Essentials of Data Science with R Software – 1
Professor. Shalabh
Department of Mathematics & Statistics
Indian Institute of Technology Kanpur
Lecture No. 53
Distribution of Sample Mean, Convergence in Probability and Week Law of Large Numbers

Hello friends welcome to the course, Essentials of Data Science with R Software – 1 in which we are trying to understand the basic concepts of probability theory and statistical inference. And now, we have nearly completed the probability part. Well, that was an overlapping part also with the statistical inference, but now we are gradually moving towards the statistical inference.

Now, let me ask you one very simple question. Suppose, somebody ask you that, how much time are you going to take from your home to college? You always tell some numerical value, say 20 minutes. My question is, take a pause and think how you obtain this number 20. Now, if you ask me, I would say, well, you are going to the college almost every day and then you try to take the observations on the time taken that mean you try to record every day for a couple of days, that time taken from your home to college.

And after that, what you have to do? Take a pause, and then think what are you doing? If you ask me, I will say, you are trying to find out the arithmetic mean. You simply take all the observations, and you try to find out their arithmetic mean. And then whatever is the value of arithmetic mean, you inform me as 20 minutes. I have a question for you. Who told you to compute the arithmetic mean? Or how you took a decision that you have to compute the arithmetic mean? Why not geometric mean, harmonic mean, median mode, etc.?

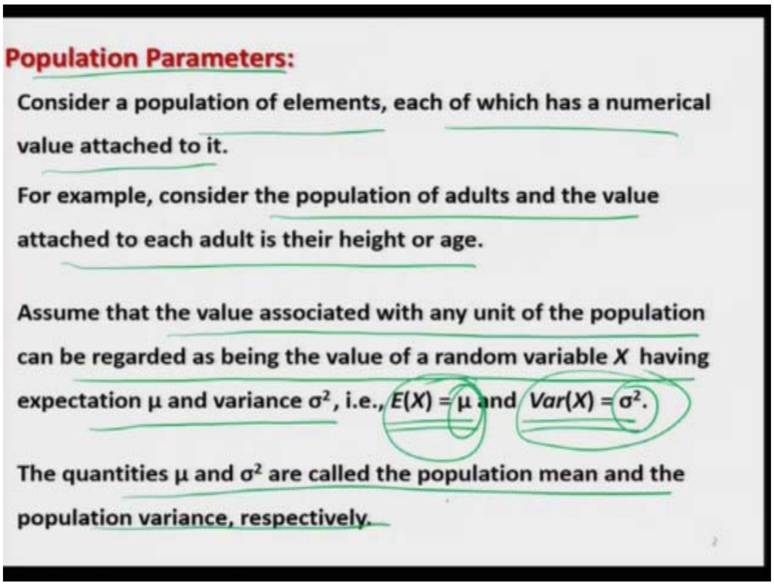
And I am not countering your decision. Your decision is good, and I do respect it. And that is a correct decision also. But my question in this course, is that we are here to find out why, how, where, how, means how we are going to do to conclude this decision or how we are reaching to this decision. Now, these are the questions, which we are going to answer in this lecture.

And then the next question is, you just took the sample mean of say 10 to 20 days, and you told me 20 minutes. But possibly, if you ask me, I will say okay, every student who is staying near about your home, they are taking almost 20 minutes to reach to the college. Is that correct? You always take it correct. But now I am talking of the all population who are staying in the neighborhood of your house.

So, what are I am going to do, with your one value that you have computed only on the basis of a small sample, I am trying to take a call which is valid for the entire population. Population of students who are going to your college from your neighborhood. And you have seen that these things are correct. These things hold true, but my question is that how and why? What is the statistical reasoning behind it? What is the scientific thought process behind it, which make us convinced that these values are correct?

And I always feel that we all are statistician. God has given us a component of a statistician that is why we are doing all these calculations very fast. The only thing is this, we do not know what are we doing and that is what we are doing in this course. So, in this lecture, I am going to talk about some results about the sample mean is convergence, and one more topic about weak law of large numbers. So, all these smaller topics we will try to answer the question which I have raised now.

(Refer Slide Time: 04:28)



So, let us begin our lecture. So, one concept is this, what population parameters? Well, that is a very small thing, very small concept, but it is very important for us to understand. Suppose we consider a population of some elements, each of which has a numerical value attached to it. For example, if you say consider the population of adults, and the value attached to each of the adult is their height or their age. So, population is simply the collection of all such adults.

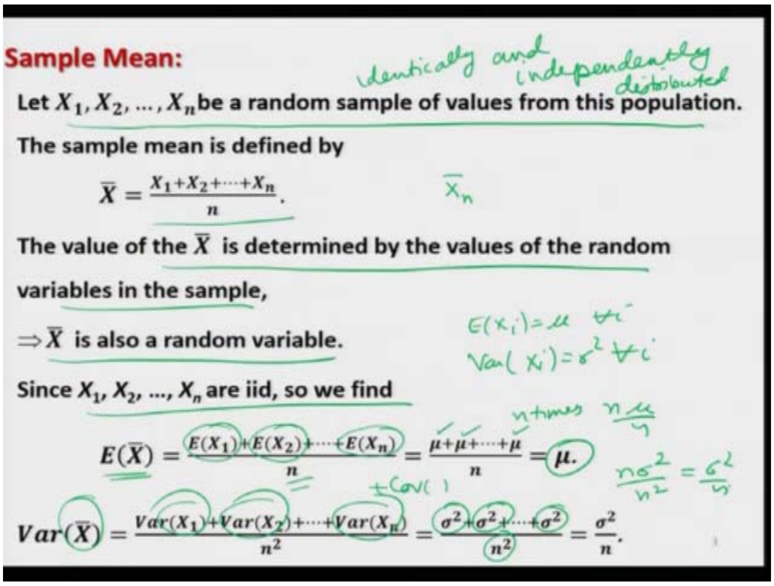
For example, if I say the population of students who are staying in your neighborhood. You can define that okay anybody who is staying within the radius of 100 meters from your home and they are going to the same college where you are going they are trying to create a population of students, and there is a value attached to every student that how much time they are going to take from going from their home to the college. This time maybe 20 minutes, 19 minutes, 21 minutes, etc..

So, what we can now assume that the value associated with any unit of the population can be regarded as being the value of the random variable X. For example, I can say that let there be a random variable X, which is the time taken by the students. Now, every student will have some value 20 minutes, 18 minutes, 22 minutes and so on. So, I can simply assume that this value is the value of our random variable, and this random variable is assumed to have mean mu, that is expectation is μ, and variance is σ^2 that is expected value of X is μ and variance of X is σ^2 that is a very simple basic assumption which I am trying to make here.

Now, if you try to see what are this μ and σ^2, they are the values which are present in the population that means if you try to take the entire population and then you try to find out their arithmetic mean or their mean value, this will come out to be same as here μ. And if you try to take the entire population try to find out their variance, then it will come out to be like σ^2. But the fact is this, both these values are not known to us that is a different aspect, but these quantities μ and σ^2, they are called as a population mean and population variance respectively.

So, they are called as population parameters; μ and σ^2 here are the population parameter of which population the population which is associated with the random variable X. And, that this population can also be characterized by some probability function also. Well, that I will try to show you, but I am just trying to give you here a hint.

(Refer Slide Time: 07:18)



So, now, we have understood the very simple concept of the population parameters. Now, we will try to understand what is sample mean. Sample mean that you know, that is simply the automatic mean. So, let X_1, X_2,…, X_n be a random sample of values from this population mean we already have assumed a random variable X. So, these X_1, X_2,…, X_n they are observed from this population, and they are random samples. Random sample means, they are identically and independently distributed.

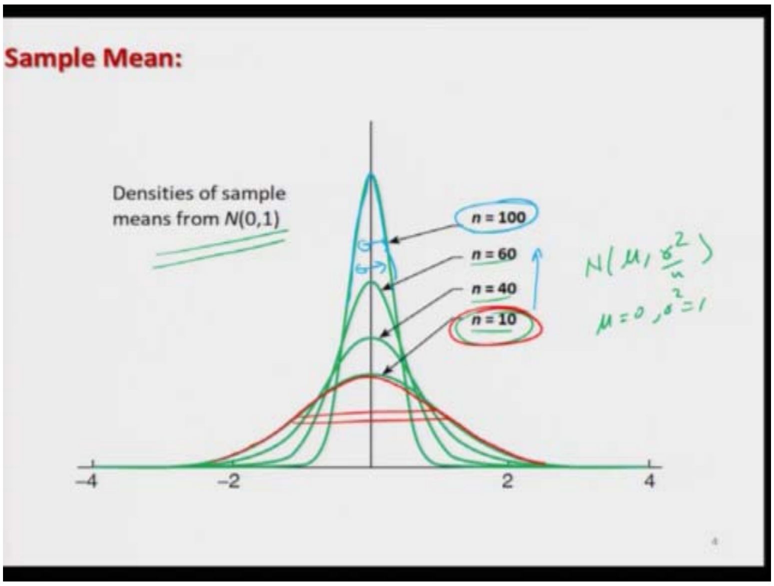
So, now, we define the sample mean, as $\bar{X}$ is equal to X_1 + X_2 plus… X_n divided by n. After some time I will try to attach here a subscript like here $\bar{X}$ here and so, that will be indicating the sample size, but at this moment, I am not introducing it here, because I just want to give you the basic concepts and basic idea first. So, now, the value of this $\bar{X}$ is determined by the values of the random variable in the sample.

So, obviously, means $\bar{X}$ is a function of a random variables, so, this $\bar{X}$ itself is a random variable. So, now, since X_1, X_2,…, X_n are iid, so we can find out the mean and variance of this sample mean $\bar{X}$. So, you can see here that expected value of $\bar{X}$ will come out to be expected value of X_1 + respectively of X_2 up to expected value of X_n divided by n. So, now, you already have assumed that expected value of X_i equal to μ for all i and you also have assumed that variance of X_i is equal to σ^2 for all i.

So, now, if you try to write down here this will become μ + μ +μ + μ … n number of time so, this will become here and μ upon n which is equal to here μ. So, you can see here that the expected value of $\bar{X}$ is also here μ. Now, if you try to find out here the variance of $\bar{X}$, so, that will become here variance of X_1 + variance of X_2 + variance of X_n + there will be some covariance term. But since you are assuming that $X_1, X_2, \ldots, X_n$ are independent, so, that covariance term will become 0 and we will have here only the variances σ^2, plus σ^2 and times, that will become here n σ^2.

And then, this here n^2 this will be here like this and you can write down here n^2 upon σ^2 that is equal to a σ square upon n. So, now you can see here that you have the observations $X_1, X_2, \ldots, X_n$ from there you obtain the sample mean $\bar{X}$, $\bar{X}$ is also a random variable because that is a statistic, and $\bar{X}$ also has got the same mean as μ as of the original population and variance of $\bar{X}$ has the variance σ square upon n.

(Refer Slide Time: 10:33)

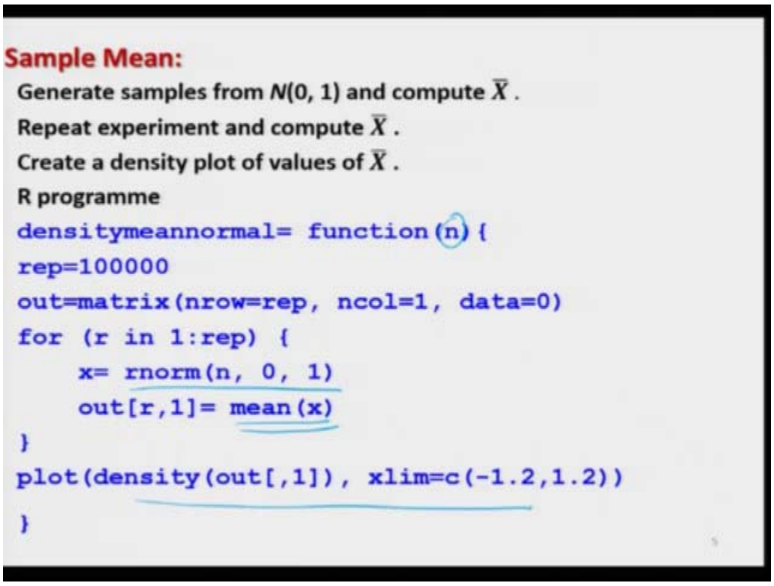


And you can see that it is not difficult to obtain it provided you know the basic fundamentals. So, now, what I try to do here that I try to generate say, a large number of samples from the normal 0, 1 population. Now, you cannot ask me how to generate the random observation from normal 0, 1. And then, for every sample, I tried to compute the sample mean. So in case if I suppose to generate 5,000 samples and can I try to compute their sample means and then I try to plot the distribution. So, this job I am going to do now, in the forthcoming slides.

So, in case if you try to simply plot the curve of normal see here μ and σ^2/n, then for $\mu=0$ and $\sigma^2=1$ just by choosing n equal to 10, 40, 60 100 you can obtain this type of hair curve. Now you try to observe what is really happening. And that is going to answer very interesting questions, which I was raising right from the beginning in the course, but I was just waiting for this slide to come.

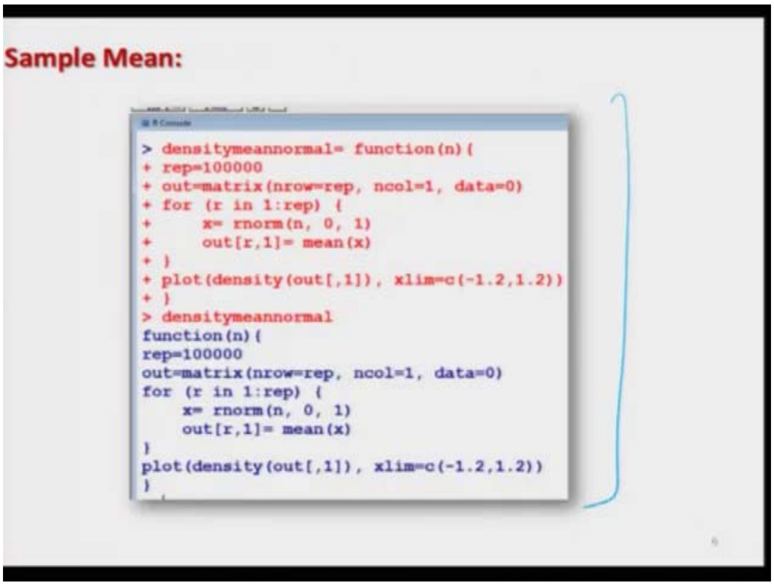
You can see here, when n is equal to here 10. Let me change the color of my pen you can see here. When here n is here 10, then your curve here is like this one. And you can see here that the variance here is very high. But in case if you try to take your n equal to 100, the curve here is this one and here the variation is much smaller. And in case if you try to increase the value of n from 10, 40, 60, 100, this variance is decreasing. And the fact you can observe that as the sample size is increasing, the variance σ square upon n will also be decreasing, but that you can see graphically also.

(Refer Slide Time: 12:26)



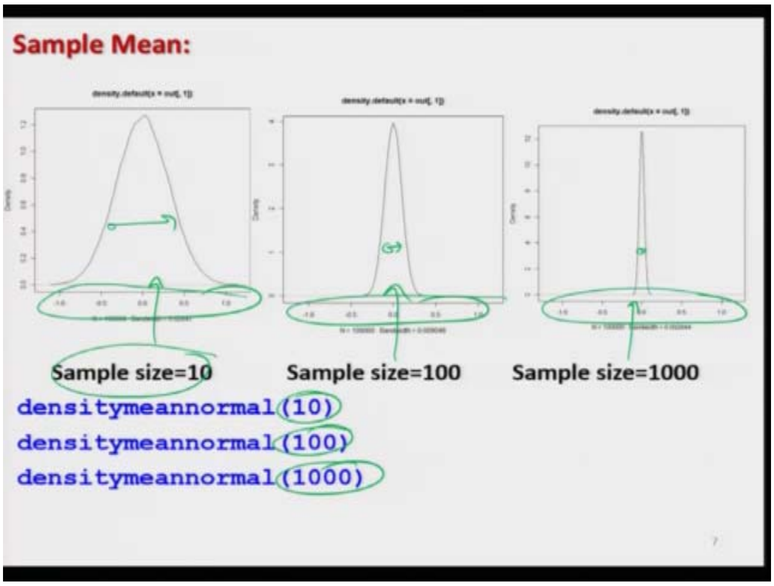
Now, let me try to do the same thing. For that, I have written here a very simple say this program in which I am trying to generate a large number of say here samples of a given size here n from the normal distribution with mean 0 and variance 1, you can see there. I have used here r norm. And then whatever is their outcome that I am trying to store as a mean of their values, and then I am trying to plot the density using this command.

(Refer Slide Time: 12:59)



So, if you try to first understand. This is a screenshot of the same program, I will try to show you on the R console.

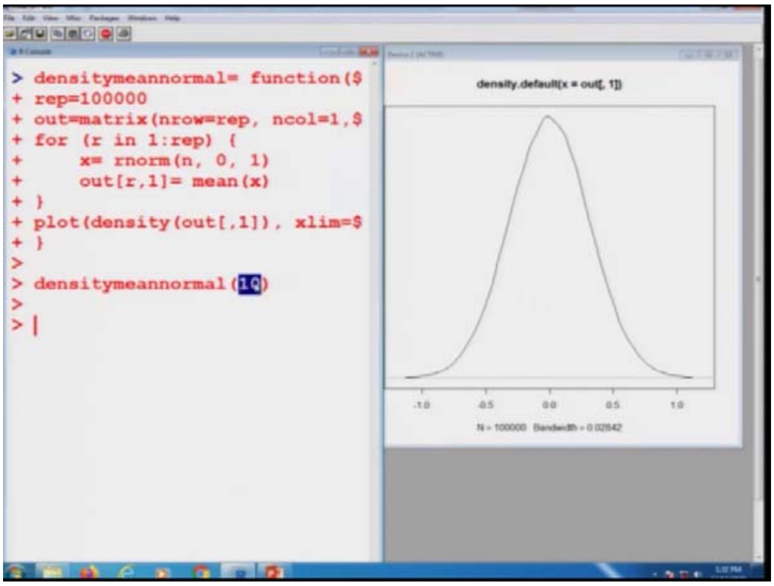
(Refer Slide Time: 13:03)

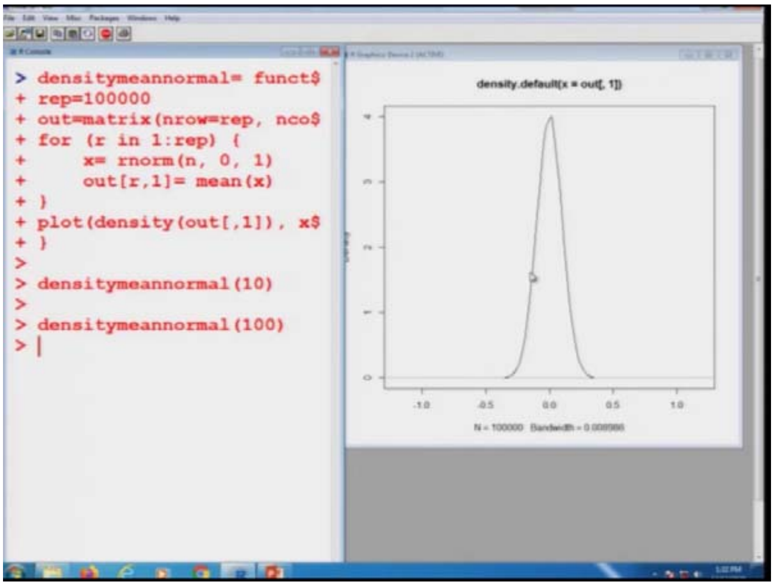


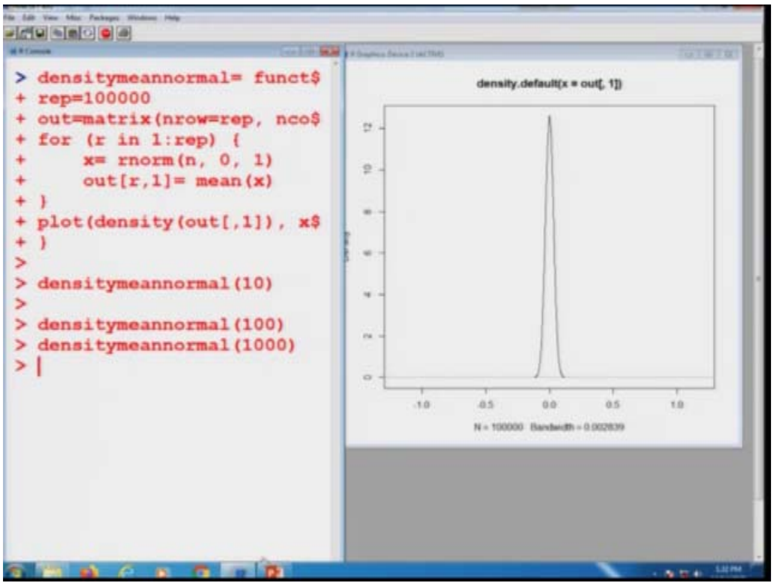
But you are going to get here this type of curve, which I have just shown you here. For example, if I try to take only the sample size to be here 10 then this curve is for here 10. And if I try to increase it to 100 or say 1000 these curves become here like this one and this one. This you can obtain in your R software also. I will show you. So, you can see here now, look for this width.

And remember one thing I have taken the value, the range on the X axis to be the same in all the figures, so that you can easily compare the widths of these curves around the mean value. So, you can see here that as n is increasing, this curve is becoming more narrower, or its variability is decreasing. So, let me try to first show you this thing on the R console, so that you are also confident that whatever I have shown you that is correct. And I will show you that whether this curve can be generated or not. So, I have copied here the program and then I try to execute this program. I copy the command to save some time.

(Refer Slide Time: 14:12)





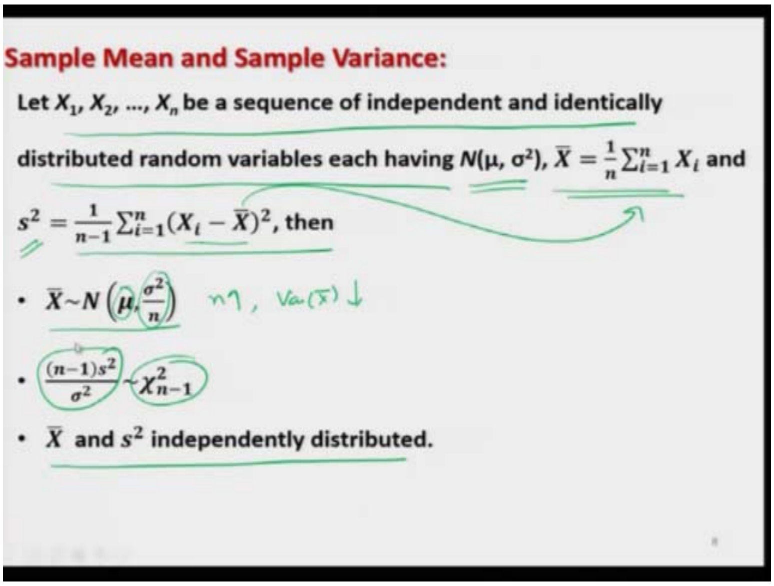


And you can see here. As soon as you give here you get here this type of curve for n equal to 10. Here in this 10, this is the sample size. And now you can see the same curve I will try to show you that if you try to make it here 100 what happens you can see here after getting 10,000 samples, it is becoming like this. And if you try to make the sample size to be here 1,000 then you can see here before your eyes that what is really happening. It is trying to generate 1,000 observations 10,000 time and then it is trying to plot this curve. You can see here now this has become very narrow.

So, this is exactly what I was trying to explain you. And you also know that if you have more information more number of observations your conclusion become better. For example, if you are going from your home to your college only 4 times, then you tell me that time and if you are trying to go for say this 100 times, and then tell me that time, you can very well understand what is the difference between the two conclusions.

In case if the sample size is more that means, if you tell me the time based on the sample mean, which is based on say 100 observation that is going to be more precise then only the sample mean based on 3 or 4 or 5 observations.

(Refer Slide Time: 15:38)



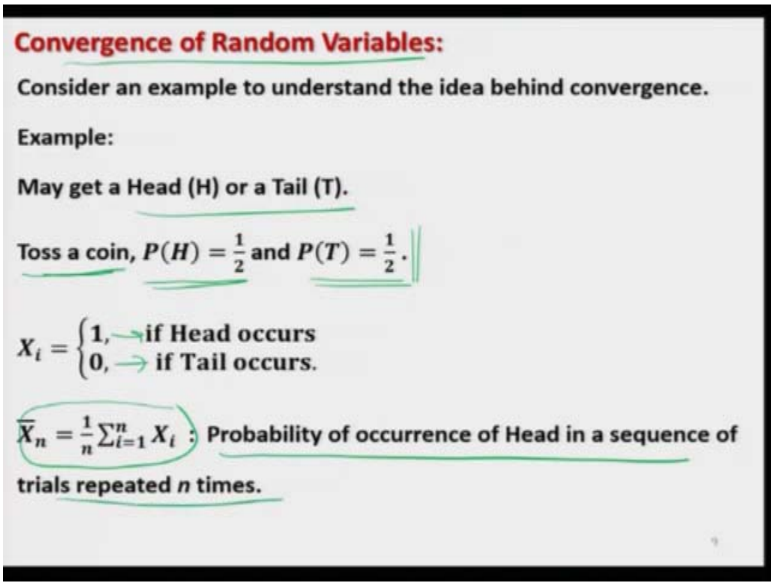
So, this is what I wanted to show you here, what happens. And if you try to remember, whenever I was trying to deal with the probability mass functions and probability density functions also, I was trying to show you towards the end the difference between the theoretical mean and the mean which was coming from those distribution. And now, you can recall that when I was trying to increase the sample size, the sample mean was getting closer to the theoretical mean.

So, now, based on that, let me try to give you here one very important result, which is related to the sample mean and sample variance that is let $X_1, X_2, \ldots, X_n$ be a sequence of identically and independently distributed random variables each having a $N(\mu, \sigma^2)$ distribution, then the sample mean $\bar{X}$ and the sample variance is $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ where $\bar{X}$ is obtained here.

Then we have here 3 important results. The first one, I already have shown you that is $\bar{X}$ follows a normal distribution with mean μ and variance σ^2/n. So, that is also clear. You can see here that if your n is increasing, then the variance of $\bar{X}$ will be decreasing. And that is the same thing that was reflected in the graphics also. The second result is that $(n - 1)s^2/\sigma^2$, this will follow a chi square distribution with n - 1 degrees of freedom. Now, you understand what is chi square. I do not need to explain you.

And the third very important result is that $\bar{X}_n$, s^2 are independently distributed. Now, I do not think if I have to explain you what is the meaning of independently distributed random variable because $\bar{X}$ as well as a small s^2, both are random variables, both are a statistic, they are a function of random variables. So, what I am trying to say here, that if you try to work under such a setup, then this $\overline{X}$ and s^2 will be stochastically independent of each other.

(Refer Slide Time: 17:48)



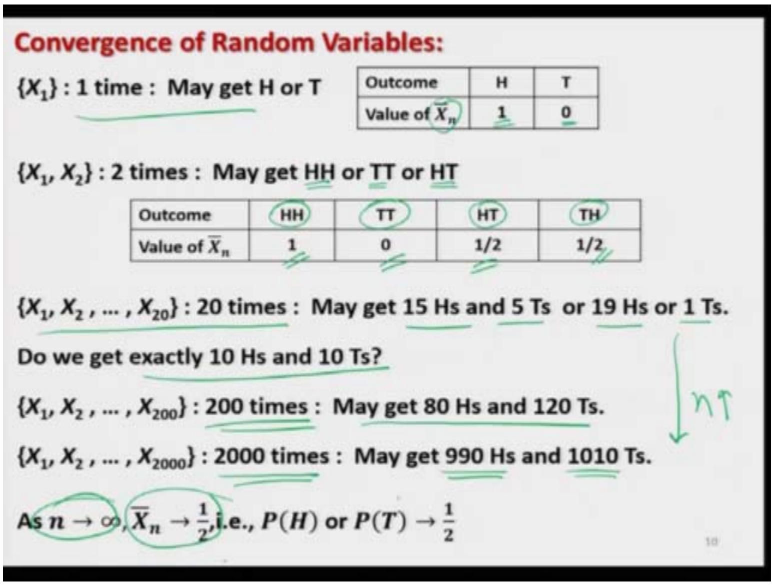
And then with that you can prove by any of the results using PDF that the joint PDF can be expressed as the product of the marginal PDF or CDF or expectation. Whatever you want to use, you can prove this result without any problem.

So now, I try to come to one more aspect that is about the convergence of random variable. Now, let me try to give you an example to explain you what does this really mean? Suppose you toss a coin and then you will get here either a head or a tail. And you always said that the probability of head is equal to 1/2 and probability of tail is equal to 1/2 . So, you can recall that in the beginning of the course, I have taken this example. Means I had a taken and then I had explained you what is the meaning of this 1/2 .

And then we have conducted some experiment also on the R console using the command sample. And we were convinced that if you are trying to repeat the experiment for a large number of time then this probability 1/2 , will be achieved as n goes to infinity. So, I am not going to repeat it, but now, I want to explain you the same thing through a different concept. And that will help you in understanding what is called convergence in probability and it will help you in other type of statistics what you are trying to do.

So, now, you can see here. In case if you want to find out here the this $\overline{X}_n$ that means the value of the sample mean based on n observation or say n number of tosses, where X_i takes value 1 if there comes head and takes value 0 if there comes tail. So, this $\bar{X}_n$ is simply going to indicate the probability of occurrence of head in a sequence of trials repeated n times. And you can recall that this was the same thing I had shown you as a relative frequency also earlier.

(Refer Slide Time: 19:39)

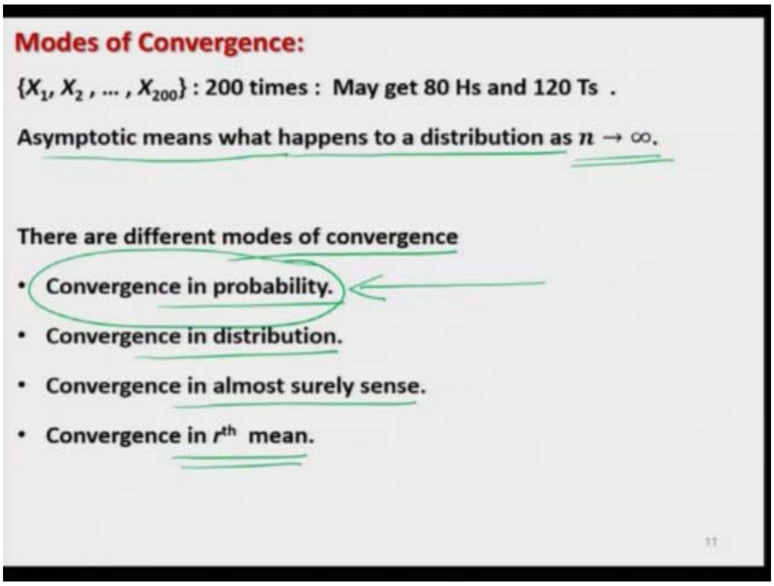


So now, what will happen that if you toss it once, you may get head you may get tail, and in both the cases the value of $\bar{X}_n$ will come out to be here. See here, 1 and here 0. And if you try to toss the coin two times, then your outcome maybe 2 heads 2 tail or 1 head, 1 tail. So, in this case, what will happen? The value of $\bar{X}$ and that is the mean will come out to be here 1, 0, 1/2 or say 1/2 . Depending on whether you are getting 2 heads, 2 tails, 1 head and 1 tail like this.

So, now, if you try to repeat this experiment 20 times and if you try to toss the coin 20 times what will happen? You may get 15 heads, 5 tails or you may also get 19 heads 1 tail or they can be any combination. I have taken just these two value to explain you. But my question is, do we exactly get 10 heads and 10 tails? The probability is extremely low. And if you try to repeat the experiment 200 times, then you may get 80 heads and 120 tails or there can be any combination.

And if you try to repeat it for 2,000 times possibly you may get 990 heads and 1,010 days. But if you try to see what is really happening, if you are trying to increase the number of tosses that means you are trying to increase the value of here n. The number of observations are increasing. And every time you are trying to find out the sample mean. So, you can see here that as n goes to infinity, the value of $\bar{X}_n$ goes towards 1/2 . And that is why we say that the probability of head or probability of tail is tending to 1/2 and that is the same thing what we have learned earlier also.

(Refer Slide Time: 21:29)

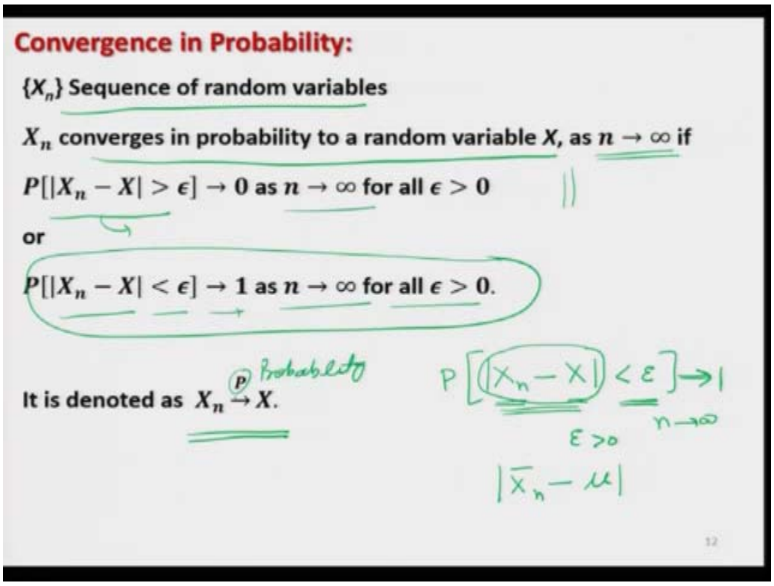


So, now, this means what? When you are trying to increase the number of observation or you are trying to increase the n, we say that these are asymptotic results. And asymptotic means what happens to a distribution as n goes to infinity. Continuously trying to increase the number of times the experiment is conducted. For example, in this case, you are trying to increase the number of tosses of the coin. So, this concept can be very well explained by the concept of modes of convergence.

So in statistics, we have mainly four modes of convergence that we try to study. One is convergence in probability, convergence in distribution, convergence, in almost surely sense and convergence in r^{th} mean. But definitely, I am not going to explain you here the convergence, but I am simply going to understand what is the use and what is the definition of this converges in probability? Because that is the concept, which we are going to use at least in our course.

But definitely, if you want to learn more beyond this course, you will need to understand these things from the book. But now, there should not be any confusion that you will have any problem in understanding the theory or trying to or there should not be any problem in convincing to yourself that why theory is needed. Means, whatever was happening in the beginning, that I have shown you that they will certainly hold true with the theory.

(Refer Slide Time: 23:06)



So now, let us try to understand the convergence in probability. What does this mean? So, suppose this X_n is a sequence of random variable, and then we say that X_n convergence in probability to a random variable s as n goes to infinity, if this condition holds true. Now, once you try to look at this condition, in the first shot, it looks very difficult, what is this? Now, let me try to explain you.

So now, if you try to see you have a random variable here, X, and you are trying to know its value based on some other random variable, which is depending on the sample size X_n. And what do you want? You ideally want that the difference between the two statistics should be ideally equal to 0. But, I mean, statistically speaking, that is difficult to achieve. So, what we try to say that, suppose the difference between X_n - X is quite small. So, now we try to say that suppose this is less than some quantity ε where ε is greater than 0.

Now, the difference between X_n - X can be positive or can be negative. So, to be on the safe side, I try to take here an absolute value. And now, what are you trying to see that as n is increasing the value of this is statistics, and its difference with this thing X that is getting smaller and smaller, which is less than epsilon.

And the probability of such an event, this is going towards 1, that means the probability is increasing means you have seen you are trying to compute here the population mean by $\overline{X}_n$.

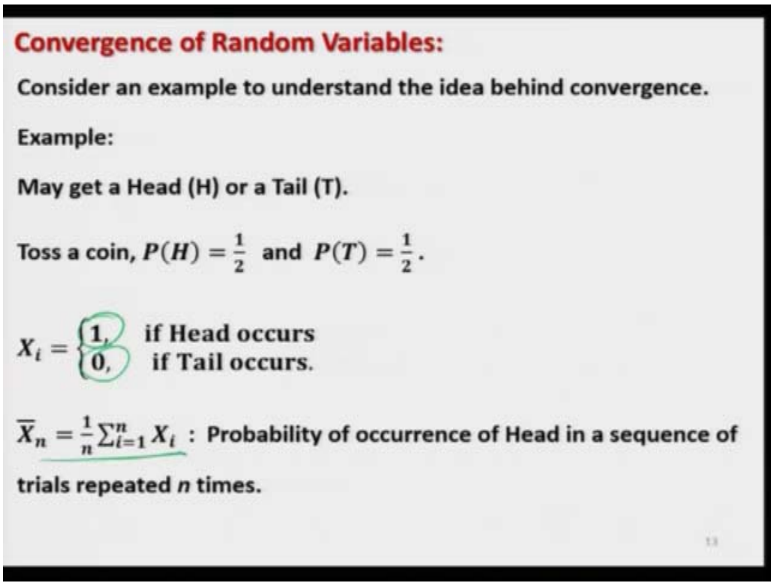
And as you were trying to increase the value of here n, the difference between $\bar{X}_n$ and μ that was decreasing. That means the values were getting close to each other. That is what I had shown you many times during the probability density functions and probability mass function.

So, now, if you try to see you are trying to say that the probability that the difference between X_n - X in the absolute sense is smaller than ε and the probability of such an event is tending towards 1 that mean this event is going to occur with higher probability. And this is going to happen as n goes to infinity. And this is what exactly I have written here if you try to see that probability that mod of X and - X is less than ε tends to 1 as n goes to infinity for all ε greater than 0.

So, this is not difficult. The only thing is this whenever you are trying to see or look at such an expression try to divide it into small component and then try to understand it what the expression is trying to explain you. And the same condition can also be written here as a probability that absolute value of X_n and - X is greater than ε it is going to 0 as n goes to infinity.

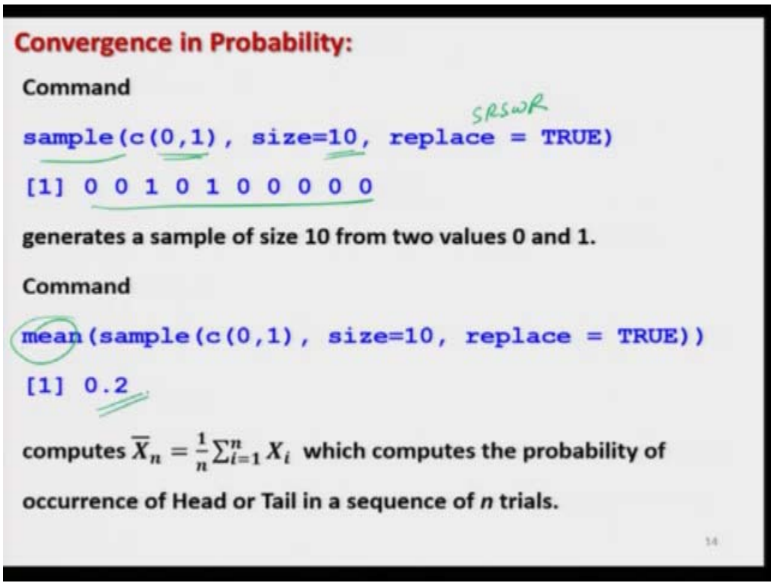
Because here now, you are trying to say that the difference between X_n and X is becoming greater and greater that means, they are becoming, means quite far away from the true value. So, the probability of such an event is going to 0 as n goes to infinity. So, this is actually the concept of convergence in probability and this is indicated like this X_n converges to X_n probability and this P is indicating the probability which is written over an arrow right arrow.

(Refer Slide Time: 26:33)



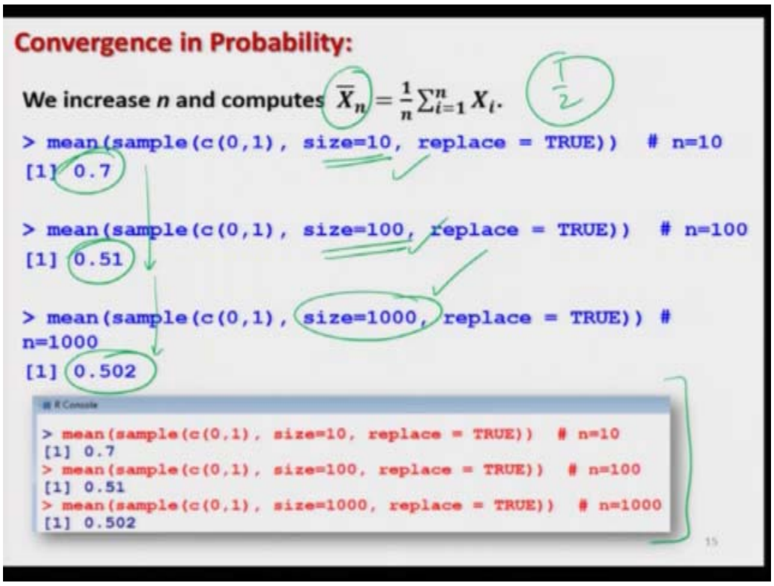
So, just try to take an example to understand the idea behind this convergence. So, I try to take the same example which I just considered. So, we are trying to toss a coin you may get a head with you may get a tail, and if it comes head the value is indicated by 1 otherwise 0 and we try to compute the value of $\overline{X}_n$.

(Refer Slide Time: 26:55)



And we try to conduct this experiment in the R console. Although, we have done it earlier so I will not to repeat it in the R console, but you can do it yourself. So, I try to obtain here a sample between the two values 0 and 1. And I try to obtain here n number of values and by SRSWR that is simple random sampling with replacement, and you get here these values. And from these values, you try to obtain the mean. So, this mean will come out to be here to 0.2.

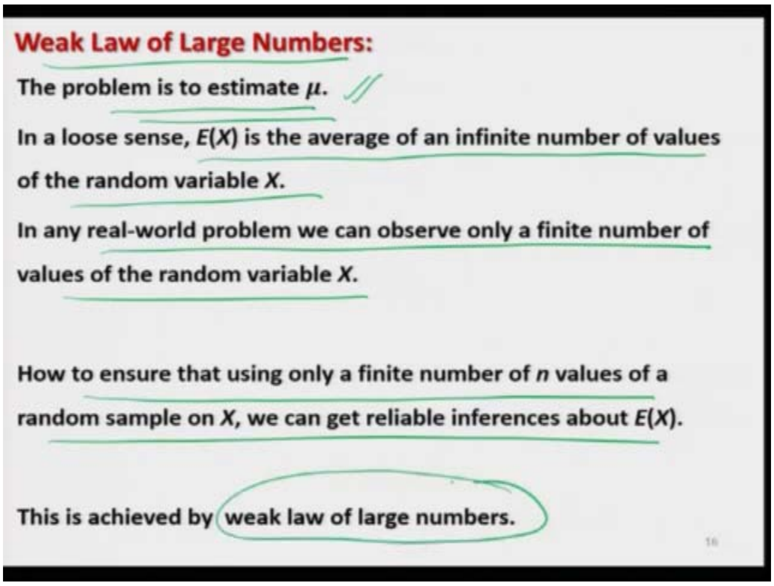
(Refer Slide Time: 27:30)



Now, I try to increase this n and I try to compute this value. So, when I try to compute it for size 10 this comes out to be 0.7, when I try to increase it to 100, this value becomes 0.51, when I try to increase it to 1,000 that means, there are 1,000 observation on which I am trying to find out the arithmetic mean of 0s and 1s, which are randomly generated, this is coming out to be close to 0.502. So, what you can observe here, you are simply trying to find out the value of the sample mean, as n is increasing from 10 to 100 to 1000.

And you can see here the true value here is 1/2 , 0.5. And mean of these value is converging to 0.5 as n is going to infinity from 0.7 to 0.51 from 0.51 to 0.50. And you can believe on me because here is the screenshot of the same observation when I conducted this experiment when I was trying to prepare the slides. Now, if you try to do it possibly you will get a different values, but the pattern is going to be the same.

(Refer Slide Time: 28:34)



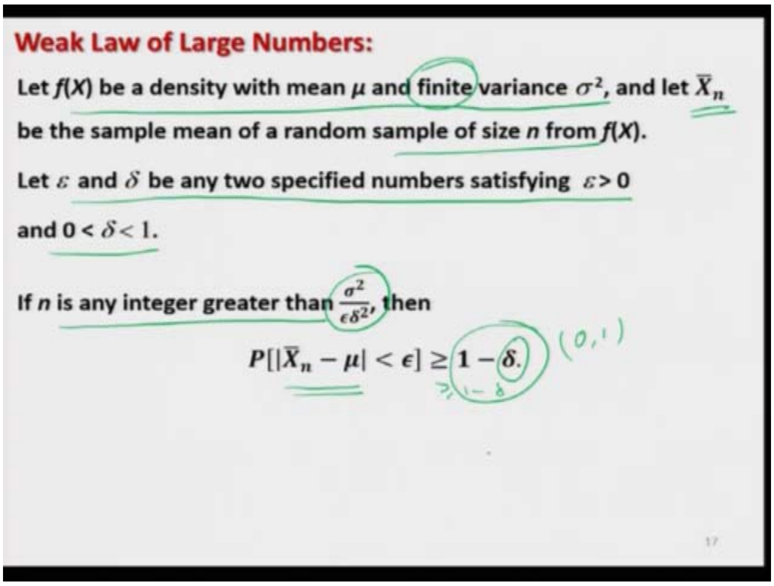
So, now, after giving you this idea of conversion and probability we try to address the problem, which I really wanted to address. And this result is called as weak law of large number. So suppose, we want to estimate the value of μ. Why? Because μ is the population mean that you do not know. This is the population parameter. You do not know that what is the average income of the people in this country, because if you want to know it, you have to go to each and every citizen of this country.

You have to ask for the monthly income and then you have to find out the arithmetic mean of all such observation that you cannot do it. So, is it when the statistics what are we really going to do? We are going to take a small sample and then we try to compute the sample mean. Well, that is a different thing that how the sample mean came into picture and why not other thing, so that I will try to address in the forthcoming lectures.

So, the problem here is now to estimate the population parameter μ, which is mean. So, in some loose sense, we can say that expected value of X is the average of an infinite number of values of the random variable X. But the problem is that in a real world problem, we can observe only a finite number of values of this random variable X. So, the question is, how to ensure that using only a finite number of a small n number of values of a random sample on X we can get a reliable inferences about expected value of X.

There the same question that if you try to go from your home to your college say in finite number of times and then you try to find out the value of the average time that is going to give me the correct value, but then you are trying to estimate it only on the basis of say 5 trips 10 trips to 20 trips and so on. So, how to rely that the time, which you have told me based on only a small number of observation is reliable? This assurance can be achieved by the concept of weak law of large number. So we try to understand it.

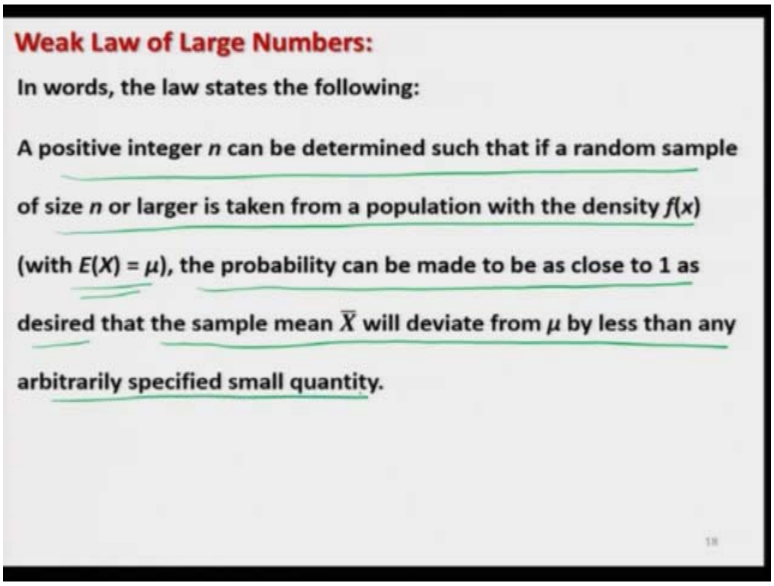
(Refer Slide Time: 30:39)



So, now, I am sure that by after looking at this slide, you will not get scared, because these are very simple thing. If you try to look at them carefully, try to divide them into small parts, you can understand them very easily. So, what I am trying to do here, let f(x) be the probability density with mean μ and variance σ^2, and we are simply assuming that the variance is finite. And we get a sample from this probability density function, and we compute the sample mean $\bar{X}_n$ based on a sample of size a small n.

And now, let ε and δ be any two specified numbers satisfying ε is greater than 0 and δ is between 0 and 1. Then, if n is any integer greater than or equal to $\sigma^2/\varepsilon\delta^2$, then probability that $\bar{X}$ and minus μ less than ε is greater than or equal to 1 - delta. What is this trying to tell you? δ is a quantity, which is a line between 0 and 1. So, do not you think that this is the probability?

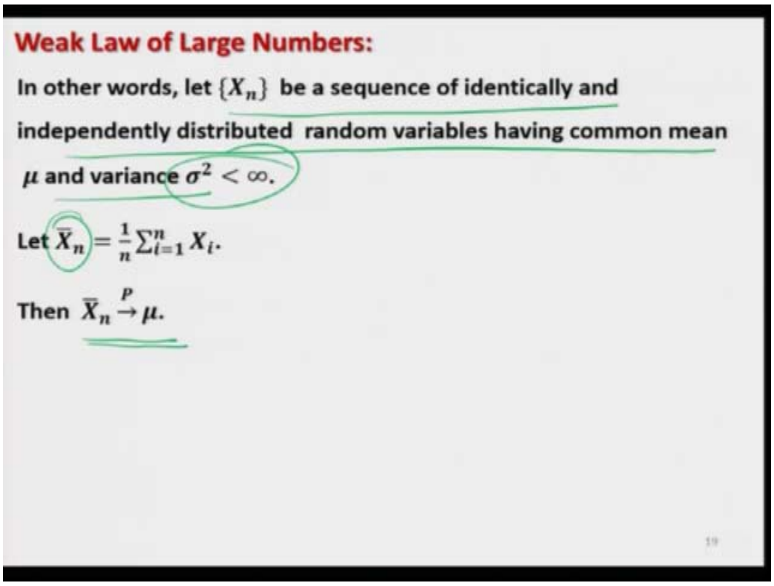
(Refer Slide Time: 31:51)



So, now, if you try to see here what is happening in very simple words, I can state the same thing as following. A positive integer n can be determined such that if a random sample of size n or larger is taken from a population with the density is small fx, which has got the mean mu, then the probability can be made to be as close to 1 as desired that the sample mean $\bar{X}$ will deviate from μ by less than any arbitrarily specified small quantity, which is here delta.

So, just by choosing proper value of your sample size, you can make this probability as close as possible to this quantity 1 - δ. And that is what you are actually doing. That was happening when you were trying to conduct the simulation in the earlier slide that you were trying to increase the size of the sample.

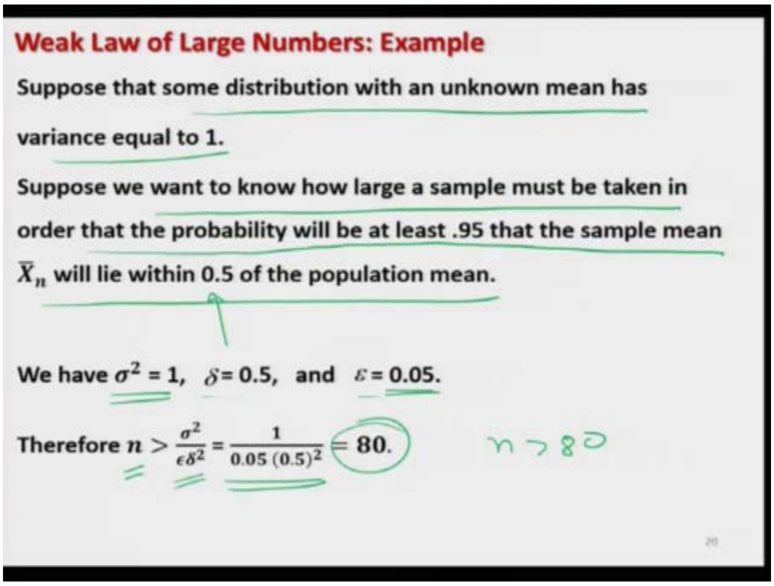
(Refer Slide Time: 32:50)



And this is a similar thing. And now, the weak law of large numbers says in a very simple language in the using the concept of convergence in probability that let X_n be a sequence of identically and independently distributed random variables having a common mean μ and variance σ^2 which is finite, that is the only condition that the variance is finite. Then the sample mean $\bar{X}_n$ will converse in probability to mu.

So, now, you can see that how this convergence in probability is going to help you in the weak law of large number and how are you going to get an assurance that whatever inferences you are trying to draw on the basis of a finite sample and they are going to be valid for a larger population.

(Refer Slide Time: 33:37)

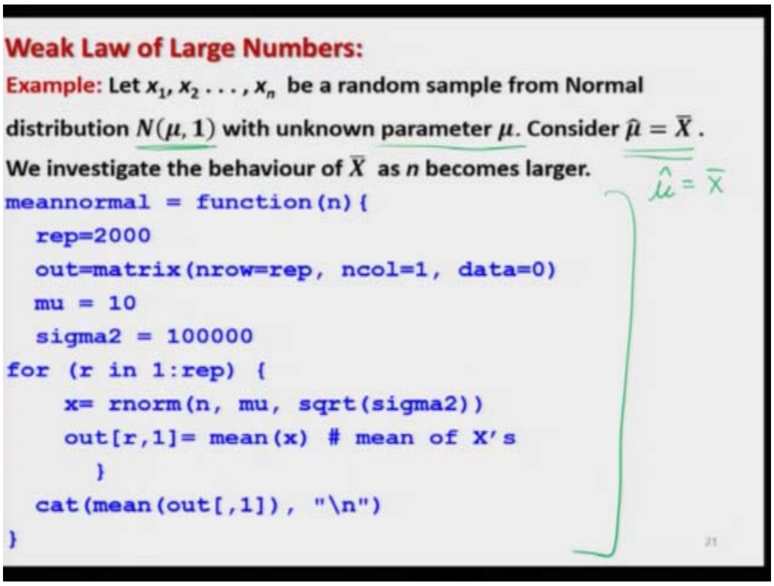


So, let me try to take an example to show you here that how are you going to find n and actually, this also helps you in finding out the value of the sample size that how many observation one can take depending on some condition. So, suppose that some distribution with an unknown mean has variance equal to 1, because that is the only condition that we are saying that variance has to be finite.

And suppose we want to know how large a sample must be taken in order that the probability will be at least 0.95 that the sample mean $\overline{X}_n$ will live within 0.5 of the population mean? Population mean is unknown to us, that we do not know, but we can always say that okay, the value should not be beyond certain limit. So, we are trying to say that it has to lie within 0.5 of the population mean.

So, here we have σ^2 equal to 1 and we can choose δ ε is equal to 0.5 which is coming from here. And suppose we choose the ε is equal to 0.05. So, now, I can choose here n say $\sigma^2 / \varepsilon\delta^2$. And if you try to substitute this value and solve it, this comes out to be n greater than 80. So, if you try to choose a sample of size more than 80, then you can achieve this.

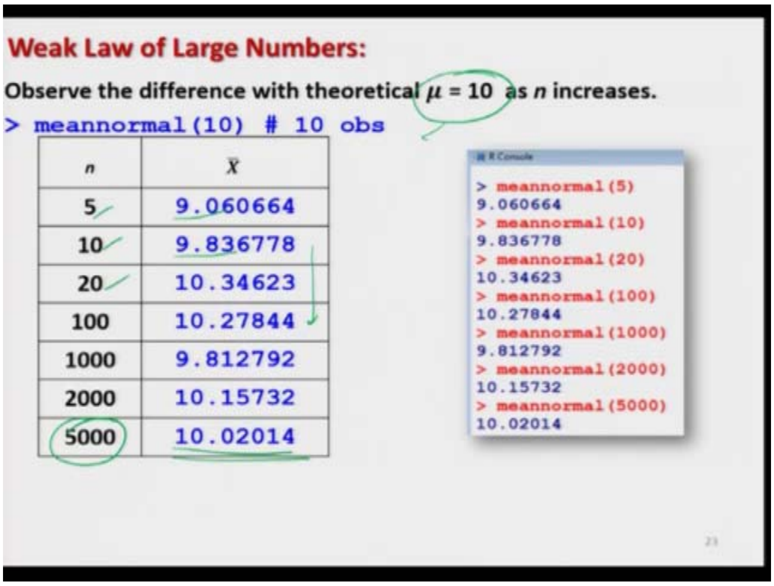
(Refer Slide Time: 34:55)



So, do not you think that this is very interesting? And if you try to want to have an assurance you can also conduct a small say simulation, which I am showing you here. So, I have done here that I have written a small program here that is very simple, where I am trying to generate the random sample from normal population with mean μ and variance 1. And this μ is the population parameter that is unknown to us.

And suppose we are trying to write down that this μ is going to be estimated by sample mean Well, I am using here a symbol $\hat{\mu}$, $\hat{\mu}$ means μ is being estimated by the value of sample mean. I will try to talk about and discuss this the symbol hat at a later stage, but in simple language, it is trying to say you are trying to find out the value of μ by sample mean. Now we try to investigate the behavior of this $\bar{X}$ as n become larger. And for that I have written here this program.

(Refer Slide Time: 35:51)

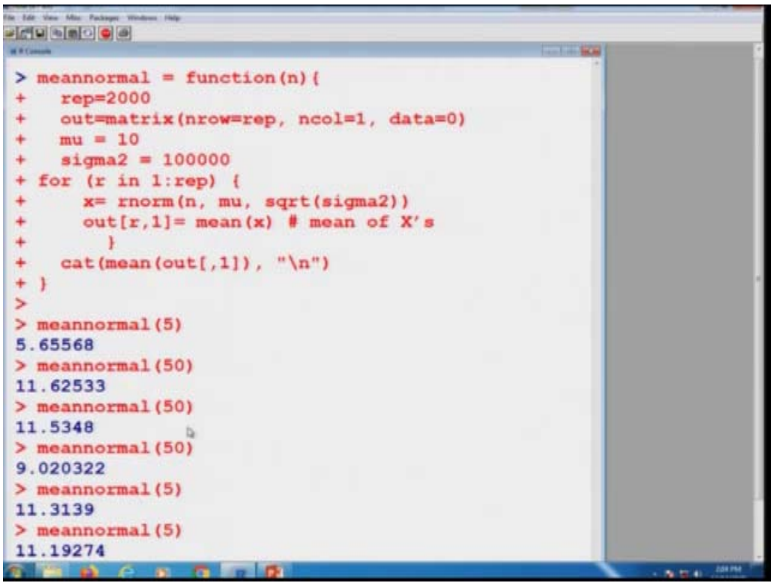


And this is the screenshot of the program anyway. But now, if you try to look at these values, what are you going to get? So, I am trying to generate the observation from a known population μ equal to 10. But for a while, I am assuming that, that is unknown to us. And then I try to choose different values of here sample size and I try to compute this μ by the sample mean. What I try to do here I try to choose here different values of here and I try to execute this program.

And then I try to find out here the value of this sample mean $\bar{X}$ when n is equal to 5, 10, 20, and so on. So, you can see here the true value here is 10, but then for 5, it is coming out to be 9 then for 10 it is 9.83. But as you try to increase the value of your sample size, the values are becoming more closer to 1, and the definition can be on the left hand side of 10 or the right hand side of the 10 that is less than 10 or say greater than 10, but these deviations are very small.

You can see here when n is equal to 5,000 the values of sample mean is coming or to simply 10.0. So, now let me try to show you this program on the R console, so that you get convinced.

(Refer Slide Time: 37:10)

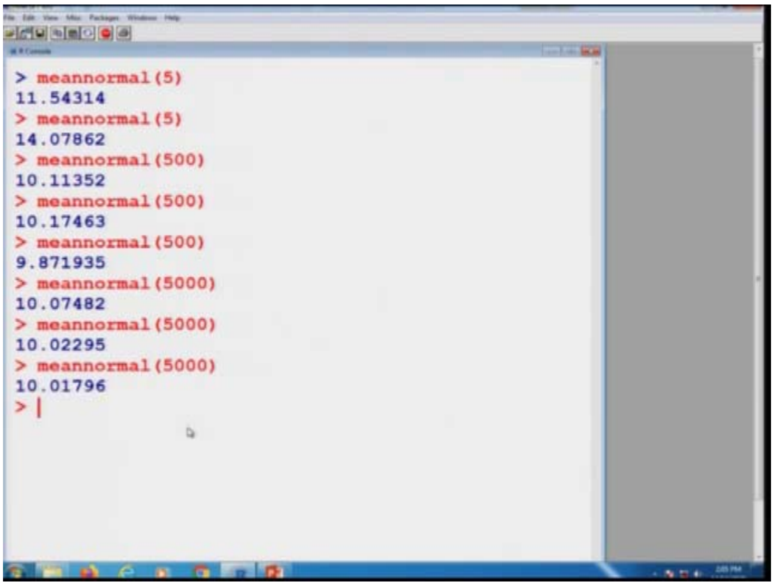


And you can also conduct these types of experiment that will give you more confidence sexually. So, let me try to clear the screen, remove this graphic so that we can create new graphic so you can see here like this, so I can execute this program by writing mean normal, which is the name of the program which I have given.

You can see here if I try to give here a sample size to be 5, which is here like this. So you can see here that you are getting here a value here 5.655. But if you try to take here, say this sample size to be 50 we are getting here a value 11.62. And if you try to repeat this experiment, you will be getting the same thing here. But if you try to repeat the same value for here for five sample size, you will see that there is a lot of variation.

I mean, the earlier value was 5.65 now it is coming out of 11.31. And if you try to change it, this will again come out to 11.19.

(Refer Slide Time: 38:38)



So, now, in case if you try to repeat it you can see here this is happening like this 11, 14. So, what if you try to I mean, the sample size to be 500 you can see here it is more closer to 10 and if you try to repeat it once again it is more closer to 10. And if you try to repeat it once again it is more to the to 10. And if you try to make it the sample size to be 5,000 here, you can see here that it is more closer to 10. And if you try to repeat this experiment, you can see here the values are pretty close to 10.

So, now, we come to an end to this lecture. Now, I want to ask you one thing. Up till now, most of you what is scared of theory, whenever from this tests point of view we used to talk about the theory, you will get scared. But now after this lecture, are you convinced that without this theory, you cannot understand what is really happening inside the data. And now, you can find out the reason that why this is happening. So, this is what I wanted to explain you beside explaining the some important result about the sample mean and related results.

So, now, it is very important for you that you please try to look at these results try to understand them. Whatever mathematically statements have been suggested try to understand them. Now, you should be convinced that yes, they are making some sense they are trying to convey something which is very useful for your data science. Because now we are moving into the statistical inference where we are we are really going to use the result in the real life.

So, it is very important for us to understand that whatever is happening we should know the reason. And as a human being once we know the reason, we feel satisfied, and more happy.

So, you try to have a look try to revise. And I will see you in the next lecture. Till then, good bye.