

# **Lecture Notes on Stochastic and Spatial Dynamics in Biology (EC 303)**

A course taken by Dr. Vishweshha Guttal, CES, IISc

Ananda Shikhara Bhat

# Lecture 01 (L): Introduction to random walks

August 6, 2021

*Lecturer: Vishwesh Guttal**Written by: Ananda Shikhara Bhat*

Individuals often move around for various important reasons such as foraging or finding a mate. A tool that is often used to model such movement is the theory of *random walks*. In this lecture, we look at a relatively non-technical introduction to unbiased random walks.

## 1 Displacement and MSD

### 1.1 Displacement in a simple random walk

Consider an organism on a discrete one-dimensional space. Let  $X_t$  be the random variable representing the position of the organism at time  $t$ . Throughout this discussion, we assume that time is discrete. Consider a stochastic process that assigns the following movement rule to the animal:

$$X_{t+1} = \begin{cases} X_t + a & \text{with probability } \frac{1}{2} \\ X_t - a & \text{with probability } \frac{1}{2} \end{cases} \quad (1)$$

The organism can be thought of as executing as a 'random walk' over the one-dimensional space. Figure 1 shows what a trajectory produced by one such process may look like in two dimensions. In this figure, the blue square indicates the initial position of the organism, at the origin, and the red square indicates the position after 10,000 time steps.

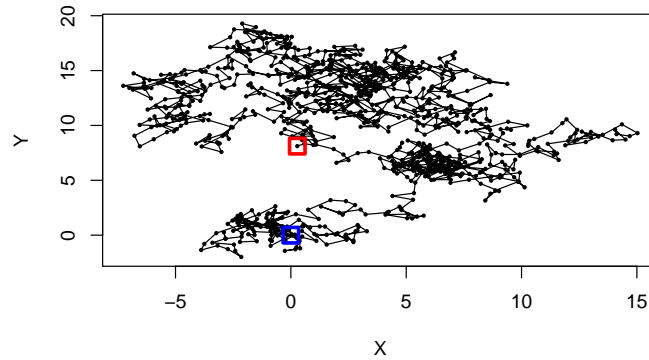


Figure 1: A sample trajectory of an organism moving according to Equation (1).

More generally, a *one-dimensional random walk* is a stochastic process of the form:

$$X_{t+1} = X_t + a \quad (2)$$

where  $a$  is a random variable with some probability distribution. In Equation (1),  $a$  is a bernoulli(1/2) random variable.

If we were to observe an organism starting at  $X_0$  and moving along one dimension according to (1) for  $t$  timesteps, where may we expect it to end up? In other words, what is the expectation value of the *displacement*  $D_t := X_t - X_0$ ? The calculation for the displacement is straightforward if one notices that  $X_t$  can be written in terms of  $X_{t-1}$  as follows:

$$\begin{aligned} \langle X_t \rangle &= \frac{1}{2}(\langle X_{t-1} \rangle + a) + \frac{1}{2}(\langle X_{t-1} \rangle - a) \\ &= \langle X_{t-1} \rangle \end{aligned}$$

Thus, applying this recursively  $t$  times, we obtain:

$$\begin{aligned} \langle X_t \rangle &= X_0 \\ \implies \langle D_t \rangle &= 0 \end{aligned}$$

In other words, the expectation value is 0, which sounds like the organism hasn't moved at all! It is easy to see that this argument is independent of the dimensionality of the walk. However, a cursory glance at 1 tells us that the organism has moved away from the origin after all. Where, then, have we gone wrong?

The issue is that since the displacement can be either positive or negative, all our calculation tells us is that if we had many such systems, the displacement would be zero *on average* (Note: such a walk is often called an 'unbiased' random walk). However, one could get  $\langle D_t \rangle = 0$  even if the organism went very far off in the positive  $x$  direction in half of the times we tried the process, and very far off in the negative  $x$  direction in the other half of the times. The expectation value thus fails to capture the notion of 'spread away from the starting point'. For this, we require a slightly different quantity.

## 1.2 The mean square displacement

Intuitively, the expectation value can be zero because realizations with a high positive distance covered 'cancel out' with those which had a high negative distance covered, since both are equally likely to occur. To account for this, we will instead define the 'mean square displacement' (MSD) of a random walk as:

$$MSD(t) := \langle (X_t - X_0)^2 \rangle$$

We can slightly simplify the term on the right using the linearity of the expectation value as follows:

$$MSD(t) = \langle X_t^2 \rangle - 2X_0\langle X_t \rangle + X_0^2 \quad (3)$$

Using the fact that  $\langle X_t \rangle = X_0$  for an unbiased random walk, we then see that:

$$MSD(t) = \langle X_t^2 \rangle - X_0^2 \quad (4)$$

Equation (4) is very general, and applies to any unbiased one-dimensional random walk whatsoever. For the random walk given in Equation (1), we have:

$$X_t^2 = \begin{cases} X_{t-1}^2 + a^2 + 2aX_{t-1} & \text{with probability } \frac{1}{2} \\ X_{t-1}^2 + a^2 - 2aX_{t-1} & \text{with probability } \frac{1}{2} \end{cases}$$

and thus

$$\langle X_t^2 \rangle = \langle X_{t-1}^2 + a^2 \rangle$$

This is a recursive equation. One can write  $X_{t-1}^2$  in terms of  $X_{t-2}^2$ ,  $X_{t-2}^2$  in terms of  $X_{t-3}^2$ , and so on. After doing this  $t$  times, we arrive at:

$$\langle X_t^2 \rangle = X_0^2 + ta^2$$

Substituting this into Equation (4), we arrive at:

$$MSD(t) = ta^2 \quad (5)$$

This is telling us something profound. It tells us that as the time increases, we expect the MSD to linearly increase, meaning that in any given realization, an organism executing such a walk is likely to be found further away from the origin as time increases. Note that this is despite the fact that the organism is *equally likely* to go in either the positive or the negative direction at any given instance(!).

### 1.3 A general expression for the MSD (by students)

Consider a random walk of the form (2). Let the random variable  $a$  have mean  $\mu$  and variance  $\sigma^2$ . We will now obtain an expression for the MSD of such a process. Since we are not assuming that the walk is unbiased, we must use equation (3) for the definition of the MSD.

Since  $\langle X_t \rangle = \langle X_{t-1} \rangle + \langle a \rangle$ , we can apply this relation recursively to obtain:

$$\langle X_t \rangle = X_0 + t\mu \quad (6)$$

We can also calculate  $\langle X_t^2 \rangle$  in a similar manner:

$$\begin{aligned} \langle X_t^2 \rangle &= \langle X_{t-1}^2 \rangle + \langle a^2 \rangle + 2\langle aX_{t-1} \rangle \\ &= \langle X_{t-1}^2 \rangle + (\sigma^2 + \mu^2) + 2[\text{cov}(a, X_{t-1}) + \mu\langle X_{t-1} \rangle] \end{aligned}$$

Substituting equation (6), we obtain:

$$\langle X_t^2 \rangle = \langle X_{t-1}^2 \rangle + (\sigma^2 + \mu^2) + 2\text{cov}(a, X_{t-1}) + \mu(X_0 + t\mu)$$

This, too, can be applied recursively to obtain:

$$\langle X_t^2 \rangle = X_0^2 + t(\sigma^2 + \mu^2) + 2t\mu(X_0 + (t-1)\mu) + 2 \sum_{k=1}^{t-1} \text{cov}(a, X_k) \quad (7)$$

Finally, we can plug (6) and (7) into (3) to obtain our final expression:

$$MSD(t) = t [\sigma^2 + \mu^2(2t-1)] + 2 \sum_{k=1}^{t-1} \text{cov}(a, X_k) \quad (8)$$

## 1.4 Dimensionality considerations (by students)

So far, we have only been considering a one-dimensional walk. What happens when an organism moves in multiple dimensions? Assuming that the choice of direction made in each dimension is independent, the MSD simply adds linearly. To see this, note that for a  $D$  dimensional random walk, the MSD is given by:

$$MSD_D(t) = \sum_{d=1}^D \langle (X_d(t) - X_d(0))^2 \rangle$$

where  $X_d(t)$  denotes the random variable for the value of the  $d^{\text{th}}$  dimension of the position at time  $t$ . However, each term in the summation is simply the MSD of a one-dimensional walk along that dimension. Thus, we see that:

$$MSD_D(t) = \sum_{d=1}^D MSD_d(t)$$

where  $MSD_d(t)$  is the MSD for a one-dimensional random walk along the  $d^{\text{th}}$  dimension at time  $t$ .