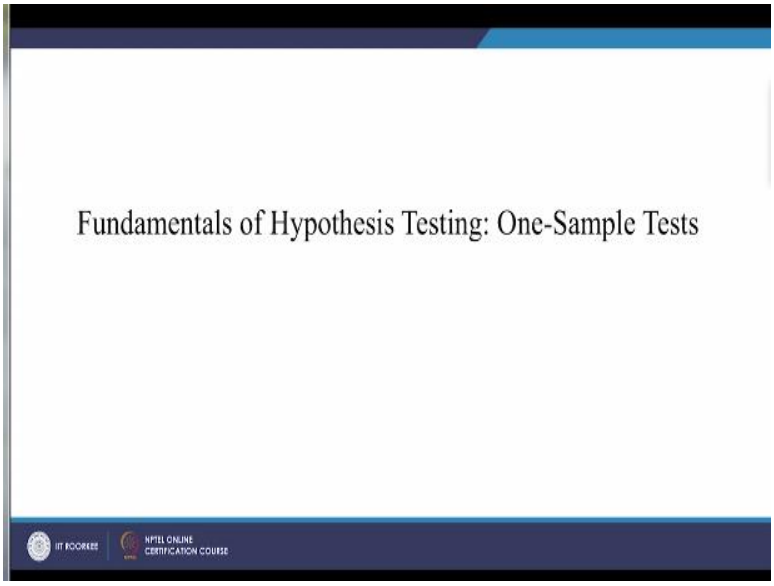
Business Statistics
Prof. M. K. Barua
Department of Management Studies
Indian Institute of Technology-Roorkee

Lecture-30
Sampling and Sampling techniques

Hello friends, I welcome you all in this session, as you are aware in previous session we were working out couple of examples on estimation. As I have told you that statistics is of two types, the first one is descriptive statistics and the second one is inferential statistics, when we talk about inferential statistics it has got two types again you can have estimation and the other one hypothesis testing.

So when I talk about inferential it means we are trying to find out something about population through sample. So that is what we have seen in estimation, we did estimate, population, mean through sample mean, we did estimate population proportion through sample proportion. We have looked that how to calculate even sample size in case of estimation of population mean and population proportion.

(Refer Slide Time: 01:42)



So, will continue with inferential statistics and today's topic is on hypothesis testing, we will get first one sample test. So, you can have 1 sample test, 2 sample test right, so let us look at what is hypothesis. But before defining hypothesis and how to test a particular hypothesis there are

several situations about which the decision makers want to know or there are several questions for which the decision makers want to know the answers.

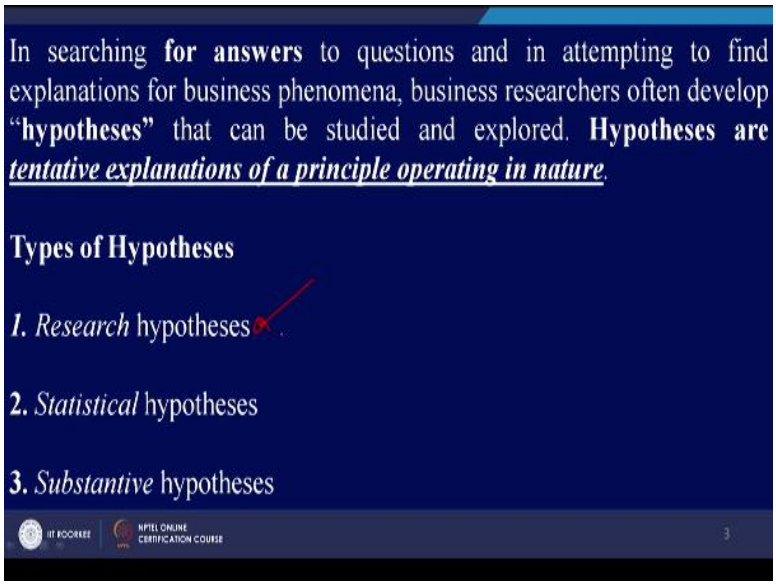
(Refer Slide Time: 02:24)

So, let us take one of the questions is what container shape is most economical and reliable for shipping a product. So, you can have different types of shapes maybe a cylindrical one or a conical one or some other shape right. So, it depends on how for the product is to be transported from one place to another, what is the nature of the product and so on right.

If it is liquid then you need a container of different shape, size, if it is gas then it is you need different type of container right. So this just 1 question let us say which management approach best motivates employees in the retail industry. So there are several approaches, so for is management is concern to motivate employees right. So in retail industry or let us say in banking industry or an insurance industry or automobile industry and so on right. How can the company's retirement investment financial portfolio be diversified for optimum performance.

So, let us say if you have got some money and you want to invest in different portfolios, so where you should invest it right. So, this again one more decision, what is the most effective means of advertising in a business to business setting. There are n number of media available these days, so where you should advertise your product it. So, these are couple of questions which decision makers want to know.

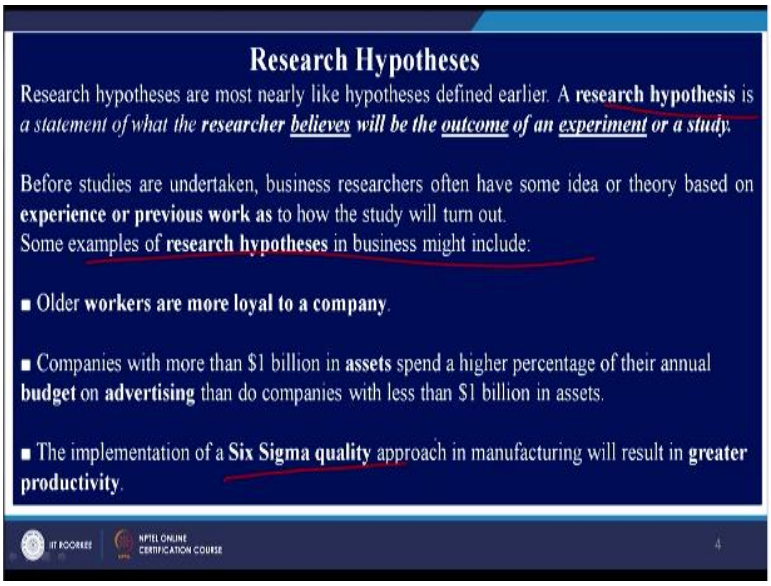
(Refer Slide Time: 04:04)



So, when you search or when decision makers search for answers to questions they often take help of business researchers and business researchers often develop something called hypothesis. And hypothesis is nothing but hypothesis is tentative explanations of a principle operating in nature. So hypothesis is something about which is there is in the nature, it is about the existing theories and practices, it is about current situation.

So what is hypothesis, hypothesis is tentative explanation of a principle operating in nature. There are different types of hypothesis, the first is research hypothesis, the second is statistical hypothesis and third one is substantive hypothesis, let us look at what is research hypothesis.

(Refer Slide Time: 05:20)



So, the research hypothesis is similar to what we have just defined, what we have defined that it is tentative explanation of principle operating in nature right. So, a research hypothesis is statement of what the researcher believes will be the outcome of an experiment. You see what happens decision makers on the basis of their experience they know the outcome of a particular action.

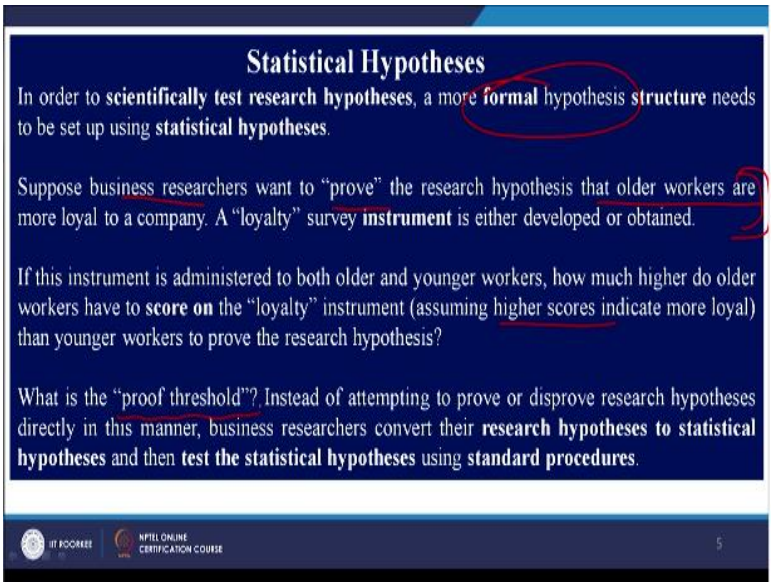
So rather than taking action and looking at output or examining output decision makers and business researchers they know what would be the possible outcome of the experiment. So, without doing experiment they know what would be the possible outcome. So research hypothesis what, is statement of what the researcher believes will be the outcome of an experiment or of a study.

And as I said just on the basis of their experience and their knowledge and previous work they know what would be the possible outcome. So let us look at some of the examples of research hypothesis. Let us say whole workers are more loyal to company in comparison to the new workers right, this is just a research hypothesis. Companies with let us say asset more than 1 billion dollars will spend more on let us say advertising.

Their advertising budget would be more compare to those companies whose assets are let us say less than 1 billion dollars right. If we implement 6 sigma then definitely the productivity would

improve, so this is again one more research hypothesis. Now we do not know whether this hypothesis is true or not we are just giving a research hypothesis.

(Refer Slide Time: 07:27)



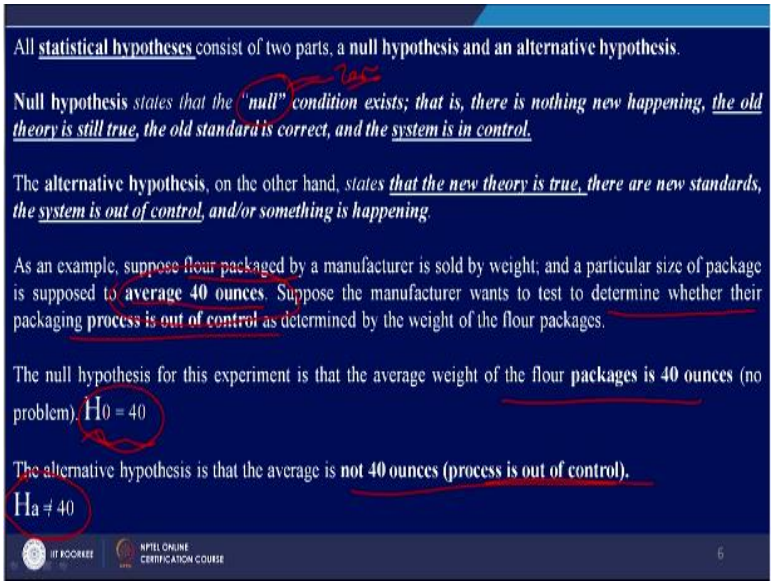
So, we have seen research hypothesis, what is research hypothesis, research hypothesis is something which a researcher believes that this would be the outcome of an experiment. Now what is statistical hypotheses, statistical hypotheses similar to the research hypothesis but the statisticians make little changes in research hypothesis from statistical point of view and they just refresh it little bit.

So, it is more formal hypothesis structure right, otherwise statistical hypothesis is same as research hypothesis. So, the moment we test research hypothesis will call it statistical hypotheses. So, business researchers want to prove the research hypothesis that older workers are more loyal to company, so this is the hypothesis or the business researcher is different, what is the hypothesis he wants to prove that more workers are more loyal to the company.

And to know this hypothesis or to test this hypothesis there will have to be let us say an instrument or a questioner. So that questioner would be given let us say old employees and new employees and let us say the questions in questioner or on 5 point scale. So, an loyalty you can have 4, 5 questions and then you compare mean of let us say 10 old workers and 10 new workers right.

So, if there is a difference then we say and whether old employees are more loyal or not right. So, the point here is let us on a 5 point scale the mean of old workers on a loyalty scale is let us say 4.5 and for new workers it is 3.5. So, can we say that old workers are more loyal. So, this decision you have to take how far or how much difference between these 2 would be considered as significant difference or not. So we have to lookout the proof threshold, what is the proof threshold?

(Refer Slide Time: 10:17)



So, there is as standard procedure for hypothesis is statistical hypotheses testing. So, basically statistical hypotheses or of 2 types you got null hypothesis and alternative hypothesis. Null hypothesis states that null condition exist, null condition exist means the current situation will prevail, the current theories will prevail nothing new is happening right, the old theories are still true, the old standards are still correct and system is in control.

So, this is about null hypothesis or when you do not expect any change in the situation is null hypothesis right. Null it is similar to 0 right 0 there is no change in the existing system that is null hypothesis. On the other hand the alternative hypothesis actually challenges the status scope. The null the alternative hypothesis will say that the old things have changed, the new theories true there are some new standards now and the system is out of control.

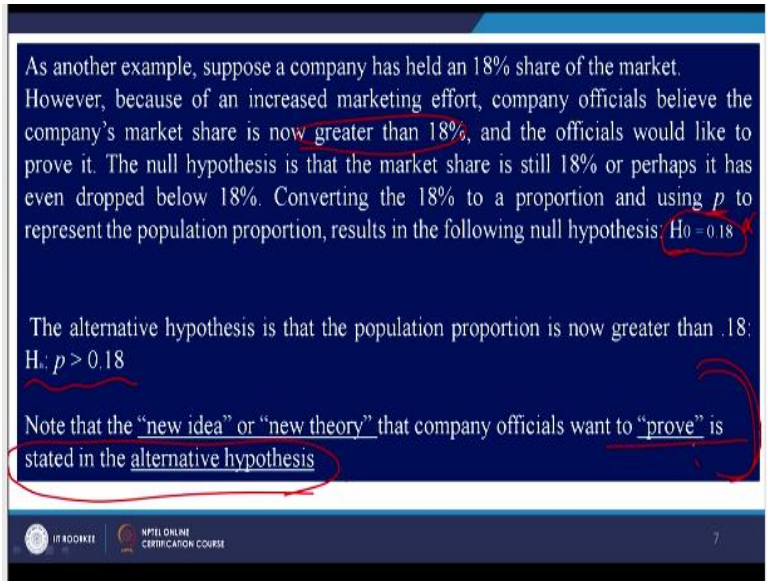
Something is happening while null hypothesis things are constant right as it is nothing is moving, nothing is changing. So, let us couple of examples of null hypothesis and alternative hypothesis. Let us say flour packaged by a manufacturer is sold by weight and a particular size of package is suppose to average 40 ounces. So, let us say you are manufacturer of flour package and the weight each package should be 40 ounce.

Suppose manufacturer wants to test whether the packaging process is out of control, so what he will do, he will test this hypothesis. He will pickup few bags, he will take the weight age of few bags he will find out it is mean and he will compare that mean with the population mean right. And then he will take a decision whether process is in control or not, so what is our null hypothesis that average weight of the flour package is 40 ounces.

And there is no problem in the system, system is when you absolutely find, so null hypothesis is 40, the weight of the package is 40 ounce. On the other hand alternative hypothesis would be it would challenge this statement and the challenge would be no the weight is not 40, it is either more than 40 or less than 40. So, this is how you would be writing alternative hypothesis Ha H0 this is null right.

So, $H_0 = 40$, Ha is alternative hypothesis is not equal to 40 right, so this is an example of null hypothesis and alternative hypothesis.

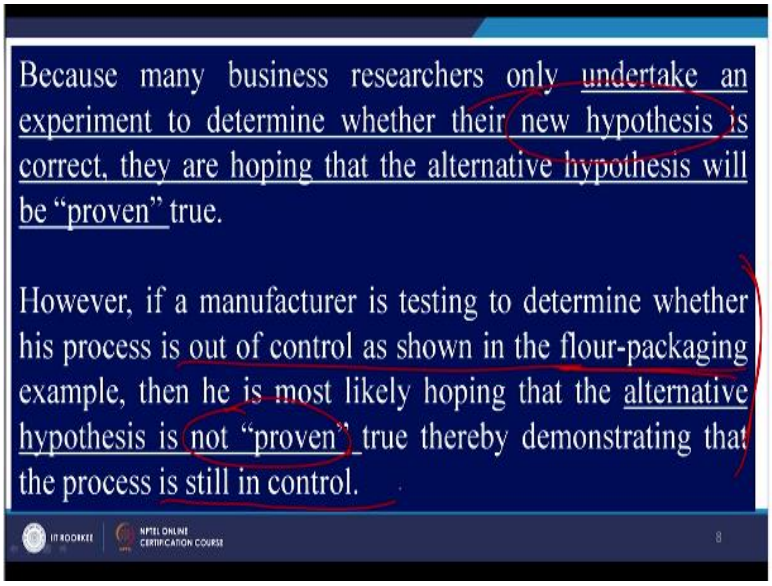
(Refer Slide Time: 14:06)



Let us look at one more example, let us say the market share of your company is 18% and you have put in lots of efforts in your business, let us say in last 3 years. And you are expecting that your market share has gone up right. And you are expecting that now the market share is more than 18% and you want to test this hypothesis. So, how to test this hypothesis let us say p is the market share or the proportion, so will say that H0 null hypothesis is market share is 18% and alternative hypothesis is market share is more than 18%.

So you may reject null hypothesis and you may not reject null hypothesis, it all depends on how you are choosing samples and after choosing sample you are comparing it with the population mean. So, here new idea, new theory that company officers want to prove is stated in the alternative hypothesis, this is something you need to keep in mind. So, whatever you want to test is to be kept in alternative hypothesis not in null hypothesis right.

(Refer Slide Time: 15:57)



So, business researchers only undertake an experiment to determine whether their new hypothesis is correct, they are hoping that the alternative hypothesis will be proven true. So they would always like to prove alternative hypothesis and the null hypothesis right. So let us look at the same example, in fact most of the times researchers want to prove alternative hypothesis right.

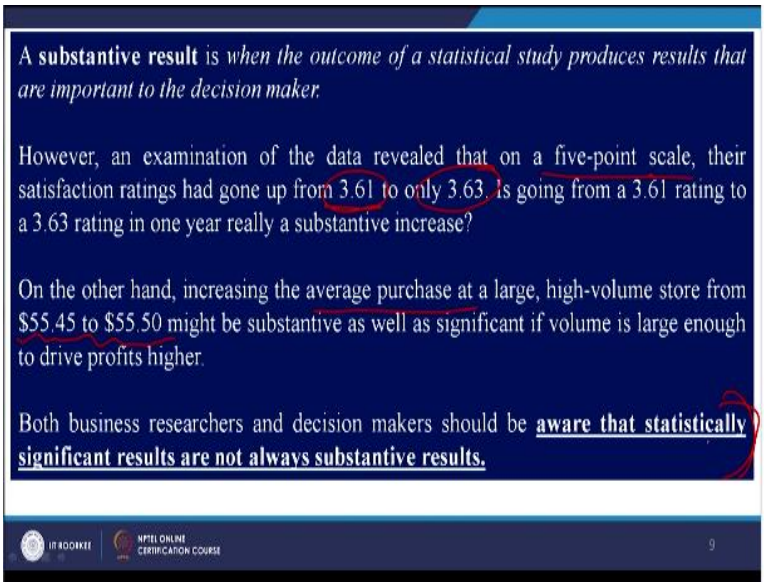
Now again it depends on situation, for example in our flour packaging problem the manager or the CEO of the company would like to prove that the process is in control, it has not moved out of control. So, he would like to ensure that the mean weight of the flour package is 40 ounces right and that he will not like to prove that the processes out of control right.

So if a manufacturer is testing to determine whether his process is out control are shown in flour packaging example. Then he is most likely hoping that alternative hypothesis is not proven, so he does not want weight to be more than or less than 40 ounces. So, he is most likely hoping that the alternative hypothesis is not proven true thereby demonstrating that the process is still in control.

So, he does not want to prove alternative hypothesis and he will ensure that the process is in control. Now the third type of hypothesis is substantial hypothesis, so we have seen two hypothesis, the first one was research hypothesis, the second one was statistical hypotheses. In

statistical hypotheses we have got null and alternative hypothesis, the third one substantial hypothesis or substantive hypothesis.

(Refer Slide Time: 18:47)



Now when the outcome of statistical study produce research data are important to the decision makers. So, there has to be substantial change in the situation, let us my null hypothesis is that the average weight of a glass is 45 kilograms right. And my alternative hypothesis is that the weight is not 45 kilograms ok, so I will pick up few students I will measure their weights now I will compare it with the hypothesized mean whatever I have hypothesized about the population.

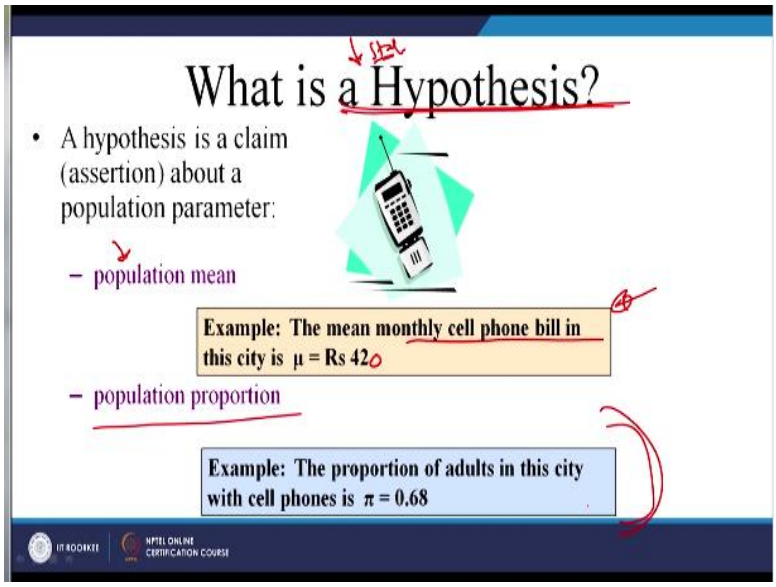
So, if there is large difference then I will say that I will reject the null hypothesis right. If population mean is closer to the sample mean then I am not reject the null hypothesis. So, there has to be as substantial or significant difference between population mean and sample mean and this decision would be taken by the decision makers or business researchers whether the difference is significant or not.

So just look at on 5 point scale where you measure the loyalty of old employees let us say you have measured loyalty of old employees 2 years back on a 5 point scale this was the average loyalty. And now it is gone up to 3.63, now is this substantial change this is to be seen right, now for some managers or for some researchers this might be this the substantial change for some of them this might not be substantial change.

So, it depends on once you test this particular hypothesis then you will have to decide whether there is a substantial increase in loyalty or not. On the other hand let us say increasing the average purchase price from let us say 55.45 dollars to 55.50 dollars. And there is a large amount of products or large volume of finish product you are selling. So, that might a substantial profit, so you have to look at whether there is a substantial change in situation or not right.

So both business researchers and decision makers should be aware that statistical significant results are not always substantial results, so this point is to be taken care of. That statistically significant results are not always substantive results ok.

(Refer Slide Time: 22:06)



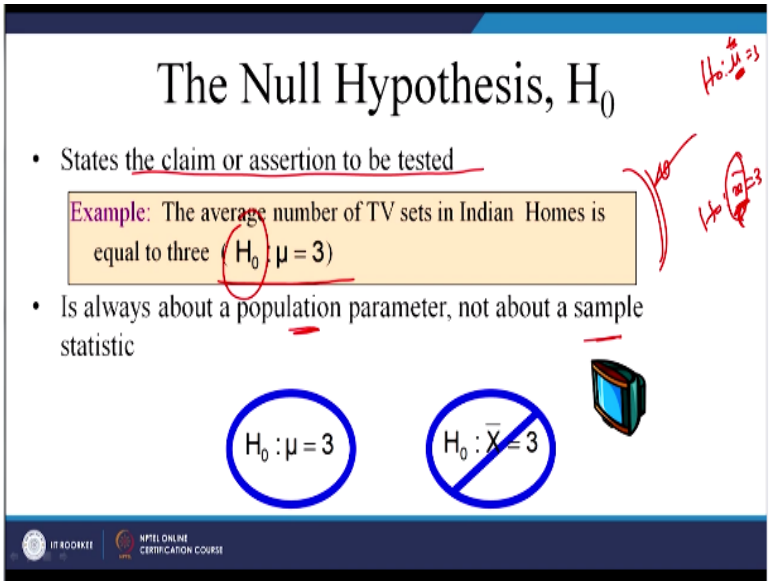
Let us look at some more points related to hypothesis, so what is hypothesis when I say hypothesis I am talking about statistical hypotheses right. So, statistical hypothesis are hypothesis is a claim about population parameter. So, decision maker or business researchers would be making certain assumption about population parameter maybe population mean, population proportion, population standard deviation so on right.

So, let us say the one of the assumptions is or one of the hypothesis is the mean monthly cell phone bill in the city is 42 rupees or let us say 420 rupees, so this is just an assumption. Now the manager of the cell phone company would be thinking that whether this mean is 42 rupees or

420 rupees or it is gone up, so this is about population parameter right not about sample statistics.

So you can have one more hypothesis population proportion, the proportion of adults in the city with cell phone is 68% or 0.68. This is just one more assumption the researcher would test this particular hypothesis whether the adults having cell phone are 68% or more than that or less than that right.

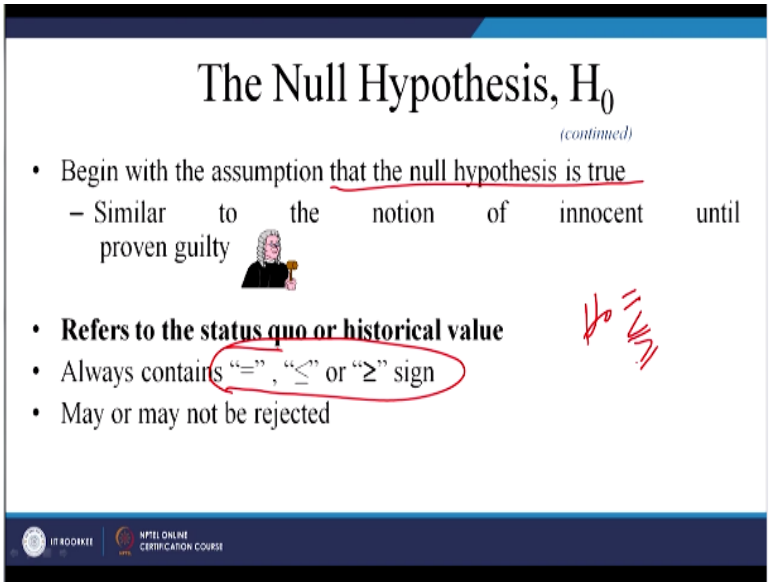
(Refer Slide Time: 23:58)



So, whenever we test statistical hypothesis as I said you have to have null hypothesis. So, generally you always have null hypothesis and alternative hypothesis, so it is about the states the claim or assertion to be tested right. Let us look at one more hypothesis the average number of TV sets will be homes is equal to 3 it is just a hypothesis and let us call it null hypothesis.

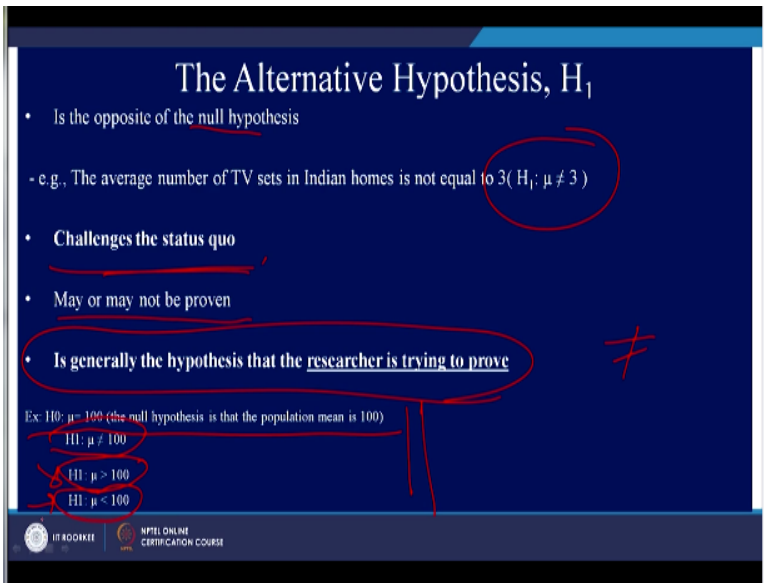
Because you want to test whether the number of TV sets in Indian homes are 3 or not, so this is null hypothesis. This is H0 and this we are talking about mean right, so this is equal to 3 right, so null hypothesis or hypothesis is always about population parameter right not about sample statistic right. So, you cannot say that null hypothesis is X bar = 3 this is not this would be incorrect, so we always use population parameter right not sample statistics.

(Refer Slide Time: 25:35)



So null hypothesis begins with the assumption that the null hypothesis true. So, it is similar to the notion of innocent until proven guilty. So, as we have already mention that null hypothesis about that old theories are still true, there is no changes taking place or there is a status quo maintain right, it refers to status quo as I said always contains these types of symbols, so your null hypothesis can be equal to or less than or greater than, less than equal to greater than equal to right, so may or may not be rejected ok.

(Refer Slide Time: 26:37)



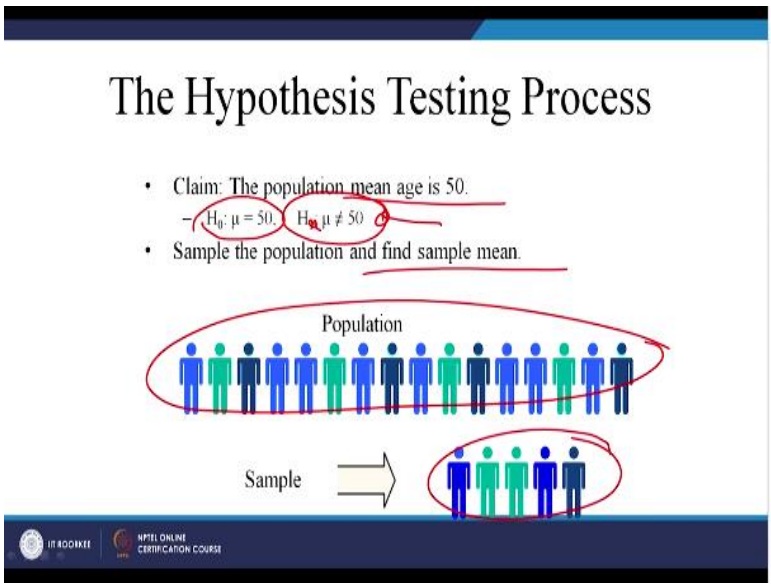
Let us look at alternative hypothesis though we have already talked about the alternative hypothesis this it challenges the status quo. So what is our null hypothesis, the average number of TV sets in Indian homes is 3 right and alternative hypothesis is let us they are not equal to 3.

So, either less than 3 or more than 3 right, so it challenges the status quo, challenges the old theories, challenges the existing things, it challenges the existing system or it challenges that the system is out of control.

Things have changed, may or may not be proven, is generally the hypothesis the researcher is trying to prove as we have mentioned earlier as well right. So, a researcher would always try to prove alternative hypothesis ok, so let us look at how do we write this null and alternative hypothesis. So, null hypothesis is 100, let us say population mean is 100, alternative hypothesis can be it is not equal to 100 right, so this is not equal to 100 or population mean is more than 100 or less than 100.

So these are only 3 percentages, so either population mean is 100. If not 100 then what could happen not equal to 100, the moment is a not equal to 100 it will have 2 more options right more than 100 or less than 100, so this is alternative hypothesis.

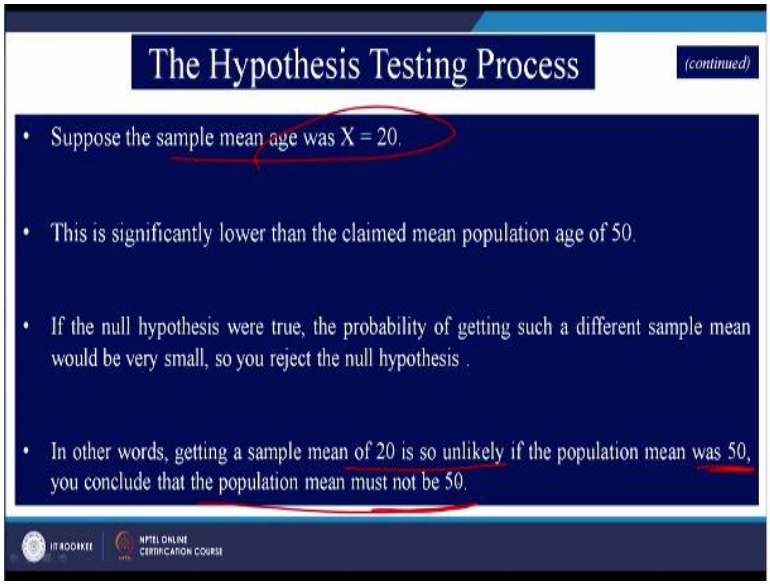
(Refer Slide Time: 28:31)



Now there is a hypothesis testing process, similar to what we have seen in case of estimation process right. In estimation process what we did, we did estimate population parameter from sample statistics. But in hypothesis what do we do, we assume about population parameter and then we try to prove or disprove that, so that is the major difference. So let us take an example let us say that the population mean is 50, the age of the population is 50.

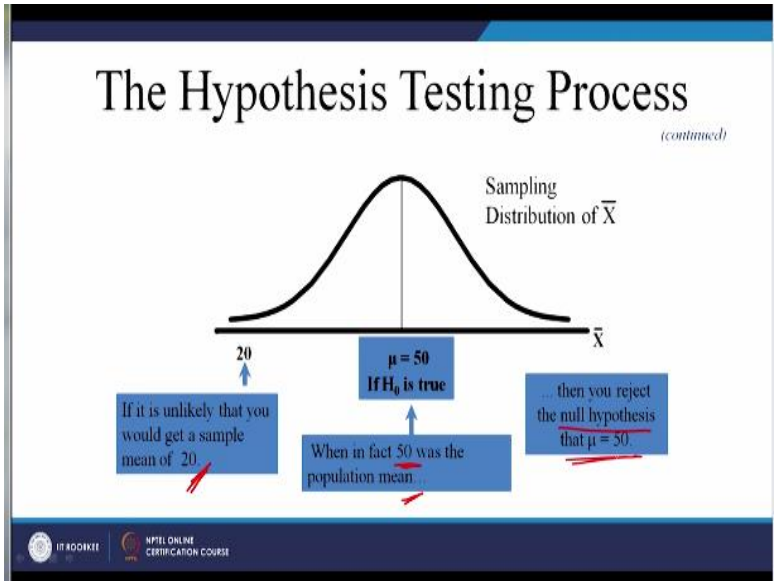
So, null hypothesis is H0: μ=50 otherwise it is not 50, so either you write Ha or H1 and the same thing right. Sample the population and find the sample mean, so to prove this that the age is not 50, will take sample from the population, so this is your population will take a sample from this population.

(Refer Slide Time: 29:51)



And then we will find out the sample mean, let us take for simplicity sample mean is 20. Now we will say that we are assuming that the population mean is 50 but sample mean is 20, so we will say that the sample mean is too small right. So will reject this null hypothesis right, so getting a sample mean of 20 is so on likely if the population mean was 50, so you conclude that the population must not be 50, so you will reject in a hypothesis.

(Refer Slide Time: 30:32)



So, you can say the same thing in this way if it is unlikely that you will get a sample mean of 20 by in fact 50 was the population mean then you eject the null hypothesis that mean is 50. So, with this let me stop here, let me mind what we did in today's session, we discussed about hypothesis, we discussed about types of hypotheses, we have seen research hypothesis, statistical hypotheses, in statistical you have seen null and alternative hypothesis and we have seen substantial hypothesis right.

And we have seen how the hypothesis testing process works, so there is a population you take a sample and you compare population mean and sample mean or let us say population proportion and sample proportion. And then you take a call whether your hypothesis was true or not, thank you very much.