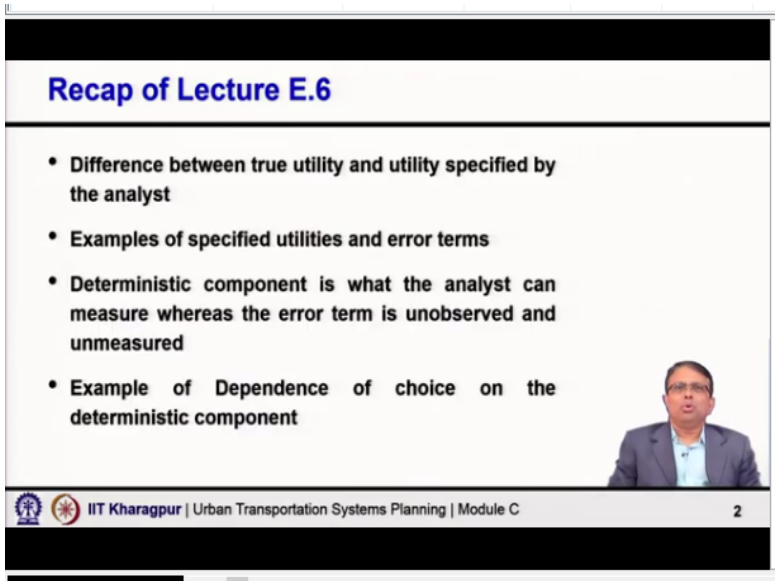
**Urban Transportation Systems Planning**
**Prof. Bhargab Maitra**
**Department of Civil Engineering**
**Indian Institute of Technology-Kharagpur**

**Lecture-37**
**Logit Choice Models-I**

Welcome to module E, lecture 7. In this lecture, we shall talk about logit choice model. Of course, logit choice models also do include or logit choice models are also included under desegregate choice models, but specifically, since we are going to discuss about logit choice models, I made the title as logit choice models. But logit choice models are also under desegregate choice models, particularly desegregate probabilistic choice models.

**(Refer Slide Time: 00:52)**



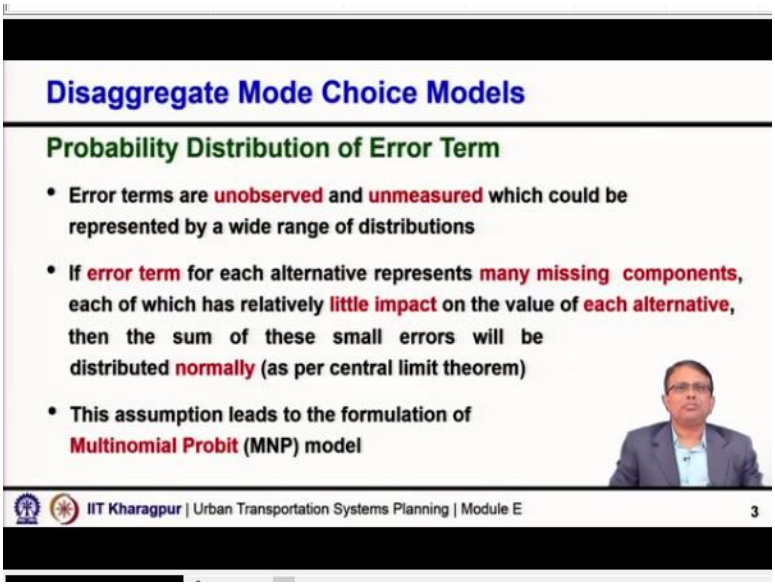

What we discussed in lecture 6 a quick recap to that, we discussed about the difference between the true utility and the utility specified by the analyst, we said that analysts have a certain limitations. So, whatever we take in the model, that utility may not be the true utility, there could be several reasons, which we explained. So, there will be a difference between the true utility and the utility specified by the analyst.

We gave you several examples to show that how the specific utilities and the true utility could be different and what are the error terms in corresponding examples? And then we say then the true utility is deterministic component of the utility plus an error term which accounts for all this missing components and this error term is an observed and unmeasured and follow certain random distribution.

And is the random variable of course and what the analysts can consider is the deterministic component of the utility. And then, we explained to you with an example or rather a set of examples, that the dependents have explained to you through a set of examples that dependence of choice on the deterministic component, the deterministic component of utility for an alternative increases.

So, the probability of choosing that alternative will increase, the deterministic component of an utility decreases. So, there will be a reduction in the probability of using that alternative. So, that dependency we explained to you with a set of examples. Now, with this background, let us continue our discussion. And we are now going to be more focused on logit models. As I said, logit models come under desegregate models and desegregate particularly desegregate probabilistic choice models.

**(Refer Slide Time: 03:15)**



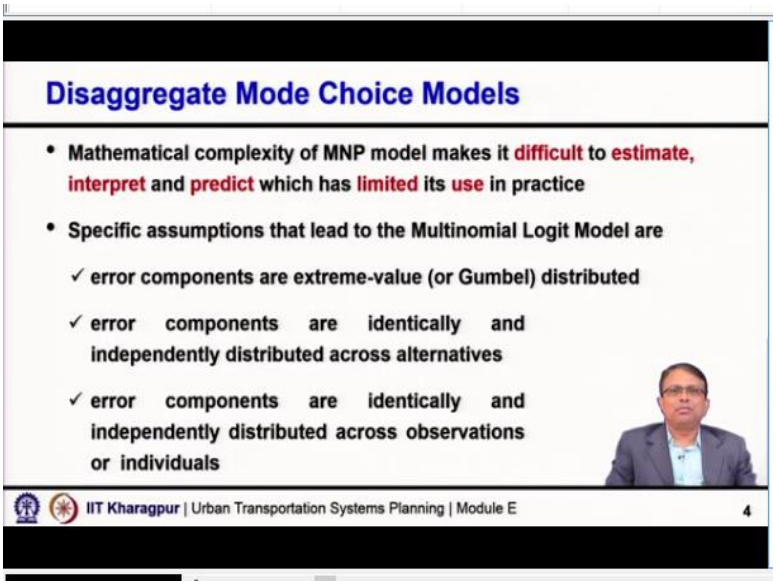

Now, let us go back to this error term. There are many things, but a few things quickly we shall try to discuss. Now, this error term is what is said unknown variable. So, it follows certain distribution. Now, what distribution we are assuming, depending on that, the type of model will be decided. So, first of all the error terms are unobserved and unmeasured, which could be represented by a wide range of distributions.

So, if we consider that this error term of each alternative represents many missing components, not one but many missing components and each of which has relatively little impact on the value of that each alternative. Then as per the central limit theorem, we can say that sum of all the small errors will follow a normal distribution. So, one possibility is of

course, to assume that this error term, this random variable follows normal distribution and normal distribution is very common.

So, anything commonly to be assumed in all such cases, it is basically the normal distribution and if we assume that this is error is normally distributed and it follows the normal distribution the variable, then it will lead to the development of multinomial probit model MNP. So, one thing from this slide if the error term if we assume that it follows normal distribution, the random variable that is the right term then the model will be basically multinomial probit even.

**(Refer Slide Time: 05:37)**




But MNP is not so popularly used in real life or in practice, because the mathematical complexity of MNP makes it difficult to estimate interpret and predict which has limited the use in practice. So, MNP models it is not that MNP models are not developed I have also developed in various works MNP models have used it, there are some specific contexts where MNP model performs better, but generally not so popular because of its complexity in their estimation.

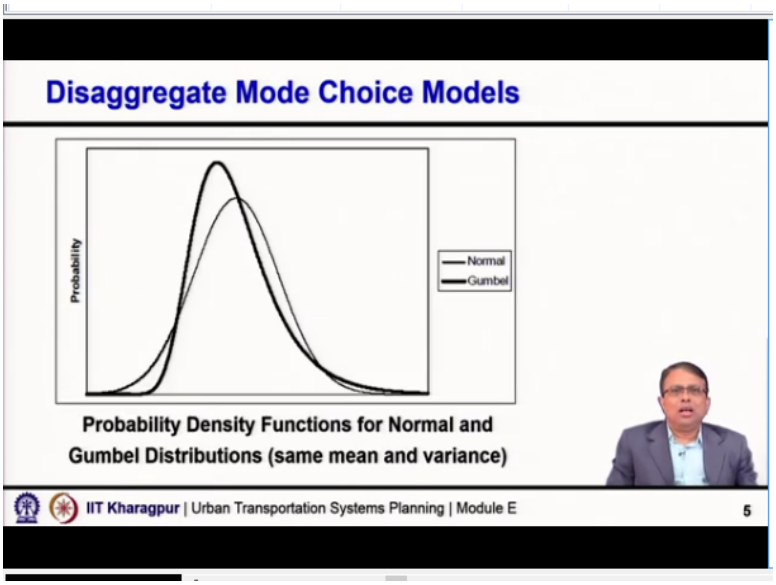
So, what is that other alternative? The other alternative is if we consider that this error term or the random variable is extreme value or dumbbell distributed then this formulation will be multinomial logit MNL. So, 2 types of logit model, multiple types of logit model are there, in this one we shall discuss binomial logit, multinomial logit and also to our side a little bit introduction on the nested logit model

So, but that fundamentally the difference between logit and probit is due to the assumption of the distribution of the error term. So, specific assumptions which have led to the development of multinomial logit models are first the error components are assumed to be extreme value or gumbel distributed, second error components are assumed to be identically and independently distributed across alternatives.

Third error components are assumed to be identically and independently distributed across observations or across individuals, observations come from individuals in this case, it is the BBL data, that we collect and we try to analyze. So, we can think that this is the error components are identically and independently distributed across observations or individuals. So, across observations and or individuals also they are identically and independently distributed.

Across alternatives also identically and independently distributed and the distribution follows extreme value or gumbel distribution that led to the development of multinomial logit or logit category of model, further you will see it could be binomial logit, multinomial logit, nested logit, further other types of developments are there which we will not discuss here in details, but a very popular widely used model at least I have used that formulation in several works, several works not one work but several works. That is the random parameter logit model or mixed logit model but which we will not discuss in this course.

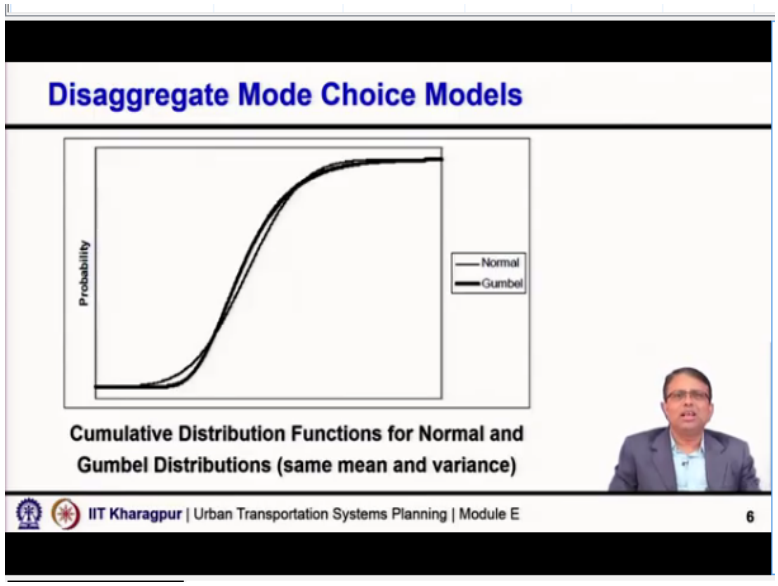
**(Refer Slide Time: 09:17)**




So, now a quick look at how the probability density functions look like for normal distribution and gumbel distribution? You can see the thin line is representing the normal

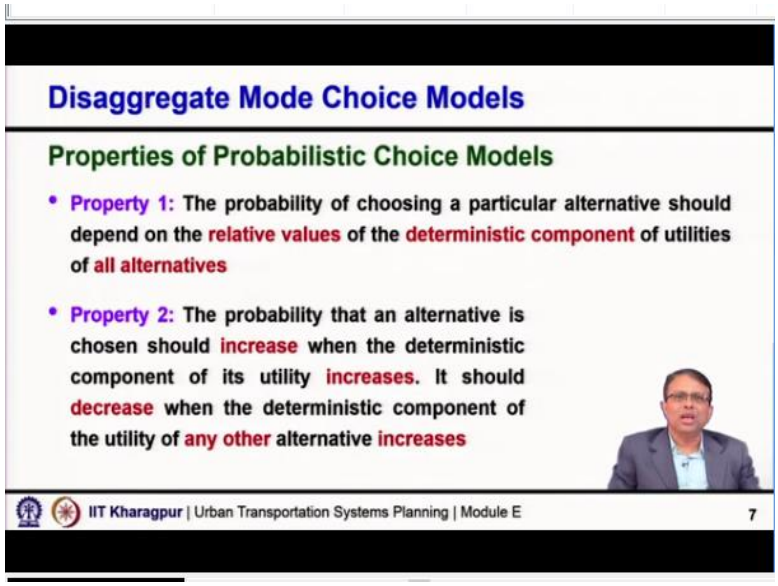
distribution and the thick line is indicating the gumball distribution. So, how the PDF or the probability density function looks like.

**(Refer Slide Time: 09:42)**




Similarly I have quickly shown you that how the CDF or the cumulative distribution function looks like for normal and gumball distributions? They are not very different, there are definitely a lot of similarities and that is why people make it I do not consider it normally distributed, but I considered the distribution as gumbel distribution and then make it computationally a little easier for estimation and application purpose.

**(Refer Slide Time: 10:17)**




So, I am not going into further details because there is a little bit limited scope for touching on different topics. Let us come to the properties of probabilistic choice model. When we said you that is basically the analyst limitation and not only the limitations remember that error

terms include what, sometimes it may be that it is the limitation as analysts, that I know that something influence the choice behaviour.

I know that I should include that but I cannot include maybe because the data is not available, that is one component which is taken care by the error term. Second is I know that it influence but I do not want to include them because they cannot predict, I said the health impact, somebody is not feeling one day, somebody has got some urgent work on one day.

Even if suppose while developing the model I collect the data meticulously from each and every person and try to say look at the calibration part I considered all such kind of influences as relevant variables but and to make my model very good, but how to predict it? Can I predict it and as I said that repeatedly I am saying the model you develop not for the busier you want to apply it in the future.

So, any variable what you are using in a model you should be able to predict those variables in the future so anything if I cannot predict and if I cannot predict accurately or reasonably accurately always there will be some certainty that even in all any variable you consider in a model it will still be approximate, but at least approximately we should be able to predict or in a realistic manner we should be able to predict.

So, if not then we do not consider. The third is that I even do not know that something influenced the choice decision. So, I there was no scope for me to even consider that. So, all such kind of things that taken care by the error term and the distribution of error, what distribution we assume for the error term, that will determine whether it will be legit or it will be a probit.

Now let us with that one let us look quickly at the properties of probabilistic choice model. There are 4 properties of that probabilistic choice model, desirable property I should say one, the probability of choosing a particular alternative should depend on the relative values carefully observed relative values of deterministic component of utilities of all alternatives, utility of an alternative depends on the values of that alternative only, it does not depend.

That utility does not depend on what is the value of travel time by bus, but what we are saying the probability of choosing car or choosing a particular alternative should depend on

the relative values of the deterministic component of utilities, deterministic component of utilities is something which we can quantify which we can measure of all alternatives, what is the probability that people will use car will not only depend on the deterministic component of utility of car.

But also will depend on the deterministic of utility of not only drive alone but also carpool, also bus or in a real situation if you have got also metro operating then metro will also come, all alternatives so we have to consider. Property 2 the probability that an alternative is chosen should increase when the deterministic component of its utility increases. This in a way we have shown you in our previous lecture towards the end we took a combination of examples to tell you that how these parking charges imposed for the drive alone?

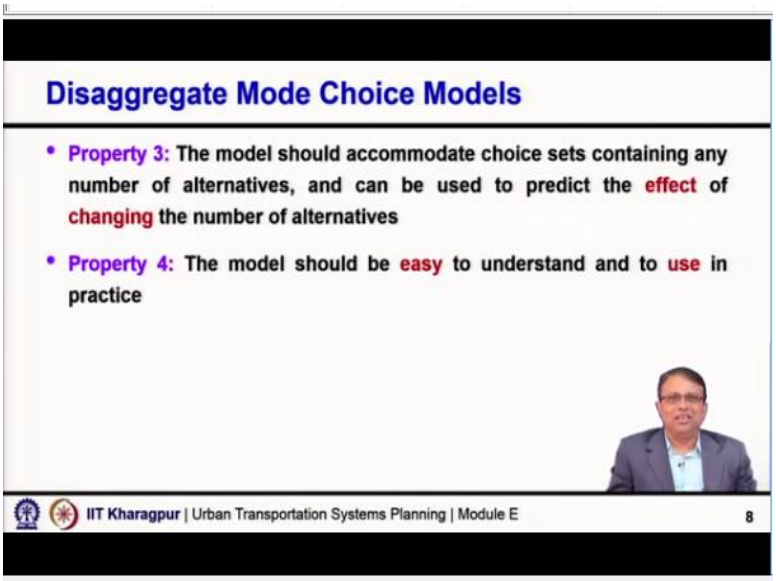
How the parking charges increases and because of the disutility increases associated with the drive alone or utility decrease and utility decreases, so you have say when even the in real sense the probability of using car also decreases because the utility decreases. So, same is true in the other way. So, if you think in a reverse way that the parking charges was 30 rupees, now it becomes 15 rupees, now it becomes free of cost what will happen?

The utility will increase, so obviously the results will be in the reverse direction, you will find the probability of using car is increasing as the utility is increasing. So, utility increases the probability of choosing that alternative increases. Utility decreases the probability of choosing that alternative decreases, but also the other part because what we say relative values have deterministic component of utilities of all alternatives. So, there is an interdependence.

So, what will happen the probability of choosing an alternative will also decrease when the deterministic component of the utility of another competing alternative increases. So, let us say A B and C, so utility A increases probability of choosing A will increase, utility of A decreases probability of choosing A will also decrease, but if the utility of B or C increases then also probability of using A will decrease, that is what I say that it should also decrease for a particular alternative say A.

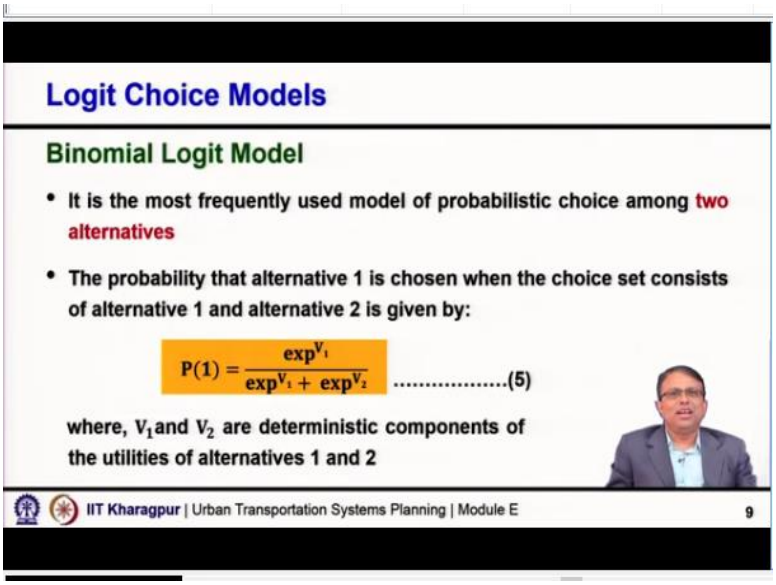
If the deterministic component of utility of any other alternative say B increases. So, if B increases the probability of using B will increase as for the first statement. So, obviously probability of using A will reduce fine.

**(Refer Slide Time: 17:13)**




Third the model should accommodate choice sets containing any number of alternatives, I should be able to handle, my model should be able to handle 2 alternative, 3 alternative, 4 any number of alternatives and also can be used to predict the effect of changing the number of alternatives. Forth the model should be easy to understand and use in practice. Ultimately practitioners will use these models. So, you should make the use model should be such that one can easily understand and use it in practice. So, all these 4 properties are desirable properties of probabilistic choice modeling.

**(Refer Slide Time: 17:59)**

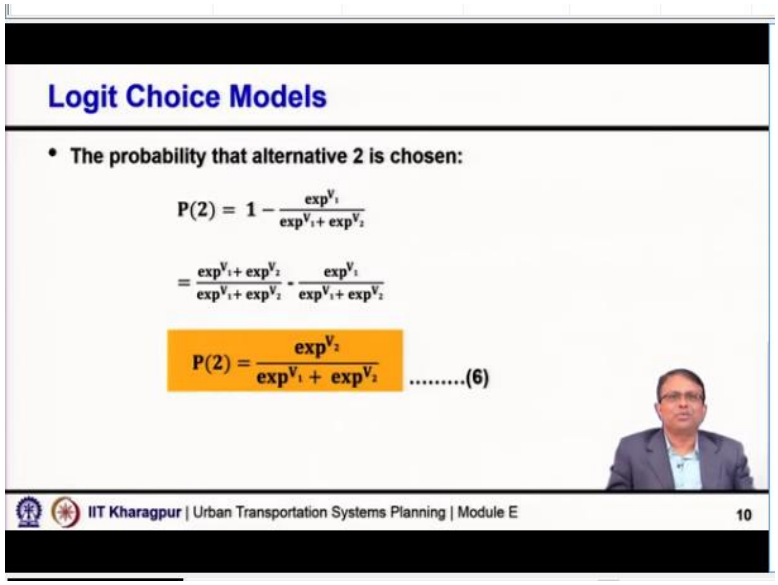


Now let us go to the binomial logit model. The word binomial means what it immediately tells you the number of alternatives 2, there are only 2 alternatives. So, it is frequently used model of probabilistic choice among 2 alternatives and as per binomial logit model what is the probability of using alternative 1 = e to the power V 1, V 1 is the utility deterministic component of utility.

For alternative one divided by e to the power V 1 + e to the power V 2. So, numerator includes exponential of the utility of that alternative deterministic component obviously and the denominator includes e to the power V 1 + e to the power V 2. So, this will give you probability of using alternative 1 that is the very fundamental formulation.

Now 2 things will be of interest here, one is to understand we are saying exponential of utility. So, we should be able to feel that how the exponential value, exponential function changes as the utility changes? So, if the value V changes how e to the power V is changing? Second how this change in V 1 and V 2 will reflect the probability? How the probability will change because of the change in V 1 and V 2. We shall slowly try to understand those.

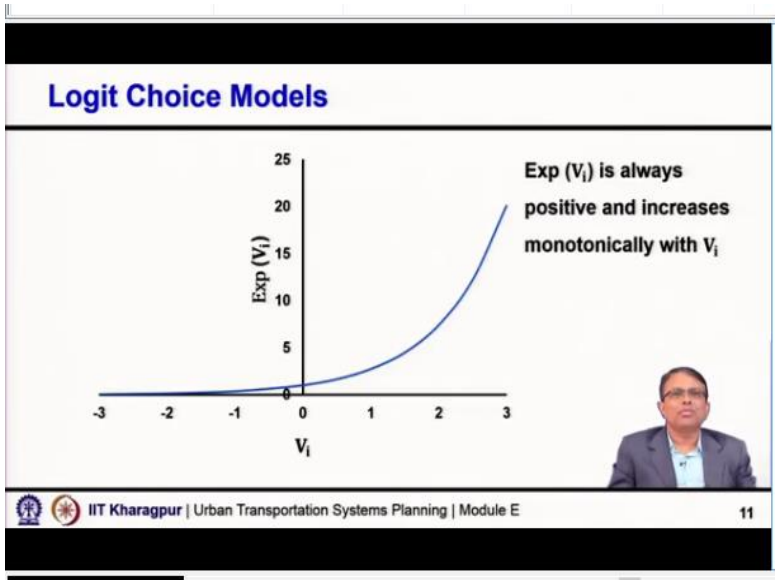
**(Refer Slide Time: 20:08)**




Before that, if I have to say what is the probability of them choosing alternative 2, obviously, we will say there are only 2 alternatives. So, probability of choosing both alternatives should be either this or the other, the total probability of making a choice should be 1. So, probability of choosing alternative 2 is 1 minus probability of choosing 1 that is what we have done here.

And if you will see that means probability of choosing alternative 2 is e to the power V 2 instead of V 1 in the first case, because that was alternative 1. So, use e to the power V 1, here it is for 2 see it is e to the power V 2 and the denominator remains same e to the power V 1 + e to the power V 2. That is what we get as well.

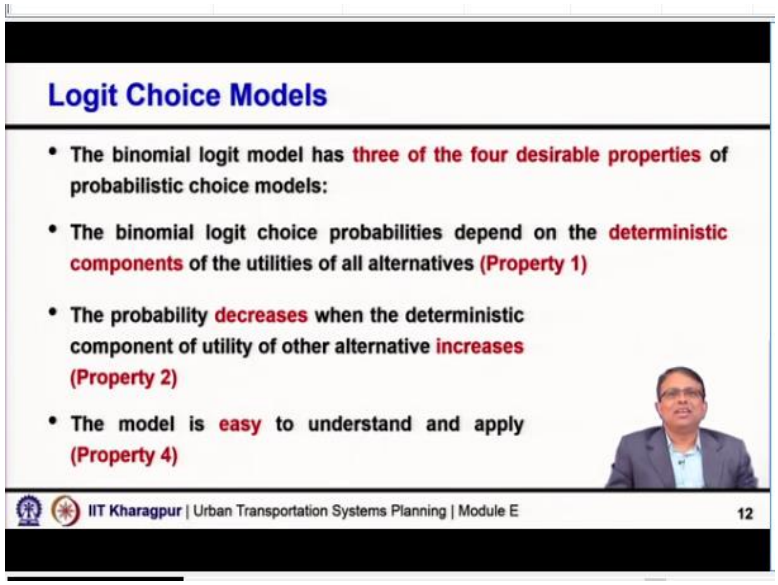
**(Refer Slide Time: 21:05)**




Now, as I said, I have plotted this approximately this is of course approximate, to give you a feel, how the exponential function looks like, I am saying if my V is changing from minus to plus or when the value of V is increasing, we say exponential function, how really the e to the power V changes, you can see that as values are increasing, there is a significantly it is the exponential value is increasing and how at what rate and how the changes happen?

That is very interesting. So, it is just for you to feel that when you were saying e to the power some utility, taking utility and taking e to the power utility, it makes a lot of difference, this is the exponential function.

**(Refer Slide Time: 22:12)**




The second is binomial logit model has 3 of the 4 desirable properties; we said about 4 desirable properties just a few minutes back. Now, you see that there are 3 properties which it is able to satisfy. First, the binomial logit choice probabilities depends on the deterministic components of utilities of alternative yes, in the equation, you have e to the power V 1, also you have e to the power V 1 as well as V 2 in the denominator.

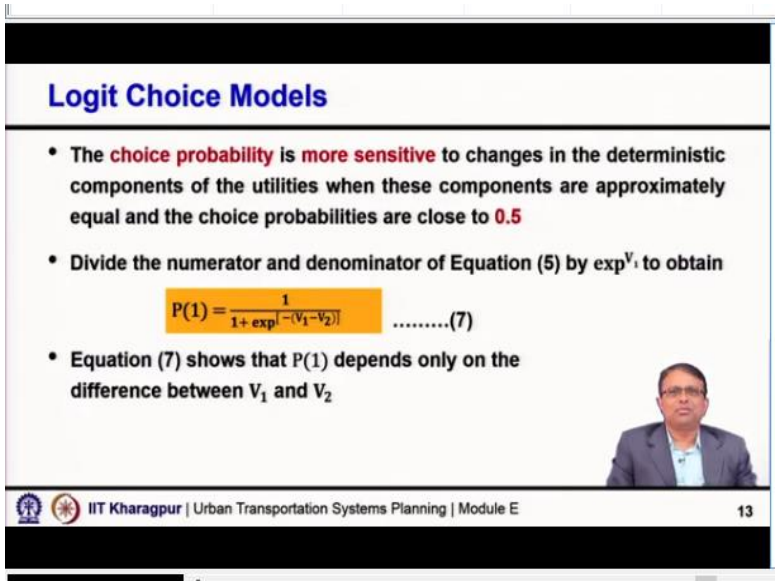
And the numerator e to the power V 1 or e to the power V 2, depending on what probability we are trying to find out. So, but probability of choosing 1 of these 2 alternatives depends on the utility of both alternatives in the equation. So, the first property is satisfied, the probability decreases when the deterministic component of the utility of other alternative increases go back, simply see this equation.

Suppose V 1 increases, so probability of choosing that alternative increases. So, utility of alternative 1 increase naturally, exponential V 1 will increase. So, overall, the probability of choosing 1 will increase, but if the other alternative, if V 2 increases, what will happen? V 1 does not change V 2 change, when V 2 is increasing, what will happen? Probability of choosing 1 will decrease.

So, the second property is also satisfied. Third is not satisfied, it should be able to take any number of alternatives, only it can take 2 alternatives, that is the third property is not satisfied, but the fourth property is satisfied I am sure if I could explain you correctly, it is not difficult to understand and difficult to apply as well. So, you can easily understand you can calculate the deterministic component of utility given that utility equation and you can apply

this simple model as I have shown to get the probability. So it is property 1, property 2, property 4 as satisfied, property 3 is not satisfied because it can handle only 2 alternatives.

**(Refer Slide Time: 24:54)**



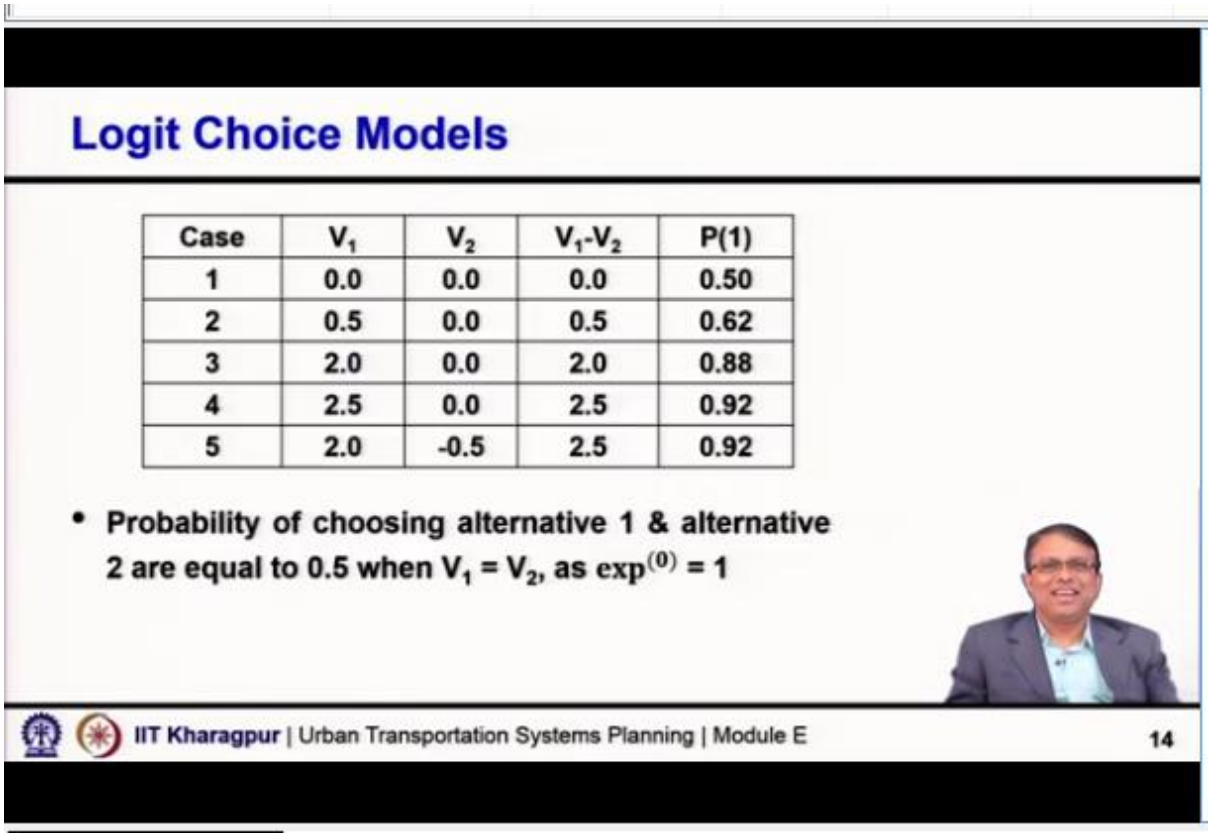

Now, what we can observe here the choice probability is more sensitive to changes in the deterministic component of utilities, when these components are approximately equal, what will happen e to the power V 1 divided by e to the power V 1 + e to the power V 2. So, if my V 1 and V 2 both are same, then what is the probability? Probability of choosing each alternative is 0.5.

Now, interestingly, at that level, if any change in utility happens either for alternative 1 or alternative 2, then the change will be very significant in the probability. But, as you go far away from this midpoint where both are equal as you go far away, the change in the probability due to change in the utility is going to be smaller and smaller. Now, one more interesting thing, just simply take the equation this basic equation e to the power V 1 divided by e to the power V 1 + e to the power V 2.

So, if I divide both the numerator and the denominator by e to the power V 1 what I get, I get then 1 by 1 + e to the power take the minus outside V 1 – V 2, this is very important. And you must remember this always, what it is saying generally because it is something in numerator, something in denominator, the probabilities the ratio of something unknowingly, we tend to think it is the ratio of V 1 and V 2, that is controlling the probability which is wrong.

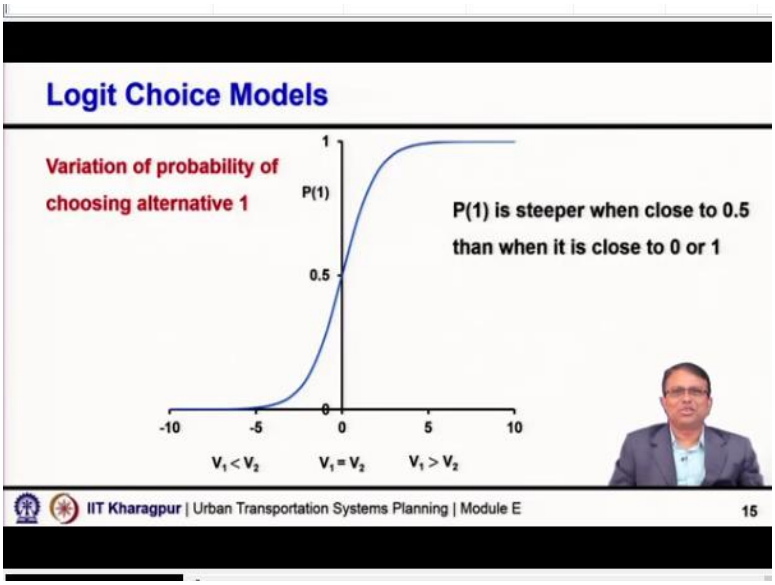
You see here, it is the difference in the utility value that is controlling the probability not the ratio. It is no where the V 1 by V 2 comes into picture. It is V 1 – V 2 that is controlling the probability.

**(Refer Slide Time: 27:21)**




So, let us take when V 1 and V 2 is equal 0 0 and say V 1 – V 2 is 0, then probability of using each alternative is 1.5. Both utilities are same, logical, very logical, utilities have 2 alternatives are same what you expect 50% people will use this and the 50% people will use the other alternative. That is what is quite logical. Now, as it changes, now, the difference become 0.5, difference become 2, 2.5 and how the probability is changing? So, now it is the time for us to see how the probability changes with the difference in V 1 and V 2. Let us do that.

**(Refer Slide Time: 28:11)**

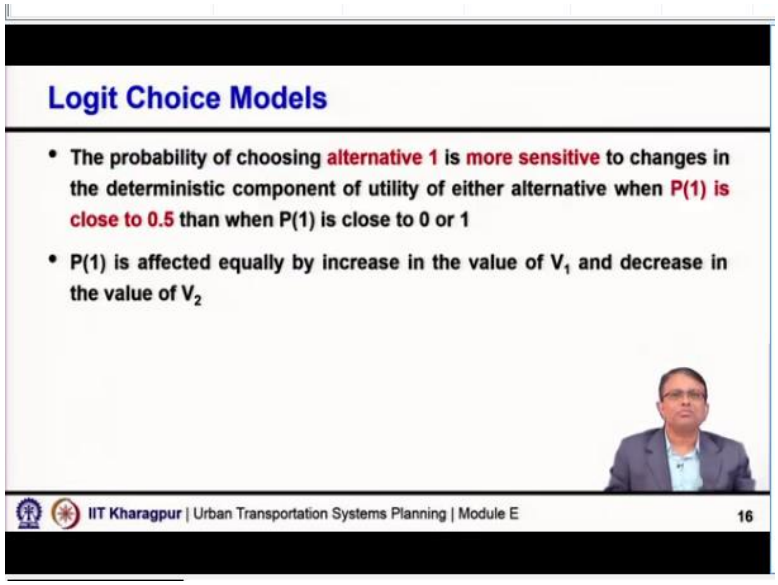



That is what I have shown, see it is look like a S graph, this is a very interesting and very realistic thing you will find in many contexts, you will find this curve is valid. This is one of the really most beautiful way of representing the things, very realistic way of representing the thing what will happen? So, take the simple thing you have bus and train equal fare, equal all other travel time, all attributes for the moment, let us consider them as identical you only same fare.

So, what will happen 50% 50% 05 0.5, now you little bit maybe make them little cheaper, what will happen, a lot of people will try to ship to them. Now, if you keep on increasing the fare progressively, you will find lesser and lesser people are shifting, for the first change how many people have shifted? Next increment of the same amount not so many people will ship, little less.

And as you went so you make a toll road increase the toll, make the toll very high, still it will not be 0. Still some people will use it, maybe the probability will come down, but it will not become 0. Also it will not become more 1 because there are other factors, there are error terms, there are other considerations. So, how the maximum change will happen around 0.5 probability or when the utility of 2 alternatives are equal. And as you go away the impact will become lesser and lesser.

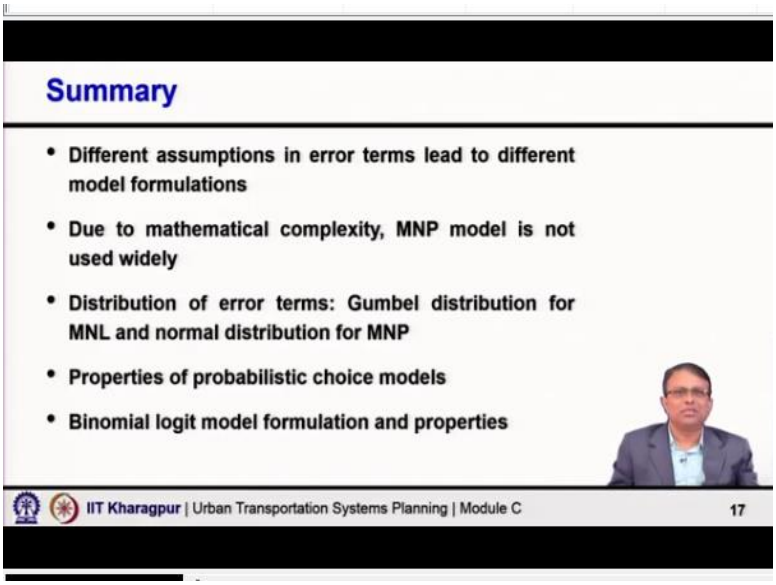
**(Refer Slide Time: 30:27)**




Now the probability of choosing as I said alternative 1 is more sensitive to the change in the deterministic component of utility of either alternative when P 1 is close to 0.5 and then than when P 1 is close to 0 or 1 obviously very insignificant change will happen when it is close to 0 or close to 1, maximum will happen when it is 0.5 and I explained it why it is so, why it matches with the real behaviour, it very well matches with that.

V 1 is affected equally by the increase in the value of B 1 and decrease in the value of if I increase V 1 or if I decrease V 2 the effect is going to the same why? Because it is the difference V 1 – V 2, the difference will matter, how much is the difference? The difference may come because of the change in V 1 the difference may also come because of the change in V 2. it does not depend on where the changes happening. It is basically the relative value that means the basic difference in V 1 – V 2 that is controlling.

**(Refer Slide Time: 31:37)**




So, with this what we discussed in this lecture, we discussed about various assumptions in error term which led to different model formulation. Normal distribution led to the development of product, gumbel distribution led to the development of logit, we said that due to mathematical complexity MNP is not so popular or not so used widely. We also said that what are the desirable properties of the probabilistic choice model.

And then we showed you what is the formulation of binomial logit model, what are its properties and we say that binomial logit models satisfy 3 of the 4 desirable properties. So, we close it here, we shall continue this discussion in the next class and talk about multinomial logit model, with this I close this lecture, thank you so much.