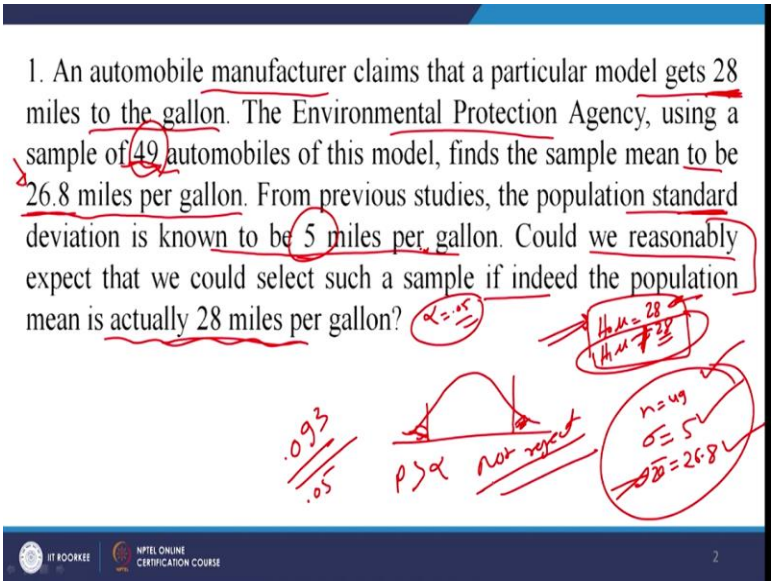
Business Statistics
Prof. M. K. Barua
Department of Management Studies
Indian Institute of Technology – Roorkee

Lecture - 37
Hypothesis Testing using Minitab

Hello friends, I welcome you all in this session. As you are aware previous session we were discussing about Hypothesis Testing and we have worked out several examples using One tail test and Two tail test. In today's session, we will work out couple of more examples but we will use Minitab software. So let us get started.

(Refer Slide Time: 00:54)



The first question is an automobile manufacturer claims that a particular model gets 28 miles to the gallon. The Environmental Protection Agency, using a sample of 49 automobiles of this particular model, finds the sample mean to be 26.8 miles per gallon. So what is the claim of the manufacturer? Manufacturer is saying that the model gives 28 miles to the gallon but when Environmental Protection Agency selected 49 automobiles, the agency found that it is just 26.8 miles per gallon. Okay.

From previous studies, the population standard deviation is known to be 5 miles per gallon. So you have been given standard deviation of population. Could be reasonably expect that we could select such a sample if needed the population mean is actually 28 miles per gallon. So we have to

test whether the automobile manufacturers claim is true or not. So you have to first of all frame null and alternative hypothesis.

In fact, in this question, you have not been given an alpha value. So if not given you can assume it let us say 0.05. In fact, I have thought you couple of methods in which you need no require even alpha value. We just use the Z value for calculation of upper and lower limits and then we should see whether the sample mean falls in those limits or not. In fact, we have seen confidence interval approaches wherein in we need to see whether the hypothesis mean falls in confidence interval or not.

So in this question you need to have first of all null and alternative hypothesis then the alpha value of course you can take this value here. The third one is you need to decide whether you would be using t-test or z-test. So how do you thing, what would be the null and alternative hypothesis? So you are supposed to frame first of all null hypothesis. So I think you can easily frame null hypothesis and the null hypothesis is this, right.

So let us say mean, population mean is hypothesis size is 28 miles per gallon. What about alternative hypothesis? Will be > type, would it be < type or would it be not equal to type. Just read this question once again. Could we reasonably expect that we could select such a sample if needed the population mean is actually 28 miles per gallon? So what you want here is you need to have your alternative hypothesis as not equal to 28, okay.

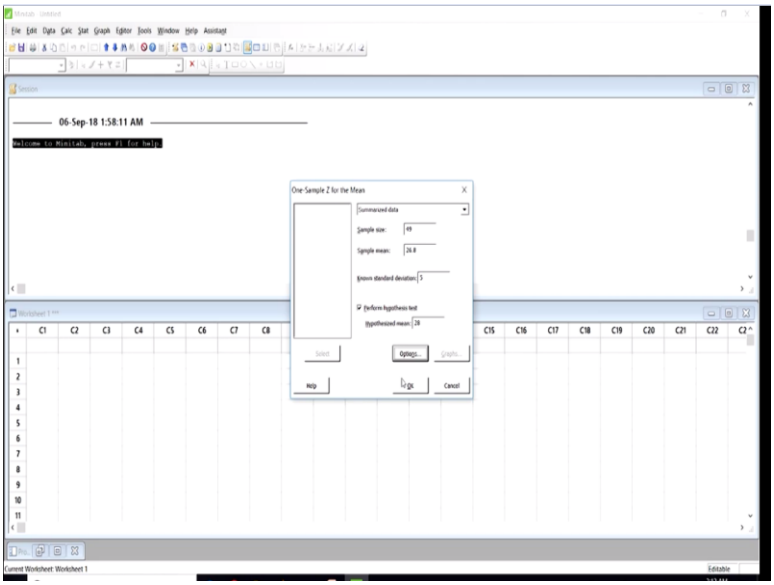
Since this is known now, you have to decide which test you would be using here. It is a z-test or t-test. Keep in mind that we have always said that we will apply t-test when there are two conditions met. First is standard deviation of population is unknown. What about this question, that about standard deviation of population, it is there, yeah it is there. Standard deviation of population is given, this 5 miles per gallon, so this known.

And of course, sample size is also not equal to 30 or < 30, sample size here is 49, so n=49; population standard deviation is 5, sample mean is this, right 26.8. And of course hypothesis mean is 28. So this is a case of One tail test or Two tail test? It would be case of Two tail test. Is

not it? Because this sign is not = sign. So this rejection region and this is another rejection region, right.

So this one is a Two tail test example. So let us solve this question using Minitab right. So we have noted down that, n is 49; standard deviation is 5; sample mean is 26.8 and hypothesis mean is 28.

(Refer Slide Time: 06:40)



Let us look at Minitab software. So first of all in this Minitab software there are two windows, two areas so this is, this is one area wherein you would be getting answer and this is your worksheet. The lower part of this is worksheet right. So let us look at how to use Minitab software. So first of all go to Stat and you should go to Basic Statistics. After Basic Statistics is 1-Sample Z test, right. Why Z test?

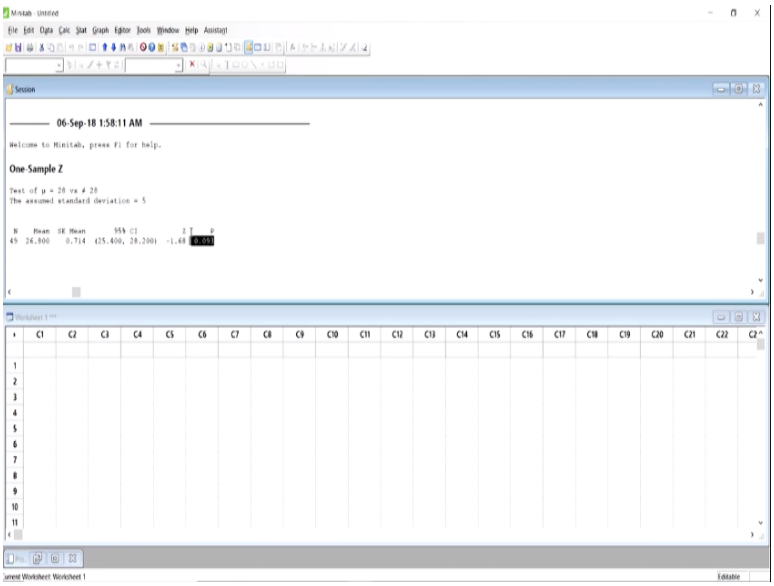
Because you have got sample size more than 13 and standard deviation of population is known as well. So here you have two options One or more sample each in a column or you can have a Summarized data. So in this case we have been given summarized data. We have been even mean; we have been even population standard deviation. Otherwise we would have, we could have solved if there were raw data given to us without having any mean and standard deviation, right. So we will click at Summarized data.

One you click at Summarized data the first value to be entered over here is Sample size. So in this case our sample size was 49. Now, let us look at the sample mean. So what was the sample mean? If you look at sample mean, sample mean is 26.8, right. So just enter 26.8 over here, right. Known standard deviation, yes it is known. Let us look at standard deviation 5, right. So standard deviation is 5. Then you need to perform hypothesis test. So you need to click over here, right.

Now, then Hypothesized mean, you need to enter over here, hypothesized mean which is will go back to our example which is 28, right. So will click at Hypothesized mean and we will enter hypothesized mean over here. Now we have to decide two more things here. The first is whether this test is left tail test or right tail test or it is a two tail test, right. So for that we need to go to options over here then Confidence level, confidence level is 95%.

So we will write 95% over here, so by default this is 95% over here and Alternative hypothesis is just see, Mean is not equal to hypothesized mean. And there are three options there, right. The first is, Mean is < hypothesized mean, the second is Mean is not equal to hypothesized mean and the third one is mean is > hypothesized mean. So in our question we will be using mean not equal to hypothesized mean, okay. So you just click it over here than okay again and finally okay. Now we will get the answer.

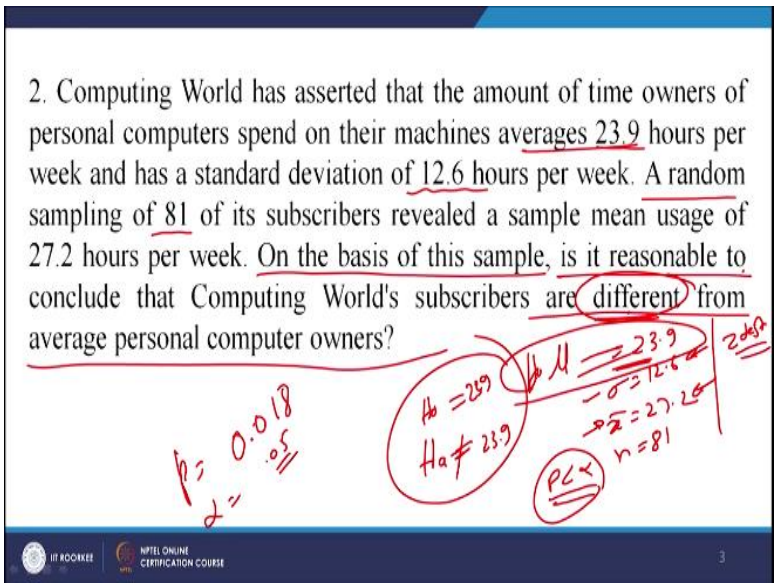
(Refer Slide Time: 10:31)



Now if you look at this, in this option here we have a hypothesized mean as 28 verses this was our alternative hypothesis not equal to 28; sample size was 49, mean 26.8 this is mean of sample, right 26 is the mean right; Standard error 0.714, so how we will get standard error? It is sigma/√10 is not it? So confidence interval is again you can get in this type of outputs, you will also get confidence interval which is 26.4 to 28.2. Z value is -1.68, right.

So P value is 0; just look at this. P value is 0.093. Is P value < alpha? P is 0.09, 0.093 and alpha is 0.05. Is P < alpha? No. P is > alpha, so we will not reject hypothesis. What is our condition? If P is < alpha null hypothesis go, it is to be rejected otherwise will not reject null hypothesis. So what is our null hypothesis here that the manufacturer claim of 28 miles per gallon is correct, whatever he is claiming is true. We have not rejected his claim. So that is our conclusion. Okay. Let us move onto second question.

(Refer Slide Time: 12:46)



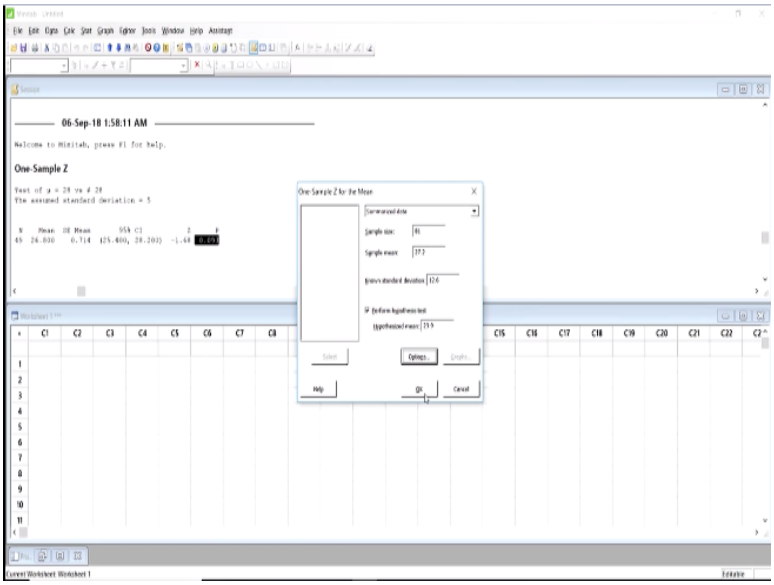
Computing World has asserted that the amount of time owners of personal computers spend on their machines averages 23.9 hours per week with a standard deviation of 12.6 hours per week. So you have been given 23.9 hours with a standard deviation of 12.6 hours. A random sample of 81 subscribers revealed a sample mean usage of 27.2 hours. So the hypothesized mean is H0 is equal to this 23.9. Standard deviation is 12.6 and sample mean is this.

On the basis of the sample is it reasonable to conclude that computing world subscribers different from average personal computer owners? Would do you think? What would be the null and alternative hypothesis here? And what would be the type of test? So null hypothesis is this H0=23.9; standard deviation is there; sample mean is there and n is 41, is not it. So you need to decide on whether you would be using t-test or z-test.

Let us human decide that as well, right. So you would be using z-test, right not the t-test. Now come to the alternative hypothesis, the most difficult part. And you would get the hint in question itself. Just read it. On the basis of the sample, is it reasonable to conclude that the Computing World's subscribers are different from; it is not whether they are more or <, right it is just mention that are they different from average personal computer owners?

So it has to be H1 or Ha to be not equal to type 23.9 and null hypothesis would be 23.9, right. Now you can solve this question again using Minitab.

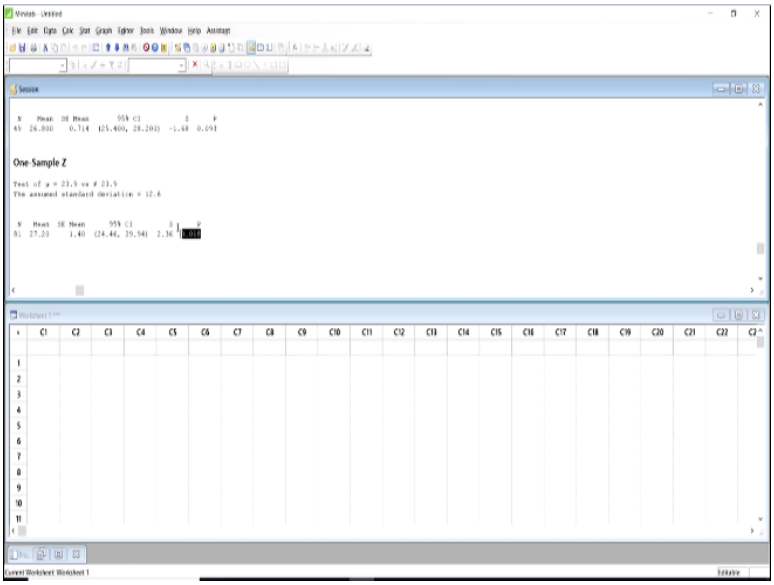
(Refer Slide Time: 15:38)



Just look at this. So I am going to Stat, Basic Statistics, 1-Sample Z test, now summarize data. Just enter all those values, Sample size is 81, Sample mean is 27.2, Standard Deviation is 12.6, Hypothesized mean is 23.9, 23.9 will go to option, we will test this hypothesis at; in fact, we have not been given any alpha value so let us test it again 95% significance level. And your

alternative hypothesis would be, Mean is not equal to hypothesized mean, you should press OK and OK.

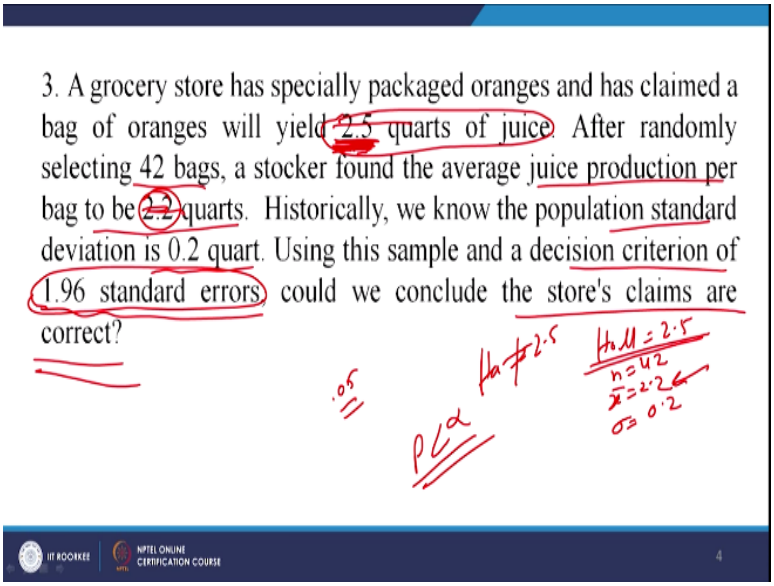
(Refer Slide Time: 16:55)



Let us look at output over here. This is your output. So N is given, your sample mean is there, confidence interval is there and P value is also there, right. So P value is 0.018, is 0.018, 0.018 is the P value and what is alpha value 0.05. So is P < alpha? Yes, P is < alpha, so we will reject null hypothesis. And what is null hypothesis? Null hypothesis is this, this that computing world has asserted that the amount of time owners of personal computers spend on their machine is averages 23.9, so this statement holds good, right.

So this hypothesis is not rejected, okay. So this is how you should work out questions using Minitab.

(Refer Slide Time: 18:12)

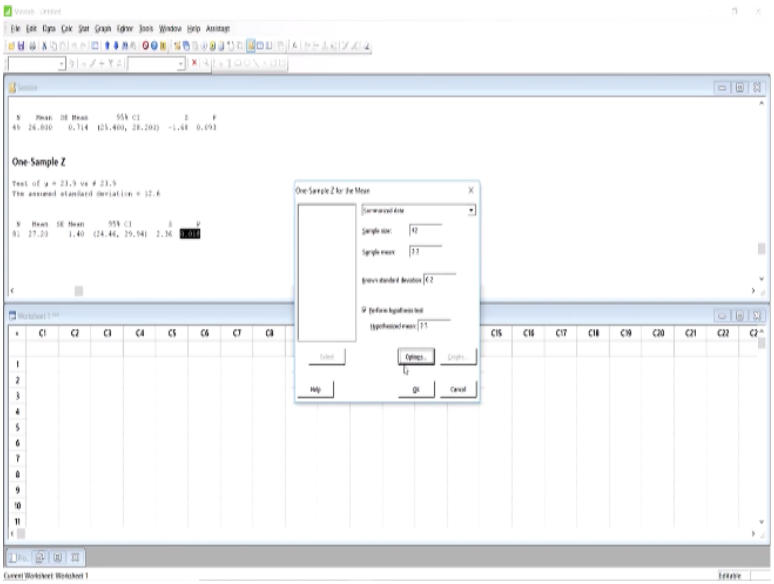


Let us look at one more example. A grocery store has specially packaged oranges and has claimed a bag of oranges will yield 2.5 quarts of juice. After randomly selecting 42 bags, a stocker found the average juice production per bag to be 2.2 quarts. Historically, we know the population standard deviation is 0.2 quart. Using this sample and decision criterion of 1.96 standard error, could we conclude the store's claims are correct? So what is the claim of store?

The claim of the store is, a bag of oranges will yield this much quarts of juice, 2.5 quarts of juice. So this is your null hypothesis, right is 2.5, right. And after randomly selecting 42 bags, so n=42. A stocker found the average juice production per bag to be 2.2 quarts with a standard deviation of 0.2 quart. Using the sample and decision criterion of 1.96 standard errors, could we conclude the store's claims are correct?

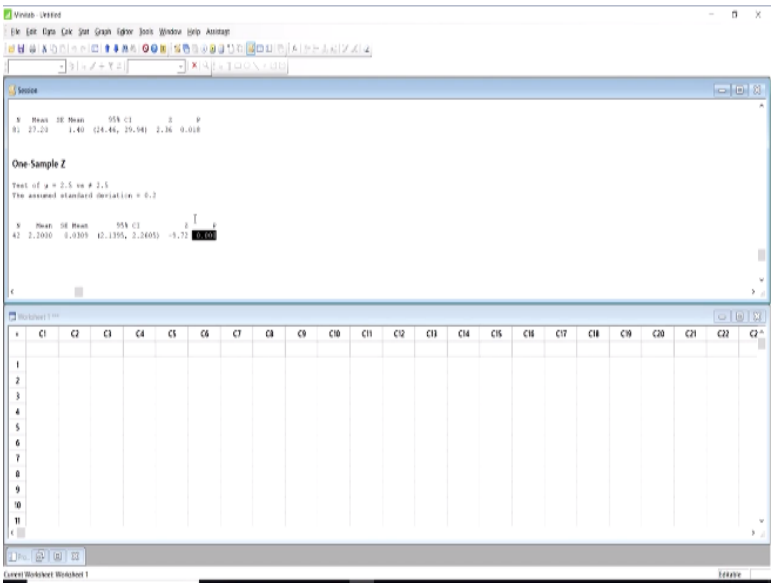
So what would be the null and alternative hypothesis? So null is this, alternative hypothesis would be is not equal to 2.5. Okay. So you can solve this question again using Minitab.

(Refer Slide Time: 20:28)



So go to Stat, Basic Statistics, 1-Sample Z okay then summarized data then Sample Size, you need to look at sample size. So sample size is 42, sample mean is 2.2, standard deviation is 0.2. We need to perform hypothesis test and hypothesized mean is 2.5. And in fact we have been given this decision criterion of 1.96 standard error which is 0.05 significance level, right. So you can have again hypothesized mean= hypothesized mean and confidence level is 95%, right. So just press OK again OK.

(Refer Slide Time: 21:53)

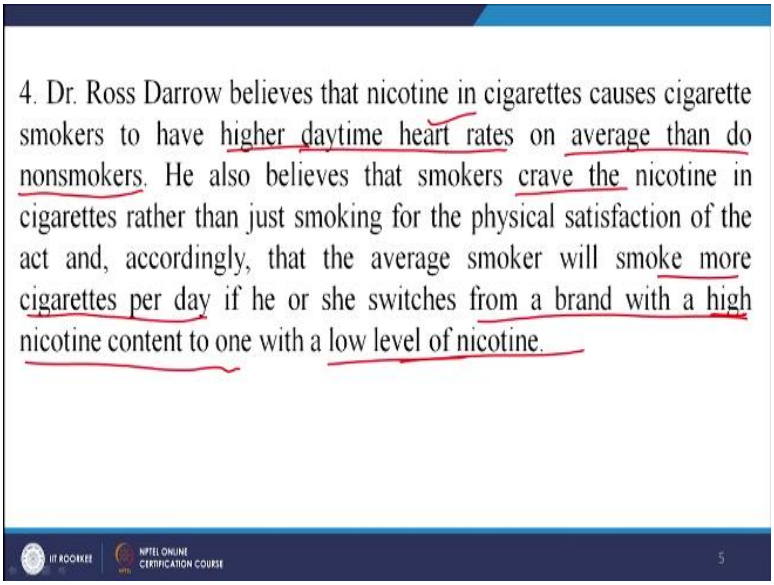


So you will get this output where mean is given, sample size it given, standard error is given, confidence interval is given, P value is given, so P value is 0. And what was our significance level, 0.05. P is < alpha so we will reject null hypothesis since P is < alpha will reject null

hypothesis. And what is our null hypothesis? We will reject null hypothesis and what was null hypothesis? That, bag of oranges will yield this much quarts of juice now.

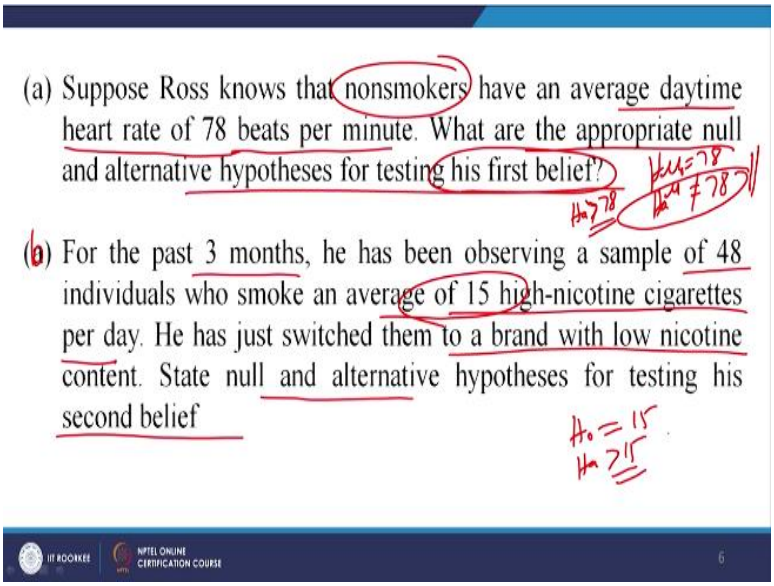
It will not result in yield in this much quarts of juice so this hypothesis is rejected, it would be either more than this or < this. So what it would be, it would be actually < 0.05 because sample mean is 2.2 quarts, right. So this is how we can work out one question on hypothesis testing.

(Refer Slide Time: 23:03)



Let us look at third question. Dr. Ross Darrow believes that nicotine in citrates causes cigarette smokers to have higher daytime heart rates on average than do nonsmokers. So smokers will have higher heart rate than nonsmokers, right. He also believes that smokers crave the nicotine in cigarettes rather than just smoking for the physical satisfaction of the act and, accordingly, that the average smoker will smoke more cigarettes per day if he or she switches from a brand with a high nicotine content to one with low level of nicotine content, so smokers would like to switch from high nicotine content to low level of nicotine content.

(Refer Slide Time: 24:13)

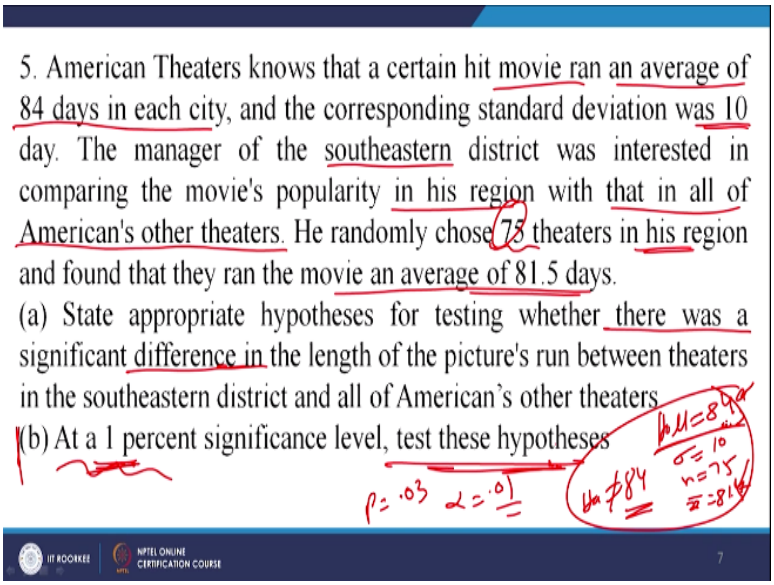


Suppose Ross knows that nonsmokers have an average daytime heart rate of 78 beats per minute is for nonsmokers, right. What are the appropriate null and alternative hypothesis for testing his first belief? So what will be the null hypothesis of course it is simple one, 78 and alternative would be; so this H0, let me write it in this way, alternative not be = 78, right. Would it be eual to, less than or greater than?

Suppose Ross know that nonsmokers have an average daytime heart rate of 78 beats per minute, what are the appropriate null and alternative hypothesis for testing his first belief. In fact, we will have to go back to the question. Initially, he said that, he believes that the smokers have got more heart beat than nonsmokers. So you will have to frame this alternative hypothesis not like this but H alternative is > 78, okay.

Right, so let us look at Part b. For the past 3 months he has been observing a sample of 48 individuals who smoke an average of 15 high-nicotine cigarettes per day. 15 high nicotine cigarettes per day; he has just switched them to a brand with low nicotine content. State null and alternative hypothesis for testing his second belief. So what would be null and alternative hypothesis now? This is now 15, right so null hypothesis is H0=15 and H alternative is > 15. Let us look one more question.

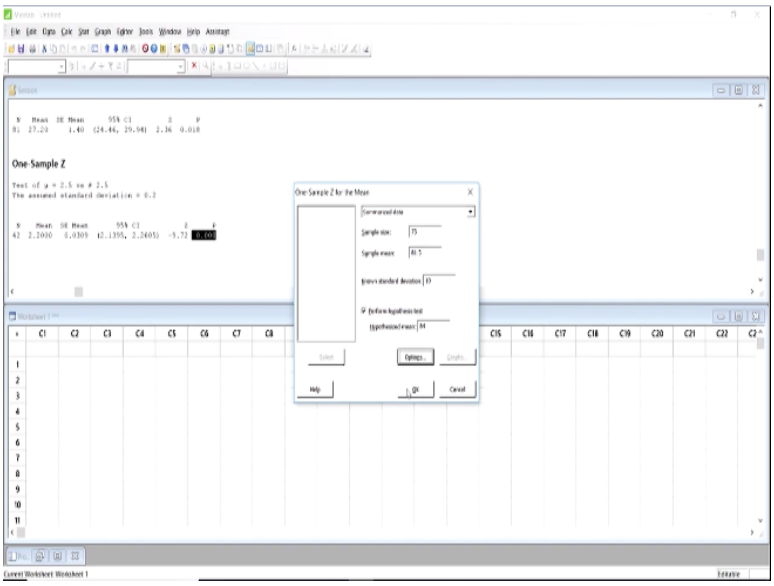
(Refer Slide Time: 27:06)



American Theaters knows that a certain hit movies ran on average of 84 days in each city, and the corresponding standard deviation was 10 days. The manager of the southern state southeastern district, the manager of southeastern district was interested in comparing the movies popularity in his region with that in all of American's other theaters. He randomly chose 75 theaters in his region and found that they ran the movie on an average this much.

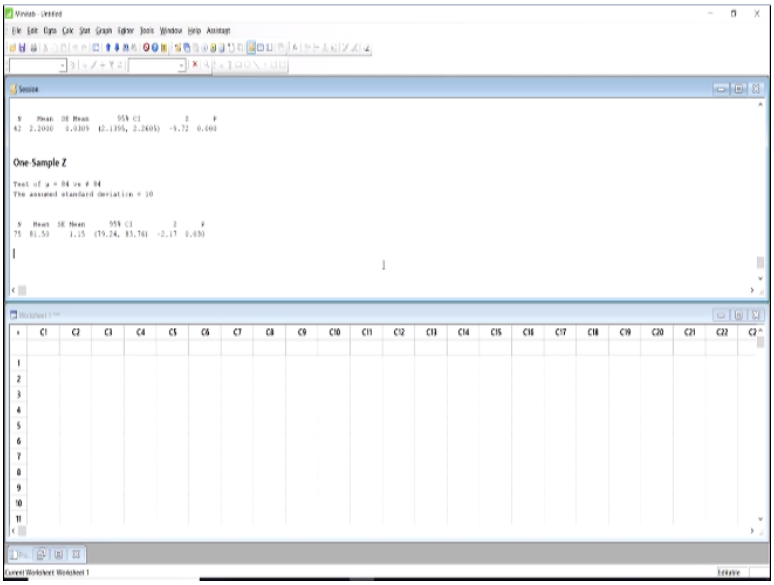
So hypothesized population mean would be what, is 84 days with a standard deviation of 10 days, sample size 75 theaters, sample mean 81.5 days. Now, we have to state appropriate hypothesis for testing whether there was a significant difference in the length of pictures run between theaters in his region as well as other theaters, right. So it would be a very simple one. So one is your null hypothesis and alternative hypothesis would be not equal to 84 days. So we will work out this question using Minitab. So we know that n is 84, standard deviation is 10.

(Refer Slide Time: 28:49)



So we will go to Stat, Basic Statistics, 1-Sample Z, so sample size is 84, sample mean is 81.5, standard deviation is 10 and hypothesized mean is 84; sample size is different right, so hypothesized mean is 84, we will see the sample size again it is 75, right. So we will go to a option, we will click at option, we will test at 95% itself, OK, click OK again OK.

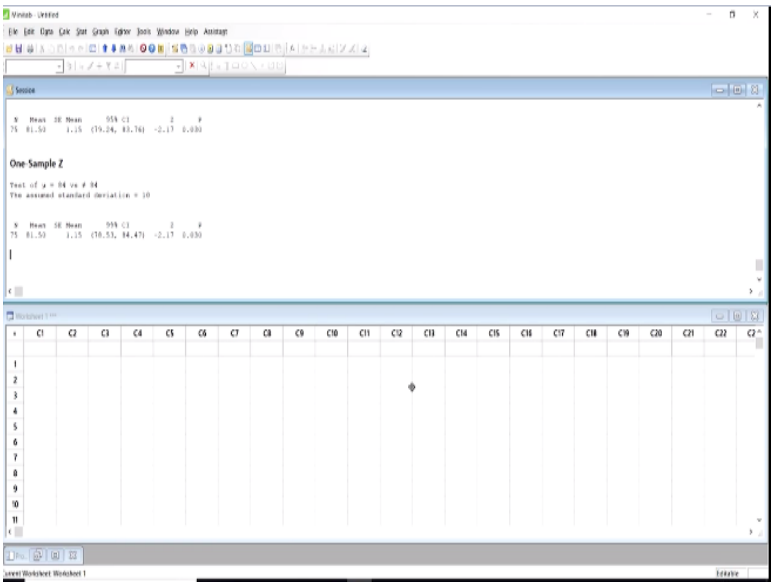
(Refer Slide Time: 29:52)



So P is 0.03 which is lies in alpha, so we will reject null hypothesis, right. So we will say that, we will reject null hypothesis it means the movie ran not 84 days, more than this or less than this, right. But our sample mean is 81.5 days it means movie ran < 84 days, right. Let us look at the second part of it. So you can test this sample hypothesis at 1% significance level. So we will

change this significance value for the same question. We will just change yeah; this is now 99%. We will click at OK, OK.

(Refer Slide Time: 31:00)



So this 0.03, so 0.03, 0.03 is the P value and alpha value is 0.1, is P < alpha? No. So we will not reject Null hypothesis. So P > alpha, so we will not reject null hypothesis; so at this, at 1% significance level we will not reject null hypothesis and at 95 we will reject null hypothesis. So today we will have worked out couple of questions using many types, software. The software helps you in knowing answer very quickly, only the point is you need to go for data entry properly, you need to understand the question properly.

You need to write alternative hypothesis carefully because your answer depends on how you are framing your alternative hypothesis or how correctly you are framing your alternative hypothesis. So these are couple of questions we have seen on one sample test. In next class, we will work out couple of examples on one Sample size test only and if time permits we will take up some questions on two sample test as well. With this, let me complete today's session. Thank you.