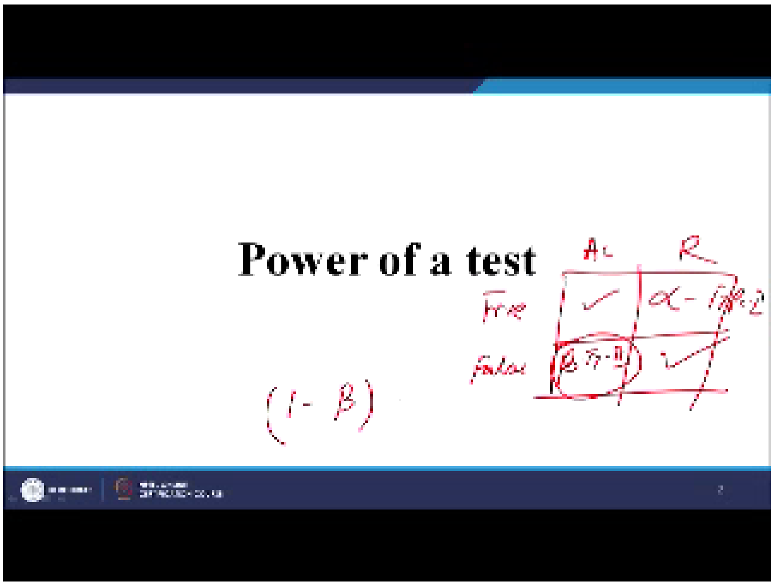
Marketing Research and Analysis-II
(Application Oriented)

Prof. Jogendra Kumar Nayak
Department of Management Studies
Indian Institute of Technology – Roorkee

Lecture - 21
Power of a Test and Sample Size Calculation
(Problem Solving)

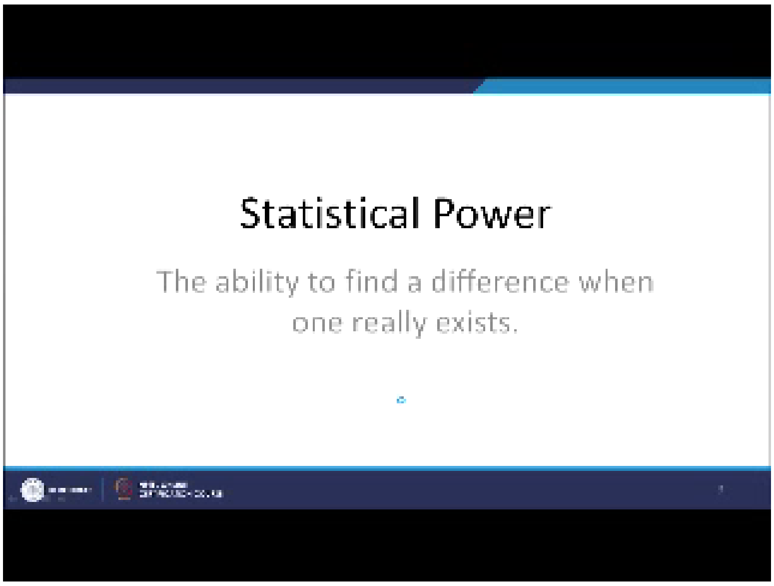
A very warm welcome to everyone out here welcome to our course of marketing research and analysis 2nd part. So, as in the last class we are discussing about hypothesis development and the errors that we usually face during the hypothesis testing so the type 1 and type 2 errors and then we came across something important equally was the power of a test. So, we had stop there and we are today continuing in this lecture 5 in the about the power of a test and then about sample size calculation during hypothesis testing.

(Refer Slide Time: 01:16)



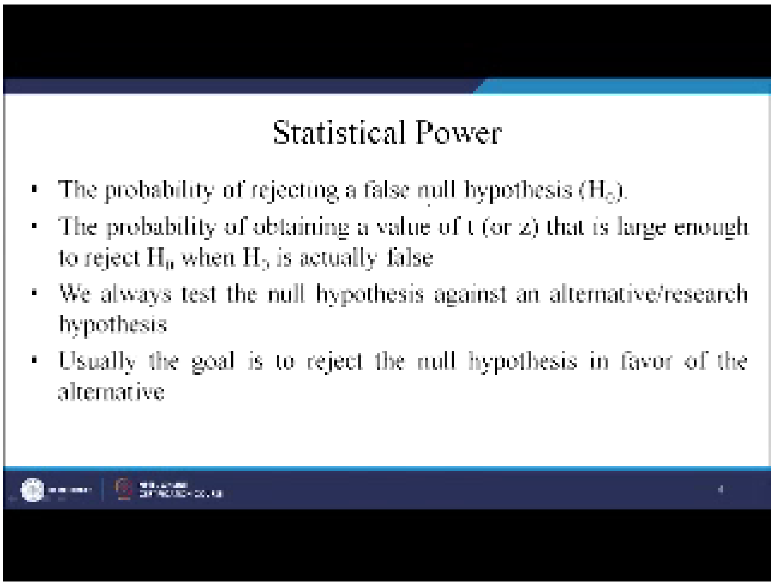
So, today we will go through it so let us see what exactly is the power of a test if you remember last time we spoke about this 2 errors one when the hypothesis was true & false and then it was accepted and rejected. So, when it was true and accepted so it was ok when it was true but rejected so it was we said alpha or Type 1 error ok, similarly when it was false but accepted it was again a mistake and we said this mistake is called beta or type 2 error ok. And this was again the correct so this when we said about this beta this beta is a type 2 error so we said 1 – Beta is what is called as the power of a test.

(Refer Slide Time: 02:01)



So, what is the statistical power it says the ability to find the difference when one really exist the higher the power of a test you can say the stronger the testes in simple terms we can understand it this way. So, let us see that we are saying the probability of rejecting a false null hypothesis so as you saw in the first initial slide.

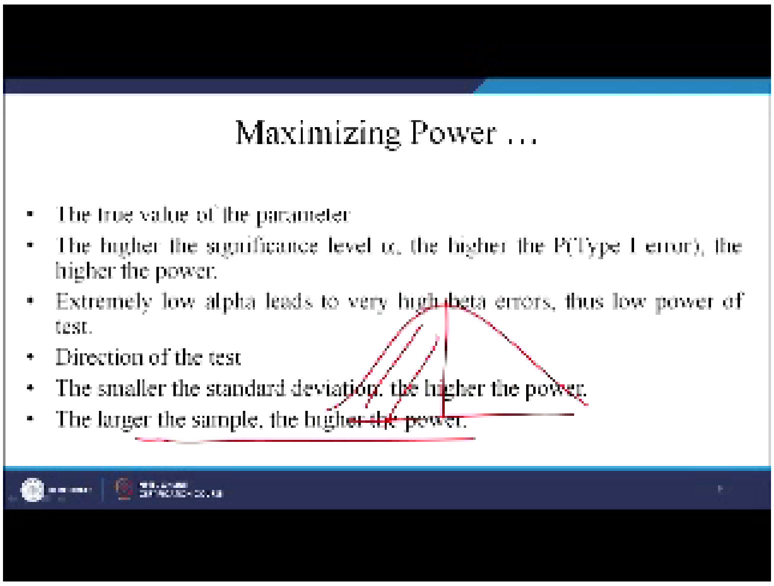
(Refer Slide Time: 02:28)



So, we said the type 2 error was accepting a false null hypothesis so, power is 1 - accepting a false null hypothesis that is what it says the probability of rejecting a false null hypothesis ok. Similarly the probability of obtaining a value of t or z whatever that is large enough to reject H0 when H0 is actually false so you need a substantial significant value so that you can claim you know this is wrong and this not to be accepted right.

We always test the null hypothesis against alternative that we have already discussed usually the goal is to reject the null hypothesis in favourite alternative that also we have discussed keep any researcher whenever he tries to conduct a researcher does a hypothesis testing the measure purpose generally speaking it is not always but generally speaking in any at least in marketing research and market related studies it is to reject the disprove or reject the null hypothesis and accept the alternative hypothesis. So, the researcher is more interested for the alternative rather than the null we can understand it this way.

(Refer Slide Time: 03:40)



What is the power let us come to the discussion power is probability so the number as any probability value lies between 0 and 1. So, if it is as you can understand if it is towards 0 it is poor but if it is towards one as closest to 1 it is better so 0 is bad in simple terms one is good ok. to make power has highest possible so the intention of the researcher is always to have the power as much as possible right why because you are visiting a false null hypothesis read so that is our intention is to make the power as high as possible right so how to maximize power.

So, the true value of the parameter if you understand the true value of the parameters the parameters are maybe the mean the standard deviation when is the true value when you can calculate you can understand and then accordingly your power of the study will increase similarly the higher the significance level is now we have also discussed that when you can calculate you can understand and then it accordingly your power of the study will increase.

Similarly the higher the significance level now we have also discussed that alpha and beta are related right so if your alpha is high beta will come down slightly similarly vice versa. Now as you have a higher alpha automatically you have a lower beta so what happens as a net result is your 1 - beta becomes higher right. So, therefore the higher the significance level higher the Type 1 error the lower is a type 2 error and therefore the higher the power ok.

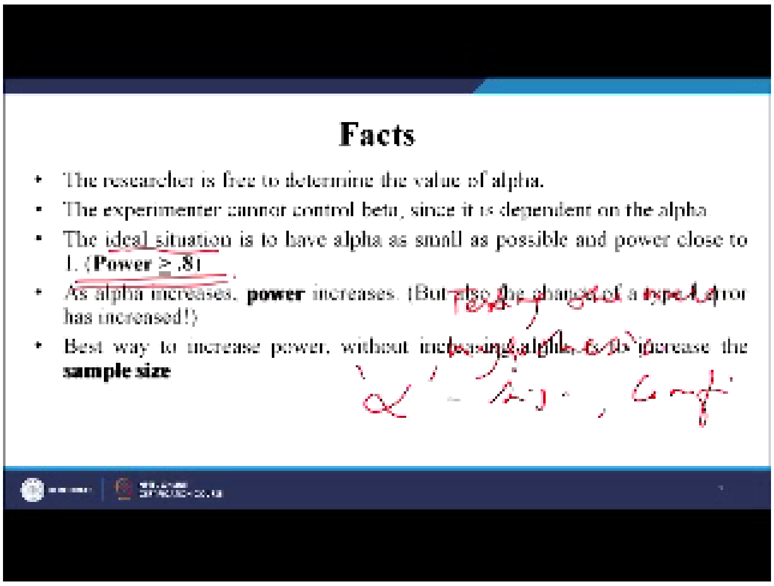
Extremely low alpha leads to very high beat errors thus lower power of the test as just I have explained if your alpha value suppose that is it very low level right 1 or .01 or something then alpha and beta are like a swing like a swing alpha and beta are like a swing so they cannot be the same if alpha is going down beta tends to go up alpha so as it tends to go up so as it tends to go up in very low alpha will lead to very high beat errors thus the power of the test will go down since power of the test was beta ok 1- beta.

Direction of the test, now the direction of the test also has an influence on the power of the test know how that happens let us see this now when you have a two directional test so that test is divided into two sides in such conditions the power the test becomes in a slightly low right in comparison when you are interested when you are very clear with the direction or the orientation of the study. So, when you are interested in orientation you are interested may be for other one side only there in that case the power of the test becomes more robust.

You can logically also think because you know the direction in the earlier case when it was two directional or non directional you did not know which side the test is moving therefore when you are clear logically the power of the test also the strength of the test improves ok. The smaller the standard deviation higher the power this is another relationship between power and standard deviation. The larger the sample higher the; more the sample size increases as sample size goes on increasing the power of the study goes on increasing right.

So, the best way to improve the power of the test is to increase the sample size you can be sure about that.

(Refer Slide Time: 07:04)

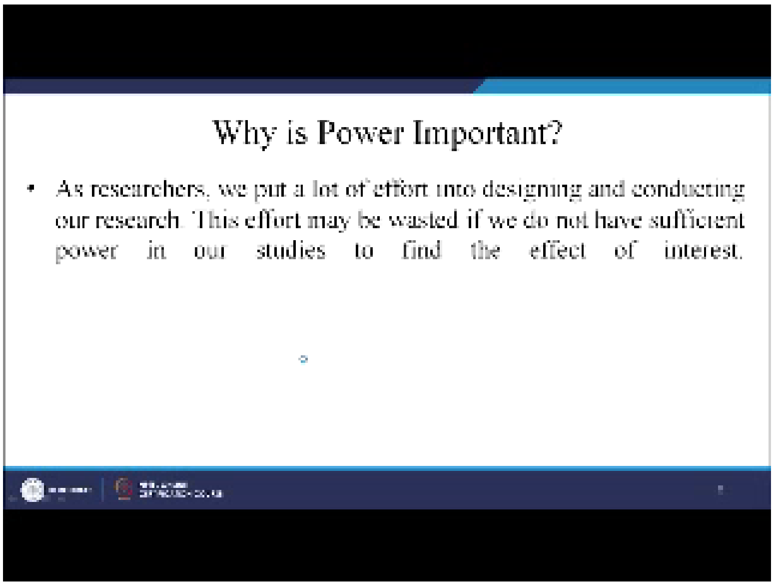


Facts: There is something is free to determine the value of alpha we have already discussed significance value of alpha. Experimental cannot control beta since it is dependent on the alpha you are always have said in the first lectures also that we are testing the null hypothesis what we are doing we are testing the null hypothesis ok. So, when we are testing the null hypothesis we are testing it at a particular level of alpha or at the significance level correct or at confidence level.

So, what we are saying is the study is dependent on the researcher who decides the alpha so as we cannot control the beta we are dependent on the alpha right the ideal situation is to have alpha as small as possible and power closed to 1 right. But very low alpha also will create very high problems. So, what you do is the ideal situation it is saying the ideal is the ideal situation you have alpha a small alpha and the power closed to 1 and statisticians have found out when the power is greater than or equal to 80% it is considered to be a very strong or a good test ok so, that you can always check.

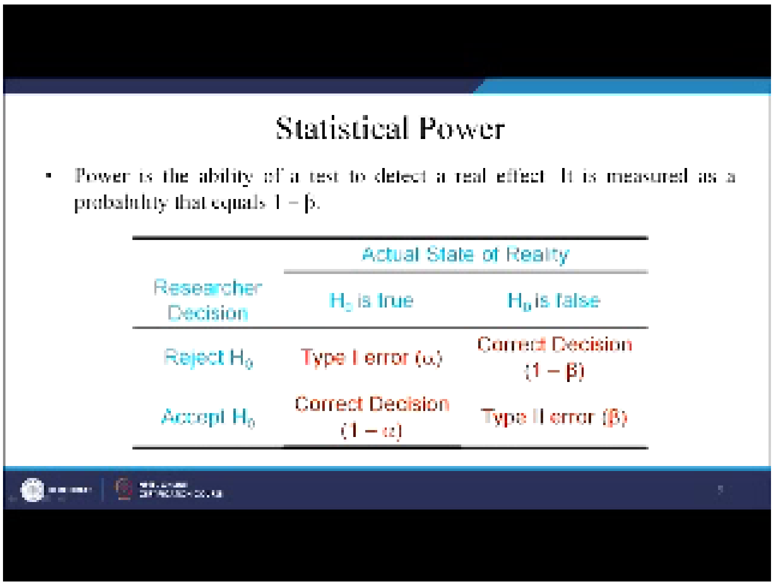
Alpha increases the power increases right but also the chance of a Type 1 error has increased ok. Best way to increase power without increasing alpha is to increase the sample size this just now I explained.

(Refer Slide Time: 08:45)



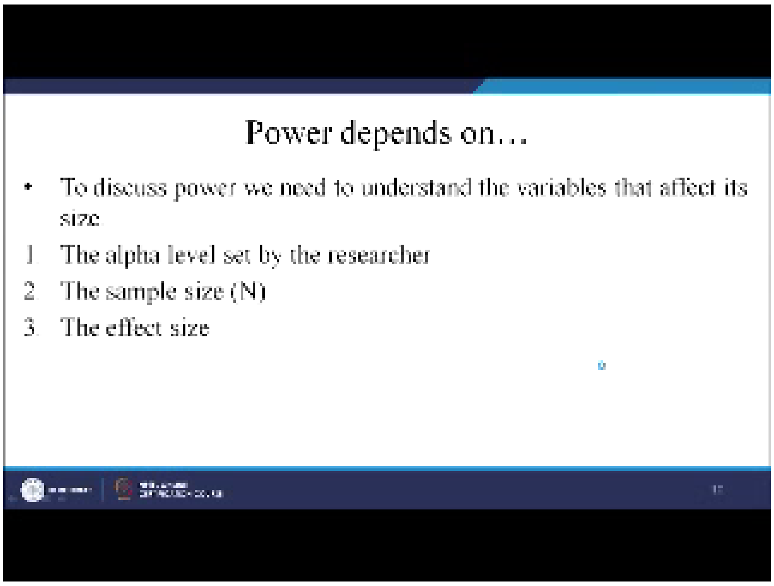
Why is power important? As researchers we put a lot of effort into designing and conducting the research. This effort may be wasted if we do not have sufficient power in our study that means you do not have the power means you cannot reject a false null hypothesis. So, when you cannot reject a false null hypothesis that means you have accepted or bad product in the market which is very dangerous right. So, this is what is the effect of interest for consumers and researchers.

(Refer Slide Time: 09:22)



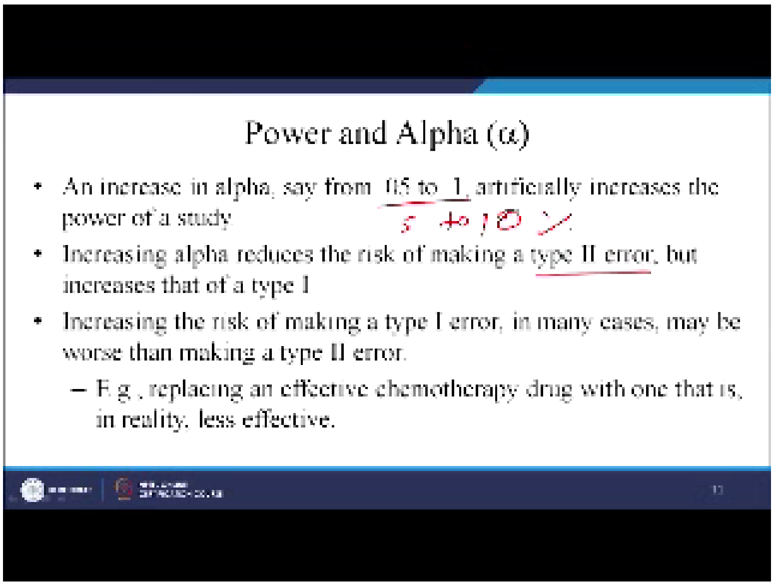
So, power is the ability to detect a real effect is measured as I said 1 - beta right so this is a situation.

(Refer Slide Time: 09:32)



To discuss power let us; we need to understand the variables that affect its size the alpha level set by the researcher we discussed it sample size and now what is this effect size let us discuss this one ok this is already been discussed but still an increase in alpha from let say 0.5 to 0.1 that means 5% to 10% artificially increases the power of a study. Now when alpha is increasing we said beta will decrease and as beta will decrease the power will increase ok. Increasing alpha reduces the risk of making a type 2 error but increases that of a Type 1.

(Refer Slide Time: 10:07)



Increase the risk of making a Type 1 error in many cases maybe worse this is interesting what is saying increasing the risk of making a Type 1 error in many cases maybe worst than making a type 2 error. What is this which condition not let us go back I had said that type 1 error is a producer's risk if you remember it is a producer's risk or a manufacturers risk. Type 2 is a consumer's risk. So in case when the researcher feels that the producers risk is to be given more importance right or is of more significance then in that case he would be more cautious to reduce the Type 1 error right.

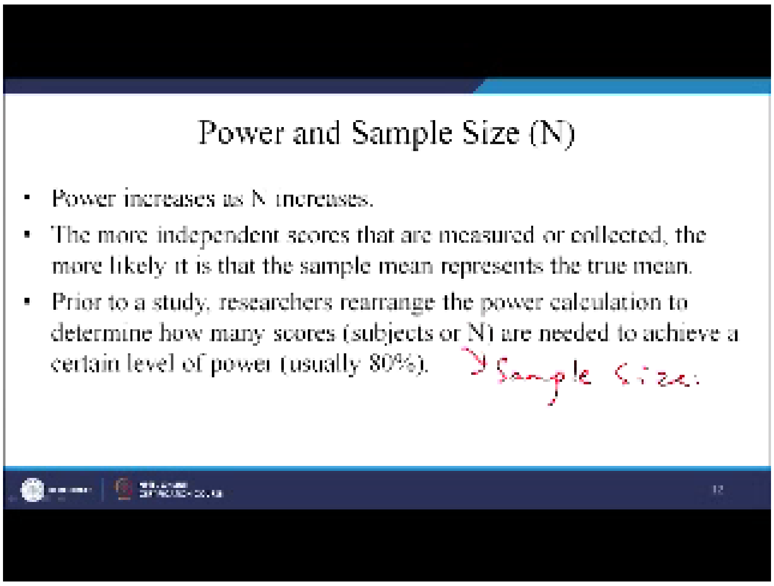
So, on other hand in some cases when the consumer litigation support the consumer files a case of litigation in you go to court and sues the company or something in those cases the companies are more bothered for example Pharmaceutical drugs companies are something in those cases there more cautious about the type 2 error. But in steel manufacturing company or you know laptop manufacturing company of something for example of that kind the general use of manufacturers are more concerned about the Type 1 error rather than the type 2 error because at least if the laptop goes wrong you have a probability or a chance of getting back the laptop and correcting it rectifying it and then giving back to the customer and even making use of giving some extra incentives so that he is happy.

But what if you are given a wrong drug to the consumer if you are given a wrong drug the person may suffer from severe disease or several disease or severe condition and he might even die or you might have some problem then he will not be able to you cannot make him happy again it is very difficult. So he would go to the court of law and maybe file a case against you right as a company ok.

Example if you see this is example replacing an effective chemotherapy drug with one that is reality is less effective. Now suppose you are replaced and effective drug against in effective one and then it is a very dangerous case right. So, there in such cases you have to be careful power and sample size power increases as N increase right I had explained. The more independence scores that are measured or collected the more likely it is the sample mean represent the true mean.

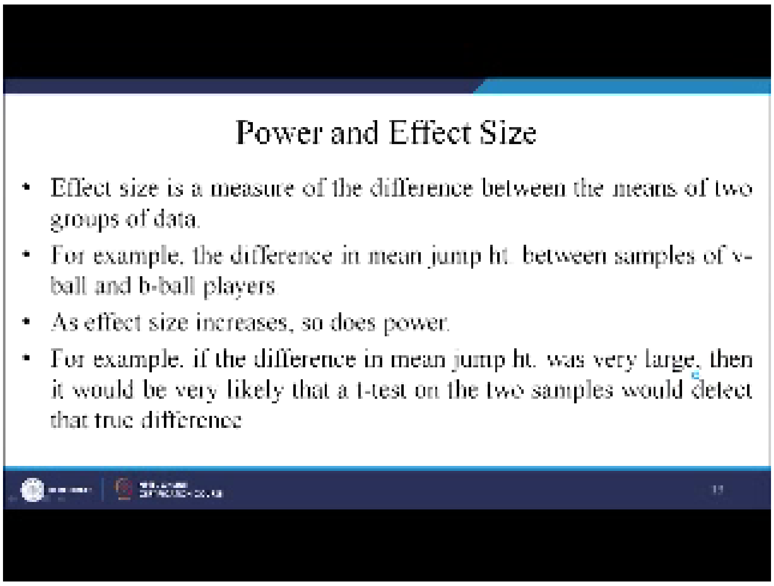
So our ideas is what is power we said having a strong test, a test which is stronger in the sense that it is explaining the right thing. Now what is the right thing? The right thing is that the sample mean is actually representing the true mean that is what is actually hypothesis testing. So, as you have more number of sample what is happening is the difference between the true mean and the sample mean will come close right as they come close we say the power of a test as increased.

(Refer Slide Time: 13:11)



Prior to study researchers arrange where is the power calculation to determine how many subjects or sample size. are needed to achieve a certain level of power usually 80% so, that differs from study to study ok.

(Refer Slide Time: 13:18)

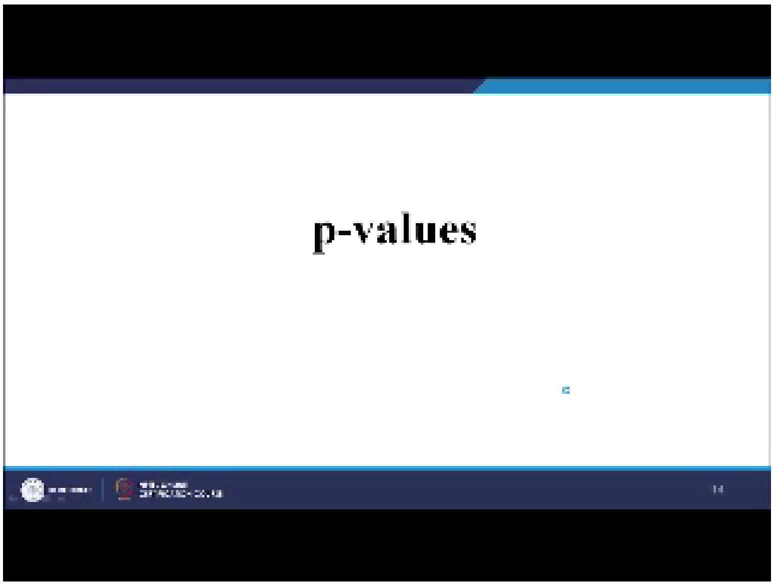


Now coming to the third thing that we were the thing what is power and effect size? What is effect size first let us understand? It is a measure of the difference between the means of the two groups of data. Now group 1 let us say and group 2 so the difference between the means so whatever the means are let us say G1 bar minus G2 bar is equal to is the difference right effect size is this difference between the means.

So, we are saying that it should be as close as possible right to represent the population that means they are from the same population. For example the difference in mean jump height between samples of Volleyball and basketball players. we have taken 2 kinds of sports as effect size increases so does power right. So, for example let us understand how do you understand as the from reality real life real term. If the difference in mean jump height was very large that means the volleyball players and the basketball player the jump right. The jump let us say the Volleyball players are jumping more or the basketball players are jumping more whatever if this difference is very large then it would be very likely that a t test on the samples would detect the true difference.

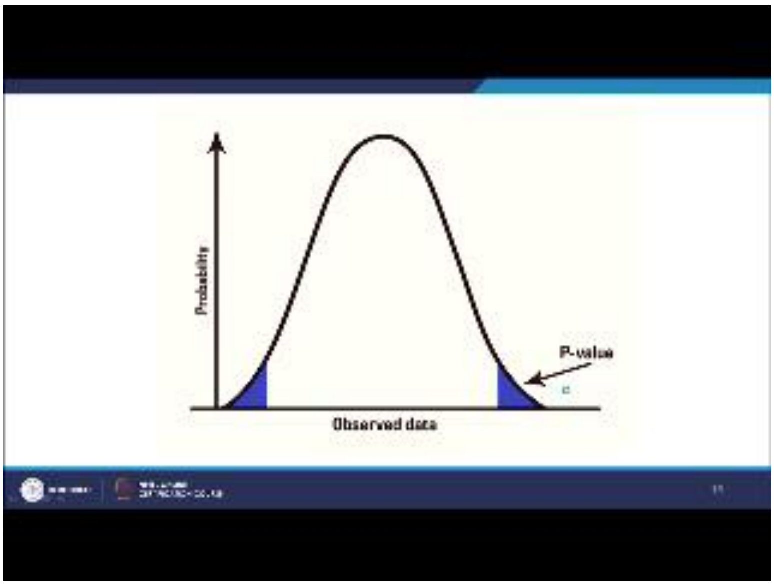
But if the defect if the difference is very less than the test would say there is no difference so the null hypothesis would be accepted the difference between the jump height and jumped by the Volleyball players and the Basketball players is actually more or less the same there is no difference rights and you would not be able to reject the null hypothesis and accept the alternate in those cases. The effect size higher is the effect size higher is the power of size so that means the power is ultimately to detect whether the null hypothesis is to his properly being rejected or not right and alternatives being accepted or not.

(Refer Slide Time: 15:14)



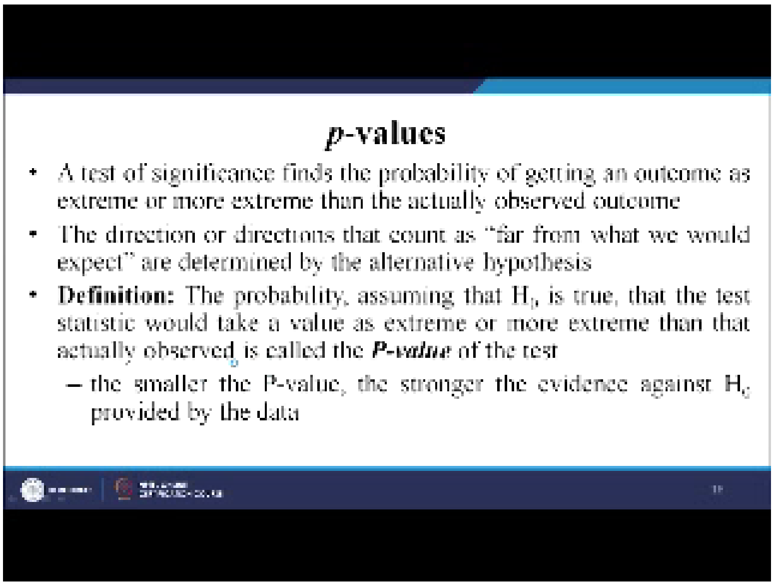
P- values I think I have already been explained.

(Refer Slide Time: 15:15)



This is just a repetition of the slide so just again just to show you. So, this is the dividing so these are the two sides so we have kept this earlier also as 5% maybe 2.5%, 2.5% for a 5% case.

(Refer Slide Time: 15:29)



So, what this p value? Now you have seen many a times that the in statistics or in any publication you must have seen that nowadays they are writing the p-value. Now what is this p-value and what is it indicate basically? The p-value actually tells is a test of significance which find the probability of getting an outcome as extreme or more extreme than the actual observed come. Now let us understand what is the meaning of that, so, if this is the test so we are saying what is the probability of the value lying to the extreme corners.

So, it says what is the probability of getting an outcome the result whatever the µ as extreme or more extreme than the actually observed µ. So, actually observed mean so it would be the mean it could be any other parameters also. What is the probability? If the probability is very, very low let us say, probability of getting this value is very low then that means it is very close to the, this end right.

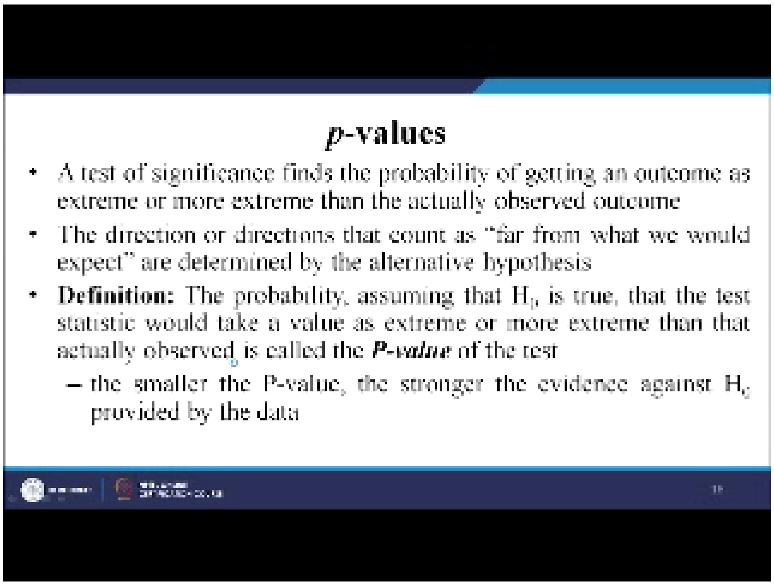
So, when it is very close to this end or something then we say the probability if it is less than 05, .05 at 5% level of significance study. So, if it is very low then we will reject the null hypothesis. What is it saying? Now let us look at it again the probability of getting an outcome to the extreme ends. If the probability of getting it at the extreme ends is high right so we will accept that the null hypothesis but if the probability of getting the value to the end at the extreme end is very low then we will reject the null hypothesis.

Let see this direction or directions account as far as far from what we could expect a determine the alternative hypothesis. The probability assuming that H0 is true the test statistic would take a value. Let us come back as you mean that H0 is true that the test statistic would take a value as a extreme or more extreme then that actually observe is called the p-value that means that means if you are you are calculated value somewhere here and your probability p-value is lying somewhere here that means you are rejecting the null hypothesis ok.

The smaller the p-value the stronger the evidence against H0 provided by the data that mean if it is a very low p-value we would reject the null hypothesis and accept the desired alternative hypothesis but is the p-value is let us say more than 5% for in generally most of the cases then you cannot reject the null hypothesis rather you accept the null hypothesis or you cannot reject the null hypothesis let us not say except, saying except null hypothesis is the wrong way of explanation so we will say you cannot reject the null hypothesis ok.

So, as I have seen let us understand this p-value what is this p-value? As it is saying the probability that assuming H0 is true that the test statistics would take a value as extreme or more extreme than actually observed value. Now how do you understand the smaller the p-value the stronger the evidence against H0 provided by the data or the null hypothesis let me explain you through a diagram this is to be very clearly understood. How do you check this p-value?

(Refer Slide Time: 19:02)



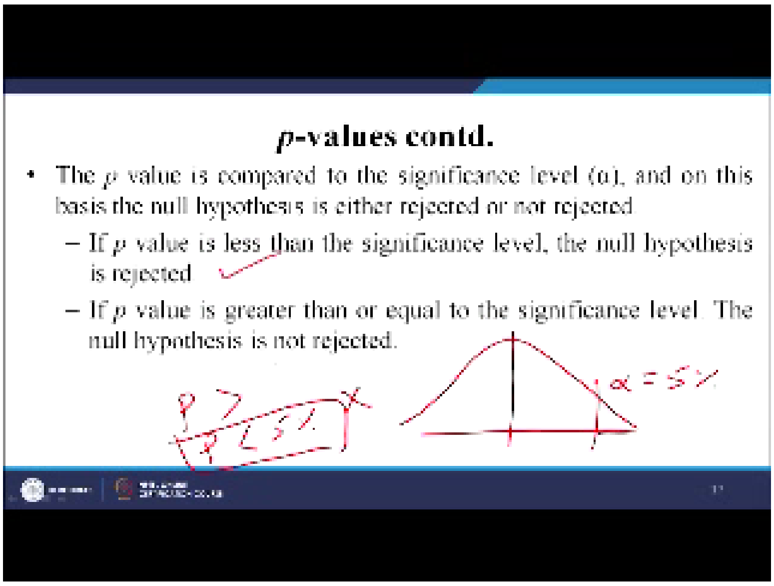
For example let us say just take this so you have calculated now how do you calculate the speed well you know you have taken a alpha of let us say 2.5%, 2.5% ok. Now let us say calculated value is maybe for example of 1.25 ok, now 1.25 plus or minus plus minus has a the area under the curve is equal to; if you go to the Z table you will find out that it is something around .394 sorry, 394. So, now .394 is the area under the curve that means what in both sides.394 and .394 so that means what .5 -.394 that is equal to something .105 something around that.

So, this is the value that you have calculated now your values laying that means somewhere around here your rejection region this is the rejection region & this is the not rejection or acceptance region acceptance ok. So, if you have value has come somewhere here beyond then that means you cannot reject the null hypothesis that is why the lower the p-value had my area under the curve this value which is dependent on this calculated value statistic would have come will it say instead of 1.25 it would have been late say 1.7 or 8 and the table value corresponding let us say is area under the curve. 497 right.

So, in this case what is happening if I am a subtracting .5 - .497 then this is giving me .0 let me do it 0.5497 so 300 so, .003, so now .003 is falling somewhere which is your the limit of significance which you are declared as a researcher was 5% so in this case lets say 2.5% here. so it is coming somewhere here to the extreme end of the distribution normal distribution then you say that I am cannot accept I cannot accept I am rejecting the null hypothesis because it is falling into the rejection zone the best clarity is very, very important for any researcher how do you calculate the p value.

How it is so different from the alpha so this is what needs to be very clear then we move on one has to the next slide right.

(Refer Slide Time: 22:04)

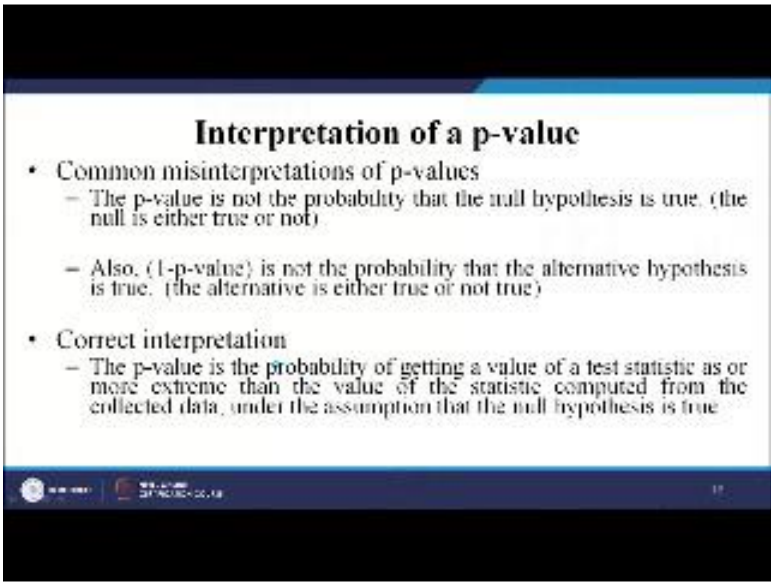


Let us written out here the p value as compared to the significance level and on this basis the null hypothesis is either rejected or not rejected. So, you have it right so let us say the alpha is this is alpha ok if p-value is less than the significance level, so your alpha is equal to 5% if p as I said if two conditions if p is greater than p is less than right. If p value is less than the significance level p is less than the significance level of lesson 5% right the null hypothesis is rejected.

So, you are rejecting the null hypothesis on basis of saying that since my probability of occurrence at the extreme ends of the parameter is less than 5% whatever significance level I have taken so I am rejecting my null which says that there is no difference no change and accepting the alternate rather which says that there is change there is a difference right. On the other hand vice versa if p is greater than alpha that means this is alphaif p is greater than alpha somewhere this side we say, I understand.

Then we would not be in a position to reject the null hypothesis and we would therefore say there is no change or no difference in the study right which is not a concern or interest actually to the researcher in which there are some special cases where we will see later on maybe where the researcher is interested in accepting null hypothesis but 99% in generally in a marketing condition any economic condition we are interested in always accepting the alternate right.

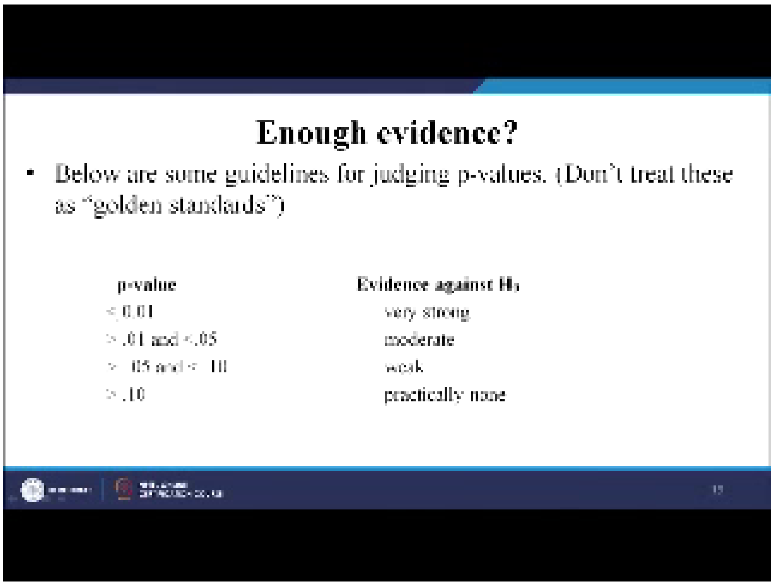
(Refer Slide Time: 23:49)



What are the common misinterpretations of p-values the values not the probability that the null hypothesis is true? The null is either true or not it does not talk anything about that right also 1 - p is not the probability that alternate hypothesis is true right. Correct interpretation is the p-value is the probability of getting a value as more extreme then the value of the statistic computed from the collected data.

So, you are saying simple terms let us not get into too much of complication the p-value says where is the outcome lying in the distribution. If the probability of it should lie at the extreme ends if it is lying; if the p-value is very, very less let us say it is less than 1% or less than 5% as per whatever significance level you taken in comparison to that. If it is very; much lesser than the significance level then we say we will reject the null hypothesis and accept our desired alternate hypothesis and otherwise vice versa the reverse.

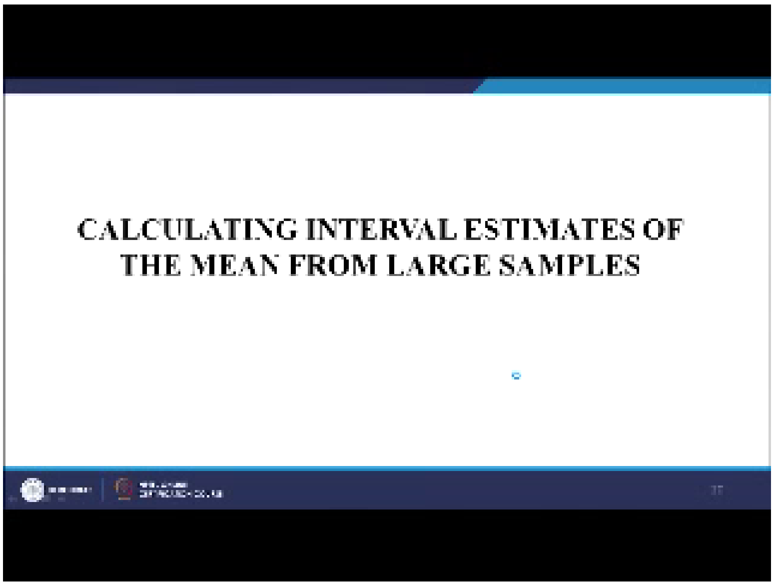
(Refer Slide Time: 24:57)



So, below are guidelines value is less than 0.01 evidence against H0 very strong p-value lies between 0.1 and 0.5 moderate right. P value lies between 0.5 and 1 weak evidence against the H0 that means you cannot reject H0 evidence against H0 means you cannot reject that means you cannot reject H0 right and if it is greater than 10 p-value then practically you cannot reject the null hypothesis.

These 3 at least this is the very clear this is; the researchers can think on these two situations right but finally we need to understand that the p-value is a very significant way or right good way of understanding whether to accept or reject a hypothesis right.

(Refer Slide Time: 25:55)

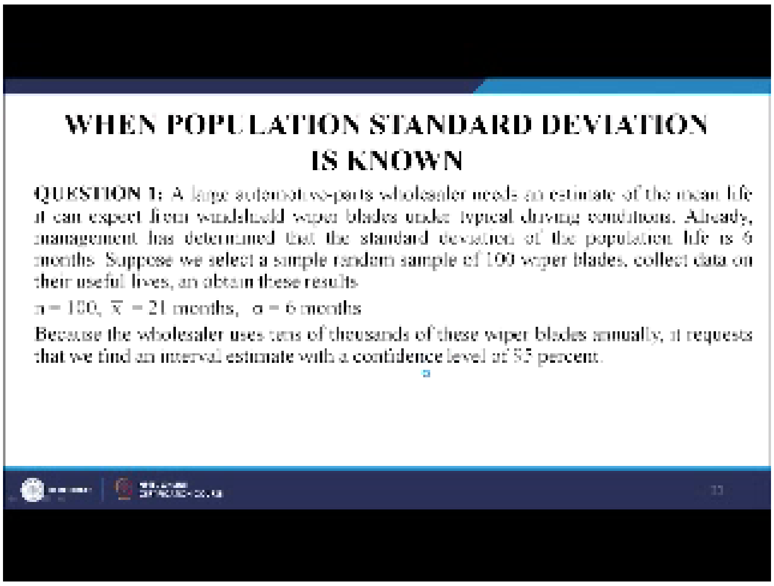


Now what will do is will start in we will get in next part which is calculating interval estimates of the mean from samples or the different conditions large sample, small samples.

So, the first case is a case where I am trying to calculate the interval estimate. Now what is this interval estimate? Now we have understood like a class interval what is the range? What is the range of the value? So there is a maximum value there is a minimum value so we are interested what is the; let say what is the class IQ? What is the minimum IQ level? What is the maximum IQ level is the average of the class IQ lying in between these two is not that is of interest.

So, this is like a line when let say you have the Mu here and you are this is a lower limit and this is upper limit and you are interested to see just whether if I know my lower and upper limit I can say whether my Mu is lying between these two are not right or which side it is tending to move right this side or upper to lower.

(Refer Slide Time: 26:54)



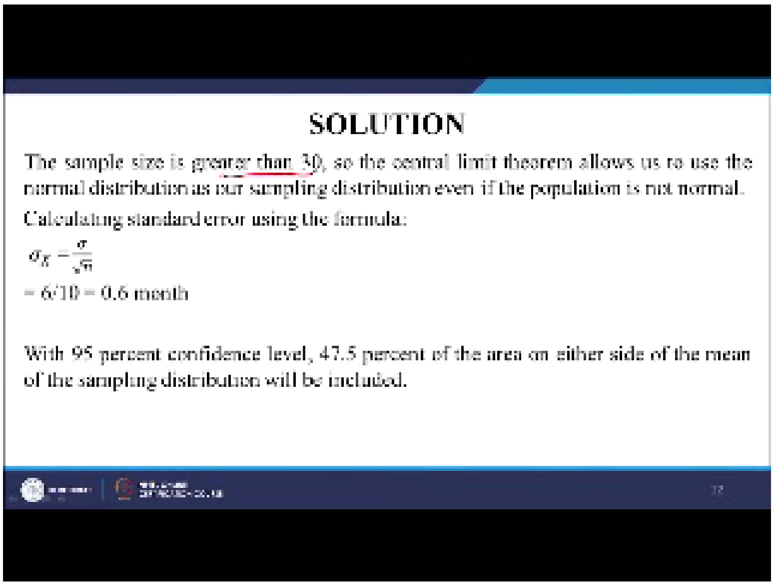
Let us see the first case, the first case is case where the population standard deviation is known now we are always; I have always been repeating the researcher is always interested to know whether the sample is actually representing the population or not is it a part of the population or not in all the studies. So go to the first question large automotive parts wholesaler needs the estimates of the mean life it can expect from windshield wiper blades ok under the typical driving condition.

The management determine the standard deviation of the population life is 6 months. Suppose we select a simple random sample of 100 blades so simple random sample of 100 collect data on the useful life and obtain x bar is equal to what is x bar? X bar is always the sample mean. The sample mean is equal to 21 right because the wholesaler uses 10’s of

1000’s of this wiper blades annually it request that you and me we find an interval estimate with a confidence level of 95%. So, it wants to know what is the life of this wiper blade at a 95% confidence level so that it can this because it is using in large bulk which can decide in the future whether to get it from the same supplier or not ok.

So, in this condition what will you do? How will you find out this interval? Let see the solution the sample size is greater than 30.

(Refer Slide Time: 22:04)



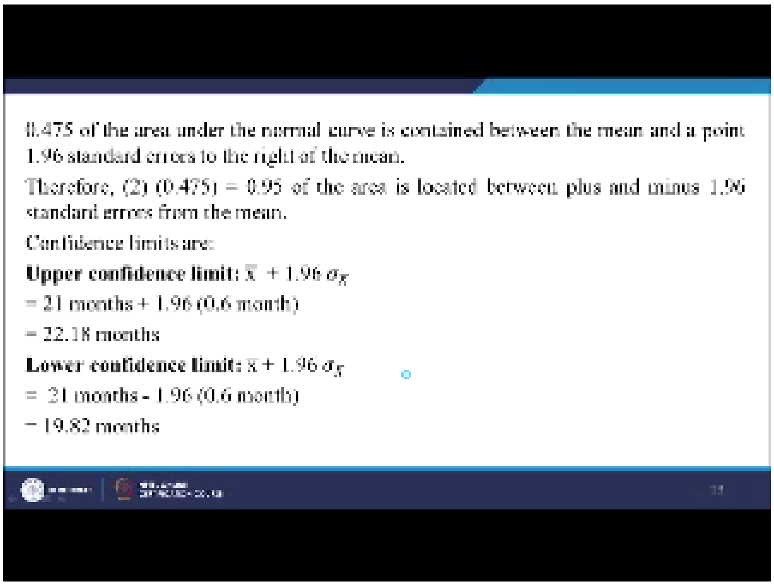
First you have to seek whether sample size is less than 30 or greater than if it is less than 30 is a case of t–test the t-distribution if it is greater than 30 it is a case of the Z distribution. So, it is greater than 30 the central limit theorem allows us to use the normal distribution. Had it been less than 30 it is not the same right. So, as it is more than 30 we are using the normal distribution and as a sampling distribution and even the population is not normal and it does not matter is says what it means is the sample size greater than 30 the central limit theorem allows us to use the normal distribution as our sampling distribution even if the population is not normal.

If the line is slightly it is not readable that means let us understand we are using the normal distribution in this case. So, how to calculate the standard error of the; how to calculate and standard now what is standard error? Standard error you may understand as the standard deviation of the population of the sample sorry standard deviation of the sample group. So, why it is required? Why should you calculate the standard error because standard error is need to be calculated because it talks about the series of samples and their distribution pattern right.

A series of samples and the distribution pattern so this is very important for you to calculate. So, now how it is calculated it is nothing but a ratio of the standard deviation upon the root over of sample size. So, in this case our standard errors of the sample for the distribution is 0.6 month 6 upon 10 so it is 6 by root over of n is 100 that is equal to 6 by 10 that is equal to 0.6 month with 95% confidence level.

Now you do not know which side of the hypothesis it should go which side of the tale of the distribution should go. So, when you do not know it is a two-tailed test. So, what we have done is the tail has been distributed into two parts 2.5 this side 2.5 % this side. So, therefore so this is actually .5 this is .5 that makes 1 correct 5 - 2.5 is 47.5% this side and 47.5% this side.

(Refer Slide Time: 30:59)



So, .475 the area under normal curve is contained between the mean and .1 and .9 that is the Z value the table value is 1.96 standard error right up to the right of the mean. So, therefore that is equal to let us find out the confidence limit 2 *.4 as I explain the last slide .95 so it lies between Z is equal to 1.96 negative to + 1.96 because this is for a 95% confidence lavel the Z value is 1.96. Now how do you calculate let us see this so $\bar{x}$ Y + Z σ. So, what is this part you have it and $\bar{x}$.

So, $\overline{x}$ versus sample mean which is 21 months +- Z is 1.96 into Sigma X is we have calculator something 6 months I believe 0.6 month. So, that gives us a value of 19.82 on the lower limit and 22.18 on the upper limit and your mean was 21 months correct, 21 month which lies between these two. So, in such condition we will say that this is the upper limit and such a condition you can even assume that the null hypothesis is to be not rejected and it lies within the boundaries of the extreme values that is a class interval. So, this is a first case this was case 1 what we will do is we will follow it up in the next lecture and will continue with some more examples and more conditions. I hope this lecture was clear to you thank you so much