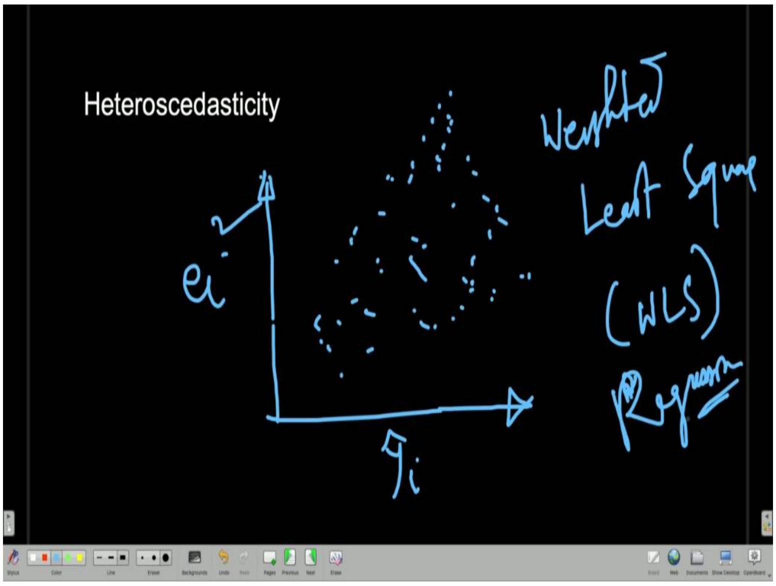
Applied Econometrics
Prof. Tutan Ahmed
Vinod Gupta School of Management
Indian Institute of Technology - Kharagpur

Module - 8
Lecture - 66
Heteroscedasticity (Contd.)

Hello and welcome back to the lecture on Applied Econometrics. We have been talking about heteroscedasticity. And in the previous few lectures, we have seen the different tests for understanding if our regression model has some heteroskedasticity problem present. Now, it is one thing to know if there is a problem, but it is another thing to actually bring some remedy to the problem. Now, looking at the nature of the heteroscedasticity or how the error term is distributed, if we can try to conceptualise it like this;

(Refer Slide Time: 01:02)



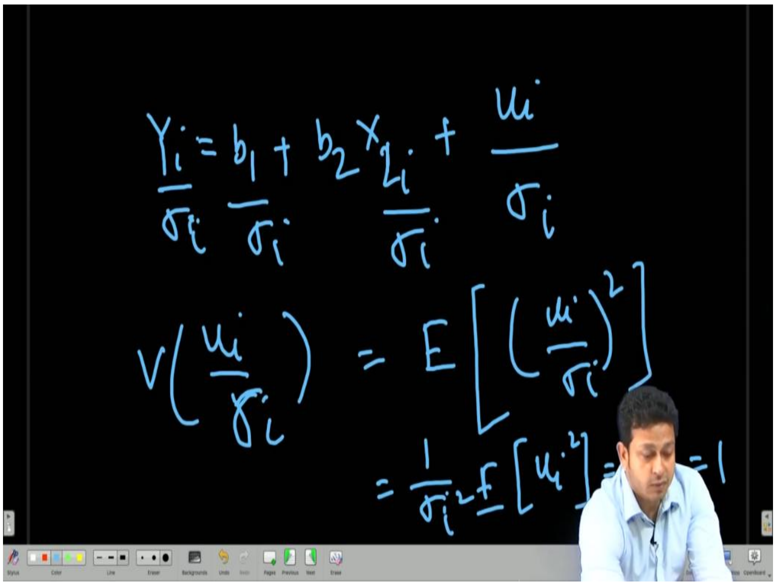
Let us say, if we look at the distribution of the error term, it looks like this usually. There could be different other patterns of the distribution, but usually it is like this, when I have my e i values here, and let us say, here I have y i predicted. Now, one thing I can do is that, I can actually see that for some of the observations, the variance of the error term; let us say, I can also write the variance actually; it is better to write variance.

So, the variance of error term is actually low or less. And in those cases, what I can do is, those observations I would actually prefer. And what I can do is, I can actually assign more weight in my regression equation to those observations where the variance is actually less.

And I assigned low weight to the variances which are actually high. And this kind of regression is actually called weighted least square regression or in short, WLS regression.

Now, it is difficult to actually understand which of the observations for which you have this variance of the error term low, and to assign corresponding weights. So, in order to actually achieve that, we adopt different means to reduce the variance. So, to understand how we can reduce the variance of the error term, we will just try to understand conceptually how that is a possibility. And let us just do some little bit of maths for that.

(Refer Slide Time: 02:46)



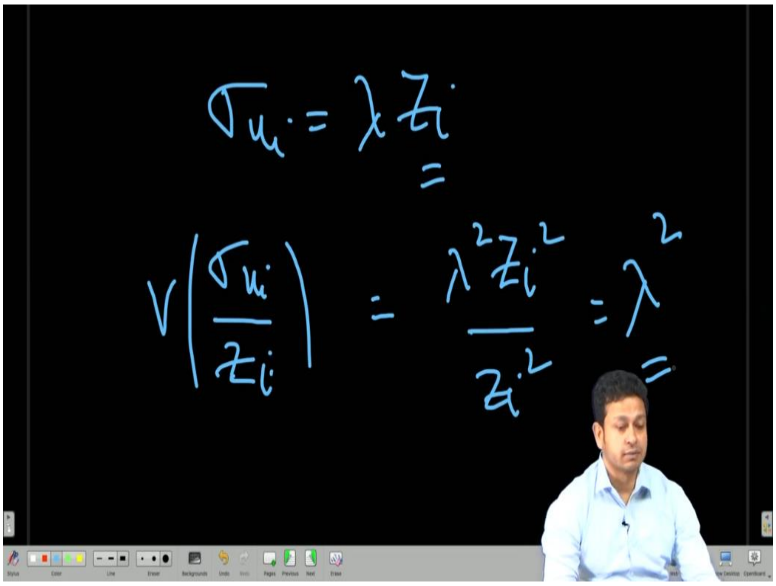
So, let us say I have my regression equation which is Y i is equal to beta 1 plus beta 2 X 2i; and I can have u i. Now, I know that the variance of the error term is sigma i square. And I want to get rid of the variance, because I want to get rid of the variance of the error term. So, one way to get rid of the variance is simply dividing all the observations with the variance or the standard deviation.

So, if the standard deviation corresponding to the error term is sigma i, and if I divide all the observations with this sigma i values, so, I can actually get rid of the variance of the error term. And how is that possible? So, let us say, if I want to understand the variance of this new term, which is u i by sigma i; my stylus is creating trouble; and this is going to be expectation of u i by sigma i square.

And we do not have the other term, because that is basically 0. And it is going to be 1 by sigma i square; and I have expectation of u i square, which is nothing but sigma i square. So, sigma i square, sigma i square cancels out. And all I have is the value 1. So, it is basically, I standardise it and I have now a constant variance for the error term, which is equal to 1. And that is something I really covered, because I want this variance to be constant.

Now, this is one way of dealing with this. And we can write the modified regression equation by dividing this whole equation with the sigma i. Now, this is all right, but it is very difficult to actually understand the variance of the error term. So, you really do not get the variance of the error term. So, since you do not get the variance of the error term, what we do is, sometimes we use a proxy.

(Refer Slide Time: 05:21)



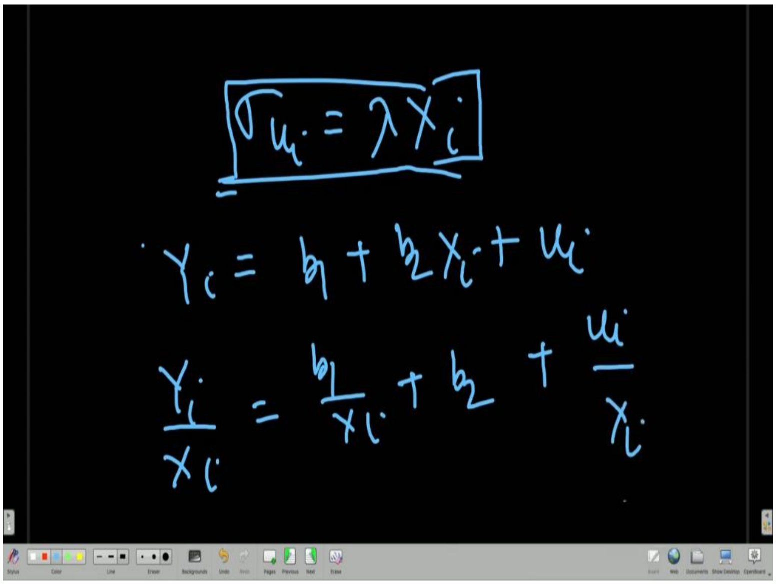
Let us say, this variance of the error term, standard deviation of the error term is actually varying with some variable Z i, which is observable. So, in our case, let us say our suspect variable. So, when we say that, okay, the error term is actually varying with some X variable; so, that is my suspect variable. So, I can say that the variance of the error term is varying with the X variable.

So, I know the X of the Z variable here, because that is given to me; but I am not given the value of sigma u i. So, I can always assume that I can replace sigma u i with Z i, with some constant term lambda. And the beauty of it; we are just going to see it. So, if we take the variance of; let us say, if I divide sigma u i, Z i; and I really do not need to know the value of sigma u i here.

And if I take the variance of this, what I will get is essentially, in the similar manner, I will get lambda square Z i square by Z i square; and which will leave me a value of lambda square, which is again a constant term. So, the moment I have a constant term, our purpose is served; we basically wanted it to have a constant variance. And even here, we do not need to have the value of lambda, because, all we need is a constant variance.

So, as long as that is satisfied, I am fine. So, that is the measure we can possibly adopt. So, using this, we can actually, just instead of Z i, we can use a suspect variable which is our X i.

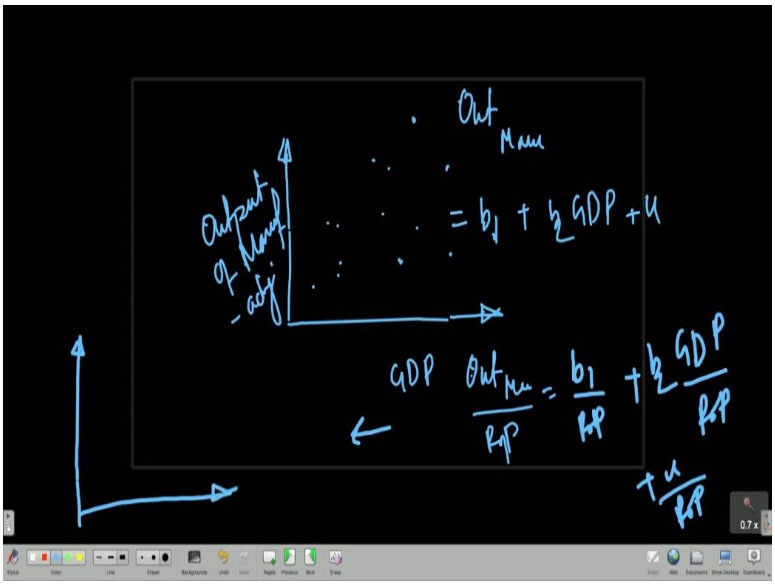
(Refer Slide Time: 07:13)



So, instead of Z i, we can also use our X i, which is our suspect variable. For example, in our previous example where we have plotted the error term with my suspect variable general education, and I saw that my general education is actually influencing my error term. So, what I can do is, I can divide the whole equation with X i. So, what I can do is basically divide the whole; so, since I do not have this, what I can do, Y i let us say, beta 1 plus beta 2 X i plus u i.

So, what I can do, because of this assumption here, I can actually divide Y i or the whole equation with X i. So, that would serve the purpose. Basically, we are following the same logic here, beta 1 by X i; here I will have beta 2; and then I will have u i by X i. So, basically, dividing by a variable which can reduce the variance of the error term, we can also divide by some proxy of the X i. So, I will give you an example of this.

(Refer Slide Time: 08:40)



So, let us say we have an assumption that the countries which are economically well-off, the share of GDP in the manufacturing sector or the output in manufacturing sector is actually proportionate to the GDP. So, we essentially assume that a country to have a good GDP would lead to a high output for the manufacturing sector. And let us say we actually plot all these observations here. And we will see something like this. And here it is GDP.

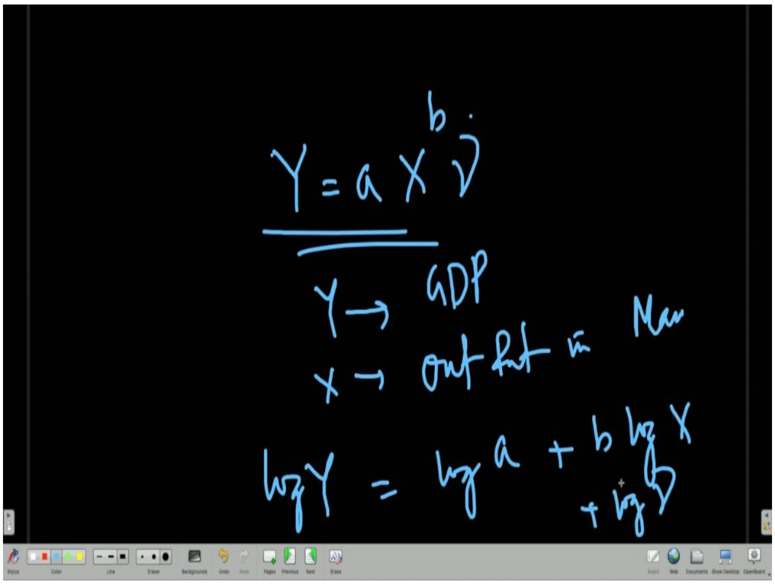
Now, the problem here is, some countries, we have smaller GDP; and for them, like if the output of manufacturing is actually influenced by GDP or not, I am really not sure, because error term is, it becomes so divergent, it is very difficult to actually understand the relationship between output of manufacturing and GDP. And when my regression equation, here is going to be, let us say output manufacturing is equal to beta 1 plus beta 2 GDP plus some error term.

Now, the same problem, again the heteroscedasticity is present and I have to have some variable to actually divide both the Y and X's. Now, I can do one thing, I can actually divide by GDP for both Y and X variables. Also I can use some other proxy that is actually indicating GDP, and which could be for example, the population. So, I can basically divide the Y and X's with population.

And here, what I can do is, output manufacturing by population; let me reduce the screen size; is equal to beta 1 by population beta 2 into GDP by population plus u by population. So, this is one way of actually addressing the heteroscedasticity problem. And once we actually do that, we will see that the heteroscedasticity is actually gone. And we will see something, the result of it is going to be something where we can have a slightly better distribution of the error terms.

Or if you want to again plot the GDP per capita and output manufacturing, you will see that the error term is actually less heteroskedastic. So, that is what we want to achieve. And we are not going to do this hands-on here. We are not going to solve it here, but we are going to give a hands-on for this. So, that is one way of actually solving it. There could be other ways. For example, you have seen previously that you can have this functional misspecification.

(Refer Slide Time: 12:05)

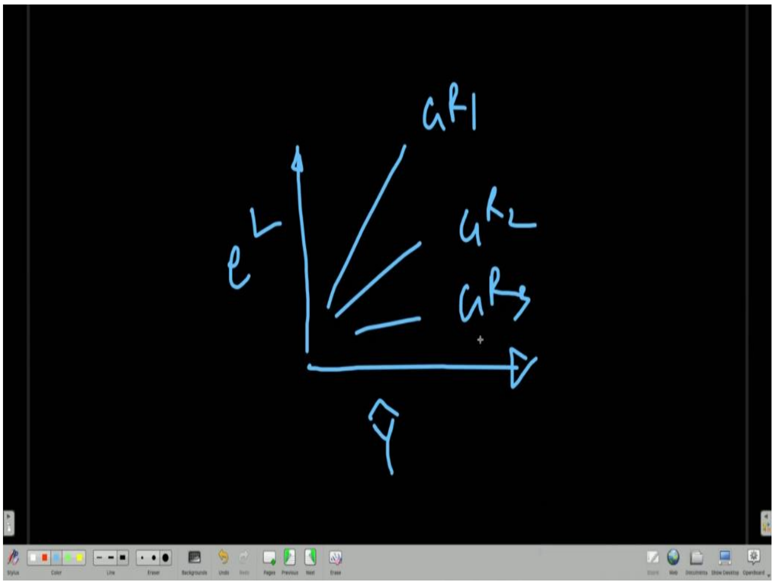


So, the functional misspecification could be like, I can have something like more of a logarithmic relationship. I can have, let us say Y may output; my GDP is equal to a X beta with some error term, where my Y is my GDP and X is the output in manufacturing. Now, in the previous case, we are thinking that Y and the output manufacturing are actually related linearly, but this is the actual functional form.

So, if this the actual functional form, so, what we have to do is, instead of running a linear regression, we have to use this logarithmic; you have to take logarithm on both the sides. So, it is going to be essentially a log log sort of regression. And then, if you perform the regression; so, we have seen, for model misspecification, the functional misspecification, we might have the problem of heteroscedasticity; but if I do something like this, that problem may go.

And if the problem goes away, then we can also get rid of the heteroscedasticity problem. There could be other ways like we can have different subgroups, and we can use different subgroups, and we can run different regressions for different subgroups. We have seen that. For different subgroups, we can have different levels of heteroscedasticity.

(Refer Slide Time: 13:42)



For example, we have seen something like this, error term square; this is our Y hat; and let us say this is group 1, group 2, group 3. So, we can also actually run the regression separately for these different groups to actually get rid of the heteroscedasticity problem. We can identify; but if there is any one group where the heteroscedasticity is high, or if we actually run the regression separately, we can actually see that completely the heteroscedasticity problem is gone.

So, we will do some hands-on on this kind of remedial measures. And with this, we will come to a conclusion about the lecture on heteroscedasticity. And that is one of the important Gauss-Markov assumptions which you need to fulfil when we want to have the best linear unbiased estimator. So, with this, we will end the lecture here.