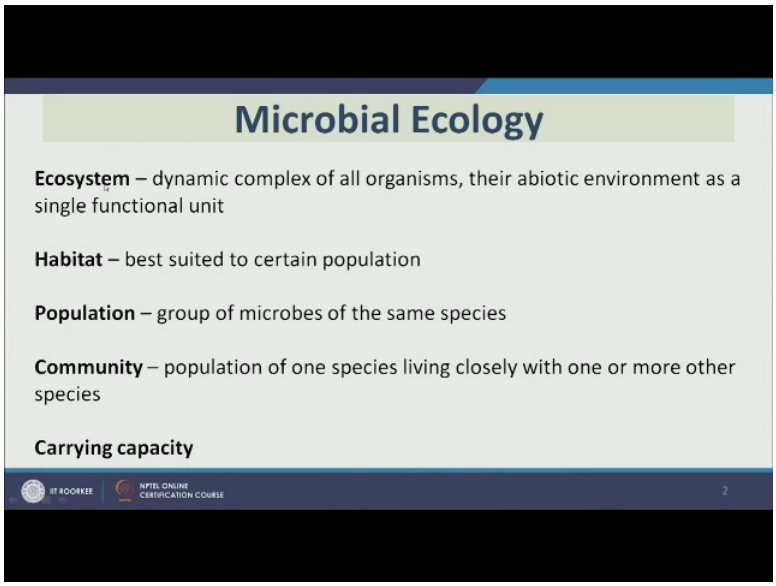
Applied Environmental Microbiology
Dr. Gargi Singh
Department of Civil Engineering
Indian Institute of Technology, Roorkee

Lecture – 16
Microbial Ecosystem I

Dear students, welcome to this lecture where we will start diving into ecological diversity in microbes and how microbes are influenced and influenced their ecosystems. And before we start understanding the environmental links that microbes have and how they influence and are in turn influenced by their environment, but we sort of already know the background off its important to understand the quantitative measures that we can use in order to give, in order to understand the progression of microbial communities. So, that is what we are going to do today in this lecture, microbial ecosystem we are going to be introduced to some certain very basic definitions of ecology and then move on to quantitative measures of diversity, already.

(Refer Slide Time: 01:05)



So, first ecosystem the basic term ecosystem is a dynamic complex of all organisms. So, microbes, higher order of lives, viruses, human beings all of them included, how they interact with each other, influence each other and they are abiotic environment. So, abiotic environment is also very important part of ecosystem and they are treated together as a single functional unit. So, if ecosystem is a system a single functional unit

that has living beings and nonliving components and they interact with each other they influence each other and together it is referred to as ecosystems.

If I say marine ecosystem then I mean a sea environment where we have certain abiotic elements such as high salinity we have water we have certain pH certain temperature and certain compounds that are present in sea water. And then we have the biotic elements we can have fishes, we can have mammals, such as dolphins and whales, we can have human beings and then we have microbes tons of them we can have viruses and plants and then together they interact with each other and together they refer to as marine ecosystem.

Next is habitat. Habitat is a way of referring to a place or in an environment that is ideal for a community to survive that is ideal for a population to live. So, it gives that population the food it requires the environment it requires. Habitat is also referred to not just the place the physical environment, but also an environment that can be created for or for a population to survive and to be happy end. Population, now this is very important in microbial ecology we are very specific about how we define population and how we define community. So, population is defined as group of microbes of the same species.

So, on species they have all the microbes within a population have to be same if I am talking about a community that has multiple species I cannot call them a population. Within species it can have different genus, different genre, but on species level all of them are same then they refer to as population.

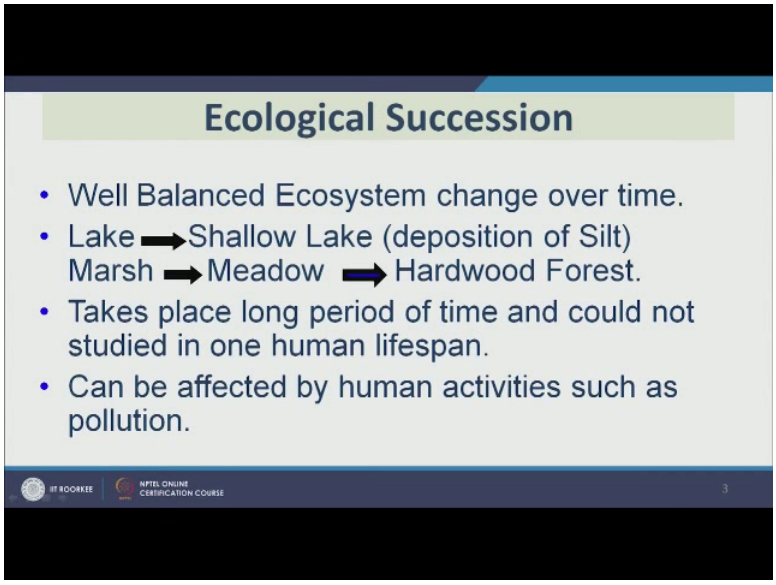
Community is when we have multiple species living closely and interacting with each other. Let us say we have multiple populations living and living with each other, but they are not interacting with each other they are not affecting each other, they are not competing, they are not contributing or anything at all they are existing independently then this is not a community it is a population.

Now, next important term that you need to understand is carrying capacity. So, each ecosystem coming back to the first definition each ecosystem which includes the abiotic components and biotic components and their interplay has an upper limit to how much life it can sustain. If you remember the logistic model of growth that we started we talked about carrying capacity. How the microbial population will eventually plateau up to a

particular point and it will no longer grow further from there that is the carrying capacity of the ecosystem. If population exceeds the carrying capacity then population will automatically decay a slow exponential decay and it will come back to the carrying capacity, already.

So, carrying capacity is another aspect of this phenomena is that, let us say there are it is a community of two microbial population and one of them is outcompetes the other. So, one of them starts increasing then in the same proportion the other has to start decreasing because the overall carrying capacity of my ecosystem is saturated.

(Refer Slide Time: 04:43)



Now, next is ecological succession. We know that microbial communities evolve with time a part of it has to do with changing, changing of the abiotic components of the environment. So, abiotic components of environment can change on a micro level or on more chemical level such as initially it was aerobic and then we had nitrate reducing environment and sulfate reducing environment and then we have methanogenic environment and fermentative environment. So, this is on chemical level how the abiotic components of the ecosystem are changing.

In fact, the entire geography might also change. For example, here I am showing and lake might turn into shallow, lake the deposition of cells, then marsh, and meadow, and then eventually hardwood forests. And we know that this has happened over time on the

surface of earth, over and over again in many most places in almost all places of the world. So, both the biotic and abiotic components of ecosystem undergo succession.

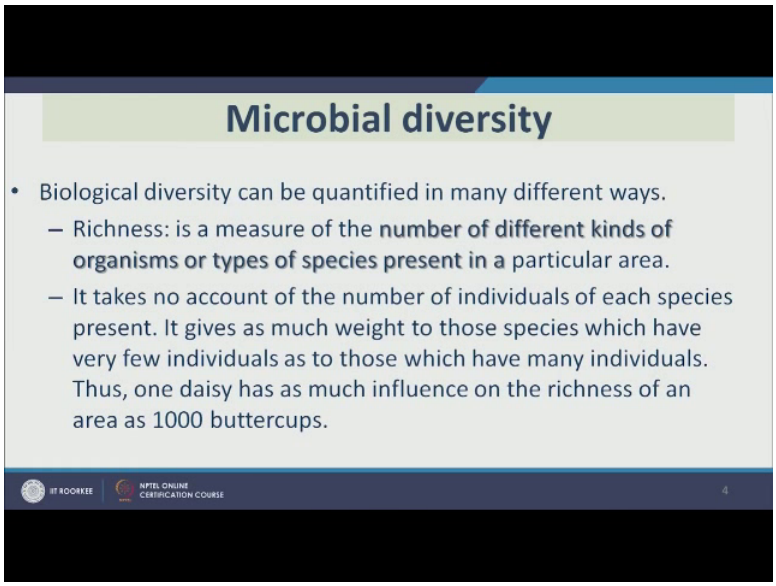
Now, note that this is not called evolution, this is called succession. So, with changing conditions the members change. You know first ever dinosaurs now they are its a mammalian world you know from reptilian they have moved to mammalian, microbial, reptilian and now mammalian and guess we do not know what will be the next you have to wait and watch I guess. But this is ecological succession and it is also evolution if new species come. For example, humans were not present always, but we are only new phenomena relatively. So, this is an evolution of life a new face of life and then humans are gone because remember everything that has the condition of evolution is that when humans are no longer fit to survive we will die as a species. So, when humans are gone something new might evolve life might evolve in new interesting forms. But succession happens when the ecosystem their community needs such changes without emerging into some new form of life and ecological succession in abiotic terms such as lake turning into hardwood forests usually happen in really large time skills.

Now, remember microbes replicate much faster than the time it would take for the lake to turn into shallow lake. So, for microbial purposes we are more interested in changing of chemistry the around and in the ecosystem. And this can be definitely affected by human activities it is actually being affected by human activities and one of the examples that I gave in a previous lecture was how antibiotic resistance has been horizontally transferred into pathogens from environmental groups that naturally carry a background and let us say not harmful levels of antibiotic resistance.

But now are pathogens have it and there has been a reason and the reason we suspect is anthropogenic reason we have increased the background levels of antibiotics and xenobiotics and other compounds that trigger microbes so as to the more eager to acquire antibiotic resistance. In fact, only the ones that have acquired will survive better or the ones that have acquired other capabilities along with antibiotic resistance will survive better. And does, we know that human activities affect their ecological succession. So, when is few decades ago pathogens were largely antibiotic susceptible now the antibiotic resistant and this expression has been facilitated largely by human activities.

Now, next thing that we want to understand is microbial diversity. Now, microbial diversity can be quantified in many many different ways, the many measures the many definitions that we have for microbial diversity and each of them points at a one particular feature of microbial community.

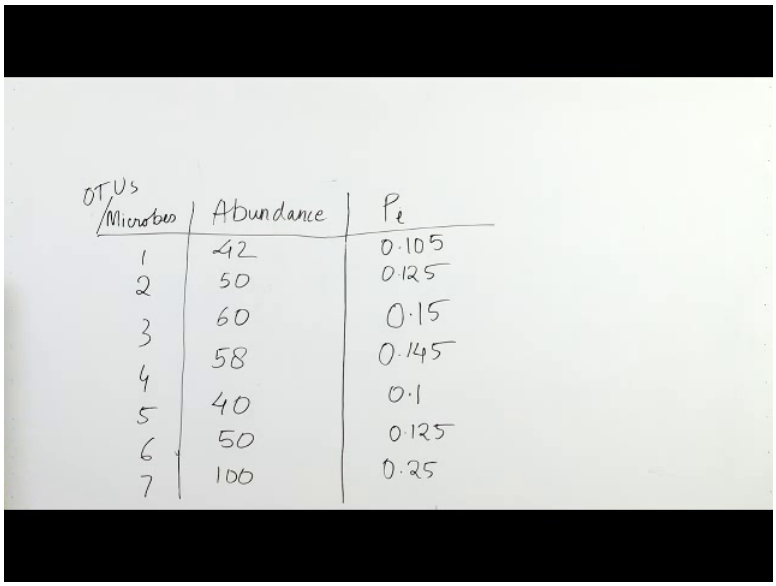
(Refer Slide Time: 07:53)



So, first one is richness. Richness tells me how many different kinds of organisms or species are present in a particular area. So, let us look at this example.

Let us say I have a sample and I sequenced the sample and I found out that there are excuse me; that there are 7 different kinds of microbes in the sample and thus the richness of the sample is 7.

(Refer Slide Time: 08:22)



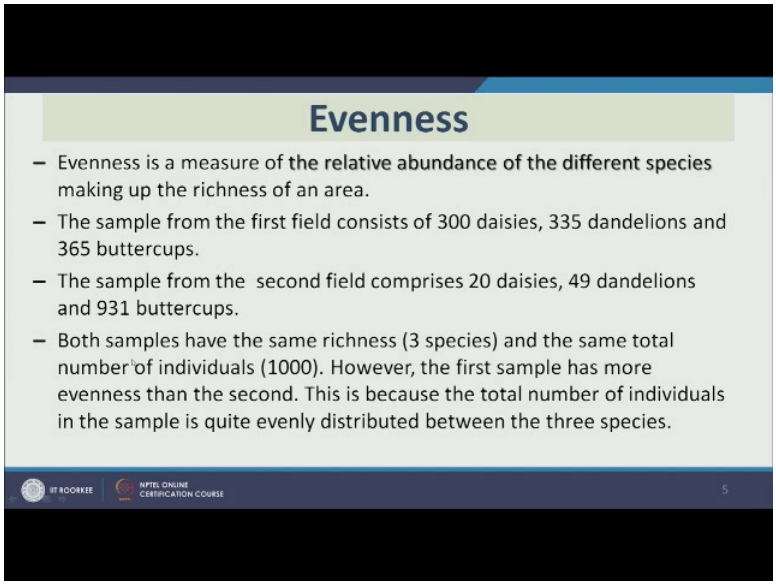
So, if I just count a number of different microbes my answer is 7, but here is the thing a simple parameters such as richness does not inform me which of these are dominant, is it in even community or is it highly skewed community where a single microorganism let us say the species number 7 out completes every one of them and dominates every one of them.

For example, university campus could have different kinds of life forms, human beings, certain animals such as dogs, cats, rats, birds, certain birds. So, if we list them together we will notice we can get the species, how many species are present and thus we get the richness. But it still does not inform us if a particular species is out competing everybody else. For example, there are more humans than birds in certain campuses. So, we can say this is mostly a human community with a little bit of other species present and this community would be different from another university. Let us say more idyllic university and I am imagining Shantiniketan or some other similar place where we have certain amount of humans, we have certain amount of animals and they all are relatively more even. So, that is a very different ecosystem than our university campus when humans exceed all other kind of animals. So, the richness in itself will not inform me, how the relationship is between these microbes or species and if it is an even distribution or not?

So, we note that richness in itself does not take account of number of individuals of each species it gives equal weight to everyone. Then, the example I have given the example I

have used in this presentation in this lecture is up for plants and trees. So, one daisy plant has as much influence on the richness as 1000 buttercups.

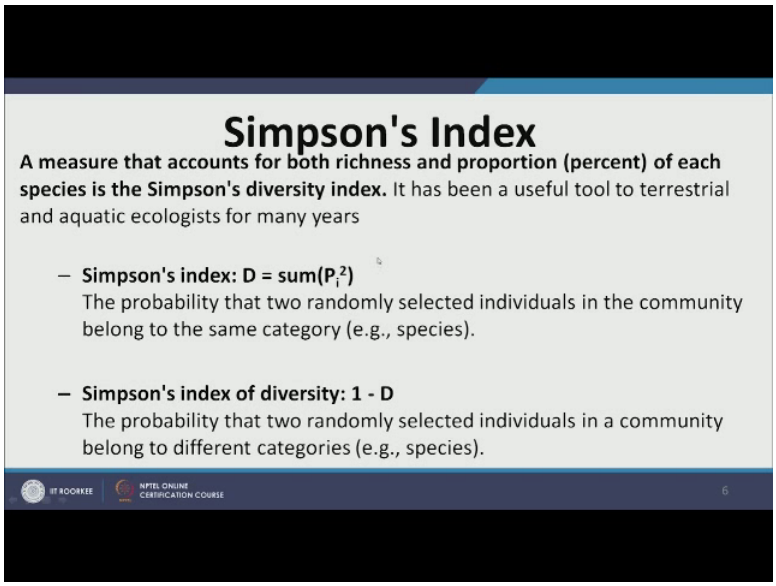
(Refer Slide Time: 10:15)



On the other hand next parameter we have is evenness. Evenness looks at the relative abundance of different species that make up the richness of an area.

So, let us say we have first we have two fields one field has 300 daisies, 335 dandelions, 365 buttercups. The second field has 20 daisies, 49 dandelions and 939 buttercups. So, both are equally rich in terms of richness number of species that they have both are equally rich both are three different kinds of flowering plants. However, we note in the first one there is a more evenness. There are 300 sum of each of the three whereas, in the other we have most of them are buttercups and few of them are daisies and dandelions. I have no idea how they got so few dandelions there are such weeds. Now, both samples have same richness and the same total number of individuals 1000 flowering plants; however, the first sample is more even than the other and this is because the total number of population is more evenly distributed.

(Refer Slide Time: 11:16)

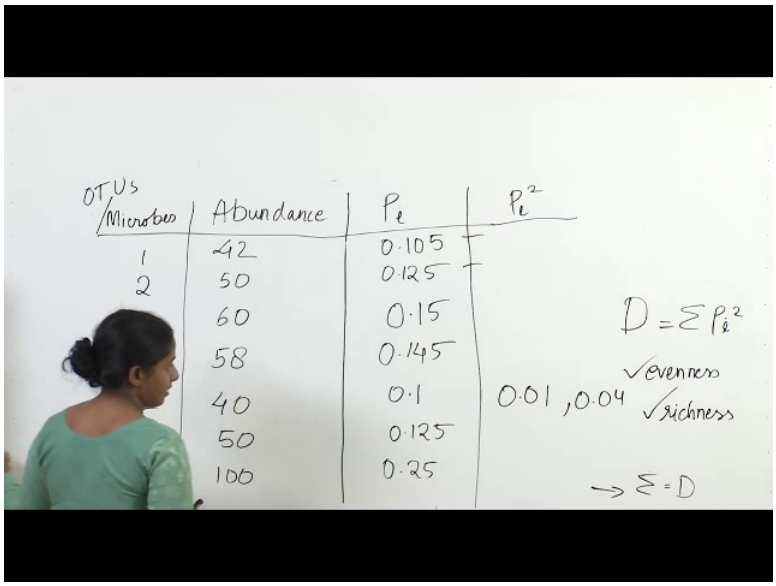


The next parameter we have is Simpson's index. Now, for a microbiologist or for a scientist or an engineer or just science literate science who is very interested in microbiology Simpson's index is very popularly used. It is a very important and potent index for diversity because it accounts for both richness and evenness. So, it looks at both richness and the proportion of each species and if that is that and then we calculate Simpson's diversity index. Now how do you calculate it? Simpson's index can is referred to as with capital D and this is by summing all P i square.

Now, what is this P i and what is this D? Let us proceed with the reason with an example. We have 7 different species of microbes and we know how many of these microbes were detected. For example, the first species we detected 42, for second we detected 50 and so on and so forth, a total of 400 were typed and we have their relative abundance. So, divide 42 by 400 the total number and so on and so forth and we get the relative abundance.

Now, in order to calculate Simpson's index or D which is summation of all P i square.

(Refer Slide Time: 12:28)



What we need to do is we need to find their square P i square which I am not going to do in the class, but you need to do them and then you can do their sum total and you can get your D. This is how you calculate Simpson's diversity index.

Now, remember I mentioned that Simpson's diversity index accounts for both richness and evenness. So, let us look at how it accounts for richness. Because it is summation of individual entities, the more the entities are the more likely the greater number, the greater magnitude of my Simpson's index would be. If I have 7 versus if I have 70 are more likely to get a higher number of Simpson's index if I have 70 versus if I have 7. Thus it does account for richness, so richness; check.

Now let us look at evenness. This is the relative abundance and when we square it, it will reduce even further. You know if something is less than 1 and you square it, it gets even smaller in magnitude. So, in this sense because we are talking about relative in abundance and what we are adding up is the square of relative abundances we are automatically including evenness.

The way it works is the closer the number is to one if it is less than one its square will be less small compared to if it is further apart. For example, the square of 0.1 would be 0.01 whereas, if somewhere I had 0.2 I do not have 0.2, but if I had 0.2 then the square would have been 0.04. So, here we have a difference of 2 and here we have a difference of 4, so by 4. Thus it accounts even for evenness. Because Simpson's index is summation of

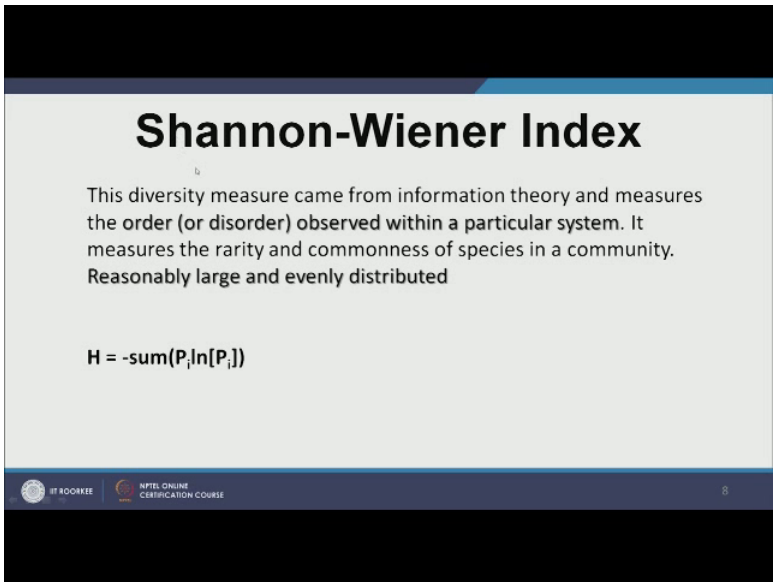
relative abundances it can never be more than 1, it will always be less than 1 and if we subtract it from 1 we get Simpson's index of diversity which is the probability the 2 randomly selected individuals in a community belong to different categories which helps us define Simpson's index as the probability that 2 randomly selected individuals in the community belong to same category.

So, if there was a way in which we could pick microbes out of a bank and let us imagine these 400 microbes were sitting in the bag and I could pick them up and see what kind of microbe they are. So, the probability that the two consecutive microbes that I will pick belong to same category will be given to me by Simpson's diversity index, subtracted from 1. So, if we subtract Simpson's diversity index from 1, then it will tell me what is the probability that two randomly selected microbes belong to same species all right.

So, now let us look at Simpson's reciprocal index it is one upon D. Now, if I take my sigma log P i square and I raise me get its reciprocal it gives me and it gives me an indication of equally common categories that will produce the observed Simpson's index. So, what it means is that, 1 upon D will tell me how many species if they are perfectly even in their distribution. So, equal number of each members of species will it required to produce same diversity. So, this gives me an idea of evenness.

If it requires the same number of or similar number of quite close by number of species then what I actually have here then I can say oh my sample is pretty even. But if it requires much less let us say I have 7, but it requires only 2 equally even species to produce same Simpson's index reciprocal then I know that that my species my community is not diverse at all, it is not even at all. So, these influenced by two parameters the equitability of percent of each species present and richness which is relative abundance and richness. For given species richness D will decrease as the percent of species becomes more equitable and vice versa.

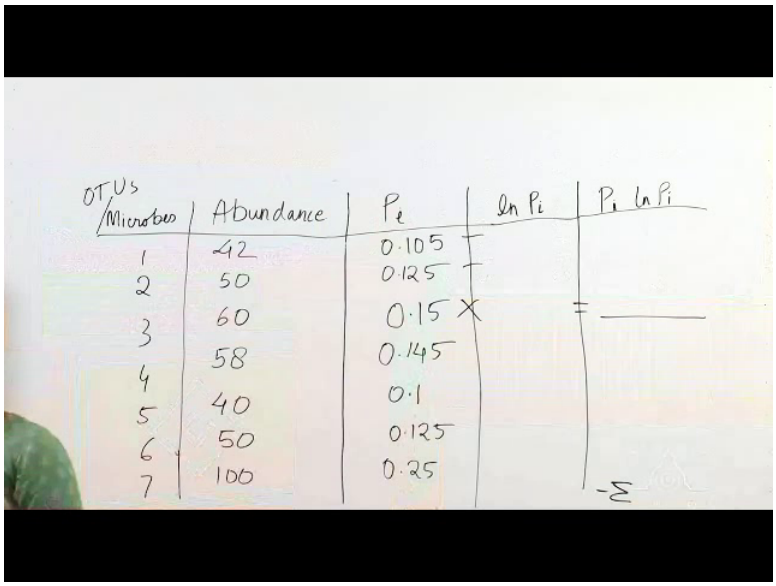
(Refer Slide Time: 16:40)



Next very popularly used indexes Shannon-Wiener index. So, this diversity measure has come from information theory and it measures the order observed within a particular system. So, it measures the rarity and commonness of a species in a community and it accounts for both richness and evenness giving more importance to evenness. So, what it will look like H is equal to minus summation of P i log P i and I have used natural logs and people use logarithmic with base 10 which is also ok.

So, to calculate Shannon-Wiener index you need to find out logarithmic of P i and the standard way is using natural log, so you need to find natural log of each of these P i.

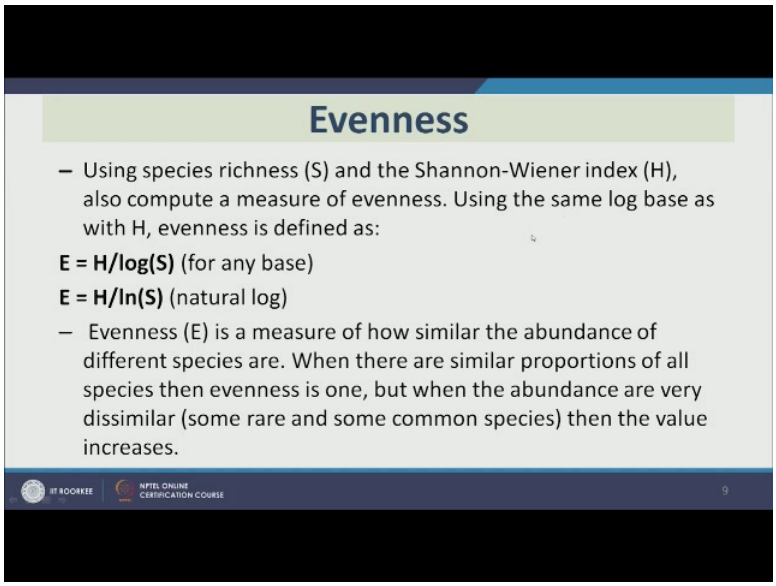
(Refer Slide Time: 17:34)



So, use your calculators and convert your relative abundance into natural logs and then you need to have another column here that will have the product of the two, product of this with this. So, you multiply these two and you write it here. Now, add all of these. So, you get your summation and then put a minus sign in front and you have calculated your Shannon-Wiener index.

So, well we have been talking about evenness this so much, so let us use Shannon-Wiener index to get good insights into evenness. So, if we know the value of species richness how many species are present and Shannon-Wiener index we can compute a parameter for evenness a measure for even this so, but make sure you use the same long piece that you use for calculating your Shannon-Wiener index. If you use natural log like how I showed you make sure in this formula you use the natural log and if you use the logarithmic with base 10, you use logarithmic base 10.

(Refer Slide Time: 18:26)

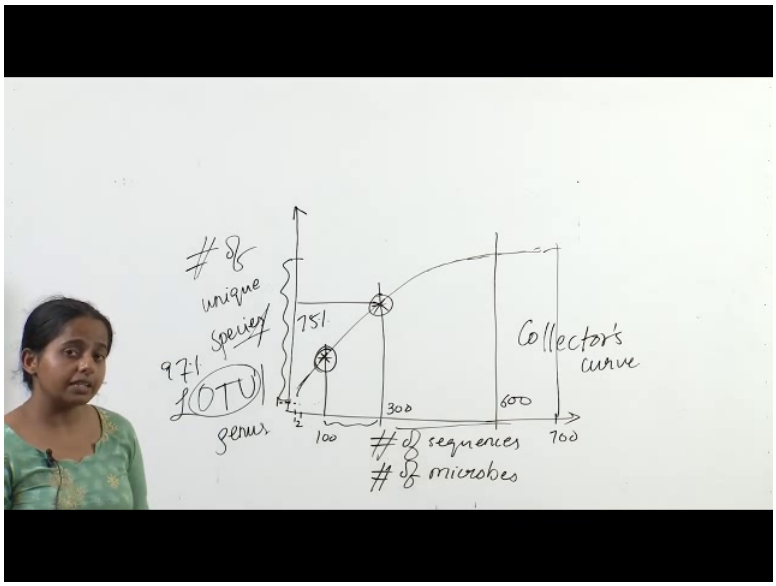


So, if you divide the Shannon-Wiener index by logarithmic of your richness you will get evenness. So, evenness is a measure of how similar the abundance of different species are we already talked about it. Then there are similar proportions of all species and evenness is 1. So, the closer even this is to 1, you know that it is perfectly even the further away you know it is not even. But when the abundances are very dissimilar then the value increases all right.

Then there are some other parameters that I want to talk to you about in this class before we move onto next lectures that will extensively talk about these parameters talk use these definitions to explain the concepts that we are going to cover. The first is collectors curve and there are multiple parameters that are used to quantitatively understand the meaning of collectors curve, but let us just try to understand the idea what collectors curve is and it is all the other name for collectors curve is rarefaction curve, but let us look into it.

Collectors curve. So, let us imagine again I can pull microbes say genetic signature or microbes from a bag and I can see what kind of microbes there and in fact, they are tools that allow us to do, but I just do not manually pull them it is a machine that does that. So, let us say I sequence microbes and this is the number of sequences that I pull up or this could also be number of microbes that I type.

(Refer Slide Time: 19:54)



So, remember in previous example we type 400 different microbes from a sample and it could also be sequences and this is number of unique species, number of unique species.

Now, does not have to be species, it can be OTU it can be something else like genus. Now, this OTU is a unique term. So, let me explain to you what OTU is before we move on to collectors curve. So, OTU stands for Operation Taxonomic Unit. So, when we sequence genetic material sometimes it is not very meaningful to give it a name like acidobacteria for example, all of them might be acidobacteria. But still we know that they are very diverse within themselves, but we cannot type them if more finely because our genetic sequences very short or we have sequenced, the parts that are more conserved. So, they are conserved across acidobacteria and we do not know what kind of acidobacteria it is. So, in that cases we use OTUs. So, in OTUs what we do is we look at genetic similarity and we decided number quite arbitrarily sometimes.

Let us say we had decided a number 97 percent. So, all sequences that are 97 percent similar to each other or more similar to each other are called as one operational taxonomic unit. So, they are referred to as one kind of microorganism. We do not know if they are similar on species level on phylum level or going more finer on genus level or strain level even we do not have that information, but we just say for operational purposes for purposes of making sense out of data we call them one operational

taxonomic unit. So, they are one kind of microbe and most of the plots that I have seen in recent days on y axis they plot number of unique OTUs.

So, let us say I pull out sequences or microbes from my back proverbial bag and I get one unique sequence of course, when I pick out my first sequence and then I pick out another and I get the same sequence. So, I picked out another, but I have still the same sequence already and then as I keep picking out more sequences or more microbes I start getting more and more unique samples, already. So, I get more and more unique sequences and here notice I am getting a lot of unique sequences and then after a while no matter how many more I pull I will stop getting more unique sequences because all the unique sequences have already been seen or most of them at least have been already seen.

So, it will plot you and this is how a typical collectors curve looks like. Well not like growth model per se, but it looks like a logistic curve, so it looks like this. So, initially there is a very fast rate of getting unique sequences because our microbial communities are very diverse and then it plateaus off. So, this is collectors curve it is also called as rarefaction curve and this is very important curve because it gives you an idea of how efficient my sequencing was or my sampling was. Let us say this is corresponding to 100 microbes this is corresponding to 300 microbes and this is corresponding to 600 microbes yeah, let us say that.

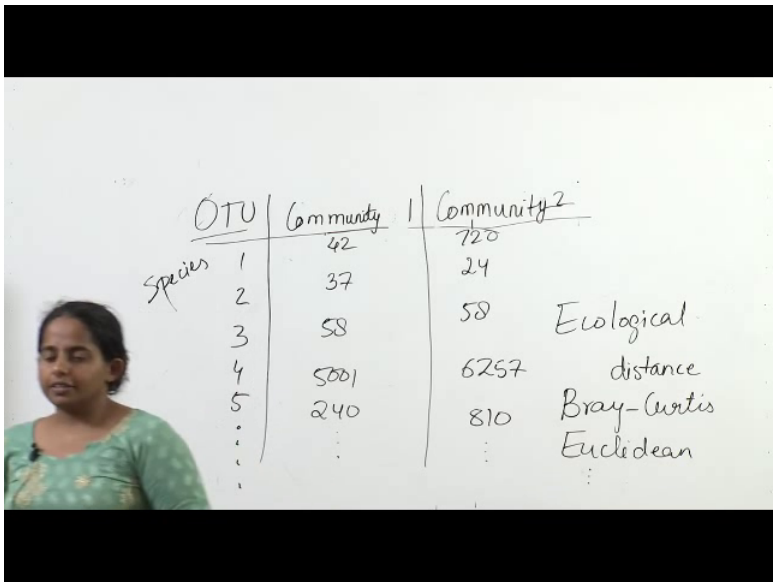
So, if I tell you that hey man I have 100 microbes that I have sequenced or I have sampled and I know I can tell you what the diversity of the sample is based on this data point. So, if you have collectors curve you can say well 100 is not enough you need to sample more to capture all the unique sequences. So, these are the maximum unique sequences present in the sample to capture all of them you need to sequence 500 more.

If I have captured 300 you know well whatever information you have is about 75 of microbial community. So, my sample in if I have sample 300 sequences or 300 microbes in this instance informs me about this percentage of microbial community. If I have sequence let us say 700 then you can say oh well you have sort of over sequence because at 600 you already hit the plateau, but you have sequenced all the unique sequences or most of them anyway.

So, this is how collectors curves are used and they are very important and useful for high throughput amplicon sequencing or high throughput whole genome metagenomics or metagenomics in general and they help me understand what is my sequencing depth sufficient or not. In other words radar sequence enough or not to get a complete picture of microbial community or am I only getting a picture of more dominant species or just the ones that are randomly sampled first.

Next word that I want you to understand is similarity and I have sort of already covered about similarity in previous lecture, when I talked about how we draw phylogenetic trees. So, let us say I have two sequences and I can look at their sequence ATGC whatever the sequences and I can match them to each other, align them with each other look for the dissimilarities and calculate the percentage differences that they have and get an idea of their similarity or their dissimilarity. But there is another aspect in which we can use our concept of similarity between microbes which is when we talk on community level, how similar or dissimilar are two microbial communities to each other.

(Refer Slide Time: 25:38)



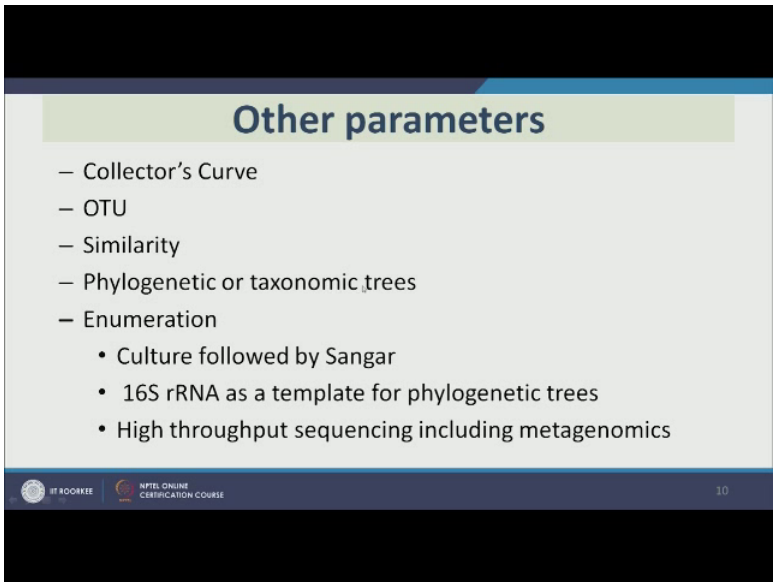
So, let us come back to our old example. So, in this example I have two populations one mistake already. So, in this example we have two communities, community 1 and community 2. Each of these communities have different species. So, these are species or they could be OTUs we do not know depending on our experimental design and we have many of them. So, I have not listed all of them, but only I filled the first 5 and each of

these communities has certain members of each OTU or species for example, the first OTU or species in community 1 is 42 members detected and in second 720 members detected.

Now, if I want to find out how similar our community one in two to each other or how dissimilar they are to each other. Then there are indices that will help me find it out and these are known as ecological distances and some of them that I can name to give you an insight and a direction to move on our Bray-Curtis distance, Euclidean distance and if you look up a Pivec if you look up a vegan package in r which is a software you will find a list of many different distances that ecological distances that help me understand how similar or dissimilar the microbes are to each other. So, we have Bray-Curtis we have Euclidean and we have many more.

The next thing I want you to understand is phylogenetic or taxonomic tree and in one of the previous lectures we have talked in length about it. If you still are not clear about what they are please go back and refer to them and in subsequent lectures I will tell you how to make trees which is it would the tree part of this title we will get more clear to you.

(Refer Slide Time: 28:32)



Next is enumeration. Now how do we enumerate microbial communities? You earlier and the first developed process was first we cultured the microbe in the lab is a pure culture.

So, there is only one microbe growing up to the strain level and then we sequence it using Sanger sequencing and then the technique came that instead of focusing on sequencing using Sanger or the entire genome we can only sequence 16S rRNA using Sanger or using some other method, but we are not focusing on the entire genome we are interested only in 16S rRNA and the sequence it and they use it to understand what kind of microbe it is. And this is very important and relevant when we look at high throughput amplicon sequencing which is a form of metagenomics for our environmental sample. And then we have which is here high throughput amplicon sequencing. This can also be done by Sanger by the way. Here we had the option of doing whole genome aerodynamics when instead of looking just as 16S rRNA we can align it with our data base and look at the entire genome to find out what kind of microorganism it is.

So, dear friends this is all for today. In the next class we will explore the microbial ecology even further and look at the diversity and look at different aspects of microbial ecology and population dynamics.

Thank you very much.