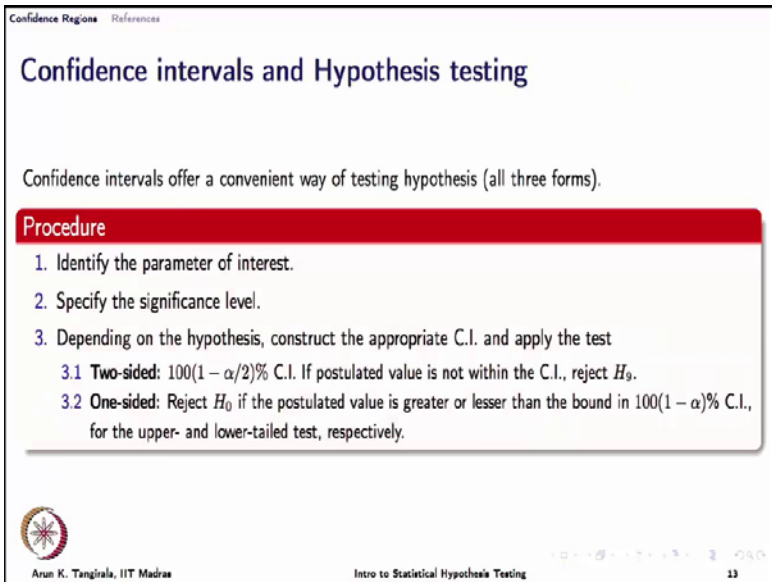
Introduction to Statistical Hypothesis Testing
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture – 14
Confidence intervals and Hypothesis testing

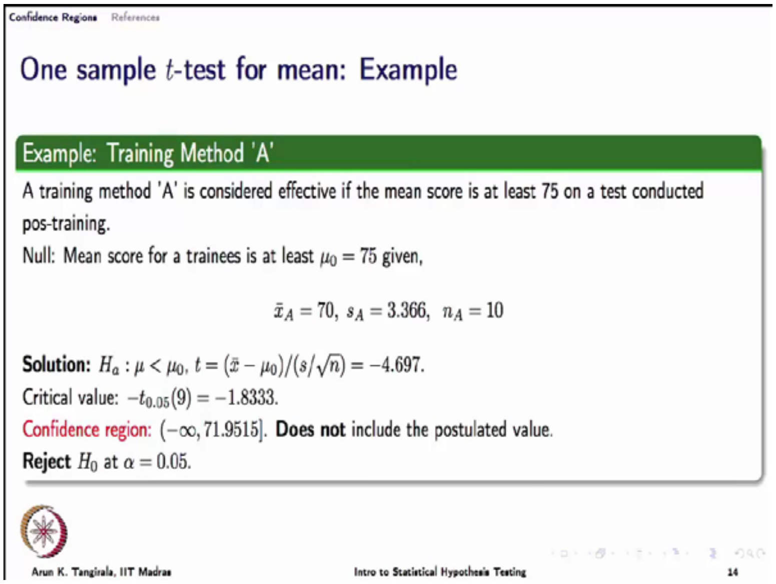
(Refer Slide Time: 00:01)



Let us go back now, to the second objective of this lecture; which is to establish the connection between Confidence interval and Hypothesis testing. I have already explained basic idea; now I am just giving you the systemic procedure, as usual identifying the parameter of interest and specific the significant level alpha.

Now, instead of going through this critical region or p-value approach directly contact the confidence region, depending on the alternative hypothesis. The alternative hypothesis is two-sided one that is of the not equal to 2 types then you construct a two-sided confidence region. If the alternate hypothesis is of the inequality type either lower-tailed or upper-tailed, then you consider one-sided region. Remember, we just derive the one-sided confidence intervals for the left-tailed one we derive the upper bound for the parameter; and for the right-tail or the upper-tail test, we derive the lower bound. And if the postulated value, let us say for the left tail test is greater than the upper bound then the null hypothesis is rejected. And for the upper-tail test, if the postulated value is less than the lower bound then the null hypothesis is rejected that is all.

(Refer Slide Time: 01:30)



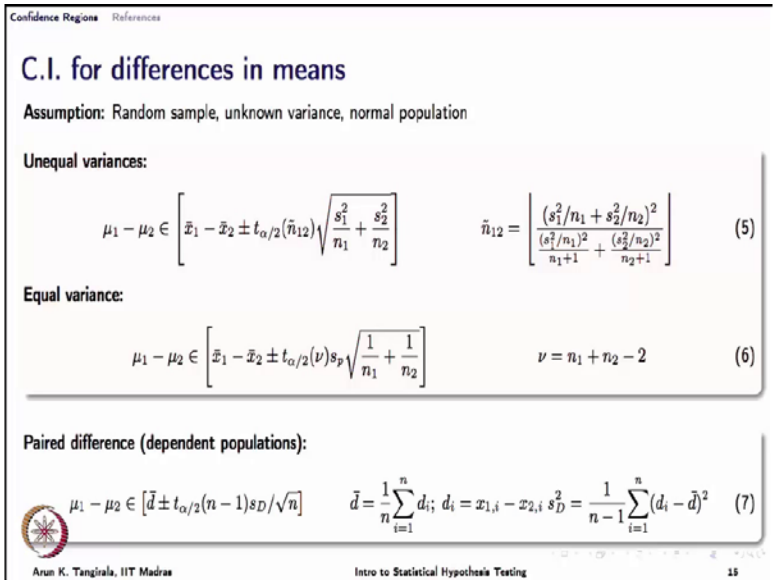
So, let us understand this again through the same example that we have seen earlier and remember I had given confidence interval and so on. So, here this is example of testing the effectiveness of a training method if you recall the data, we have gone through all those detail earlier. What I want to draw your attention to is a confidence region that I have contracted now this is remember what is alternate hypothesis mu is less than 75. So, which means it is lower-tailed or a left-sided test, therefore, I construct an upper bound for the mean; that means the maximum value for the truth is 71 point something if the null hypothesis is true that is well that is how you can think of.

But here we do not say if the null hypothesis is true and so on. We are saying for the data with the sample mean as the estimator, the maximum possible truth is seventy-one point something, but if the postulated truth is greater than the upper bound then I reject the null hypothesis and that is the case here. The postulated value 75 and 75 is not the one of the possible truths what are possible truths anywhere from minus infinity to 71 point something, it is a very wide confidence interval, but at least it is not wide on both sides that is very important; at least one bound should be finite. So, 95 percent confidence, I can say the truth is anywhere between minus infinity and 71 point something; however, the postulated value is 75. Therefore, it is more likely most likely that the null hypothesis is not correct. Therefore, you reject the null hypothesis and that is the answer that we got even with the critical value analysis.

The critical value approach says look the if the null hypothesis is true there we anchor thinks around the null hypothesis that is the prime difference between conducting the hypothesis test the way we have learned until now using the critical value approach and the confidence region approach. In the confidence region approach, we do not really stick to the null hypothesis at all; we just use the data and estimator and construct a confidence region. And now bring null hypothesis into picture and ask, what is the postulated value? If the postulated value false into the confidence region, then we fail to reject the null hypothesis, very simple. So, here there is a subtle, but very important difference between in the way we conduct our hypothesis test. In a critical value approach, we say if the null hypothesis is true then the extreme value that I am willing to tolerate is minus 1.8333.

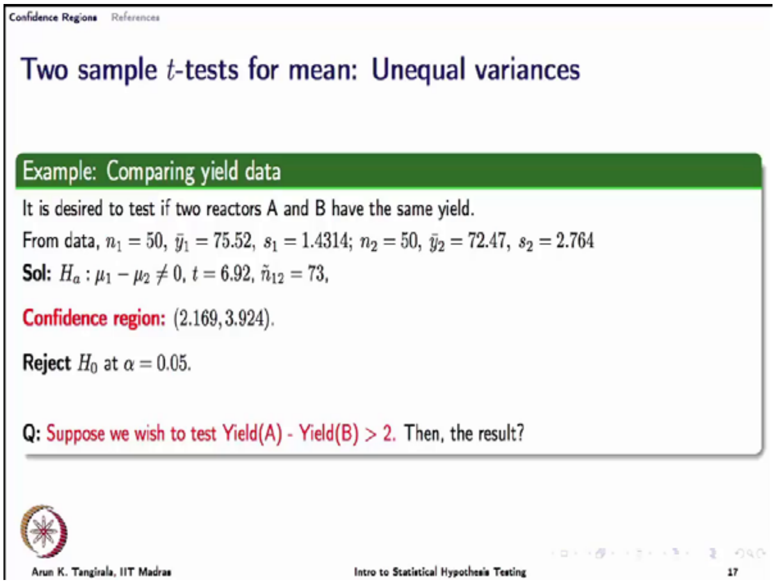
Now, the observed statistic is left lower than or more extreme than what I am willing to tolerate, therefore, I rejected null hypothesis. But now suppose I want to ask what if the suppose I want to re-postulate the truth and say is mu not equals 71, suppose the null hypothesis now change to mu not equals 71, will I reject null hypothesis. Well, again I have to compute the observed statistic and then compare definitely then the null hypothesis would not be rejected; but first region approach, I do not have to compute a statistics at all I can say that 71 is one of the possible true, therefore, the null hypothesis cannot be reject, alright. That is the difference that you should observe between using the confidence region approach and the standard approach for hypothesis testing. And this is the advantage that holds for all kinds of hypothesis test for other parameters too.

(Refer Slide Time: 05:25)



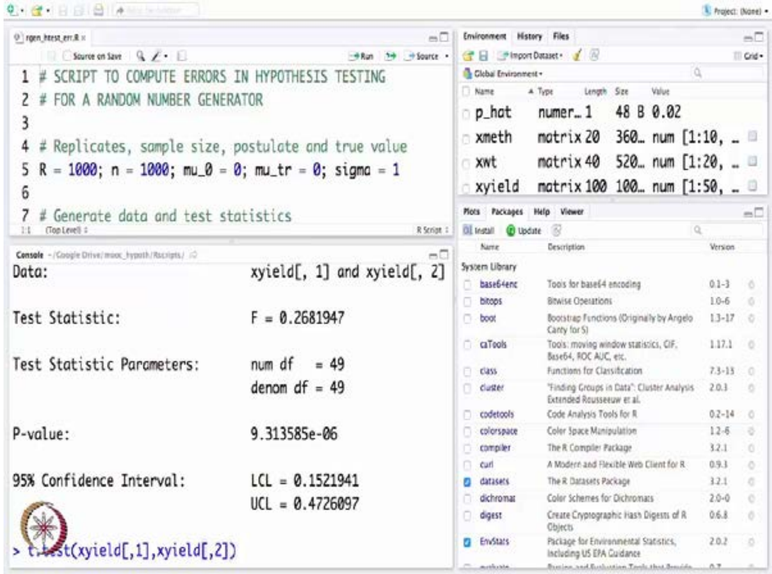
Now, until now we have talked about means naturally we move on to confidents intervals for differences in means one again you can go though the procedure and starting with the sampling distribution for the sample means as the case may be unequal variances, equal variance, paired difference and so on. Pick the sampling distribution appropriate one, write the probability region and then write the 100 times 1 minus alpha percent confidence region. So, I have essentially done that and reported the confidence regions for you. All the symbols are something that we have seen before in the sampling distributions and also in hypothesis testing.

(Refer Slide Time: 06:02)



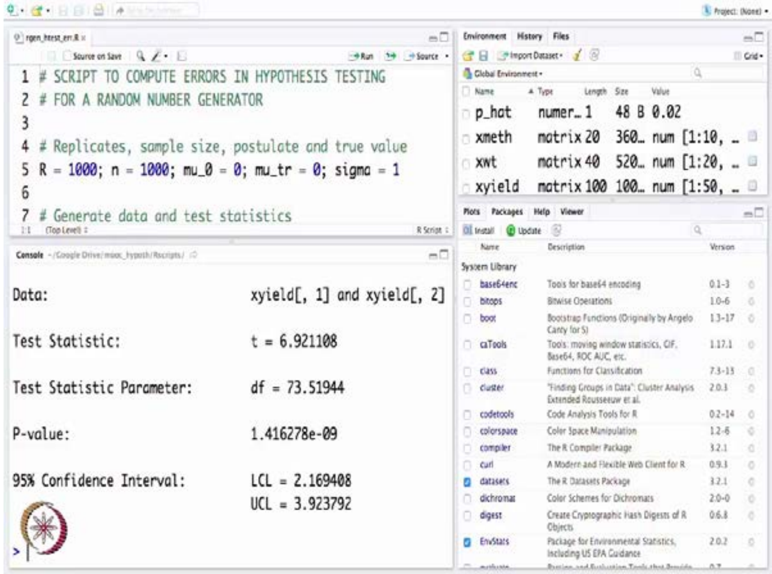
We will quickly move on to the example for the case of unequal variances, where we were comparing yield data. Remember I had the reported the confidence intervals for you. For the yield data, the null hypothesis was that the both reactors give you the same yield and the alternate hypothesis, no, there is a difference we do not know which one gives me more yield, but I do not care, I am interested in either. So, then it becomes a problem of contracting two-sided confidence region; as you can see, I had given the two-sided confidence region. How did I arrive at this, use this equation that I have on the top for the unequal variances case - equations 5, and then contract the confidence region plug in all the values and you will find. In fact, I showed you also last time in R the confidence region.

(Refer Slide Time: 07:02)



If you recall, when we construct the t dot test for the yield data; in fact, if I were had to construct the hypothesis test a standard way I have to specify the postulated truth, but I am if I am only conducting confidence region all I have do is specify the alternate hypothesis. And here the default alternative is that is a two-sided one and a variance dot equal is false; that means unequal variance case. I do not have to worry about the postulated difference at all.

(Refer Slide Time: 07:42)

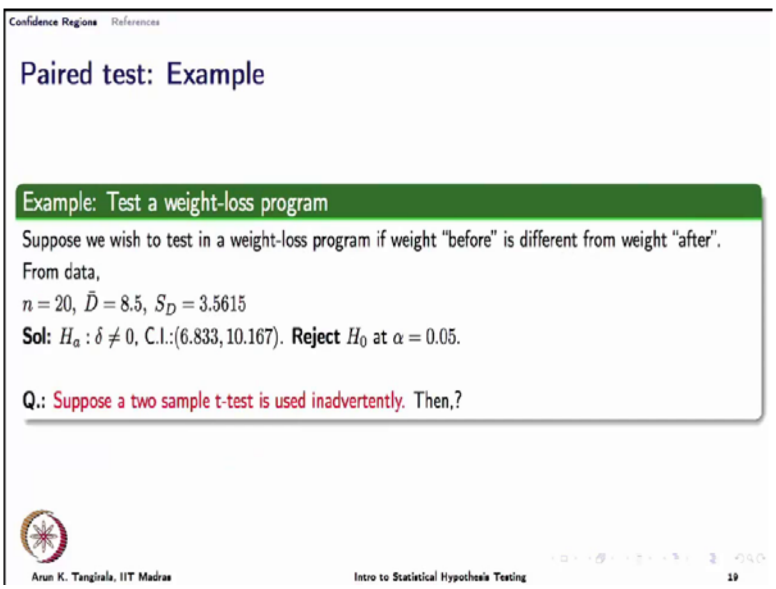


So, let us do that, and you see the confidence interval coming out of the t dot test here, which is what I have reported for you in the slides. In fact, you can see even if I change the postulated value, the confidence interval remains the same, alright. What would change is of course the p-value. So, you see if I change the null hypothesis then the confidence interval remains the same, and the observed statistic is going to change; obviously, because observed statistics depends on the postulated value. What I am trying to show here is the confidence interval construction is independent of the null hypothesis, which what I have explained detail earlier that is the advantage of using the confidence interval approach, alright.

So, let us go back to the problem the postulated difference is 0 whereas the confidence region for the differences is given there, and 0 is not contained in the confidence region, this 95 percent confidence region, and therefore we reject the null hypothesis. Now, we can ask the question, what if I want to a conduct hypothesis test where I want to ask that the difference between the yields of A and B reactors are 2. And let us say the alternate hypothesis is that yield A gives me 2 percentage more than the yield of B. So, what is the difference now do I have to recalculate the confidence region, no, I do not have to. Now the postulated different is 2. So, again I go back to the same confidence region, and ask if 2 is contained in the confidence region, it is not. Therefore, I once again reject the null hypothesis that the difference between the yield of A and B is 2 in favor that is not equal to 2. But suppose alternate hypothesis is greater than 2 that I have asked here then what kind of confidence region should I construct, then I have to construct a one-sided confidence region, where I construct the lower bound. So, here earlier I have just answered the question when the alternate hypothesis yield A minus yield B is not equal to 2. But suppose I want to answer yield A minus yield B is greater than 2, then we will have to construct a one-sided confidence region, and I leave that as a small exercise to you and see if this postulated difference now is contained in the one-sided confidence region.

So, to summarize, the confidence region depends on the alternative hypothesis that is a type of alternative hypothesis, whether it is inequality type or the not equal type; that means, whether it is a one-sided one or a two-sided one, but it does not depend on the postulated value. Once you construct a confidence region, you can test with all possible postulates so long as the alternative hypothesis is of the same type that is all very good.

(Refer Slide Time: 11:20)



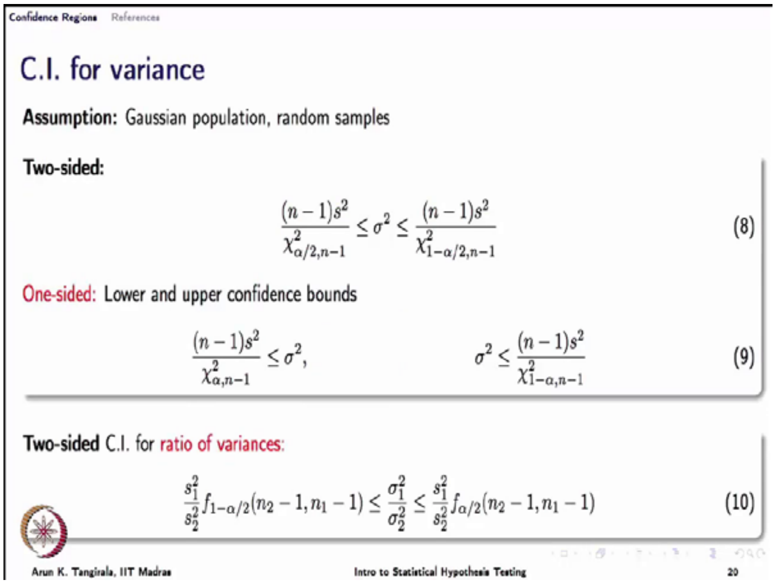
So, let us quickly move on to the paired test thing where we look at the weights in a weight-loss program before and after. And here to contract the confidence region we use this expression in equation 7, provided we are looking at a two-sided one and this is that of a two-sided one. We want to see if there is a difference between the weights in a weight-loss program. And the confidence region is reported here using the formula that I had given in equation seven. And the postulated difference is 0;

that means, there is no difference between the weight before and after, but the confidence interval does not contain that, and therefore, we reject the null hypothesis and conclude that 0 is not a possible truth. And, now you can ask what can I say that the different is may be let us say 8 and the alternative hypothesis difference is not equal to 8 then in that case the null hypothesis would not be rejected because 8 is a possible value.

But if you were to ask the difference between let us say even the difference between the weights before and after is 10 kg’s, very nice, still the null hypothesis would not be rejected. But suppose the alternate hypothesis of the type the difference is greater than 10 then I will have to construct a one-sided confidence region, and then answer that question of hypo the alternative hypothesis which is the weight difference is greater than 10. So, again I leave that as an exercise to you. And also what you should do is construct a confidence region, assuming like we did in the hypothesis test case. And I would ask you to do this as well there, treating the samples before and after as independent; that means, we apply a 2 sample T test there we revaluate that is we revealed the critical value, recalculate the critical one and we use a different kind of a statistic.

Here as well, suppose you were to treat this in advertently as case of the populations being independent of each other before and after, then the confidence region construction would be different. In what way, instead of using equation 7, you would probably use equation 6 or equation 5 depending on whether you assume equal variability or unequal variability, clearly the confidence intervals would be different. And what I want you to do is do is go ahead and calculate those confidence regions, either you can use the formula here or I have the data set there uploaded for you on the website; from where you can run the t dot test and obtain the confidence intervals. Setting variance dot equals false and paired also to false. And see you reject the null hypothesis or fail to reject null hypothesis, very good.

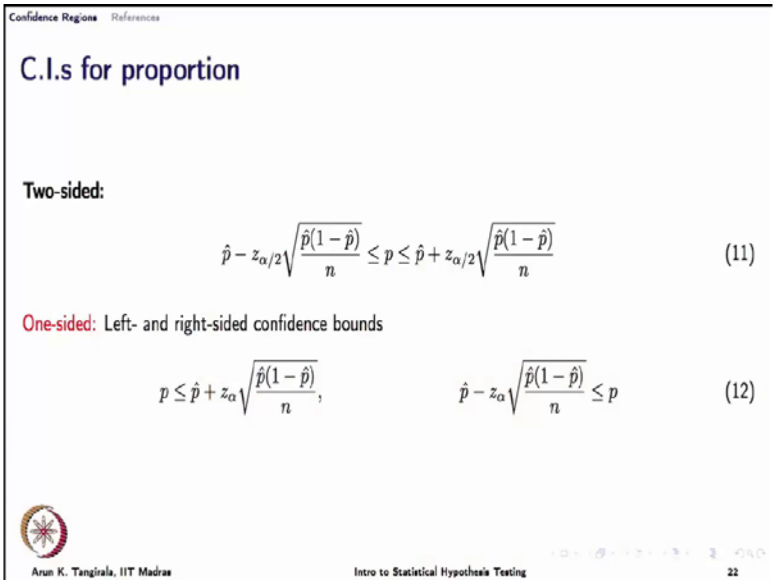
(Refer Slide Time: 14:20)



Now, let us move on and talk about the confidence intervals for other parameters. This is going to be a swift thing, we do not just (Refer Time: 14:27) breeze through this because as I said the derivation of this confidence region is now kind of standard. Look up the sampling distribution of the estimated right the probabilistic region write the confidence region, and you would obtain the confidence intervals that I have given you in equation 8, 9, and 10 and so on. So, once again here we have a two-sided confidence region and a one-sided confidence region for the variances. And we can also give a two-sided confidence region for ratio of variances.

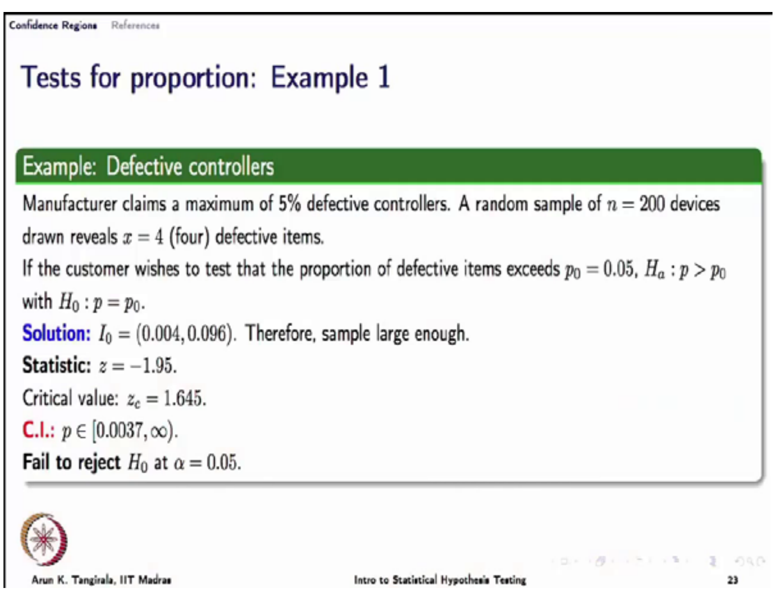
So, let us recall an example quickly where we are comparing variability of yields, right. And we have gone thought this example in the previous lecture and we had rejected the null hypothesis based on the critical value approach. Now, if I use the confidence interval approval, the confidence region for the ratio of the variances is 0.152 and 0.473. Why did we construct this kind of a confidence region, because alternate hypothesis is of the not equal to type. So, it a two-sided confidence region, and it does not include the postulate value which is one, that means, variance are equal therefore, we reject the null hypothesis. So, excellent, we are actually now hopefully becoming experts in the confident interval based hypothesis testing. And finally, the confidence intervals for proportion again to read this out would be like reading small sermon will not do that.

(Refer Slide Time: 16:02)



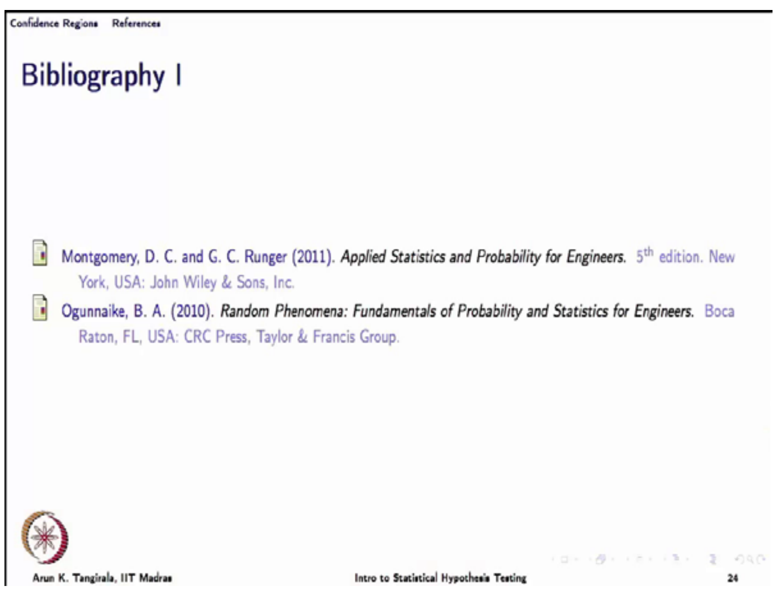
Again you will you can actually read, derive these expressions confidence region two-sided and one-sided starting from the sampling distributions. Remember that all the assumptions that we made in deriving the sampling distributions whole here as well that means, if any of those assumptions are violated, these confidence intervals are not valid any more, you will have to re-derive them.

(Refer Slide Time: 16:30)



And let us look at this example for the defective controllers, where we looked at this motivating example when the manufacture claims a maxim of 5 percent defective controls and so on. And there we had failed to reject the null hypothesis based on the critical value approach; and here the test is actually a not a two-sided one, because alternative hypothesis is of the inequality type p greater than p naught and p naught is 0.05. Therefore, I need to contract a one-sided confidence region, which involves contracting the lower bound on the proportion; and the lower bound returns out to be 0.0037. I have just calculated that in R by using this expression in equation twelve all right. And it turns out that the postulated value is within the confidence region postulated value is 0.05 that means, it is greater than the lower bound. So, fine, I fail to reject the null hypothesis. In fact, the manufacturer could have even claimed 0.01 and still the null hypothesis would not have been rejected, correct so that is the thing here with the confidence region, using the confidences region for hypothesis testing.

(Refer Slide Time: 17:56)



And with this, we come to a close of course, I could have discussed the differences in proportions, but again it would be a routine thing. What I would advice is go to read up book on by Montgomery and Runger, or the other book look up the confidence region for the differences in proportions work out the example that we did in hypothesis testing. And see if you arrive at a same conclusion that we made with the critical value approach or the p-value approach. Now, there is one point that I want to make, it was easy to derive the confidence regions because we knew the sampling distributions that is step two is critical. If I do not know sampling distribution, then there is no way I can derive a confidence region because I cannot right the probabilistic region. Here it was easier to derive the confidence regions because analytical expressions for the sampling distribution are available.

Now in practice, in reality, there may be many situations in which we do not have the sampling distribution expression at all. In linear regression, we have; but in non-linear regression may be for large sample sizes you have, but in very complicated estimation exercises which are not uncommon today; it is quite difficult to derive the sampling distribution analytically. Then how does one construct confidence region, does one contract at all, well, yes, and that is though the use of Monte Carlo simulations or you can say bootstrapping method or surrogate data analysis and so on which typically forms a topic in an advanced course on estimation or advance coursed on statistical inferencing today.

We are not going to talk of bootstrapping in this course, but I thought you should know this especially those of you who are conducting research in some kind of an advance data analysis or advance statistical inferencing. So, hopefully you enjoyed this lecture; in the next lecture, we are going to look at hypothesis test in the linear regression scenario where we will setup the hypothesis null and alternative hypothesis for the slope and intercept. And also for correlation to begin with and then at the same time, we also talk of confidence intervals then ask in a linear regression, what are the typical hypothesis test not all of them, but a few of them that we conduct. Followed by the closing lecture, where will talk of sensitivity of hypothesis test to one of the main factors, which is sample size. Ok then.

See you in the next lecture.