Probability Method in Civil Engineering

Prof. Dr. RajibMaity

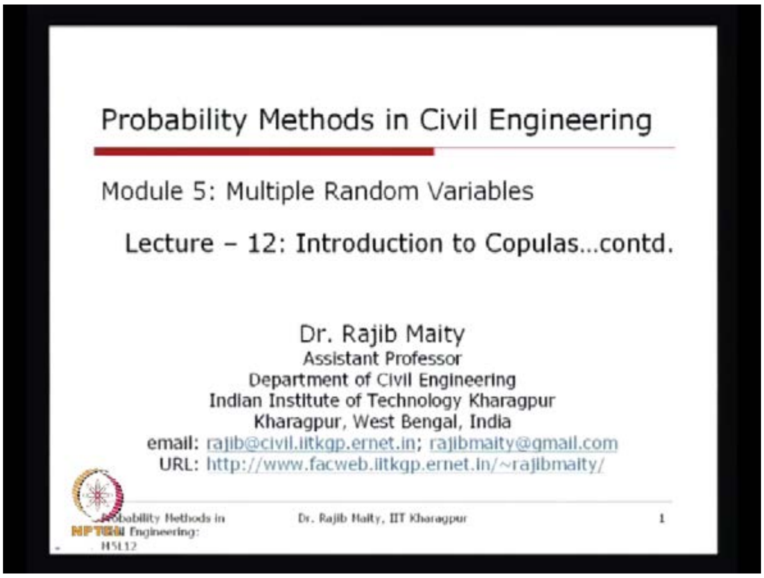
Indian Institute of Technology, Kharagpur

Lecture No. # 30

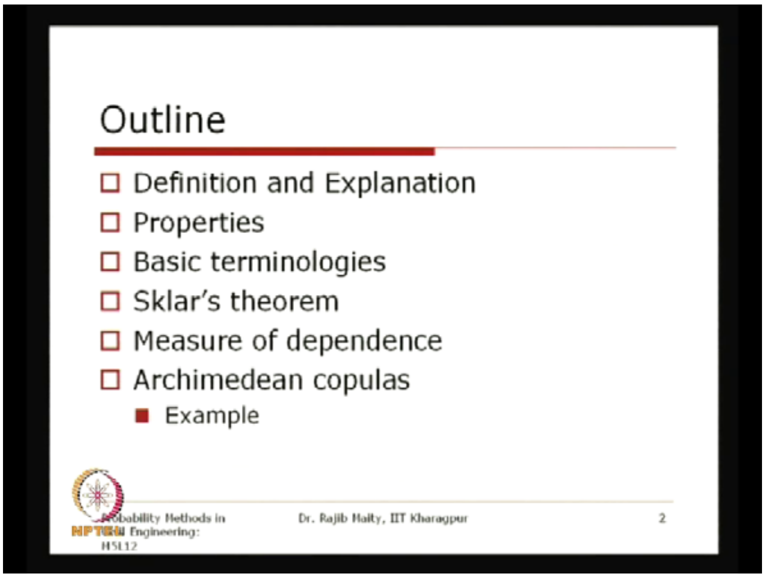
Introduction to Copulas (Contd.)

Hello and welcome to this lecture onintroduction to copulas. We started thissectionthis copula in the last class andwe covered up to the dependence, and today's class also we will cover and we will continue the same topic. In this lecture, we will start with their dependence properties, and after that, we willdiscussspecifically about one specific group of copulas which is so farused and has seenapplications incivil engineering particularly in the hydrology and water resource engineering. So, we will discussthat class of copulas and we will seesome problemsafter that.

(Refer Slide Time: 01:04)



(Refer Slide Time: 01:07)



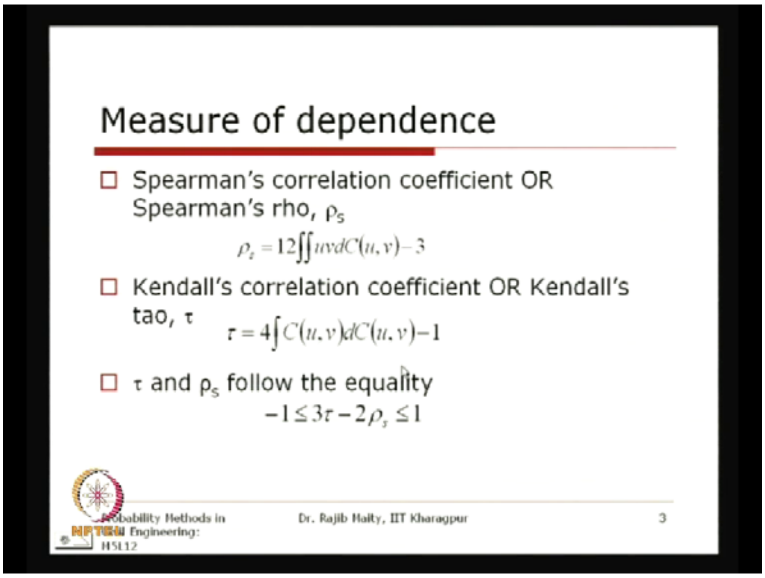
So, ouroutline if you seethat outline of thislecture today thatwe have already have you know that in the last class, we havecovered this definition and explanation of what is this theory of copula, and after that, we havediscussed some of the properties of the copula function that it shouldsatisfy to, to, be a copulafunction, andas you know that when we aretransforming from the one-dimensional to the multidimensionalcases, so there some of this basic terminologies or the meaning of the basic terminologies also should be extended to the higher dimension.

So, regarding that, we have discussed some of, some moreterminologies.For example, the grounded we have discussed, we have discussed what is the age, volume and what is the two increasing, increasing, function. So, those are the properties that should be satisfiedby a function tobefore we can declare that function to be a copula.

And after that, we also discussed about the Sklartheorem, and while discussing this one, we have, we told that this is basically the, basically themost important theorem in this one, where we can use the function copula to, to, develop the joined distributionfrom their marginal distribution.

And we also started the discussion on thismeasure of dependence, and today, basically we will start from this measure of dependencefrom the copula function and also we will see that how to get thosedependence parameter from the samples, and after getting that sample dependence, then we will go and we will see that how those parameters can be used, means from those parameters, how we canobtain the parameter for the copula, copula, function. So, that we will see, and after that, we will seeexample on the,on these copulas.

(Refer Slide Time: 03:04)

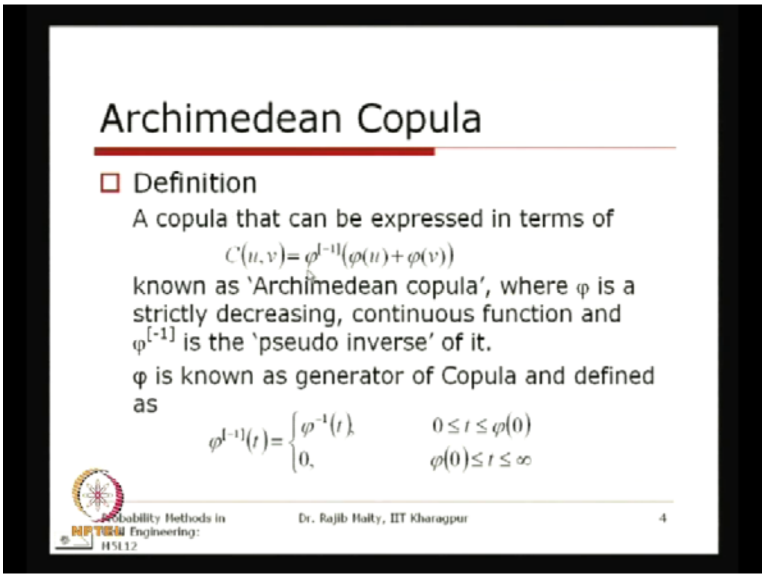


So, as was, was, discussed in the last class that there are two mainmeasure of, measure of dependence. Those arewe will be using in this and these are the expression for those. First one is the spearman's correlation coefficient or the Spearman's rho and this rho s is equals to the twelve times of this double integration with respect to the copula function minus 3. So, this expression is gives you the Spearman's rho, and similarly, the Kendall'scorrelation coefficient or the Kendall's tauwhich is the 4 times the integration of this copula function with respect to thatfunction c minus 1, is gives you the Kendall'stau, andwe have also, we have alsostated that this tau and this rho s followthis relationship means through this inequality; inequality this is. So, minus 1 less than equals to 3 times tau minus 2 times rho s, which is less than plus 1.

So, this 3 times tau minus 2 times rho s is bounded byminus 1 2 o plus 1, but what is more important here is that this is, this the expression that we have explained here. This is in general for all the copulaswhatever maybe the class. Now, you know that even though this we can explain it very, mathematically we can explain it rather easily, but for some of the cases, it will be better if we get some more, some more, what to say, thateasier form of thisfunctions.

So, that can say those are basically the, in the mathematical point of view, those are more easily can be applied and once such case is the Archimedeancopula that will bediscussingafter this. How these expressions means, this expressionwill be changing depending onwhat the copula in hand.

(Refer Slide Time: 05:14)



So, this Archimedean copula, so one special class of thiscopula, the definition of this Archimedean copula is stated through this that a copula that can be expressed in terms of this c u v, this is the copula function, is equals to that phi inverse of phi u plus phi v.

ow, you know that for the bivariate case, there are two parameters, there are two variables - one is the u and v, and now, there is a function which is that phi. So, it should be, it can be, it is the first one is that phi u and the second one is the phi v. So, their functional form of this phi is is same. So, if we add them and thenwe take 1 inverse, this inverse is known as the pseudo inverse will explain in a minute.

So, if we can explain that full, full function in this form, then that class of copula is known as Archimedean copula. Now, this phi, this, this function phi a, this is a strictly increasing continuous function. So, you know that this strictly increasing meansit is, if that, if we take two, two values - one is the e 1 and u 2, and if the u 2 is greater than e e 1, then phi u 2 should be greater than phi e 1. So, that is the strictly increasing andmeans strictly increasing. When we are saying, that time that equality sign also is not there.

So, earlier in, when we are discussing the two increasing, you know that we are, we are, we are discussing, or in case of the one-dimensional that single random variable, variable, while describing the c d f, we are talking, we are explaining that it is non decreasing.

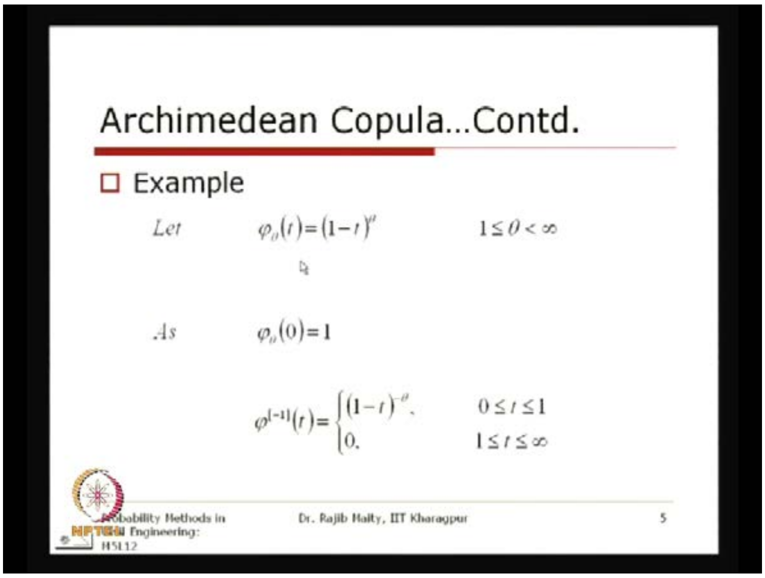
So, the difference between the non-decreasing and strictly increasing is that, non-decreasing meanstheir equality sign is included. So, that if that, the second value is greater, that is, u 2 is greater than the e 1, then if the functional value at u 2 is greater than equal to that the functional value at e 1, then we say it is non decreasing. So, that equality sign is also included.

Now, the strictly increasing means that equality also not there; it has to increase if the attribute increases. So, that is, so, this is the functional property of this phi. This is that strictly increasing and this is a continuous function, and that phi in straight bracket written minus 1. There is a phi inverse; it is known as the pseudo inverse of thatsame function phi, andthis function actually you know that nowonce we can, we can, we can, express that full copula function in this form; that means, here one function, that is, most important is this phi function, and if we know that what is this phi function, then we know what is the copula function also.

So, that is why this phi function is known as the generator of this copula. So, phi is known as generator of the, of the copula, and this, this, pseudo inverseas we are just express that this pseudo inverse of this phi, of this inside whatever we have written that phi u plus phi v. So, this pseudo inverse is explain in this form. So, this pseudo inverse is equals to x inverse. This minus 1 when you are write without the straight bracket, it is the inverse function only.

So, it is the inverse of that function as long as this t isin between 0 and 5 0. So, whatever the function we get, if what is the value at 0 of that function from 0 to that value as long as the still t lies, then we take simple inverse of that function, and beyond that range, that is, from phi 0 to the infinity, it become it should be 0. So, this is actually what is the pseudo inverse means. Now, we will see some of this example of this Archimedeancopula along with that generator phi,then it will bemore clear.

(Refer Slide Time: 09:35)

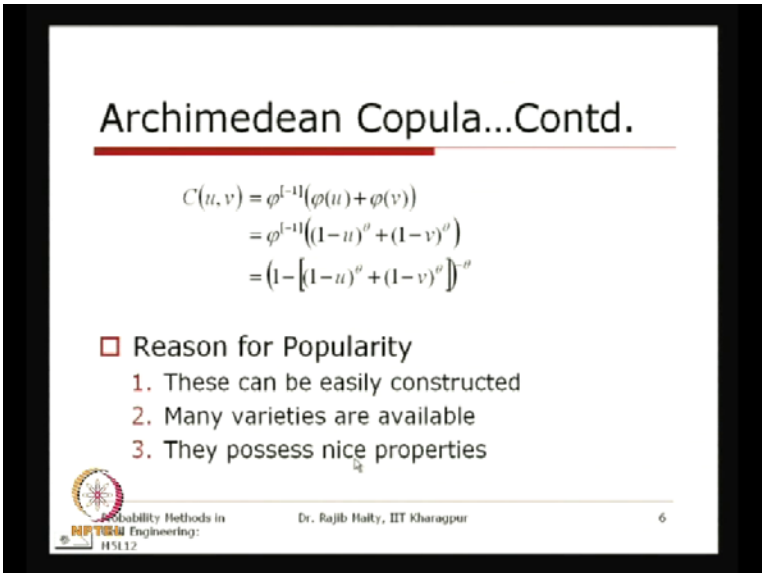


So, so if we just take one example of thatgenerator, suppose that, whatthat phifunction that we are taking, that we are, just now we havestated is this 1 minus theta 1 minus t power theta.Now, theta is one parameter. We will come to this parameterfew minutes later. So, suppose that this functionwith one parameter, theta is explains through this 1 minus t power theta, where this theta can vary from minus 1 to infinity.

Now, if we take thatphi 0 meansphi at t equals to 0, then we can see that this,this, the value of this function should be equals to 1. Now, if we follow that the pseudo inverse of this function, that is, it is 0 to that phi0, that is, phi value at 0, then this is becoming 1. So, we have to take the inverse of that function at 0 to 1, and from 1 to infinity, it should be 0.

Now, if we take the simple inverse of this phi t, then it is 1 by, you know that it is 1 by this phi function. So, this will be 1 minus t power minustheta, and for the restarea, that is, from 1 to infinity, it should be 0. So, this is the complete definition of the pseudo inverse of the functionphi t as explainedthis functional form. Now,this function, if this function changes, obviously our pseudo inverse also will change.

(Refer Slide Time: 11:22)



Now, if I want to know, what is the copulathrough this generator? Now, this c u v what we have to write is that c u v is equals to that pseudo inverse of that function, in, into that phi u plus phi v.Now, this is just we are changing that, that, attribute, that is, it is phi u. So, it is 1 minus u power theta and 1 minus v power theta, and for this, for this form, we have to take it is pseudo inverse.

Now, if you want to take its pseudo inverse, then we have to take that 1 minus t power minus theta. So, this is, so, this is 1 minus whatever there in this full expression, this expression, thenfullthat power minus theta. So, this is explain, and for the rest of the places, it is 0, and this one you know, that is, from this 0 to 1.

So, this functional form what we got is one copula and this is definitely one Archimedean copula. Why it is Archimedean Copula because we can explain it through this one, and for this Archimedean copula, that generator is 1 minus t power theta. Now,now, we will see this. These functions are having some, some, names, depending upon who has first invented this, this, function. Proposed this functiondepending on that, there are many such Archimedean, Archimedean copulas, and you know that for eachfunction, what is the,the, one functionthat is controlling everything is this one - this phi,phi, t which is the generator of this, of this function.

So, depending on that, if the phi t changes, we will geta newArchimedean copula. Now, this Archimedean copula has found itspopularity in different field ofresearch including inthe different areas of Civil Engineering particularlyif you see, so, so far we have seen many application in this hydrology and water resource engineering, and why this Archimedean copula has becomethe reasons?Why these are becomepopular is that it,it is, it can be easily constructed as you have seen that one example just now we have discussed.

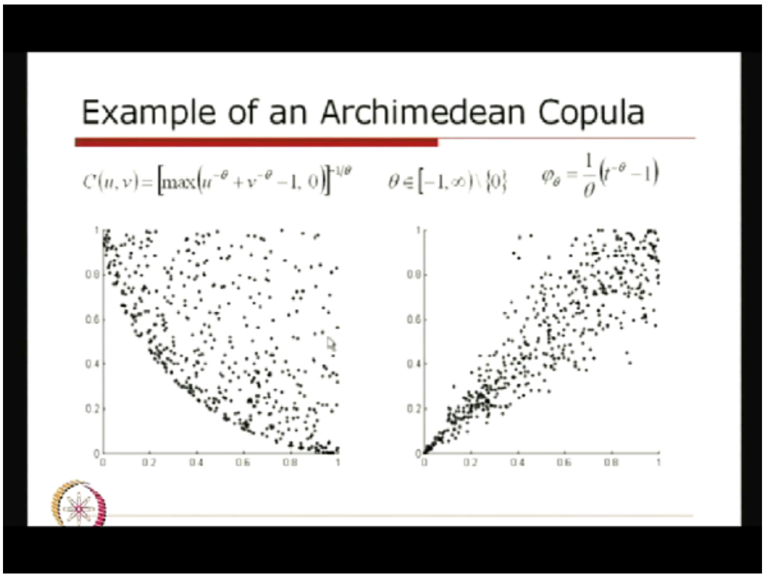
So, these things can be easily constructed, and for this one, the many varieties are available depending on its dependence. So, what is the nature of dependence that it can or the range of dependence that it can, it can capture. For,for the timing I can mention that there are theta functions that I was mentioning. This is actually the measure of the dependence,dependence, between the two random variable that you are consideringhereand there reduce vitiate is u and v.

So, their dependence, so, this is the range of their dependence. Now, so, so as we are we can havemany varieties. So, depending on our requirement whether there is a positive dependence, negative dependence, very narrow range of dependence. So, we can, we can pick up whateverthe, which Archimedean copula we,we, wish totake. So, so, as so, there the lot of varieties is available.

In this class, this Archimedean copula is one class among which different copulas are available. So, we can pick up depending on ourrequirement of theproblemand they possess some nice properties. So, we will just explain, because as I told that this theta is here is the parameter which is the, which is controllingthe dependence between u and v and justfew slides before we have also discussed that Kendall's tau and which is the measure ofdependence.

Now, through this Archimedean copula, we can show, we will just show in a minute that this can be related to each otherthrough a niceexpression and that actually is giving that formthat, how to get, how toobtain that theta from thattau. So, our steps will be nowif we have the data first of all, we have to see that data follows which distribution, and from the sample data, we can, we can alsocalculate that what is their Kendall's tau orthose measures measure of dependence, and from that tau, we have to determine that what is the parameter for the copula that we have, that we have taken, and if we know that copula, then the full form of this join distribution is known to us.

(Refer Slide Time: 16:01)



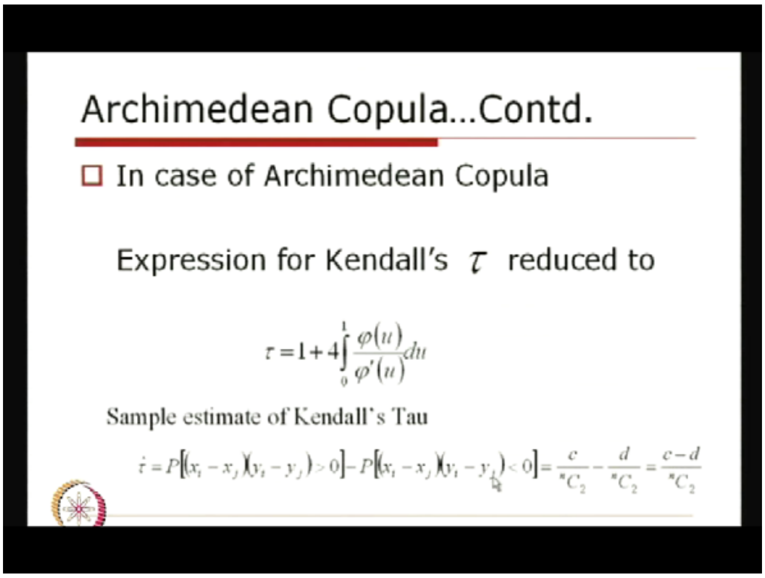
So, this is, this is one, this could be the completestep, and depending on this theta as I was telling that, theta it actually depends onwhat value of thistheta, that is, the dependence measure what value of the theta you wish to take, and depending on that, it canchange a lot. Say for example, here another example of this copula function is shownhere, which is the maximum of u power minus theta plus v power minus theta minus 1, 0. So, maximum of these 2 power 1 powerminus 1 by theta, and here, this theta, that is, the dependence parameter which belongs to this minus 1 2 plus infinity excluding 0. This is the meaning of this full form. This is the dependence can range from minus 1 to plus infinitythat 0 value is excluded in his range.

And if as it is an Archimedean copula, you know that there should be one generator. So, this phi function obviously that t is the attribute, which isnot shown here just for thissimplicity. It is phi theta;theta is the parameter of this dependence. So, this generator of this copula is that 1 by theta into t power minus theta minus 1. Now, here, the two figure that you can see.

So, this we can generate.I will just explain how we can generate this values; this is the combination of these u and v. So, scatter plot between u and v so that this x axis is your u and y axis is your v. Now, you can, we can assume some thetajust to see that how their, how theirscatter plot look like so that we can assess some kind of dependence.

So, this is for this point eight or so, this theta value, whereas this one is for, if we change the theta 2 say point, sorry, if we change the theta 2 five, so which is a high dependence, dependence parameter is higher. So, this is for the theta equals to phi and this is for theta equals to 0.5. So, you can see that when the theta is less, then we can get the,the, dependence pattern can change significantly depending on what the, what the,the, theta parameter, the,the, dependence parameter is opted for, and this is one of the positive thing of having a wide range of dependence. You can see here, it is from the minus infinity to the plus, sorry, this is from minus 1 to plus infinity. So, the wide range, so, we canmodel depending on whatever the data is available to us, we can model a wide range ofdependence between the two random variables.

(Refer Slide Time: 19:00)



So, now, we will seethat howthis, this, Kendall's tau that I was talking that this Kendall's tau is relatedto theirparameter that theta parameter if it is an Archimedean copula. Now, remember that whatever the expression that we are going to, going to express is that this is for in case the copula is the Archimedean, Archimedean copula.

So, earlier the expression that we have seen that is for the general as I, as I mentioned that that is for the general expression for any copula, and from there depending on the whatever the special properties that class is having, we can reduce that relationship to some other form, some other easiermathematical form.

So, here, the example for that Archimedeancopula. So, the first line very very straight forward is that,in case of Archimedean copula, so,in case of Archimedean copula, this Kendall's tauis reduced to this form, that is, 1 plus 4 times integration from 0 to 1 that phi u by phi prime u. This phi prime u is that first order derivative of that function with respect to u, and this phi u you know that this is that generator of thisfunction and that we will integratedefinite integration from 0 to 1 and that 4 times plus 1 gives you the form of this tau.

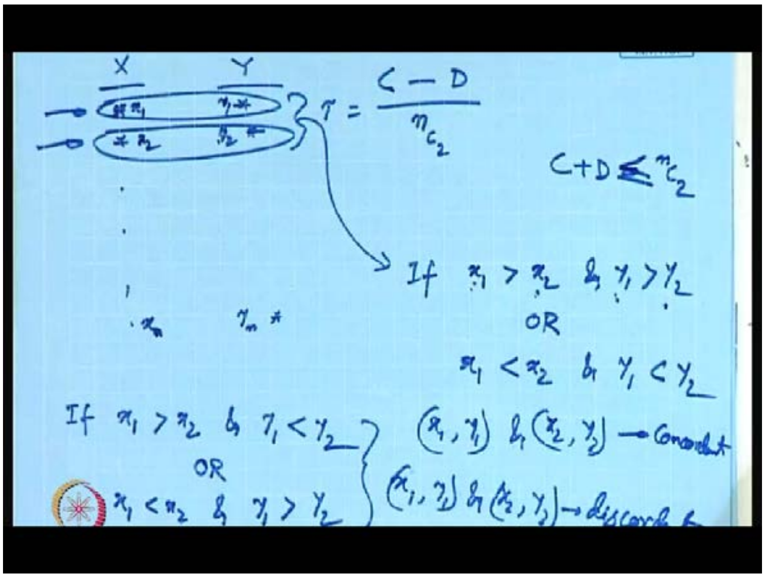
Now, you know that this, in this form, there is the parameter theta is there. So, that theta will be there in this expression and left side there is tau. So, here, what we can see is that there will be a relationship between this tau and the theta, where the theta is your,the measure ofthe parameter that is capturing the dependence between two random variables.

Now, so, what we have to do? First, from the sample is that, we have to estimate what is the value of tau and that is a Kendall's tau, and this Kendall's tau if we know from this data, then we can estimate that sample estimate. We can get the sample estimate of the theta and that theta we can, we can, usefor the complete development of the joint distribution.

Now, we will take some time to, to, explain what is, how we can get that sample estimate ofKendall's tau. This sample estimate ofKendall's tau, this is tau cape is equals tothatx i minus x j multiplied by y i minus y j greater than 0 minus x i minus x j into y i minus y j less than 0 for all i j - where i and j are not equal. So, this may be the mathematical expression. Let me explain that what it, what it, meansthrough two.

Basically, the first probability that we are talking about this is the concordant pair. So, x i y i, y x j y j these are the two pairs of thisof the sample, and if this, if this relationship holds good, that is, it is greater than 0, then we say that it is a concordant pair, or if this relationship holds good, then we say this is a discordant pair.

(Refer Slide Time: 22:37)



So, I will just explain here how thisthing from the data how,how, how, it looks. So, suppose that there are two that randomvariable that we are talking about and we are we are having some data available to us. So, this is the first entry. The first entry of this x and the first entry of y. So, this is paired. So, this is basically what we are talking about. This is say the x 1 and this is y 1. Similarly, this is x 2 and this y 2, and like thatwe are having in data set x n and this isy n.

Now, to know that, what is theKendall's tau between these two series is there. So, what we have to take? We have to just pick up two pairs out of this n pairs. Now, two pairs out of these n pairs means so you can, you know that so totalpair that we will have is that n c 2. There is n combination 2. Out of n pairs like this, how many ways you can pick up? Two pairs. So, this is the total possibility.

Now, after you pick up say for example, you are taking the first pair and the second pair. Now, you have to know whether these two pair are concordant or discordant. Now, if this x 2 suppose, so the, if I just take this to 1, so, if say thatx 1 is greater than x 2 and y 1 is greater than y 2 or x 1 is less than x 2 and y 1 is less than y 2. So, if this one increasing, this should also increase, or if this is decreasing, whether this is also decreasing or not? So, that, that trend that is the why it is changing. Whether, in both the cases,it is changing in the same direction or not? If that is so, then that means this is what is explained through this mathematical form. Then these two pair x 1, these two pair means x 1, y 1 and that x 2, y 2. So, these two pair is should be the concordant; these 2 pair is said to be the concordant.

Now, if the opposite thing,thing, happens, so that if that x 1 is greater than x 2 and y 1 is less than y 2 or x 1 is less than x 2 and y 1 is greater than y 2. So, if this relationship holds good, then the same, then that pair x 1 y 1 and x 2 y 2 is known as discordant. So, now, if this one, so, now, if you see that now how to test it through that, supposewe can just seewhether if just by visual inspection also you can see, or what we can do is that we will, if we take that x 1 minus x 2, if we just want to combine both these condition together, then we can say that if the x 1 minus x 2 multiplied by y 1 minus y 2 is greater than 0.

You see if it is greater than this and this is greater than this, then the multiplication also will be greater than this, or if this is less than this and this is less then this, then also x 1 minus x 2 multiplies by y 1 minus y 2 will be greater than, will be greater than 0. That is why it is written here that if that x I minus x j multiplied by y I minus y j is greater than 0; that means, it is a concordant pair.

And if the opposite thing happens,for example, here what we have shown that it is greater than and it is less than, it is less than and it is greater than, then what will happen? This is the discordant. So, their multiplication x 1 minus x 2 multiplied by y 1 minus y 2 will be less than 0.

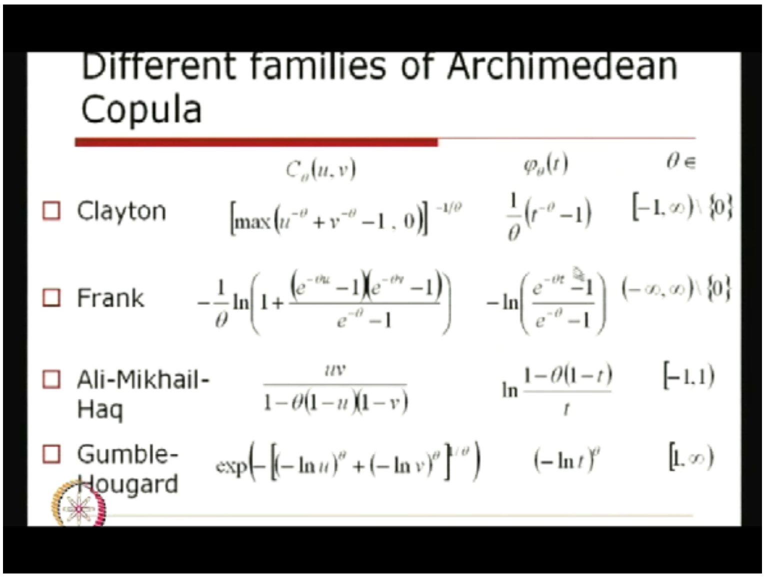
So, in this case, this is the, this is, this is discordant. Now, this sample, this Kendall’s tau is the difference between the probability of concordant pair minus probability of discordant pair. Now, out of this n, out of this n pair, that is available to us. We know that there are totalway that we can select two pair is the n c 2. Now, out of this n c 2 choices, there could be few choices will be your, will be the concordant and few choices will be discordant.Suppose that total number of concordant pairs is c and total number of discordant pair is d.

So, way the definition says there is a difference between the probability of the concordant pair minus the probability of the discordant pair; that means the probability of concordant pair. From this sample, what is that? It is c by m c 2 and probability of discordant pair is d by n c 2. So, if we get this 1, then this will give the estimate for thusthe Kendall’s taufrom this sample.

Now, there could be one thingcan also happen, that is, if that x 1 is equals to x 2 or that y 1 is equals to y 2, so, because this is a sample, this can happen. So, in those cases, we generally excludethose, those, pairs. So, that is why, so, the c plus d is in case in the ideal, ideal cases if both are the continuous random variable and the,the, probability of being exactly equal tothe other output is,is, 0 means from the theory point of view, that from the sample, it can happen that x 1 is equals to x 2 or some pair is equals to x 2.

So, if they are not equal, there is no entries are equalfor both the cases, then ideally c plus d is equals to your n c 2, but as there could be somepossibility of the equality, then this is basically c plus d is less then equals to n c 2, 1 or 2 combination that has to beignored. So, after ignoring that one, you will get that c minus d divided byn c 2 and that is giving you that what is the estimate of thisKendall's tau and this is what is explained. Here, this c and this d are the number of concordant pair and this is the number of discordant pair. So, this is the c y n c 2 is nothing but your, the probability of concordant pair and this is the probability of discordant pair. Their difference is gives you the, that Kendall's tau.

(Refer Slide Time: 29:53)



Now, we, form the sample, we got how to get their, their, sample; how to know their sample estimate of the Kendall's tau. Now, we will see that how this sample estimate can be linked to thattheta, because we have already seen the expression that, if that theta, theta, and tau are related to that one, one, form in case of the Archimedean copula, and before that, we will just discuss about fewArchimedean copulas. So, we are in the class of Archimedean copula. What we are showing you now? The different possibilities few slide back,I mentioned that there are many varieties are available and their range of the dependence also varieswidely.

So, here, some four examples, four different types of Archimedean copulaare picked upfor the discussion. The first one is the Clayton copula, for which the form of thiscopula is the maximum ofu power minus theta plus v powerminus theta minus 1,0. So, out of this two, which one is maximumthat powerminus 1 by theta. So, this one just few minutes back also we discussed, and so, this copula name is your Clayton copula.

Andthis one asits generate is the form of itsgenerator is your 1 by thetamultiplied by t minus,minus, t minus theta minus 1. So, this is the generator of this copula, and the last column is giving you the range of theirdependence as this is already, already, stated that this range is from this minus 1 to plus infinity excluding 0.

The second one is the frank copula, for which theexpression for the copula function is minus 1 by theta log natural of 1 pluse power minus theta u minus 1 e power minus theta v minus 1 divided by e powerminus theta minus 1, and for this frank copula, itsgenerator is minus log natural of e powerminus theta t minus 1 divided by e power minus thetaminus 1, and its, the range of this dependence is minus infinity to infinity excluding 0. Now, as you know thatas this generators are given and generator function is given and these are the Archimedean family. So, what you can do is that, you can take this, this, generator and,and, fit into thatform, this form and you can, we can see that whether you are get that copula form or not.

So, we have, in the example, we have discussed this copulawhich you will see here now. This is the, so, so that one is 1 copula which is not listed here. That is also one of these Archimedean copula.

So, the next one, the after this Clayton and frank, it is the Ali MikhailHaq, andthis copula form is u v divided by 1 minustheta into 1 minus u into 1 minus v and its generator is log natural of 1 minus theta into 1 minus t divided by t, and if you see its range, it is from this minus 1 to plus 1; 0 is included, but this plus 1 is,is, it is an open,open, boundary and minus 1 is a close boundary. You know this open boundary means it is not exactly equal to one, but, but, mathematically it isvery close to 1.

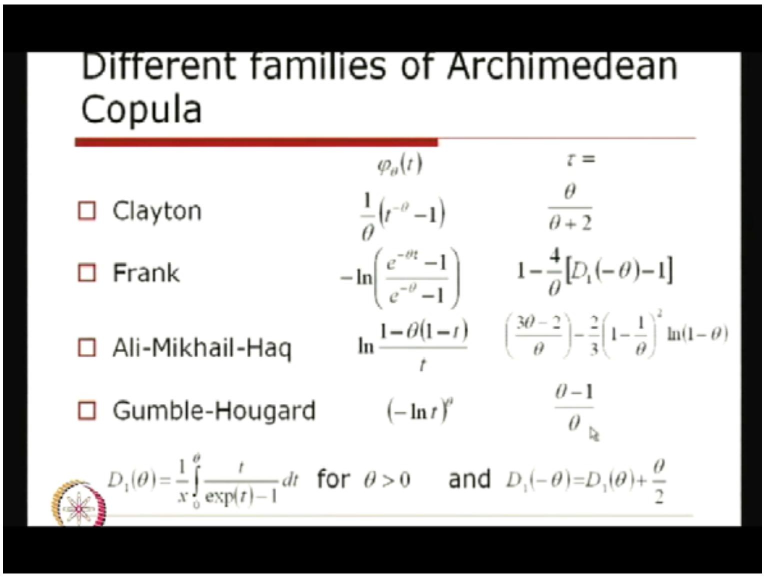
So, this is the, this is, this istheir dependence range. So, the difference between this first two, that is, the Clayton and frank is that for this two that dependence range is very wide, but the 0 was excluded. For the Ali MikhailHaq, the dependence range is very narrow, but 0 is includedhere.

And the last one that is the gamble haggardcopula, for which thatcopula form is exponential of minus of that minus log natural u power theta plus minus log natural v power theta whole power 1 by theta and its generator isminus l n t powertheta and its range of dependence is the 1 to infinity. So, you can see here, this is one example where only the positive dependence can bemodeled.

So, even the positive means it starts from 1 only, so, not even from 0. So, the only the positive dependence can be measured. For the Ali MikhailHaq, the dependence range is range is narrow, but it is close, but it isnarrow at 0, but it includes 0. For the frank, it is the most wide range from the minus infinity plus infinity, but the 0 is excluded. For the Clayton, it is minus 1 to plus infinity. Again that 0 is excluded.

Now, we will see how this one, this, this, functionwe have seen that this one divided by its firstderivative definite integral from 0 to 1 4 times and it has beenadded with 1 to get that tau to relationship with this tau, but I amtelling is this one. So, this Kendall’s tau equals to 1 plus 4 times 0 to 1 phi u by phi prime d u. So, we will just seefor this 4 Archimedeancopula how thatthing is,is, related with the tau.

(Refer Slide Time: 36:08)



Here is that thing, here is thatexample that is for the Clayton, we have seen just in the previous slide that its generated is 1 by theta t power minus theta minus 1, and if we fit in that equation, we will get that this tau is equals to theta by theta plus 2.

Similarly for the frank, it is thatminus log natural of this e power minus theta t minus 1 divided by e power minus theta minus 1. So, these are the generators which you have alreadydiscussed here. Now, here, what you are showing that with this one, if we fit this,this, generator in this equation, then how this relationship between tau and theta come, and for the frank, it comes that this is tau is equals to 1 minus 4 by theta d 1. This isfunction again D 1 of this minus thetaminus 1.

So, so,you can see that this is the full expression with respect to theta and which is equated totau. Now, this D 1, before I go to that one, this d 1 is known as thisfirst order. This one stands for the order the first order d y function.Now, what is this first order dy function is explained here, is 1 by x 0 to theta T by exponential t minus 1 DT in case that theta is greater than 0; otherwise, if for the negative 1 that theta minus, sorry, d 1 minus theta is equals to D 1 theta plus theta by 2.

So, if it is negative, if the theta is less than 0, then we have to convert it to this form. Then this one will beexpress through this one; basically for thenegative one, this 1 plus theta by 2 it will come. So, from this expression, you can see that may be this one will not be, will not have any close form solution, but numerically you can solve that one, but for the other one, the first one the Clayton we can this tau if we know, then it is very easy to solve what is the value for thetheta.

Similarly, for the Ali MikhailHaq, that the generator islog natural of 1 minustheta into 1 minus t divided by t, and for this one, that tau is equals to 3 times theta minus 2 divided by theta minus 2 by 3 into 1 minus 1 by theta whole square log natural of 1 minus theta. So, this is the full form ofthat tau is equals to this one, and last one is the Gumble Haggard, for which the generator is minus l n t power theta, andfollowing the same equation, this tau will be equal to theta minus 1 bytheta.

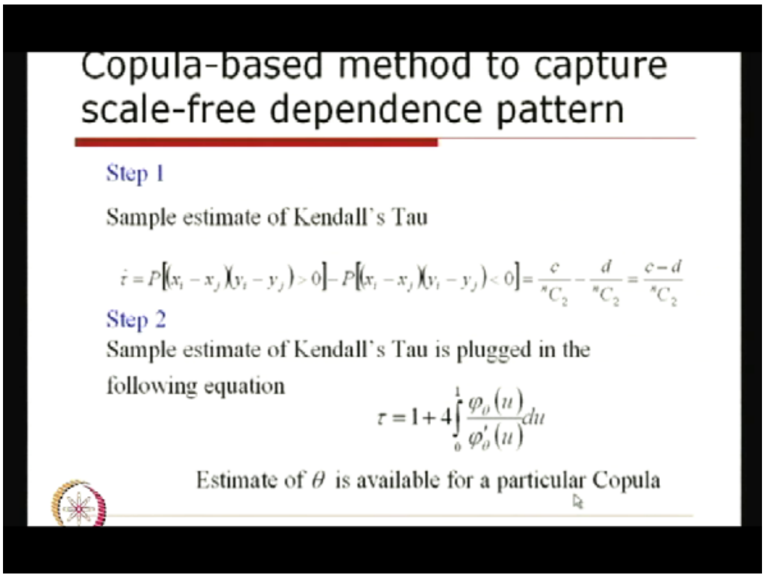
So, now, it is known that from the sample estimate, we can, we can have that tau first;then we can have the estimate of this theta and that theta we can fit into that whatever the copulafunction tentatively used. That you can fit and then we can get the joint distribution. Now, after getting the joint distribution, sometime it is essentialthat.

So, if you use only one copula function, obviously, you will get 1, and now the depending on the choice which copula function you are using, so, that way, that your, that joint distribution may change. So, this is one of the question mark, but there are some methodologies available, for which how to test thatfitness which copula is the, is,is, is, the fitting the data.

For example, this is also not uncommon in the sense in the in the single random variable. If we have a data, we generally go for the test that how this data is going to fit in the, for which distribution. So, that time also we have to pick up some severaldistribution and we have to see that for which distribution the data is fitting properly.

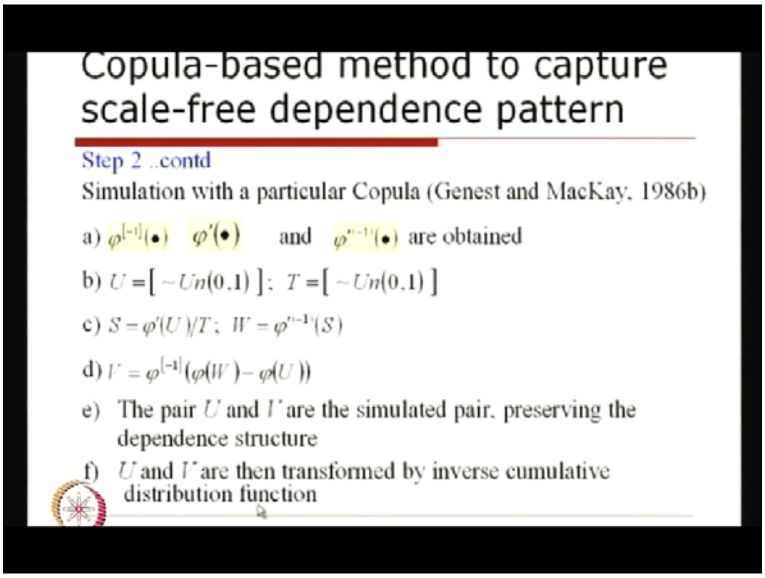
Similarly, here also if we just takemore than one copula function, then we can test that which 1 is fitting the best. We will take one example of this, of thevita logy, but problem to discuss that whatever the theory that we have explained so far and we will see thatthis theory is applicable. In a step wise fashion, we will explain, and finally, it is going to sum up this prediction problem, but so far as the copula is concerned, this iswe can, as long as we can get that their joint distribution. Up to that, we will, we will see, and after that, one slide is there for that prediction one, but that is thefurthertaking of thatparticular outcome.

(Refer Slide Time: 41:11)



So, thatthe example that we are going to discuss is on this copula base method who capture the scale free dependence pattern. So, this one as we havediscussed that first one, the first step that we should do the, whenever we are having some data, we will we will estimate the, what is the sample estimate of the Kendall's tau of this data. So, this is the form that we have discussed, and in the step two, that sample estimate of the Kendall's tau in plugged in the following equation, because as we are using thatArchimedean copula, so, you know that this equation should be plugged in that tau estimate. So, we will get the estimate of this theta for a particular copula.

(Refer Slide Time: 41:55)



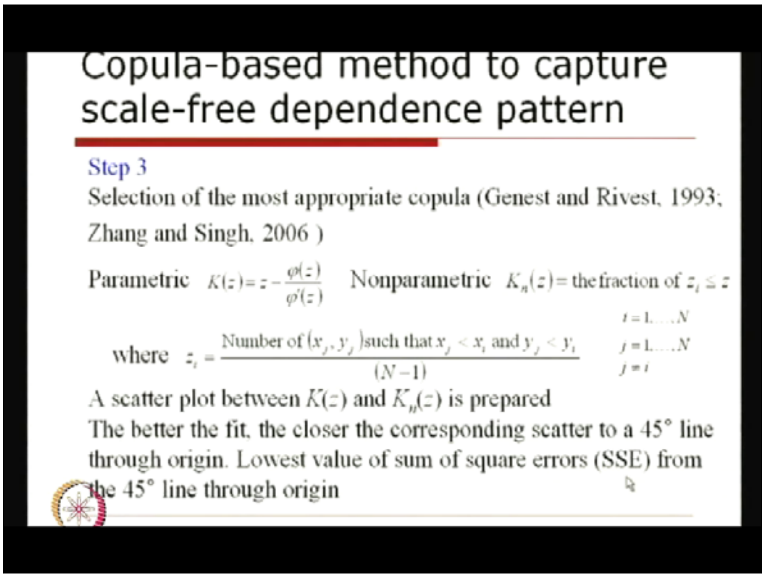
Now, what we have to do? We have to generatesome of thesimulation of the particularcopula. So, one copula we have taken; we know what is this dependence parameter. Now, if we can simulate thatthat, so one joint distribution is there, we have to simulate that one preserving their dependence pattern. So, this one is that, is the, are the steps.

So, first of all, we know what is this generator, what is this generator, that is the phi, and from that phi, we have to obtain that this pseudo inverse. Its first derivative and its Inverse of the first derivative, and then, we generate two series - one is U which is one uniform distribution, which between the 0 and 1, and another one is the tuniform distribution again from 0 to 1.

So, this u and t will generate independently. So, thus,thus,U and t have totally independent from each other. Now, we will get another one, another series out ofthis U and t, which is that s is equals to that first derivative of this, of thisgenerator of that u divided by that t. So, each point if I just divide and then from that s if we take that its inverse of the first,first, derivative of that s, then we will get 1 series that is W.

Now, this v is equals to that pseudo inverse of that generator phi w minus phi u and this will give you another new series which is v. Now, the pair u and this newly generated pair v, these are the simulated pair preserving their dependence structure. Now, this dependence structure share means that through the parameter theta which is barred in this, in this generator function. Then u and v are transformed by that inverse simulative distribution function to get their in the originalscale.

(Refer Slide Time: 44:40)



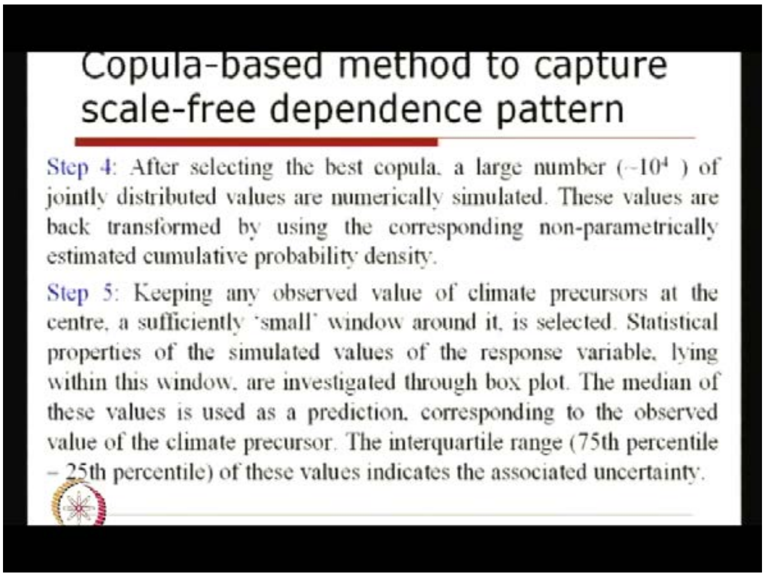
Now, this original scale means that first of all say that we are having the two data set which follows, which may follow some distributionor gamma whatever. So, from that distribution,I can first get their reduced vitiate which are u and and,and, v. Now, from that now whatever we have generated here that u and v, these are those reduced variety but they preserve their dependence structure. Now, that u and v can be transformed back through that inverse simulativedistribution function to get that in that original, original, scale of the random variable.

Now, in the step three as I wasdiscussing that the selection of the most appropriate copula, so, that is, there are two estimate - one is that parametric, that is, the k z is equals to z minus phi z divided by phi prime z. Now, phi prime is the first, first derivative of that generator putting that z as the variable. So, this one is the parametric estimate of that cumulative distribution and the non parametric is also we have to use it from that data that is available to us. That is the fraction of z I which is less than z. You know that for this is basically the cumulative distribution how we get from this data.

Now, this z i is equals to that number of x i j i such that, x i is less than x j is less than x i and y j is less than y i divided by n minus 1, where this i varies from 1 to n; j varies from 1 to n, but I and j are not equal to each other. So, in this way, we will get both the estimate that k z and this k n z. One is parametric; another one is non parametric.

Now, thisnow, this,this, k and z is scatter plot. If we just prepare, that is, k z and k n z if the, if we, if wehave selected thatcopula correctly, then these two should be the, should be in such a way that it should be a straight line through the origin, which next 45 degree angle with the origin if these two are ideally same. So, it says that better the fit, closer the corresponding scatter to a 45 degree line through origin. Lowest value of the sum of squire error from the 45 degree line through the origin. So, that has to be picked up as the best one.

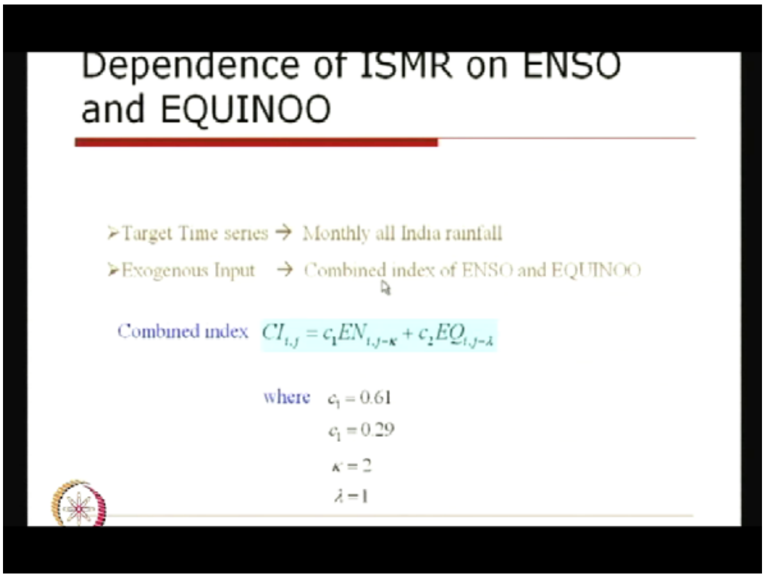
(Refer Slide Time: 46:33)



Now, afterselecting the best copula, a large numberof the jointly distributed values are numerically simulated. These values are back transformed byusingthe corresponding non parametrically estimated cumulative probability density keeping any observed value of the that here, that is the climate, hydro climaticproblem. Any the climate precursors at the center is sufficiently small window around it is selectedstatistical properties of the simulated values of the response variable lying within this window are investigated through the box plot. The median of these values is used as a prediction corresponding to the observed values of the climate precursors the interquartile range, the 75 percentile to the 25 percentile of these values are indicated of the associated uncertainty.

So, this is the last step as I was telling that is the further taking through this one, but as long as you get their joint distribution and you are able to simulate the large number of, large number ofthat joint pair preserving their dependence structure. So, this, this preserving the dependence structure is the most important thing here that we are telling. So, we will just go through quickly that problemthat we aretalking about.

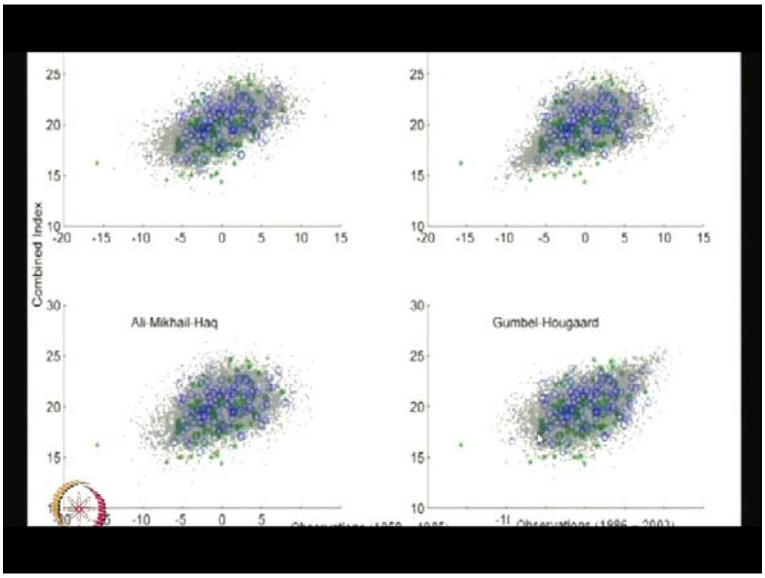
(Refer Slide Time: 47:55)



So, this you have to, for the detailed discussion, you have to refer to one journal paper that I will tell you later, but for this 1 is the there are two time series has beentaken.So, the problem is basically the bivariate here. The monthly all India rainfall is the one time series and other one is the combine index between this Eons and Equinoo. This Eons is the Elion southern oscillation of all the tropical pacific ocean and this is the equatorial Indian ocean oscillation, which is a atmospheric phenomenon over this tropical Indian ocean. So, there arebackground how we get thiscombined index, this combined index for the I-eth month and j-eth year is your that, that, kappa month lag of this eons info index and this lambda month lag of this equino index multiplied by c 1 c 2 and this values also obtain from other analysis c 1 c 2 and this kappa and lambda.

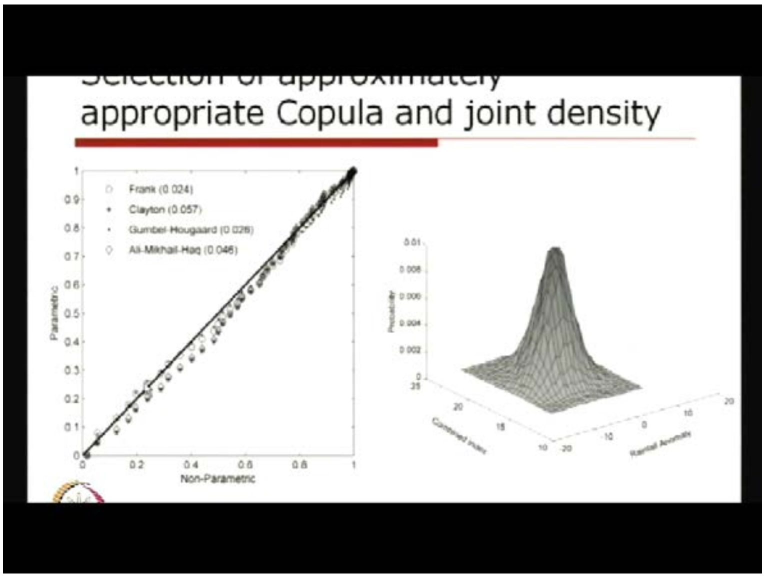
So, finally, using this parameter and using these indices, we get one that inanother time series which is c I which is known as the combined index. So, we have the all India, monthly all India rainfall and we have the combined index. So, two time series is available to us.

(Refer Slide Time: 49:15)



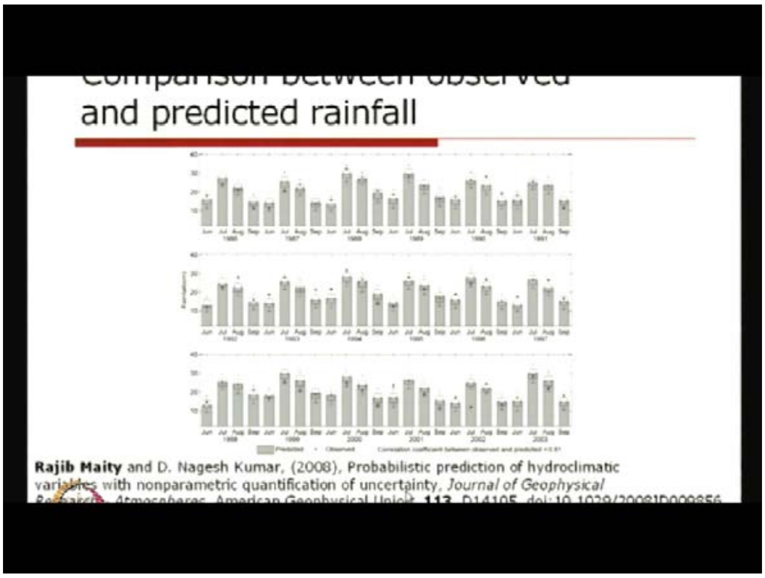
Now, we are followed those steps and we have used four differentArchimedean copulas here and these four panels that you can see are for the four different copulas - one is the Frank Clayton,Ail MikhailHaq and Gumball Haggard. Now, you can see that, what you can see here are the three differentcolours - the first one is the, is, is,is, the blue circle. These blue circles are the data points that has been used for whatever the calculation we need for this Kendall’s tau and its theta estimation, and from the respective copula, how we get their joint distribution, and after that, we have simulated through that jointdistributionthat we have developed and that simulationis a shown by this grey cloud, this dots, grey dots, and over that, we have over led some unforeseendata by this, by this, green star. So, this was basically used for the validation. So, this iscovered by this cloud. So, similarly, for all these three copulas, we have used the same procedure and you can get this one. Now, the second is the, which one we should pick upfrom this one.

(Refer Slide Time: 50:35)



So, for that one, we have just a non parametric estimate and the parametric estimate and we plotted through this scatter plot andtheir SSE - Sum of Square Error - from this 45 degree line isobtained and these are the values is shown here. So, this one you can see that the frank one is the best, because this is the closest to the 45 degree line, and using this frank, this is how the joint distribution between the rainfall annuallyand the combined indexlookslike.

(Refer Slide Time: 51:08)



Andafter this one, the finally what we have seen is that? This is that prediction on to the last step that we are talking about, which is may not be thein the main focus of this theory of copula that we are discussing in this class, but for this, furtherof this whatever the problem we have discussed for,for, the references, you can refer to this publicationthe probabilistic prediction of hydro climaticvariable with non parametric quantification of uncertainty.It is in journal of geophysical research atmosphereby American geophysical unionthis reference you can refer to.

So, what I want to mention here at the end of thisthe brief introduction of this copula is that, this copula itself is a wide areaand it is not possible to capture the, everything in these two lectures that we have covered here. So, further studies you can have in the different textbook that is available forthis on this copula, and so far is the application is concernedon this on this copula is that, is that, this as very recent and mostlyin,in, in, the research area. So, one of thusexample is shown here and you can refer to some of these recentpublication for itsvarious application of this copula in the different field ofresearch includingthe different field of Civil Engineering. Thank you very much.