Data Analysis and Decision Making – II

Prof. Raghu Nandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur

Lecture – 01
Utility Analysis

A warm welcome to all my dear friends. A very good morning, good afternoon, good evening to all of my dear students and friends who are taking this DADM – II course which is Data Analysis and Decision Making, part – II. And, if you have gone through the introduction which basically covers the topics of nonparametric decision making and I would not go into the introduction once more, but just mention the few important points.

Those concepts will which will cover would be utility theory, then this concept of multi criteria decision making, multi attribute decision making, utility theory, then DEA which is data envelope analysis, AHP, ANP and then we will cover the concepts of TOPSIS, Elektra, Mac Beth all these methods then we go into the concepts of different type of metaheuristic techniques, simply without going to the mathematical details their algorithms and how we solve the problems and try to give examples for each of them.

And, switching back even though it would not be irrelevant event, but still I would like to mention is that, in the first part of the DADM – I, we have covered the bulk of statistics and in the third part of DADM – III which will be the third for in this series we will cover the different concept of optimization, operation research and such related topics.

So, once again a very warm welcome to all my dear friends and as you know this course total duration is 30 hours; 30 hours means 12 weeks which would basically be converted into 60 lectures. Each week will basically have 5 lectures and after each 5 lectures we will have an assignment. So, in totality will have 12 assignments and obviously, there would be one final examination and the final examination will consist of all the coverage which will go through lecture one to lecture and the last one which will be the twelfth one and assignments as you know it will basically cover the weeks lecture only and the marks and all these things would be collated and given accordingly.

So, my name is obviously, many of you may be doing this course or taking my course the second time, but still I will try to introduce myself. I am Raghu Nandan Sengupta from the IME department at IIT, Kanpur. So, again a very warm welcome to all of you and let us start the discussions about DADM – II.

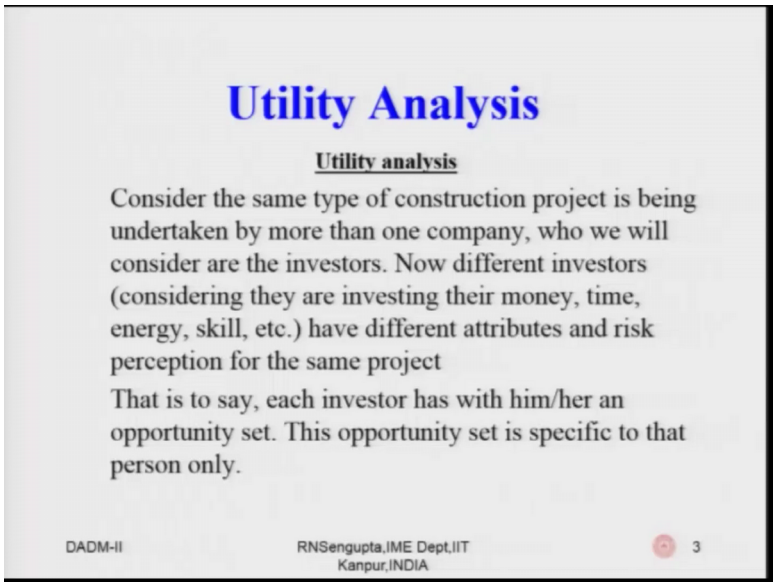
So, before we start DADM – II generally or the concepts of the nonparametric methods I will try to discuss the utility theory, try to wrap up the concepts of utility theory and how they are utilized, the concept of different type of decision making which you make, the concept of stochastic dominance concept, the first order, second order all these things would be utilized in a very simple manner, we will try to cover them as we proceed.

So, this is lecture – 1, and with without much ado let me start one. So, in utility theory what we mean by utility theory? Utility theory is some concept where we try to gain some value by taking a decision. So, the decision can be buying a car, can be decision may be say for example, I want to optimize on the cost it can be decision on trying to optimize the different trying to build up a dam and we have different resource constraints it can be man, material, manpower or time. So, you want to basically optimize on all the fronts.

So, if it is a single optimization problem we find it is easy to solve because you do not consider any other ideas to be optimized; whether maximum or minimum that is a different question, optimize means we are trying to get the best value. But, when it comes to the multi objective one then it becomes a problem because trying to optimize one obviously, does not give you the same picture as when you get try to get the concept of trying to optimize more than one or two or three. Because trying to optimize the objective function one basically may increase or decrease depending on what your idea is the second objective, while a second I am trying to basically optimize the second one you will basically try to increase or decrease the first objective. So, you have to basically make a compromise accordingly.

So, these ideas of utilities in a very simple sense would be utilized in this concept.

(Refer Slide Time: 05:08)



So, consider that we are in the project management phase and the same type of construction of project is being considered undertaken by some investors, some engineers, some construction company and we will consider them as investors because they are putting their money there. Now, different type of investors different type of people the different companies are investing their energy, skill, money, manpower and time everything and have different attributes or different type of qualities based on which they will try to analyze the project.

So, if I am analyzing a project say for example, of buying a car my buying the car may be different from second person who is buying the car because my main concern may be the cost. The second person main concern may be say for example, the space the car. Car has because he or she may have different big family, friends to help them in the transportation work or they want to have a pool car, where they want to go to office or it may be a person who wants to have his budget constraints or not there, his space constraints are not there, but he wants to buy an expensive car show to his friends or her friends that yes, I have a car. So, obviously, there are different attributes based on which a decision can be made.

Another decision can be say for example; I want to buy a house. So, when I am buying a house my main decision of buying the house may be obviously, is the cost, but other things can be how near is it to the bus stop, how near is it to the metro stop or how safe is

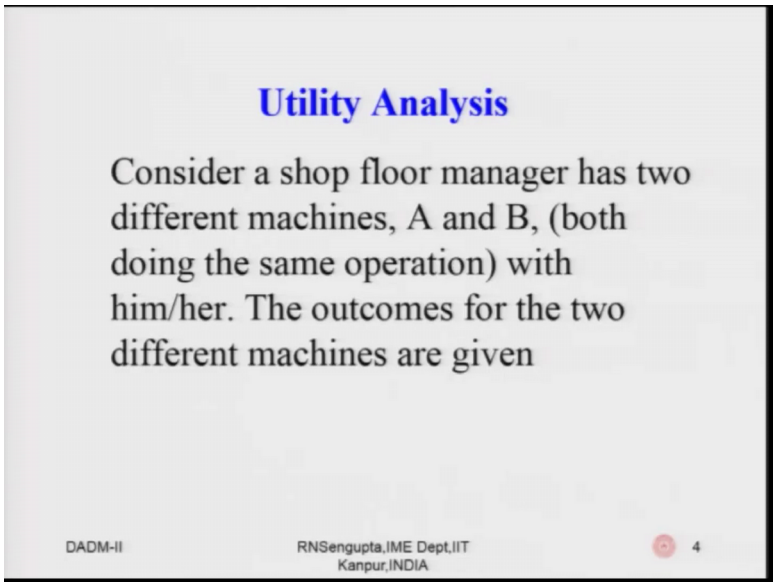
this locality where I will be staying or it can we say for example, how close is the market shop or how close is basically the hospital or maybe how close is this school. So, there would be different attributes different ideas based on which I will take the decision.

So, this in this example which you are considering or the discussion which you are having they have would have different attributes and different type of risk perception. So, the word risk I am using for the first time, but the risk would basically mean in a very simple sense, the overall loss the person is would be facing. It can be risk if you know from very simple statistical sense can be sigma or sigma square which is basically standard deviation or variance, in the concept of finance it can be beta and they can be different type of different risk concept also.

So, that is to say each of the investors or the companies we has with him or her a specific opportunity set. So, whatever the decision sets which you have for yourself would be different from me. So, we are basically calling an opportunity set based on which we will take the decision such that it meets our criteria of taking the best set of values for the opportunity set which basically optimize risk our main objective function which we are aiming at.

So, this opportunity set is set is specific to that person only. So, as I said it will change from person to person. So, if my opportunity set is basically the car ABC, now other opportunity set for that person may be say for example, car AB and D, C is not there for him or her. So, obviously, the opportunity sets are different and the criteria is based on which we are going to take the decision we means me when I am buying the car and the other person when he or she is buying the car would be different based on which they will take the decision to maximize or minimize their overall utility; utility or the value which you want to have by buying and that particular product.

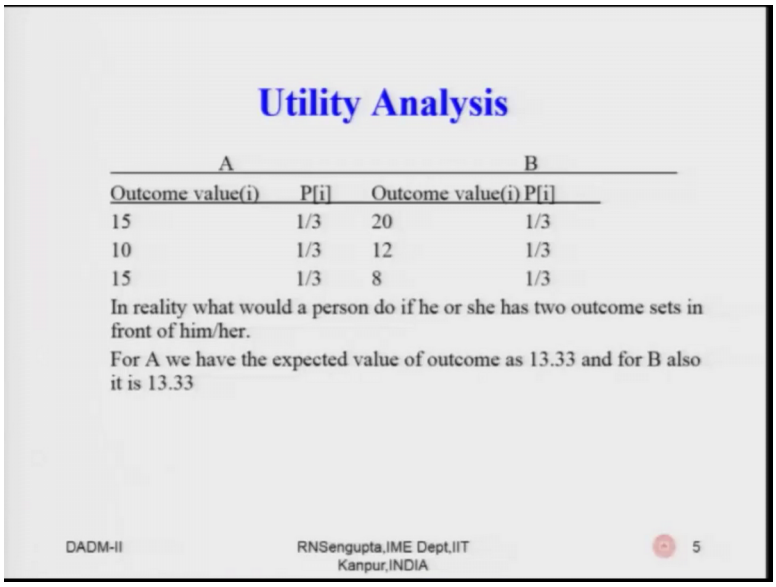
(Refer Slide Time: 08:52)



Consider the shop floor manager has two machines. So, let us discuss in a very very simple sense. Consider a shop has the machines A and B. So, both are doing some same type of operations it can be lathe machine, it can be CNC machine, it can be a turret, it can be a grinding machine, it can be any welding machine whatever it is.

And, the outcomes for this two different machines are given. So, outcomes means the amount of productivity or the amount of work the machines are able to do in the normal circumstances.

(Refer Slide Time: 09:28)



So, so, consider machine A has an output. So, the values which is which you see is so called in column well I will just highlight it with the pointer, on the first column and the third column these values I will come to this specific the reason how the values are calculate for the time being consider these values are given. Obviously, they would change from person to person, from machine to machine, from decision to decision, but why they change I am not going to consider it immediately.

So, consider the values are given and this the third the second column and the and the fourth column which is P i's are the corresponding probabilities. So, probabilities also would be dependent on the person, but for the time being we will consider them to be given again as I said for the as for the values. So, machine A has an outcome of 15. 15 can be a unit, can be in rupees, can be in yen, can be in dollars, can be in whatever Canadian dollars and UK pounds can be say for example, in liters, can be in grams, can be in kgs whatever it is.

So, the outcome values are given in and as 15 with probability one third, the next term of the outcome is 10 with the probability of one third and the last outcome is 15 again with the probability of one third. So, now, the question we would be asking that why there are two 15's we could have merge them and basically said the probability is two third the reason is that depending on the decisions whatever you take the values are coming. So, I may take decision one they are also the values 15, I may also take a decision 2, they are also the values is 15.

So, the different type of roots which I take or different type of decisions I take are different, but the end results are coming out to be 15, that is why basically two different rows for machining. Similarly, if you go to machine B there are values of 20, 12 and 8 with the corresponding probabilities again as one third one third one third which means all the probabilities are different. So, here the 20, 12 and 8 are difference; obviously, there would not be any confusion in the case if we had any confusion for two different 15's for a and then they being written indefinitely with the same probabilities a one third one third.

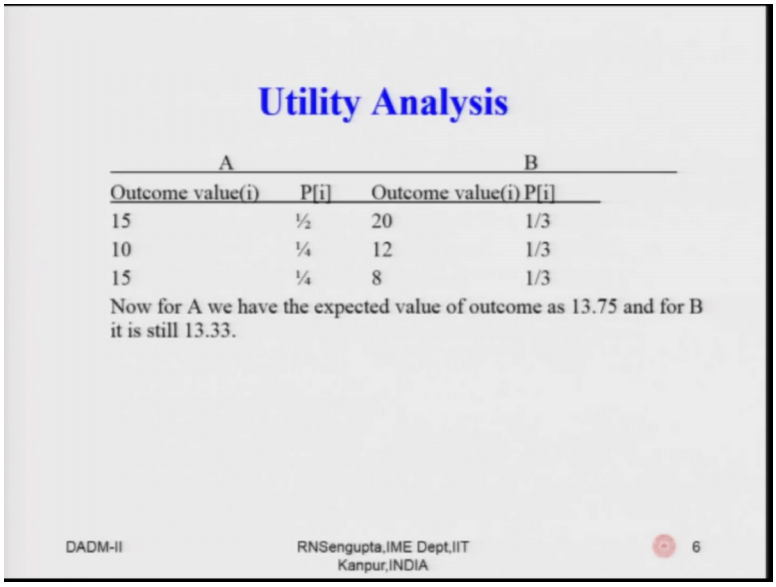
Now, in reality what would the person do if he or she has these two outcome sets that the machines which are producing and I want to find out. So, consider these 15's are values

in some units and they are net worth or the values which are cruise to that decision maker.

Now, if the person has to take a decision for machine A and machine B, for A we would basically have the expected value or the outcomes which will be given by 15 into one third plus 10 into one third plus 15 into one third, the value comes out to be 13 and 33, 13 and one third. And, for machine B it will be 20 into one third plus 12 into one third plus 8 into one third, again the value comes out to be 13.33. So, obviously, many would be thinking that we are take we that is very simple we are to basically taking the expected value and trying to basically find out which value gives us the maximum value.

Yeah, that is true, that is what we are trying to follow. But, obviously, we will see that there would be different connotations of trying to find out the expected value different connotations of trying to find out a different rule based on the fact that the expected values are also same. So, obviously, intuition would lead us to take a take us a different decisions based on the fact that the expected values are the same. We will consider that very soon.

(Refer Slide Time: 13:14)



Now, consider the scenario is a little bit changed, in the sense the values for machine A and B machine B remains the same which is for machine A it remains as 15, 10, 15 and for B it remains 20, 12, 8, but the corresponding probability for machine A has now changed. It is no more one third one third one third for all these three decisions or the

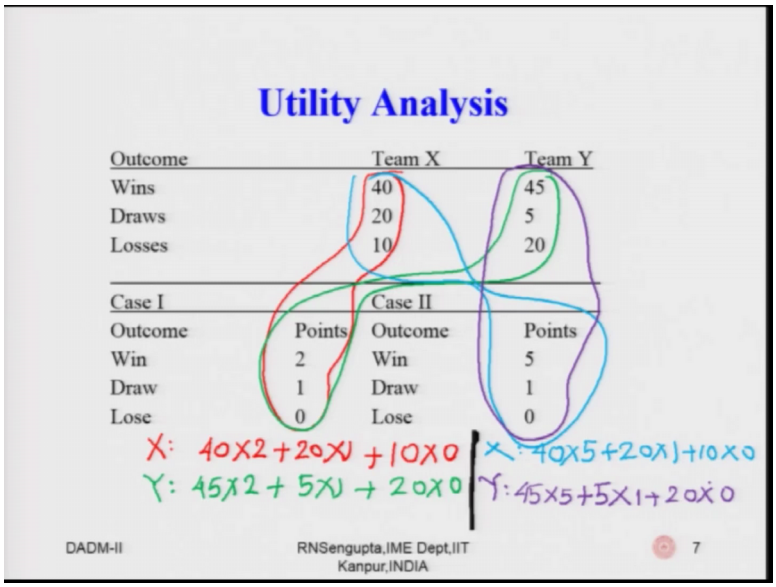
values, but now the problem is are 50 percent, 25 percent, 25 percent which is half one fourth one fourth respectively. While for machine B the probabilities remain as it is as one third one third one third.

Now, if you again recalculate the values comes out to be like this. Now, for A we have the expected value of the outcomes as 13.75, again we calculate 15 into half plus 10 into one fourth plus 15 into one fourth that value comes out to be 13.75 while for B the values have not changed. So, obviously, it will continue to be new minus 13.33.

Now, in this case if the question is again asked to you which decision would you take when you are considering machine Aa and machine B; that means, you can only buy one machine and you have to take a decision accordingly; obviously, your answer would be in the first case you are indecisive because both of them give the 13.33 values.

Obviously, no other criterias were being considered to take up the decision. While for this second example or the variant of the first example which is the one which is being shown on the slide you will definitely say that I will take the decision where I basically prefer machine A because the value it is coming out considering there are no other criterias based on which I can take the decision is 13.75. Hence I will basically take A over B for which b is 13.33.

(Refer Slide Time: 15:00)



Now, consider a different example. So, these are all just different above hypothetical examples which I am considering in order to make you understand. So, consider there is a match being played and out of all the teams I consider a say for example, X and Y and want to basically give you a notion that how the ranking system can change. So, obviously, that would have some implication that how utility can be utilized; in the since the concept of util can be utilized.

So, consider team X and team Y and in they play at the same type of tournament. But, in case 1 and case 2 which basically the overall points based on which the ranking is decided is different. Say for example, the organizers of the game before they start the tournament they would say that for each win there would be 2 points, for each draw there would be 1 point and in which loss they would be 0 point. So, the game continues and in this case all the teams play amongst themselves. There is a final, semi final, quarter final, semi final, final and now the rankings to be done ranking is to be done.

And, we find out the number of wins draws and losses for the teams X and Y respectively and these values are team X wins 40 draws, 20 loses, 10; while team Y wins 45, draws 5 and loses 20. So, in this case if you basically want to find out the ranking based on which you will basically put a X above Y or vice versa then what you will do? You simply multiply for team X it will be this one which I will just show in a very simple way.

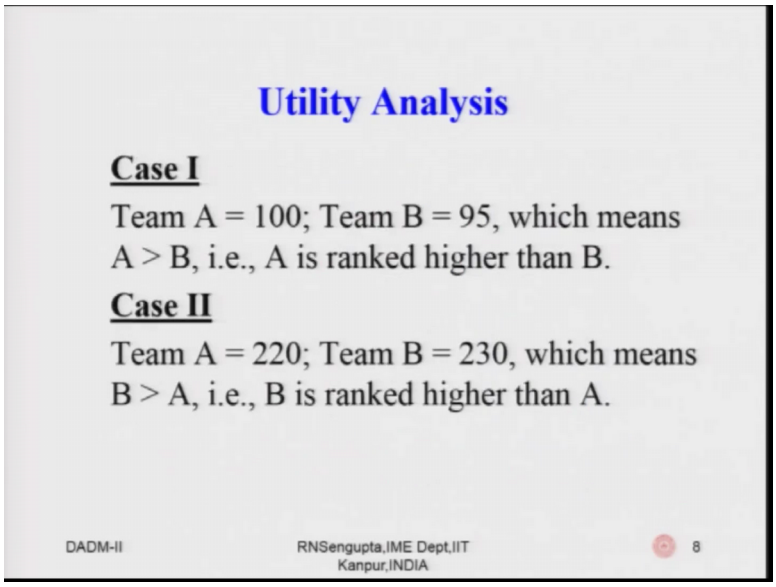
See you will for case 1, you will basically multiply for X it will be 40 into 2 plus 20 into 1 plus 10 into 0. So, that will be for X and for Y it will be 45 into 2. So, in the first set I multiplied these values. So, they are multiplying. So, let me change the color. So, it will be much easier for. So, in the second set what a dream in the second set the multiplication would be this. So, that would be Y 45 into 2 plus 5 into 1 plus 20 into 0. So, based on that you will basically do the ranking.

Now, suddenly as the game ends suddenly due to some error on the organized part they suddenly announced no; no that was not the case, the wins draws and losses would not be given as 2, 1, 0 has already been decided now the rank the point system has changed. So, now, and because say for example, hypothetically consider the sponsors who are sponsoring this football match or this hockey match whatever it is. They said no the ranking system would be based on a new scoring scale which is for a win up team will

get 5, for a draw they will get 1 and for a loss again it remains as 0. So, obviously, everybody's very very pert up. So, they have to again recalculate the rankings for all the teams and obviously, do it similarly for X and Y.

So, now what we will do? So, then I will come to the actual collections later on. So, in this case now team X would basically do the ranking. So, I am using a different color and. So, that would basically be yes for X to be 40 into 5 plus 20 into 1 plus 10 into 0. Similarly, when you come to team Y the calculations would be again I am trying to use a different color Y would be 45 into 5 plus 5 into 1 plus 20 into 0. So, when you do these calculations, you will basically have these values.

(Refer Slide Time: 20:11)



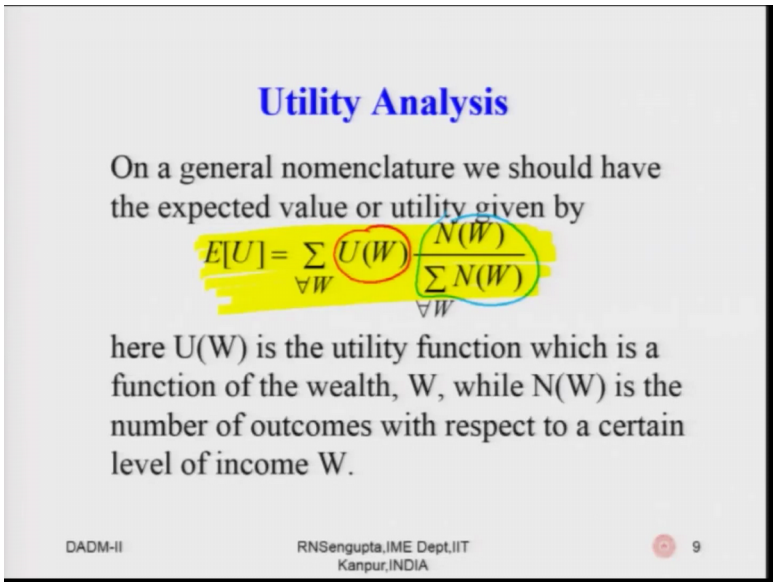
So, let us do the ranking based on when the organizers were had first case example of the ranking. So, team A would basically get a point of hundred team B would get a point at 95 which means that A would be declared considering all the other teams rankings are scores are less they will get a ranking of say for example, A being just the winner and B is say for example, the runners up. So, A would be ranked higher than B.

Now, as I said there is a confusion the sponsor says that they would not give the money if the rankings is not revaluated based on the points which I mentioned as you saw in the last slide was case 2. So, based on that, again that ranking is calculated recalculated. Now, whether A and B would be, but would be the first and second I am not going to comment on that because that would change because other teams would definitely get a

different set of ranking and some may basically it is possible some of the teams may be ranked higher than A or B, but I am only comparing sorry not A A and B, X X and Y. So, I am not comparing the overall first and set second I am only company the relative ranking of X and Y. So, A A and B.

So, teams A would basically have a point of 220 and team B would have a point of 230 which would mean because the calculation what I have shown in the last slide which would mean that B would be ranked higher than A. So, B is ranked and the ranking system has changed now no more A is higher than B, but B would be placed on a higher level than A.

(Refer Slide Time: 22:07)



Now, whenever we are doing this utility calculations if you remember I and I did mention and many of you must have got the hint is that we need to basically find out the expected value and based on that we proceed.

So, on a general nomenclature we would basically have the expected value would utility given by the multiplication of the utility function multiplied by its probability and then we will sum it up. Now, when I am saying basically the utility function and the corresponding probability they would mean the utility function based on the value which it occurs accrues to you are the basically the decision maker. So, obviously, what decision I take would be different from your decision because the utility function would change even if the values as are the same.

So, if you see the formula. So, if you see the formula the expected value of utility expected value is given by the multiplication of sum you will sum it up later on, but the multiplication of utility which is this one which is this one multiplied by the corresponding probability of that particular utility which is this one and if you remember the probabilities were given.

So, obviously, in this exam in this formula it is now clear that you will basically have a probability mass function of probability density function based on which you will try to follow the find out though the corresponding probability of P i and the UW which is the utility would basically be based on the overall net value which you would try to very consider try simply consider it is a type trying value you are trying to invest.

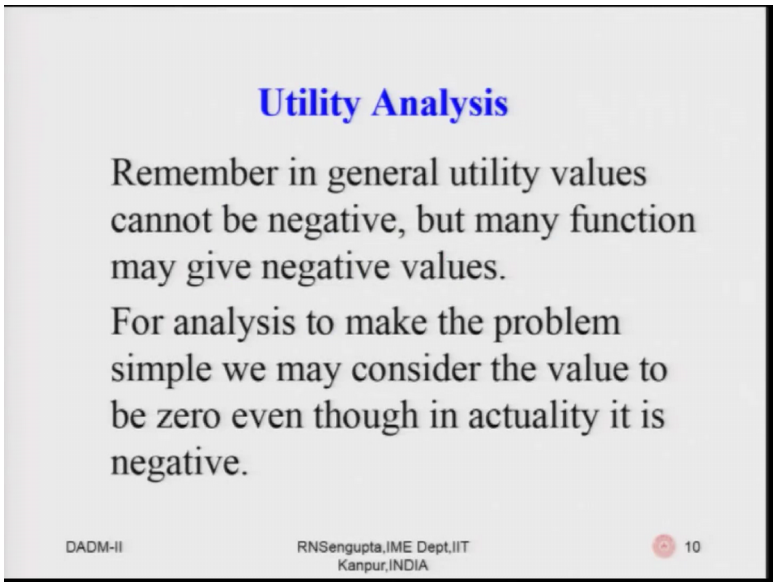
Now, the values which were given for the teams A and B or for machines A and B those values were assumed to be the ultimate values based on the utility which as the calculation has been done. So, obviously, there is an underlying utility function in which you will try to basically put your net worth based on which you are taking the decision and once you basically get the decision the values which it accrues to you is basically the utility based on which I will be doing all the calculations.

Here, UW is the utility function which is a function of the wealth, when we consider the wealth as W, now this wealth would be I am using a very generic symbol or a variable W it can be amount of money, it can be amount of material, it can be amount of a different type of manpower which way you to utilizing time can also be a factor, but we will consider in the very simple sense this is the wealth W.

While NW is basically N a function is basically the number of outcomes which basically accrues to that only decision. So, say for example, if I am taking a decision where the outputs are 15 if you see and the corresponding probabilities are given. So, the numerator in that probability is basically the NW which I am talking about and obviously, the denominator is the summation of all the NW or the all the different type of events which are occurring based on which I am taking I am trying to take a decision with for the utility.

Now, once you multiply each and every term you sum them up and the value which you get is basically the expert that we expected utility and as you remember I did mention and if you have seen in the slides you basically try to maximize the expected utility.

(Refer Slide Time: 25:58)



Now, remember in general utility functions cannot be negative, but we will consider negative values also. So, our neutral function negative means they are not worth taking the decision, so; obviously, we will try to shun away from the decisions. But, we will consider later on that there would be decisions where if two different negative values are there you will try to take the value which basically is least negative; least negative means it is more towards the post positive value.

But, in many of the cases depending on the decisions it is possible that you have to take a decision where the utility is basically negative. There would be other circumstances based on which are which would force you to take a decisions accordingly. But, many of the functions in the utilities may have negative values. In general depending on the problem we will ignore the negative values and consider them in to be 0 or we would not ignore the values and continue take their values with the negative sign and proceed with the calculation.

Now, if you if you remember the last slide what I showed it basically summation of the utility which we utility as I mentioned is a function and my utility being multiplied utility function be multiplied by the ratio which is NW by summation of m m NW. So, what that, this is basically a chance or probability or a relative frequency? So, obviously, the question would come from your side.

So, if there is a chance this is a probability does it have a probability distribution? My answer would be yes, we will try to utilize this probability distribution in order to calculate the expected value of the utility. So, what do you want to find out it basically the expected value utility where you consider the utility as a random variable and basically multiply that with it is probability mass function or probability density function and for try to find out the overall expected value of the utility. So, for analysis to make the problem simple we will consider the value to be zero even though the in actuality the value may not be zero.

So, with this I will conclude this first lecture for DADM – I DADM – II sorry, and continuing the discussion of utility 3 with more examples later on.

Thank you and have a nice day.