,.$$

We can apply the linearity of the inner product to expand and obtain

$$d^2 = \langle \mathbf{x_1}, \mathbf{x_1} + \mathbf{x_2} + \cdots + \mathbf{x_n} \rangle + \langle \mathbf{x_n}, \mathbf{x_1} + \mathbf{x_2} + \cdots + \mathbf{x_n} \rangle + \cdots + \langle \mathbf{x_2}, \mathbf{x_1} + \mathbf{x_2} + \cdots + \mathbf{x_n} \rangle \,.$$

If we apply the linearity of the inner product again, then we will expand these n inner products into n^2 inner products of the form $\langle \mathbf{x_i}, \mathbf{x_j} \rangle$,

$$d^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \langle \mathbf{x_i}, \mathbf{x_j} \rangle \,.$$

We will now apply a fundamental result in probability called the "linearity of expectation", which states that for any two random variables x and y, the expectation of their sum $\mathbb{E}(x+y)$ is equal to the sum of the corresponding expectations $\mathbb{E}(x) + \mathbb{E}(y)$. When we take the expected value of both sides of the equation above and apply the linearity of expectation, we obtain

$$\mathbb{E}(d^2) = \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbb{E}\left(\langle \mathbf{x_i}, \mathbf{x_j} \rangle\right) .$$

For any i, $\langle \mathbf{x_i}, \mathbf{x_i} \rangle$ is just the length of the vector $\mathbf{x_i}$, which is equal to 1. On the other hand, the expected value of the inner product of two random unit vectors is equal to 0, so when $i \neq j$, $\mathbb{E}\left(\langle \mathbf{x_i}, \mathbf{x_j} \rangle\right)$ is equal to 0. Therefore, the right side of the above equation consists of n terms that are equal to 1 and $n^2 - n$ terms that are equal to 0, and so $\mathbb{E}(d^2) = n$, which is what we set out to prove.

We make a couple of notes about the above proof. First, we did not use anything about the random walk being two-dimensional in this proof; therefore, it holds whether our particle is walking in two, three, or any number of dimensions.

Second, we technically did not show that the expected value of d is $\sqrt{n}$, but rather that the expected value of d^2 is n. It is not true that $\mathbb{E}(d)$ is equal to $\sqrt{n}$, but rather that as n grows, $\mathbb{E}(d)$ grows like $c \cdot \sqrt{n}$ for some constant factor c. A proof of this fact is beyond the scope of this work, but it can be shown that as n tends toward infinity, $\mathbb{E}(d)$ tends toward $\sqrt{(2/\pi)} \cdot \sqrt{n}$.

Who knew that the mathematics of random walks could be so complicated!