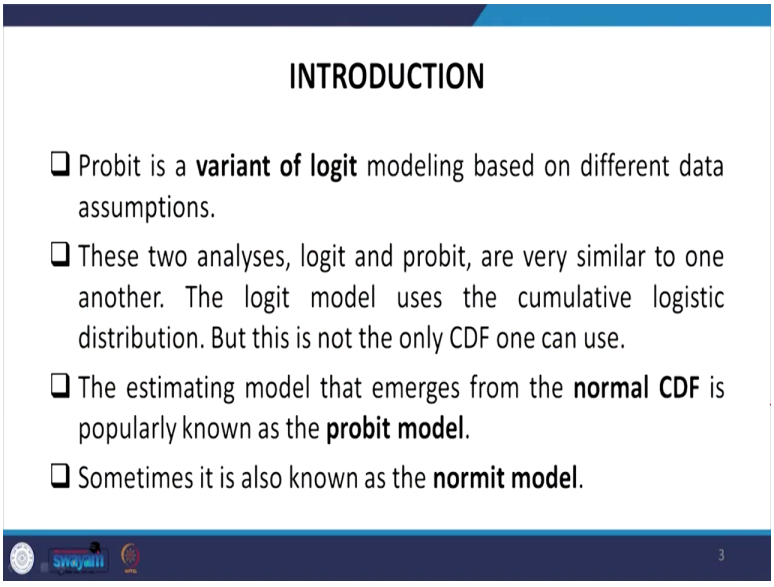
Handling Large Scale Unit Level Data Using STATA
Professor Pratap C. Mohanty
Department of Humanities and Social Sciences
Indian Institute of Technology, Roorkee
Lecture 35
Binary Response Models - IV

Welcome once again to this module of Handling Large Scale Data Using STATA. So, far we have been trying to understand ways to go into the deeper understanding of quantitative analysis of qualitative data. So, in the last 3 lectures we have been explaining the qualitative variables, qualitative dependent variables. And LPM and Logit has already been discussed but now onwards we are going to discuss, in this particular lecture we are going to discuss Probit and Tobit model.

Probit where we wanted to understand whether the error distribution or the distribution function follows normality or not. In Tobit somewhere we are going to discuss whether there is censoring of that particular data. and we will also compare all those 3 models together; whether LPM, Logit and Probit, which one is better, why it is better?

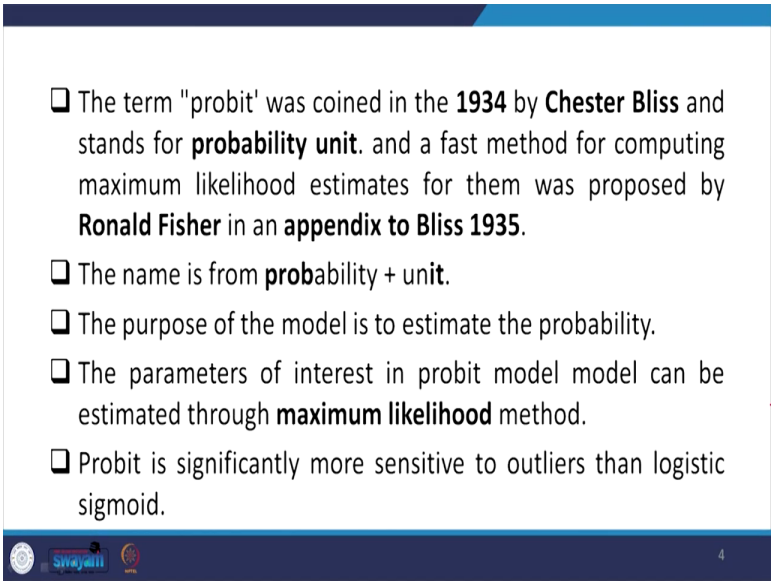
(Refer Slide Time: 01:34)

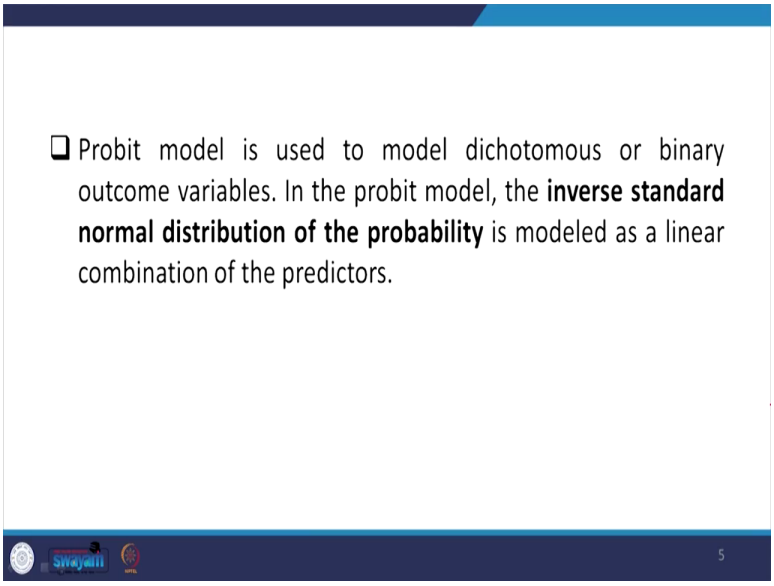


So, coming to the understanding of Probit regression so I will take a quick wrap up of the understanding since many aspects of Logit and distribution function has already been explained. So, Probit is a variant of Logit, modeling based on different data assumptions. These two analyses that is Logit and Probit are very similar to one another but the Logit uses a cumulative

distribution function whereas the Probit uses normal density function. So, coming to the estimates of the Probit, since it follows a normal density function we wanted to mention that Probit model is also called as a normit model.

(Refer Slide Time: 2:34)



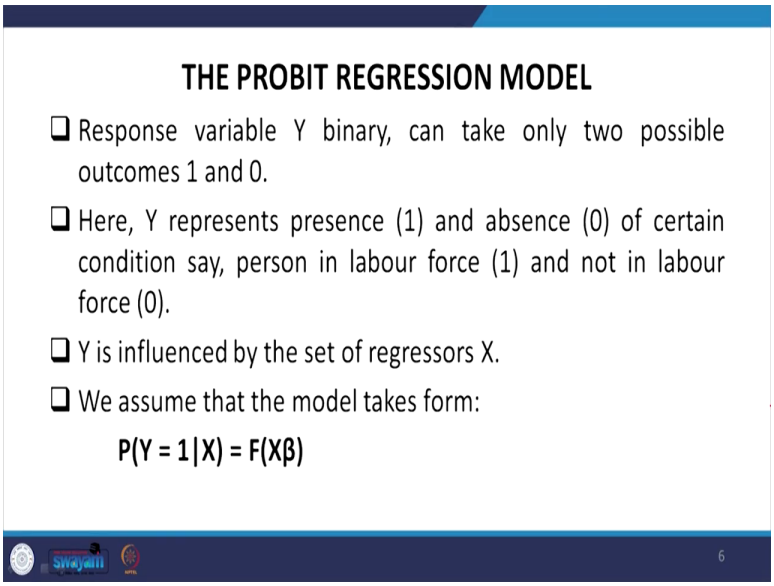


This was initially coined by Chester Bliss in 1934. Then the Probit stands for probability unit. And basically the first maximum likelihood estimates were initially proposed by Ronald Fisher in an appendix to Bliss of 1935. So, the name is from probability plus unit. That is why it is called Probit, in short it is called probability unit.

The purpose of this model is to estimate the probability value. The parameter of interest in Probit model can be estimated through maximum likelihood method also. Probit is significantly more sensitive to outliers than that of Logit because distribution of the probability change in the outliers is very less in case of logistic distribution whereas the probability change in case of the outlier in case of Probit distribution very sharp.

The Probit model is used to model dichotomous and binary outcome variables. In the Probit model the inverse standard normal distribution of the probability's model is a linear combination of the predictors. So, inverse of the standard normal can be modeled in terms of linear combination.

(Refer Slide Time: 04:15)



- Where,

P denotes probability,

F denotes cumulative distribution function (CDF) of the standard normal distribution.

Parameters β's are typically estimated by maximum likelihood method.

- It is also possible to motivate the probit model as a **latent variable model**.
- Suppose there exist a latent variable (auxiliary random variable).

7

- $Y^* = X\beta + \varepsilon$

Where Y^* is unobserved and $\varepsilon \sim N(0,\sigma^2)$

- The Y can be viewed as an indicator for whether this latent variable is positive:

$Y = 1$, If $Y^* > 0 = X\beta + \varepsilon > 0$ i.e. $\varepsilon > -X\beta$

0 otherwise.

$$P(Y^* > 0 \mid X) = P(Y = 1 \mid X) = P(\varepsilon > -X\beta)$$

$$= P\left(\frac{\varepsilon}{\sigma} > \frac{-X\beta}{\sigma}\right)$$

8

Probit regression model especially for the response variable Y binary can take only 2 possible outcomes that is 1 and 0. Here Y represents 1 as presence and absence with 0 of certain conditions such as person in labor force to be coded as 1 and not in the labor force to be 0. And Y is influenced by set of regressors that is X.

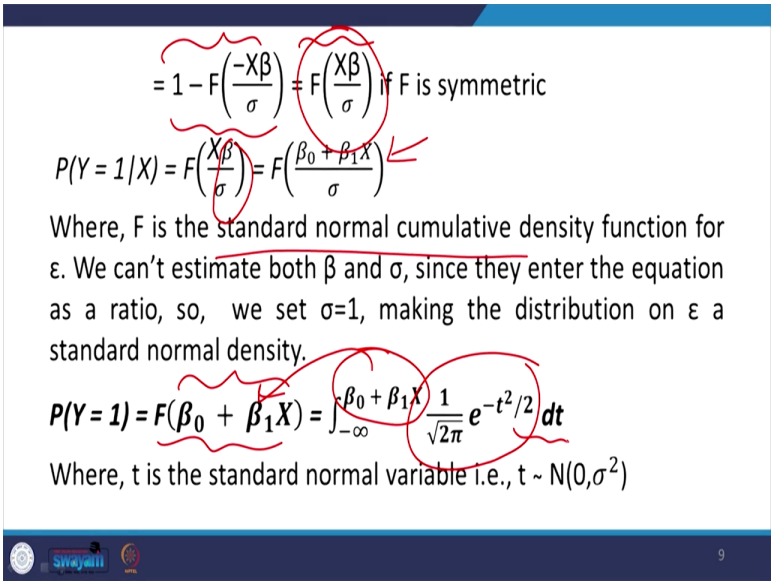
We assume now the model takes the form like Y with the probability of success given X is nothing but the probability density function or the probability function, cumulative function of X beta where P denotes probability and F denotes the cumulative distribution function of the standard normal distribution. Parameter betas are typically estimated by maximum likelihood

method. Here we also wanted to mention that it is possible to motivate the Probit model as a latent variable model since probabilities of the variables are estimated the probabilities and its expected values could be defined as not the exact values rather representative values called latent variable model.

Suppose there exists a latent variable or auxiliary random variable then Probit can be also most fitted. Then the model looks like this. Y star is the unobserved variables with error distribution of epsilon with 0 and Sigma square as standard variance. Then Y_i's viewed as an indicator of whether this is a latent variable is positive if Y is equal to 1. And if Y star is greater than 0, that is nothing but this is greater than 0, if and only if that is the case then the error will be greater than that of the negative value of the estimated values. if it is 0 then in that case the probability of error values could be estimated till that of the XB limit.

So, basically in the density function or the probability density function, if you can estimate till that particular time is more important. Based on this equation if you simply divide the standard deviation in both the side the probability limit is defined here. If any specific value till that value if you are calculating and their probabilities you are estimating the probability limit of this error is estimated through this. We are going to mention.

(Refer Slide Time: 07:43)



So, in the total density function, cumulative density function since the entire probability is 1, so 1 minus till is nothing but the cumulative probabilities of X beta upon sigma. And if these two are equal that means the distribution is perfectly symmetric. If the both sides are equal, this side as well as this side, basically probability of success in one side as compared to another side if they are equal that means the distribution is bell-shaped and symmetric.

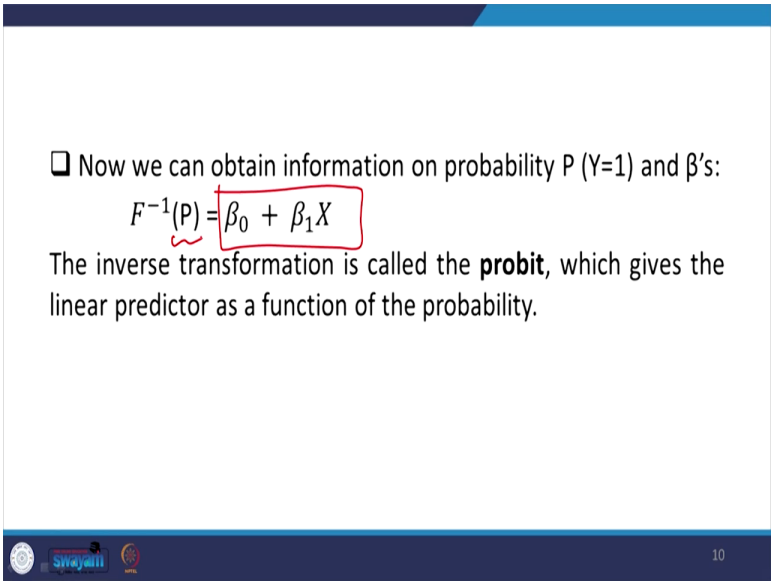
So, probability of success given the explanatory variables is nothing but the density function or the cumulative distribution function. Denominator with F nothing but X beta upon sigma. So, that is basically beta not plus beta 1 X, we are offering here, divided by sigma. So, this is nothing but, F is nothing standard normal cumulative density function for the level of error values.

We cannot estimate both beta and sigma since they enter the equation as a ratio. So, here they are coming as a ratio. It is not individually separated, alright. So, we set sigma is equal to 1 making the distribution on epsilon or error a standard normal density function. When this is 1 so basically we are referring to a standard normal density function.

If the probability value with the success of occurrence in the dependent variable is 1, probability of success is 1 that means we end up with this estimation. So, we are basically estimating till that level this estimated value with the Z value of minus infinity to infinity, probability values from 0 to 1.

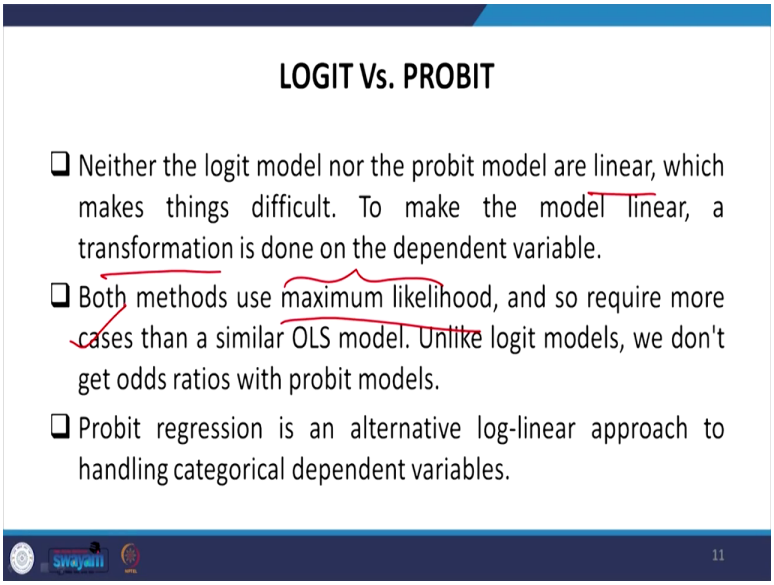
So, we are estimating on the diagram till this level which is equivalent to our estimated value in the model. Since it follows a normal density function we are referring to a normal density function equation. This is a normal density function and since integration we are using so the delta change in that distribution is also attached. So, 1 upon square root of 2 Pi e to the power minus, usually it is to be Z Square upon half of, Z Square upon 2 but here T distribution is considered because of the limiting in the sample. So, the density function equation is clearly given for the estimation, where t is the standard normal variable with its distribution follows 0 and sigma square.

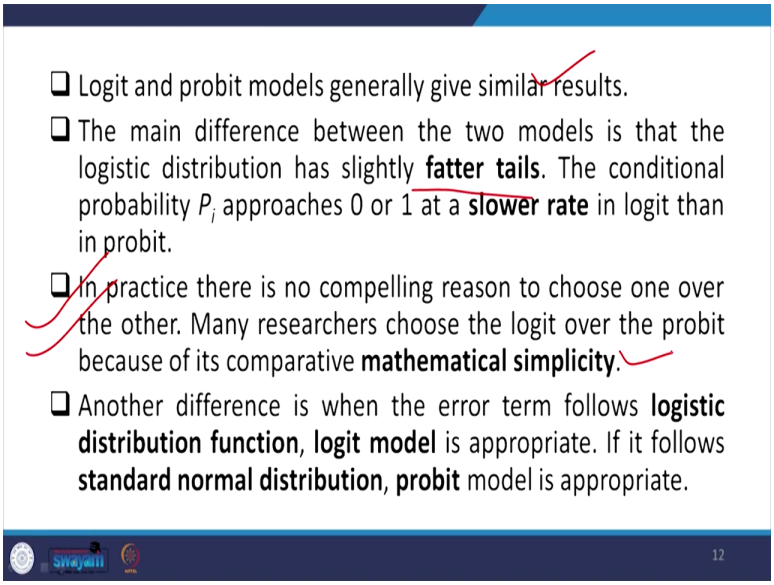
(Refer Slide Time: 10:51)



We can obtain information on probability that is of occurrence 1. That is the inverse of the probability is nothing but the linear function of it. That can be estimated. The inverse of the probabilities will give a linear equation and the interpretation whatever the probabilities values you get, if we take inverse value of it is basically called beta 1 or beta not. The inverse transformation is called the Probit value. Basically the P value is nothing but the inverse transformation which gives the linear predictor of a function of the probability.

(Refer Slide Time: 11:35)





Let us compare logit, probit and LPM together, let us first compare logit model or probit model both are not linear. So, which makes things difficult to make the model linear we make some form of transformation. So, log transformation or inverse transformation in either of the case that is on the dependent variable is going to make it linear.

So, both methods use use maximum likelihood estimator this is one, so linearity aspect I discussed, transformation is important, so we require more cases then the similar OLS model. In order to make likelihood estimator, it usually go by iteration, iteration require more number of cases. OLS less number of cases can derive a result a better result but both the models require, more number of cases that is another important aspect to be mentioned.

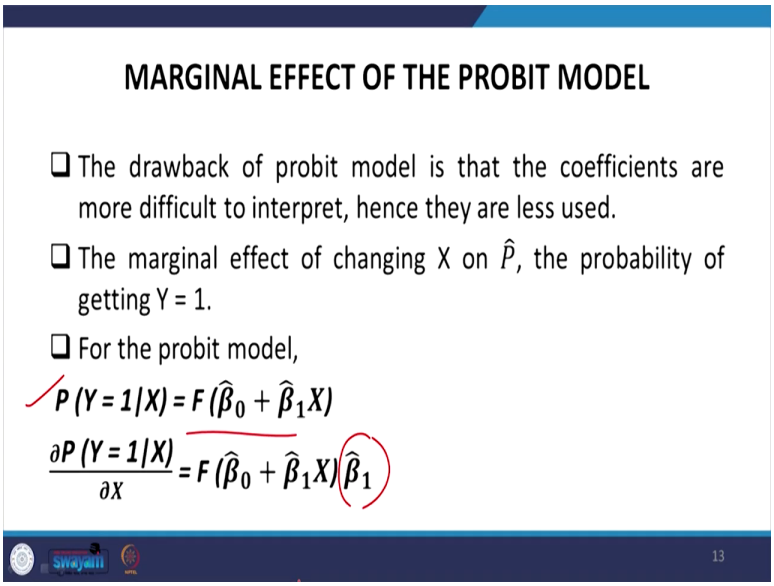
Logit model do not mention online logit models we do not get odds ratio with probit models. So, probit never gives odds ratio it gives the probability value. Probit regression is an alternative log linear approach to handling categorical dependent variables. Question always come which one to be taken, we have it clearly, that there is no such major difference except the distribution function. Logit and probit models gives similar result.

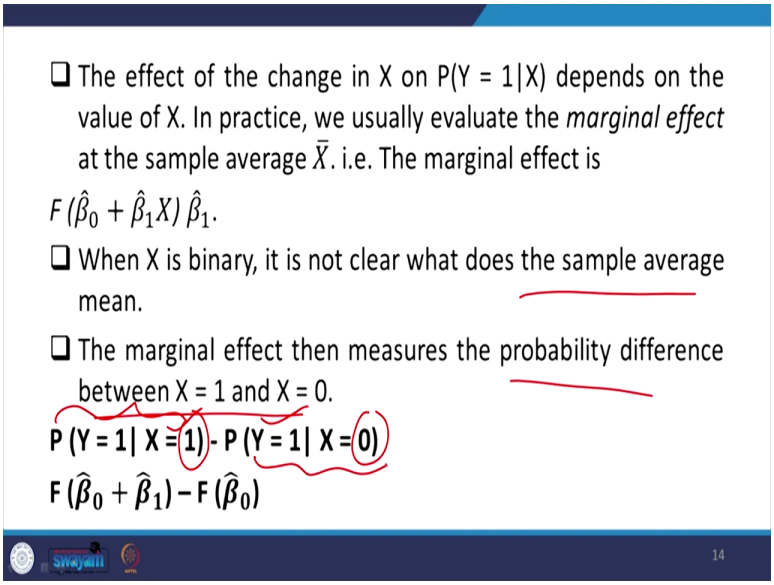
The main difference between the two models is that the logistic distribution has slightly fatter tails. So, Logit has fatter tails. The tails are more fatter whereas in case of Probit the tails are very sharp, and when it approaches to 0 or 1. The end is at a very slower rate. The conditional probability P approaches 0 or 1 at a very slower rate in Logit than that of Probit. Since the tales are flatter approaching very fatter at the end the slope is very very less that is the reason why the

change the end nearby 0 and 1 in case of Logit it is very slow whereas in case of Probit there is a sharp change.

In practice there is no completing reason to choose over the other. That is important. And do not get confused. In practice there is no such hard and fast rule to go by any single model many researchers choose the Logit over Probit because of its comparative mathematical simplicity. Logit is preferred. Another difference is when the error term follows logistic distribution, Logit model is appropriate whereas if the distribution is of standard normal function Probit is most appropriate.

(Refer Slide Time: 14:44)



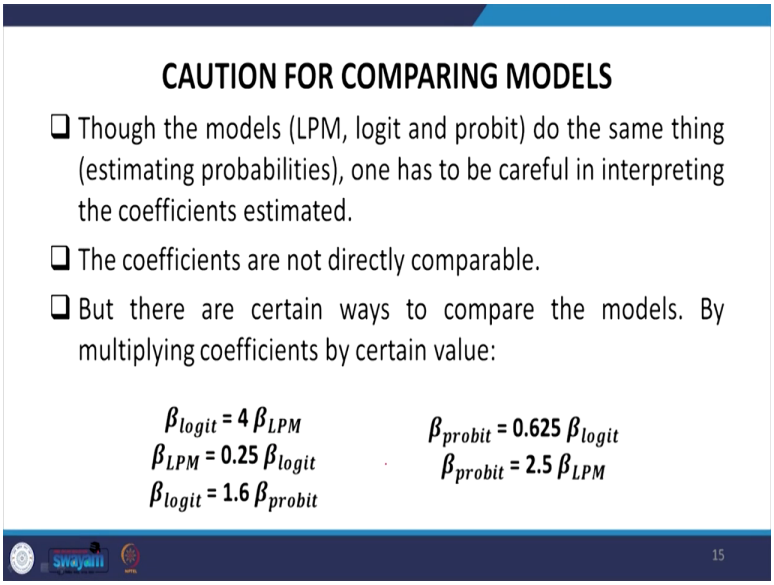


Marginal effect of the Probit model and Logit we have already discussed by taking first derivative. Similarly in this case also P value is expressed with this. We are taking dP/dX. beta 1 will separate due to dX and F prime or F, you can write down a F prime or F, it is up to you, how you are writing to indicate the first derivative.

The effect of change in X on P depends on the value of X. In practice we usually evaluate the marginal effect at the sample average of X bar that is marginal effect conditioned upon the sample mean of each of the variable. When X is binary it is not clear what does the sample average mean? When the X is binary it is very difficult to interpret.

So, the marginal effect then measures the probability difference between 1 and 0. So, that is basically probability of 1 given X equal to 1 minus probability of 1 when it is 0. So, 0 and 1, if you compare the probability difference that is basically called the marginal effect. The marginal value we are going to get in this case is nothing but, if it is categorical it is not the average that matters rather the probability of occurrence in both the case. when you subtract it is basically marginal effect.

(Refer Slide Time: 16:18)

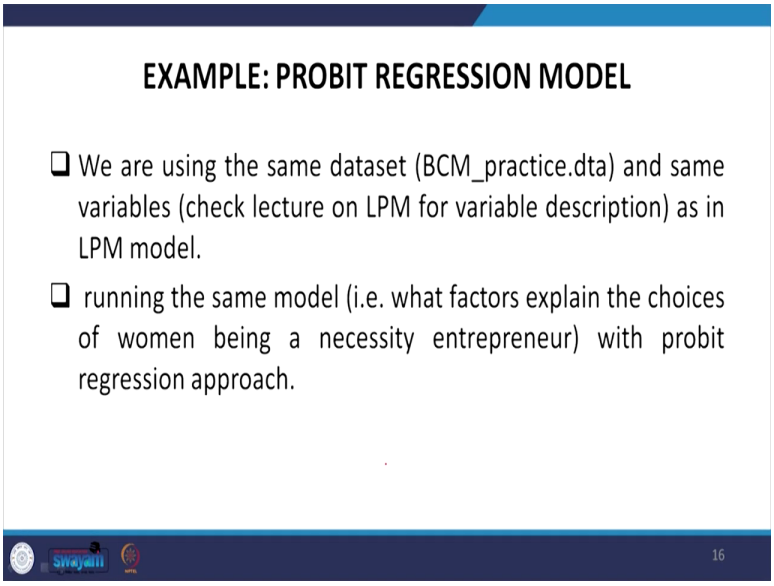


Some caution we need to follow while comparing models specially Logit or Probit or even LPM. Though the Logit model, though the model like these three do the same thing, estimating probabilities, one has to be very careful in interpreting the coefficients estimated. The coefficients are not directly comparable. So, there is no standard technique followed just to compare, But some rough approximations values are there to compare.

But there are certain ways to compare the models as I just said by multiplying coefficients by certain value. Like beta value of the Logit is four times of the LPM. Whatever the coefficient LPM gives if you multiply by 4 it gives the Logit value. So, similarly one fourth of Logit is nothing but the LPM.

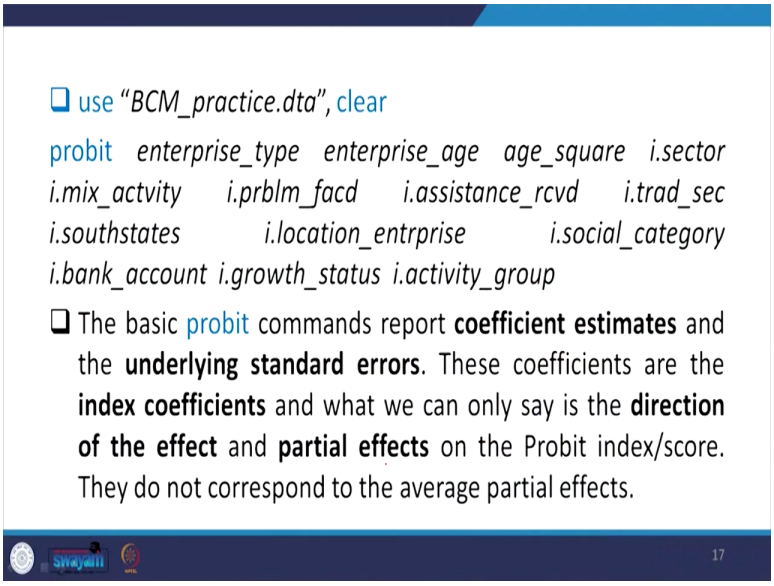
So, regarding Probit and Logit, Logit value is more than that of the Probit value. So, that is 1.6 times of Probit value. Probit gives probability value but in this case it gives likelihood estimator. But in that case, in both the case more or less are the same but difference is usually of 1.6 times. Similarly if you just take, divide 1 upon 1.6 times, it is 0.625. Comparing Probit with LPM Probit value is much higher, that is 2.5 times than that of the LPM. So, these are some comparable estimates made by different experts.

(Refer Slide Time: 18:09)



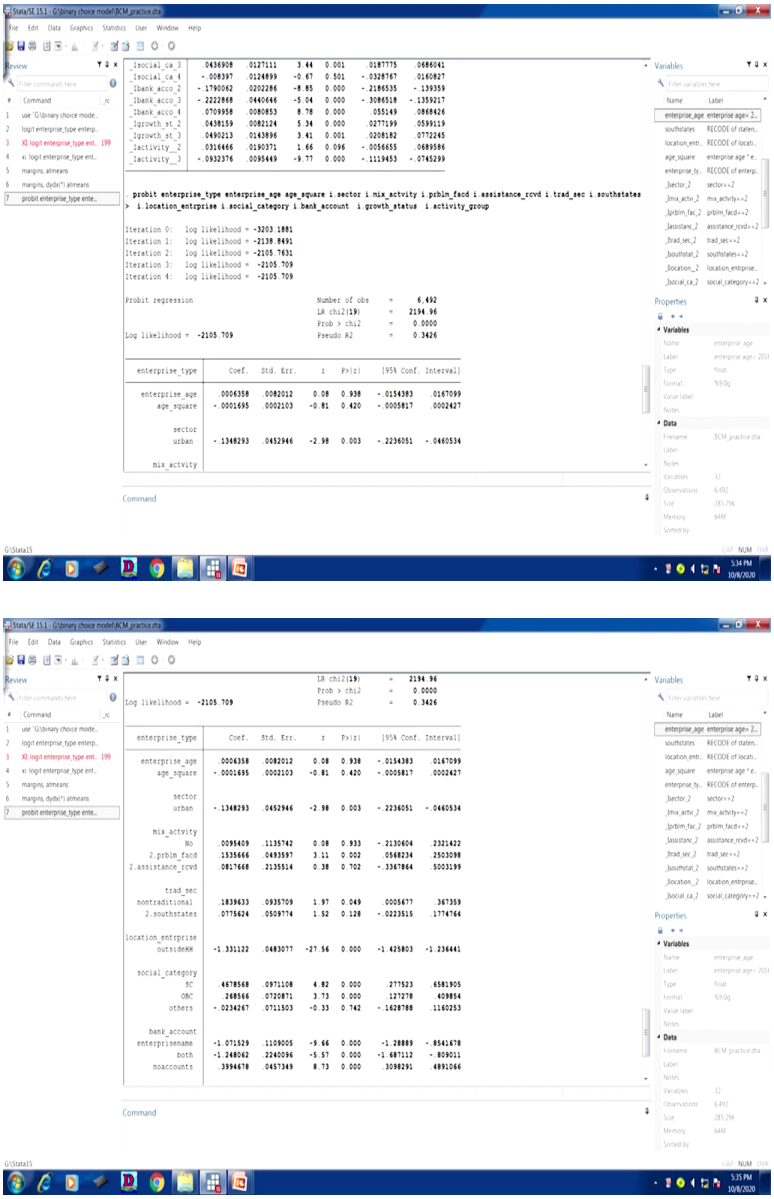
We can also compare our coefficient by taking all the coefficients together. So, we have a practice dataset that is BCM practice we are going to provide you. We are going to check all three models together and running those commands for the same dependent variable that is women being necessity entrepreneur with Probit regression approach as the first. Then we will compare.

(Refer Slide Time: 18:49)



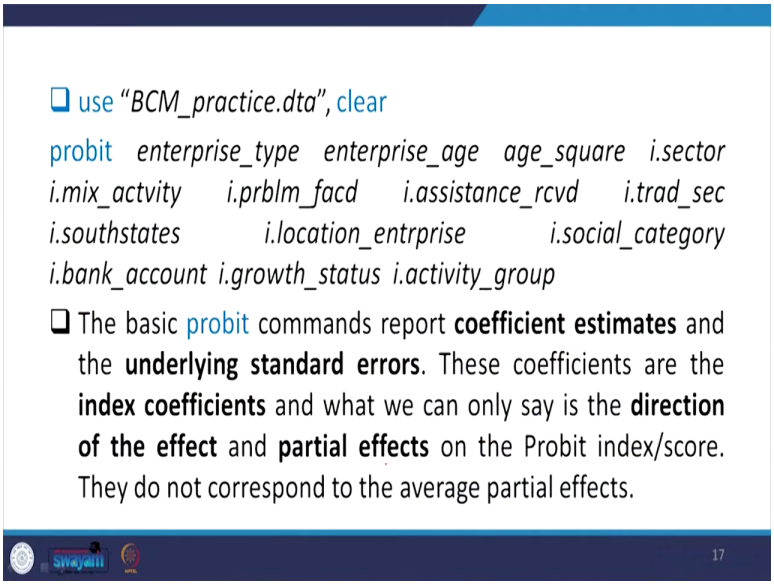
Then we will discuss their coefficient estimation, their standard errors then direction of the coefficient and the partial effects. We will discuss some of those things right now, alright.

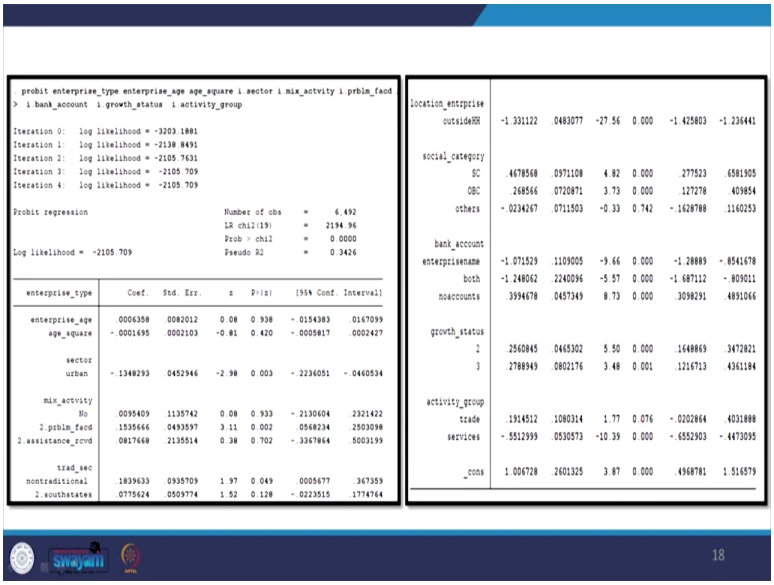
(Refer Slide Time: 19:28)



So, this is where we are going to operate. It is here. So, we will apply the same data. Then we go to the estimation on the same data. We are simply deriving; we have derived the Probit result. So, Probit result is here. first we need to check this value, the Chi square value and its significance level. These are first requirement. Then we come to the interpretation but it only gives the coefficient. But marginal effect is more important.

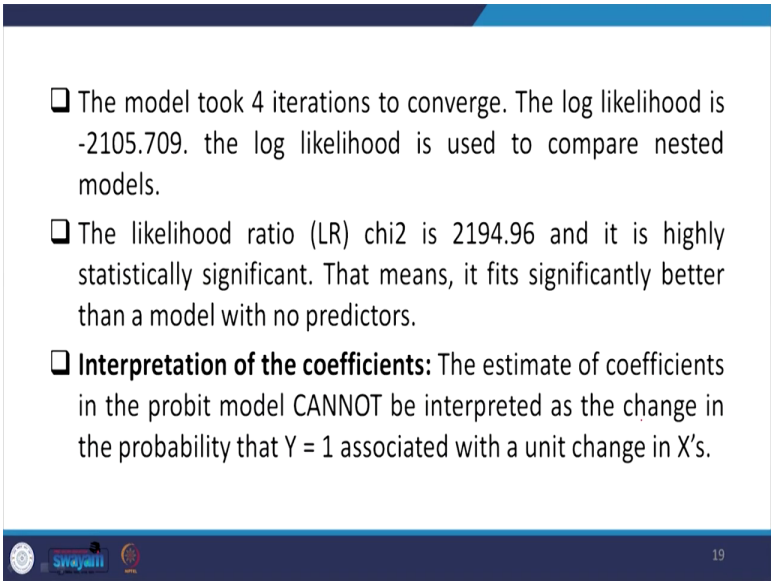
(Refer Slide Time: 19:42)

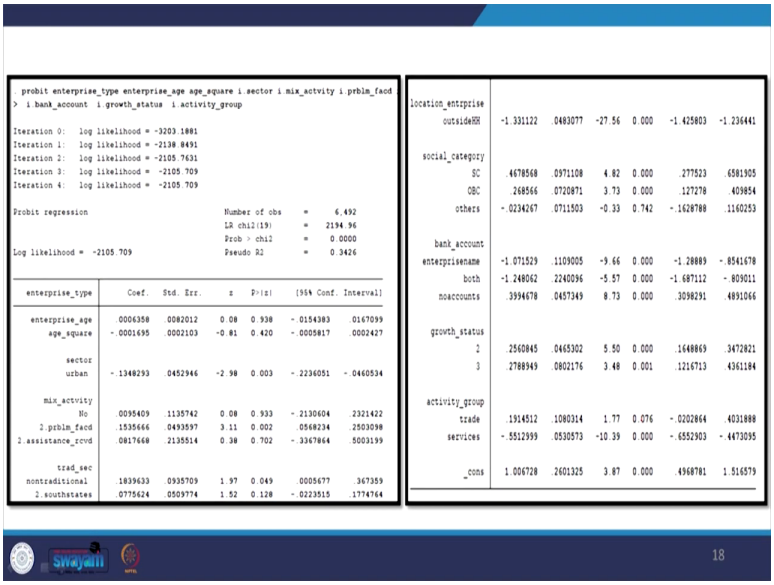




We have to go to the slide. So, coefficients are not important. We have to find out the partial effects. That is more important. This is what the result we derived.

(Refer Slide Time: 19:58)

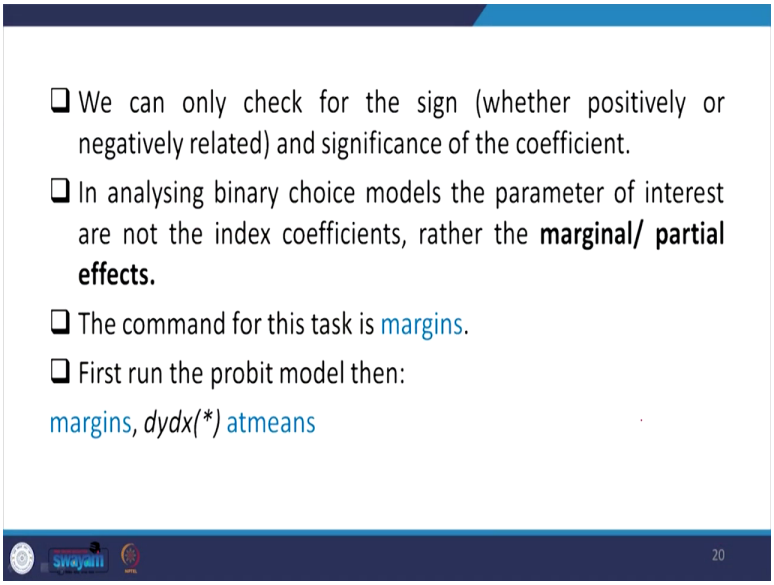


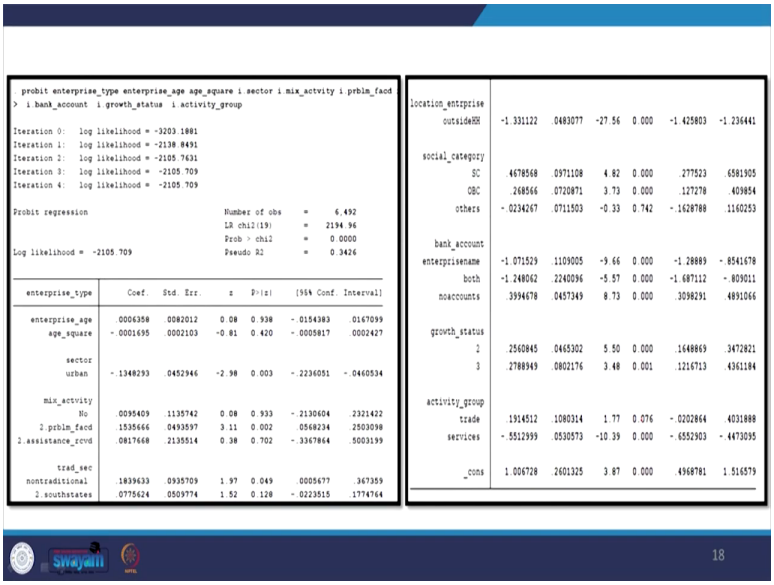


The model took 4 iteration at this moment, 4 iteration to converge. The log likelihood is of minus 2105. It is given here on the log likelihood on the very beginning. The log likelihood is used to compare some nested models if you have. The likelihood ratio that is Chi square test is of 2194 and highly significant that I have already discussed. this fits significantly better then a model with no predictor. This is what it is interpreted.

Let us interpret the coefficient. The estimate of the coefficient of the Probit model cannot be interpreted as the change in the probability that Y equal to 1, associated with the unit change in X. That is not going to be interpreted because of the categorical values.

(Refer Slide Time: 21:01)

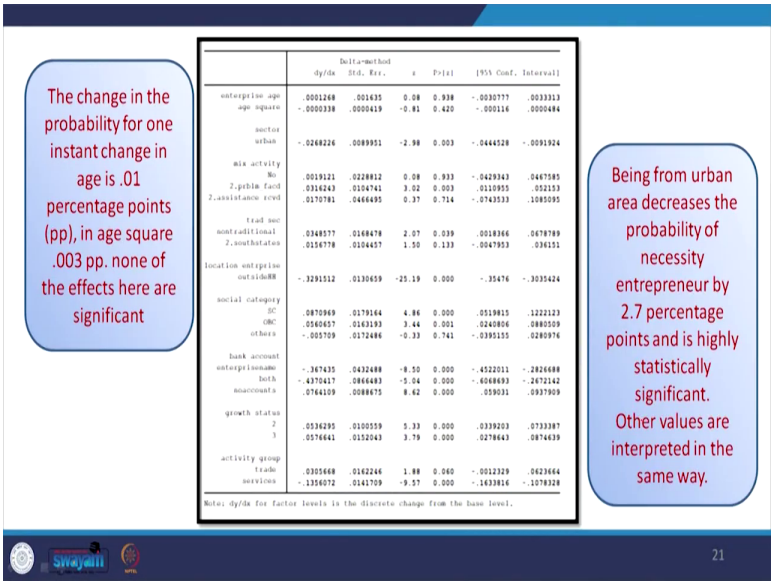




We have to take the marginal values. And we can only check for the sign whether positively linked or negatively linked. I think coefficient has already been mentioned; positive, negative in the coefficient table in the very first column.

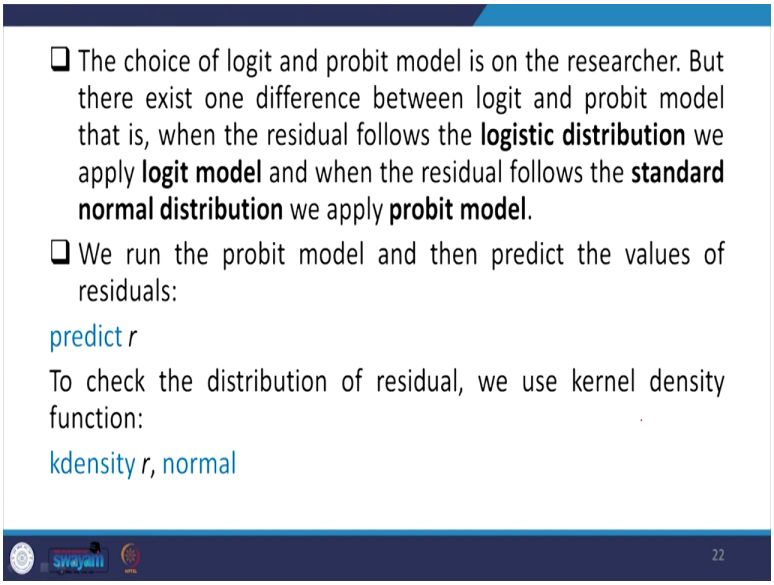
Then coming to analyzing binary choice model, the parameter of interest are not the index coefficient rather the marginal or partial effects. The command for this task we are already dealt in Logit. Here we can also go by margin with the same command dY/dX asterisk atmeans. We will get the result like this.

(Refer Slide Time: 21:43)



The result here is interpreted like the earlier way we interpreted. Here also the same way we will discuss. Like for age this is now 0.1, 0.01 percent points. Age square it is 0.003. Similarly for urban and rural comparison 2.7 times for urban it is 2.7, 0.027 is there. So, in percentage it is 2.7 percentage points lower probability of being necessity entrepreneur. So, you can go through further details and let us go further.

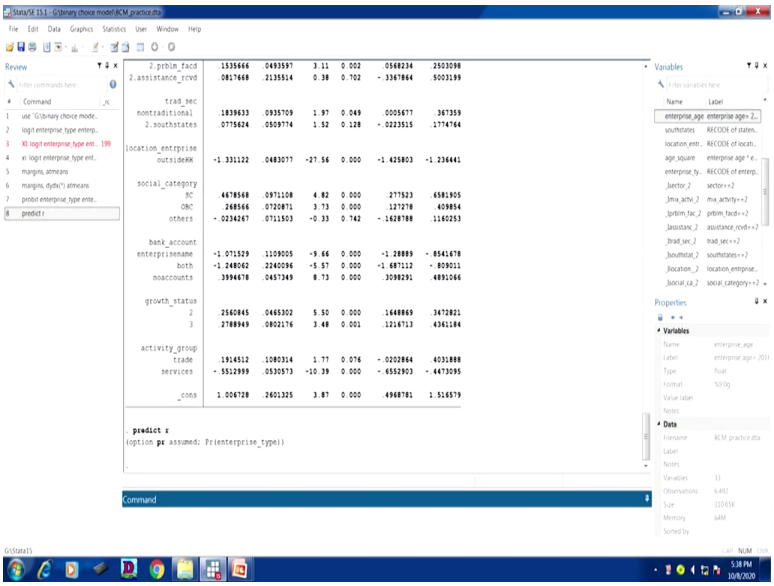
(Refer Slide Time: 22:14)



The choice of Logit and Probit model is on the researcher but there exists one difference between Logit and Probit model, Logit model, when the residual follows the logistic distribution you apply Logit model. When the residual follows a standard normal distribution we should apply Probit model.

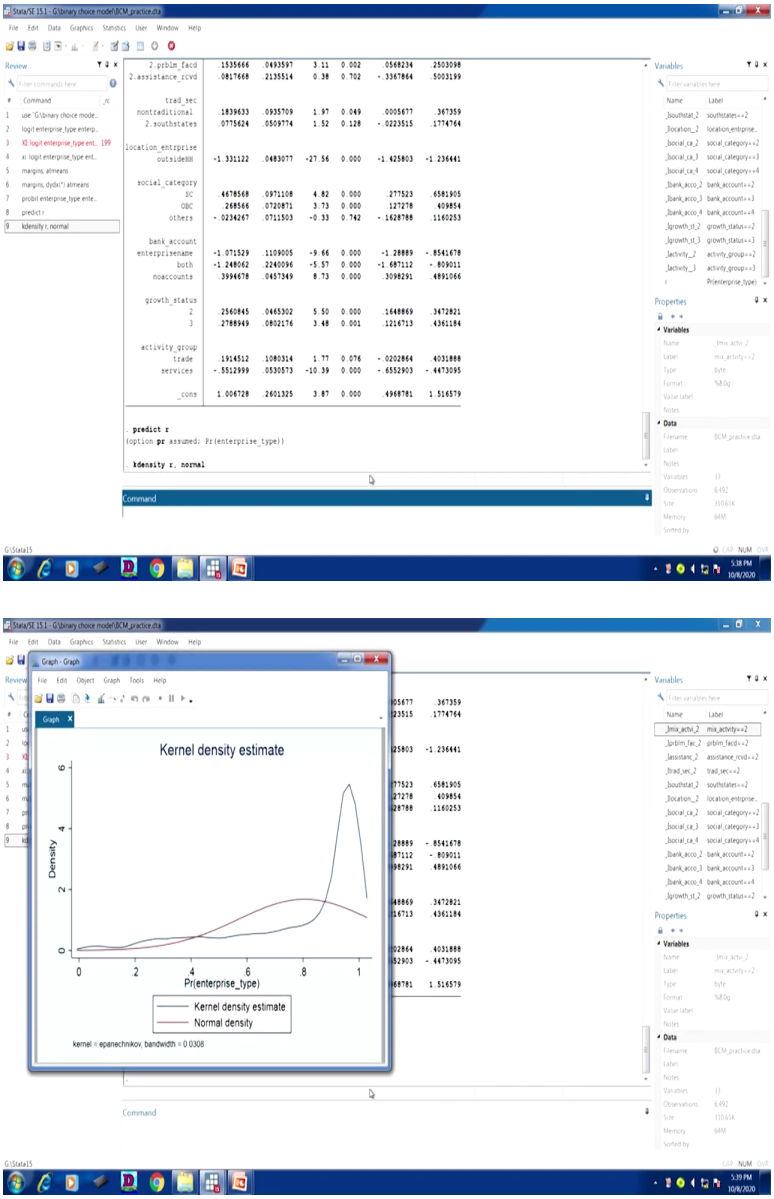
We run the Probit model and then predict the values of residuals. We will go by the first, we have already done the Probit regression. We will predict the residuals. Then we will go by kernel density function, kernel density of its normal function. Then everything will be fine.

(Refer Slide Time: 23:03)



So, we will go by predict r that is our residuals, have been predicted. We will go by Kernel density. So, the k density, r, a value predicted, the predicted values of r have already come.

(Refer Slide Time: 23:32)

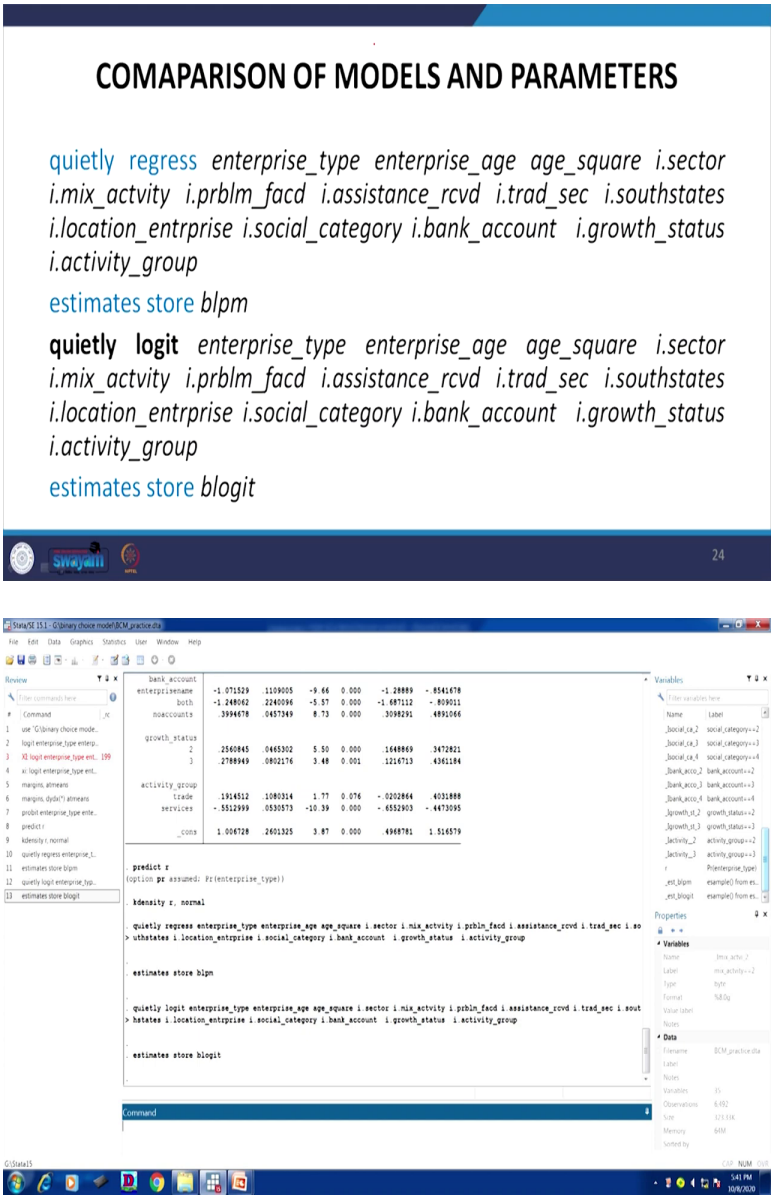


We are giving k density of r. we are going to give a normality test of it. It gives a diagram. It is now estimating or deriving. Still it is running. Look at the diagram, the graph of the predicted values that is highlighted in blue color. The red one is the normal density function, the standard normal density function. The estimated one is in blue or black color.

Lookat very clearly. The graph seems asymmetric. It is not symmetric. It is nowhere closing to the normal distribution. The normal density function is not overlapping with this. Or nowhere near to the normal density function. So, that is the reason why we are not accepting these

distributions to be normal. So, in this case it is suggested that we should apply Logit model. Probit is not the best fit. Now, this is what we said already.

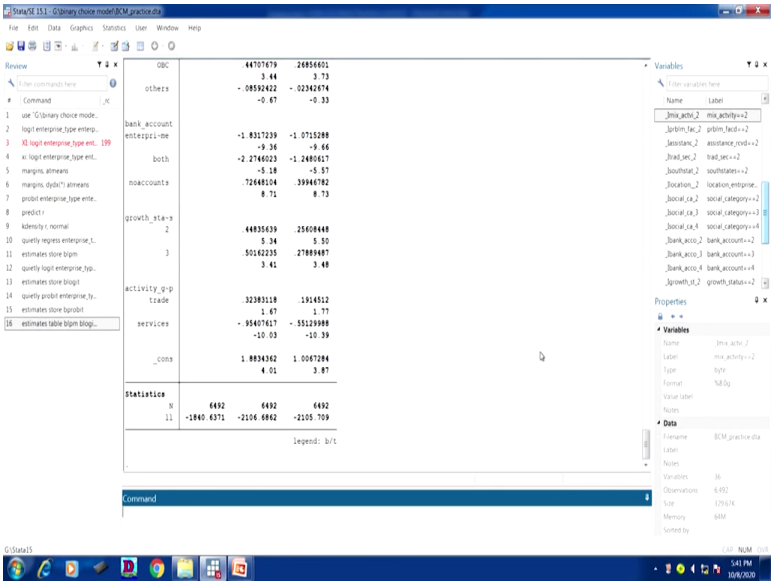
(Refer Slide Time: 24:40)



We are going to compare all the models together and to decide what is the best fit and how we should go for it. Alright, so we have three. Like if you go by this command quietly regress with the LPM model, for the LPM model, quietly regress if you do it, what we do? We can quickly do that. Let me show you. We can quickly copy this command and show you how it works. All those... So, this basically Logit and LPM we are running together.

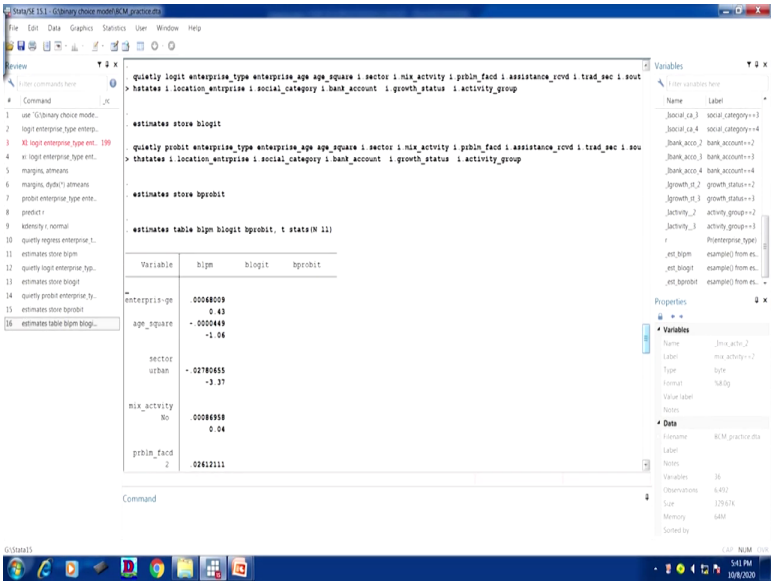
Alright, so both the results, here you can see the new estimated values has already been derived for LPM and for Logit. So, we are going to show the Probit as well. We are also comparing all those three together. In this command we are going to compare as well.

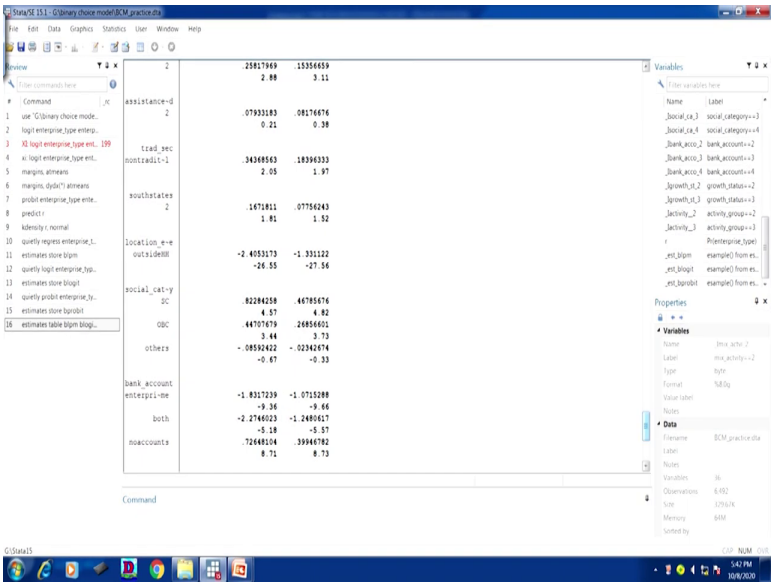
(Refer Slide Time: 25:57)



Alright, first command is for Probit. You can see from this here Logit blpm is for LPM. Then second one is for Logit then third one is for Probit.

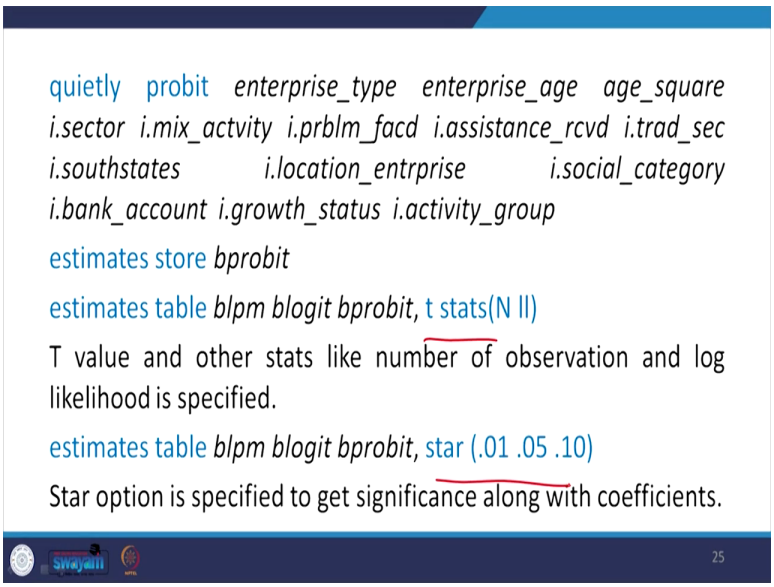
(Refer Slide Time: 26:18)

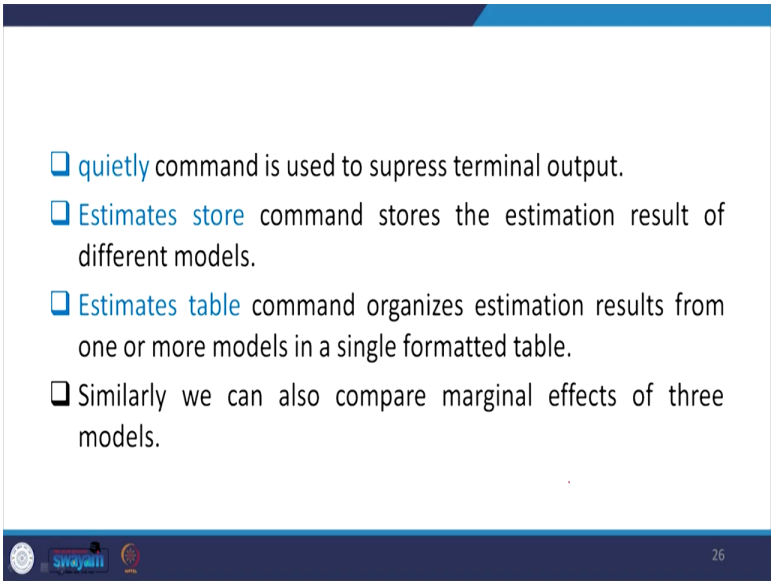




So, after that we are comparing all these three together. This is for LPM, linear probability model then Logit then Probit. from all those models we have got a comparison clearly. First column is for LPM than second and third order Logit and Probit respectively. Each of the coefficients can be compared and how they differ each other. That some of them we have already discussed in the previous equation, that 4 times related, 2.6 times related that you can compare and find out. Alright so this is what we have done.

(Refer Slide Time: 26:49)

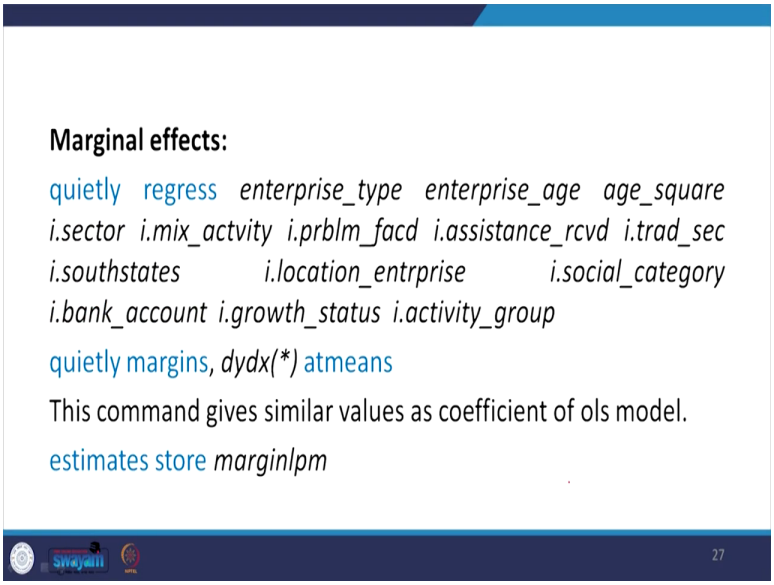




Here I just wanted to mention that T value and other stats like number of observation and log likelihood is specified very clearly. So, here first of all we specify with T and, T and its log likelihood values. But you can also specify its significance level. If you specify the significance level the significance with star mark is also going to come. The star mark against the coefficient will be there. Once the star mark is there you can compare which coefficient which variable is significant in which particular model and it will be easy for you to decide alright.

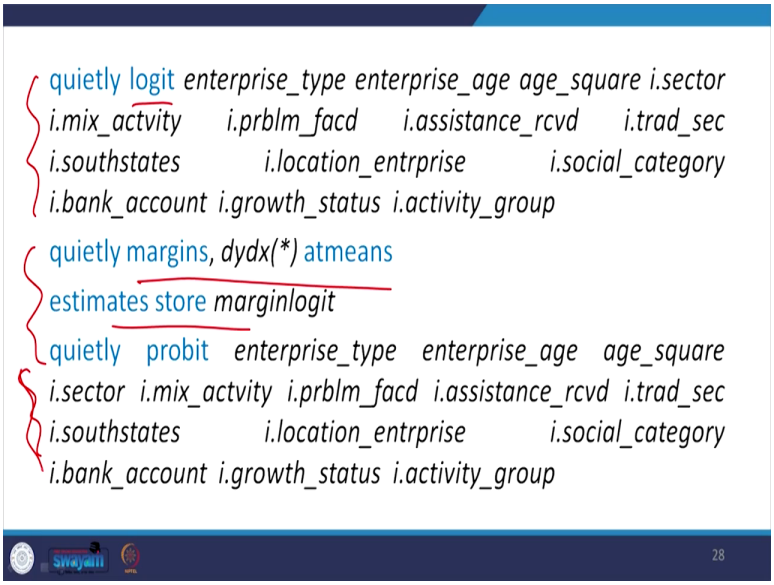
So, that is the reason why quietly command is used to suppress terminal output. And these estimates store basically gives the estimation result of different models together. And estimates table particularly. store keep it in the model and table keeps in all those models together. Alright, so then basically in single format table.

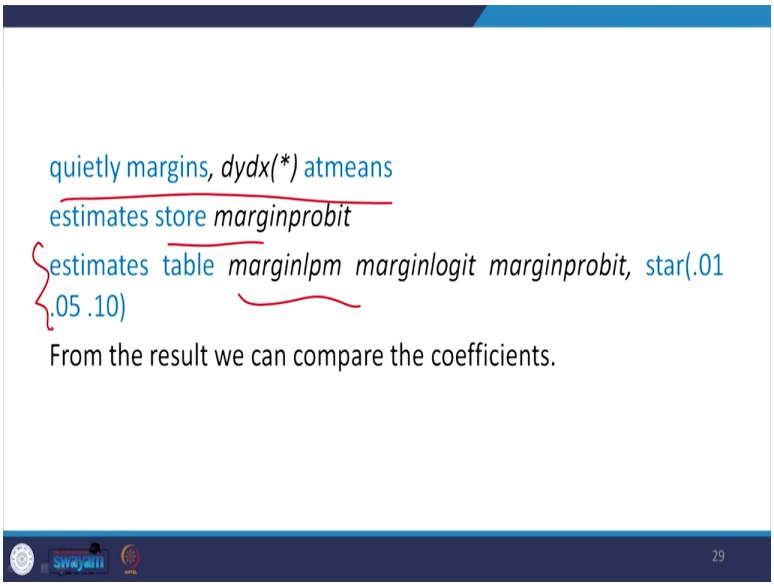
(Refer Slide Time: 28:13)



Similarly we can also compare marginal effects of these three models. Similarly we can type quietly regress with i, quietly regress that is basically the LPM model. Marginal effect is with the same regression model itself whereas after the regression only in case of LPM we have to quietly margin. Margin and dy/dx star atmeans that will give you the marginal effect. But usually that is not necessary in case of continuous data. And LPM since it considers to be continuous series, though the dependent variable is binary that is why we are running marginal effect. Then similar command estimates store margin LPM it will save.

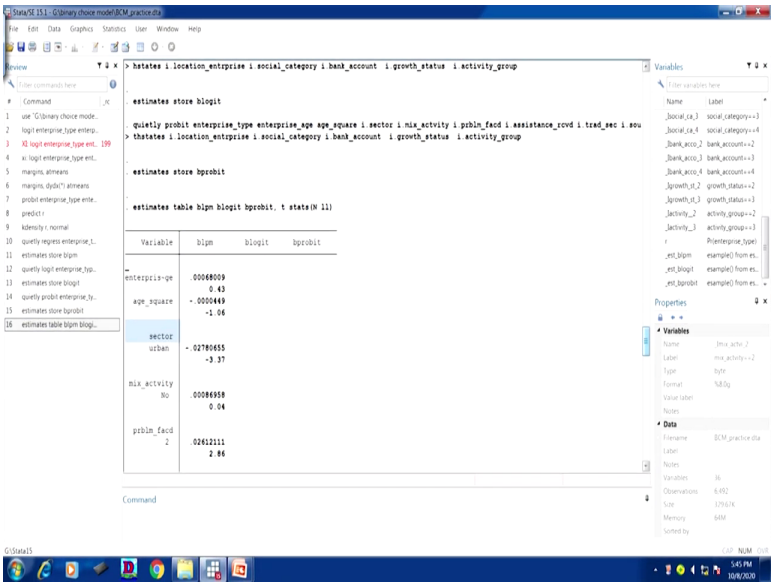
(Refer Slide Time: 29:01)





Then we go for quietly logit, then with all those commands for Probit as well. Probit after running Logit then we discuss about the marginal. Then we also store it, store will save it. Similarly after Probit we will go for the next slide its marginal value atmeans. Then we store it. After storing all those things we can take all those three models together. When you take three models together will be able to decide which one is better.

(Refer Slide Time: 29:49)



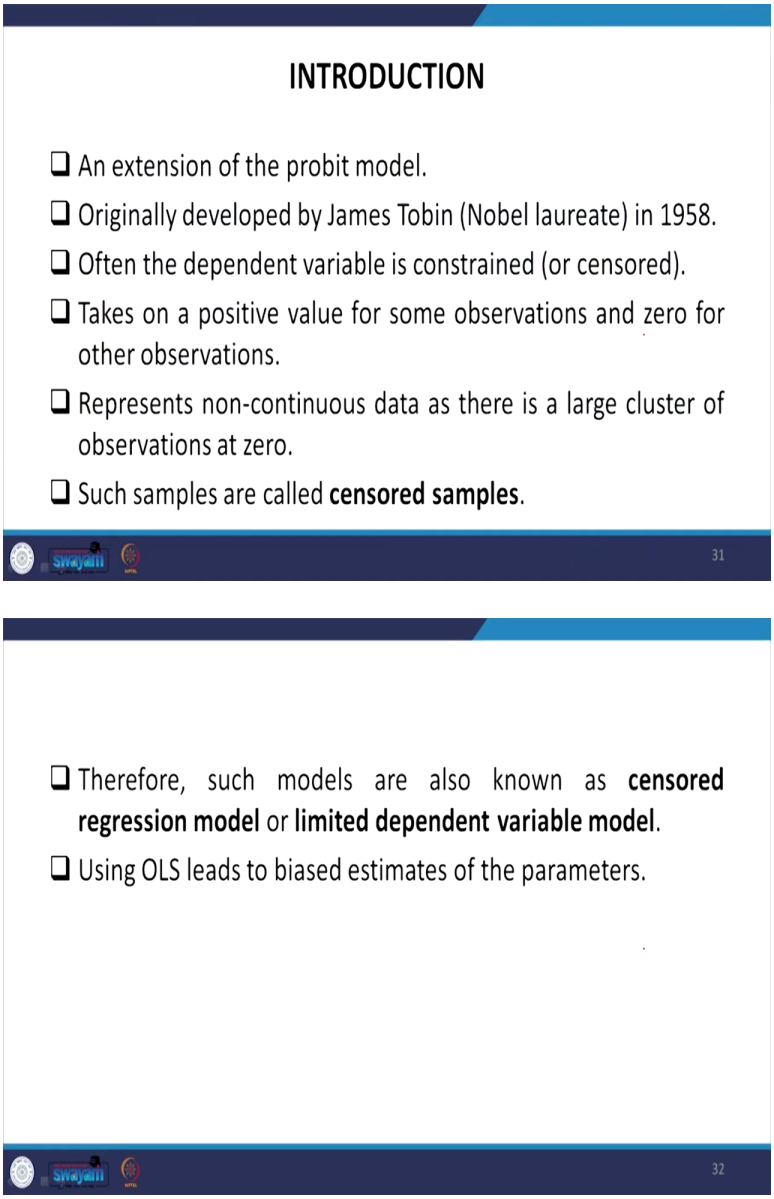
Likewise we have just shown in 3 columns. Here, 3 columns if you are confused, here only coefficients we have shown. But in our other slides we have already mentioned that if you run by the marginal effect the marginal Coefficient will be visible. Alright, so I am not running several times because of paucity of time. So, with that we finished the discussion with LPM, Logit and Probit. And we decided which one is most fitted based on their coefficient, their significance level, based on the distribution function and so on.

Coming to Tobit regression model. The last model in this series is on Tobit which is also an extension of the Probit model. Originally developed by James Tobin as you might have heard as

Nobel Laureate in 1958. Often the dependent variable is constrained. Many cases it is censored. There is a limit.

Like those who are estimating some policy implementation of a particular like PDS scheme for example and that will be applicable for people with below poverty line. So, all the income earners are not going to be considered. Somewhere you have to censor. You have to define a benchmark level of income. To that only you are going to give. If you want to give more number of people in that caveat then you can change that censor limit. And accordingly we can run the regression.

(Refer Slide Time: 31:23)

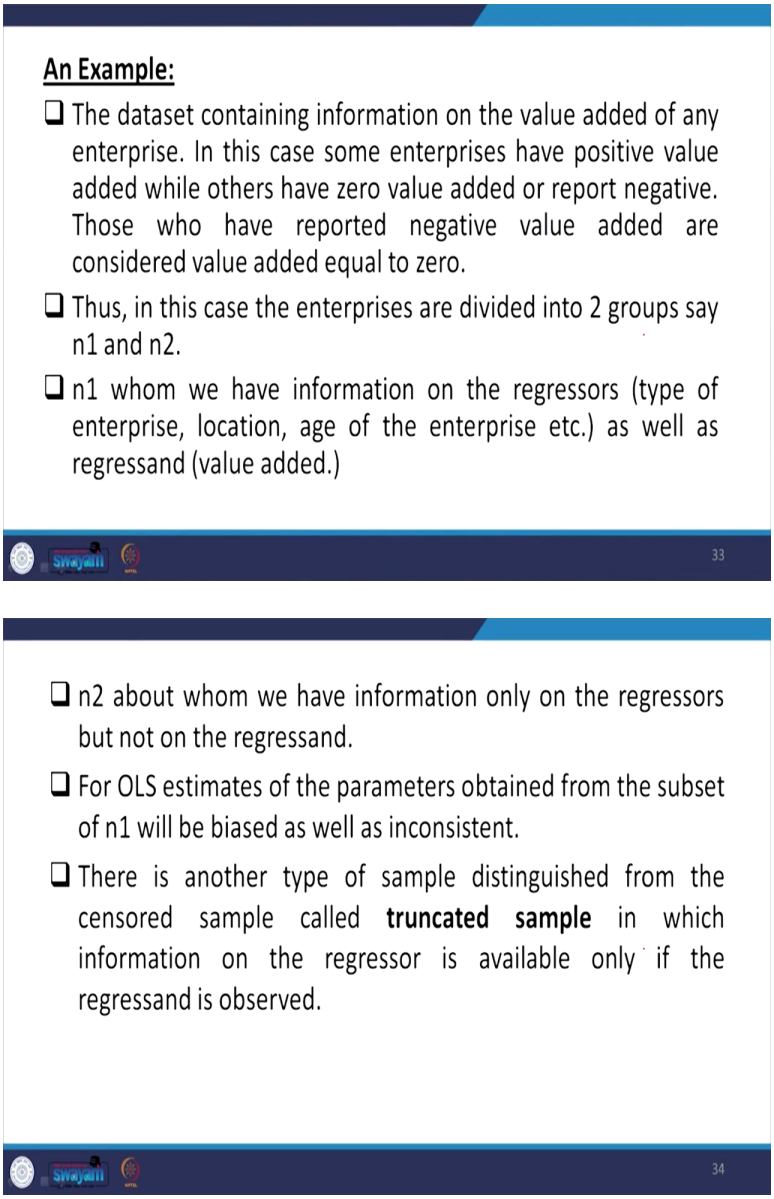


So, usually it takes on a positive value for some observation and others as 0. If you have censored then censored values are considered to be 0. And these represent non-continuous data as there is a large cluster of observation at 0. Such symbols are also called censored samples. Basically those are considered to be 0 are called censored samples.

Therefore such models are also known as censored regression or also called limited dependent variable models. Using OLS , method leads to biased estimates because, in any case this has categorized into 0 and 1. These are not a continuous series all totally though the non-censored

variables or the values are of continuous, could be continuous and different than that of the existing Logit and Probit model.

(Refer Slide Time: 32:22)

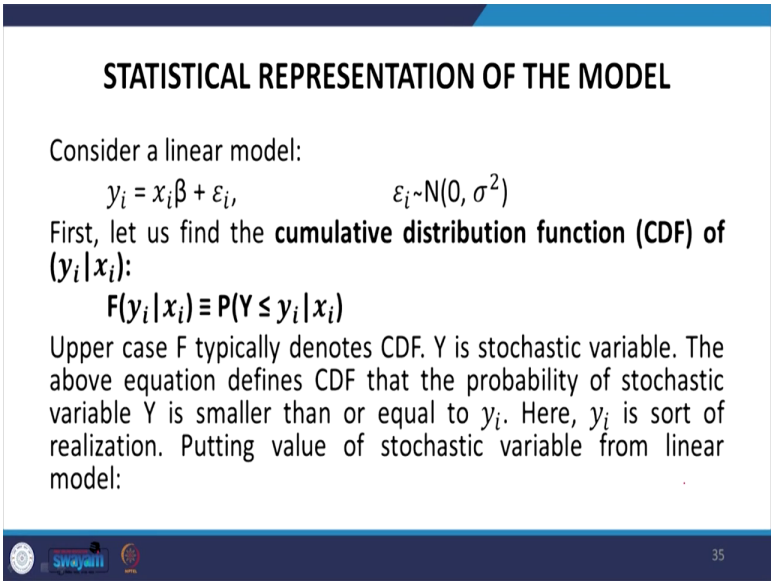


So, let us understand in terms of example. We have already referred the dataset of the enterprises context. The dataset containing information on the value added of any enterprise. In this case some enterprise have positive values added while others have 0 value added. Or they report negative.

So far as the value added of the enterprises is concerned, enterprises and their value added figures are available in the 73rd round. Those who have been reported negative value added are considered value added equal to 0 because no question of negative value added is valid. You can make it to 0. So, thus in this case enterprises are divided into two groups, say n1 or n2. So, n1 whom we have information on the regressors like type of enterprise, their location, age of the enterprise as well as regression that is value added we are discussing.

So, n2 is all about whom we have information only on the regressor but not on the regressand because it is of either 0 values or negative values, or no information. So, for OLS estimates of the parameters obtained from the subset of n1 will be biased as well as inconsistent. So, there is another type of sample distinguished from the censored sample called truncated sample and information on the regressor is available only if the regressand is observed. So, when only regressand is observed then in that case we get the truncated sample.

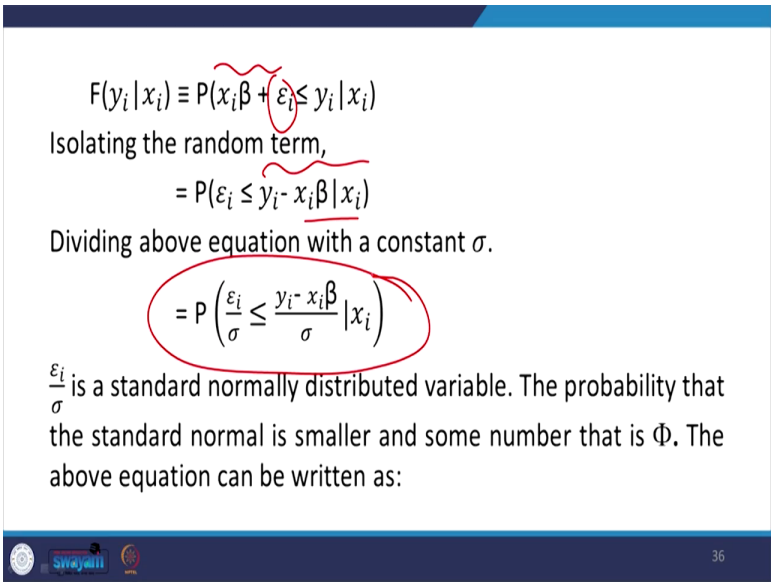
(Refer Slide Time: 34:01)

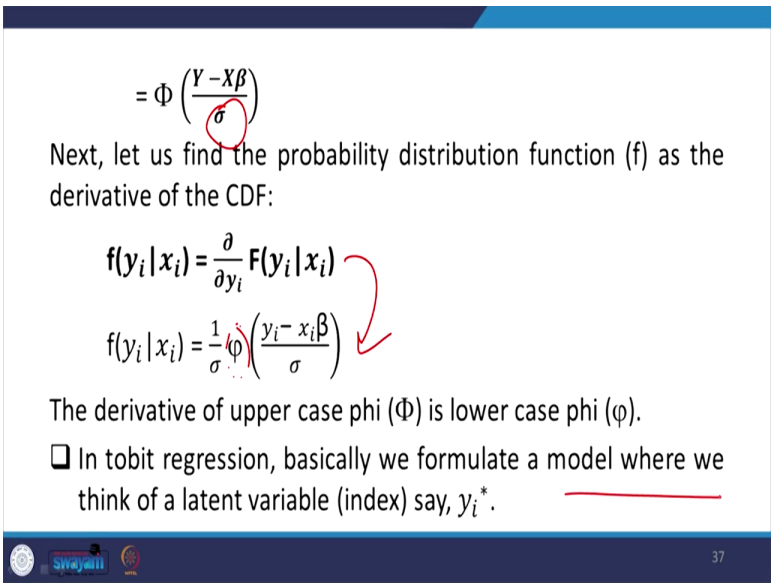


So, in some mathematical interpretation we can simply identify that model is Y of xi beta of epsilon, the error distribution is distributed with 0 and sigma square variance. So, let us find the cumulative distribution function of yi given xi where we are limiting the y with certain level of yi given xi.

So, here a limit is given correctly, less than or equal to that is going to be valid in our case. Or entire yi is not going to be defined in this particular distribution. The uppercase that is uppercase of F typically denotes CDF function or the cumulative distribution function. Y is stochastic variable. The above equation defines CDF that the probability of stochastic variable Y is smaller than or equal to that yi we are defining here. Here yi is sort of realization. Putting value of stochastic variable from linear model.

(Refer Slide Time: 35:13)



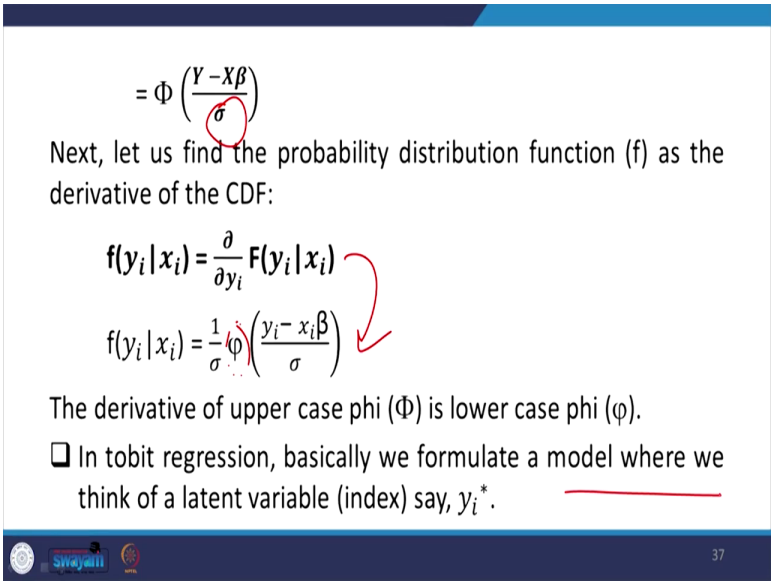


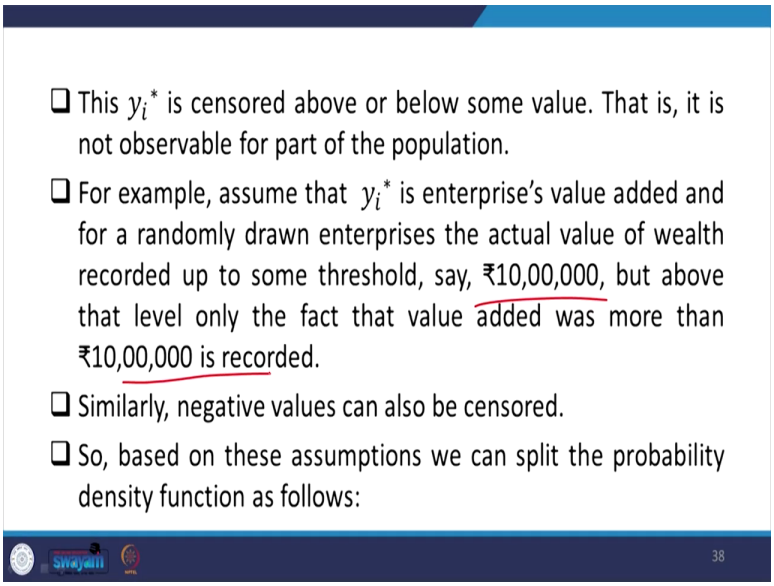
Similarly after discussing this, Y can for the disaggregate to its beta coefficients and epsilon with the limit. Error can be defined. Error distribution is defined as yi minus of this estimated

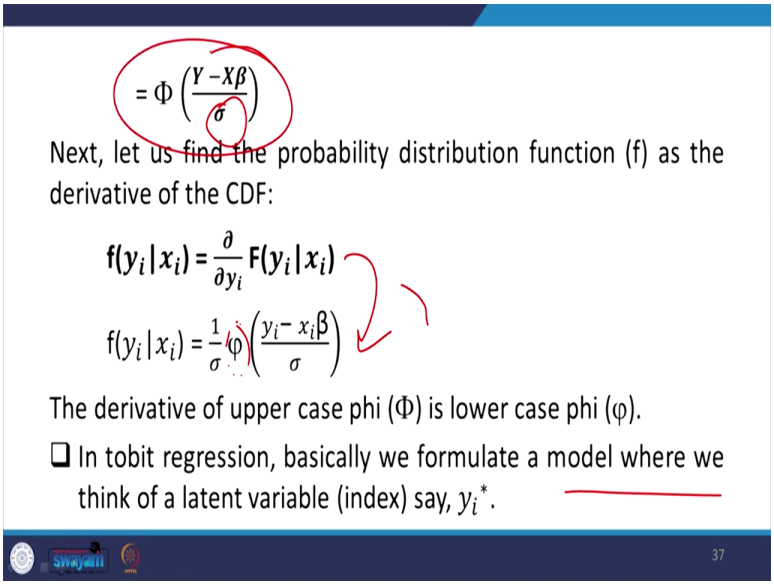
coefficients. The distribution where we are going to estimate given xi can be defined with this probabilistic value.

This is the epsilon with respect to the standard deviation is a normal distributed variable. The probability that the standard normal is smaller or some number that is indicated with phi. The above equation can be written phi, and in terms of phi as a function of yi minus x beta divided by sigma which we have defined just a minutes back. That is the probability distribution function.

(Refer Slide Time: 36:06)





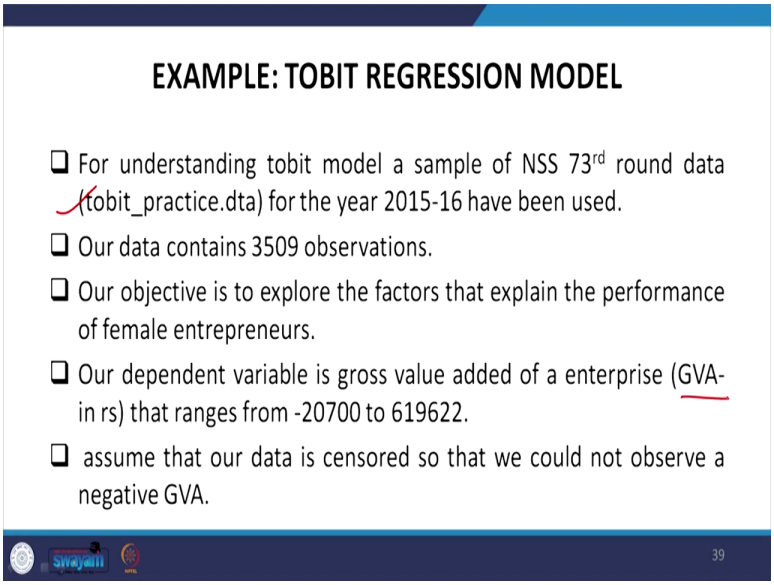


By taking their derivative we get the marginal value. And the marginal value since sigma is there and these are constants. So, this is a small phi and indicate the marginal difference of the equation. This follows from here. And in Tobit regression basically we formulate model where we think of a latent variable, say that is yi, latent variable on observed variables since probabilistic issues are attached to that variable. So, that is why we are mentioning as latent here. So, yi star is censor above or below a certain value. That it is basically not observable for part of the population. And so part of the population since not observable we can censor it, alright.

In this example like yi star is enterprise value added and for randomly drawn enterprises actual value of the wealth recorded up to some threshold let it be 10 lakhs but above that level only the fact that the value added was more than 10 lakh is recorded. Similarly negative values can also be censored if above value and below value or negative value can also be censored. Whichever we require for the necessity of the model we can censor accordingly. Based on the assumptions we can split the probability density function like the way we defined here, like this, alright. So, there are different limits. Like we can define 1 if it is less than certain limit, 2 if it is with another bracket, we can categorize their probability limit.

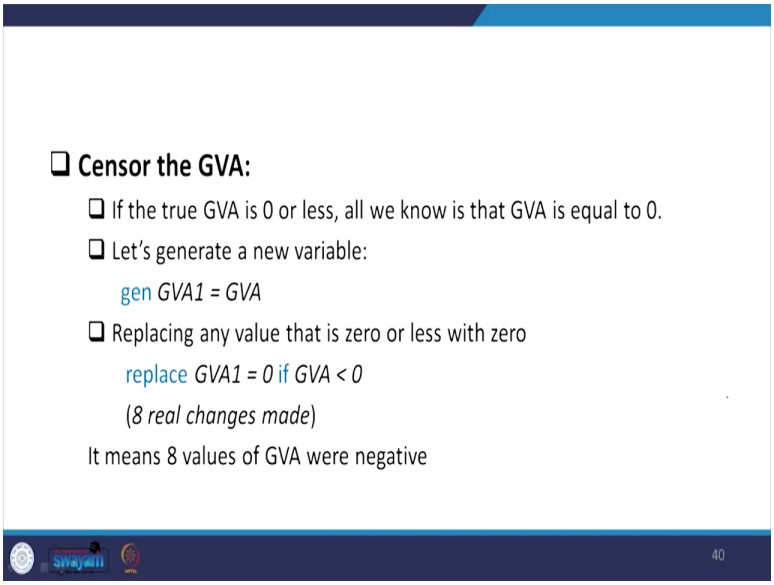
Let us come to the example to estimate it correctly. How you can estimate it? Somewhere we have to understand that there are some limits defined out of the total data we can censor at the below, at the top, or at the medium. So, you can give their coefficient accordingly.

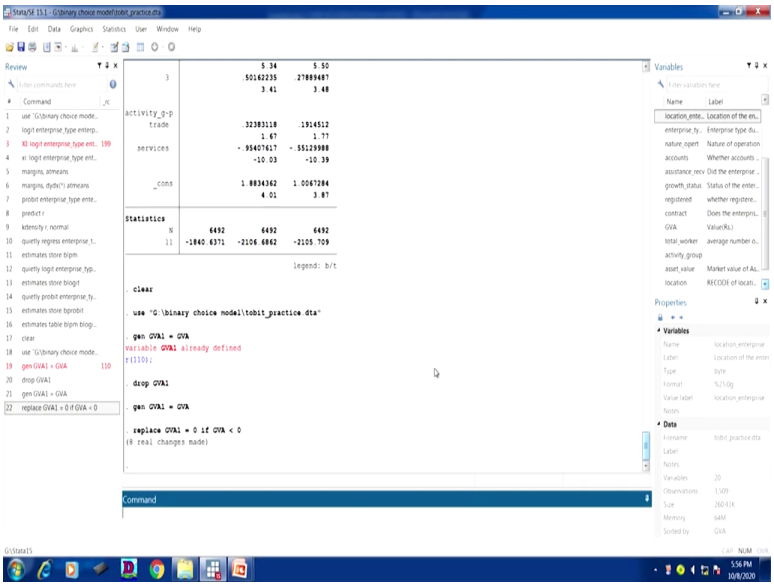
(Refer Slide Time: 38:34)



So, from the 73rd round of the data we have a practice dataset called tobit practice dataset. We are going to open right now. So, our data contains 3509 observations. Our objective here is to explore the factors which explain the performance of female entrepreneurs with a certain caveat. So, dependent variable here to be estimated with a censor value for its gross value added of the enterprise. Gross value measured is in rupees that ranges from minus 20700 to 619622. These are the distribution. Since minus values are there so it is a concern for us. We need to censor it.

(Refer Slide Time: 39:18)

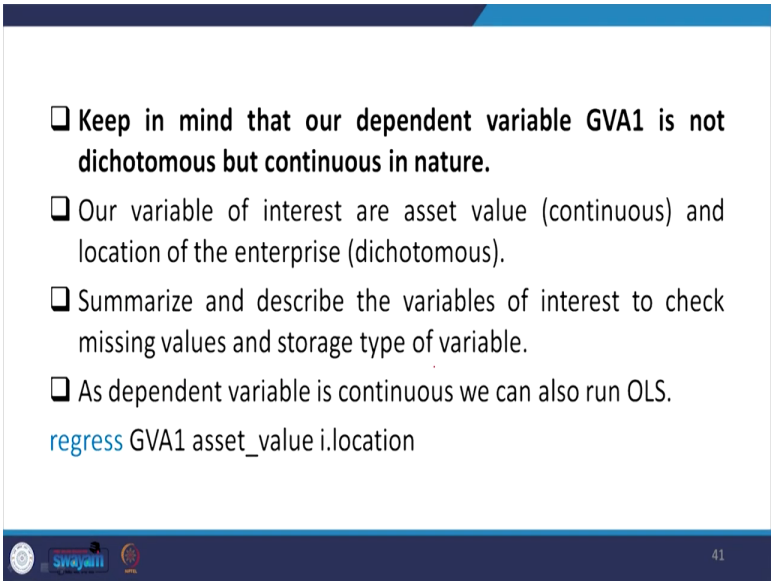




So, how could able to censor it? In the data, so let us open the data first. Then we can able to discuss. So, first of all, just a minute, alright. So, we are going to open the data that is tobit practice. This has already been opened. We are going to do something very quickly, since GVA has already been defined we need to compare GVA 1 with the value if GVA has less than 0 or negative values. Let us redefine the GVA with a new variable called GVA1. If that is having GVA less than 0, alright.

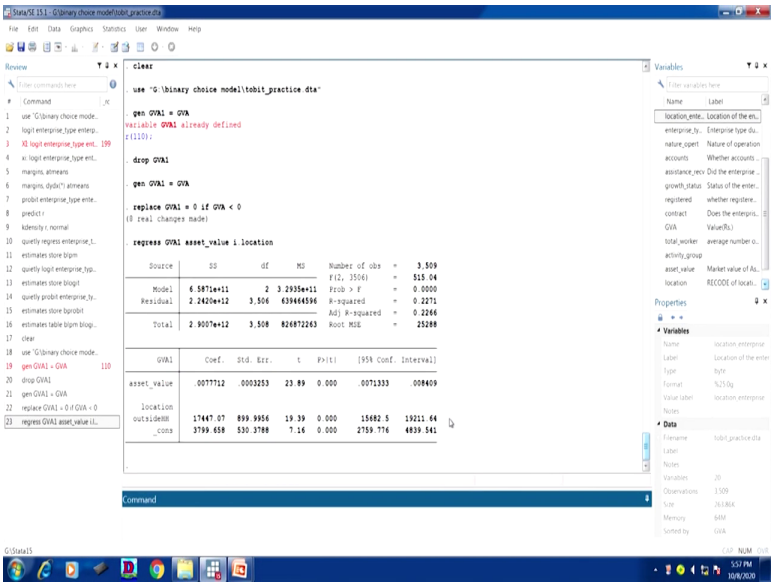
So, it is here. So, we are defining a value here, then this. So, GVA1 we have already defined so we can drop that variable at this moment. That is there, alright. Once again we are going to run the same command. So, GVA1 is defined now. GVA1 is here. It is visible at the end, GVA1. And now we need to replace GVA1 is equal to 0, we want to censor it with 0 value if GVA1 is less than 0. So, very quickly we have got 8 real changes. That means 8 cases where the values are negative. those have been converted to 0, isn't it? So, let us move on and we wanted to find some other interesting interpretation.

(Refer Slide Time: 41:17)



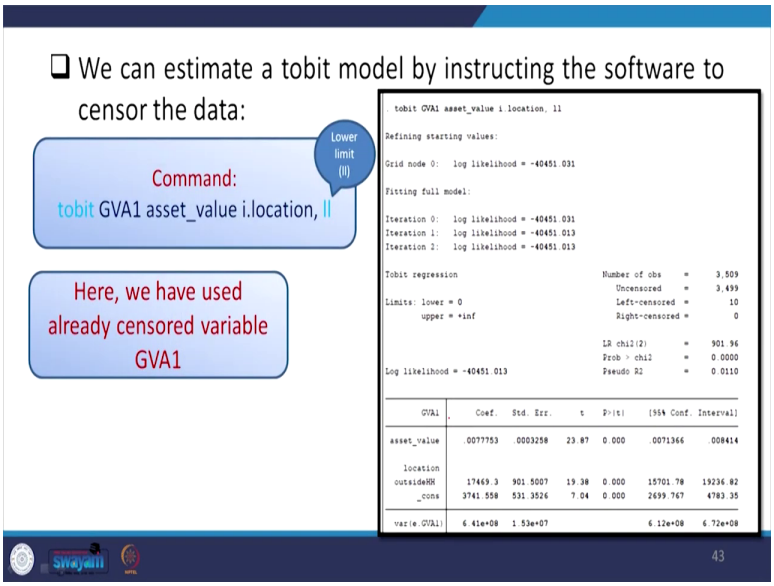
Keep in mind that our dependent variable that is GVA is not dichotomous but continuous in nature. So, that is basically 0 and above, alright. That is interesting to note but usually in Logit and Probit it has to be mandatorily dichotomous. So, our variable of interest here is asset value. That is continuous and also location and any other interesting variable you wanted to include you can include. So, we are going to test with this one as the dependent variable GVA1 right now. So, will simply go by ordinary least square method. Since it's a continuous variable so let us compare how Tobit and these two can be compared.

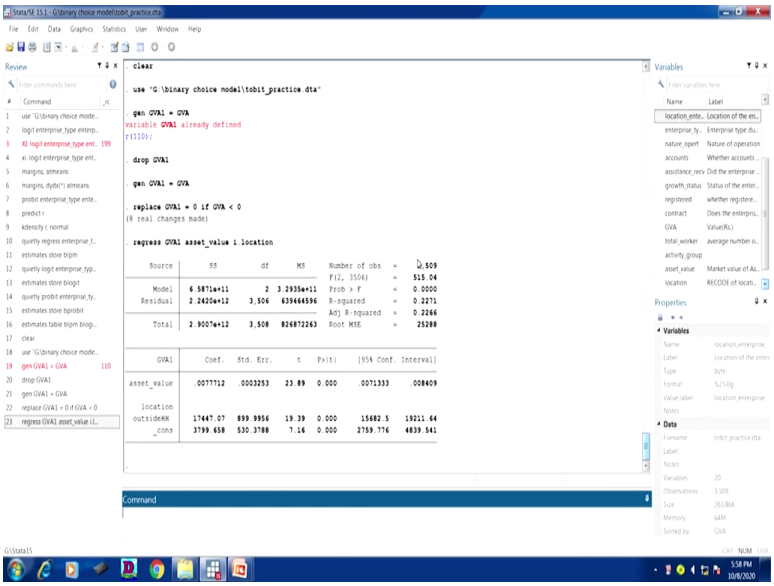
(Refer Slide Time: 42:11)



So, this is our ordinary regression coefficient that has been defined with respect to asset value, location of the enterprises, alright.

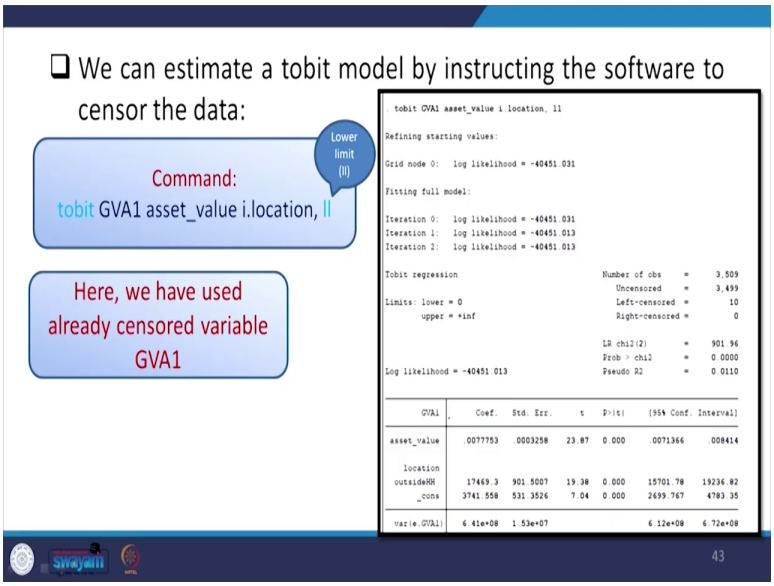
(Refer Slide Time: 42:22)

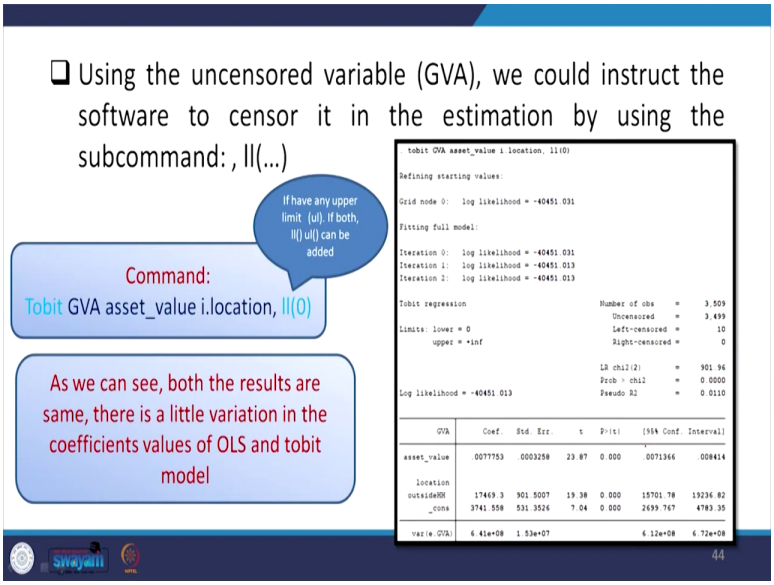




Let us go to the explanation through Tobit. So, we are going to define through Tobit now. We have defined already for ordinary least square. So, let us define through Tobit. So, Tobit is going to give little different result and that is interesting to note. And here the Tobit, our command is here, or at the end we have a lower limit, lower limit in general we have written. One specific lower limit we can also mention. So, when we have defined this ll, at the end ll, not 11, please read this is ll, lower limit. In general the lower limit we have defined. Now, this is the Tobit result and this is the OLS result.

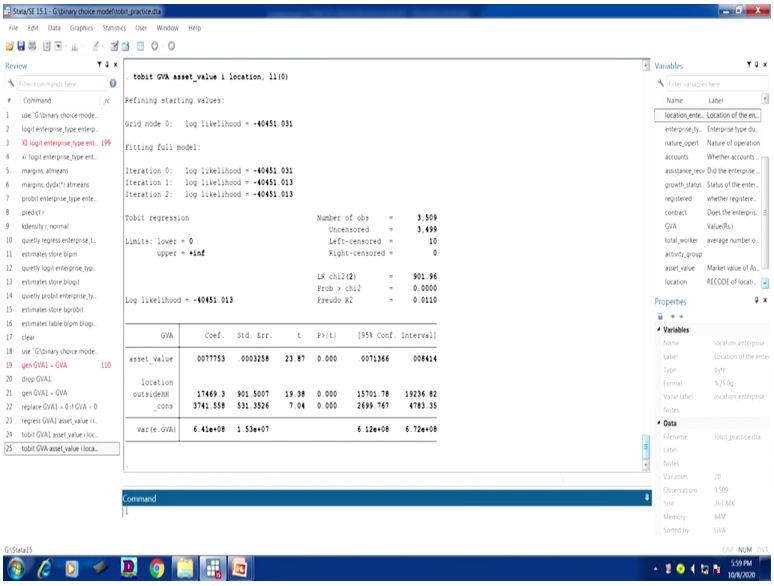
(Refer Slide Time: 43:12)

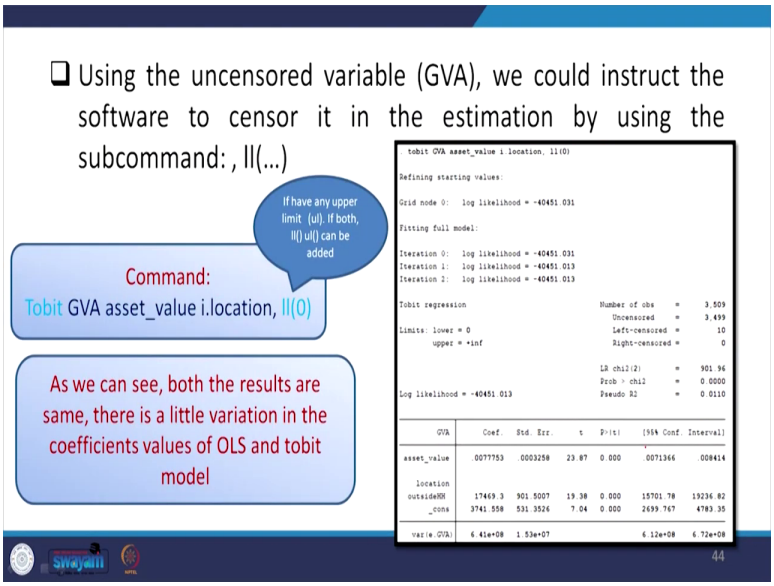




So, what we will do? We will discuss. Here we have used already a censored variable that is GVA1. GVA1 is a censored variable. Using the uncensored variable GVA, we can go by the GVA also like this. So, GVA, uncensored variable so how it looks like? We can compare.

(Refer Slide Time: 43:39)

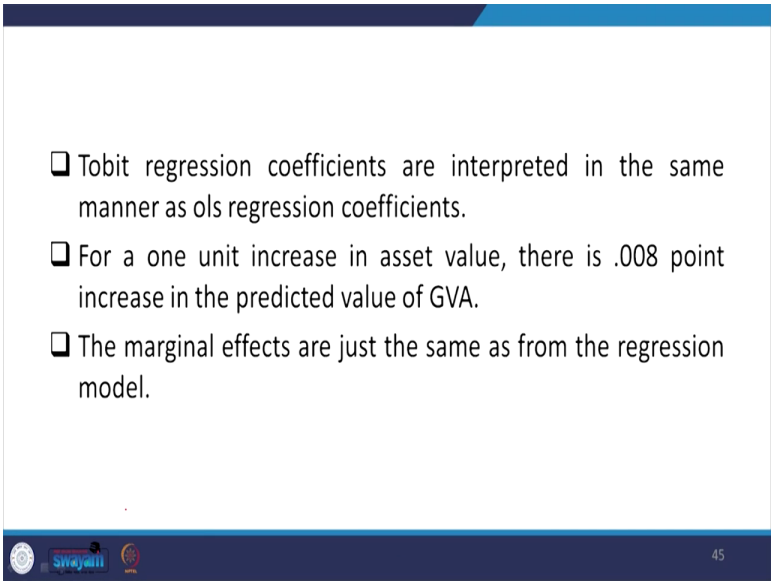


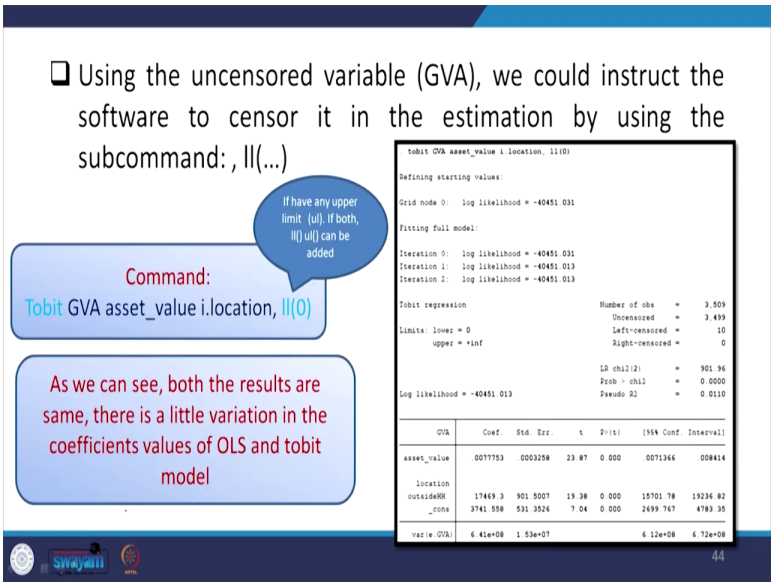


GVAs are uncensored variable. So, in this model it is not GVA1, this is only GVA where the negative values are also included. what is interesting to note here is that, please mark it, here we are mentioning in the command is lower limit. If you have any upper limit you can write down with at the end ul as the upper limit. If both are there, lower limit and upper limit with the particular value, here we are considering 0 as our lower limit because below that we have already censored.

So, upper limit if you have a particular value maybe 66000 or may be 50000 above we are not going to discuss that limit you can set. As you can see in the both the models that there is a little variation in the coefficient values of OLS and Tobit.

(Refer Slide Time: 44:34)





Tobit regression coefficients are interpreted in the same manner as the OLS coefficients. For one unit increase in the asset value there is 8 percent, 0.008 percent increase in the predictive value of GVA. So, here it is the case. With the asset value case 0.00777 so roughly 0.008, alright. So, that is the interpretation.

Coming to this we wanted to say that, not a problem we are just finalizing. For a one unit increase in asset value it is of 0.008 point increase in the predicted value of GVA. The marginal effects are just the same as from the regression model. So, basically all other test you can do it the way we have already done it. So, rest of the details you can experiment and find out. If you have difficulties please come up with the doubts. We will be very happy to explain you in detail. So, all the qualitative dependent variable models, all the dummy variable models have already been completed. So, this is all for today in this lecture. From the next week onwards we are going to unfold discussion of panel regression using STATA. So, thank you very much.