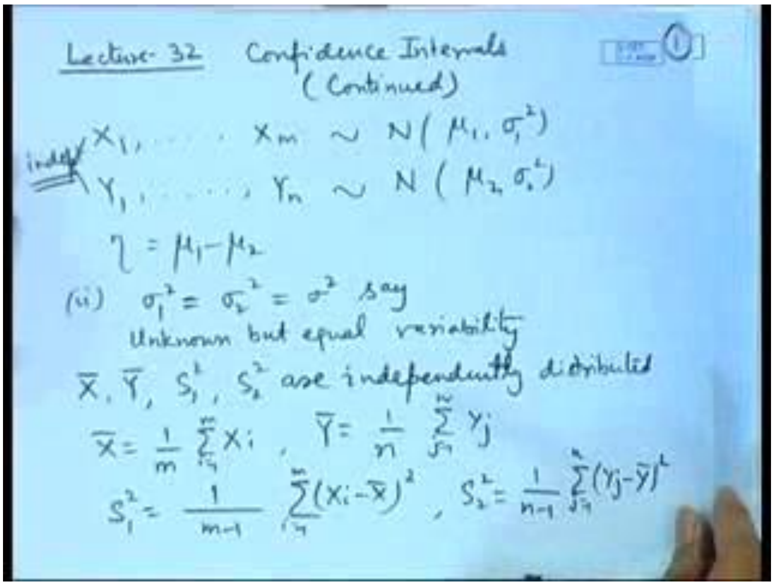
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 63
Confidence Intervals – III

We continue our discussion on the Confidence Interval Estimation. Let me repeat the set up here.

(Refer Slide Time: 00:35)

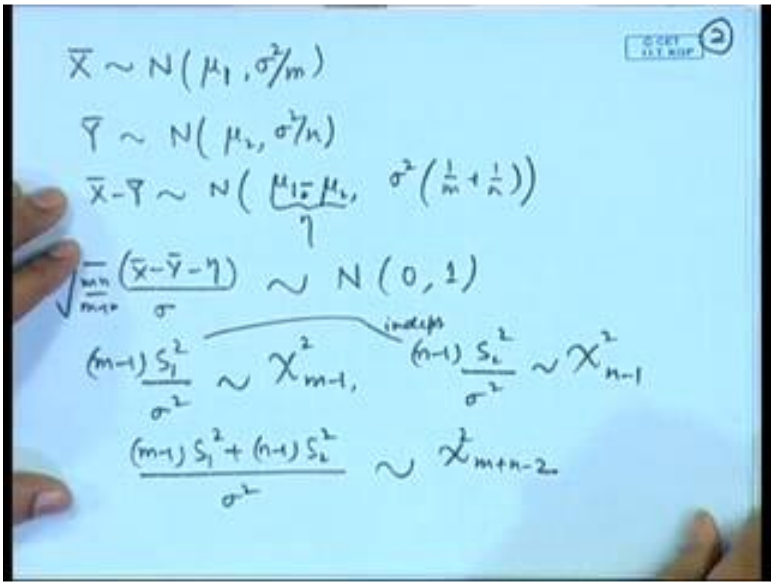


We are interest in the comparison of the means of two normal populations. So, we have a sample X 1, X 2, X m from normal mu 1 sigma 1 square; Y 1, Y 2, Y n is another independent random sample form normal mu 2 sigma 2 square population. These two samples are taken to be independent. So, we are interested in the confidence interval for mu 1 minus mu 2, let us call it eta. So, we have earlier found out the confidence interval for the situation when sigma 1 squares and sigma 2 square are known. But in general the sigma 1 square and sigma 2 square may be unknown and we may be required to find out the confidence interval.

So, we take the case two that sigma 1 square and sigma 2 square are unknown but equal, that is unknown but equal variability. Now this type of situation may arise for example, you are looking at two brands of certain product. Now the variability of the say average life for example, it may be same, but average life themselves may be different. In such cases this model is useful. Let us look at the analysis of this. So, as we have seen that the sampling distributions of X bar, Y bar, S 1 square, S 2 square will be of interest here. So, X bar, Y bar, S 1 square and S 2 square are independent. In the sampling form normal distribution we know this fact independently distributed. Here X bar is 1 by m sigma xi i is equal to 1 to m, Y bar is the mean of the second sample that is 1 by n sigma yj j is equal to 1 to m.

If we consider the sample variance of the first sample that is 1 by m minus 1 sigma xi minus X bar whole square i is equal to 1 to m and S 2 square is equal to 1 by n minus 1 sigma yj minus Y bar whole square that is the sample variance of the second sample.

(Refer Slide Time: 03:22)

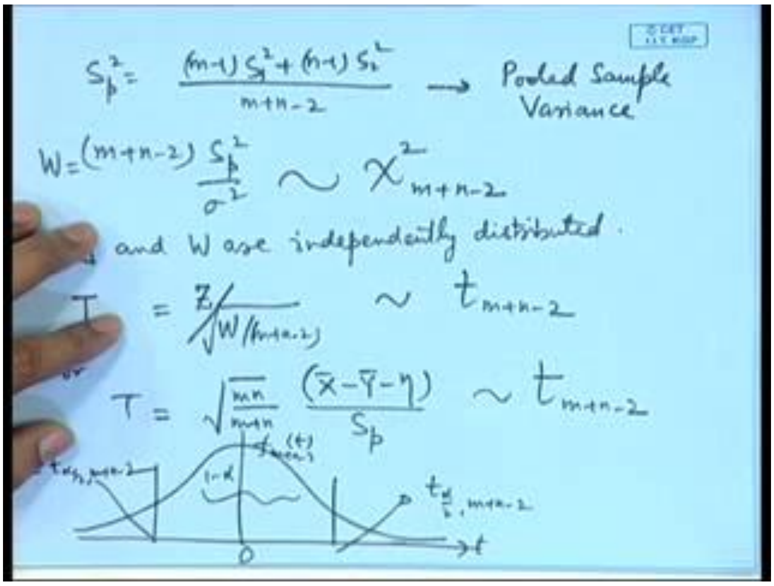


If we consider these quantities then we have the following observations; that is X bar follows normal distribution with mean mu, mu 1 and variance sigma is square by m. So, here sigma 1 square and sigma 2 square both are same. Then Y bar follows normal mu 2 sigma square by n. If we consider here X bar minus Y bar that will follow normal with mean mu 1 minus mu 2 and variance will be sigma square 1 by m plus 1 by n.

So, if we want, this is the quantity eta. So, we get X bar minus Y bar minus eta divided by sigma and root of this that is root of mn by m plus n that will follow a standard normal distribution. However, this involves the unknown parameter sigma also, so we cannot straight away use it as a pivot quantity. So, we need estimator for sigma also. So, we can get it here by considering m minus 1 S 1 square by sigma square follows chi square on m minus 1 degrees of freedom and n minus 1 S 2 is square by sigma square follows chi square on n minus 1 degrees of freedom.

Once again these two quantities are also independent. So, I can add these and we get m minus 1 S 1 square plus n minus 1 S 2 square divided by sigma square that follows chi square distribution on m plus n minus 2 degrees of freedom.

(Refer Slide Time: 05:17)



Let me define a quantity S p square that is equal to m minus 1 S 1 square plus n minus 1 S 2 square divided by m plus n minus 2; that is pooled sample variance. If we use this pooled sample variance then what we are having is m plus n minus 2 S p square by sigma square is following chi square distribution on m plus n minus 2 degrees of freedom.

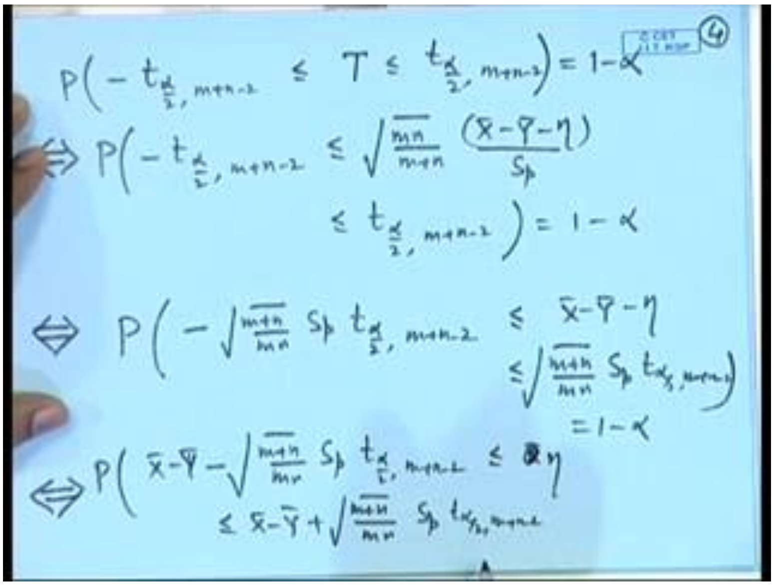
Now, we have the distribution of X bar minus Y bar minus eta divided by sigma multiplied by a constant as a standard normal distribution, and let me call this quantity as say Z and I have a quantity let us call it say W this is having a chi square distribution. Another thing we can notice here this is that Z involving only X bar and Y bar and W's involve in only S 1 square and S 2 square that is S p square. So, Z and W are independently distributed.

So, if they are independently distributed I can look at the distribution of Z divided by W by m plus n minus 2 square root that will have t distribution on m plus n minus 2 degrees of freedom. So, this quantity is equivalent to root mn by m plus n X bar minus Y bar minus eta divided by S p. So, that follows t distribution on m plus n minus 2 degrees of freedom.

Now let us observe: given the samples xi's and y j's we can evaluate X bar Y bar and S p and this involves the parameter eta for which we need the confidence interval and the distribution of this quantity is free from the parameters of the distribution. Therefore, this value t can be considered as a pivot quantity and we can make use of this to construct a confidence interval for eta that is mu 1 minus mu 2.

So, we look at the t distribution it is symmetric about the axis then it is symmetric about 0. So, this is f m plus n minus 2 t. So, we look at the point here this point is t alpha by 2 m plus n minus 2, and we have on the left hand side the single point that is minus t alpha by 2 m plus n minus 2. So, this intermediate probability is 1 minus alpha.

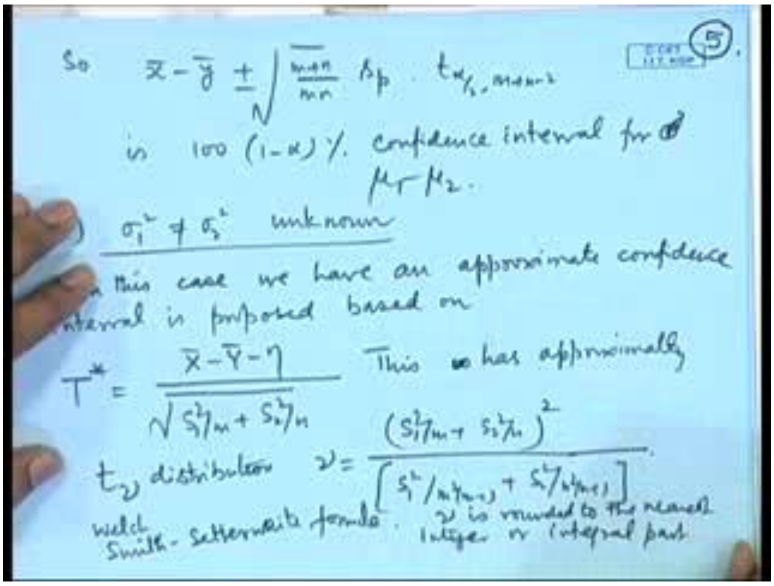
(Refer Slide Time: 08:37)



And we are in a position to write a statement that probability of minus t alpha by 2 m plus n minus 2 less than or equal to t is less than or equal to t alpha by 2 m plus n minus 2 that is equal to 1 minus alpha. So, expanding this t and then adjusting the terms we will be able to construct a confidence interval for mu 1 minus mu 2. So, t is here a square root of mn by m plus n X bar minus Y bar minus eta divided by S p; that is less than or equal to t alpha by 2 m plus n minus 2 that is equal to 1 minus alpha.

So, this is equivalent to root m plus n by mn S p t alpha by 2 m plus n minus 2 less than or equal to X bar minus Y bar minus eta less than or equal to square root m plus n by mn S p t alpha by 2 m plus n minus 2; that is equal to 1 minus alpha. So, this means that X bar minus Y bar minus root m plus n by mn S p t alpha by 2 m plus n minus 2 less than or equal to X bar less than or equal to eta less than or equal to X bar minus Y bar plus; that is equal to 1 minus alpha.

(Refer Slide Time: 10:58)



So, in the situation when the variances of the two populations are unknown, but equal the confidence interval for mu 1 minus mu 2 is obtained as X bar minus Y bar plus minus a square root m plus n by mn S p t alpha by 2 m plus n minus 2. So, this is giving a 100 1 minus alpha percent confidence interval for mu 1 minus mu 2.

Notice here is that since the variances where assume to be equal we are making use of a pooled sample variance. Now one may ask a question that in place of this suppose we consider simply S 1 is square or S 2 is square only, because in that case also we are getting a variable which is having a distribution free from the parameters, so why not use only this or only this. The question is that if we use only say S 1 square then the degrees of freedom that we will get for the t variable will be m minus 1. So, if we get only m minus 1 then in that case the interval will be having the width X bar minus Y bar plus minus a square root m plus n by mn.

Now this term will not come here rather we will have S 1 only this coefficient will not come here, here we will have only S 1 and the degrees of freedom will be m minus 1. Naturally the length of the interval will increase if we have less degree of freedom. So, in order to get more accuracy or you can say more precision we need a smaller interval with the same confidence coefficient. Therefore, it is beneficial to use more information here.

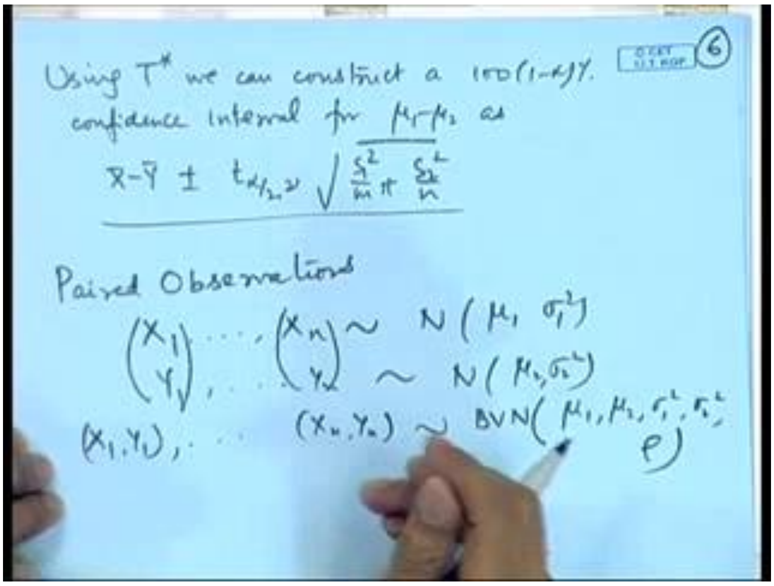
Let us take the case when both mu 1 and mu 2 may be unknown. Then let us look at the procedure here that has helped us to create this confidence interval. The procedure that we adopted was that the distribution of S p square by sigma square that is chi square and the Z variable that we utilized that also as a sigma in the denominator. So, we were able to get rid of this. If the variances are not equal then in the first place we will be getting sigma 1 square here and here we will get sigma 2 square, so when we add the two terms in the denominators I will get S 1 square by sigma 1 square and here S 2 square by sigma 2 square. And the same thing will happen with the Z also, where we will get sigma 1 square by m plus sigma 2 square by n.

So, in no way by taking the ratios I can get rid of sigma 1 square and sigma 2 square. Actually it turns out that there is no exact confidence interval; that means, the interval which is having the length a shortest length and as well as a fix confidence coefficient; that means, a distribution free term here not getting. In this case this is known as a Behrens-Fisher situation. So, we will consider this case sigma 1 square is not equal to sigma 2 square and unknown; that means, they are completely unknown. In this case a approximate an approximate confidence interval is proposed based on; let us call it t star that is X bar minus Y bar minus eta divided by square root S 1 square by m plus S 2 square by n.

So, how this has come? In the first case where sigma 1 square by m plus sigma 2 is square by n was there we have simply replaced sigma 1 square and sigma 2 square by their unbiased estimates. So, it was proved by Welch etcetera that this is having has approximately t distribution on nu degrees of freedom, where nu is given by S 1 square by m plus S 2 square by n whole square divided by S 1 square by m square into m minus 1 plus S 2 square by n square into n minus 1; which is by Welch and it is known as a Smith Satterthwaite formula.

Now this need not be an integer. So, nu is rounded off to the nearest integer or integral part. That means, suppose it is turning out to be 11.3 second we take only 11.

(Refer Slide Time: 17:04)



So, using this one can write a confidence interval; using T star we can construct a 100 1 minus alpha percent confidence interval for mu 1 minus mu 2 as X bar minus Y bar plus minus t alpha by 2 mu a square root of S 1 square by m plus S 2 square by n. This will be the confidence interval. Then there is no information about the equality of sigma 1 square and sigma 2 square.

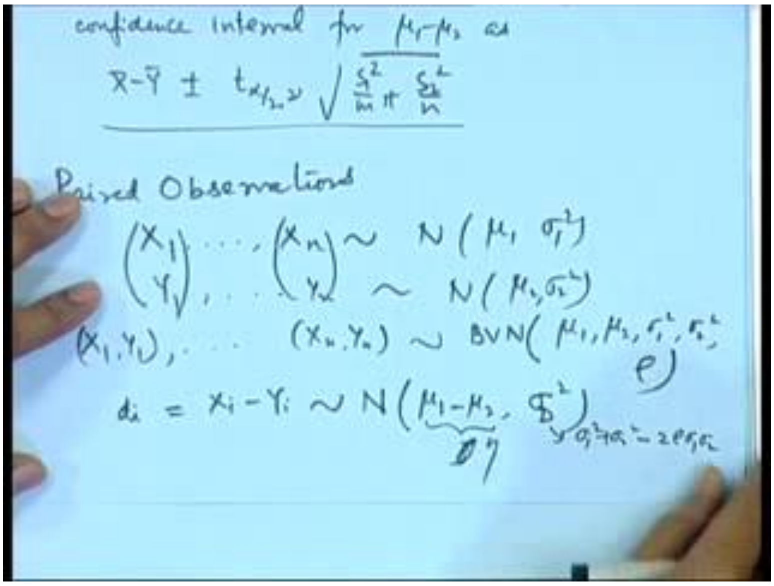
Now there is another situation which occurs quite frequently; for example, we are considering the comparison of the two training procedures. So, suppose there are two training procedures for certain learning. So, we select say 10 pupils and we give them instructions using one training procedure a test in conducted to measure the outcome of that. Now, for the same set of 10 pupils another learning procedure is important for a fix period of time and another test is conducted.

Now the scores are not independent because our subjects are not independent, same set of people has been selected. For example, it could be some weight reduction procedures like the fatty people are there and we are giving them certain weight reduction program. So, by taking certain procedure for one month their weight is reducing by this much. Now for the same set of people another procedure is adopted then how much weights have been reduced. So, we compare the same set of people with respect to their course.

So, here this is related to paired observations. So, here although you are saying X 1, X 2, X n say follow normal mu 1 sigma 1 square and Y 1, Y 2 Y n they follow normal mu 2 sigma 2 square, but actually the sample has not been selected in this way because these observations may be paired. So, basically the model becomes that X 1 Y 1, X 2 Y 2, X n Y n this is having some sort of bivariate normal distribution with parameters mu 1 mu 2 sigma 1 square sigma 2 square and some correlation coefficient rho may be there. Once again we are interested in the interval for mu 1 minus my mu 2 that is we want to look at the difference in the average effectiveness etcetera.

A simple procedure for this is obtained by using the linearity property of bivariate normal distribution, because we know that if the random variable X Y is having a bivariate normal distribution then any linear combination a x plus b y is again having a univariate normal distribution.

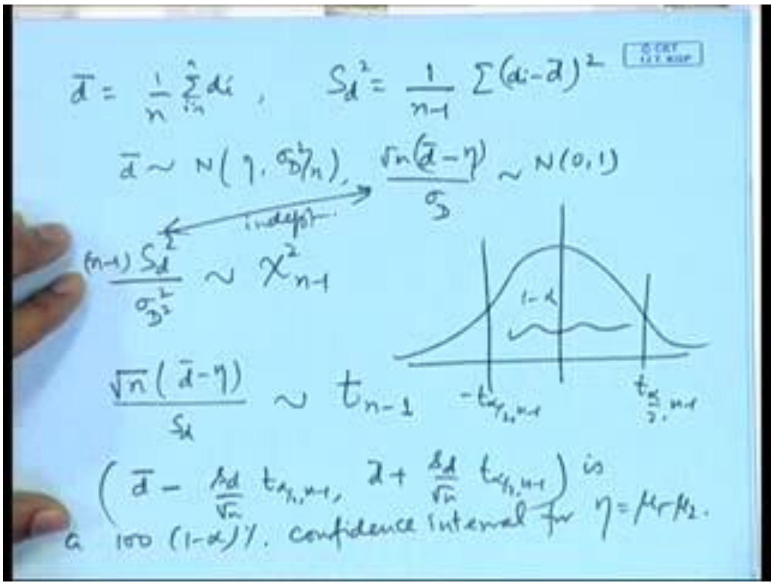
(Refer Slide Time: 20:33)



So, here if I make use of say observations let me call it di that is equal to xi minus yi, then that will follow normal distribution with mean mu 1 minus mu 2 and some variance let me call it sigma d square; actually it will be sigma 1 square plus sigma 2 square minus twice rho sigma 1 sigma 2. So, sigma 1 square plus sigma 2 square minus twice rho sigma 1 sigma 2, will let me call it sigma d square. That is not important here because they are all unknown and we need only an estimate of this, because we are interested here in the confidence interval about mu 1 minus mu 2.

So we can make use of; now this looks like a problem of the confidence interval for a mean of a normal distribution which we have done in the first place.

(Refer Slide Time: 21:30)



So, we can consider say d bar as 1 by n sigma di i is equal to 1 to n, and we consider S d square as 1 by n minus 1 sigma di minus d bar whole square. Say if we look at this then we can see the d bar follows normal eta sigma d square by n and from here we can get d bar minus eta root n by sigma d follows normal 0, 1. Also n minus 1 S d square by sigma square by sigma d square that will follow chi square distribution on n minus 1 degree of freedom. And once again these two variables will be a statistically independent. So, using this we can write a square root n d bar minus eta divided by S d that will be having a t distribution on n minus 1 degree of freedom.

Now, observe this function here it is involving the random variables that are observations xi and yi's, d bars are the mean calculated from the differences and S d square is calculated as the variance of the difference observations. And here the parameter of interest eta is appearing and sigma e is squares etcetera are absent here. So, this can be used as a pivot quantity and we get a confidence interval by writing down from the distribution of the t on n minus 1 degree of freedom. So, this probability is 1 minus alpha and we get d bar minus S d by root n t alpha by 2 n minus 1 to d bar plus S d by root n t alpha by 2 n minus 1. So, this becomes 100 1 minus alpha percent confidence interval for eta that is equal to mu 1 minus mu 2.

So, we observe here that all these cases are differently handled; that is when we observe a sample we have to look at carefully. So, if the variance is known to us then we have some procedure, if the variances are unknown but we suspect that the variances may be equal then we have another procedure, if the variances are completely unknown then we have another procedure. On the other hand if the sampling is not done in the independent fashion; that means, we have correlated observations then we have may arrange the data in a paired way and then we can apply a pairing formula.

So, the confidence interval for the same parameter mu 1 minus mu 2 when we are sampling from to normal populations it is dependent upon the situation. We have to a statistician has to carefully see that which type of method will be adopted here for finding out the confidence interval, otherwise you will be coming up with the faulty conclusions.