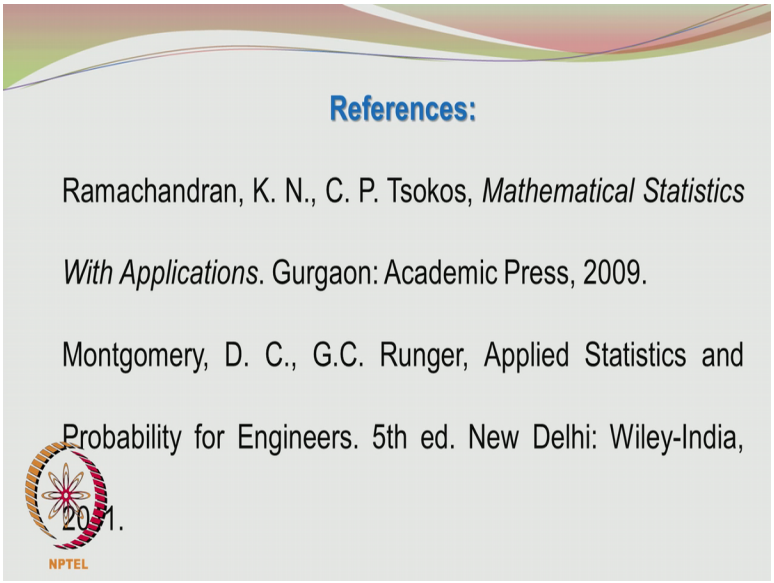
Statistics for Experimentalists
Prof. Kannan. A
Department of Chemical Engineering
Indian Institute of Technology - Madras

Lecture - 22
Hypothesis Testing - Part A

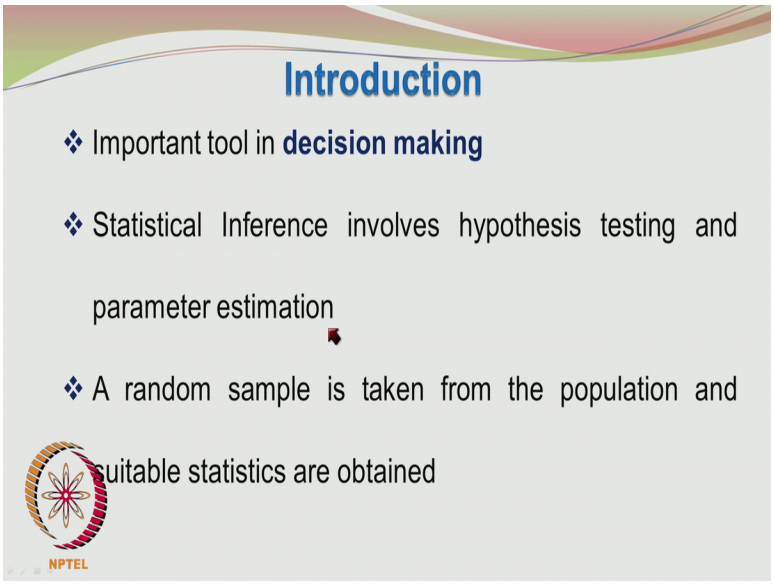
Hello and welcome back, in today's lectures we will be looking at Hypothesis Testing, so we will be defining different types of hypothesis, and how to handle different problem statements, posed the questions raised in the problem statements in form of suitable hypothesis. And then define a test statistic use the appropriate statistical distribution find the probability, and based on the probability we make certain conclusions.

(Refer Slide Time: 01:10)



So the references for this topic are the books written by Ramachandran and Tsokos, Mathematical Statistics with Applications from academic press. And the prescribed text book by Montgomery and Runger, Applied Statistics and Probability for Engineers, 5th edition from Wiley-India.

(Refer Slide Time: 01:37)



So where do we apply this hypothesis testing procedures, it is an important tool in decision making, the person making a decision maybe the manager, it may be the owner of a company or it may be even a researcher. What decision does the researcher make, he performs experiments finds the response as a function of the variables he changed during the course of experimentation.

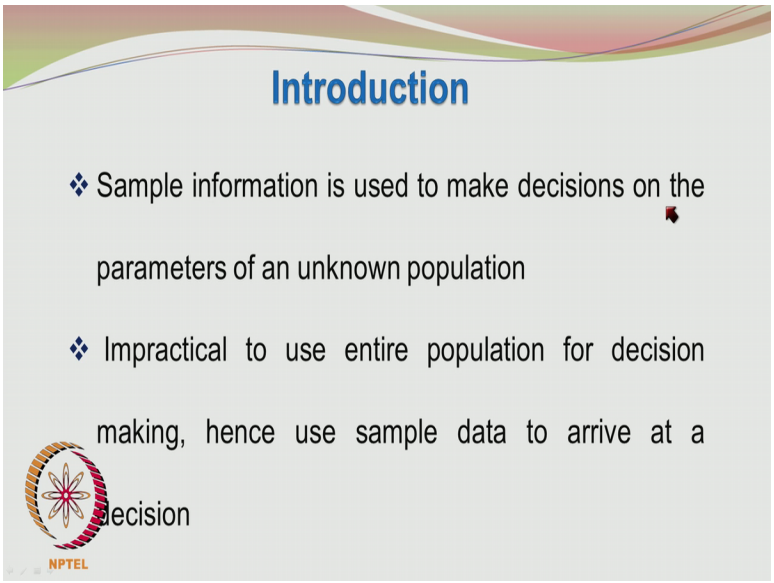
And the decision he makes are which of the variables are important and which of them do not have a bearing on the response of the process. So there can be different types of decision makers. So statistical inference involves hypothesis testing, I really have not told you what is meant by hypothesis, I will do so shortly and parameter estimation. A random sample is taken from the population and suitable statistics are obtained, again we are going to use the random sample.

Since the random sample is being used in so many places in so many different ways, it is important that we ensure that the sample we have taken is indeed random, and the sample elements are independent of one another, and follow the same probability distribution. Again we have to emphasize that we do not know the parameters of the population, we do not even know the nature of the population whether it is having a normal distribution or it is having a skewed distribution.

The random sample is used to make decisions on the sample distribution of the means, the sample distribution of the variances and so on, we have seen that the normal distribution, the t-distribution may be used in the tests carried out on the mean the enquiries made on the population parameter mu, we use the chi-square and the f distribution to make enquiries regarding the variances of the population, and also the ratio of variances.

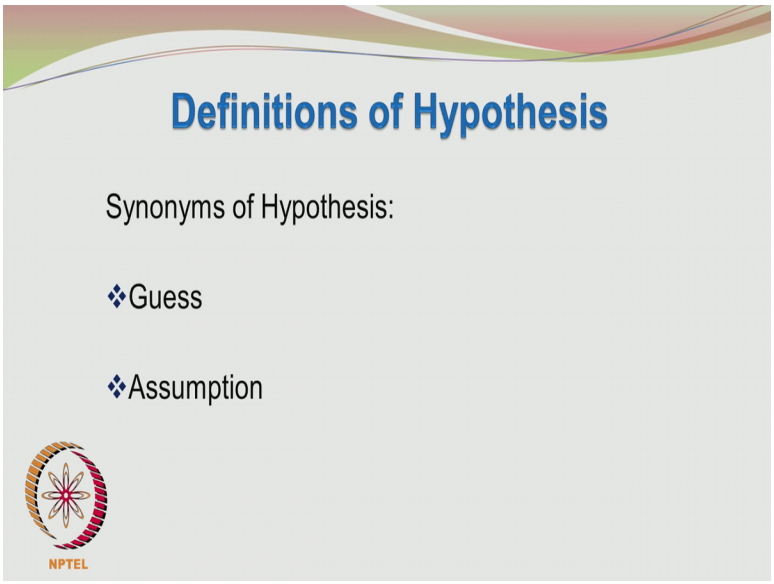
So the sample statistics in the form of sample mean and sample variance have a lot of significance. Another important thing is even though we are using the sample information we are always querying about the population parameters.

(Refer Slide Time: 04:46)

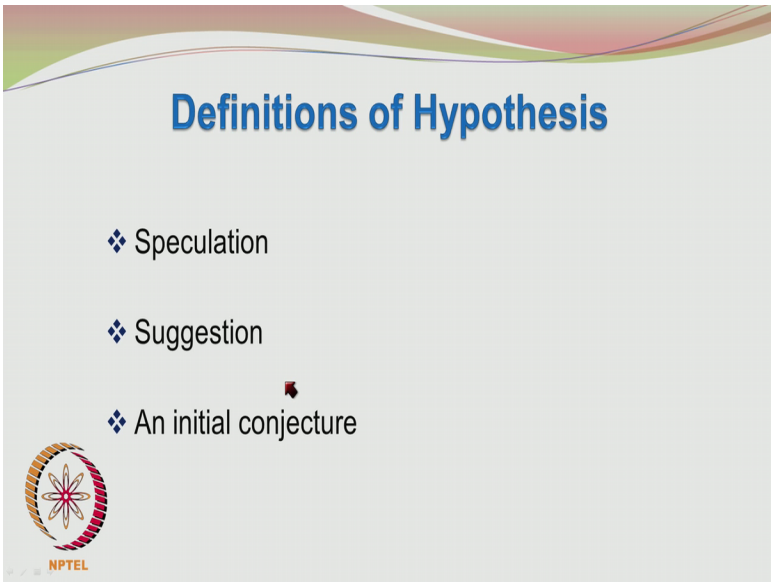


So the sample information is used to make decisions on the parameters of an unknown population, we do not know the mu and sigma square, and so we use the sample information as point estimates of the population parameters, and then we make certain decisions, how we make the decisions? We will see shortly. Obviously, it is impractical to use the entire population for decision making, so we are forced to resort to sampling and use a sample information to arrive at a decision, usually only one sample is taken from a given population.

(Refer Slide Time: 05:35)

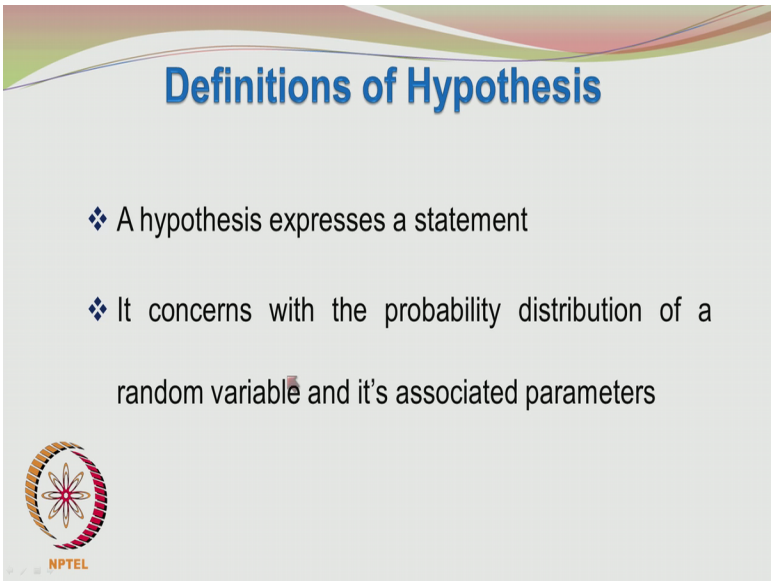


(Refer Slide Time: 05:55)



So I was just introducing the term hypothesis testing, what is really meant by a hypothesis? If you look at the dictionary some of these synonyms of hypothesis are guess, assumption, speculation, suggestion and initial conjecture.

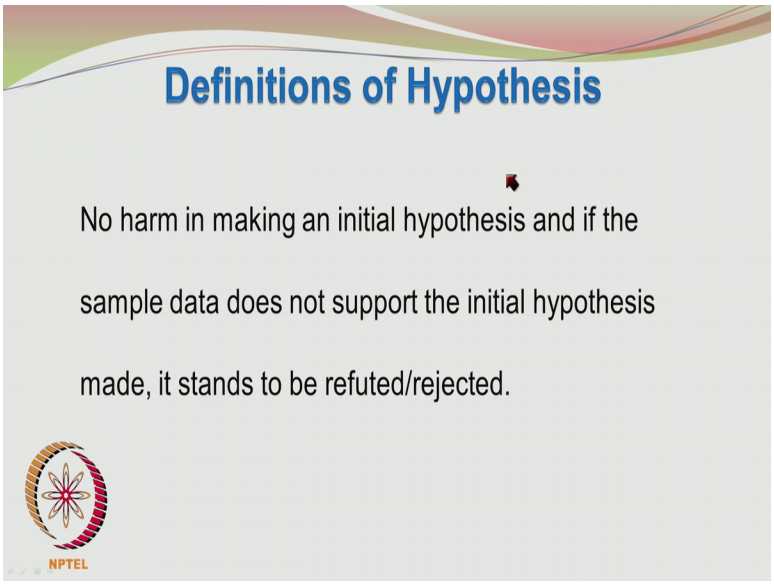
(Refer Slide Time: 06:00)



Basically, a hypothesis expresses a statement it concerns with the probability distribution of a random variable and it is associated parameters, so the hypothesis deals with the population parameters either mu or sigma square. Again I am emphasizing when you are writing down a hypothesis or a statement you are involving the population mean or the population variance, you are not using the sample mean and sample variance, when writing down the hypothesis okay.

Roughly, you can see that hypothesis will involve mu or sigma squared, it will not involve x bar and S squared in the hypothesis statement, even though the further analysis will be done with the sample mean and the sample variance, the x bar and S squared do not find a mention in the hypothesis statements. So I said hypothesis statement and then I am saying hypothesis statements, the reason for that is there are 2 hypothesis statements okay, and neither of them involved x bar or S squared both of them involved either mu or sigma square right.

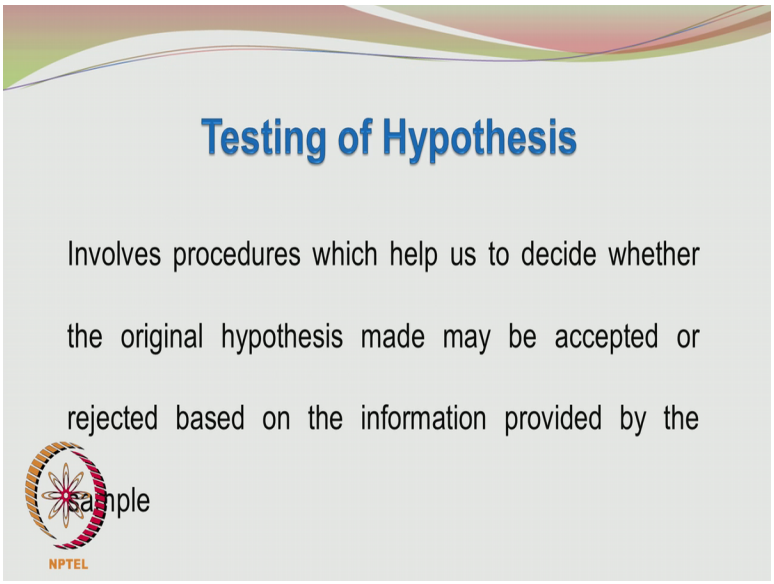
(Refer Slide Time: 07:41)



So doubt that maybe in everybody's mind is, we know alright hypothesis is a statement or a guess or a speculation, but it does not mean that all the people would make the same kind of hypothesis statements. Well the hypothesis statement is usually binary guilty or not guilty, it is having a population mean value of 50 units say 50% marks or it is either >50% or <50%, so the hypothesis statements are rather simple.

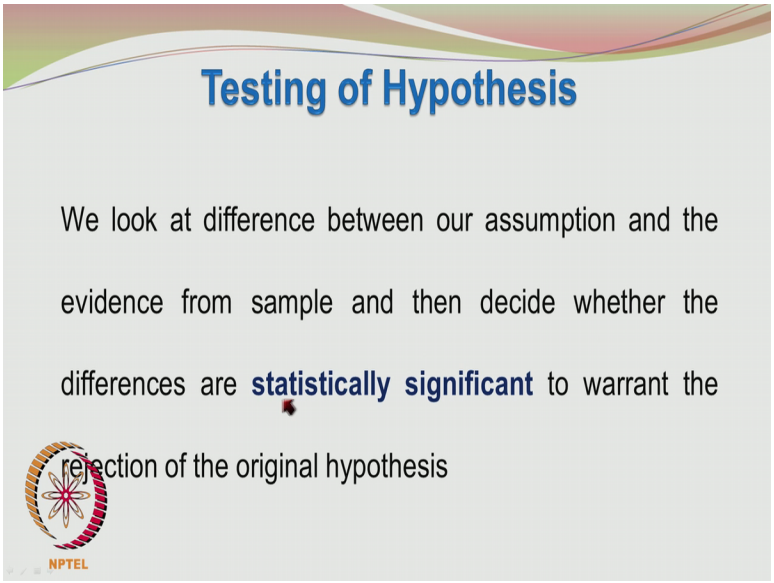
So you can go ahead and make a hypothesis statement without bothering too much, and if the sample data does not support your initial speculation or initial hypothesis, then the hypothesis statement stands to be refuted or rejected, then you can always go back and look at the problem history and problem data, and come up with the new hypothesis. The preliminary hypothesis test you carried out, will give you valuable clue as to what should have been the proper hypothesis statement.

(Refer Slide Time: 09:18)



So the hypothesis testing involves procedures which help us to decide whether the original hypothesis made may be accepted or rejected based on the information provided by the sample, so you are using the information given or available in the sample to make decision whether the original hypothesis statement is correct or not correct.

(Refer Slide Time: 09:55)



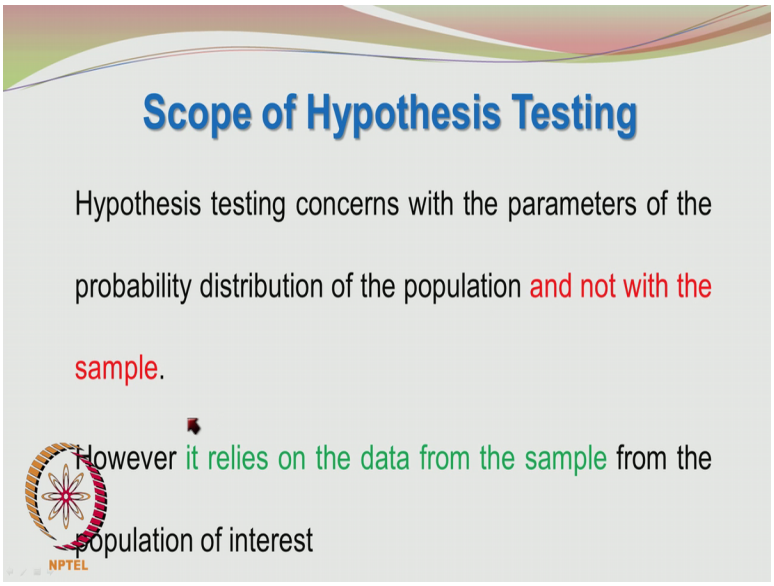
So we look at the difference between our assumptions and the evidence from the sample, and then to decide whether the differences are statistically significant to warrant the rejection of the original hypothesis. For example, you may be thinking that the nuclear reactor supplied by a company is going to give you a mean power output of 2.3 gigawatts, but the plant is actually giving a mean power output of only 2 gigawatts okay.

So you are having 2.3 gigawatts as the population parameter mu, whereas x bar is telling you it is only 2 gigawatts, so the difference is-0.3 gigawatts. So whether this difference is statistically significant enough to conclusively say whether the mean output from the reactor is 2.3 gigawatts is an incorrect statement or it is the correct statement. So coming again we look at the difference between the x bar and mu, x bar is 2 gigawatts and mu is 2.3 gigawatts, so the difference is -0.3 gigawatts.

So we can see whether this difference of -0.3 gigawatts is statistically significant enough through hypothesis testing procedures, if it is statistically significant enough then we can say that okay look this reactor which the company has supplied cannot give a mean power output of 2.3 gigahertz as originally claimed. However, if the test shows it is statistically insignificant.

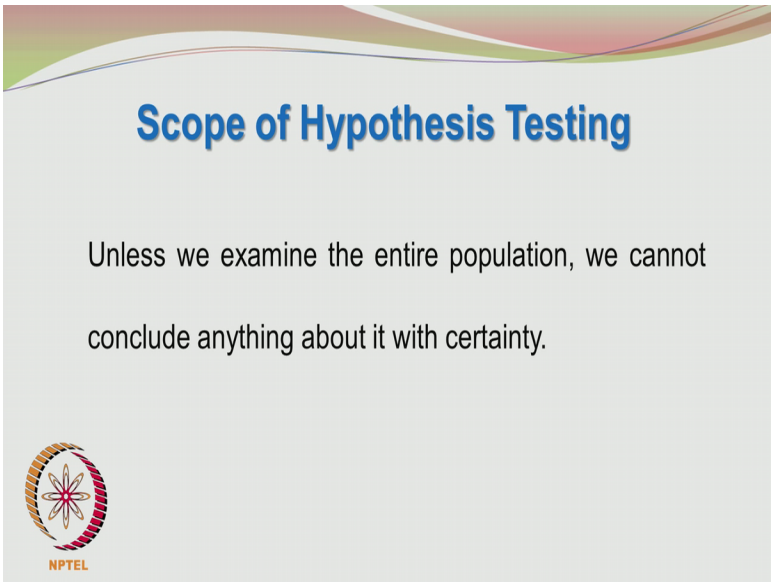
Then what we say is based on the data which is available to me I cannot really conclude that the reactor supplied by this company XYZ company is not supplying average power output of 2.3 gigawatts as claimed okay, there is insufficient evidence for me to actually reject the company's claim that the mean power output is 2.3 gigawatts.

(Refer Slide Time: 12:27)



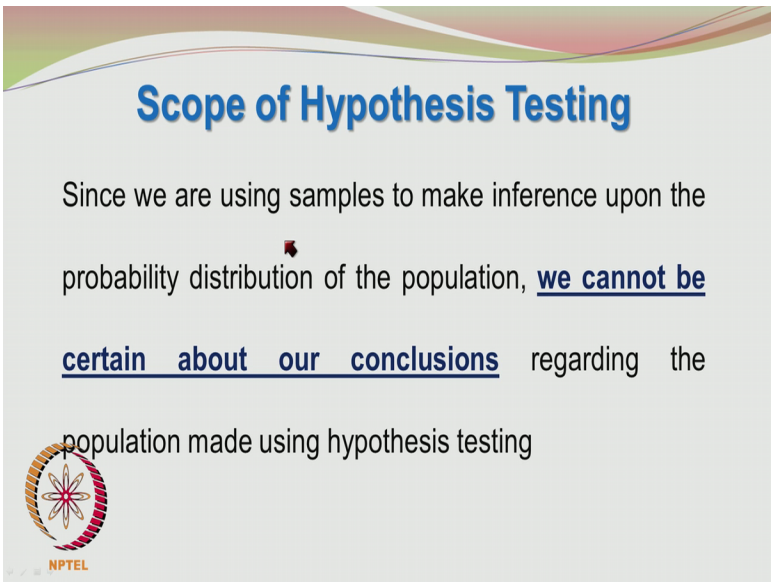
Again I would like to emphasize that hypothesis testing concerns with the parameters of the probability distinction of the population and not with the sample. However, it relies on the data from the sample from the population of interest.

(Refer Slide Time: 12:44)



Always please remember that the hypothesis testing is all subject to some uncertainty, unless we examine the entire population, we cannot conclude anything about the population with 100% accuracy or certainty.

(Refer Slide Time: 13:08)



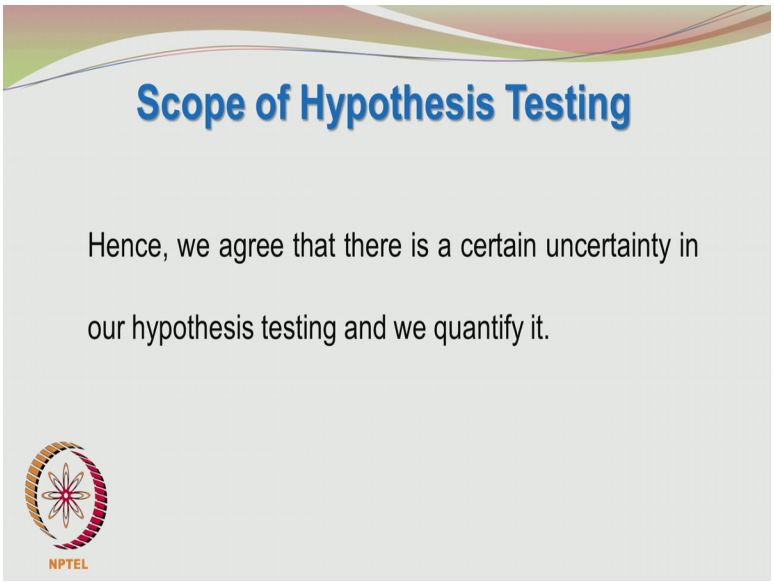
Since we are using samples to make inference upon the probability distribution of the population, we cannot be certain about our conclusions regarding the population parameters made using

hypothesis testing such a small typo here I will correct it okay. Since we are using samples to make inference upon the probability distribution of the population, we cannot be certain about our conclusions regarding the population parameters made during hypothesis testing.

We are speculating on the parameters of the population, can the mean value be this much, can the variance be this much, and these are the population parameters values. So after we take the information or evidence provided by the sample, we do some test procedures and then we make a conclusion, based on the evidence provided by the sample I cannot conclude that this population parameter mean can be 50 units.

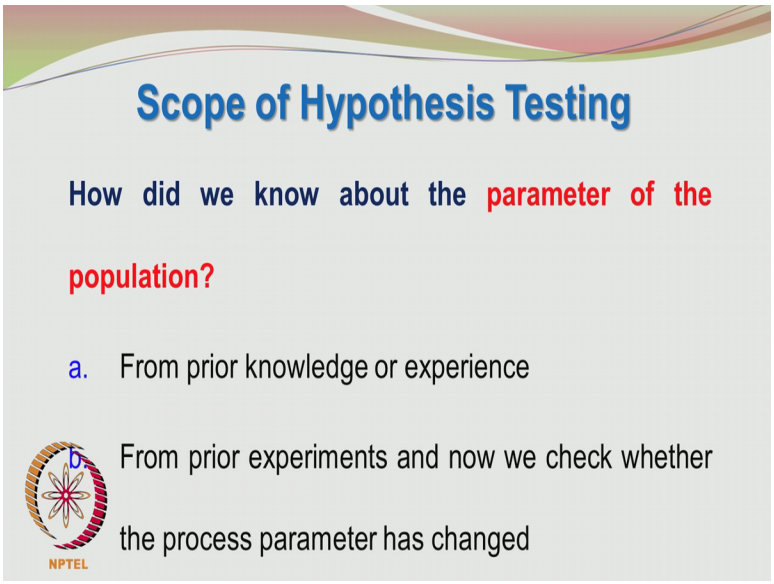
Or I can say that right based on the evidence provided by the sample, I can sample or accept the original problem statement that the mean is indeed 50. So you are making a decision based on the evidence provided by the sample, but this position cannot be considered to be conclusive or 100% fail proof.

(Refer Slide Time: 14:53)



Hence, we agree that there is certain uncertainty in our hypothesis testing and we quantify it.

(Refer Slide Time: 15:02)

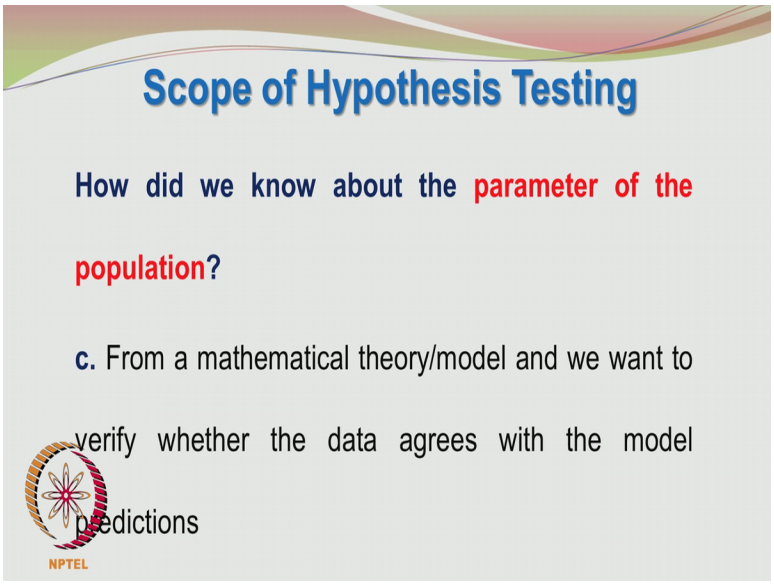


So whenever we are speculating on the population parameter, the mean of the population can be this much or the variance of the population can be this much, on what basis we are doing the speculation. We make these estimates or assumption regarding the population parameters from prior knowledge or experience, from prior experiments, and now we check whether the process parameter has changed.

Suppose we are having a machine which is operating at a certain power, other settings of the machine are also fixed it is giving a average particle size of 50 millimeters, so we are using the machine for a long time under has been working very well, and so the average is very very close to 50 millimeters. So we can say that mean=50 millimeters, suppose some modification have been made to the machine, maybe it went for maintenance check or some part was replaced.

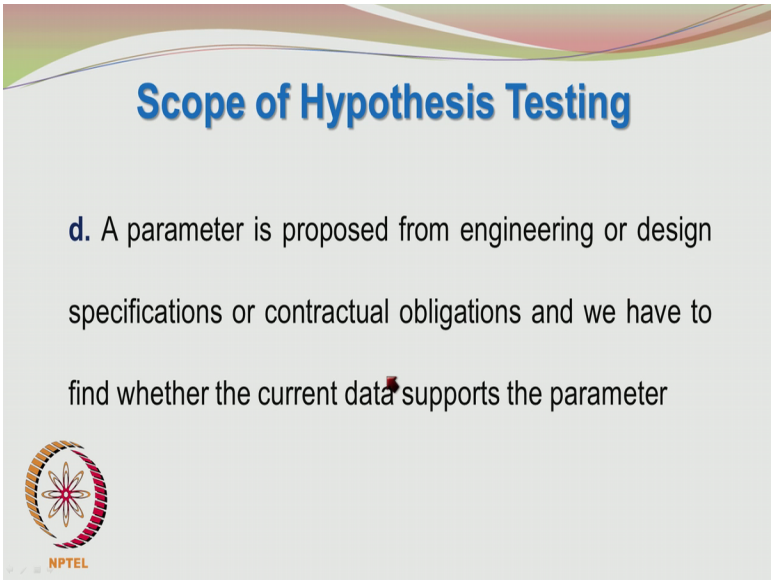
Then again the machine is grinding the raw materials and providing particle sizes. Now we want to know whether the previous mean value of 50 millimeters is again met by the machine, or because of the process modifications it is producing particles which are lower than the original mean value or greater in size than the original mean value. So we know from prior knowledge or experience, from prior experiments we have collected our accumulated large amount of data. And after some modifications we want to see whether the process parameter has changed.

(Refer Slide Time: 17:23)



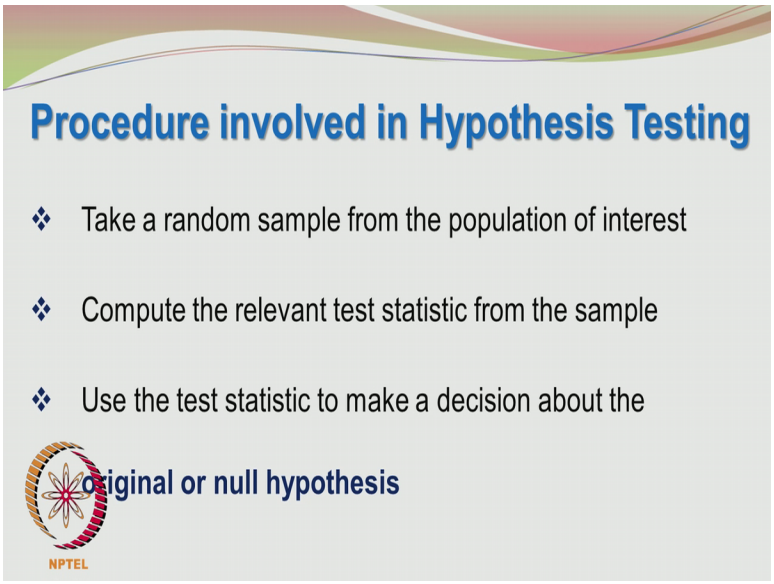
Also the parameter maybe set based on a mathematical theory or a model, and we want to verify whether the data agrees with the model predictions.

(Refer Slide Time: 17:42)



A parameter is proposed from engineering or design specifications or contractual obligations, and we have to find whether the current data supports the parameter. For example, the XYZ Company was guaranteeing 2.3 gigawatts from its nuclear reactors as the average power output, so that is based on contractual obligations or design specifications, and then we actually monitor the reactors performance and take a random sample and find the mean power output. We can check whether the mean power output of the reactor is actually 2.3 gigawatts or lower.

(Refer Slide Time: 18:32)

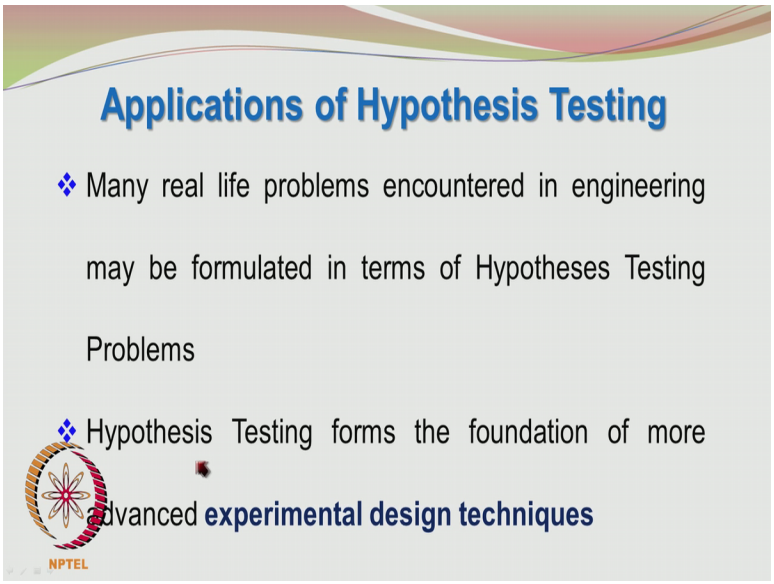


So what is the procedure involved in hypothesis testing, we first take a random sample from the population of interest, again this is the most crucial step sometimes its importance is underestimated, you cannot blame the final decision or the final outcome, if the sampling was not done properly. Compute the relevant test statistic from the sample, the relevant test statistic we are familiar with so far are the sample mean and the sample variance.

We really have not looked into any other test statistic, but the sample mean and sample variance would do just fine for us including our analysis of design of experiments. Well you must be curious what are we doing here now mean where is it going to tie up with design of experiments, I am eager to learn design of experiments, I would suggest you to be a bit more patient we will be definitely looking at design of experiments very shortly.

And you will really appreciate all the concepts we have learnt so far when you see them being applied in design of experiments. Use the test statistic to make a decision about the original or null hypothesis, so I have introduced the new term original or null hypothesis. Let us now get into the defining of original or null hypothesis.

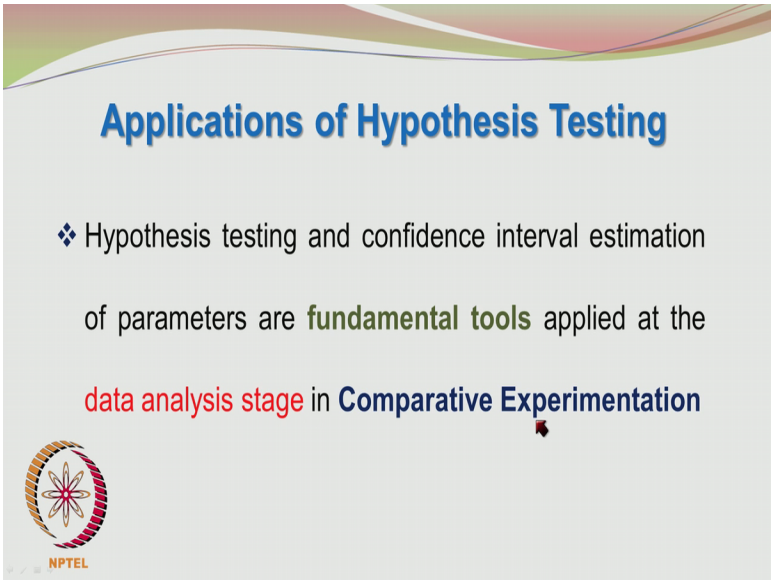
(Refer Slide Time: 20:23)



So many real life problems encountered in engineering may be formulated in terms of hypothesis testing problems. Hypothesis testing forms the foundation of more advanced experimental design techniques, that is what I told you just a short while back that whatever we are learning now hypothesis testing, the t-distribution, Chi-square distribution, the f distribution will find immediate applications in the design of experiments.

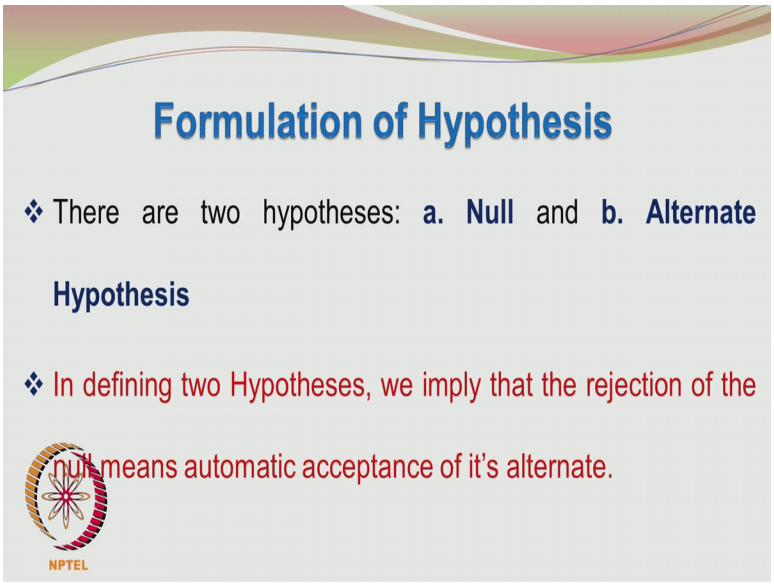
In design of experiments as well as in linear regression tools, we will be extensively seeing the application of the t-test, the f test and also the 95% confidence intervals.

(Refer Slide Time: 21:18)



So hypothesis testing and confidence interval estimation of parameters are fundamental tools applied at the data analysis stage in comparative experimentation.

(Refer Slide Time: 21:27)



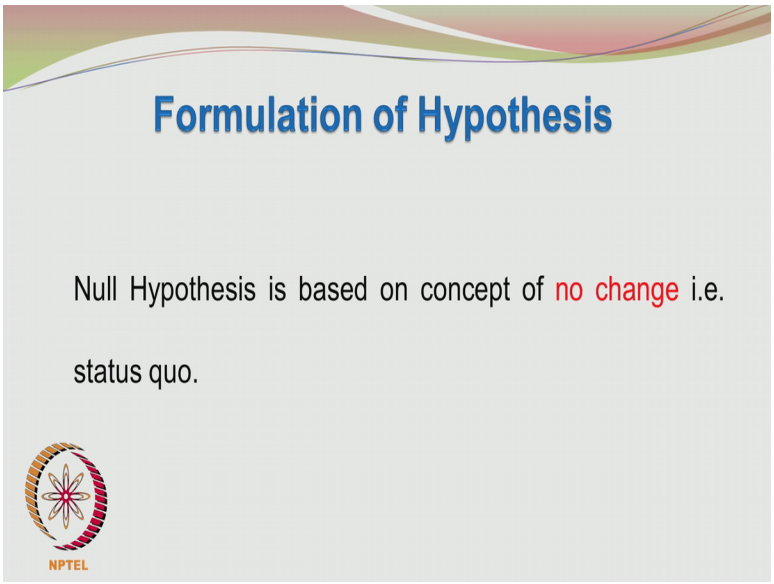
There are 2 types of hypothesis, one cannot exist without the other. The first hypothesis which is usually made or initially made is the null hypothesis, and the second hypothesis is its contradictions the alternate hypothesis. In defining 2 hypothesis, we imply that the rejection of the null means automatic acceptance of its alternate. So why do we make the 2 hypothesis statements, so any null original or initial statement or hypothesis is a speculation.

So we cannot claim that it is 100% accurate, to allow for the existence of an alternative to the original statement we have made, we propose or state the alternate hypothesis. There are different ways of disagreeing, the one way is a kind of neutral manner saying that I do not agree with you. The second way of saying is what you are saying is not correct actually, the performance is better than what you are saying.

And the third way of rejecting the original statement or rebutting the original statement is the performance is lower than what you are claiming. So there are different ways to disagree one is simply saying it is different or being bit more specific saying it is better than or less than. So normally we are not talking here of philosophical arguments, we are talking about entities that may be quantified into numbers and mathematical formulae.

So there are 2 hypothesis, one is the initial original or the null hypothesis, and the second is the alternate hypothesis which is so stated that it contradicts the null hypothesis.

(Refer Slide Time: 24:12)



How do you propose the null hypothesis? This is a very interesting issue here, you are going to a company after finishing your B.Tech or M.Tech, you are very enthusiastic and bubbling with energy, and you find the process being done in the industry. And you feel look this process is not being done correctly, if I made this modification the process will run much more smoothly or much more efficiently.

Then you go and tell it to the management the management of course welcomes new ideas, but that would mean also a lot of commitment from their side in terms of manpower, time, money, and they also do not know how the customers or the client would react to it. So the management would be rather inclined to let things continue as they were, probably not as efficiently as you are considering, but there is no immediate shutting down of the planet or revamping of the process.

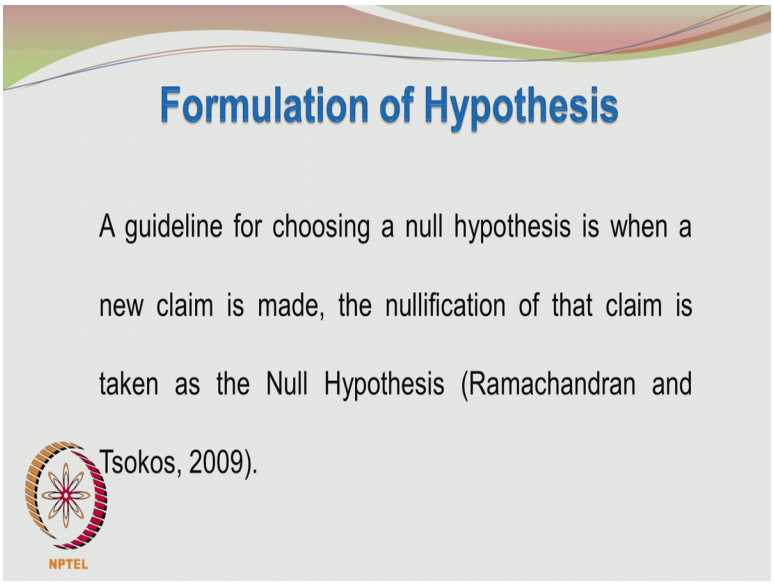
So things will continue to run as usual, so you have to provide a strong convincing evidence for them to stop everything and make the process modification. Even when you were doing experiments in the laboratory for your M.S or Ph.D. research program, you have to be skeptical, you have to assume that none of the variables you are investigating is going to affect your

process response. Well this is a sort of in contradiction or opposition to our aim of doing the experiments.

We want to clearly show that some of the variables or all the variables we are investigating is impacting the process in a strong manner, but a true researcher will always be a skeptic, he will have to say initially that there is no change or status quo is being maintained. Even if you change the variables, it is not going to affect the response, of course this may not be true. Because obviously when you change something there is going to be an effect on the response.

Judge sitting in the court may have to make the most important decision of all, whether to allow a person to continue to live or to be sent to be gallows, because he has apparently committed a crime. So the judge's initial attitude will be look this person is innocent, the prosecution has to provide evidence beyond the reasonable doubt, so that person may be convicted. So the null hypothesis is sort of stabilizing one, it keeps us grounded and it is based on the concept that there is no change or status quo is being maintained.

(Refer Slide Time: 27:59)

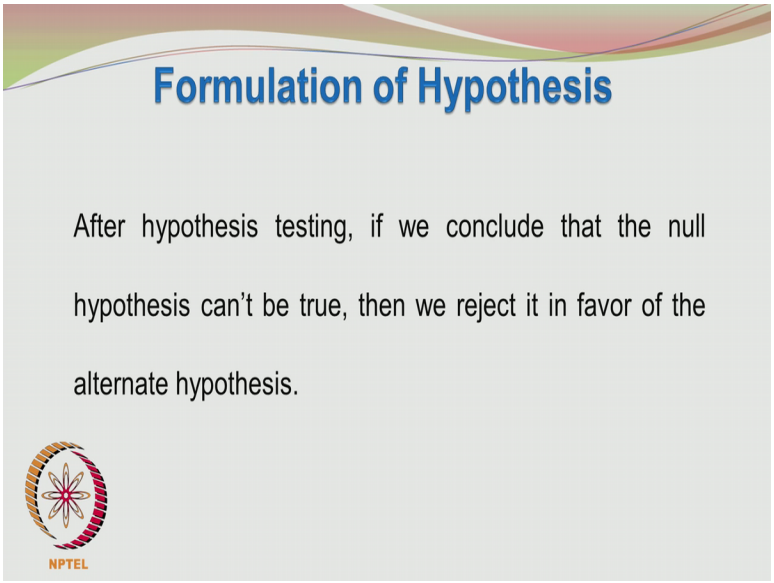


A guideline for choosing a null hypothesis is when a new claim is being made, the nullification of that claim is taken as the null hypothesis, according to Ramachandran and Tsokos. The nullification or the contradiction of the claim is taken as the null hypothesis okay, so the claim is

I am going to increase the plant performance by 10% thereby making you a profit of 1 million dollars per annum.

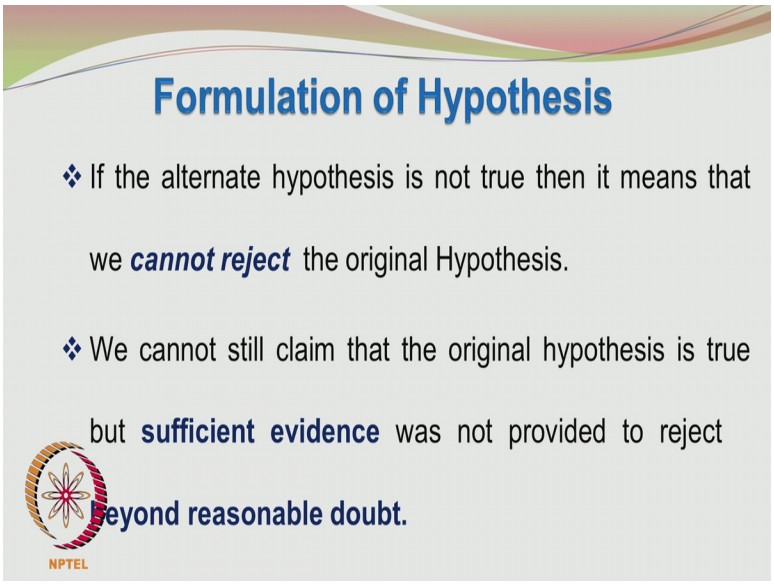
So there is a new claim being made the management is cool to the idea and say that look your claim is not going to make a difference, the plant will continue to run as before. So the null hypothesis is always the nullification or the neutralization of the claim that is being made.

(Refer Slide Time: 29:08)



Anyway we cannot live in speculation, we have to try out different ideas, and then we have to come to the conclusion. After a hypothesis testing, we conclude that the null hypothesis cannot be true, then we rejected in favor of the alternate hypothesis.

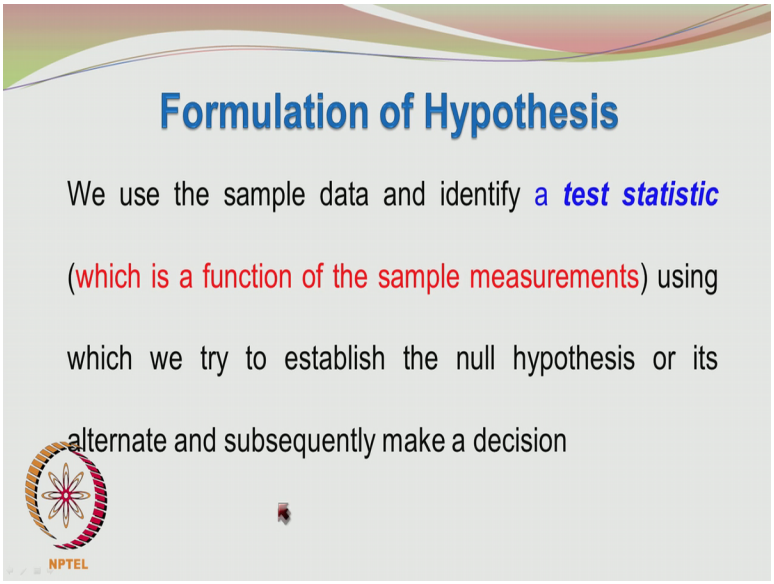
(Refer Slide Time: 29:35)



If the alternate hypothesis is not true, then it means that we cannot reject the original hypothesis. We are not making an absolute statement that the null hypothesis is correct. We have to say perhaps in legal terminology that enough evidence was not provided to establish beyond a reasonable doubt that the null hypothesis is wrong. I am repeating, enough evidence that could established beyond the reasonable doubt that the null hypothesis is wrong was not provided.

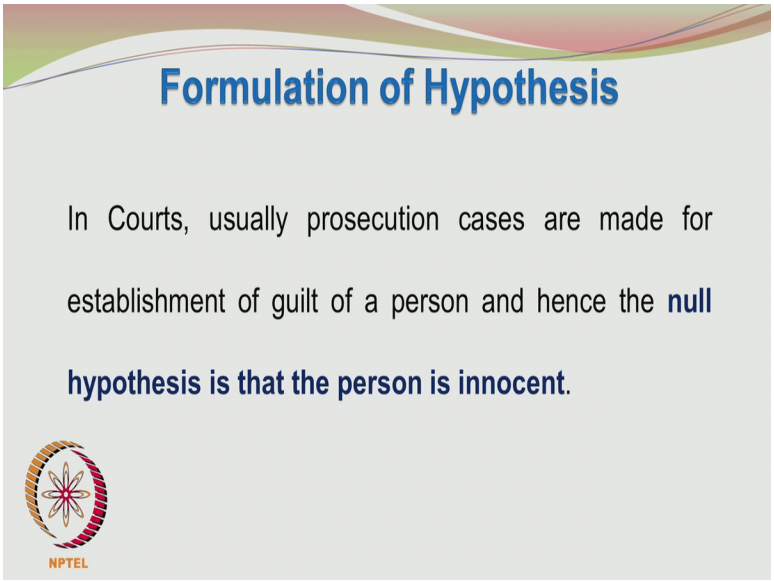
Hence, there is no sufficient evidence to reject the null hypothesis, we are not claiming or we are not stating that the null hypothesis is true, we are only saying that sufficient evidence was not provided to reject the null hypothesis.

(Refer Slide Time: 30:58)



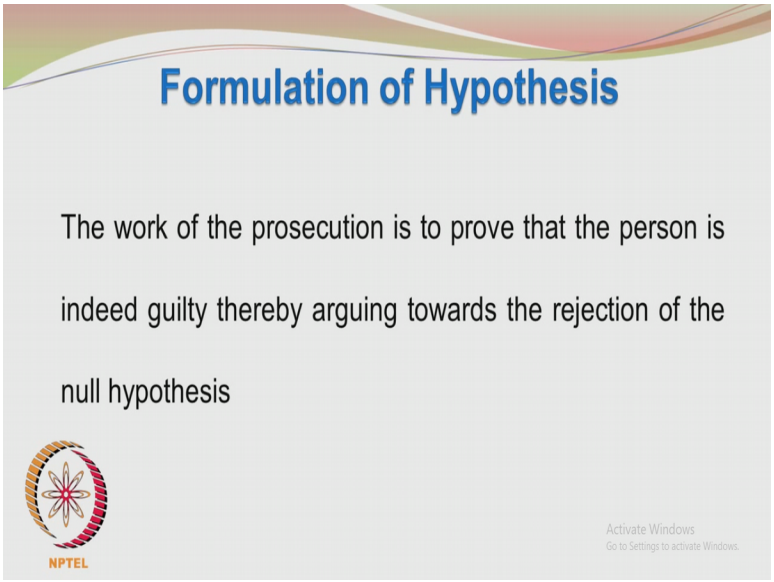
We use the sample data and identify a test statistic which is the function of the sample measurements using which we tried to establish the null hypothesis or it is alternate and subsequently make a decision. So far we have seen, what is meant by a null hypothesis and what may be the forms that may be taken by the alternate hypothesis, now we have to see how to actually go about carrying out the hypothesis test procedure.

(Refer Slide Time: 31:35)



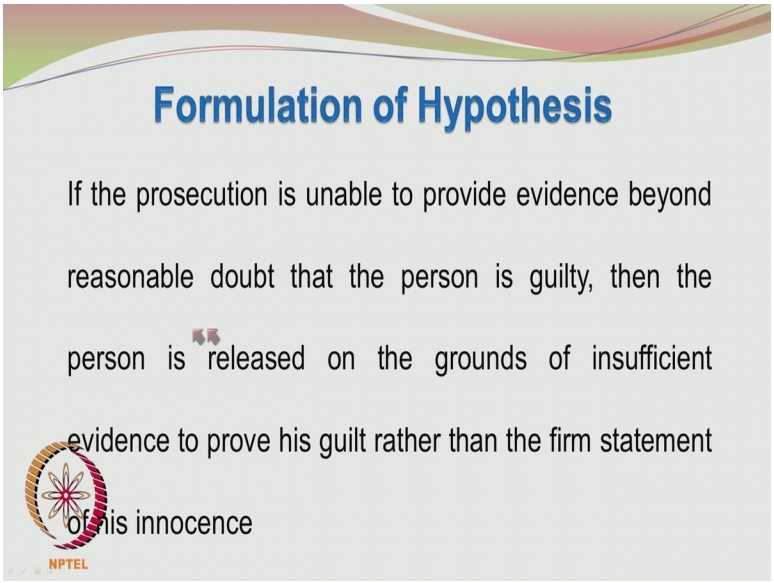
So when you formulate the hypothesis statements we can take an example from the legal side, so as I said previously the null hypothesis is the person is innocent.

(Refer Slide Time: 31:52)



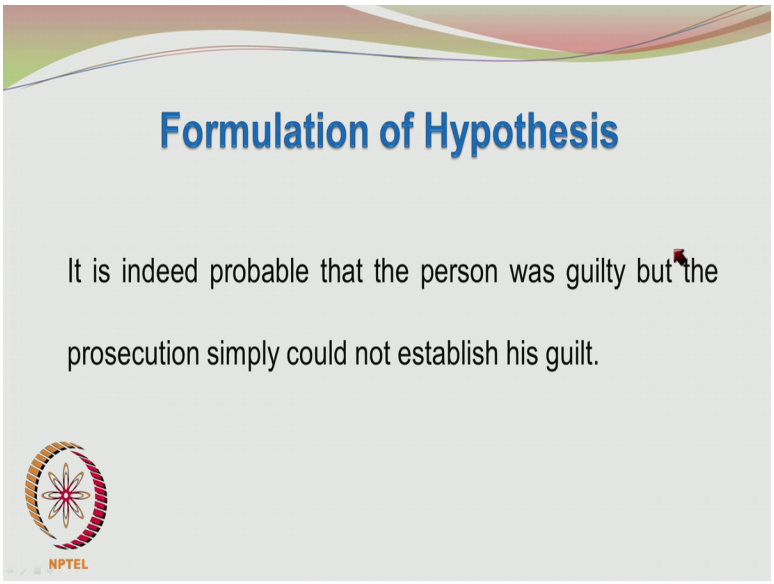
The prosecution vehemently wants to prove the case, and send the charged person to jail, they have to come up with suitable and clinching evidence. So they are working towards the rejection of the null hypothesis. The management is making a null hypothesis that the process modification is not going to make a difference, whereas you the originator of the idea will be actively looking for suitable evidence to indeed prove that your concept is correct, and the null hypothesis can be rejected.

(Refer Slide Time: 32:39)



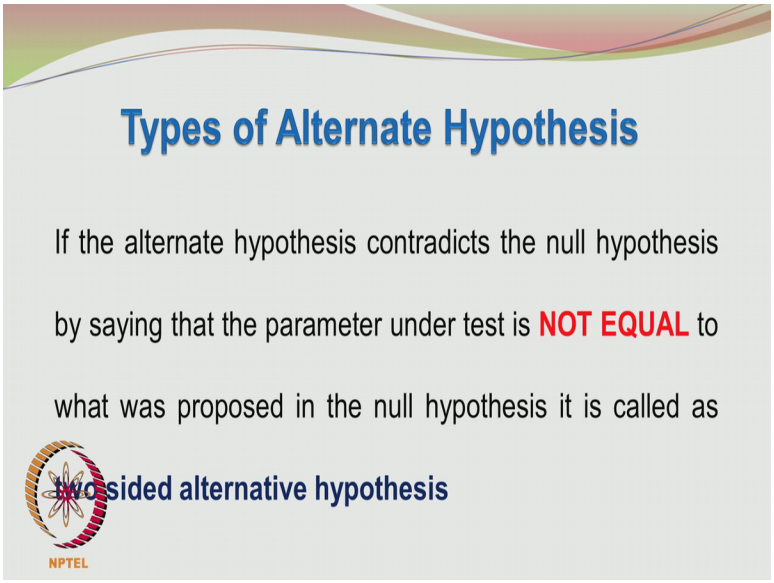
So if the prosecution is unable to provide evidence beyond reasonable doubt that the person is guilty, then the person is released on the grounds of insufficient evidence to prove his guilt rather than the firm statement of his innocence.

(Refer Slide Time: 32:57)



It is indeed probable that the person was guilty, but the prosecution simply could not establish his guilt.

(Refer Slide Time: 33:01)

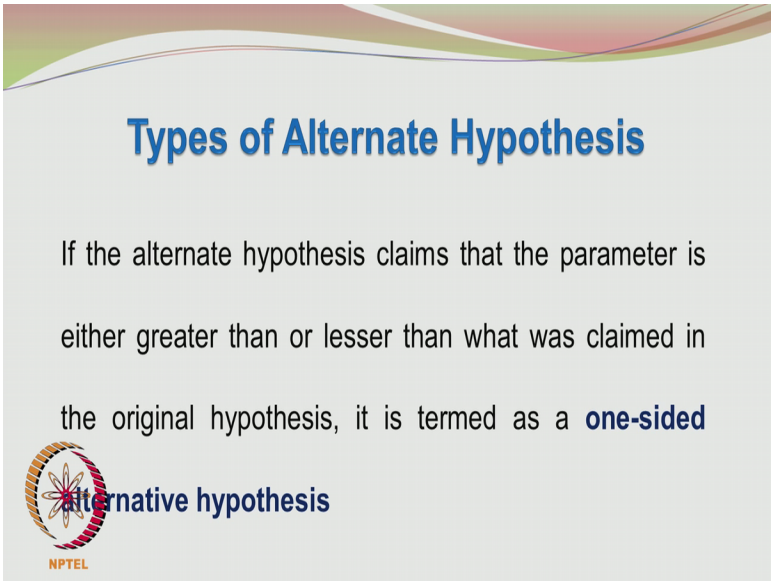


So there can be different types of alternate hypothesis, if the alternate hypothesis contradicts the null hypothesis by saying that the parameter under test is not equal to what was proposed in the all hypothesis, it is called as a 2 sided alternative hypothesis. The null hypothesis proposes a statement that the population parameter is equal to a certain value, the alternate hypothesis considers the possibility that based on the evidence provided the population parameter may be less than the proposed value or greater than the proposed value.

Because we are talking about random phenomena, when there is a negative deviation on one occasion, on another occasion there may be a positive deviation. So in order to account for such kind of deviations, we are using the alternative hypothesis in terms of not equal to, the original null hypothesis may say that the population mean value is 50. The alternate hypothesis maybe the population mean value mu !=50, under such situations we allow for the possibility that the population mean may be >50 or <50.

So we are allowing for the 2 sided possibility, so it is called as a 2 sided alternative hypothesis.

(Refer Slide Time: 34:48)



If the alternative or alternate hypothesis claims that the parameter is either greater than or lesser than what was claimed in the original hypothesis, it is termed as a one sided alternative hypothesis. For example, in the nuclear reactor case, you are having a null hypothesis of mu=2.3 gigawatts, the sample mean is showing you value of 2 gigawatts, so the company of the industry which is using that react sees red and claims that no the parameter value of mu of 2.3 gigawatts is not correct, the actual average power output from the reactor is <2.3 gigawatts.

It will not even consider the possibility that the mu average power output can be >2.3 gigawatts, it will say it will be <2.3 gigawatts. So this is the one sided alternative hypothesis using the less than sign in the alternate hypothesis statement. Alright, what about the alternate hypothesis

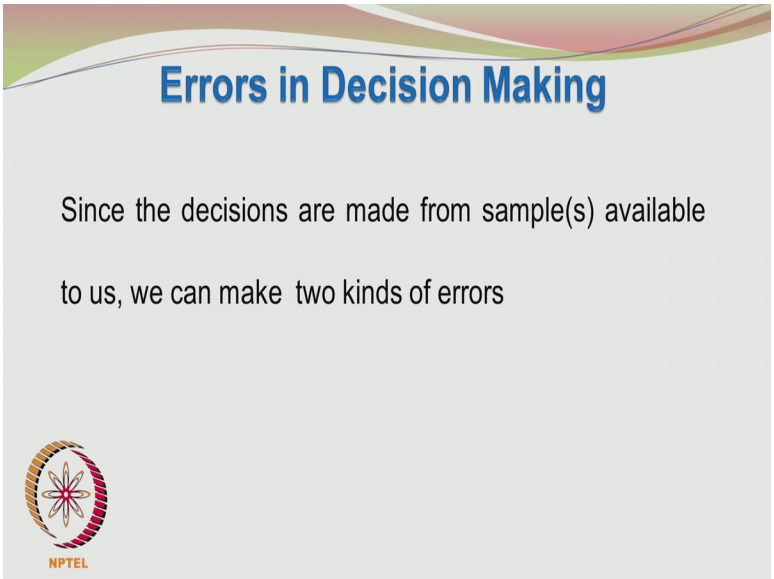
involving greater than. For example, you are monitoring the air quality in a particular place where recently that has been lot of industrial activity.

So the null hypothesis which the new industrial companies will be interested in this that there is no change in the pollution levels, the average pollution levels are the same as they were originally before the industrial activities started. I am using the advanced or state of the art pollution control measures, so I am not letting out really any toxic gases to the atmosphere, so the mean population levels of the pollutants in the ambient air is the same it is unchanged.

But the pollution monitoring agency maybe more interested in proving that the mean value cannot be what was originally, it has indeed gone up. In which case you are going to use the greater than sign in the alternate hypothesis. So the null hypothesis will be mu=mu 0 the original or the base line value, the alternate hypothesis for the industrial pollution problem is mu>mu 0.

So after the industrial activity has started there has been an increase in the pollution levels that is what the environmental pollution agency will be looking for, it is really not going to be interested in the possibility mu<mu 0, it is interesting only in mu< mu 0. So again we will be going in for a one sided the alternative hypothesis, and that would involve the sign of greater than in the present case okay.

(Refer Slide Time: 38:28)

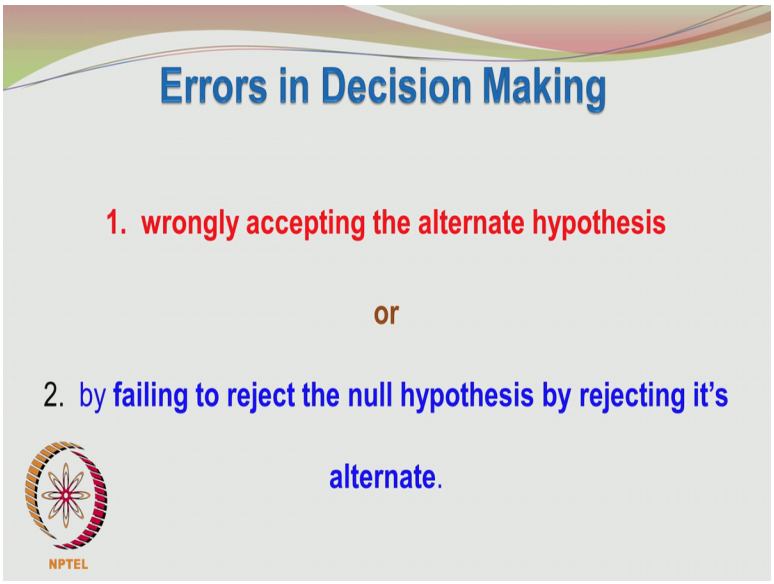


Okay, so after the judge has passed a decision regarding the innocence of the accused okay either way he may found him not guilty or may have found him guilty. So there will be always nagging question did I released the person wrongly or did I send an innocent person to the person, so these kind of doubts will always be there. So whenever we have made a decision with based of intentions we still may have the doubt that our decision was not correct.

So how to quantify these kind of errors in decision making. From a common sense point of view, we make a decision after giving a suitable margin okay, so that we do not make the wrong decision okay, the margin may be quite lenient. For example, if students are writing an exam, our expectation would be the students appeared at least 60% marks to pass the course, for example the exam is so easy in my opinion they should get at least 60% if they have really understood something from this source and then pass it.

But we set a pass mark of 40% just to be on the safe side, so that no person who has reasonably understood the course is failed.

(Refer Slide Time: 40:31)



So the errors in decision making are wrongly accepting the alternate hypothesis, so the null hypothesis in fact was correct, but you have accepted the alternate hypothesis wrongly. The second one maybe failing to reject the null hypothesis by rejecting its alternate, which of these 2

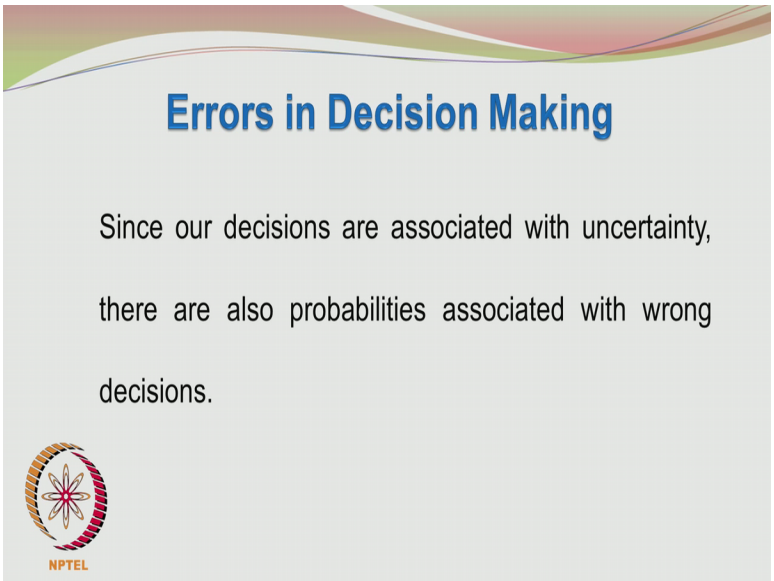
is a more serious error. Let us again take the court case, and the first error wrongly accepting the alternate hypothesis would mean the person was in fact innocent.

But the judge based on the evidence presented to him wrongly accepted the prosecution arguments and sent the innocent person to jail that is a very serious issue. Or the management got carried away by your technical presentation and accepted your idea spent a lot of money, and unfortunately the idea did not really work okay, that was not any significant or noticeable improvement in the process after the modifications were carried out okay.

The second error is to fail to reject the null hypothesis by rejecting its alternate, so what it really means is the person who was indeed a guilty, but the judge released him okay, so what really happened was the guilty person got away free. When you compare the 2 types of errors, the first one is pretty serious okay, the second one is also an error but it is not as serious as the first error. The first error an innocent person was sent to jail or he was hanged or whatever.

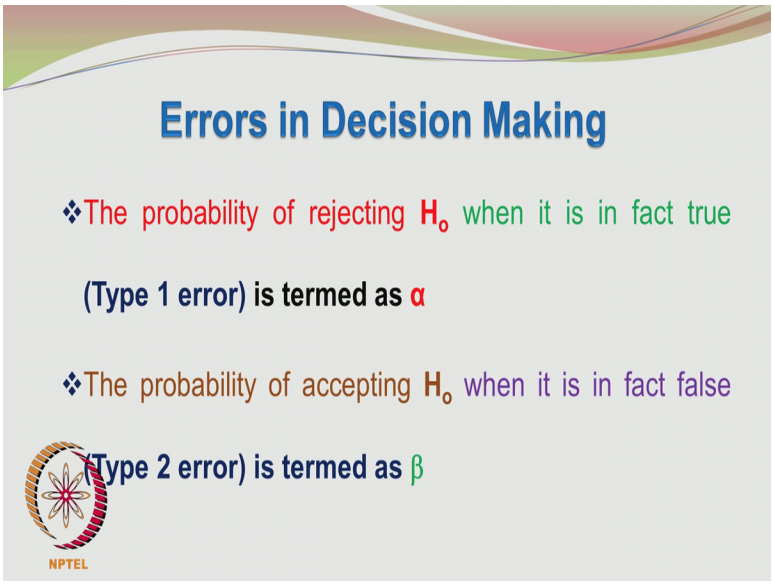
And in the second case the guilty person got away, so which of these 2 is more serious you yourself will be able to answer.

(Refer Slide Time: 43:05)



So our decisions are associated with the uncertainty, there are also probabilities associated with wrong decisions.

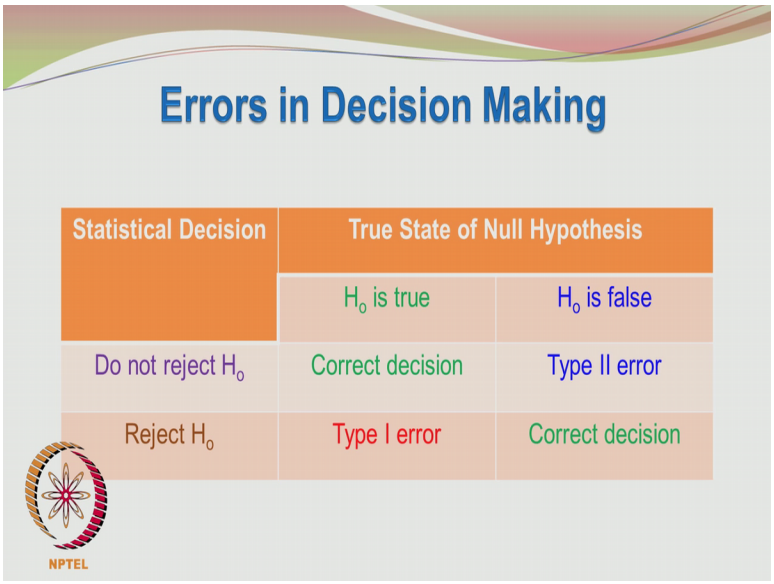
(Refer Slide Time: 43:14)



The probability of rejecting the null hypothesis, when it is in fact true is called as the type 1 error and it is termed as alpha, we have seen alpha earlier when we were constructing the confidence intervals. The probability of accepting at H0, when it is in fact false is the type 2 error and it is termed as beta. So probability of rejecting H0 when it is true is called as the type 1 error and it is denoted by alpha.

The probability of accepting H0 when it is in fact false is called as the type 2 error and it is termed as beta, maybe I should write the probability of failing to reject H0 when it is in fact false is called as type 2 error.

(Refer Slide Time: 44:30)



So when we come back after a small break, we will be looking at this table, and we will see after a few minutes.