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

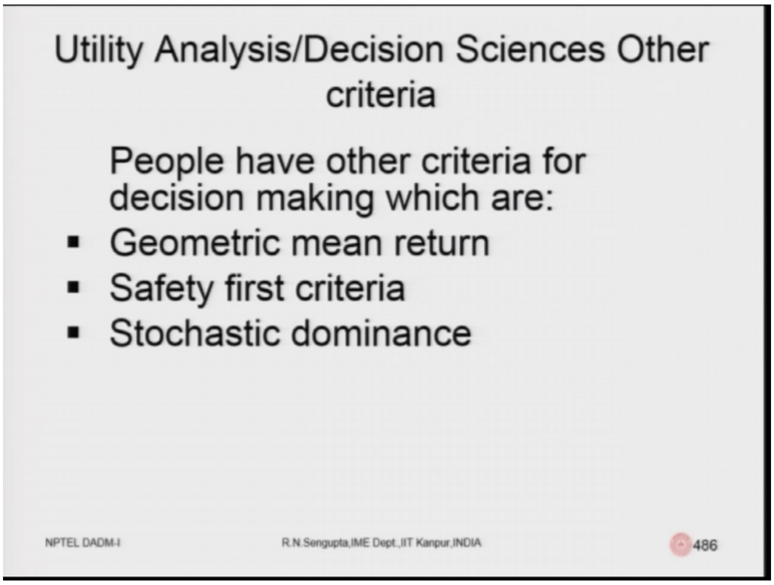
Lecture – 49
Utility Analysis

A welcome back my dear friends, a very good morning good afternoon and good evening to all of you and this is the DADM which is Data Analysis and Decision Making course under the NPTEL Mooc series. And if we know as you know this is the course duration is basically for 12 weeks which is 30 hours; 30 hours basically gets converted into 60 lectures because each lecture is for half an hour. And each week we have 5 lectures which is half an hour each and we are in the 49 lecture which is basically the last, but one lecture for the 10 th week.

Now, if you remember in the last lecture we were basically discussing about the concept of equivalence, that given decision deterministic one you want to basically balance it with the probabilistic one. And I did use the word of fair gamble, but it can be need not be a fair gamble where the corresponding properties can be p and 1 minus p. But, what I actually meant was that the expected value of the deterministic one which is technically the certainty value based on which we will try do the calculation and break them into the gamble, would be such that the expected value of the of the gamble would also be equivalent to this.

That means, if the certainty value c then the u of c multiplied by 1, 1 is the probability will be equal to u of w 1 into p 1 plus u of w 2 into 1 minus p 1 because, 1 minus p 1 and p 1 are the probabilities. And we can continue doing it and break into more gambles or reduce them with respect to the expected value and discuss it further.

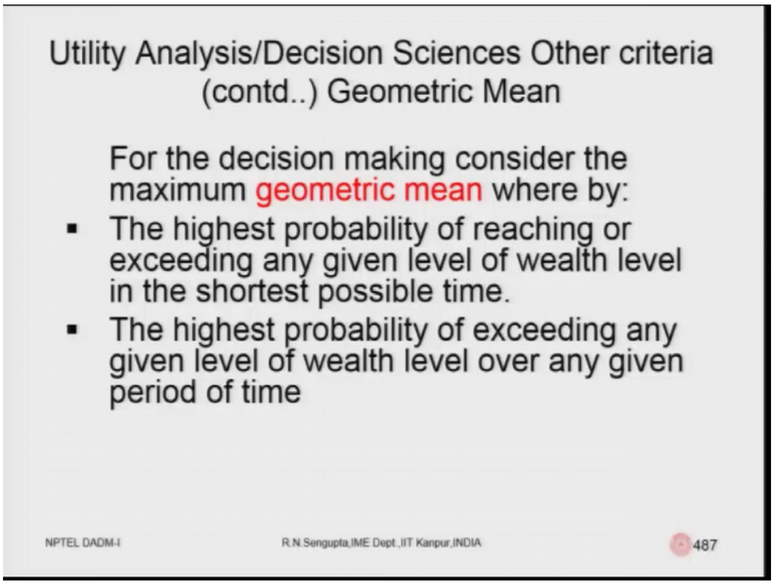
(Refer Slide Time: 01:57)



Now, further considering the discussions utility or under or decision sciences the other criteria's which can be utilized for ranking are basically geometric mean safety first criteria and stochastic dominance, so I will go through them slowly. Now, we generally know for any general any to find out the central characteristics of a distribution, we generally use the mean the median and the mode in another mean we have the geometric mean the arithmetic mean and the harmonic mean.

So, you are trying to basically use the geometric mean for as a tool in order to find out the utilities acquiring from any decisions. So, what we mean by geometric mean what we mean by safety criteria stochastic dominance I will basically discuss.

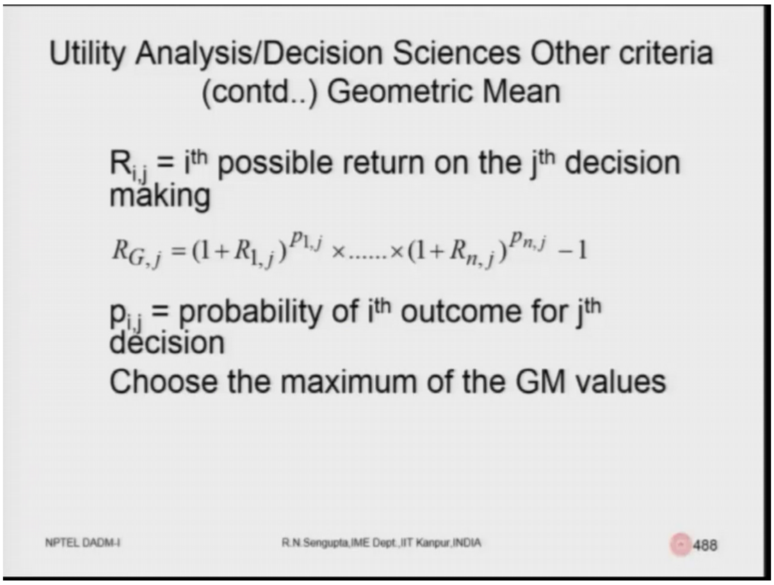
(Refer Slide Time: 02:47)



Now, for the geometric mean for the decision maker for the decision makers, consider the maximum geometric mean is the case when the highest probability of reaching or exceeding any given level of wealth is possible in the shortest period of time. So, we will basically try to utilize that value of the of the functional form of the mean which is the geometric mean, such that it will give us the highest utility or the highest expected value of the utility.

For the decision making considering the maximum geometric means would also mean that the highest probability of exceeding any given level of level over any given period of time would also be possible. So, I am basically trying to utilize the concept of geometric mean, in order to consider that what is the highest probability of reaching or exceeding that and the highest probability of exceeding any given level of wealth. So, there is one given level of wealth and what is the what is the highest probability over and above of that and another case is basically what is the highest value attainable based on this we will take the decision.

(Refer Slide Time: 03:55)



So, for the utility analysis or the decisions, so consider R ij is the ith possible return on the jth decision and the probabilities are given as R and the returns is basically given by i R suffix j G. So, G is basically returns based on geometric point and the js are basically the suffixes based on which here I am trying to consider. So, that is the decision making for which number and ith is basically possible set of such outcomes inside one decision.

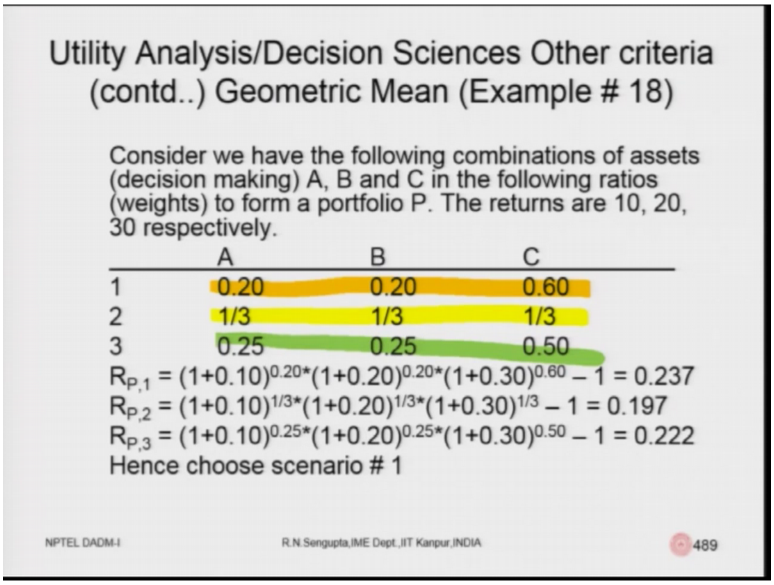
So, consider there would be j such decisions each would basically have i of them under each R of this so called decision making. Now, I will consider the values of corresponding to the first second third fourth returns for say for example, the first decision making second decision making so on and so forth accordingly. So, what I need to find out is basically the probabilities.

So, I will basically add up 1 plus R1 into j into p 1 into multiply it dot till one plus R a into j into p n minus n comma j minus 1 because, I am trying to basically find out the probability accordingly. So, p i js would be the probabilities based on which I will do.

So, for each decision the probabilities would basically sum up to 1. So, if I am considering say for example the jth decision or the first decision the corresponding probabilities would be p 11 for the first case p 11 p 1 2 no p it would be p 11 then p 2 1 and p 3 1 p p 4 1 so on and so forth.

Where the first suffix would basically mean the possible returns under that decision making and one the second one would basically imply the decision making which is there. We will choose the maximum probability or the geometric mean value based on which the calculation can be done.

(Refer Slide Time: 06:07)



Now, consider we have the following combinations of acids of decision making which are considered as A B C in the following ratios or the or the weights to form a portfolio the returns are 10 20 and 30 respectively and in this case this 1 2 3 which you have are the decisions which is j is equal to 1 j is equal to 2 and j is equal to 3.

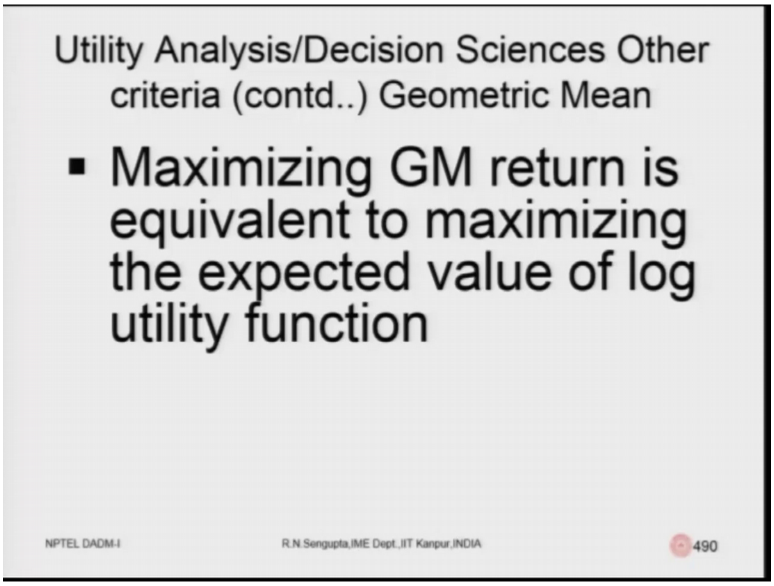
So, under j is equal to one you will basically have A B C combined in some proportions under j is equal to 2 again you will basically have A B C combined in certain proportions, under say of example 3 you will basically have A B C combined in certain portions.

So, the proportions based on which you are trying to do the calculation for 1 2 3 are basically given a long the row. So, it is 20 percent 20 percent 60 percent for A B C under decision 1, one-third one-third one-third for A B C under decision 2 and 0.25 0.25 and 0.5 which is 25 percent 25 percent 50 percent for A B C under decision 3.

So, the corresponding this values of the returns for the geometric mean would be given as per the norm when you are trying to find out. So, if you remember the returns for given so these are basically the corresponding probabilities. So, probabilities are given as 20 percent 20 percent and 60 percent 33.33 or one-third one-third one-third and the last case being 25 to 45 and 50.

So, these values are coming as 0.2 0.2 0.6 for the 1st case one-third one-third one-third for the 2nd case and 0.25 0.25 and 0.5 for the 3rd case. Now, the returns are basically given for the case of the A B C as 10 percent 20 percent and 30 percent, so when you utilize that the values which you have found out for R p 1 is 0.237 R p 2 which is 0.197 and R p 3 is 0.222 based on that you will be able to rank them accordingly to find out what is the best decision or the choice based on the fact that you are trying to take the geometric mean of the returns for your decision. Now, we basically consider the utility analysis decision sciences other criteria's which is the geometric mean.

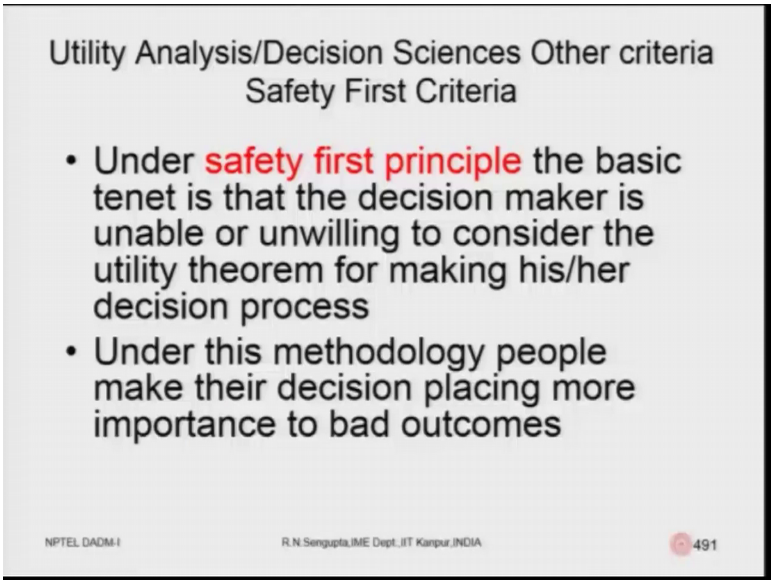
(Refer Slide Time: 08:45)



So, maximizing geometric mean returns is equivalent to maximizing the expected value or the log utility function. So, when you are taking the log utility function in the same case if you remember we have considered the quadratic utility function and the returns to be normal and they were basically equivalent (Refer Time: 09:55) of a we are trying to basically maximize the geometric returns it will be equivalent to maximizing the expected value of the log utility functions.

So, hence if we are trying to take a decision trying to maximize the geometric return in one hand and trying to basically maximize the expected value of the log utility function they would basically give you the same result that is an if and only if condition. So, if you use criteria one to buy to select and then again to use criteria 2 to select, where criteria 1 is maximizing the geometric return mean return and criteria 2 is basically maximizing the expected value of the log utility function, so both of them would basically give you the same result. In the similar way when you use the returns to be normal or trying to basically use the quadratic utility functions in both ways you basically calculate you will get the same result.

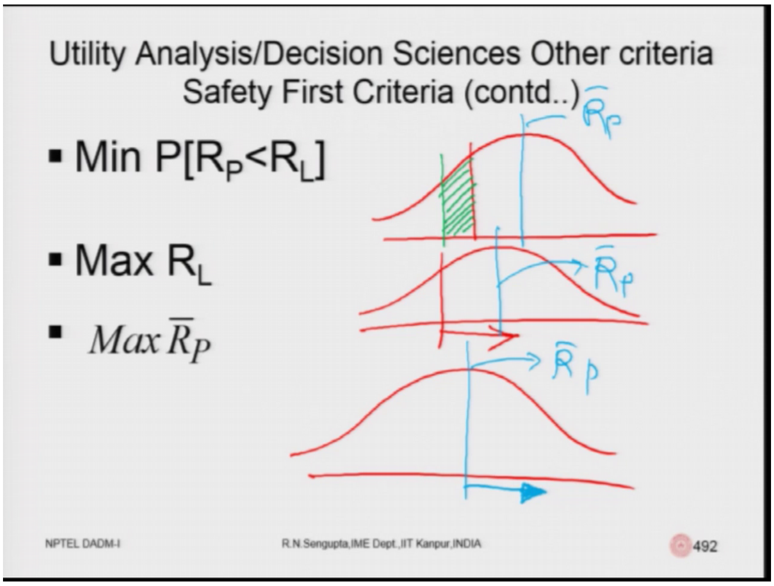
(Refer Slide Time: 09:55)



So, under the first principle safety principle the tenets is that the decision maker is unable or un willing to consider the utility theory for making his or her decision process. So, here I am basically trying to analyze the decision making, under the methodologies for which the person is considering the utility theory is such that it is a safety first criteria comes into the decision making process.

So, in this safety first principle criteria the analysis is like this. In the decision making is because here the utility function will be such that the person considering the utility function would always be very risk averse. In the sense that he is trying to basically take a decision such that he wants to avoid the risk on trying to make the minimum loss and under this methodology people will make the decision placing more importance to the bad outcomes, that means probabilities may be the same for the good and the bad, but they will try to basically analyze the decisions. Where bad decisions are important for them based on which they are going to take the decisions which will be effective for them.

(Refer Slide Time: 11:19)



So, when we are considering the safety first criteria I did mention the concept that what is R P RL and R p bar. So, I will basically go into details considering you have a decision where the decision is not only of 1 criteria or 1 only 1 set. Consider you are taking a decision and in the sense that the decision is basically for buying the set of stocks or buying a set of projects whatever it is or investing in projects. Then R P would be the basically the so called portfolio being formed from this decisions and you want the probability of portfolio of this decisions to be as less as possible then a fixed values which you have.

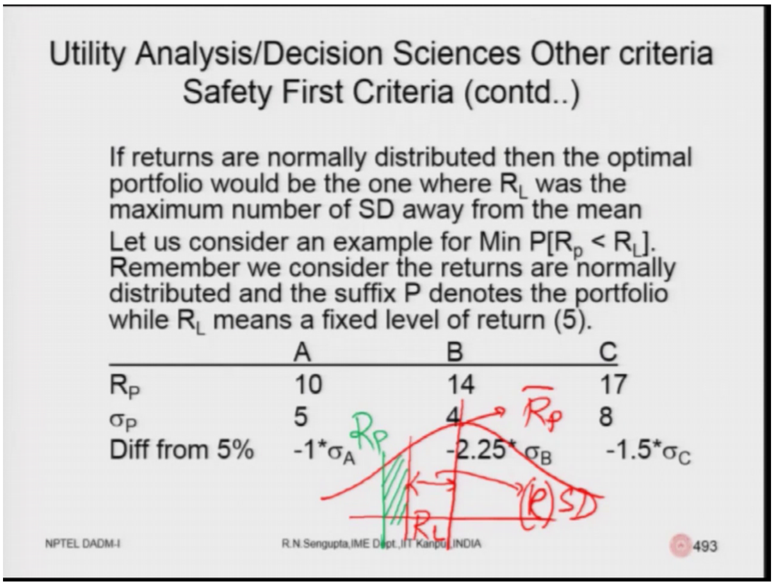
So, let me draw the diagrams for the first one, so it will be in the case if this is R 1 and consider if R P falls on to the left hand side consider this sorry let me use a different color. So, it falls on the left hand side you want to basically minimize this probability that is all, because less the probability better if it is on the right hand side, obviously you will try to maximize it push it as far as possible to the right.

In the case of maximizing RL what we are doing is this is RL, push RL to the right hand side as far as possible such that automatically the concept of trying to minimize the distance between R P and RL onto the left side, would be such that R p would also be automatically stretched on or pulled on to the right hand side.

So, in the sense that we are trying to basically maximize our overall return and finally when I taking the R P value with a bar which is the expected value what is here is basically this which I would I did not draw. But, I will mark it here with blue colors so this is the average value this is the average value.

So, what is happening this is the for the third diagram this is the average value. So, you are trying to basically push it on to the right hand side. So, one case you are trying to minimize the area between RL and R P, where R P is on the left hand side the second case is basically trying to push RL on more on to the right. So, that automatically R p also moves and in the third instances basically you are trying to basically move the average value such that the whole distribution shifts on to the right, because higher returns are on the right hand value.

(Refer Slide Time: 14:13)



So, consider that if the returns are normally distributed I am considering returns are normal remind that, then the after one portfolio of the decisions who would be the one where RL or some threshold was the maximum number of standard deviations which is away from the mean value. So, what you are trying to basically consider is that if the distribution is there with an average value R p bar and RL is something where some value on to the left hand side, we I want to find out that what how many such standard deviations RL is form the mean value.

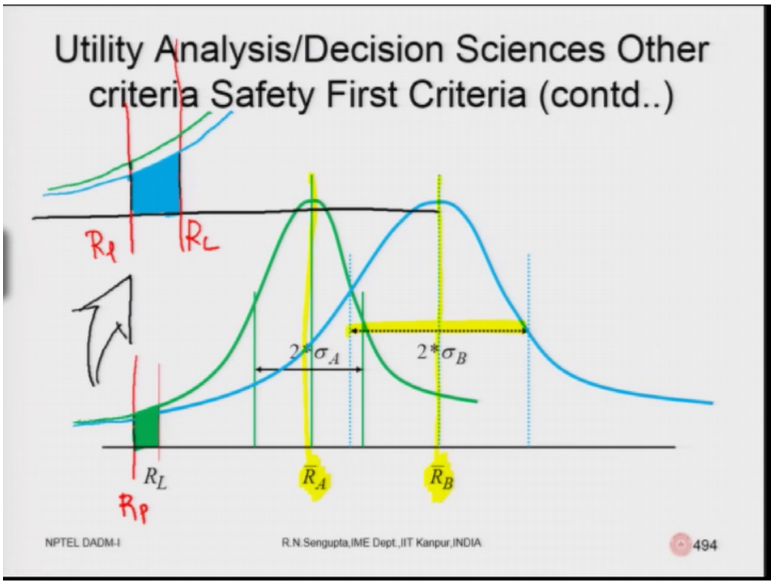
So, if I am able to draw I will draw it here. So, even though it is a there are lot of information but I will draw it here, so it is easy for. So, this is the value so this is the R r R average R p value and this is R r. So, I want to find out this is how many such standard deviations, standard deviations into k. So, this k is the number of standard deviations which is there on the left hand side.

So, let us consider an example for the minimization of probability of R P less than R L. So, this R P value which I said was on the right hand side left hand side you want to minimize this right area. So, this is RL this is R P this is R p bar. So, let us consider an example for which we try to minimize probability of R p less than RL, remember we consider the returns are normally distributed and the suffix P denotes the portfolio where RL means the fixed level of returns which we have kept are 5.

So, some threshold values so with R P value being of A B and C being 10 14 and 17 the standard deviations are given for A B C as 5 4 8, then the difference from the 5 percent would be for the case when you are trying to analyze RL which is 5; 5 would be one standard deviation would be.

So, on the left so it will 2.25 standard deviation and it would be one point 5 standard deviation with respect to the values of A B C, when we combine them obviously the portfolio will have a different mean and different standard deviation. So, consider these are investments only one asset or they are a some portfolios which we have renaming them as A B C, so that does not matter basically I am trying to get the concept clear.

(Refer Slide Time: 16:59)



So, when we analyze the safety first criteria, so these are the distributions which we have those blue color distributions are normal one, again the color green color is the normal one. Only remember one thing for both the cases the mean values are different which is fine because, for the green one returns are on an average lower and for the blue one returns are on an average higher. So, hence I will use the highlighter, so this is the average value which is R bar A this is the average value which is R bar B.

Now, what is important is to note down that the standard deviations are different also. So, for B it is much for flatter, so hence that plus and minus standard deviation the overall but the area covered would be the same if we convert them to a standard normalized value. So, this is the overall area or the values from minus sigma minus sigma plus to plus sigma.

So, which is for sigma corresponding to B the B distributions because that is what the distribution has been drawn for the blue one. And for the red one it is basically minus sigma A to plus sigma A and obviously it is seems to be narrow because it is much more compressed. But, when I am trying to find out the corresponding value how far it is, I would obviously consider the consider that the R P value let me draw it R P somewhere here R P value is somewhere here.

So, what I will be interested would be like this, if I am considering the case of the green one. So, let me blue one so let me use the blue color it will be easier. So, I will I will I will draw it accordingly so, now let me draw it and try to analyze the areas probabilities, if I consider the blue part I will write 2.

So, I will be most interested to find this area because, that is the area between the RL value and R P this is R P. So, I can I should write it down for better clarification for all so this is RP. So, this blue area is the we want to minimize the and when I want to find out I have removed it, so it is say for me to draw when I want to find out with respect to the green one, because the distributions are same for 2 different portfolios decisions.

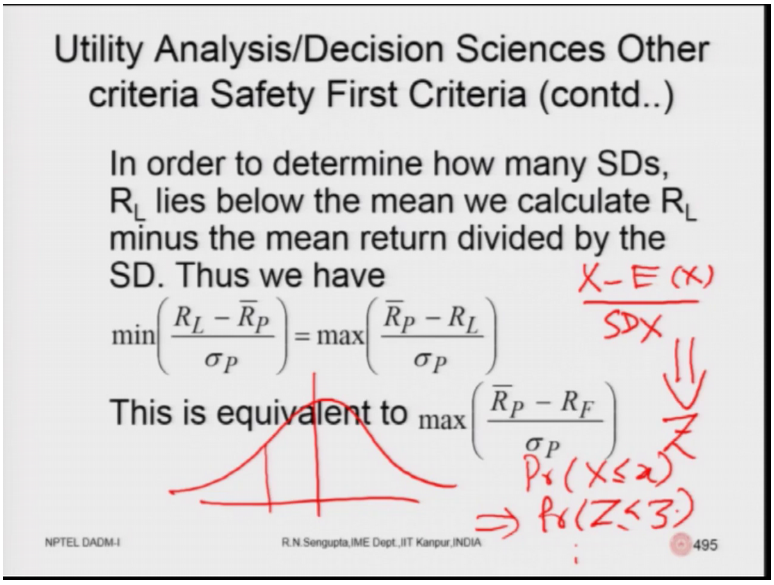
So, it will be now be this area, so the areas would be different and minimization would happen based on the area being covered. So, if I expand this I think I will be able to give a better picture let me try; let me try to draw.

So, this would be the first is theirs the black line which is this one, then is the RL value and is the blue line I am just put drawing this left portion. Then is just the green line then is the RL value, so what I have drawn is this.

Now, expanded this in order to make you understand I will mark these values, so if you pay attention you can understand. So, this will be R l this will be R P, so when I want to find out the area the green portion would be first I will draw the green portion.

So, this is will be the so you press the green one, so corresponding to green graph of the normal distribution. So, let me first erase it done then I want to basically do it for the blue one. So, the blue one would be this so obviously those areas would be different. So, this is just a way of trying to analyze the decisions so you can basically draw it accordingly.

(Refer Slide Time: 22:49)

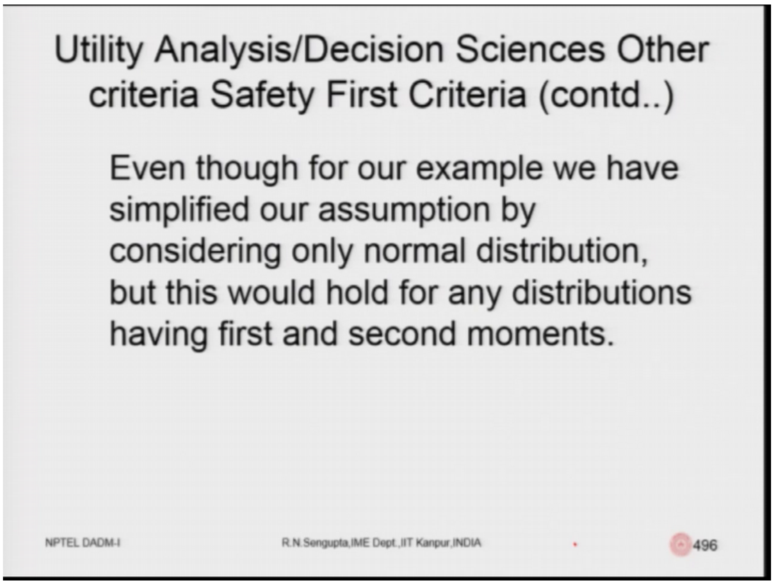


Now, in order to determine how many standard deviations RL lies below the mean value which is R p bar, we locate RL minus the standard deviation divided by the by the standard which we are trying to basically normalize them normalize them. If you remember you always use this formula, if it is X the random variable x minus X vector value of X by the standard deviation X that would be normalized and give you the capital Z which is the standard normal deviate.

Thus we want to minimize RL less than R P being less than RL, so obviously this is being normalized RL minus R P is the mean value of the return divide by sigma P, so that would basically also give me my maximization. So, if I basically take the minus sign outside it will be maximization of R p bar minus RL divided by sigma P.

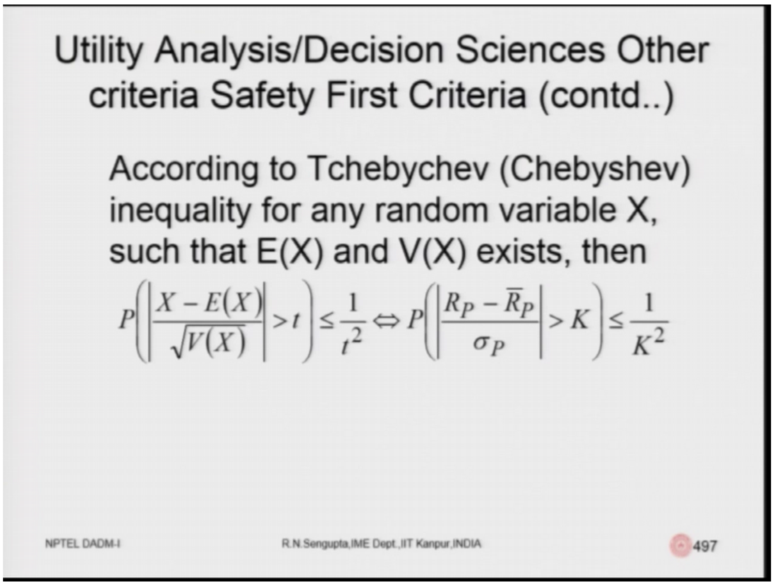
So, this would be equivalent to the case that if we replace R L by the risk free interest rate from the bank it will be R T bar minus R F divided by sigma p. So, what we are doing is that we are considering the normal distribution and trying to find out what is what is the overall area which is being covered. So, this my apologies so initially it was this then it gets converted into, so this is capital X 1 x capital Z small z so they are just transformation.

(Refer Slide Time: 24:37)



So, even though for an example we have simplified assumption by considering only normal distribution, but this was whole true for any distribution having the first and second moments accordingly, so we can solve the problems accordingly.

(Refer Slide Time: 24:53)



Now, according to Chebychevs inequality, so if Chebychevs and (Refer Time: 25:00) can be also utilized I will consider that in more details later on. So, according to Chebychevs inequality for any random variable having expected value of E X and V X, we can find out the corresponding amount of deviations which you will have or what is the maximum amount of deviation which you will have, so based on that you can take the decisions likewise.

I will explain that in the later task, so with this I will end this 49 lecture and considering more details further on. I know that I am basically finishing a little bit fast but only for this portion because if I try to basically consider that it will take a little bit of more discussions which I will try to basically keep it in the last class for the 10 th lecture.

So, it will be much more easier for me to explain and go slowly such that you also have a feel at in at in sync into the discussions and hopefully by to the 50th lecture we will finish off the decisions sciences. And continue start continuing the relevance of this would come out later on when you consider the multivariate distribution on the other cases.

Have a nice day and thank you very much for your attention.