

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

Lecture - 40
Introduction to Module II

Hello and welcome back to the lecture on Applied Econometrics. We have completed module I. Now we are going to enter into module II. Now module II is going to be an exciting journey, because in module II, what you are going to do is we are actually going to build econometric models. So we have been talking about econometric models, but we really have not built anything yet.

What you have done so far in module I, and let us all do a little bit of recap. We essentially learned different tools and techniques. And we have seen how we can actually, you know sort of do statistical decision making. And we will see going forward in module II, when we build this econometric models, this decision making processes and the tools that we have learned, they are going to be used extensively.

So it is very important that we have a firm grip on what we have learned so far. Now when we talk about statistical model building or econometric model building, so what we really do? Essentially what we do is we actually sort of build our models from data, okay. And what we do with data, we try to see the relatedness among different variables, right?

So we have like, independent variable, a bunch of, you know like, sort of bunch of independent variable and a dependent variable and try to see the relationship among the dependent and independent variables. So it basically explains some phenomena, okay. And this phenomena we actually get from say, economists who got you know the journal papers in economics.

For example, we can see that you know how mother's participation in labor force actually influences child's education let us say. So that could be an important economic question. And we sort of build a model around that. And how do we do

that? Again, we basically get data. And then we try to see how the variables are related, how we can actually explain that.

Now when I say this one thing we need to remember, and that is this model building, this econometric model building which we are going to cover in module II, we still remain in the paradigm of correlational thinking. We are not yet there in the causal thinking part. So we still think correlationaly. So basically, we will derive or sort of our inferences are derived from the corelatedness among the variables.

So we will just try to see that in this module. Now essentially, in this module, how we are actually you know sort of going to progress will you know there is a there are, there is a checklist and in that checklist, we will see how you know like, we do a diagnosis of the models, okay. So when you build a model, how we actually do is we essentially need to ensure certain things are taken care of, okay.

And what the certain things are? We will see that there are assumptions, there are many assumptions. And we know that whenever we are going to build an econometric model or any model per se, we need to have some assumptions, because model is always representing the reality in abstract, right? And when the reality is represented in abstract, we have to have assumptions, right? We have to have assumptions.

And as an econometrician, we need to know what these assumptions are. And when we have these assumptions, how they are limiting us or you know, so how they are limiting us. And that way, we sort of get a sense of the models, okay. So we know this model is giving these assumptions. So what are the assumptions? Let us say we will talk about linearity in parameter, okay. Just in the next few lectures, we are going to see that.

So what it really mean and why really it matters, okay? Or for that matter, let us say we will talk about, or there should not be any sort of pattern in the error term. Why there should not be any pattern in the error term okay, and how do I really detect that? And if I can detect that, how do I really address that? Alright. So you know there could be a question.

Let us say we do not have a variable, which we think is important, but we do not have data. For example, you know talent. Talent of an individual and his wage. So we really do not have any data on talent. So how do I really you know address this problem when I do not have a data of variable. We do not say that we can 100% address a problem, but there are ways around how we can actually look into it.

Or there could be like variables, independent variables, which are related to each other. Now when the independent variables are related to each other you know we ideally do not want them to be highly related when they are actually you know explaining a dependent variable. But when they are related, how do I actually look into that? How do I really solve that, okay?

So these are the things that are assumptions. When you talk about the assumptions we will actually address this kind of problems that we come across. And this diagnosis we need to do to see how robust our model is. So we will learn that diagnosis part in this journey. Now what are the terms that we can come across? We will come across terms like model specification, functional specification.

We will talk about multicollinearity. We will talk about heteroskedasticity, autocorrelation, and so forth. In this module, we are also going to talk in detail about the dummy variable. So that is going to be very important when we you know talk about model building and when it deal with the categorical data, in statistics, or in the kind of problem which address in the domain of economics, lot of data are actually categorical data, okay. So we will just see that.

We also said that, we will talk a little bit about the relatedness between machine learning and econometrics. And actually, they are quite related. And, you know we do not have enough literature right now that explains the relatedness between these two disciplines. And I personally would love to, you know sort of explain that.

But given this, you know this limitation of the course, we will briefly touch upon that, and we will touch this topic in the context of when you explain least square method and stochastic gradient descent, okay. So we will see how these two discipline can

actually talk to each other. So we sort of touch upon that. Finally, we are going to follow Christopher Doggett's textbook and Gujarati's textbook, for this module.

And just the last bit of advice, when we do this module is that, keep in mind the models. That when we talk about the models, we always go back to the literature. So always the models, the idea of a model comes from, say existing literature or a phenomena that you have seen, right? Let us say just the fact that education and wage, we know that they are related.

And there is a literature, how education is actually, you know help you to build the human capital, and that actually helps you to get higher wage. So that way, we will sort of always whenever you build a model, it is not a one stop process, it is a continuous process. We always, you know having a conversation with the literature or you know existing, you know the research that is being conducted.

So we always have, we always sort of, you know look at our model in the context of what is being done or what is done already or what is being done right now. So that way, we sort of, you know understand the econometric modeling part. Throughout this module, we are going to use small examples, and these examples, we will numerical example.

So for to explain the small concepts, we will be using small data sets, and little bit of coding to actually explain this concepts. So with this, I sort of end this introduction to module II and looking forward to an exciting journey in module II. Thank you.