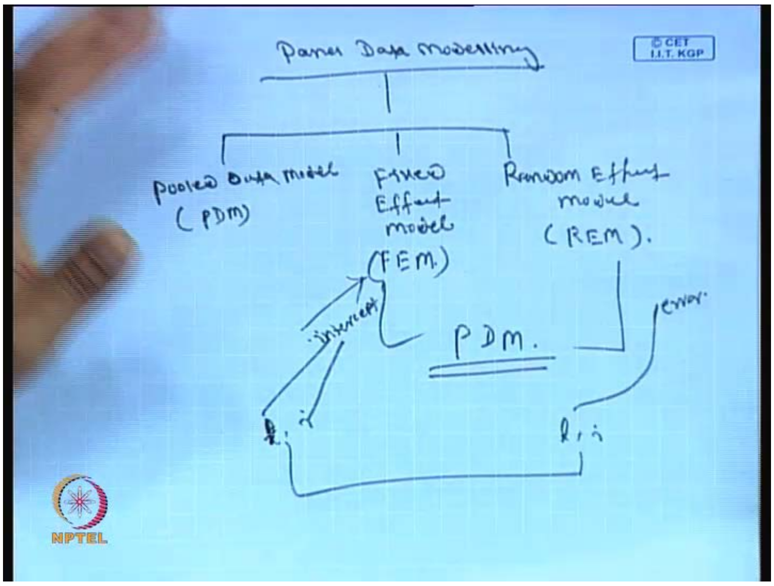
**Econometric Modelling**
**Prof. Rudra P. Pradhan**

**Department of Management**
**Indian Institute of Technology, Kharagpur**

**Lecture No. # 31**
**Panel Data Modelling (Contd.)**

Good evening, this is doctor Pradhan here. Welcome to NPTEL project on econometric modeling. Today, we will continue the panel data modeling. So in the last lectures, we have just entered our discuss little bit about the panel data setting. So in the last lectures, you can say we highlighted, you can say various types of data, then how we move to panel data, and what is the advantage of panel data, and what are the disadvantages of panel data. So today, we will specifically highlight the estimation part of the panel data model. So, suppose a panel data model is concerned, there are three different modelling setup; one is called as a pooled data models, then fixed effect models and random effect models.

(Refer Slide Time: 01:15)




So basically, panel data modeling, so we will together have three different setup, we have three different setup altogether; one is called as a pooled data, then this is fixed effect model fixed effect model, popularly called as an FEM, Fixed Effect Model. Then this is pooled data model, popularly known as PDM, then this is random effect model, popularly known as REM, Random Effect Model. So altogether, there are three different models, pooled data models, fixed effect models, and random effect model. So pooled data, specifically we are just pulling the cross sectional and time series observation without having the impact of, you can say, you know individual effects, and you know time series effect, so we are just you know just observing or pulling the data and estimating the process. So, it is not a pure feature of panel data settings, so pure features of panel data setting we can observe from this fixed effect model and random effect model.

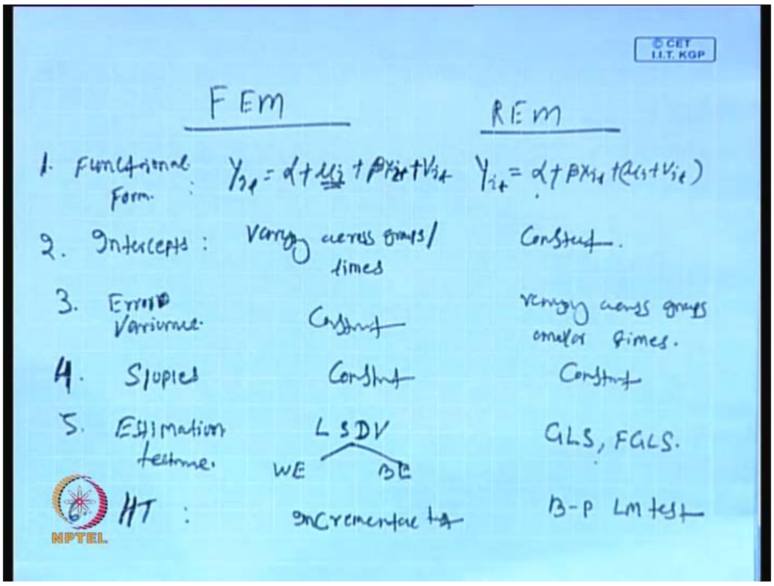
So, before we highlight the detailed estimation process of pooled data models, random effect model and fixed effect model, we briefly highlight the exact difference between random effect models and you know fixed effect models, because these two are the purely represented as a panel data modelling purely represented as a panel data modelling. In one case, the impact on error term this is in this fixed effect term, this is on intercept, the impact is on intercept that means the time and i impact t and i impact will go to intercept. And in this random effect t and i impact will go to the, you know error terms; this time, it will go to the error terms.

So that means, in this panel data setting, we basically highlight the difference between this fixed effect model and random effect model. So fixed effect model, since panel data is there, so there is sample structure is i t, so the i t impact will go to the intercept and t i part will go to the error term. If we observe means the moment we will use i t then obviously sample size will be especially very high and it is just like multiplication of i into t that is the total number of observations we have to observe in the case of panel data setting.

So obviously, it is extremely advantage for us to go for the estimation, but in the mean times, when we will go for clubbing the time series and individual aspects means cross sectional aspects then obviously there is some error component complexity will start. That complexity can be observed through intercept term and can be observed through error terms. If it is observed through intercept then it is called as a fixed effect model and it is if it is observed through error component then it is called as a random effect model.

So now, we will see so that means this particular to this particular to derive fixed effect model and random effect model are very important for this panel data settings. So before you go for the estimation, let me highlight the technical difference between this two random effect model and fixed effect model. So, let me highlight the difference between fixed effect model and random effect model.

(Refer Slide Time: 05:14)




So this is called as a fixed effect model and the random effect model. Now, you see first is with respect to functional form first with respect to functional form, this is with respect to functional form. So we have already mentioned, so for fixed effect model, so the model will be written like this Y i t is equal to alpha plus mu i plus beta X i t plus say v i t. So this is how the fixed a functional form of the functional form of the, you can say this fixed effect models. So in the case of random effect model, this can be written as a Y i t equal to alpha plus beta X i t plus mu i plus v i t, this is how the random effect model.

Then second difference with respect to second difference with respect to is with respect to intercepts, so second difference with respect to intercepts intercept alpha into alpha. So that means, alpha no intercept is mu i here and here it is not there, so that means it is varying it is varying across varying across groups or you can say times and or both it can be times. Here, you know it is intercept is constant; absolutely it is constant in the case of random effect model.

Now the difference can be observed in many angles. So, first observation is for functional form with respect to functional form. Then second difference, we can observe through its impact of intercept and third observation, we have to observe through third observation, you have to with respect to error variance, this is error variance with respect to error variance. What is this error variance? So this is constant here, now it is varying across groups or time groups or time or N times

Then fourth difference, we can find through slopes fourth difference we can find through slopes. This is constant and slope will be constant, if it is vary then it is called as a dynamic panel data models. We will be discuss in later stage. In the meantime, this is slopes and constants, so five, and then estimation technique estimation estimation technique this is. What is the technical difference between estimation technique?

We will use here, least square dummy variable technique least square dummy variable technique that to you know within effect and between effect. Then here, we will use generalized least square method and feasible generalized least square methods. So this side means random effect model can be analyzed on the generalized least square method, means can be estimated through generalized least square method or feasible generalized least square method. And sixth is sixth difference we can observe through hypothesis testing. So in that case, we will apply incremental test then we will apply B-P L m test. So these are the complete difference between this fixed effect model and random effect model. So that means, you know when we will go for panel data modelling, we have two different setup. One is fixed effect model and random effect model. So in the case of fixed effect model, we can write start with the functional form Y i t, which is called alpha mu i plus beta X i t plus v i t then U i t equal to alpha plus beta X i t mu i plus v i t.

Now this model can be you know distinguish means they can different from each other with respect to intercept. In one case, intercept will vary with respect to groups or individual time that is what is called as a fixed effect models. In the case of random effect model, the intercept is absolutely constant. In the case, again the another component which we will make the difference is error variance, the random for random effect models error variance will be very, but in the case of fixed effect model, it will be remain constant. In the case of slope coefficients, both fixed effect model and random effect model are absolutely constant.

Then estimation technique, in the case of estimation technique for which the effect models, we will apply least square domain variable technique that to within effect and between effect, so that means the technique which we have used in the case of dummy variable modelling. That to dummy dependent structures, because here we are whatever we are discussing in the panel data setting here, all are in a linear model here, we are not using any a non-linear models. The way we have discussed in the case of dummy variable modelling, where we have used logic model and profit model where these are purely a, you know non-linear model. Where you know a binary search model or linear probability model is specifically linear in nature, but in other case, it is purely non-linear in natures. When you have dummy independence technique dummy independence technique, where we have also discussed entire means, the entire discussion wise with respect to linear model only.

So that particular, because dummy dependent dummy independent model means that is just like a say multivariate regression modelling, where we are capturing something extra issues there, but the a modelling approach is completely same. For instance, we are just likely means one of the standard assumption, we observe in that particular structure setup is that. We have to apply O L S technique and provided all this parameters should be linear in nature all this parameters should be linear in nature, but which is not the case, in the case of dummy dependence model that to logic model and profit models.

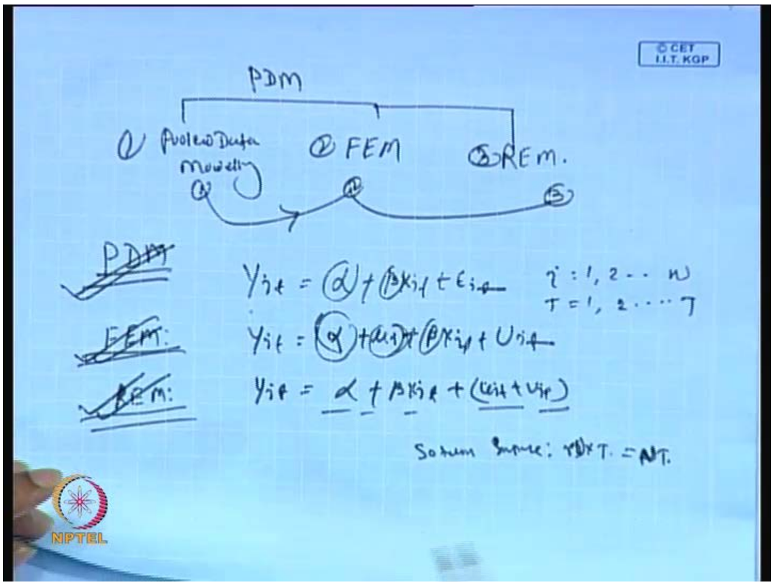
So here also, we are just means this particular panel data modelling is the detective part of this dummy independence technique. So that means, we have discussed and dummy independent variable modelling, which is not means panel data is just you know borrowed from that particular technique. If you are very clear and if you understand very clearly the panels you know dummy independent technique then obviously to understand the panel data modelling is not so difficult or to you know estimate the panel data modelling is also not difficult. But if you have any, you know constant understanding particularly on the dummy independence technique then obviously you will face problem here in the panel data modelling that we need linear setup.

So in any case, I have already given the signal how, what is the exactly panel data setting? How you have to enter to the panel data setting? What are the advantages and what are the disadvantages and what are the techniques, we have to apply? What are the basic models in the case of panel data settings? So that to pooled effect pooled models, then pooled O L S models, then you know fixed effect model, then you know random effect model. Then just now, we have highlighted the difference between fixed effect model and random effect model that to with respect to its functional forms and the rule of intercept, rule of slopes, then you know it is estimation process, then error variance and you know say techniques and it is hypothesis testing, so it is completely different.

On the one case, this slope is usually we are assuming here constant, because we are taking the linear functional form of this models. So means, we are discussing panel data setting, we are just clubbing the time series and cross sectional time. On in the same times, we are assuming that all linearly related to each other. Just we are observing, how cross sectional unit have an impact, when we will club the panel data time series data. And when what is the impact of time series, again when we will club the cross sectional observation with time series that is our objective of this panel data setting. With this basic difference between random effect model and fixed effect model, we like to proceed here the estimation process, you know panel data setting.

So how this, you know techniques can be used and what is the technique estimation procedure? So how, we have to get out the means come out with this an estimation estimated outcome, so that we have to discuss here. Now the basic theme is here, you see here.

(Refer Slide Time: 14:41)



So what we have discussed now, this panel data modelling is basically in three different setup, one is called as a pooled data pooled data modelling, then fixed effect model, then random effect model. So this is one, this is second, this is third. Let us first highlight, what is this pooled data model then we will move to random fixed effect model and random effect model which is very well connected. So you proceed first this one, you first go with this one, then you move to this one, then you move to this one, this sequence is very perfect.

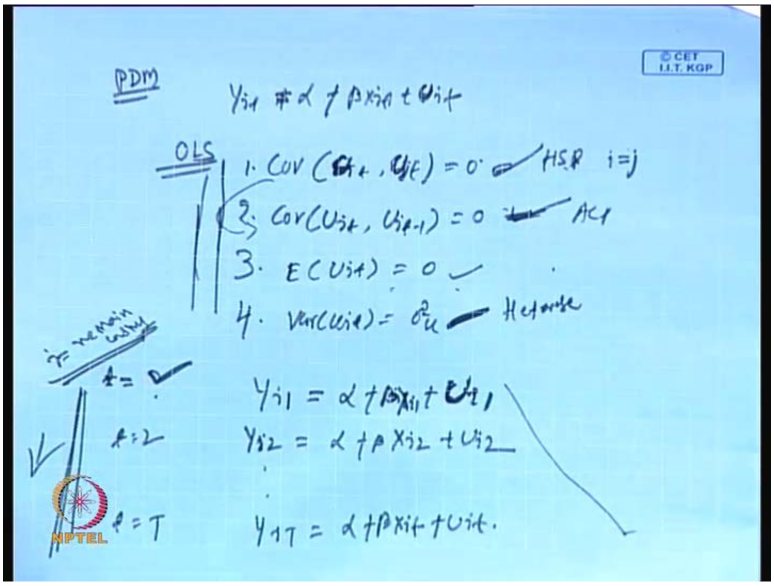
So means what is my suggestion is that, when you are investing some problems with respect to panel, I think it is better you have to follow this particular path. You observe this particular problem or investigate that particular problem with pooled data modelling and then you have to proceed for fixed effect modelling, then you have to go for random effect modelling. Then means, after that you will get the complete picture of penal data analysis. You can means, once you have all these results then you can come to a conclusion, what is the exact safe of this particular structure and how panel data is very handy to investigate that particular you know problems.

So now, what is all about this pooled data setting? I will just observe here. One thing that is what, let us start with pooled data modelling, this pooled data modelling. So what is this pooled data modelling? pooled data modelling is just like here, Y i t equal to alpha plus beta X i t plus summation E i t, this is pooled data modelling the general format. Similarly, in fixed effect modelling fixed effect modelling, we have like this Y i t equal to you can say alpha then we can you see mu i t mu i plus beta X i t plus U i t. then random effect model will be Y it equal to alpha plus beta X i t plus you know U i t plus V i t, so this is how the complete setup of these three models.

So the basic framework of these three models is like this. Now, we like to know how to estimate all these things. So means, we like to know what is this alpha and beta component, we like to know this alpha then mu component and then beta component then also this U observation, V observation. These are the things we have to observe, but we remember one thing here. when we will discuss all these things, there are certain constants or condition, because here we are observing that i equal to 1 to up to N and T equal to 1 to up to t. so this is how so total sample size so total sample size is total sample size in each case is i into t N into T, so that is what is called as N t, this is N t, this is total sample size.

So now, we will little bit know something about the estimation procedure of pooled data modelling, fixed effect modelling and random effect modelling. Let us highlight what is all about this particular issue.

(Refer Slide Time: 18:15)




We start with pooled data modelling first. So pooled data modelling the structure is Y i t equal to alpha plus is equal to alpha plus beta X i t plus E i t. This is how the usual structure of you knows as usual you know structure of cross sectional modelling or you can say then time series modelling. So, it is just clubbing the data time series and cross section, but we use O L S simply. We are not studying the, you know the various aspect here. So this is how the pooled data modelling is all about.

But pooled data will be handy, when it has to means once we will apply the O L S, then O L S is standard assumption. With the basis of standard assumption, we have to go for the estimation. So, what are the standard assumptions here? Now, the basis standard assumption is that covariance of summation i t summation j t summation it is epsilon i t or it is better you put here U i t, so then U i t then U j t is equal to 0, this is standard one assumption.

Second assumption is that covariance of U i t and U i t minus 1 is equal to 0 that is what correlation problems, this is heteroscedasity problem. then third is the error of E error of mean of error term U i t must be equal to 0, this will be lead to heteroscedasity problem, this will be lead to autocorrelation problem and this is sum of the mean deviation mean from the central point should be equal to 0 that will has to be satisfied.

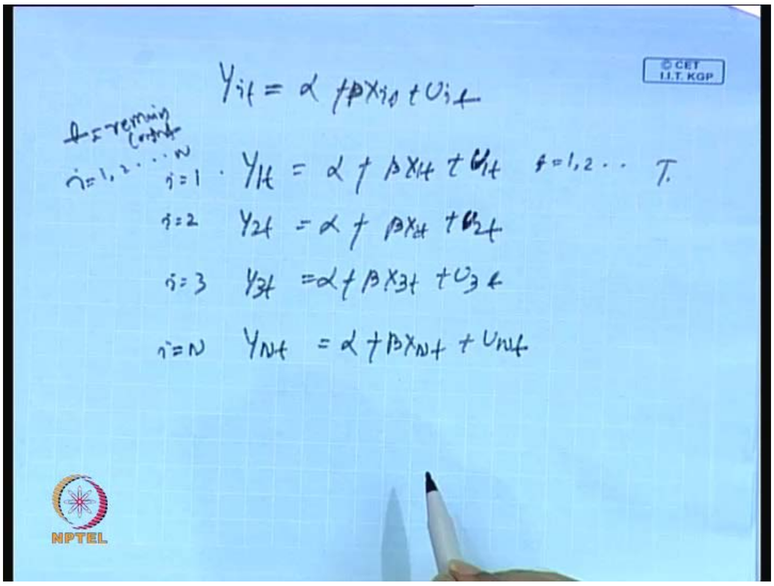
Four, then variance of variance of U i t should be sigma square U. in fact, this is you know heteroscedasity problem in fact this is heteroscedasity problem, if i naught equal to i equal to j, if i naught equal to j then it will turn down like a autocorrelation problem, this is like a autocorrelation problem. Now, in the if all these condition are satisfied then you know we can spending it is means then we can estimate by spending with time period only. Let us see there are two ways you can estimate this process. You know, first keeping i constant you the t variations and keeping t constant you make the i variations then we will see how is this structure altogether.

Let us start with the first, you know varying t and i remain constant, so you see here t and t remain constant, if t remain constant then the model will be Y i 1 equal to alpha plus beta U 1 plus summation U i 1. So that means, here t is varying if we will start varying t then obviously when t equal to 1, then we will find Y i 1 equal to alpha plus beta i 1 beta X i 1 plus U i 1, this is first model. Now, when t equals to 2 then Y i 2 is equal to alpha plus beta X i 2 plus U i 2, it will continue. Now when t equal to t then obviously, Y i t is equal to alpha plus beta X i t plus U i t, so this is how the structure is all about.

Now you see the model is represented is here, Y i equal to alpha plus beta X i t Y U i t. Now, the condition is that here, you have to estimate the model and by default while X is the easiest technique to apply. So for applying O L S conditions means O L S technique this condition has to be satisfied. Let us assume that this condition is satisfied, and then obviously the model will be categorically grouped like this. That means, one case here we are putting i remain constant i remain constant, if you put i remain constant then obviously t is making variations then obviously the structure will be like this.

So now in another case, keeping t remains constant, if we will make i variation then obviously we will get another setup. Let us see if t is varying t remain constant and i varying how is the set.

(Refer Slide Time: 23:18)



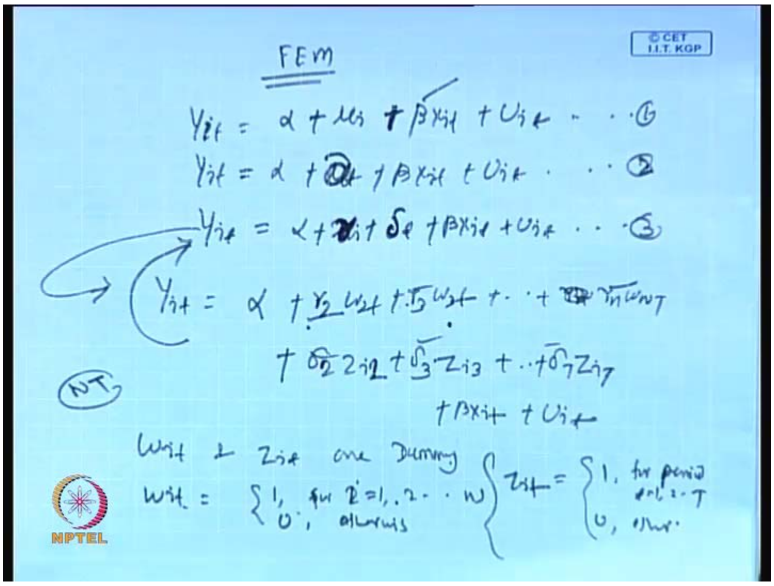
Now, same structure we will start with you now, Y i t equal to alpha plus beta X i t beta X i t plus U i t. then when t remain constant when t remain constant and I will vary 1 to up to N then what is the setup? Then the model will be Y 1 t is equal to alpha plus beta X 1 t summation 1 t plus t equal to 1 to up to t, so this is first case.

So, when i equal to 1, this is the structure. When i equal to 2 then Y to t equal to alpha plus beta X 2 t plus summation, this is U it is beta, we put U no U 1 to so this is U to t U to t.  Then when i equal to 3, then Y 3 t equal to alpha plus beta X 3 t plus U 3 t. So this is how the proceeds. When i equal to N, because i varies from 1 to N, so Y N t equal to alpha plus beta X N t plus U N t, so this is how the procedure is all about.

So now you remember here, the moment you use you know pull that analysis. Then obviously you have to apply O L S technique means say, if provided your condition must be satisfied whatever we have discussed the covariance of U i upon Z t remain constant. And you know t upon t minus 1 t minus 2 keeping i remain constant, it should be equal to 0. And you know variance of error term must be constant and obviously mean of the error term should be equal to 0. If all these conditions are satisfied then you can simply apply the O L S technique. But if you apply O L S technique, then obviously there are two different setup models we will means, it will give you signal of two different aspects of modelling. In one case, i equal to constant t vary and another case, t remain constant i will vary. If i constant t vary then that is the setup and if t is constant, i will vary then this is the setup. This is how the complete setup of the modelling.

Now with the help of information, you can just apply the O L S technique and have the estimated results. then as usual you have to go say you know specification test overall fitness of the test and also you go for other problems, which means you have to test like this heteroscedasity, autocorrelation issue, then multicollinearity issue and so on. This is how the altogether about the pooled data panel data modelling. Now, in fact before going to panel data modelling, pooled data modelling the basic structure of pooled data modelling must be known to us, that is how we have discussed.

(Refer Slide Time: 26:30)




Now, we will move to fixed effect modelling. So, what is the structure of fixed effect modelling? What is the fixed modelling? So fixed effect modelling is general framework is like this, so Y i t equal to alpha plus mu i plus beta X i t plus U i t beta X i t, here we are assuming slope remain constant, both the cases we are assuming slope remain constant. So once slope remain constant, then obviously the structure is like this way.

So when the model can be written like this way also Y i t equal to alpha plus, in fact mu t keeping i remain constant plus beta X i t plus U i t. This can be written this is this can this is form one; this is second form you can write. Then there is another form format, you can use that is equal to alpha plus mu i, it is better you put here. This is delta t, let us say this is delta t, so mu i this is delta t plus beta X i t plus U i t, this is third format. So any format you can use for this fixed effect modelling.

So now, we use this particular model, because it is more compact to analyze the panel data modelling that to fixed effect modelling. When we will go for fixed effect modelling, then obviously the impact of individual and intercept time series will go to particular intercept terms. We like to know what is the impact of you know how this impact of individual and times will go to the intercept. If it is purely going to intercept, then we are observing through mu i. And if you are going purely towards the time series, then obviously it is delta t. And if your observing to, then you will club it mu i plus delta t.

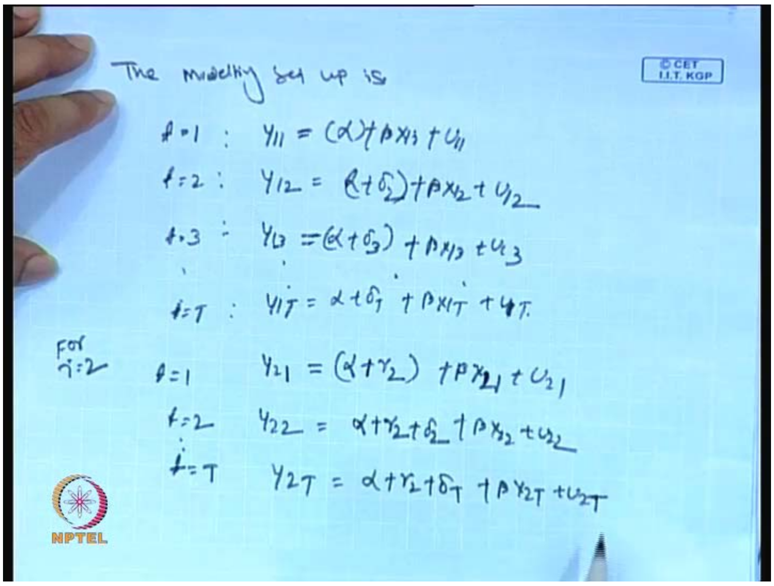
Then, if we will you know classify this particular structures means, if we will put it in an elaborate way, then the model will be like this. Y i t equal to alpha plus like this you know let us say mu 2 or gamma 2 w 2 t plus gamma 3 w 3 t plus gamma N w N t, then plus delta 1 delta 2 Z i t Z i 2 plus delta 3 Z i 3 plus continue delta t Z i t plus beta X i t plus U i t.

That means what we have done here. Here, what is here W i t are dummy here, W i t and you know Z i t are dummy. So this is about the dummy application used here in the panel data setting. So, one observed through time series issue, another is observed through cross sectional issues. What is W i t? W i t equal to you know it is equal to 1. if i, you know i equal to 1 to up to N, so this is W i t so this is W i t. Then it will be equal to 0, otherwise. Similarly for Z i t, it is equal to 1 for period only. For period that is t equal to 1 to up to t and 0 otherwise. This is how the structure is all about this particular setup.

So now, what you have to do? You see here, so here there are total observation is N t. these are the means these are the parameters which we like to estimate. So that means we have taken alpha to you know W i t is the cross sectional observation. So cross sectional observation, in fact we have put here mu i. Instead of mu i, you put it here gamma i, so that it will be easy to understand, it is better to put gamma i here, then it is delta. It is already delta, so delta t that is how, because it should be very consistent with that particular, because we just expanded this particular model with this particular model.

So with the having this particular structure, now you know there are there are three different aspects here. Keeping t remain constant, if i will vary, how is that setup and keeping t remain constant, it will i will make vary, then how is that setup. The way we have discussed in the case of pooled data analysis. Let us see here, so how is that setup?

(Refer Slide Time: 32:04)

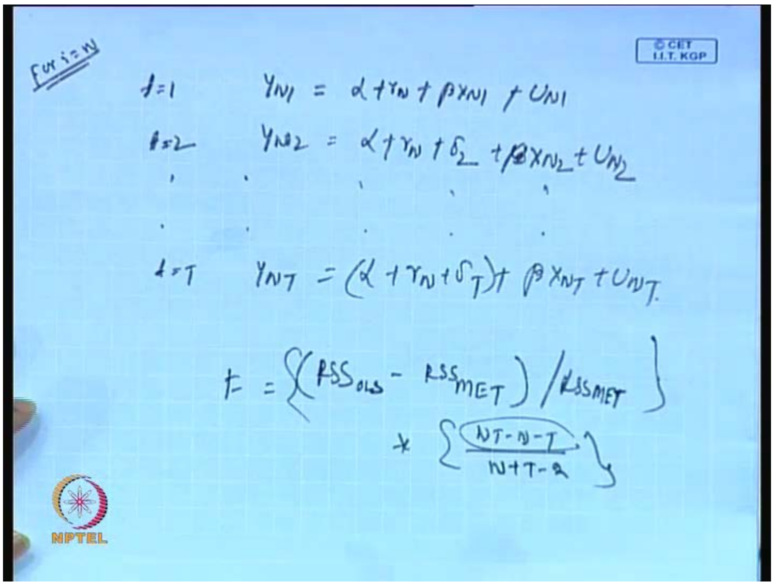



So now this modelling setup is like this the modelling setup modelling setup is like this modelling setup is here, for i equal to 1 so that means what we will do, we will make i variation and t variation. So accordingly we will see how is the modelling be various for the estimation angle only. when i equal to 1 or i equal to 1, when t equal to 1 then it will be Y 1 1 equal to alpha plus beta X 1 i plus U 1 1, so this is this is first model.

Then when t equal to 2 t, then Y 1 2 equal to alpha plus delta 2 plus beta X 1 2 plus U 1 2, this is another model. Then, when t equal to 3 t then obviously Y 1 3 equal to alpha plus delta 3 alpha plus delta 3 this should be one intercept. This will be coming to intercept format, then plus beta X 1 3 plus U 1 3, it will continue. then when t equal to t, then obviously U 1 T equal to alpha plus delta t plus beta X 1 T plus U 1 T, this is how the first setup, when i equal to 1.

Similarly, when i equal to 2 when i equal to 2 then no, for i equal to 2 when t equal to 1 then the model will be Y T 1 equal to alpha plus gamma 2 plus beta X 2 1, this is called this is right, then beta X 1 beta X 2 1 plus U 2 1. When t equal to 2, then it will be Y 2 2, then obviously, it will be gamma plus alpha plus gamma 2 plus delta 2. It is both will added now plus beta X 2 2 plus U 2 2. Similarly, when it will continue, when t equal to t, then Y 2 T is equal to alpha plus gamma 2, because i equal to 2 here, then delta t plus beta X 2 T plus U 2 T. this is how U 2 T for this is i equal to 2 only.

(Refer Slide Time: 35:11)



Then similarly, it will continue then for i equal to N when for i equal to N, then T equal to T 1, T equal to N means Y N 1 equal to alpha plus gamma N plus beta X N 1 plus U N 1. Similarly, Y N means T when T equal to 2 or i equal to N, then Y N 2 equal to alpha plus gamma N plus delta 2 plus U beta plus beta X N 2 plus U N 2, it will continue. Then, when T equal to T, then obviously Y N T is equal to Y N T is equal to alpha plus gamma N plus delta T plus beta X N T plus U N T. This is how the model has to be expanded. So that means, as usual the format is with respect to means, we are targeting the various on individual, variation time series, variation on intercept. Obviously, we have gamma i and delta T gamma will gamma i will take the care variation of individual units and delta T will take care the time series effect.

Ultimately in one case, when we will make i will vary and T will vary the model modelling structure will be completely means, it will go accordingly. So what we have done here, keeping for putting for i equal to 1, then we are making t variation then we get the setup. Then i equal to 2 then T vary, then you will get another setup. When i equal to N and T vary will get another setup. This is how means, like this we will get N number of means N number of modelling setup.

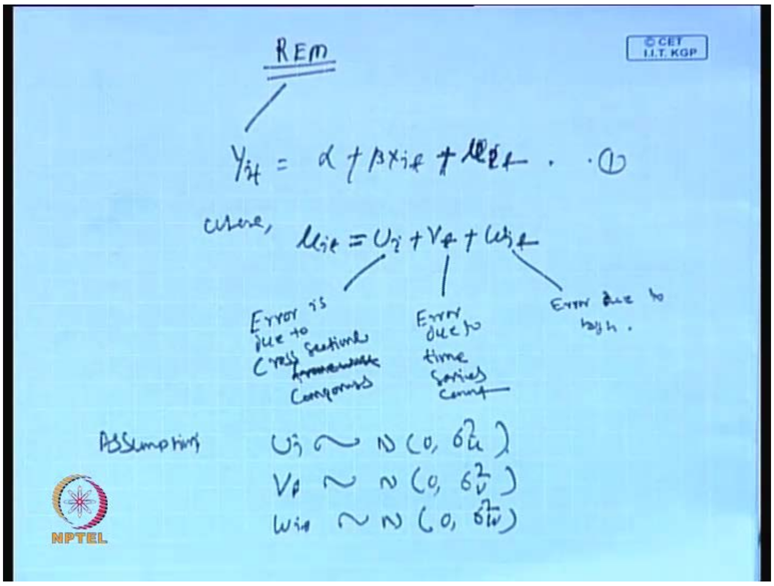
So, with this we have to estimate. So this is very complicated case and it is also variant testing in the same time, because you just make a look when the particular original particular equation. Then you make i vary or T vary, accordingly you will get subsequently different sets of models. So it is very interesting, if you have mathematical knowledge then obviously you will just follow this structure and you create so many mathematical models means statistical models.

So now, the thing is that it is not only enough to just derive the statistical form of the model, keeping i or T 1 time it is constant and another time it will vary. It is same time, we like to know what sigma the exactly our objective here to represent all this structures. Basic objective is to know this a statistical significance of these particular models. So, we like to know all these parameter has to be statistical significant and the overall fitness of the model should be statistical significant.

So far as overall fitness of this model is concerned, we apply here F statistics. We have to apply here, F statistic. F statistic formula is here, R S S ordinarily least squares minus R S S (( )) M E T. Residuals, one is restricted another is unrestricted and divide by R S S unrestrict M E T restricted and unrestricted models into the degrees of freedom and T minus N minus t divide by N plus T minus 2 N minus N minus 2, this is how the degrees of freedom. This is how the F statistic has to be calculated.

So this is ordinary least square that is which we have derived from pooled data analysis. And then this restricted versus unrestricted, then obviously we will get the F statistic. That F statistic has to be statistical significant. if this is the calculated value, if the calculated value is over take the tabulated value with particular level of significant, then that model is somewhat very structure and it can be used for you know prediction or forecasting. So this is how the, similarly, we are we can discuss the importance of the parameter or weightage of parameters of that particular signal. That means, we can observe which particular unit, individual unit is more efficient than the others. Similarly, which time periodic is more effective than others? This is how we have to study the individual impacts. Likewise, we can also discuss the, you know random effect model. So what is this structure of random effect models?

(Refer Slide Time: 40:18)




You see likewise a fixed effect model, so we will proceed random effect model. You see in the case of fixed effect models, we are targeting the impact of cross sectional observation or individual variation or time series variation on intercept. So the moment when we will apply time series variation on individual, you know variation in the intercept, some kind of impact will go to the error terms. We like to know what is the exact variation on error terms. When the impact will go to error term, then how is that setup? So that setup will discuss with the form of the random effect model.

What is the general format of random effect model? the random effect model basically will write like this way, Y i t is equal to alpha plus beta X i t plus, you can say, you can put it here mu t mu i t, let us say mu i t, this is general format of random effect model. However mu i t can be written as U i plus V t plus, you can say W i t, so that means, this is equation 1 and where mu i t can be that means the impact will directly observe from the error term here.

So this particular structure is error for sectional component, this is due to this error is due to individual time series means due to cross sectional cross sectional framework, this is due to cross sectional cross sectional framework, this is due to time series framework time series, it is better to put cross sectional components not framework components, it is due to error due to cross sectional component. This is error due to time series component time series component and this is error due to both error due to due to both ok.

Here the standard assumption is that here the standard assumption is that, for U I, it will be normally distributed with 0 mean and unit variance. For V t, it will be normally distributed with 0 mean and you need variance. Similarly, for W i t, it will normally distributed with 0 mean and you know unit variance. So this is how this structure is you can say you can say observe.

So that means, what we have observed here, in the case of, you know random effect model which is you know one partition of this fixed effect model. because means it is very integrated process, first we observe the pooled data case, where we are just applying the means, we are integrating the variable together cross section, observation to we are clubbing the observations with you know time series and cross sectional. Then we directly apply O L S technique and observe the estimated parameters and over all fitness of the model.

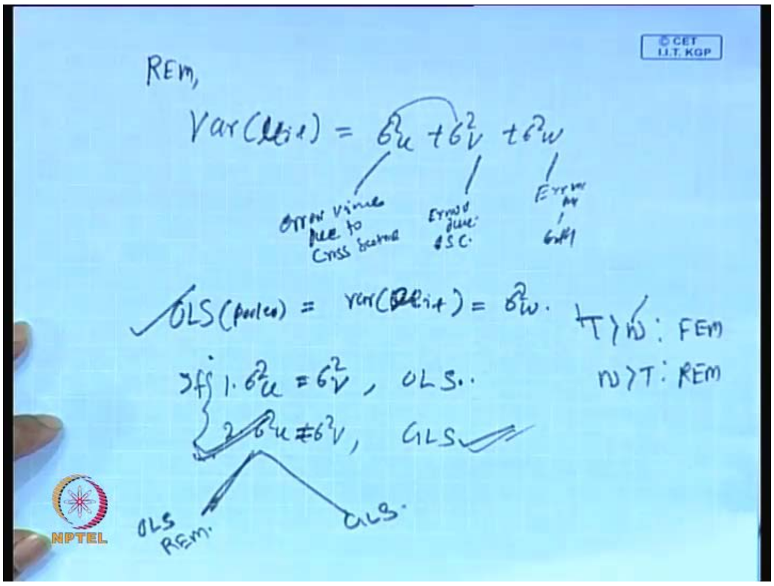
Now, when we will you know integrate cross sectional observation with the time series observation. Then obviously, some kind of problems we will face means, obviously by default there will be some problems that is you know the variation of individual impact cross sectional impact and the variation of time series impact. That we will capture through the, you know fixed effect models and random effect model. When these cross sectional observation and time series observation, we are targeting with the intercept. That means, when we means we are coming to a point that so individual observation and time series time series observation when we will integrate, and then there will be some impact on this entire models.

These model can be means further impact on the entire model. That can be investigated again with respect to the impact on intercept and the impact on the error terms. So, when we will discuss with the impact on intercept, then obviously that particular model is called as a fixed effect models. So when we are targeting the impact of individual variation and time series variation on error term, then that will observe through random effect model, what we have already mentioned here.

So the simple simplest form of these a random effect model is like this, Y i t equal to alpha plus beta X it plus mu i t. so where mu i t equal to U i plus V t plus W i t. so this U i is error due to cross sectional component. This is V t is error due to time series component. Then you know W i t is error due to both. Now on an average, you know the deviation you know deviation effect of time series is randomly represented by V t here.

Now, we have to see means, before we go for this estimation, the condition is that you should be normally distributed with mean 0 and unit variance.  Similarly, V t means 0 and having constant variance. Similarly, W i t normally distributed with mean 0 and constant variance. With this particular setup, for this means in this particular setup, so what we will observe is.

(Refer Slide Time: 46:11)




So the for this particular effect models for this particular random effect model, we like to we are very keen on the total variation of error terms. So this is variation of error terms U i t means total variation, that is not U i t, this is mu i t, this is total error variance is equal to sigma square u plus sigma square v plus sigma square w, so this is error variance, so this is error variance  due to cross sectional component due to cross sectional component; and this is error variance, error variance due to time series component due to time series component; and this is error variance due to both. So this is how you know for the means, if you will go by pooled data analysis say if you go by pooled data analysis, then O LS pooled is simply equal to variance of mu i t variance of mu i t variance of mu t is equal to sigma square w only.

So now, there are two cases here, so there are two cases. If sigma square U is equal to sigma square v if sigma square U equal to sigma square v, then obviously you will use

OLS; if sigma square U equal to sigma square v, then you will go for OLS estimations. If sigma square u not equal to sigma square v, then you have to go by G LS technique, you have to go by G L S technique. So that means, here there are two different cases here. so one case means, total variance of error term is equal to sigma square u sigma square v and sigma square w. So sigma square v is the error variance due to time series component, sigma square is error variance due to cross sectional component and sigma square w is error variance due to both.

So now, so far as technique wise concerned, we will either we can apply O L S technique or we can apply the G L S technique. If the variances of this and this are equal, then we will apply O L S technique and get this get the error variance. And if you know, if the sigma square U and sigma square v are not equal, then you have to go for G L S technique. But you know, what is, if for the case 2, so we have to proceed by two steps again. First you apply O L S on random equation model, you apply OLS on random equation model, and then calculate the R s s to estimate the sample variance and calculate the R s s in a sample variance. Then apply again, G L S technique here, and then you apply G L S technique here, by using this error variance through the random effect model. This is how the structure of random effect model, in the case of panel data setting. So altogether, we have three different sets of panel data modelling. So, one we have discussed pooled data setup, then fixed effect modelling, then random effect modelling.

In the case of pooled data modelling, we are just clubbing the data without storing the impact of time series component and cross sectional component. But in the case of means, once it is not authentic or it is not you know most feasible, because the moment we will club the time series data and cross section data then obviously there are certain impact will go to the entire models. So that impact means time series impact, variation of time series impact and variation of cross sectional impact.

So to that entire time entire you know panel data model, once the impact will go it may be go to error terms that is with respect to individual. That is what we written here with respect to gamma i and it may go to also with respect to, because of the time component that we have represent here delta t. And it can also go to the error term altogether. Now, if it will go to the intercept either in the form of mu Y means gamma i or delta t or both, then obviously it is called as a fixed effect model. And if it will directly go to the error component, then it is called as random effect model.

Obviously, we have we have in fact discuss the entire issue about this three techniques means, how we have to go with pooled data analysis, how we will go with fixed effect modelling and how we will go with random effect modelling. We have also discussed the various estimation procedures of the pooled data technique. And we have also discussed the, you know we have very discussed with respect to fixed effect modelling and also we have discussed with random effect modelling.

So now, you know, in fact you know the fixed effect modelling is more interesting than random effect modelling, because we are observing intercept with respect to individual dimension and you know time series dimension. That is how, because the parameters are too much lengthy. You know, we have just moved, you know gamma i and delta t, so gamma i is for instance, it has lots of integration with domain variable. When we will say gamma i equal to 1, 2, 3, like this, then you apply gamma 1 with unit matrix. So your first component will be one, then rest of it 0. And second, if gamma equal to 2 then you know the factor will be first 0, then second vector will be unit, other items will be 0, similarly we will proceed. This is how the structure all about of means is a well connected with dummy variable technique, that to dummy independent technique.

Similarly, it can be observed through time series issues also, delta t. For t equal to 1, then obviously you multiply delta 1 into unit matrix, where first component will be unit and others will be 0. And similarly, when delta 2, then second unit will be 1 then others will be remaining 0, so accordingly it wills proceeds. So be careful about that issue, when we will add one after another, you know individual units and time series units. Then obviously, your parameters will start increasing and in the same time the sample will get reduce. So, this is one aspect of, one features of fixed effect modelling. So in another case, you have a random effect modelling, where the direct impact will go to the error component. Where the, you know parameters are very less in numbers, parameters are less in numbers.

So in that contest, you must be very means, you must find out the feasibility of this two models. So means, the problem is that whether you have to apply the fixed effect model or random effect model or pooled data model, in fact, there are few tricks I have already mentioned, so it depends upon the various sense of T and N, for instance it is like that way. If T greater than N then it is true is fixed effect models. If N greater than t, then you apply the random effect models. T represents time observation and T represents N represents cross sectional unit.

Now, when t greater than N, then obviously N will dominate. In that case, fixed effect modelling is this is usually thumb rule. If N greater than T, then you have to apply random effect model. but some but, it is not at all, if you will apply both the model and find out model will give the best results with respect to particular problem setup. So with this, you know we mean most of the components we have discussed in the panel data setting. So that to pooled data analysis fixed effect modelling and random effect modelling.

So far as a use is concerned or you know application is concerned, now with respect to particular problem, with respect to sample observation, that to you know individual unit, cross sectional unit, then time series unit, you have to make a decision. Besides, there are so many statistical test are there. So the test will give you an indication which particular model is more effective for that particular problem. So, with means out of all means whether it is you know N greater than T or T greater than N or you can say parameters high parameters are less or time periods is high time periods are is less.

So in all these cases, it is okay, but in same times, the accurate investigation or authentic investigation will be come forward. If we will apply statistical test like you know, lots of L M test, F test, and etcetera are there. You have to apply all these test and you find out the reliability of the model. Then you can choose which model is best for your analysis and that to with your particular problem. With this, we have we have finished this particular panel data modelling

Thank you very much have a nice day.