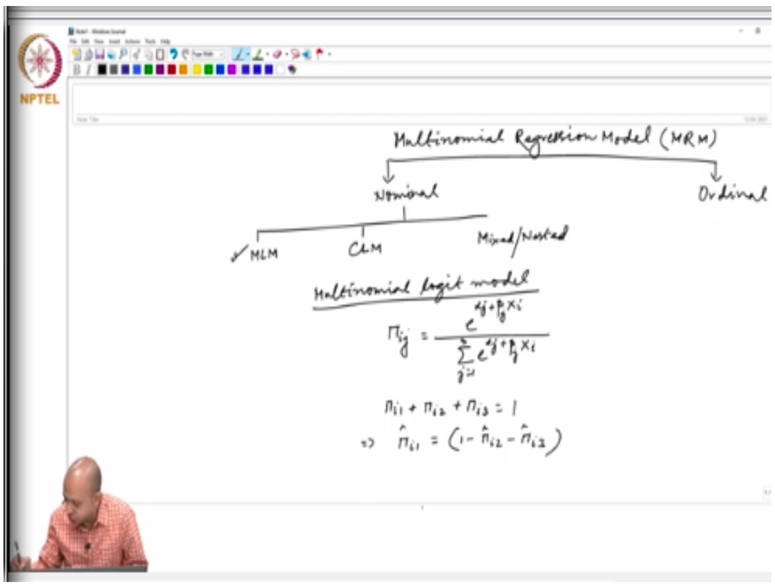
Applied Econometrics
Prof. Sabuj Kumar Mandal
Department of Humanities and Social Sciences
Indian Institute of Technology-Madras

Lecture-35
Multinomial Regression Model-Part II

We are discussing multinomial regression model.

(Refer Slide Time: 00:23)



So, this is multinomial regression model and in sort this is MRM multinomial regression model. And we said that in multinomial regression model basically your dependent variable assumes more than 2 values, so that means it is a polychotomous choice variable. Instead of being the bivariate situation which we have discussed earlier in the context of logit and probit model.

And then this multinomial regression model basically it is of 2 types, namely the nominal regression model and ordinal one. In nominal the choices are not ordered that means there is no value judgment attached but in the context of ordered MRM the options are ranked or ordered according to their preference or value. So, we started discussing with nominal multinomial regression model.

With the nominal we have a multinomial logit model MLM or conditional logit model CLM or mixed logit model instead of mixed I will say that this is mixed or nested multinomial logit model. And we started discussing with multinomial logit model, so this is multinomial logit model. And our model was

$$\Pi_{i,j} = \frac{e^{\alpha_j+\beta_j x_i}}{\sum_{j=1}^{3} e^{\alpha_j+\beta_j x_i}}$$

$\Pi_{i,j}$ which is basically probability that the i^{th} individual chooses the j^{th} option, this is our model.

And we said that here the j subscript is attached with the parameters α and β because the parameters they vary across choices. So, here if we have 3 choices let us say the example what we are talking about a +2 graduate is thinking about to go for a 2-year college or 4 years college or no college. So, that means the individual who is not opting for a 2-year college obviously attached a differential weightage to let us say the income variable as compared to the individual who is actually opting for the 2 years college.

So, that means that the coefficients vary across choices 0 college, 2 years college and 4 years college, so depending on what weightage you are attaching to a particular explanatory variable either option 1, 2 or 3 would become most probable to you, that is what it means which is not like the case earlier. So, that is what we need to think in mind that here α_j j, since j is basically j stands for choice, so our coefficients α and β they vary across choices.

And in multinomial logit model we are trying to explain why this is again basically chooser's specific data. So, that means depending on individual's socioeconomic and demographic factor they will select either option 1, 2 or 3. And we said that we are considering

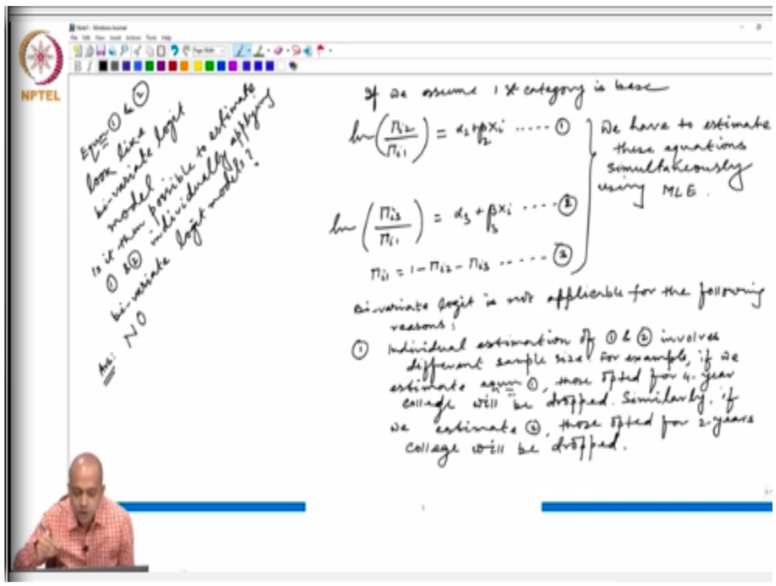
$$(\Pi_{i,1} + \Pi_{i,2} + \Pi_{i,3}) = 1$$

sum of the probabilities = 1. So, that is why instead of estimating this probability individually and separately which is actually not possible.

The moment we estimate $\widehat{\Pi_{i,2}}$ or $\widehat{\Pi_{i,3}}$, $\widehat{\Pi_{i,1}}$ would automatically be estimated by this property which is equals to. So, from here what I can say that $\widehat{\Pi_{i,1}}$ is basically $\widehat{\Pi_{i,1}} = (1 - \widehat{\Pi_{i,2}} - \widehat{\Pi_{i,3}})$

So, one category we have to always take as base and other 2 categories probability will always be calculated in terms of the base category.

(Refer Slide Time: 06:37)



So, we say that $\widehat{\Pi_{i,2}}$ if we assume that $\widehat{\Pi_{i,1}}$ is the base category if we assume let us say $\widehat{\Pi_{i,1}}$ or first category is base then what we will determine? We will basically determine

$$ln\left(\frac{\Pi_{i,2}}{\Pi_{i,1}}\right) = \alpha_2 + \beta_2 x_i$$

So, and let us say this is equation 1

$$ln\left(\frac{\Pi_{i,3}}{\Pi_{i,1}}\right) = \alpha_3 + \beta_3 x_i$$

, this is equation 2 and

$$\Pi_{i,1} = 1 - \Pi_{i,2} - \Pi_{i,3}$$

is let us say this is equation 3.

So, we have to estimate these equations simultaneously using maximum likelihood method MLE. Now the question that we raised yesterday towards the end equation 1 and 2 look like bivariate logit model, so this is like bivariate logit model. Because if we recall in bivariate logit also what we are estimating

$$ln\left(\frac{P_i}{1 - P_i}\right) = \alpha + \beta x_i$$

That means in bivariate logit model we are basically estimating the relative probability of success over failure.

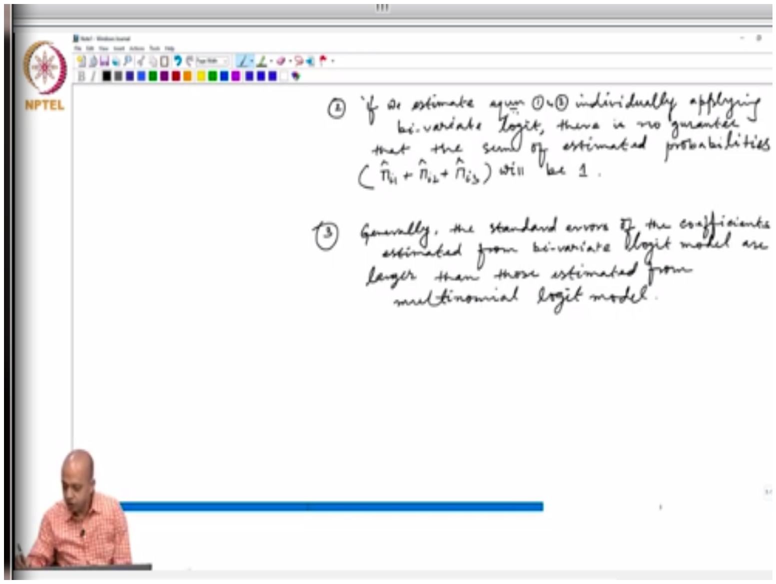
Here what we are estimating? The relative probability of a 2 years college over no college and in equation 2 relative probability of a 4 years college over no college. So, 1 and 2 looks exactly like a bivariate logit model, then the question what we raise? Is it then possible to estimate 1 and 2 individually applying bivariate logit model? Using bivariate logit model? And the answer is, basically answer to this question is no. We cannot estimate this equation 1 and 2 individually and independently using a bivariate logit model. Why this is no?

Why bivariate logit is not applicable for the following reasons? What are those reasons? Number 1, that if we estimate, so that means estimation of individual estimation individual estimation of 1 and 2 involves different sample size. For example, if we estimate equation 1 that means what we are estimating? Probability of a 2 years college over no college. So, those who have opted for a 4 years college, their information is not there in this equation, so from the entire sample those observation will automatically be dropped.

So, if we estimate equation 1 those opted for 4-year college will be dropped. Similarly, if we estimate 2 then those opted for 2-year college will be dropped. So, that means equation 1 requires only that part of the sample, equation 1 involves information from those individuals who opted for a 2-year college. Equation 2 involves those individuals who opted for a 4 years college. Here we said that 2 is for 2-year college and 3 is for 4-years college and 1 is for no college.

So, that means estimation of these equations individually involves a different sample size which is actually non-desired, that is the first reason why bivariate logit is not applicable in this context.

(Refer Slide Time: 15:48)



Then secondly, if we estimate equation 1 and 2 individually applying bivariate logit, there is no guarantee that the sum of estimated probabilities which is basically

$$\widehat{\Pi_{\iota,1}} + \widehat{\Pi_{\iota,2}} + \widehat{\Pi_{\iota,3}} = 1$$

the sum of estimated probabilities will be 1, there is no guarantee for that. Because I am considering them an individual equation and I am estimating from a completely different sample size as well.

So, they are like estimating an equation from a given sample size, they are independent and as a result of which when I take the summation$\widehat{\Pi_{\iota,1}} + \widehat{\Pi_{\iota,2}} + \widehat{\Pi_{\iota,3}}$, there is no guarantee that this sum will be equals to 1, that is second reason. And the third reason is that generally the standard errors of the coefficients estimated from bivariate logit model are larger than those estimated from multinomial logit model.

So, if we estimate the equation separately applying bivariate logit model then the standard errors of the coefficients estimated from a bivariate logit model are larger than those estimated from the multinomial logit model. So, these are the 3 important reasons for which these equations even though the multinomial logit model ultimately reduces to equation 1 and 2 which look like a bivariate logit model, we cannot estimate them separately.

We need to estimate these 3 equations simultaneously, that means we actually have to estimate a multinomial logit model applying the maximum likelihood method, that is the reason. So, this is what we need to know before we actually proceed for the empirical estimation of a multinomial logit model applying a dataset.