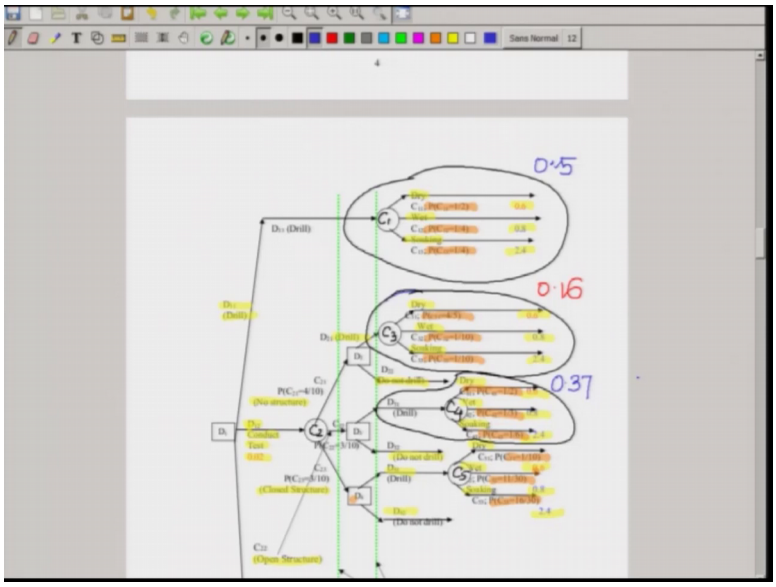
Data Analysis and Decision Making - II
Prof. Raghu Nandan Sengupta
Department of Industrial & Management Engineering
Indian Institute of Technology, Kanpur

Lecture - 20
Decision Trees

Welcome back dear my dear friends; a very good morning, good evening, good afternoon to all of you. And this is the DADM II which is Data Analysis and Decision Making course under the NPTEL MOOC series. And, this course which is for 12 weeks which is for 30 hours and each week we have 5 lectures, each lecture being for half an hour. And, after each week we have assignments and we are at the last day for the 4th lecture which is the 20th lecture. And as you remember, I am not using slides for this particular class for the last class also.

Because the diagram was so, huge that it was much better for me to draw it in word and then share in a PDF format and do the analysis accordingly. My name is Raghu Nandan Sengupta from IME Department IIT Kanpur from India. So, I will start with the same slide where I left in the 19th lecture. So, I will just recap very fast. So, you had a company and the company wanted to is in oil production. So, if it in oil production it can either do nothing; that means, though it is not a profitable business so; obviously, this arm would be true I have already highlighted.

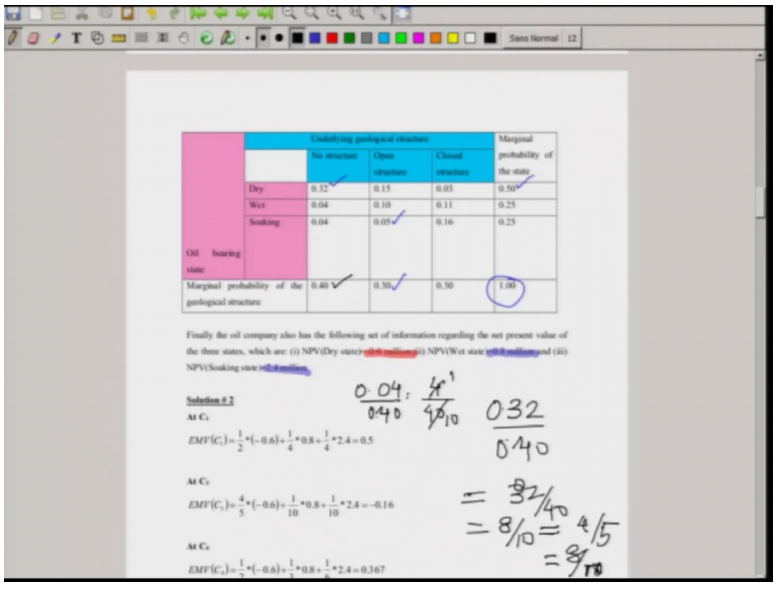
(Refer Slide Time: 01:35)



So, I am not going to again put any another colors. So, I will just hover the pen accordingly. So, this is the arm where you do nothing and the case can be where you drill, but the drill arm would be given by D 1 1. And if we drill obviously, you find the structure to be dry wet soaking and the probabilities which I mentioned in the last class I will come to that later. And the values which with net present values was minus 0.6 plus 0.8 plus 2.4 are given out for the dry wet soaking. Now, let us go through the values of probabilities, I will mark the probability highlight them using a different color; so, let me use the orange one.

So, in you can differentiate; now if you see the drill part what does it mean? In the drill part you immediately go under the drill. So, there is no conditional aspect; that means, you did not conduct any tests or you did not basically do some experiment or did not do any marketing survey or something nothing was done, you immediately went to the drill. So, the drill factor whatever the probabilities are given are unconditional. So, if they are unconditional the mark the probabilities; the probabilities are point half which is 0.5, one-fourth which is 0.25, one-fourth which is 0.25. Now, remember these values if you remember in the table; so, I think I should come back to the table again.

(Refer Slide Time: 03:21)



If I am a little bit slow going, repeating it please bear with me. The values which were there 0.5 0.25 0.25 are the unconditional probabilities of a dry wet and soaking and that is exactly what we have drawn in the diagram, which is dry wet soaking. Hence, the

probabilities are given as 50 percent 25 percent a 25 percent. Now come so, this total arm I am just marking it without putting any color. This whole arm is has been delineated nicely with the probabilities and the net present values and what are the arms. Now consider you conduct the test.

So, when you conduct the test this is given as D 1 2 and when you conduct a test you have already incurred some money and if you remember I did mention that value incurred is 20,000. So, its 0.02 with an red color; that means, is negatives its cost which is going out of your pocket. Now, when you conduct the test you found out there were 3 structures of the geophysical properties; no structure, one was open structure and another was basically found out to be closed structure. So, the arms are this once you reach let me mark it will be easy for us to so, this was D 1 1. So, let me mark it so, this becomes C 1.

So, that is why C 1 1 C 1 2 C 1 3 this becomes C 2, this becomes C 3 this becomes C 4 this becomes C 5. So, it will be easy for us to mark now and the D 1 D 2 D 3 D 4 are already marked. So, hence it will mention C 1 1 C 1 2 and C 1 3, now coming back to this arm where, you do not you conduct test. So, no structure you see the problem the arm is C 2 1 and the probability is given as 4 by 10 so, from where does this probability 4 by 10 come; it means that you have conducted the test and no structure was found. So, let us go back. So, conducted the test no structure was found probability was given as 0.4.

So, this is the probability which is have which is 4 by 10; you conducted the test and it to found the open structure. So, these are unconditional, they were not predefined by anything and conducting test was an thing which you have done with probability 1. So, the probability for open structure if you see was 0.3 which is 3 by 10, let us check it is 0.3 yes it is 0.3. Similarly, when you go for a closed structure open structure closed structure probability is same. If you go there the probabilities for the closed structure is 3 by 10 and for open structure is also 3 by 10 3 by 10 3 by 10 0.3 and these arms are C 2 2 and C 2 3.

Now, when you have considered these so; obviously, it means that you have considered no structure, open structure and closed structure and then in another under each scheme you drill or do nothing. So, if you drill which is basically for the first case no structure

and drill is D 2 and then you can arrive at C 3. So, at C 3 again you find out that is the dry or a wet or a soaking well. Now, the probabilities which are given are to be noted down carefully. So, the arms are C 3 1 C 3 2 and C 3 3 and the probabilities are corresponding to the fact four-fifth one-tenth one-tenth. And why they are so, let me mark them yellow color four-fifth one-tenth one-tenth.

Now, let us consider what decision do you go through, no structure and then you drill. So, let us go here, no structure and then you found out. So, no structure it is the this column which you have. So, find out what are the probabilities 0.32 by 0.4. So, as I said that if it is no structure and then it is dry so, the probabilities would be four-fifth. If you find out no structure wet value would be similarly when it is no structure soaking values one-tenth. So, mark these values four-fifth one-tenth one-tenth four-fifth can also be written as 8 by 10 in order to make the denominator same.

So, let us see whether they match yes they match. So, if it is no structure and then you find a dry well it was four-fifth which is 8 by 10; then if you find out no structure drill and find out wet the probability was one-tenth, then you find no structure soaking the probabilities was one-tenth. So, the probabilities are given as four-fifth one-tenth one-tenth which is the condition probabilities as found out and given in the thus the matrix which was there. The values which were there the present values were again I am repeating was minus 0.6 plus 0.8 plus 2.4 and; obviously, 1 arm which is D 2 2 is that do not drill do nothing.

Now, come to the case when it is open structure. So, if it is open structure again you drill which is D 3 1 and you do not drill which is D 3 2. If you drill again you have dry wet soaking, the returns net present value of being for dry wet soaking being minus 0.6 plus 0.8 plus 2.4 and the probabilities are given. So, let me first highlight the probabilities half one-third one-third. So, this is for the case when you have open structure let us go back open structure. So, the values are without doing a calculation 0.15 by 0.3 which is 15 by 30 which is half 0.1 by 0.3 which is 1 by 3 is one-third and then again value of 0.05 by 0.3 which will come out to be 5 by 30 which will be 1 by 6.

So, values are as I mentioned half 1 by 3 1 by 6 are the conditional probabilities based on the fact open structure has been found out and then you find after drilling you find out dry wet soaking. And then D 3 2 is basically you do not drill finally, come to the closed

structure when I am hovering my electronic pen. There once you drill you find out dry wet soaking which is D 4 1 because, D 4 is basically the arm decision which you reach once you have conducted the test. And, then found out closed structure and then you are drilling and another D 4 2 is basically you do not drill. Now, for dry wet soaking the probabilities are given I am just repeating and I will coming back again one-tenth 11 by 30 16 by 30 let us see how they are obtained. So, this is opens closed structure closed structure is the right most column.

So, the corresponding values are 0.03 divided by 0.3 it will be 3 by 30 which is 1 by 10 wet will be 11 0.11 by 30 is 11 by 30 and soaking would be 0.16 by 0.3 which will be 16 by 30 which is 8 by 15 So; obviously, 8 by 15 I just simplified it. So, if we check the values are exactly as I mentioned 0.1 10 0.1 by 10 11 by 13 16 by 30 and the net present values are as mentioned for this arm also, when you find out dry wet and soaking are minus 0.6 plus 0.8 plus 2.4.

Now, with these values given you have to do the decision take a decision; then the decisions are to be taken in such a way that the vertical green lines which I have drawn are the stages at which you take the decision. Now, with the diagram clear I will now start the calculation, remember (Refer Time: 13:27) thing there are many simplistic assumptions which I have assumed.

But, I am trying to give you a feel that how you can utilize that accordingly; remember one thing for all this problem both for the first part and this problem we are only taking expected value as the best criteria of making decision. It can be suppose that say for example, you have the variance also. So, I want to basically rank them with respect to the mean and the variance or the expected level of the variance. It can be I basically rank them with respect the ratio of expected value by variance rank them from the highest or the lowest and take it accordingly. Or, it can be I want to find out the reverse like if you remember when you took the d a we took the concept of efficiency and maximize it and in other case we took the inverse of the efficiency and minimized it.

So, in this case we can take the ratio the variance to the expected value rank them from the lowest to the highest and take the one which is the lowest, they can be different combinations. So, I am only going through the simplistic part I expected value; now let us go to the calculations so, before the calculation so, solutions. So, at C 1 you are take

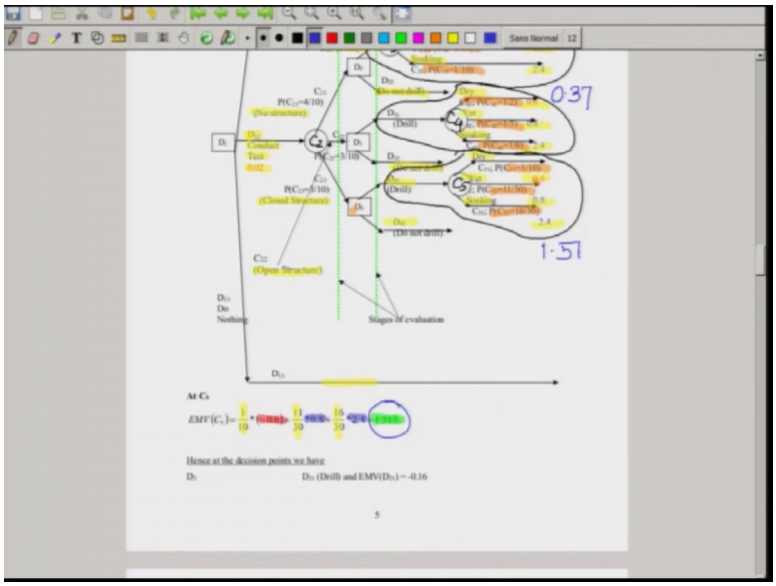
you are going to take C 1 is the upper arm where you have not conducted any test you just went to the drilling and the probabilities were given as 50 percent 25 percent 25 percent. Net present values was minus 0.6 plus 0.8 and 2.4 the probabilities were given. So, this value was minus was there and plus was there for other 2 arms, probabilities were given I will just utilize yellow color probability is given as 50 percent 25 percent 25 percent; the value comes out to be 0.5.

So, the 0.5 value is at this stage this whole arm this is point the value which we calculated which is 0.5; now let us go I am just hovering going up and down. So, please bear with me at C 3. So, where is C 3? So, this is arm this is arm is C 3. So, how do I do? I multiply it four-fifth into minus 0.6 one-tenth in to plus 0.8 one-tenth into 2.4. So, let us see does it comes like that. So, again the values are minus this is plus plus probabilities are given. So, I should use a different color to give the answer let it be green no sorry sorry. Now, note down one now note down one thing which I should not mark. So, I am going to give the answers in green, but mark it accordingly and why I am doing that you will understand.

This value 0.5 which was the first term C 1 which was here is positive. So, I will mark with the blue circle now let us check C 3. So, the C 3 values were we calculated for four-fifth into minus 0.6 one-tenth into plus 0.8 one-tenth into plus 2.4. If you calculate the value comes out to be minus of 0.16 which should be marked with red because this is negative. So, now when I compare I will go 1 by a so, when I compare this 1 I have a value of plus 0.5. So, I might think I should note it down for better understanding. So, this value for C 3s comes to be minus 0.16 when I come to; so, I have to come to C 4 now. So, C 4 value is here again minus this is the value net present value these are plus 0.8 and 2.4 and the values are given half one-third one-sixth.

What is this half one-third one-sixth? Are the probabilities if you see half where I am hovering the yellows colored pen half one-third one-sixth. So, these are the probabilities I multiply them the value comes out to be 0.367, 0.37 consider what this is also positive. So, I can mark with the color blue. So, 0.37 consider 0.37 is coming so, I go take this. So, it will be I am only taking 2 pieces of decimal 37.

(Refer Slide Time: 19:21)



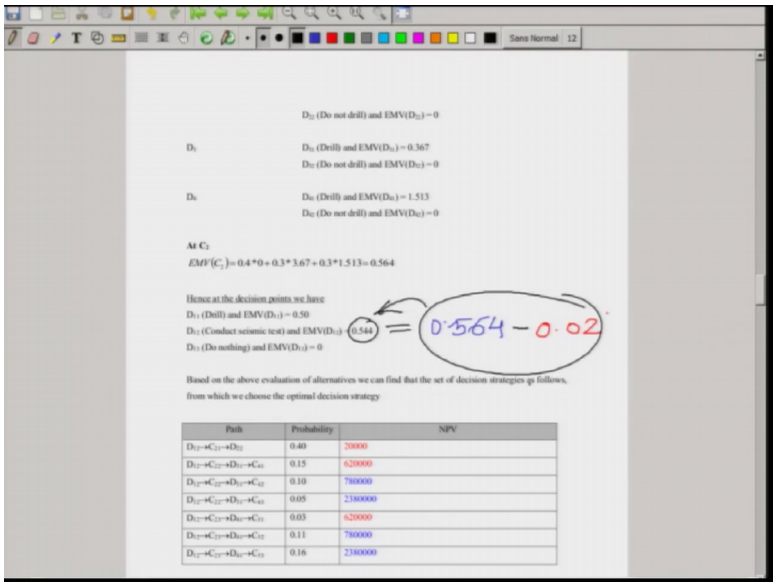
Then I go to C 5 so, C 5 I go down. So, probabilities are given as minus 0.6 not the probabilities the values plus 0.8 2.4 one-tenth 11 30 16 30. So, these are the probability which you have already calculated one-tenth 11 30 16 and 30 using the condition values. And, the value which you find out comes out to be 1.5 this is positive. So, circle with a blue color so, this one answer to be 1.51. So, these I have marked as blue. Now, if you see now I have to take a decision. So, if you want to take a decision how would you do that? First I will explain and then come up with the calculation.

So, for C 2 if I take I will compare C 3 with respect to D 2 2 so; obviously, when it take the expected value it was minus 0.1 6 and the other value for do not do anything do not drill is 0 so; obviously, I would not do anything. So, this whole arm C 3 would be not there because, we would not consider because that is a negative value. If I compare C 4, C 4 is plus 0.37 when I compare with do not drill it is 0 so; obviously, it is positive I will take 0.37 as positive and consider that. And, when I take the values of C 5 the C 5 value would be is already 1.51. I will consider that when I because when I consider that with respect to do not drill it is positive so, 1.51 would be considered.

So, the first arm has 0 second arm was 0.37 third arm has 1.54, but when I go into the last stage; last stage means here for this case when I consider C 2, C 2 is going to consider 0 with a probability of four-tenth because that is the no structure probability has been found. 0.37 with the probability of 3 by 10 and 1.51 with the probability of 3 by 10.

So, 3 by 10 3 by 10 and 4 by 10 are corresponding to the probabilities which have been found out uncondition one these are these values. That means, no structure open structure close structure corresponding to the fact whether they are dry wet soaking does not matter this is unconditional. So, hence the decision would be at D 2 it is minus 0.16 so, this is as given and note do not drill is 0. So, D 2 has a negative value. So, we would not take it at D 3 which is already given here.

(Refer Slide Time: 22:17)



We have calculated the value comes out to be 0.37 I am only taking 2 places of decimal and D 4 I am just hovering going up and down. So, again I am requesting please bear with me at the value of D 4 which is here I am considering a value of 1.51 with respect to 0. So, it is 1.513 with respect to 0. So, the net present value at C 2 which is this as I said its 0 into the values of 4 by 10 0.37 into the probability of 3 by 10 and 1.51 into the probability of 3 by 10. So, let us see it matches 0.4 which is 4 by 10 into 0 because here it was negative. So, the whole value was 0, 3 by 10 is it will be 0.367; 0.367 3.37. So, this is 0.3 by 10 into 0.37 and another thing arm the last arm is this arm is 1.51 into probability of 3 by 10 which is exactly here.

So, the actual value comes out to be 0.564 which is 0.57 so, at C 2 the value is 0.57. Hence, at the decision trees we have expected value D 1 1; D 1 1 is this one it comes out to 0.5 which you have already calculated. How do we calculate? Is half into minus 0.6 one-fourth into plus 0.8 one-fourth into 2.4 which is given 0.5 conducts a (Refer Time:

23:50) value is coming out to 0.544 544 would be this value is 0.564. So, I will take 0.54 and the other arm is 0, the 0 is basically do nothing in this arm. So, technically I have 0.5 for this drill if I conducts it miss test it will be; my another thing which I missed please bear with me I completely forgot. This 564 you have to subtract that value of 0.02 which is the sunk cost which you have made for drilling.

So, this value this is no its not an error it is it is the value which you have found out; this will help have 0.564 minus 0.02. This value when I find out this comes out to be 0.54 and this 0.02 is here is here I will just mark it with the black so, this was the value of his 0.02. So, once I have it I find out these values. So, based on the above (Refer Time: 25:12) I want to find out the set of decision strategies as follows for which you want to take.

Now, let us go one by one and through this stages. So, if I follow the stage D 1 2 C 2 1 D 2 2. So, let us see D 1 2 C 2 1, D 1 to C 2 1; C 2 1 would be this one and the path which 1 to D 2 2, D 2 2 will be this 1. So, if I consider the whole path the overall value the actual value would be basically 20,000 is the sunk cost which you have already taken.

So, if I follow and obviously, the probabilities would be given as 0.4 which would be given here; this is the probability which I have which is 0.4, 4 by 10; if I consider D 1 2 C 2 2 D 3 1 and C 4 1 D 3 1 C 4 1 D 3 1 and C 4 1 this arm. If I consider this the corresponding probability would be the multiplication of the probability which we will consider. The net present value would be given by 6,20,000 depending on the total amount of value which I have.

So, if I consider these values accordingly I can find out the table which will give me the overall path, their probabilities and the net present values and once we find out the net present values you can calculate it. So, there are a few just 1 or 2 steps of the calculation left; I will try to wrap it up in the 21st lecture. And then start off the new concepts as required.

With this I will end the 20th lecture and have a nice day. And, thank you for all your attention and stay well.

Thank you.