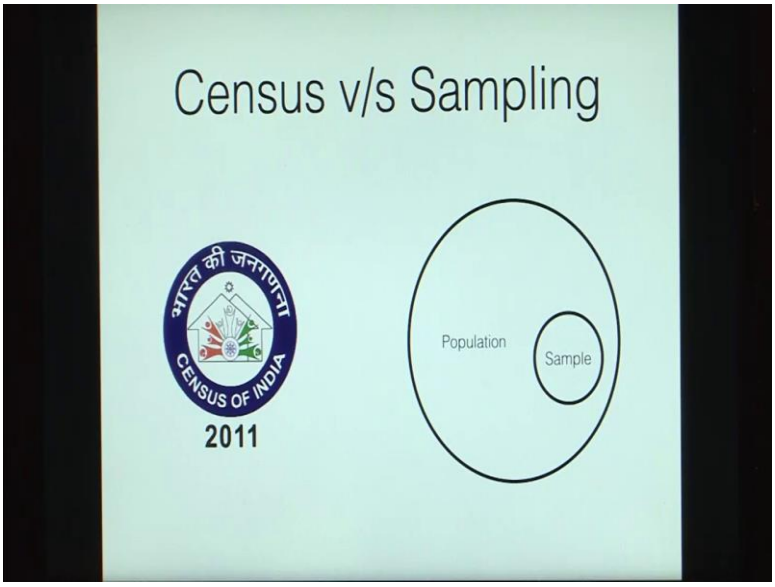
Forest Biometry

Prof. Mainak Das
Dr. Ankur Awadhiya
Department of Biological Sciences & Bioengineering & Design Programme
Indian Institute of Technology, Kanpur

Lecture - 35
Forest Sampling

[FL]. Today, we are going to have a look at forest sampling. So, let us begin.

(Refer Slide Time: 00:22)



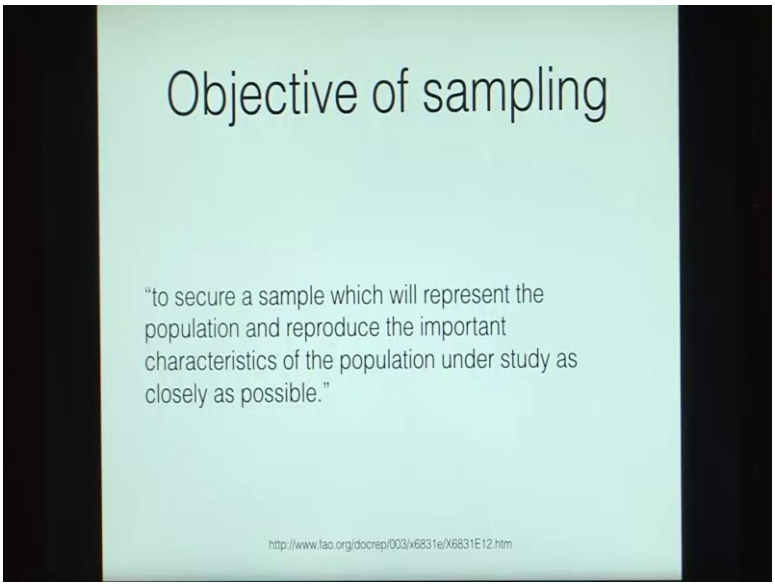
We all know what a population census is. So, every 10 years the census of India is undertaken in which all individuals that are there in India or the citizens of India are counted and a number of different parameters regarding those for instance their sex, their religion, their ages, etcetera are also noted down. Here we are talking about forest sampling. So, what is the difference between census and sampling.

So, if you look at the slides now, so when we are talking about a census, a census measures each and every person that is there in the population. So, we are not going to leave any person when we are undertaking the census of India. Similarly, when we are doing a tiger census for instance, so a tiger census is done every 4 years, and in a tiger census all the tigers will be counted. On the other hand, in the case of a sampling if we

have a huge population, we are going to take a small sample of the population. And we are only going to take parameters from this sample not from the whole of the population.

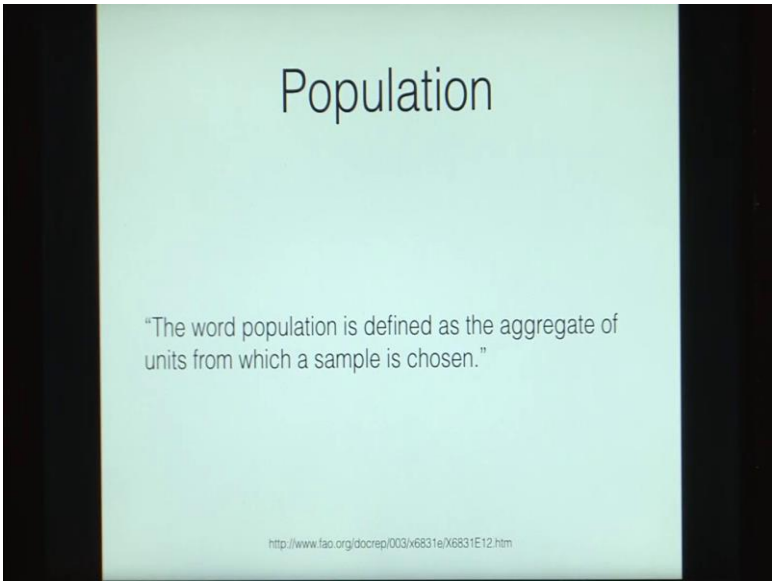
So, we will take a representative sample. So, for instance suppose you have 60 students in a class, and if you take a sample of say ten students and you try to figure out how much time do this spend in studying. So, we will take this sample, will take their parameters and will see that that is an average value that we can use for the whole population as well, so that is the basic difference between census and sampling. In the case of census all the individuals are counted; and in the case of sample just a small proportion of your population is counted or its parameters are measured.

(Refer Slide Time: 02:19)



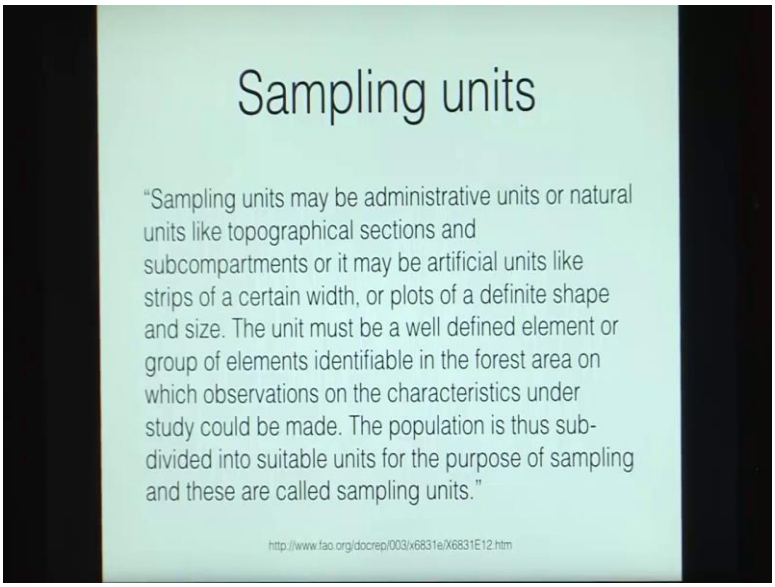
So, what is the objective of sampling? So, the objective of sampling is to secure a sample which will represent the population. So, we want a representative portion of the population. So, to secure a sample which will represent the population and reproduce the important characteristics of the population under study as closely as possible. So, what we are trying to do here is we are trying to take a representative small portion of the population, we are trying to measure its parameters and then we are trying to get a value for the population as a whole. So, we will get an average value for the whole population. When we are talking about population and samplings, how do we define a population?

(Refer Slide Time: 03:02)



So, the word population is defined as the aggregate of units from which a sample is chosen. So, for instance in the case of forest, we have n number of trees. So, all these trees will together form a population, so that is the aggregate of all different units one unit in this case is one tree. So, the aggregate of all the trees in the forest from which we are going to chose a sample is called the population.

(Refer Slide Time: 03:31)

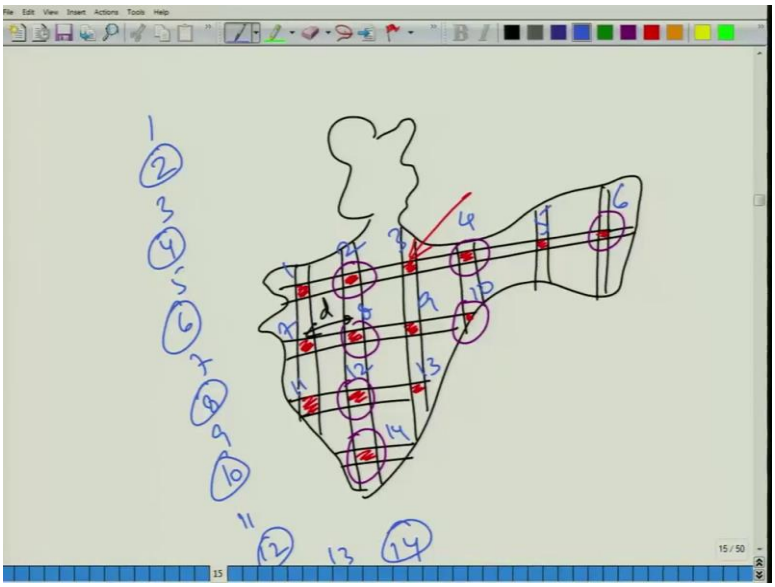


Next, we can define the sampling unit. So, a sampling unit may be an administrative unit or a natural unit like topographical sections and sub compartments or it may be artificial units like strips of a certain width, or plots of a definite shape and size. What we mean

here is suppose in the case of taking a sample of the population of India, so we can go for administrative units. So, for instance, we can take a Tehsil as an administrative unit that will be use for sampling. So, Tehsil can be a sampling unit or it can be a natural unit. So, for instance, you can say that the whole of the Satpura region.

So, Satpuras are hills of Madhya Pradesh. So, the whole of this Satpura eco region can be called a sampling unit. So, it will might be a an administrative unit or a natural unit for a example topographical sections. So, a Satpura is a topographical section or its sub compartments or we can even take artificial units such as strips of a certain width.

(Refer Slide Time: 04:44)



So, what we mean by that is in the case of India, we can just define our artificial units as these strips that are at a fixed distance d from each other. Then we can also define some horizontal strips again see at a distance of d 1 from each other, and then we can say that these portions will be our sampling unit. So, this is a completely artificial way of getting samples, because it is not corresponding to our a towns or cities or villages or even any natural areas or any topographical areas.

So, for instance, this point might be a city, it might be a town, it might even be a barren land also it might fall in a mountainous region, a hilly region a plane track may be it might be even fall on a river. So, in this case, we are taking an artificial unit which is called a sampling unit in this case. The unit must be a well-defined element a group of

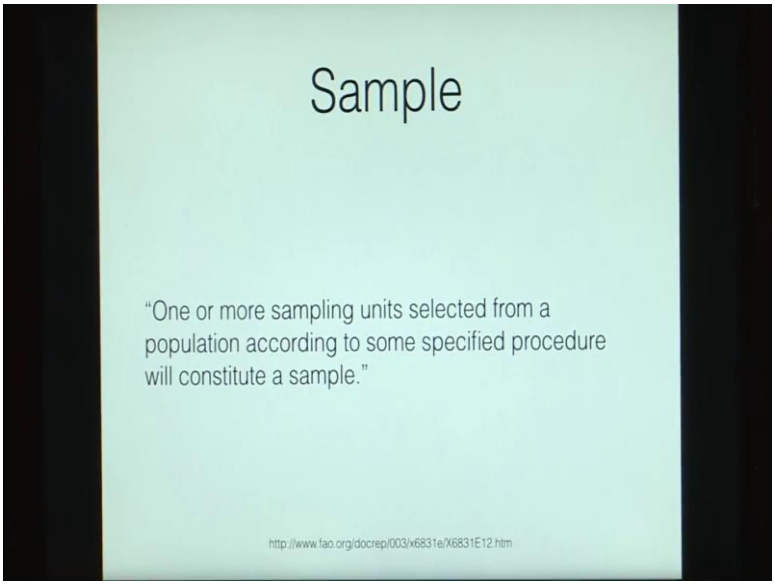
elements identifiable in the forest area on which observations on the characteristics under study could be made.

So, whenever we defining a sampling unit, it must be a well-defined elements; so whenever we are talking about a sampling unit, we must be able to define it properly. We must be able to go back to that unit again and again, if that is a good sampling unit. And when we have a well-defined sampling unit then we can take observations on the characteristics under the study. The population is thus sub-divided into suitable units for the purpose of sampling and these are called sampling unit.

(Refer Slide Time: 06:40)

Next, we can define a sampling frame. So, when we take a list of all the sampling units for our population then this list is called a frame. So, it is called a sampling frame.

(Refer Slide Time: 06:54)

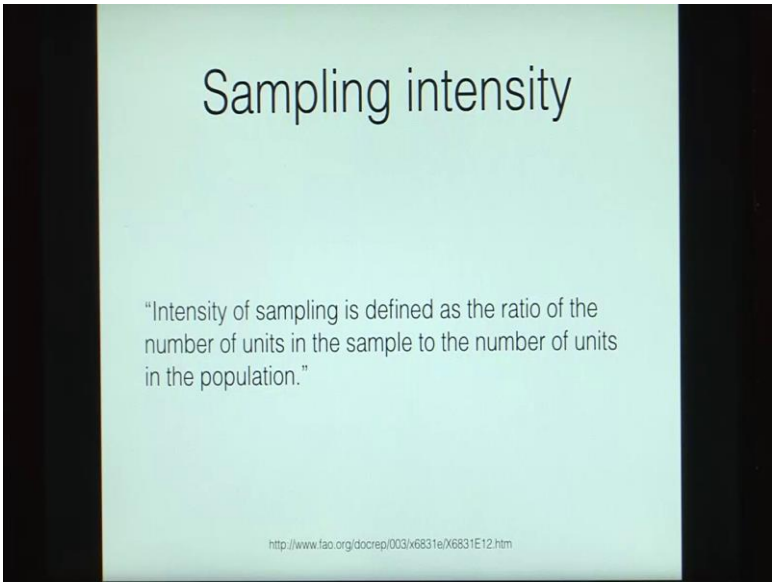


Now, once we have define those, we can define a sample. So, a sample is one or more sampling units selected from a population. According to some specified procedure and that will constitute a sample. So, for instance when we talked about this map and we had defined so many sampling units then we can say define that we are going to take every second sample sampling unit as a sample.

So, in this case this will be our sample. So, here we had 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14. So, we have defined fourteen sampling units the list of all this fourteen sampling units. So, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, so this list will be a sampling frame and if we say that that we are going to take every even value as our sample. So, our sample would comprise of these sampling units. So, this will become our sample.

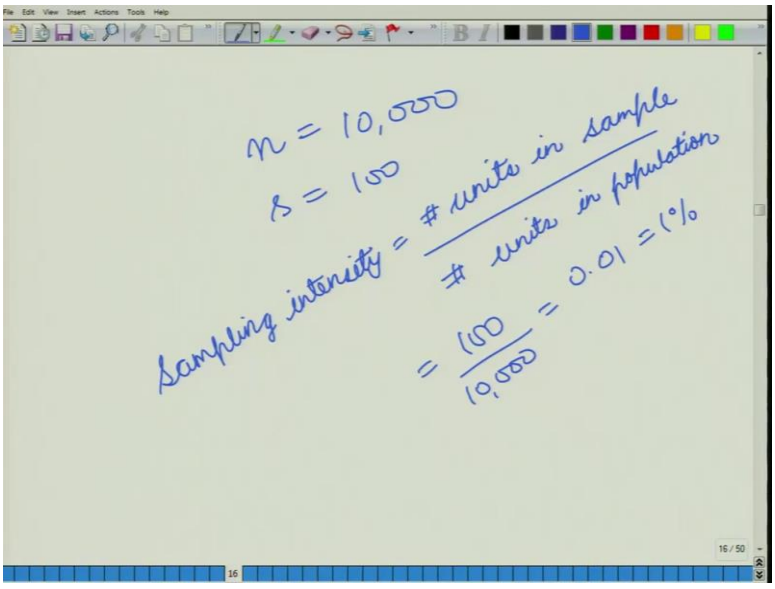
So, coming back to this slides one or more sampling units selected from a population according to some specified procedure now that procedure might vary. So, it might be a random sampling or we might say that we are going to take every second value or say every third value or every nth value. So, your specified procedure will vary, but once we have selected your sampling units they form a sample.

(Refer Slide Time: 08:41)



Next, we can talk about the sampling intensity. Now, a sampling intensity or the intensity of sampling is defined as the ratio of the number of units in the sample to the number of units in the population.

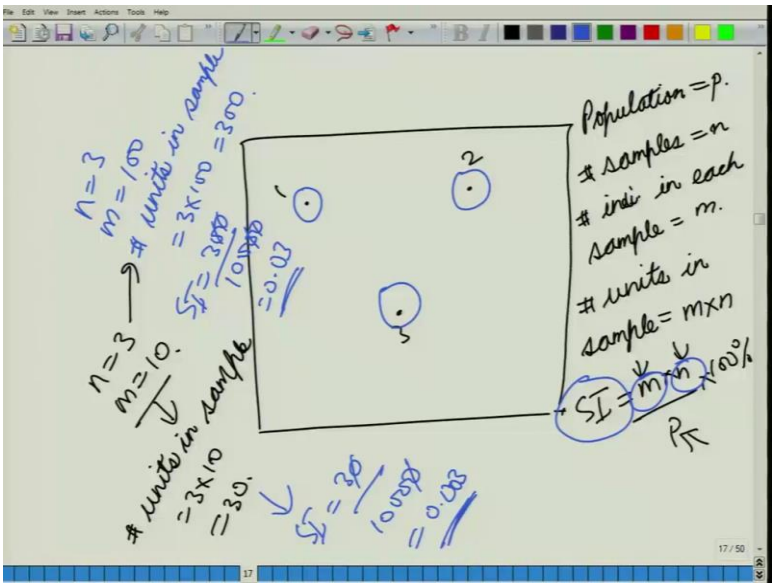
(Refer Slide Time: 09:01)



So, for instance, in a forest the total number of trees is say 10,000 and in our sample we have taken 100 trees. So, what is the sampling intensity? The intensity of sampling is defined as the ratio of the number of units in the sample. So, we have the sampling intensity is the number of units in sample divided by the number of units in population.

So, in this particular case the number of units in the sample is 100 the number of units in the population is 10,000. So, will get a value of 0.01 or we can also write it as 1 percent. So, in this case our sampling intensity is 1 percent.

(Refer Slide Time: 09:56)



So, this sampling intensity can vary because suppose in this population we have taken say three samples. So, let us say that your population is p number of samples is n and number of individuals in each sample is m. So, the number of units in sample will be m into n because we have n number of samples and each sample has m individuals. So, total we have m into n number of individuals and the sampling intensity will be m into n by p into 100 percent, so that is our sampling intensity. Now sampling intensity can be increased by increasing m or by increasing n or by reducing p. So, now, in most cases our p is constant.

So, in a forest that has 10,000 trees we can do nothing about the number of tress that are there in the forest, but we can vary m and n. So, for instance in place of taking these three samples 1, 2, 3 and suppose each sample had 10 units. So, here we are defining n is equal to 3 and m is equal to 10. So, now, suppose so the total number of sampling units or the number of units in sample is 3 into 10 is 30. Now, suppose we increased n and suppose in case of taking our three samples we took two other sample so 4 and 5. So, we have change n is equal to 5, keeping m constant. So, the number of units in sample becomes 5 into 10 is 50. So, in this case the sampling intensity was 30 by 10,000 in this

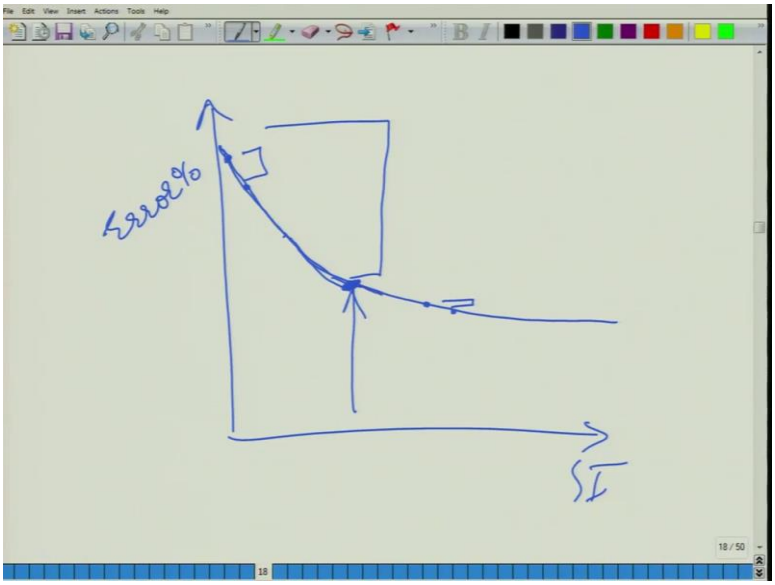
case our sampling intensity becomes 50 by 10,000. So, in this case it was 0.03; and in this case it becomes 0.05. So, the sampling intensity has increased by just increasing the number of samples that we took.

At the same time, we can also increase the number of individual in each sample. So, suppose we keep the number of samples constant as 3, but we in place of taking small sample we took larger size sample. So, suppose m was 100 in this case, so we are not taking any new samples. But in place of these small samples, we are considering larger sized samples. So, in place of having 10 individuals, now we are having 100 individuals. So, the number of units in sample now becomes 3 into 100 is 300 and the sampling intensity becomes 300 by 10,000 is 0.3. So, it increase from 0.003 to 0.03. So, essentially our sampling intensity could be varied by increasing the number of individuals in each sample or the increasing the number of samples that we have taken, so that is about the sampling intensity.

Now, how does sampling intensity matter to us, because suppose our sampling intensity was 100 percent? So, we are measuring each and every unit in the population. So, in that case our sampling will become equivalent to a census. So, for a 100 percent sampling, intensity it is a census. But when we reduce a sampling intensity, now why do you want to reduce a sampling intensity because if w can get the same amount of data by not measuring 10,000 individuals, but by measuring 100 individuals only then it is much better to just take our data for the 100 individuals.

So, by increasing our and suppose we keep a sampling intensity of zero percent. So, zero percent means that we are not measured any individual. So, the amount of errors that we are getting if you do not measure any individual will be very large because that will be any arbitrary value; if we are measuring all the individuals then our error becomes goes down to 0. So, as our sampling intensity increases the amount of error decreases.

(Refer Slide Time: 14:56)



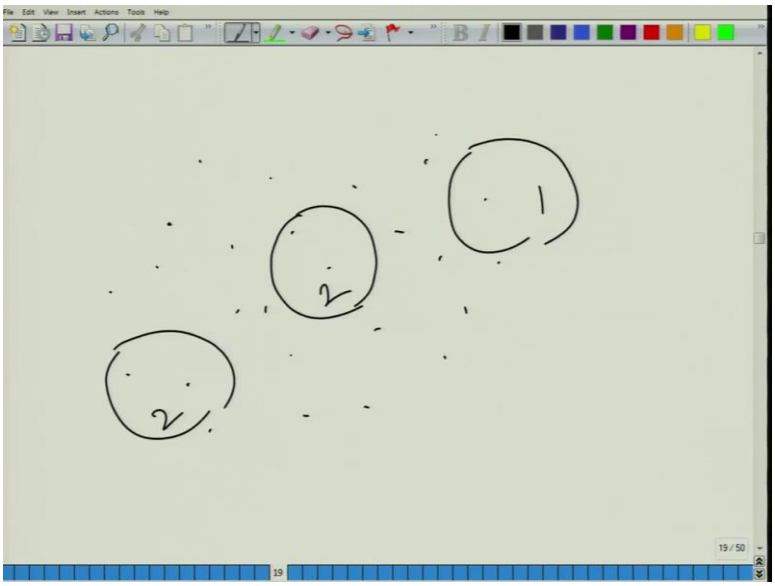
So, if we for instance took error percent versus the sampling intensity, so our error will reduce, but it will not go on reducing completely after while it will flatted out. So, if you are taking these many sample, so these many samples, the amount of error is varying a lot; but if you take these many samples or you take these many samples, this hardly any amount of difference in the percentage of error that will get. So, it is always better to take a sampling intensity at a point at which your amount of error has decreased considerably and now with every addition in the sampling intensity the amount of benefit that you are getting is now reducing.

So, it is always good to have your sample size in such a way. So, the sample size and the number of individual or your total number of units in the sample in such a way that the error is less, but the sample size is also not very large, so that is the utility of sampling intensity.

(Refer Slide Time: 16:08)

So, when we talk about these samples, so we can have different kinds of plots. So, our plots in which we are taking our samples could be circular, they could be rectangular they could be in the form of strips, or they could be topographical units, especially those that are used in the case of hills. So, when do we take a circular plot, when do we take a rectangular plot, when do we take a strip like sample and when do we take a topographical unit for our plots.

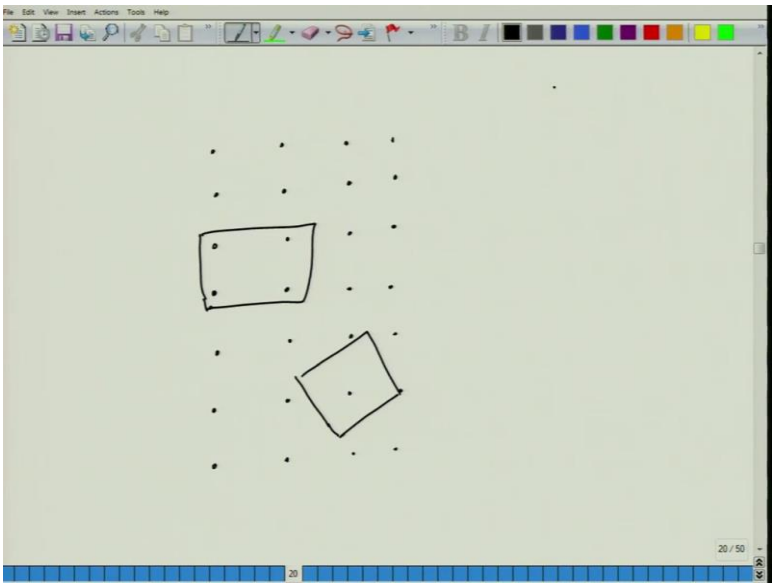
(Refer Slide Time: 16:38)



So, suppose our individuals are distributed randomly. So, these are our units that we want to take a sample of. So, in this case, if we took a circular plot, so a plot here

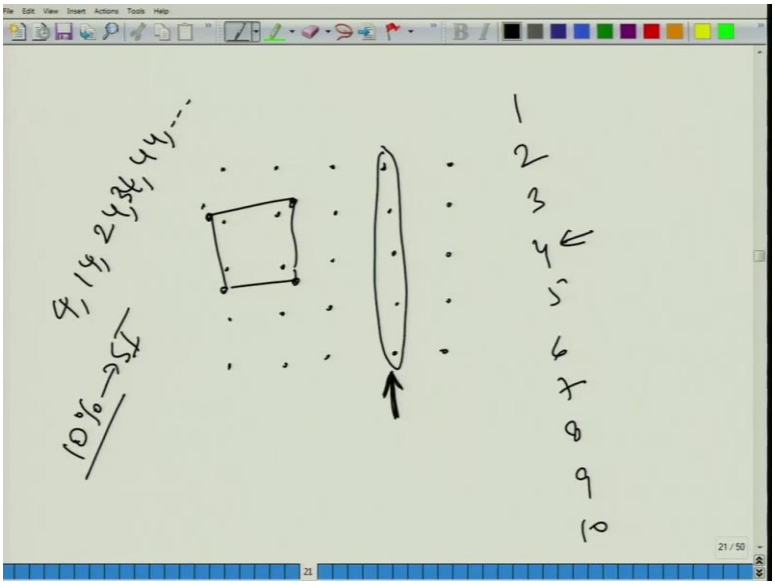
suppose a plot here a plot here. So, in this case, we have two individuals here two here one here. So, circular plots are preferred here because our samples do not have any directionality.

(Refer Slide Time: 17:09)



But for instance, our trees in the case of a plantation were arranged like this. So, in this case, we always prefer to take a rectangular sample which is oriented towards to a particular direction. So, for instance, if we took this to be our sample that would give us one reading, but if we suppose oriented it in another direction, it will not be a representative sample in this case, because here all our individuals are arranged in a rectangular fashion. So, it makes much more sense to take a rectangular plot; at the same time, we could also go for a strip.

(Refer Slide Time: 18:03)

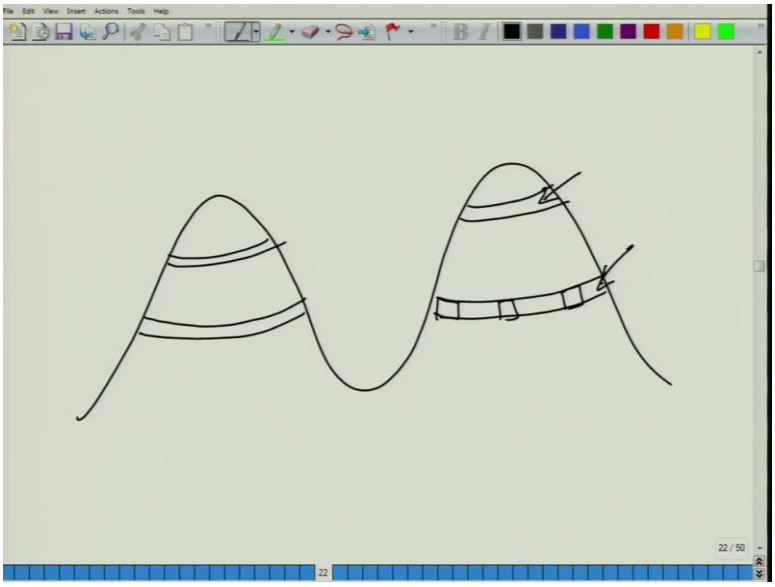


So, in the case of a strip, so these are our planted trees and suppose we want to figure out the survival percentage. So, in this case we can go for a random value. So, we take any number between 1 to 10, so 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. So, we chose any random number suppose we take four. So, once we have decided on that random number we take every tenth row with this number. So, we will first take the 4th row, then the 14th row, then the 24th row, then the 34th row, then the 44th row and so on. So, in this case, so for instance this is the fourth row.

So, we will measure all the units that are there in this row then we will leave nine other rows then we will go to the 14th row, then the 24th row. So, when we are doing that what we are doing essentially is that we are taking ten percent of the individuals that are there in the whole population as our sample. So, our sampling intensity is ten percent now when we are taking our when our individuals are arranged and this strip like fashion if we take such strips it is much better than taking rectangular plot because in the case of a rectangular plot you will have two determine its four coordinates.

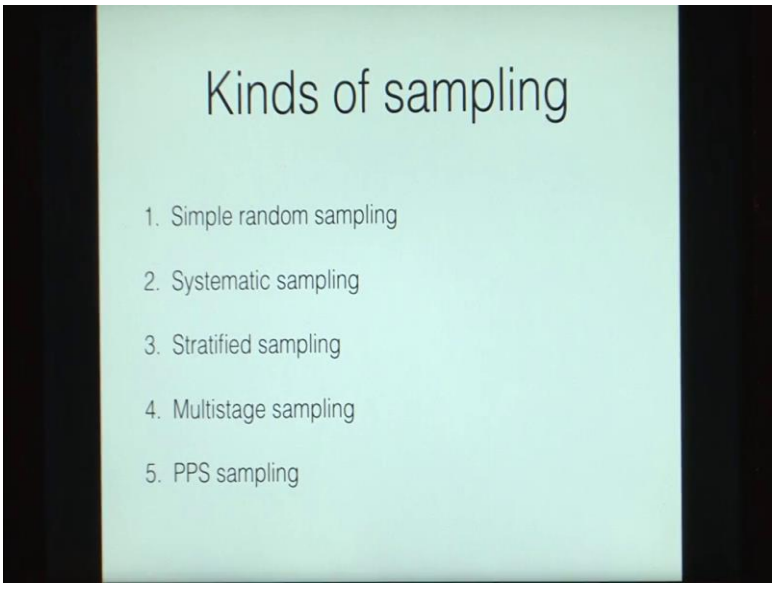
So, these four coordinates will have to be determined and then will may be take a rope all around it and then measure the individuals inside. But when we are going for this strips we can just go directly to the field and measure every ten row and get the parameters for the individuals, how many individuals are alive and how many individuals are dead to get a survival percentage for our whole plantation. So, a strip plot is preferred in such cases.

(Refer Slide Time: 20:12)



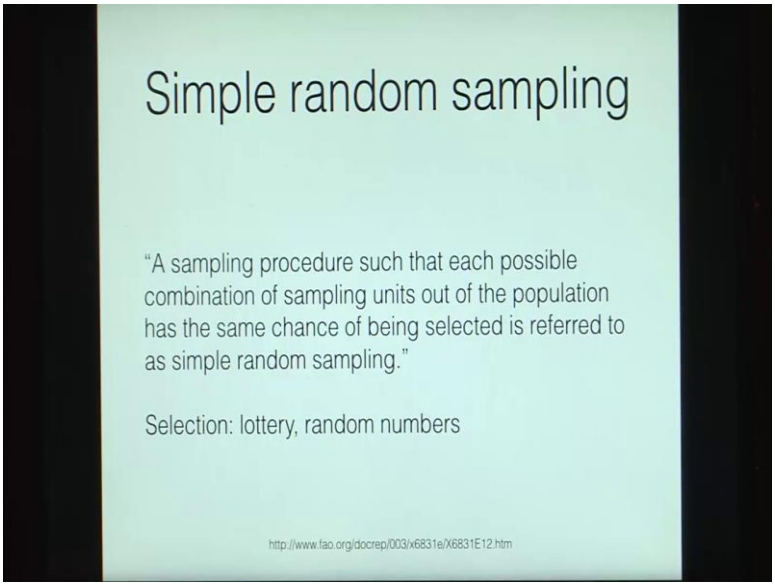
Now, in the case of hills, so suppose we have a hill like this we can go for topographical units. So, when we see topographical units, we can take these strips that are parallel to the ground. So, when we are taking this strip, the values are going to be very different from this strip, but whether we take a value here or a value here or a value here all these values will have similar parameter values. So, a topographical unit or a topographical plot is used in the case of hilly regions.

(Refer Slide Time: 20:50)



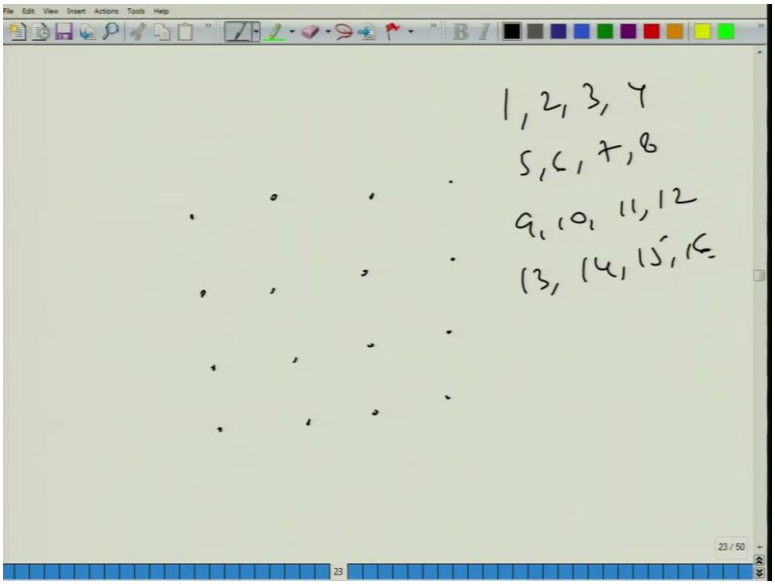
And that being said there are a different kinds of sampling that we can define. We can go for a simple random sampling, a systematic sampling, a stratified sampling, multistage sampling probability proportional to size sampling. So, let us look at these kinds of samplings.

(Refer Slide Time: 21:07)



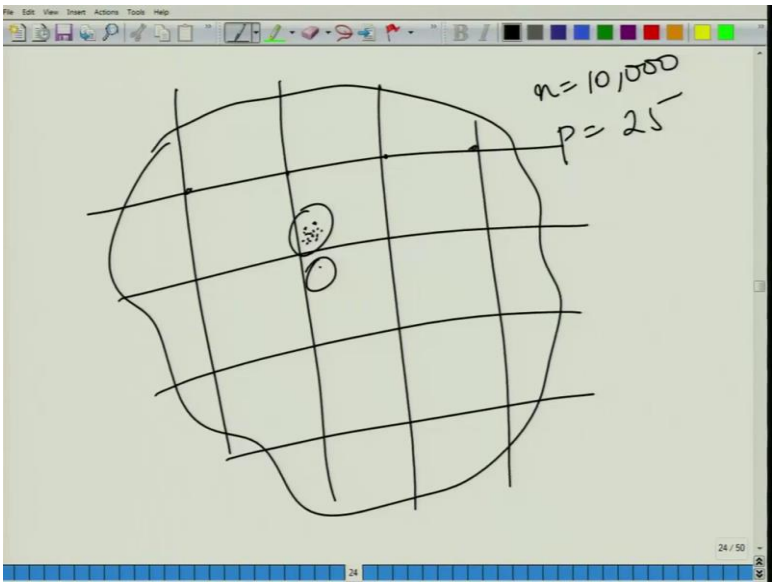
Let us begin with the simple random sampling. Now, a simple random sampling procedure is a procedure such that each possible combination of sampling units out of the population has the same chance of being selected; and this is refer to as a simple random sampling. So, for instance, if we did a selection by lottery or say a selection by random number, it will be called a simple random sampling.

(Refer Slide Time: 21:36)



So, what we are doing in this case is that we are defining all the units. So, in this case let us go for these exchange units 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16. Now, if we took a random number or suppose we put all these sixteen values in the form of chits and we close those chits put them into a basket and the then ask somebody pick those chits from the basket. So, all these individuals have an equal probability of getting into the sample. So, this becomes a simple random sampling. So, to do this it is most important, first to list out all our individuals. Now, in some cases, people might go for a shortcut and a shortcut is something like this.

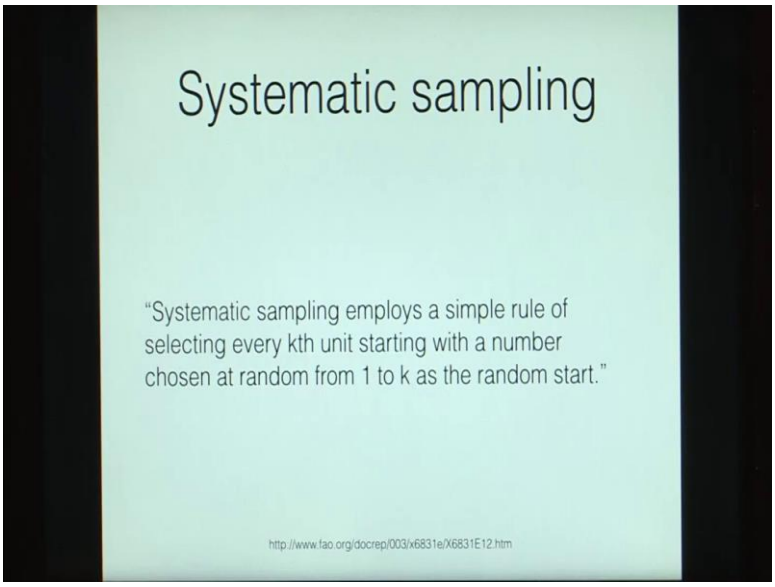
(Refer Slide Time: 22:31)



So, suppose this is our forest and in place of listing out or enumerating all the individuals we just divide this portion into say coordinates then in this case we have so the suppose the number of individuals say 10,000 the number of points in the coordinate. So, these points suppose they are say 25. So, in this case, we can take these coordinate points at random. And then we can take the our tree that is closest to the point as a our random sampling, but such a shortcut is not considered to be a correct procedure, because suppose you have trees that are close together here, and suppose you have a tree that is separated from here.

So, when we are taking this individual or whether we are taking this individual or this individual, so all of these will be closed together to this point. So, the probability of a point from this cohort to be selected into our sample is much greater than that of an individual tree. So, this will not be considered a simple random sampling. So, in the case of a simple random sampling, you first need to list out all your values and then you need to draw a sample out of it either by using a lottery system or by taking random number.

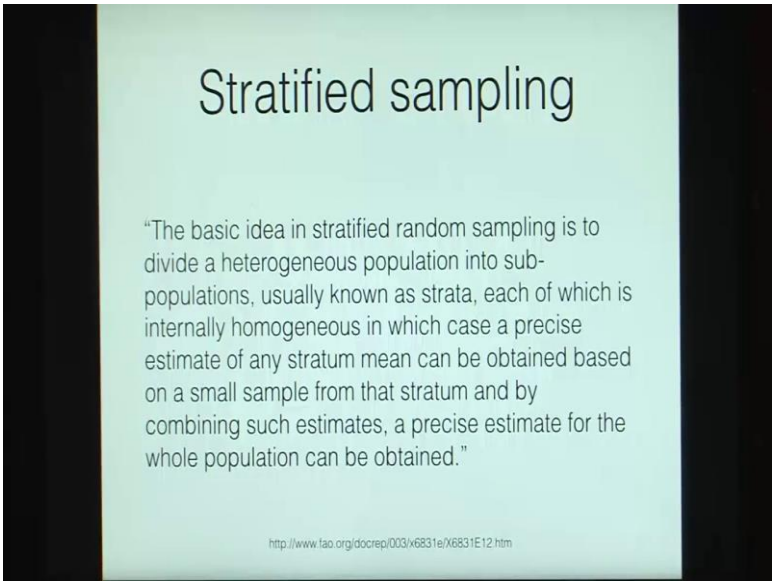
(Refer Slide Time: 24:04)



Now, the next sampling procedure is known as a systematic sampling. So, coming to the slides. So, systematic sampling employees a simple rule of selecting every kth unit starting with the number chosen at random from 1 to k as the random start. So, when we said that we are going to select every second value or every third value or say in the case of measuring the survival percentage, we took a the first value as any random number we

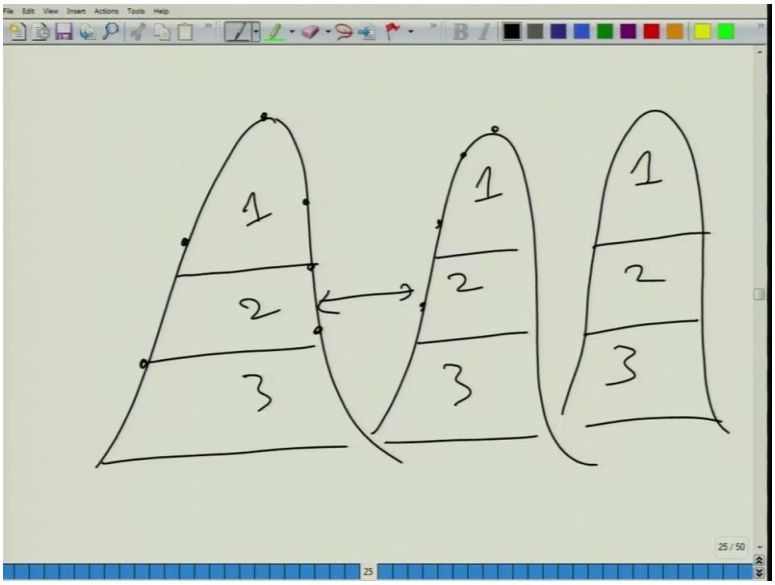
did 1 to 10, and we said that we are going to measure every tenth strip. So, that sort of a sampling is called a systematic sampling.

(Refer Slide Time: 24:41)



Another sampling method is called a stratified sampling. So, in the case of stratified sampling, the basic idea is to divide a heterogeneous population into sub populations know as strata. Each of which is internally homogeneous in which case a precise estimate of any stratum mean can be obtained based on a small sample from that stratum. And by combining such estimates a resized estimate for the whole population can be obtained. So, what are we doing in the case of a stratified sampling.

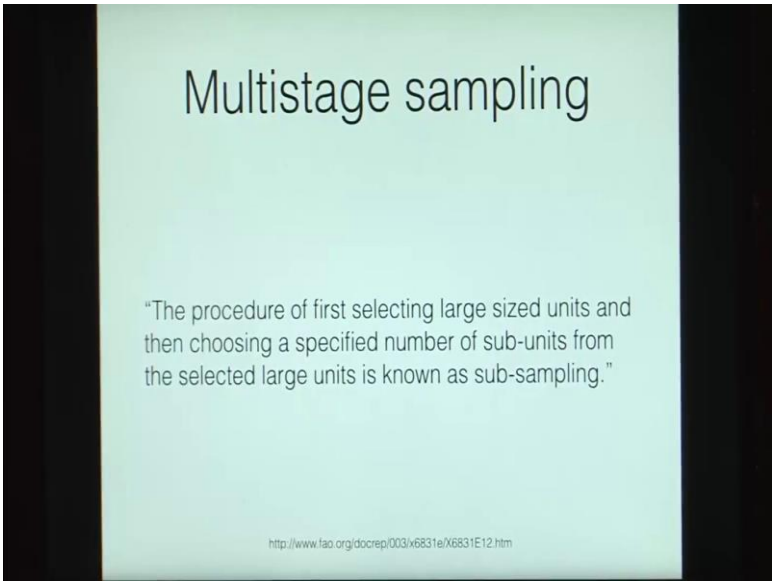
(Refer Slide Time: 25:14)



Suppose you have a hilly region, and you want to find out the average height of a point. So, you have this hilly country and you want to find out its main height from the sea level. So, one method could be to go to take all different sampling units that can be there and then to define and to go for your simple random sampling, but there can be a better way. So, a better way could be to divided into strips like this. So, all these values with 1 are at a similar height, all the values at 2 are at a similar height, all the values that are 3 are at a similar height.

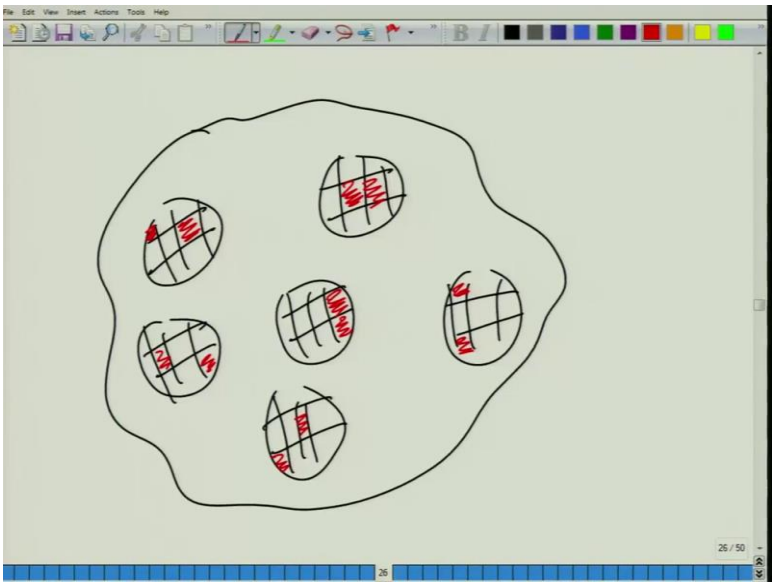
So, now what we could do is to take the average height of a sea level for all the two values and multiply those with the area that is under this height of 2. So, we have divided our population into a number of strata we are taking a main height for each strata and then we are using the size of the strata to get an average height over the main sea level. So, this type of a sampling is called as stratified sampling. So, for instance in the case of our forest, we can take the values of the trees that are close to the rivers the values for the trees that are close towards the edges of the forest and for those trees that are towards the center of the forest. And then we can use these area the areas along the strips of river, the areas towards the edges and the central areas, and use these as multiplication factors for the main heights for all these three strata, so that will be called a stratified sampling.

(Refer Slide Time: 27:05)



The next sort of a sampling is called a multistage sampling. So, in the case of a multi stage sampling we first select large size units and then chose specified numbers of sub units from the selected large units to go for a sub sampling.

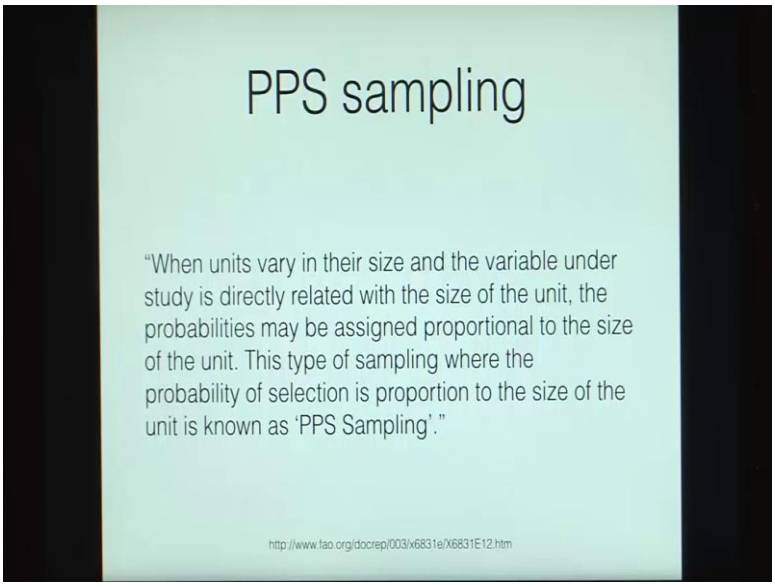
(Refer Slide Time: 27:25)



So, what we are doing in this case is that in our large forest area, we are first taking large size samples. So, suppose these are our samples then out of all these samples, we are taking subsamples that are called subsamples. So, suppose we say that we are going to have two subsamples from each of these. So, suppose in the first one, it is these two; and the second one, these are together; and third one, these are like this; and the fourth one, it

is like this. So, this sort of a sampling in which we first define our sub populations and our sub units and then our large units and then we define or get our to our sub units is known as multistage sampling.

(Refer Slide Time: 28:16)



Another way of sampling is the probability proportional to size sampling or the PPS sampling; so coming to the slides now. When units vary in their size and the variable under study is directly related with the size of the unit, the probabilities may be a assigned proportional to the size of the unit; this type of sampling where the probability of selection is proportional to the size of the unit as known as PPS sampling. So, essentially PPS sampling is what we saw in the case of point sampling. So, in the case of point sampling say with a wedge prism the probability of selecting a larger size tree in our sample is greater than the probability of selecting a smaller sized tree in the sample, so that sort of a sampling is known as a PPS sampling or the probability proportional to size sampling.

So, in this lecture we saw different kinds of sampling and what sampling actually means. What do we mean by a population a sample, a sampling unit, a sampling frame and so on? So, sampling can be utilized or is utilized in the case of forestry because our population are very large and we need to get our data quickly and easily and with ease, so that is all about sampling for today.

Thank you for your attention. [FL].