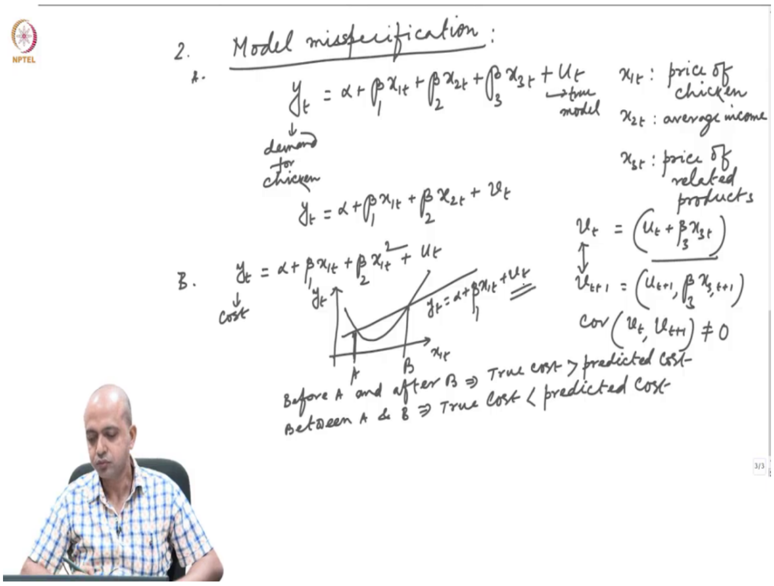
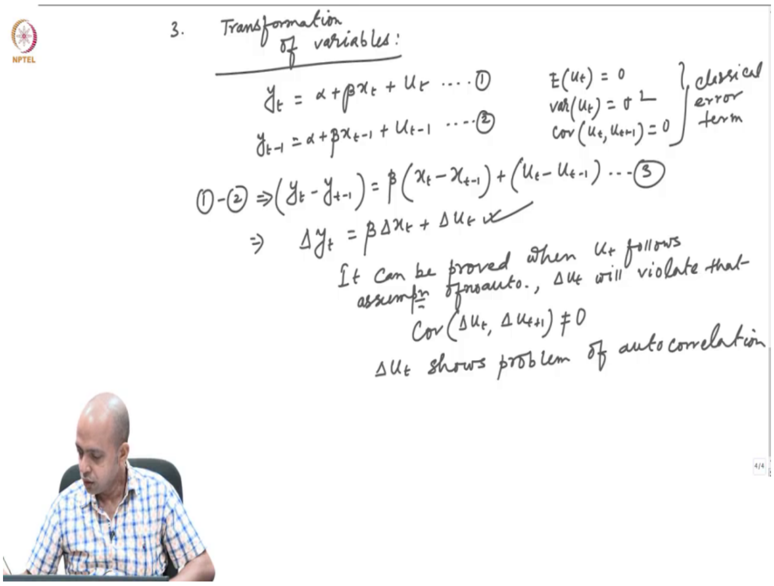
Introduction to Econometrics
Professor Sabuj Kumar Mandal
Department of Humanities and Social Sciences
Indian Institute of Technology, Madras
Lecture 49
Relaxing the assumptions of CLRM-Multicollinearity and Autocorrelation Part – 6

(Refer Slide Time: 0:25)





Another reason for which autocorrelation problem may arise is due to transformation of variable. How? Let us say you are estimating this yt equals to alpha plus beta xt plus ut and for some reason you are taking the first difference of both yt and xt. So, that means, if this is equation 1 and we assume that ut follows all the assumption of classical linear regression model that means, expectation of ut equals the 0, variance of ut equals to sigma square and covariance between ut and ut plus 1 is 0.

So, all these properties are satisfied that is why we say that this ut is basically the classical error term. Now, if you take the different yt minus 1 equals to alpha plus beta xt minus 1 plus ut minus 1 and then if you take the difference between 1 minus 2, yt minus yt minus 1 equals to beta xt minus xt minus 1 plus ut minus ut minus 1. So, that means, I can say that delta yt equals to beta into delta xt plus delta ut.

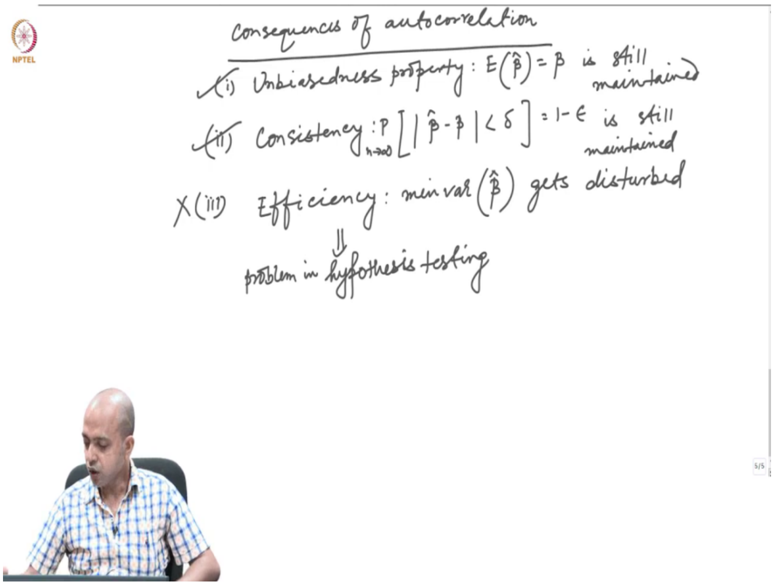
Now, it can be proved that if ut follows all this classical linear regression model assumptions that means, delta ut will not follow those assumptions. So, it can be proved when ut follows assumptions of autocorrelation that means delta ut will violate that. That means covariance between delta ut and delta ut plus 1 is actually not equal to 0.

So, this error term does not follow the assumption of classical linear regression model so that means delta ut shows problem of autocorrelation. This I will say that it can be proved when u t follows assumptions of no autocorrelation actually because ut is a classical error term and the beauty is if ut does not follow this classical assumption, then this first difference error term will follow the classical error term.

That is why we will discuss in a later part; we will show that actually this first difference transformation works better to solve the problem of autocorrelation. But if your original model ut follows the classical assumption that means, if ut does not show any kind of autocorrelation problem and if you take the first difference, then the error term in the first difference model that will show some kind of autocorrelation problem.

So, these are the reasons for autocorrelation and there are some more reasons which are given in your textbook, but these are the prime reasons for autocorrelation problem. Now the next question is consequence.

(Refer Slide Time: 6:35)

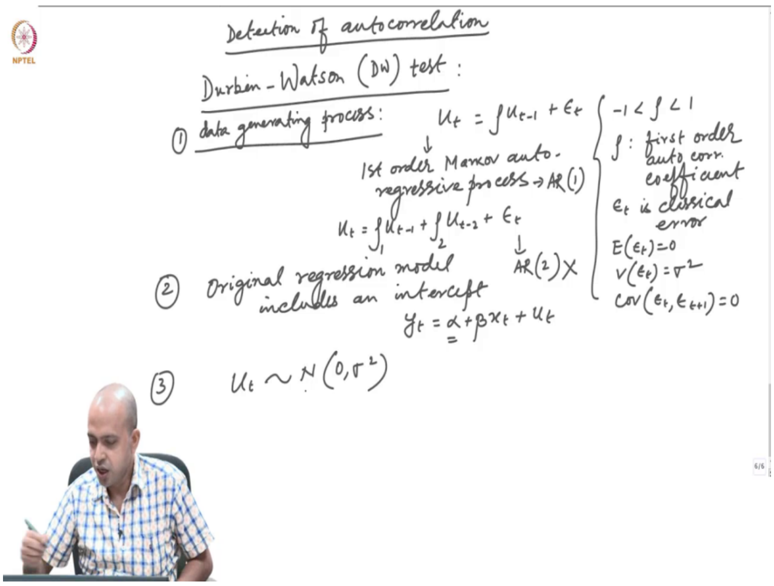


What would be the consequences of autocorrelation? So, consequences of autocorrelation: Now the unbiasedness property means, expectation beta hat equals to beta is still maintained. There is no problem with unbiasedness property even in presence of autocorrelation. Similarly, consistency which means beta hat minus beta less than delta and probability of getting such beta is 1 minus epsilon when n tends to infinity is still maintained.

So, unbiasedness is there, consistency is there but efficiency property that means minimum variance of beta hat gets disturbed. It is no longer the minimum variance. So, this is not maintained and when efficiency that means, minimum variance property gets disturbed your major problem would be in hypothesis testing because hypothesis testing.

Because if the variance of beta hat is not minimum that means your standard error is not minimum. That means, the t statistic what you are going to compute out of this standard error will show some kind of artificial value and as a result of which the decision making out of that inaccurate test statistic will be also erroneous. This is going to be the consequence of autocorrelation. I am not going into the detail of proving that why this is not maintained or why that is maintained because that is beyond the scope of this particular discussion.

(Refer Slide Time: 10:21)



So, what next is how will you detect autocorrelation problem? Detection of autocorrelation and the detection of autocorrelation there is a popular way of doing it, than a specific test statistic developed by two famous econometrician Durbin and Watson that is why the test name is called Durbin-Watson test; in short it is called DW test, before we talk about this Durbin-Watson test, what we need to know they have assumed a specific data generating process.

And what is that data generating process? Data generating process shows the error term is actually ut equals to rho ut minus 1 plus epsilon t, epsilon t that means you see are error term ut is correlated with an here rho is between plus and minus 1 and rho is called first order autocorrelation coefficient and this epsilon t is classical error, that means, it follows all the assumptions of classical regression model in a sense expectation of epsilon t equals to 0, variance of epsilon t equals to sigma square and covariance between epsilon t and epsilon t minus 1 or t plus 1 is 0.

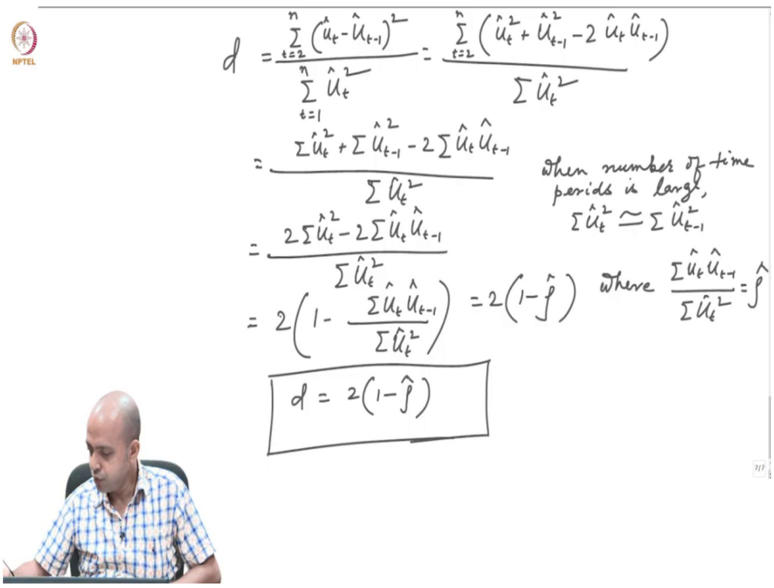
So, this is the notation and assumption what we said and when error term that data generating process shows the error term follows this particular path. The name of this path is called first order Marcov auto-regressive process. Why first order? Because see here we are assuming that there is only one period lag in the equation of error term, ut is correlated only ut minus 1, whether ut is correlated with ut minus 2 or ut minus 3, that actually we do not know. So, assume since only first order Marcov auto-regressive process.

In short, this process is also knows as AR 1 process. So, if we write the equation as ut equals to let us say rho 1 ut minus 1 plus rho 2 ut minus 2 plus epsilon t, then this is called AR 2 process because in the equation I have included the second period lag of the error term also. Since ut is regressed on its own previous value that is why the name called auto regression or auto-regressive process and Marcov was the econometrician who first introduced the term that is why the name called Marcov first order auto-regressive process.

So that, we must remember before we basically proceed about the auto-regressive process. So, this is the assumption that Durbin and Watson they basically made. So, that means, Durbin-Watson process can detect only first order autocorrelation. They cannot detect the presence of second and higher order autocorrelation.

The second is the original regression model and includes an intercept. That means, yt would be equal to alpha plus beta xt plus ut. So, the presence of intercept should be there and they also assume that ut actually follows a normal distribution; with 0 mean sigma square variance. It follows a normal distribution. So, these are the main assumption that Durbin and Watson made and based on this assumptions they did a Durbin and Watson test statistic d.

(Refer Slide Time: 17:40)



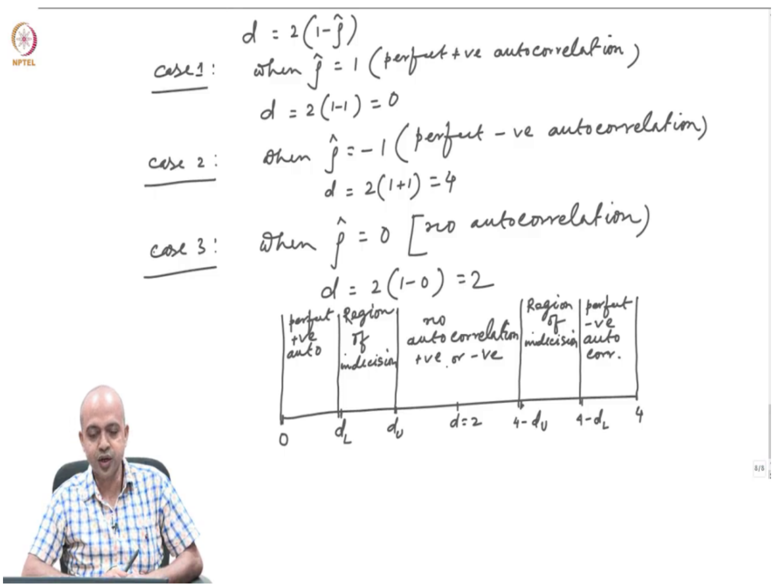
This small d is called the Durbin-Watson test statistic. Like the chi square and F you have learned about several other test statistics. Another test statistic developed by the Durbin and Watson is called DW test statistic which is used for detecting autocorrelation and how they have defined this? They have defined this as summation ut minus ut minus 1 hat square, where t is running from 2 to n divided by summation ut hat square t running from 1 to n.

Now see the numerator. Here the observation is 1 less than the total, why? Because while taking the difference, one observation is lost that is the reason. That is why I am saying t runs from 2 to n while this is 1 to n. Then if you expand this, what you will get? You will get summation t running from 2 to n then ut hat square plus ut minus 1 hat square minus 2 into u t hat into ut minus 1 hat divided by summation ut hat square.

This again I can write if I take everything outside, then what will happen, this would become sum of ut hat square plus sum of ut minus 1 hat square minus 2 into sum of ut hat ut minus 1 hat divided by sum of ut hat square. Now, when t is sufficiently large, when number of time periods is large, then what happens? Sum of ut hat square is almost like sum of ut minus 1 hat square because there is difference only for one observation. So, it is almost equal when your number of time period is sufficiently large.

So, if you substitute ut for ut minus 1, then that would become 2 into ut hat square minus 2 into ut into ut minus 1 hat divided by sum of ut hat square and this would again become 2 into 1 minus sum of ut hat ut minus 1 hat divided by sum of ut hat square equals to 2 into 1 minus rho hat. how we have defined this rho hat? Basically sum of ut hat ut minus 1 hat divided by sum of ut hat square equals to rho hat. So, this rho hat is actually the first order autocorrelation coefficient that is how we have defining it. So, that means, we can say d equals to 2 into 1 minus rho hat. This is an important relationship, d equals to 2 into 1 minus rho hat.

(Refer Slide Time: 22:38)



Now if we use the d equals to 2 into 1 minus rho hat. Then we can derive several cases to get different values for the d, depending on which particular value we are assigning for rho hat. Let us say this is case 1. When rho hat equals to 1; that means perfect positive autocorrelation, then what will happen to d? d would be equal to 2 into 1 plus 1 equals to 4. So, when rho hat equals to 1, d equals to 4.

Case 2: when rho hat equals to minus 1, that means, perfect negative autocorrelation, what will happen to d? d would become 0. So, this is 1 minus 1 so, this would become 0 and then this would become 2 into 1 plus 1 equals to 4 and then case 3 when rho hat equals to 0, so that means, no autocorrelation, then d equals to 2 into 1 minus 0 equals to 2.

So, depending on the values of d, we can easily understand whether there is any presence of autocorrelation or not. So, a value of d in the neighbourhood of t, neighbourhood of 2 will indicate that there is no autocorrelation problem in your dataset and if d tends to either 0 or 4, then that will tell you there is presence of either positive or negative autocorrelation problem. So, a lower value of d will indicate presence of autocorrelation, a value in the neighbourhood of 2 will indicate there is no autocorrelation and a value which tends towards 4 that also indicate that there is negative autocorrelation.

Now in a simple diagram if you like to represent this, let us say this is d equals to 2. So, the values of d will lie between 0 and 4. This Durbin-Watson test statistic d, it does not follow the known type of familiar probability distribution like t, chi square and F, rather Durbin and Watson developed a completely different type of table for this Durbin-Watson test statistic and that only depends on total number of observation and the number of explanatory variables that you are including in your model.

So, depending on that value what you need to do once you estimate a model, you have to see they have given a lower limit of let us say this is dL, lower limit of Durbin d value, then let us say this is du, then let us say this is 4 minus dL and this is 4 minus du. So, when you are having the values like these, in this region when your values are in the neighbourhood of 2, then what is the case? The case is no autocorrelation. So, here I am saying, in this region there is no autocorrelation-positive or negative.

And then, when the values are in this region that means, towards 4 what it says? There is perfect negative autocorrelation and when the value lies below dL that basically says that means, when it is tending towards 0, then it says perfect positive autocorrelation. So, that means, if you write your H naught, the null hypothesis is basically the presence of autocorrelation. So, sorry here it is perfect positive autocorrelation and then here it is perfect negative autocorrelation and here it is no autocorrelation.

So, that means, there are two regions, what will happen if you are here and here? These two regions are called the region of indecisiveness so that means, you cannot take any decision that is the problem of this Durbin-Watson test procedure. There are two regions when the value of d lies between dL and du or between 4 minus du and 4 minus dL, then actually these are called region of indecision. I am not able to take any decision.

But if you are here less than dL or in all other region let us say in between 4 minus du and 4 minus dL, then you can easily take your decision whether there is positive or negative autocorrelation. So, this dL actually you can think of a tolerance in terms of tolerance, how much you can tolerate for autocorrelation? So, this is the lower limit of the tolerance and if it goes beyond that that means, it shows there is some kind of positive autocorrelation.

Similarly, if you are here then that also shows there is perfect positive autocorrelation, perfect negative autocorrelation and when your values are in the neighbourhood of 2, then that basically shows there is no presence of autocorrelation either positive or negative. So, this is how with the help of the Durbin-Watson test statistic you can actually decide about whether your dataset suffers from autocorrelation or not.

But please keep in mind that Durbin-Watson test procedure can detect only first order autocorrelation, whether there is AR 1 or not. Because they assume the data generating process is Marcov first order auto-regressive process. It cannot determine presence of higher order autocorrelation.

So, we are closing our discussion with this today and tomorrow what we will do, we will take one dataset and then we will see what are the consequences, we will estimate the model using our Euler’s technique and we will see in presence of autocorrelation if we still estimate the model using Euler's technique, what are going to be the consequences and we will also see how to estimate this Durbin-Watson test statistic and decide about presence of autocorrelation or not, that we will discuss in our tomorrow’s class.

We are closing our discussion today upto this much. We have learned what is autocorrelation, what are the consequences, and how to detect autocorrelation. We have defined only the popular measure, Durbin-Watson test statistic and we will discuss another measure to detect higher order autocorrelation in our next classes.

Thank you.