Evolutionary Dynamics

Supreet Saini

Chemical Engineering

Indian Institute of Technology Bombay

Week 07

Lecture 34

Hi, welcome back, everyone. Let us continue our discussion of studying these transition probabilities and from there get to the fixation probability of a new mutation. So, we will start with the number line representation that we have been discussing, and we will define one variable that will help us calculate these fixation probabilities. So, our system could be in any one of these states from 0 to n, and there are a total of N minus 1 possibilities for our system to exist in. So, the system could exist in any one of these N minus one states given by zero to N. And just as a reminder, let's say this is the state of the system.

And what I'm interested in is the state of the system, which is the number of B individuals in the population. What I am interested in is the following question: What is the probability that a new mutation reaches fixation? This is the problem that I am interested in. And by new mutation, I mean that this mutation has just occurred.

That means the state of the system is here. This B-type individual is the new mutation that has occurred. And there is only one such individual carrying that mutation. So, the state of the system is i equal to 1. And what I am interested in is

Understanding that, what is the probability that the system reaches i equal to N? So, I can also reframe this question as: what is the probability that the system goes from i equal to 1 to i equal to N? These two are equivalent questions to ask. Now, what we also realize is that the transition of the system from i equal to 1 to i equal to N cannot happen in one step. Every Moran step allows a jump of only 1 or 0. So, if the system is at i, you can go to i plus 1, you can go to i minus 1, or you will remain at i. No other moves are permissible.

So, it has to be a step-by-step procedure through which the system moves from i equal to 1 to i equal to N. I'm going to define—let's for this—and this is the variable that is going to be extremely important in the next few videos. It's: let xi be the probability that the system reaches N when starting from i. So, which means that x1 is the probability that the system starts with i equal to 1. And reaches i equal to N, which is the same as this. So, I hope this is crystal clear to everyone that these three things are exactly the same.

The first way of posing this is: what is the probability that a new mutation reaches fixation, which means a new mutation has just entered the population. I am at state i equal to 1, and I am interested in the probability that the system will transition to i equal to N. That is the first way of saying this. The second way of saying this is: what is the probability that the system goes from i equal to 1 to i equal to N? Here, I was saying the same thing by stating that this is a new mutation. New mutation meaning i equal to 1.

I am saying the same thing here a bit more mathematically: what is the probability that the system transitions from i equal to 1 to i equal to N? So these two are equivalent. Now I am defining this transition probability in terms of a variable because eventually I want to find an expression for this probability, and that is going to be given by x1, where xi is the probability that the system starting from any i is able to reach N. So this transition probability is xi: the probability that the system goes from 1 to N is x1, the probability that the system goes from 2 to N is x2, and so on and so forth. Depending on the value of i, I am defining the starting point of the system as i, and I am asking what is the probability that the system will move from this starting point and reach N. So, in the context of its equivalence with these two questions, when I say I am defining a variable called x1, I am asking that the system starts at i equal to 1 and reaches i equal to N. So these probabilities are all the same.

x1, the probability that mutation reaches fixation, and the probability that the system goes from i equal to 1 to N. I hope that's clear to everybody. Okay. So, let me see. Right. So with this, now,

Let's go back to the number line and let me ask a few questions. i minus 1, i, i plus 1, N minus 1. Remember, xi is what is the probability that the system starts at i and reaches N. Which means when I'm talking about x1, the question I'm asking is: what is the probability that the system starts here and ends up here?

What is the fixation probability of a new mutation that I'm interested in? This is the eventual goal—to get to this part. So, with this definition, I want you to take 15 seconds and think about the question: What is x0? What is the value of x0?

I won't say anything and give you 15 seconds to think about this question. So, what you should realize is that X naught means I am asking: What is the probability that the system starts at 0 B-type individuals and reaches N? Remember that we are studying the system where no new mutations are allowed. So, if the system is starting here, there is no mutation that can occur in the system. And hence, if there is no B-type individual in the population to begin with, the chances that you will have N B-type individuals is just equal to 0.

I hope that makes sense because this number line represents how many B-type individuals are there in the population. And since no mutations are permissible in the model that we are making, we have to have at least one B-type individual for there to be any chance that B-type individuals can reach fixation. In X0, we are saying that there is no B-type individual. What is the chance that B-type individuals will take over the population? Obviously zero, because there is no B-type individual to begin with in the population.

We will take another 15 second break and I want you to tell me that what is xN. Let us pause for another 15 seconds. What you should realize that this question is asking is that what is the probability that system starts at N and reaches N. So the starting point of the system is this point itself. The system is at N, which means all individuals that are present in the population are of B type only. There is no A type individual.

So the system, if it's already at N, it's going to remain at N because it can't move from there. There is no A type individual in the population. Hence, the probability that this system starts at N and reaches N is simply equal to 1. I hope these two transition probabilities are clear to you. Another way to look at this is that what is the chance is coming back from here to here even possible for this system.

Obviously remember Moran process says that from I you can either remain at I go to i plus 1 or go to i minus 1 that is true of internal nodes on this number line but at this terminal node which is the last one you can't go to N plus 1 because we are not allowed We are not allowing the population to exceed N. Population must always remain N. So that's not allowed. It can remain at N and via the probability PNN which is going to be sum of two which is probability of B born and B dying which can happen because all N individuals are of B type plus probability of a born probability of a born and a dying that cannot happen because since there is no a type individual a cannot die and since there is no a type individual a cannot be born because a can only be born from a type individuals that are already existing in the population at this point there is no a type individual in the population hence this transition probability is zero hence this is the only way that the system can remain at N

On the other hand N to N-1 would be allowed for that to happen I have to calculate the probability of B dying, probability of B death which can happen and probability of A being born. Probability of B death is some finite number. Using the expression that we have been calculating, we can find that out. But probability of A being born is equal to 0 because A will be born from an existing A individual only. But since there is no A existing individual in the population,

another a individual cannot be born hence this probability is 0 and since these two are multiplying transition probability from N to N minus 1 is also equal to 0 n to n plus 1 is not allowed in that the rules hence the system will always stay at n always stay at the state N under the rules that we are operating at and hence that the system is at xN and it will continue to remain at N is simply equal to 1. So x0 and xN are zero and one. So we know these extreme transition probabilities for these two endpoints. However, what we are interested in is x1 when the mutation first makes an appearance. To understand that, we will do a slightly tricky mathematical representation

i minus 1, i, i plus 1 going all the way to n. So, remember that xi represents the probability that the system starts from i and reaches xN, which represents the transition from this i if i am actually talking about i to n this transition the probability of this transition happening is represented by xi the probability of this transition happening is represented by xi plus 1 and the probability of this transition happening is x i minus 1 and so on and so forth i can the probability of this transition happening is x1 this one happening is x2 this one happening is x0 which is 0 and so on and so forth so imagine that my system is at state i so what is the probability that I will get here. That is xi.

I know that. By definition, it is just xi. But what happens when I start a Moran process from this state of i? When I start a Moran process from here, one of three things can happen. Either the system will go from i to i plus 1, which happens with a probability p i to i plus 1. But once this happens, the system is here. And what is the probability that from here I reach N? That is given by x into x i plus 1. So, let me write this in words.

This is the probability that from i, the system reaches N. But this route to reaching N has to begin somewhere, and it has to begin with one of these three steps. Probability of i to i minus 1, or it remains here, which is probability of i to i, which we saw is the sum of two probabilities. So while this is a straightforward expression, xi starts from here. After what happens after one Moran process?

After one Moran process, with probability pi to i minus 1, the system finds itself at state i minus 1. With this probability, the system finds itself here. And if it finds itself here, what is the probability that it will reach N? That is simply x of i minus 1. That is one of the three possibilities that can happen when we start this random walk from I. Alternatively, the system can start from I and after one step remain at I, which happens with probability PII.

But then if the system after one step is at I, what is the probability that from here it will reach N? That is simply xi—that's the definition of xi. So that's just xi. But the system could also take this step with probability pi i plus one, i to i plus one, and then what is the probability that once it has reached here, it will reach N? That is simply xi plus one.

So I get this relationship that xi is dependent on the three transition probabilities and their respective xi's. So you will see that xi is a function of all of this but also a function of xi itself. And depending on the value of I that I choose, I can write this equation for every value of I. So let's write a few of these equations. If I write this for x1, the equation becomes x1—that means this quantity, which is the quantity we have been saying repeatedly is the quantity that I am most interested in—is simply equal to the transition probability of P1 going to 0 because i is 1, so i minus 1 is 0, times x0.

$$x_0 = 0$$

$$x_N = 1$$

$$x_i = P(i \to i+1)x_{i+1} + P(i \to i)x_i + P(i \to i-1)x_{i-1}$$

Plus transition probability of 1 to 1 times x1 plus transition probability of 1. Let me write that again. Transition probability of going from 1 to 2, i to i plus 1 times x2. Now, we discussed in the previous slide that x0 is just 0. So, I am going to plug that in.

And this term simply vanishes. So I get x1 as p11 x1 plus p12 x2. Now I can take the x1 to the other side because remember x1 is the quantity that I am interested in. So I get 1 minus p11 x1. x1 is equal to p12 x2, which tells me x1 is simply equal to p12 x2 divided by 1 minus p11.

Now, there is a problem here. I got an expression of x1, but that expression is actually dependent on x2. So, unless I know the value of x2, I cannot know the value of x1, and hence I need to find x2. So, what we will do is we will write this equation for x2 because remember by plugging a value of i, I can find out. By plugging in a specific value of i, I can find out this equation for a specific value of i. So it turns out that since the quantity I am interested in is x1, but to find x1, I need x2.

So how do I find x2? Let us write this equation first: xi is equal to transition probability i to i minus 1 times xi minus 1, plus transition probability i to i times xi, plus transition probability i to i plus 1 times xi plus 1. And we found out from the first equation that x1 is simply equal to—let me just take a look—p12x2, transition probability p12x2 divided by 1 minus p11. And I know everything. I know the transition probabilities because I know the respective growth rates.

I know the numbers in the population. The only thing missing for me to find out x1, which is the variable of my interest, is knowing the value of x2. So for this reason, I will write the equation for x2. So let's see what that gives us. That tells me that x2—now I'm going to put in this equation x2.

I is equal to 2, and that tells me x2 is equal to probability of 2 to 1 times x1, which is fine, plus transition probability of 2 to 2 times x2, which is fine. I can just take this to the other side, plus transition probability of 2 to 3 times x3. I can—I have an expression of x2, x1. I'm sorry, let me say that again. I have an expression for x1, which is a function of x2.

So instead of this x1, I can substitute this expression. So this equation here has x2 here. This will also have an x2 here. And this already has an x2 there. So these three terms have x2, but I can't solve for x2 because it turns out x2 is actually now dependent on x3.

So unless I know x3, I can't find x2. And unless I know x2, I cannot find x1 because x1 can only be found by knowing x2. And x2 can only be found by knowing x3. And similarly, and so on and so forth, x3 will only be known if I know x4, and xi will only be known if I know xi plus 1. So, I am stuck here.

Except for the fact that we already found out that xn is equal to 1. xn is simply equal to 1. So what that means is the kind of recursive relationship that I am getting here. To find xi, I need xi plus 1. That is what I am seeing from the equation.

If I have x2, I can find x1. If I have x3, I can find x2, and so on and so forth. So, what that means is that if I have xn, I can find xn minus 1. Using that, I can find xn minus 2. Using that, I can find xn minus 3, and so on and so forth.

I keep going back until I find the variable x1. This comes from a system of linear equations that I can write up. So, this system of linear equations is: I know x0 is equal to 0. I know x1 is equal to transition probability p10 x0 plus transition probability P11 times x1 plus transition probability p12 x2. That is the x1 equation.

The x2 equation is: Transition probability p21 x1 plus transition probability p22 x2 plus transition probability p23 x3. Similarly, xi is equal to pi,i-1, xi-1 plus Pi,i xi plus probability pi,i+1 xi+1, going all the way to xn-1 as probability of transition from n-1 to n-2 xn-2 plus probability of n-1 to n-1 times xn-1 plus transition probability from n-1 to n times xn. The last equation is xn, which is just equal to 1.

That is what we started with. We know the endpoints. So these are n plus 1 equations. And the number of variables, because you should be able to see that 1 to n and then one more for 0. So there are these n plus 1 equations.

And there are n plus 1 unknowns because each one of these has to be solved for. And of course, for two of them, the answer is straightforward. But these n minus 1 equations have to be solved simultaneously to get the answer. These n minus 1 equations. And that is what we will look at in the next video onwards.

Thank you.