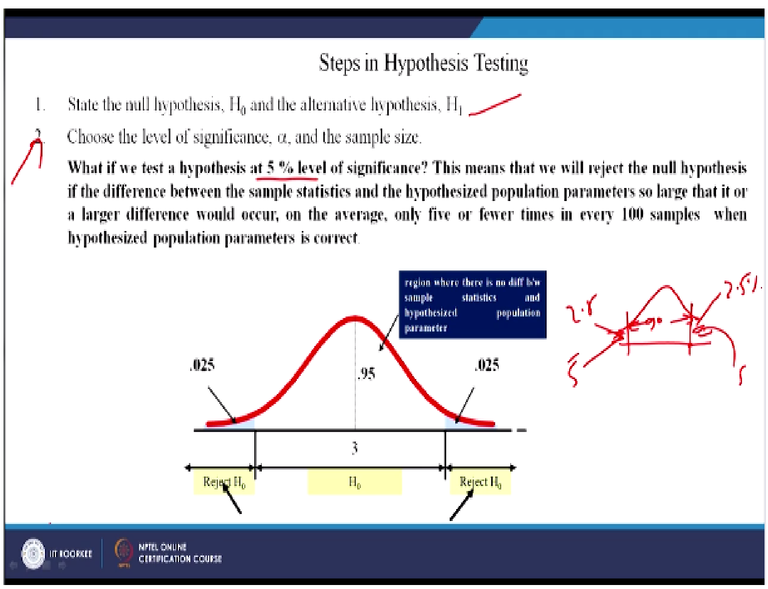
Business Statistics
Prof. M. K. Barua
Department of Management Studies
Indian Institute of Technology-Roorkee

Lecture-32
Hypothesis Testing Process-II

Hello friends I welcome you all in this session as you are aware in previous session we discussed how to test a given hypothesis. We have seen couple of steps and in today's session we will revise those steps and then we will move further. So, in hypothesis testing there are 6 steps the first one is need to frame null hypothesis and alternative hypothesis, this is the first step right then choose significance level and sample size.

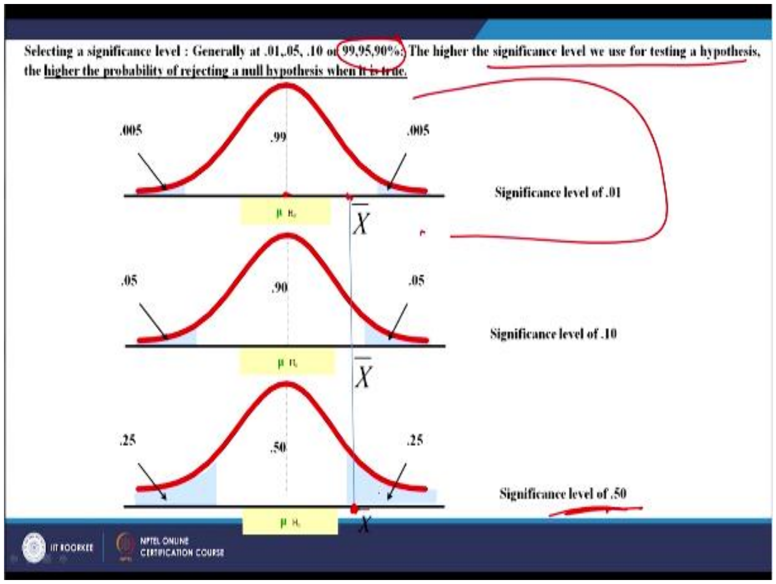
So, whether you want to test your hypothesis at let us say 90% significance level or 95 0r 99 and appropriate sample size should be selected in step number 2 right.

(Refer Slide Time: 01:17)



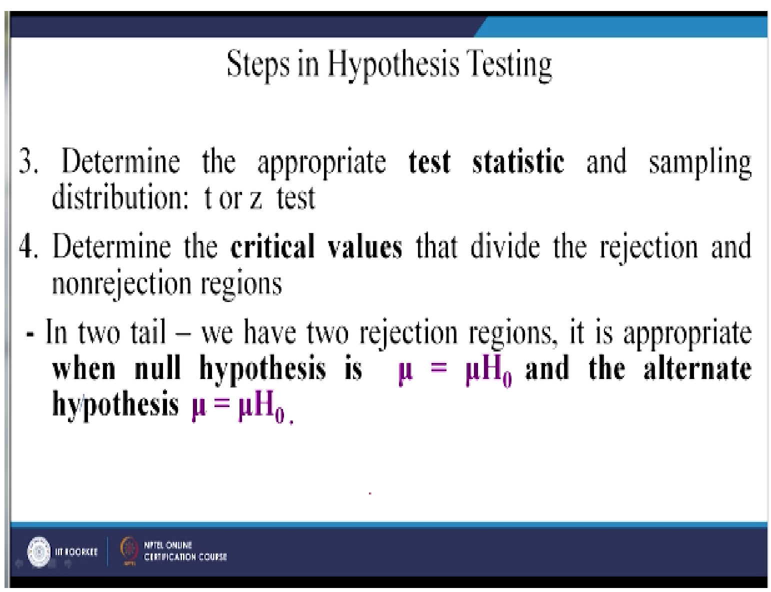
This is what we have discussed in previous class the meaning of 5% significance level, so if you want to test a hypothesis at 90% significance level. Then these 2 tails you will have 2.5% and the remaining is 95 or if you want this to be 90% and this would be 5 and this would be 5 right.

(Refer Slide Time: 01:47)



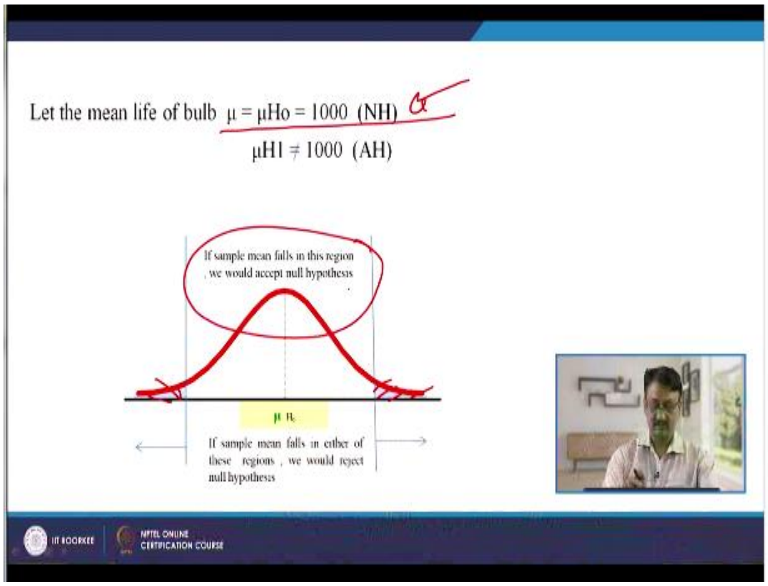
So, generally it is suggested that you should test any hypothesis at 99, 95 or 90% significance level. So, the higher the significance level there is a possibility that you are rejecting a null hypothesis which is true, this what you see in previous class as well. When you test a hypothesis at 99% significance level you are not rejecting a null hypothesis but at 50% significance level or when alpha is let us say 50% then you are rejecting the null hypothesis because this is the sample mean which is in rejection region right.

(Refer Slide Time: 02:31)



Determine the appropriate test which is to be applied if standard deviation of population is unknown and sample size is 30 or less than 30 you use t test otherwise Z test. Determine critical value that will divide on the rejection and non rejection regions. So, you will come to know which regions are rejection regions and which one is not rejection region right. So, in two-tail test you will have 2 rejection regions in 1 tail test you will have 1 rejection region.

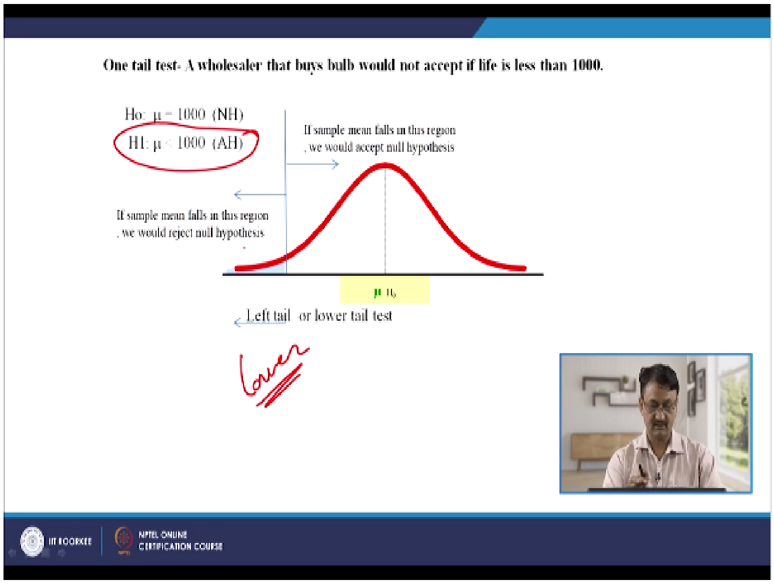
(Refer Slide Time: 03:07)



This what we have seen as well in previous session, so let us say there is a manufacturer who is claiming that the mean life of the bulb is 1000 hours. Now if you are a customer then would you by those bulbs for which life is less than 1000 hours ok we will come to this particular part of the which is in next slide. So, let us look at, so either so you frame hour here a null hypothesis like this.

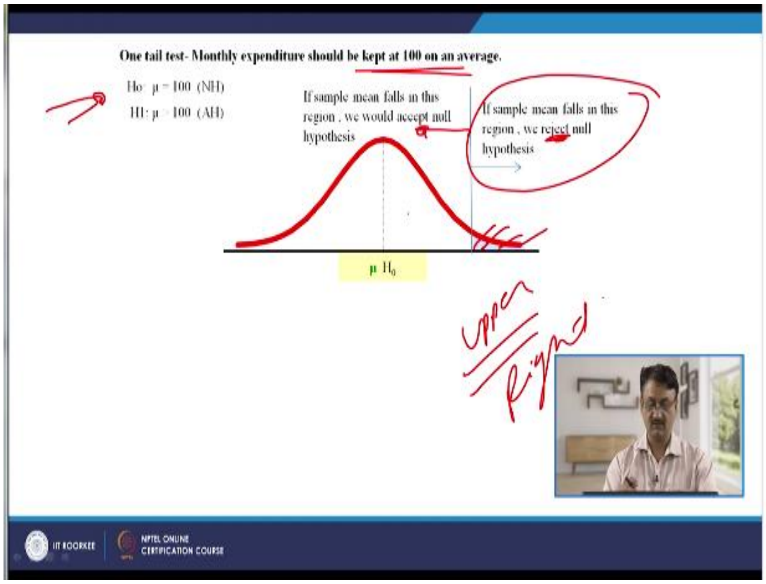
The mean life of the bulb is 1000 hours and the alternative hypothesis it is not 1000 hours right. So, you will have these 2 rejection regions right, so this is this region is a region where sample mean falls and will not reject null hypothesis right. Let us look at the same example, so manufacturer is claiming that the mean life of the bulb is 1000 hours and you have being a customer you would not like to have those bulbs for which life is less than 1000 hours.

(Refer Slide Time: 04:24)



So, what you want to prove, you want to prove this one, that you he will not reject or he will reject those bulbs for which life is less than 1000 hours. So, this is one tail test and this is also known as left tail test or lower tail test right. So, in one tail test there will always be one rejection region.

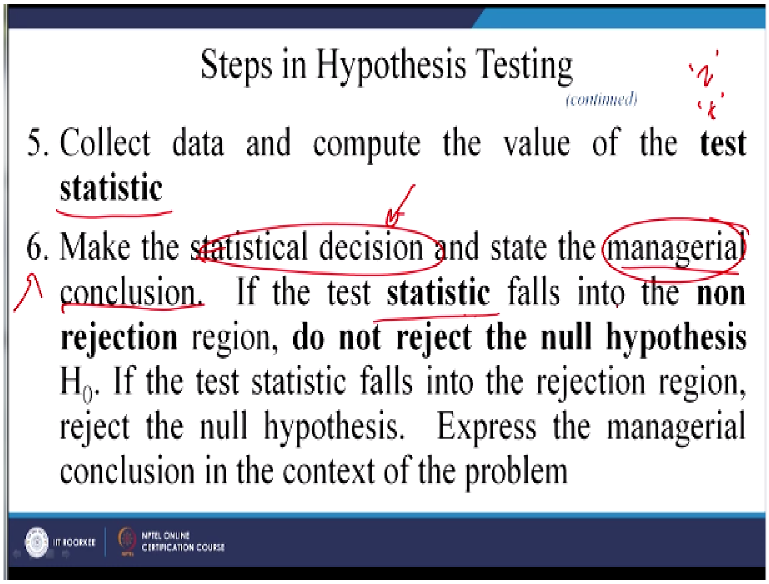
(Refer Slide Time: 04:46)



One more example on one tail test this is an upper tail test right or right tail test, so null hypothesis is this. That the let us say in this example you are manager of a particular sales area and there are let us say 50 salesman working under your supervision right. And you have given them a patrol allowance of 100 rupees per day. So, being a manager you should be worried about only those expenditures where which are more than 100 rupees.

So, this is a case of upper tail test right, so this is your rejection region and this is acceptance region, this side is acceptance region ok.

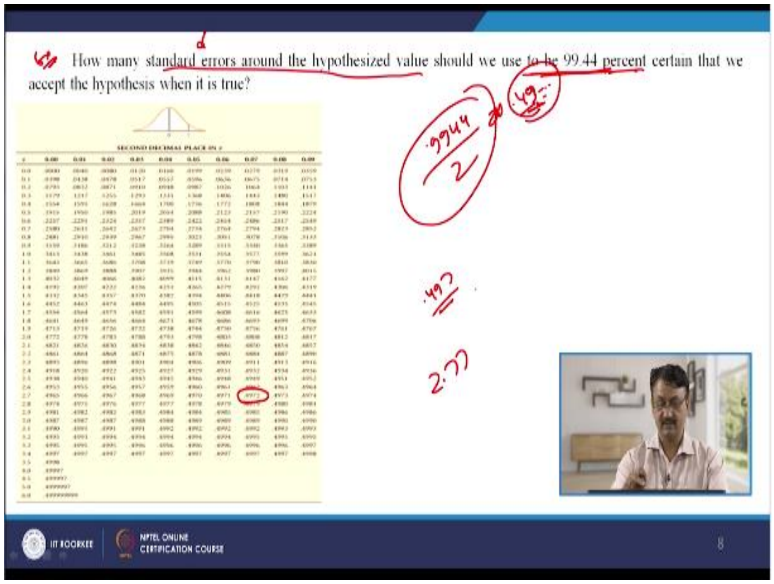
(Refer Slide Time: 05:42)



So, the next step is collect the data and compute the value of test statistic, test statistic. So, what is test statistic whether it is Z value or t value depending upon question right. So the next step is name the statistical decision and state the managerial conclusion, so statistical decision means you should either reject, null hypothesis or you do not reject. Manager and application means or a managerial conclusion means you need to take certain action after taking statistical decision.

So if you reject a null hypothesis then you are suppose to take some action otherwise the current situation will prevail. So, if the statistics sample statistic falls in non rejection region do not reject if falls in rejection region reject it right.

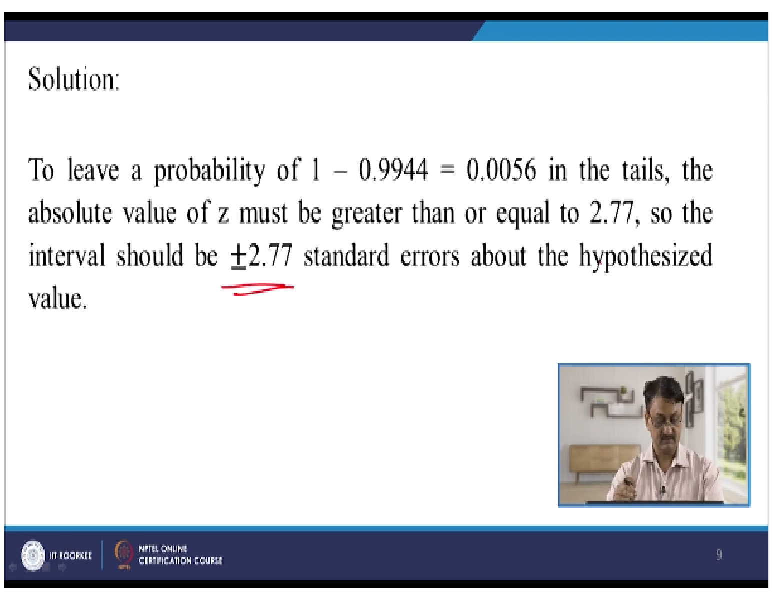
(Refer Slide Time: 06:47)



Let us look at couple of examples, so this is an example wherein the question is like this how many standard errors around the hypothesized value should we use to be 99.44% certain that we accept the hypothesis when it is true. How many standard error around the hypothesized value should we use to be this much percentage certain that will accept the hypothesis is when it is true.

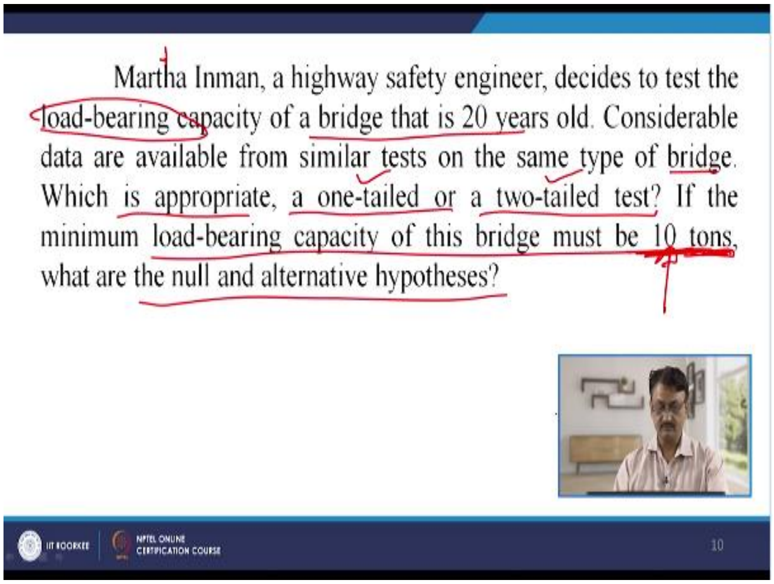
So, for this you would be requiring Z table and in Z table if you look at in fact you can do it like this 0.9944/2 right, so sum 0.47 something like that right. Now look at that area in this table, so 0.47 right it is somewhere here 0.47, so exact value you need to look at.

(Refer Slide Time: 07:56)



In fact this is 2.77, so 2.77 is this, so in fact the this is actually nothing known this when you divide this value it comes around 0.497 and so on. When you divide 0.9944/2 you will get 0.497 right and the Z value is 2.77 right ok. So the interval should be this much standard error about the hypothesized value right.

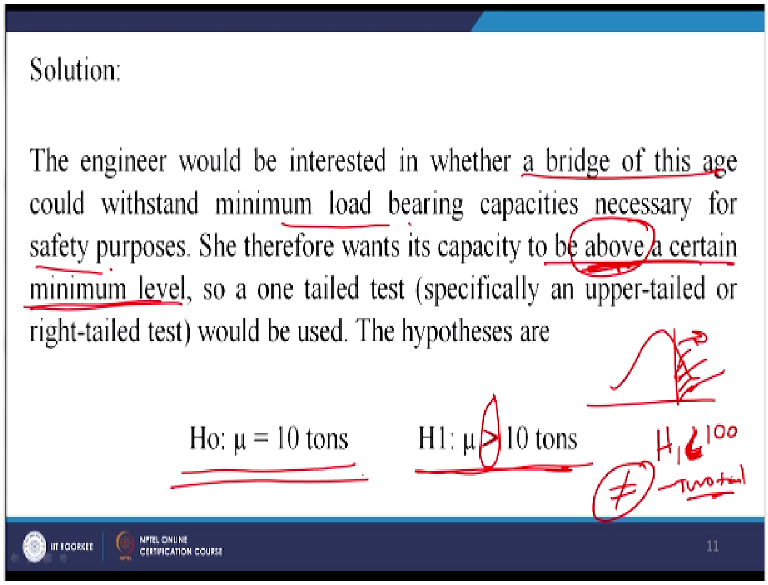
(Refer Slide Time: 08:39)



Let us look at one more question Martha Inman, a highway safety engineer decides to test the load-bearing capacity of a bridge that is 20 years old. So we have so the he use to check load-bearing capacity of a bridge right, considerable data are available from similar test on the same type of bridge, so some data are already available which is appropriate a one tail or a two-tail test, if the minimum load-bearing capacity of this bridge must be 10 tons.

So this lady is safety engineer she has to decide whether one tail test is appropriate or two-tail test is appropriate and what she has to do. If the minimum load-capacity of the bridge must be this much tons, so what are null and alternative hypothesis. So what would you do, here this is a case where the bridge has to have certain minimum load-bearing capacity and that minimum value is 10 right. So she has to check whether the bridge can sustain this much capacity or not right.

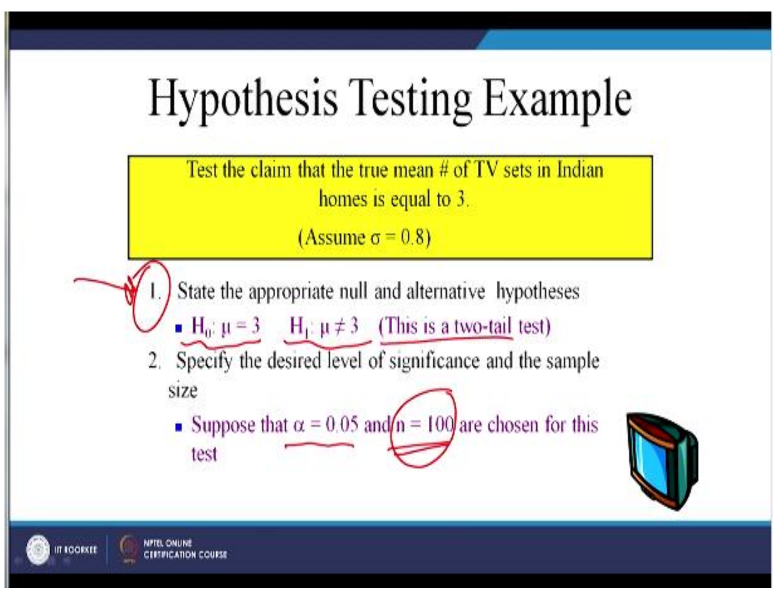
(Refer Slide Time: 10:08)



So, she will have null hypothesis like this it is equal to 10 tons and alternative hypothesis mu is more than 10. So, the engineer would be interested in whether a bridge of age could withstand minimum load for safety purposes or not. She therefore wants it is capacity to be above certain minimum level, so that is why because she wants the capacity to be above certain minimum level and what is that minimum level this 10 tons right.

So, alternative hypothesis is this right, so it will look like this type of distribution right this is what rejection region right.

(Refer Slide Time: 11:06)



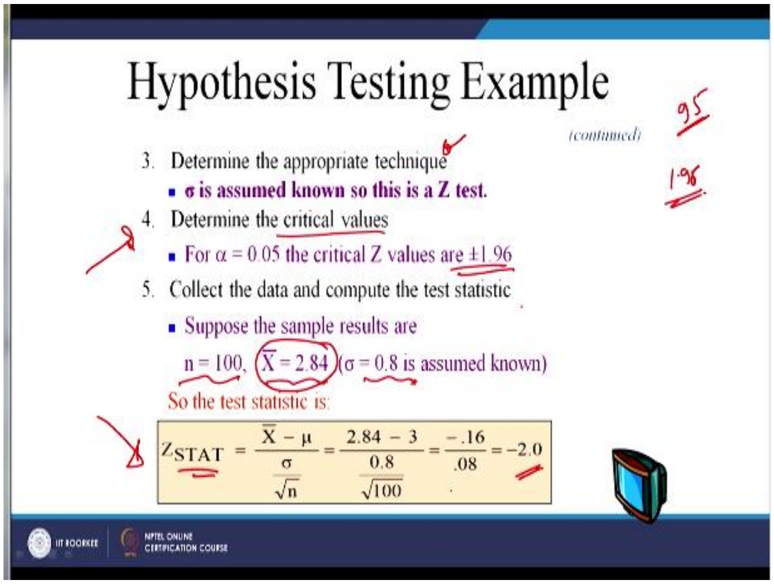
Let us look at one more example, we want to test claim whether the average number of TV sets in Indian homes is 3 or not right. So, our null hypothesis is this average number of TV sets in Indian homes equal to 3 and alternative hypothesis is not equal to 3. So which test will be using it is a two-tail test, how to decide whether you have to use two-tail test or one tail test, the sign of the alternative hypothesis would decide.

For example this is greater than type in alternative so that is why you are using upper tail test. If let us say upper let us say an alternative hypothesis is like this in some other example ok. Then this is a case of lower tail test or left tail test right and if these the sign is not equal to type in alternative hypothesis. Then you will be using two-tail, so test keep in mind for less than or equal to type of sign in alternative hypothesis you will be using one tail.

If it is not equal to sign then you would be using two-tail test right, so this is a case of two-tail test right. So, this the first step, you need to frame null hypothesis and alternative hypothesis, second is specify the significance level. So, in this question alpha is 0.05 right and 1-alpha is **is is** 95% right, n is 100. So, our null hypothesis is that average number of TV sets in Indian homes is equal to 3 we do not know whether there are actually 3 TV sets or not.

So, what we did it we just collect a data from 100 homes and we counted number of TVs in this 100 homes and then we took the average.

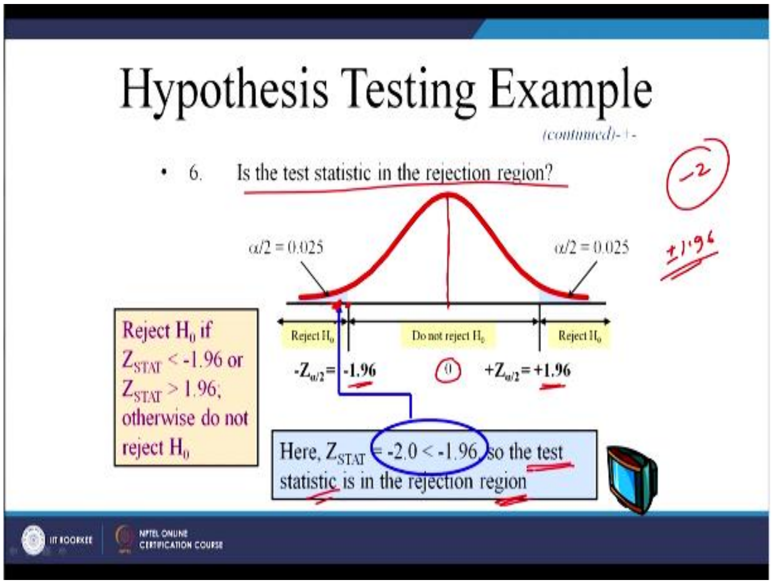
(Refer Slide Time: 13:23)



So, the average was this 2.84 right, so let us move onto state number 3 determine the appropriate technique or appropriate test to be used. So, in this case we would be using Z test why because sample size is more than 30 and we have been given standard deviation of population. The next step is determine critical value from Z table right since alpha is is 5%. So, the table value would be we checked earlier as well it is 1.96 ok, so this is your critical value is -1.962 to +1.96.

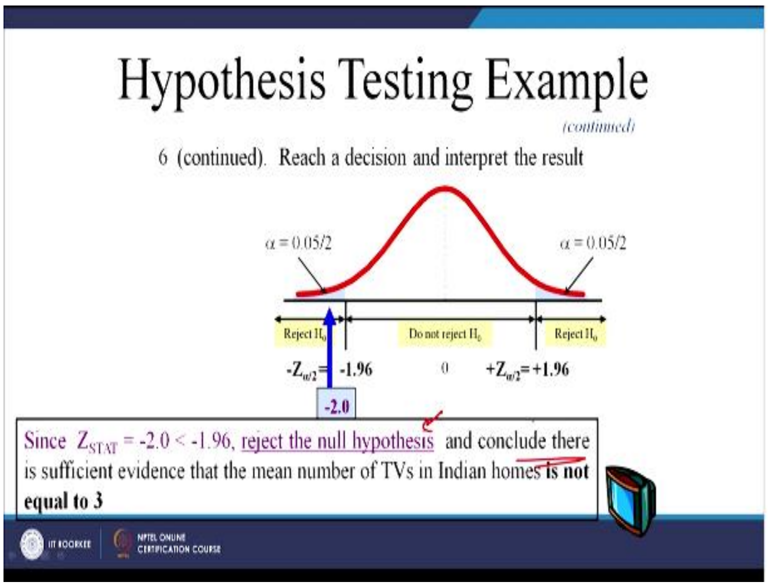
Then collect the data, so we collect a data from 100 homes and we found that average number of TV sets is this right 2.84. of course this was already available with us the standard deviation of the population. So, Z statistic would be calculated and this value is -2 ok. Now what, what is the next step we have to take a call whether we need to reject this hypothesis or we do not reject this hypothesis, this means null hypothesis. So, is this test in this rejection region you have to check this.

(Refer Slide Time: 15:06)



So, this is your Z value 0 right, this is Z value +1.96 and -1.96 from table we have got the right. Now the calculated Z value is -2, now is -2 falls in this region in ±1.96 it is outside this region, so it is somewhere here right just see this. This is -1.96, this is -2 which is in rejection region, so we will reject the null hypothesis, so the test statistics is in rejection region, so we will reject null hypothesis right.

(Refer Slide Time: 15:55)



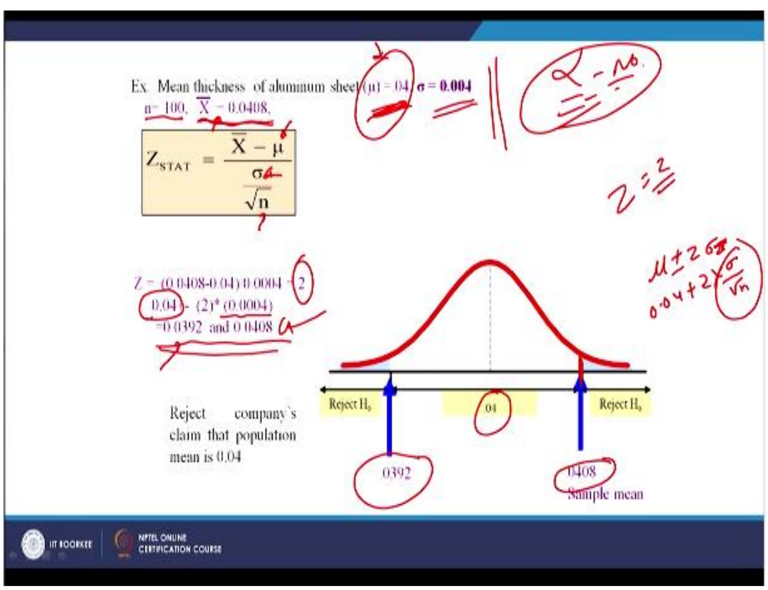
And after rejecting what we conclude, we conclude that the average number of TV sets in Indian homes is not equal to 3, when I say it is not equal to 3 it means more than 3 or less than 3. So, in this case what it is, is it more than 3 or less than 3 it is less than 3 because our sample mean is

2.84. So, we will say that we will reject the null hypothesis and we say that the average number of TV sets in Indian homes is not equal to 3 but less than 3.

Let us look at one more example, let us say you have to built you are suppose to let us say construct our stadium and for construction of stadium you need some aluminium sheets ok. So, you have selected a vendor who is ready to supply you aluminum sheets, now since you have to construct a stadium the weight of the aluminium sheets should not be too much or the sheets should not be too thick. Otherwise, the structure will collapse because of weight of the aluminium sheets.

On the other hand the weight or the thickness of the aluminum sheet should not be too thin otherwise the structure will let us say it would not be a robust structure. So, the thickness of the aluminum should be proper and that aluminum thickness sheet as I said should not be too thick and too thin. So, you have selected a supplier to who would be supplying you aluminum sheets right. So, that supplier has claim that the mean thickness of aluminum sheet is 0.04 mm or centimeter whatever you want right.

(Refer Slide Time: 18:04)



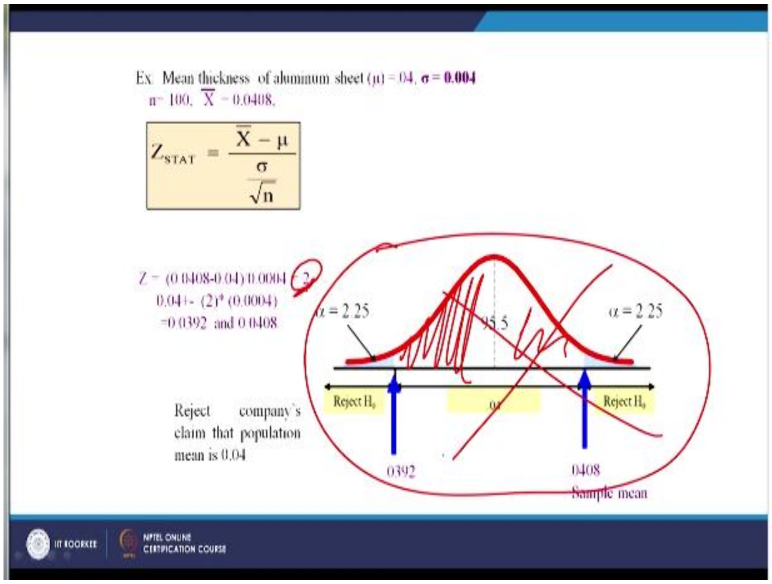
So the supplier is claiming that the mean thickness of aluminium sheet is 0.04mm with standard deviation of 0.004 and you want to test the claim made claim which is being made by the supplier. So, what you have done we have just select a 100 aluminium sheets and you just when you calculated it is the mean thickness was 0.0408 right what is the claim of supplier 0.04 aluminium.

And when you took a sample of 100 and you calculated mean thickness is 0.0408, so first of all in this case you have not been given alpha. So no alpha in this case, so how to solve a question like this, so first of all calculate Z value, Z is x bar-mu this is sample mean-population mean, population is standard deviation and sample size, so this is true right. Now once you are done with calculation of 2 what you should do this is your population mean.

So, population mean ±Z into this standard error, so mu is this right 0.04 + Z is 2 into this, this is standard error of mean. So this would be what this is sigma/root n right, so this becomes 0.0004 ok when you multiplied by 2 when you calculate range this is the range. So, this is your hypothesized population mean, this is the lower limit and this is upper limit right, so 0.039 to 0.0408.

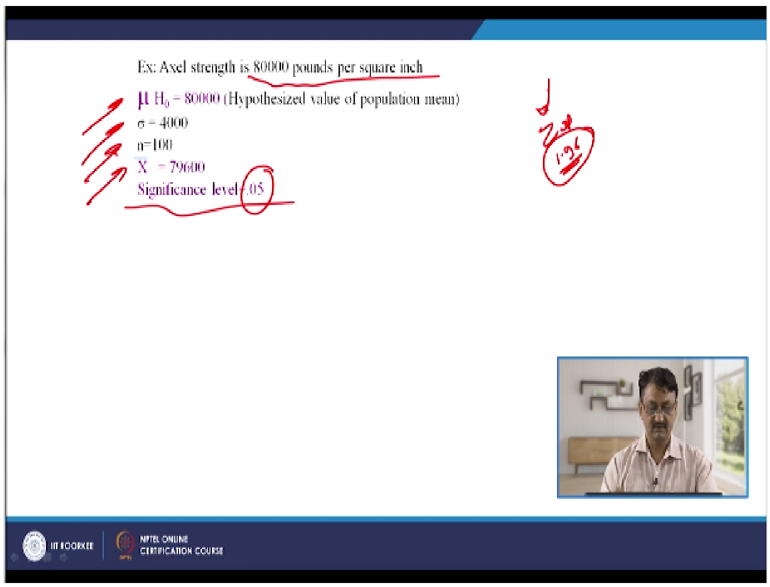
Now after finding these limits you need to check whether your sample mean is falling in this range or not. So our sample mean is exactly on this upper limit, so you will reject the null hypothesis since it is not in between right it is just on upper limit, so you will reject the null hypothesis right ok. So, this is a question which we have solved without having alpha value, so any question on hypothesis testing can be solved if you do not have even alpha value right.

(Refer Slide Time: 21:25)



Let us look at the same question and what you can do here is since you know the value of Z right. Now when Z=2 we can easily find out this area from table right, so it would be approximately 95.5. So in fact if you wish you can do this calculation otherwise is not needed at all right.

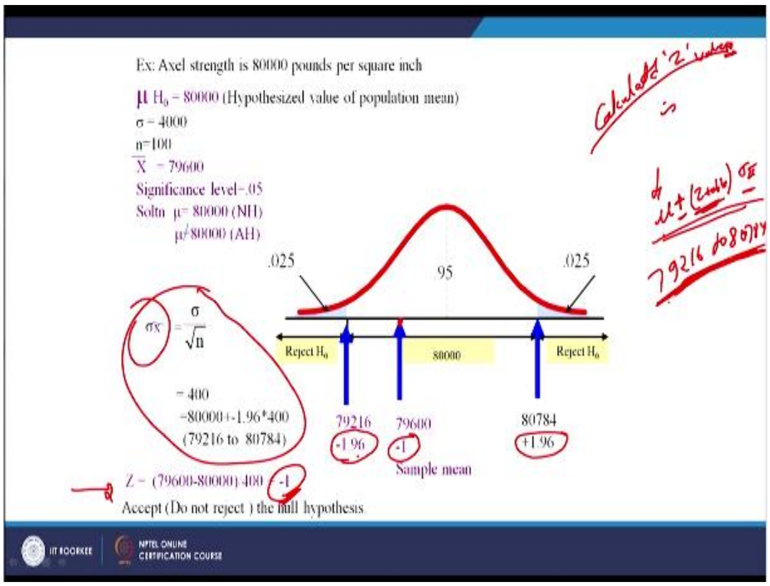
(Refer Slide Time: 22:01)



Let us look at next question let us say the strength of an axel is 80000 pounds per square inch. So, this is our hypothesized population mean which we are measuring in terms parts per inch right, so strength of axel is this. And the alternative hypothesis is that the strength of axel is not this much, so this is a case of two-tail test right. So, you have been given hypothesized mean you have been given standard deviation of the population sample size and sample mean and significance mean alpha is also given right.

So, how to solve this question, the first thing is you can easily do it calculate Z value and compare it with table value which would be 1.96 for this alpha value Z value would be 1.96 and you would be calculating Z as well.

(Refer Slide Time: 23:14)

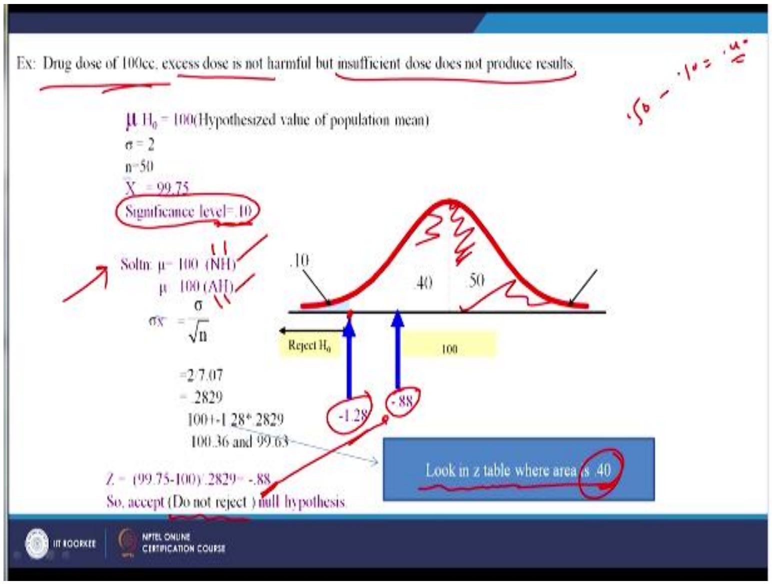


So, let us look at this, so the Z value which you are calculating is this, this is -1 right, now -1 is Z value which is here and the table values are 1.96 + and 1.96 – and + right. So, this Z value is in between critical values, in other words the calculated value, calculated Z value is in between table value right or critical values. So, we will this is in non rejection region, so we will not reject the null hypothesis right.

Now there is one more in which you can calculate the same answer, now this we have rejected null hypothesis after looking at calculated Z value and critical value right. You can do one more thing; you know this standard error of mean just 400 right, so population mean ± Z table value right into standard error. Now this will give you again some range, which is 79216 to 80784.

Now just check is your sample mean lying in this region, is sample mean falling in this within these limits 79600 yes, so we will not reject null hypothesis right. So, we have seen couple of ways of deciding whether you should reject or not reject null hypothesis right. So, if alpha is given just calculate Z and comparative with table value if not then you can have just you will always be having population mean ± Z table critical value into this right and check whether sample mean falls in this range or not this is other way of doing it right.

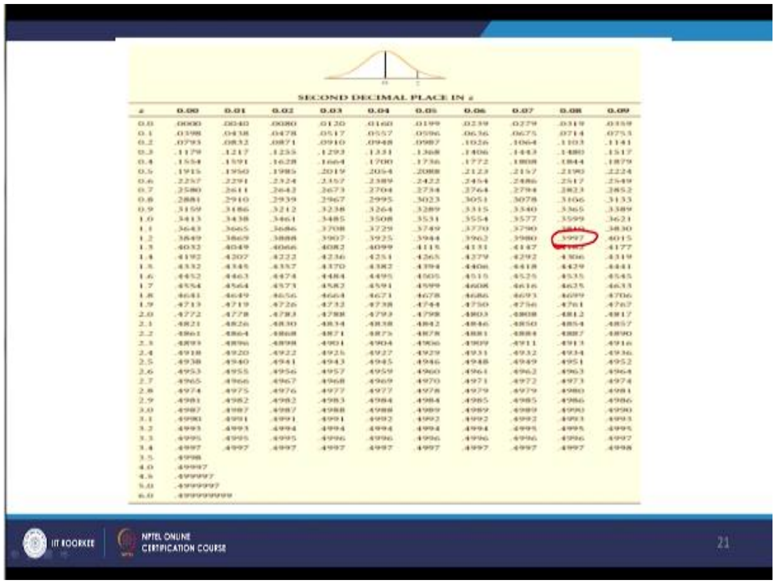
(Refer Slide Time: 26:10)



Let us look at a question on one tail test let us say there is a patient in hospital and you are supposed to give him certain medicines right. So, if you are giving a drug dose of 100cc to the patient and if you give excess dose when it is not harmful for the patient. But insufficient dose does not produce desired results, so what kind of null and alternative hypothesis we will do different. We know that excess dose is not a harmful but if it is insufficient dose then that would not produce sufficient or that would not produce desired results right.

So we have to ensure that the dose size should not be less than 100 right, so this is your null hypothesis and this is your alternative hypothesis NH null hypothesis AH is alternative hypothesis right. So, simple thing since this is a case of 1 tail test right, so this 50% area is this and this side is remaining 50% area significance level is 10% right. So, we are interested in this critical value, so look at in Z table where area is 0.40 right why 0.40 because this is one tail test right and our significance level is 0.10, so 0.5-0.1 it is 0.40 right.

(Refer Slide Time: 28:14)



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	[illegible]	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	[illegible]	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998									
4.0	.49997									
4.5	.499997									
5.0	.4999997									
6.0	.499999999									

So, let us look at Z table and area is 0.40, so 0.40 it is approximately 0.40 right, so 1.28 ok. So your calculated value is sorry -0.88 and table value is -1.28, so this is your calculated value which is here and this is your table value. So, is this calculated value in rejection region or non rejection region, so this is in non rejection region, so will not reject null hypothesis, so we do not reject null hypothesis ok.

So, we do not reject null hypothesis means what we do not require any action now right because we have not rejected null hypothesis ok. So, we are saying that the dose size is 100cc right this is the conclusion. So, let me summarize what we have done in today’s session we have looked at all the 6 steps of hypothesis right from framing of null hypothesis and alternative hypothesis to selection of alpha.

Then to decide whether Z test or t test is to be used then you need to calculate the Z value then you need to take a decision whether. We calculated Z value is within rejection region or not and then take appropriate decision. We have looked at couple of examples on how to decide whether a question has got two-tail test or one tail test, to decide this you are suppose to read the question very carefully because hint will be there in question itself.

If and if these sign in alternative hypothesis is let us say not equal to type then it would be a two-tail test why we call it two-tail test. Because there are 2 rejection regions and when the sign in alternative hypothesis is less than type it would be a lower tail test and if it is greater than type it would be an upper tail test. Let me stop here, in next class we will have some more questions, thank you.