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

Lecture – 50
Chebyshev‘s Inequality

A welcome back my dear friends and students, A very good morning, good afternoon and good evening to all of you. And as you know this is the Data Analysis and Decision Making course under the NPTELMOOC series and for the total course duration is 30 hours which is basically 60 lectures, each lecture being for half an hour and the total week duration is 12 weeks.

And this is the 50th lecture which is means we are ending the 11th week lecture so obviously, you would after this you will take the 11th week assignment. And each week as you know there are 5 lectures each being for half an hour and my name is Raghu Nandan Sengupta from the IME department IIT Kanpur.

Now, we were discussing about this in concepts of utility in decisions and why utility considering the distributions were important. Initially we considered the concepts of utility in the framework that the distribution of the utility function plus the functional form of the utility function, were important in order to understand, the way you model it that, whatever that the utility was. And our main concern was to basically find out the expected value of the utility based on which you will take a decision whether, you will rank the utilities from the highest to the lowest or basically, it rank the utility from the lowest to highest considering that you want to find the variance.

Now; obviously, the question will come that why variance because, if the utilities are same then; obviously, you will try to basically find out the utility for which the variance is the least or else another way can be when you normalize them, how do you normalize them? This I am just mentioning for your own from information. I am sure you would have understood by this time considering that you have already done about three to four lectures of utility theory.

And the reason why you rank them from the lowest to the highest is basically say for example, you are trying to normalize them, as I said you are you are trying to take the

ratio of the variance to expected value. So, the normalization would be done. So that means, per unit returns what is the variance, so that you would basically rank them from the lowest to the highest.

In other case, you want to find out the expected value or the returns per unit of variance then, then you will find out the ratio of the expected value with the variance and rank them from the highest to the lowest. Later on we saw that different utilities gives different rankings because, that was a perception based on which for decisions like say for the football match if you remember, there were different points given for the different tournaments and the ranking changed.

Then, obviously also the concept of functional form of the utility whether, it is linear, whether it is quadratic would have an effect the how the ranking is done based on the fact you are trying to find out the expected value. Then we also saw that given the utility to be fixed, if that outcomes are different. And they would also give us different answers because at the end of the day you remember, the expected value of the utility was basically calculated as integration or summation of the utility function, which is u w into the pdf or the probability of the utility function.

So, our probabilities were basically based on outcomes and the utilities were basically the utility function based on which we are doing the work. And obviously, to mean that one person would have one certain utility, the other person will have a certain different utility for the same situation. And obviously, the probabilities would also change accordingly.

So, generally there are 2 important facts as I am mentioned in again, one is the utility, one is the probability distribution or the outcomes distribution based on the utility function. Then, we consider the concept of certainty value, certainty value being certain value where, the expected value of the certainty value considering the probabilities 1 would be equal to the expected value of the fair gamble.

The fair gamble is where the probabilities are half and half and you basically have the outcomes of w 1 w 2 in the wealth perspective or the units are not important, based on w 1 w 2, I find out the utility of u w 1 using the functional form of u w 1 and using the functional of u w 2 based on I do the calculation and I find out what is the expected value of that fair gamble.

We also considered, I am going to the recap please bear with me. So, we consider the concept of non satiation, non satiation means that more I give you, more you want, so obviously, it would mean that the first derivative of the utility function would always positive. And then the point came up it will well that that is fine but how do we basically categorized human beings based on the perception of risk which may be, I am a risk lover, I am a risk in different person, I am a risk hater.

So, for to analyze that the very best methodology which I think would be best for you to understand is to consider how the graph for the utility is increasing. So, if it is increasing under increasing rate or is it increasing at a constant rate or is increasing under decreasing rate; if it is increasing under increasing rate, obviously that would mean the second derivative is positive, it is increasing at a constant rate it will mean the second derivative is 0 and if it is increasing at a decreasing rate it will mean the second derivative is a negative.

Now, based on that we later on proceeded to find out that what if we want to find out the actual characteristics of the utility function is it possible; the answer was yes, we use the concept of absolute risk aversion property and relative risk aversion property. Absolute risk aversion property and the relative risk aversion property was basically in the absolute sense, how we love or hate or are indifferent to risk. In the relative risk aversion property also in the relative sense you try to analyze, how we love risk, how we are indifferent to risk, how we hate risk.

Then based on r a and r, which is absolute risk aversion property and relative risk aversion property we found out the concept of what is a prime and r prime, which is the first derivative of a and r with respect to w, w is the wealth. And then, we considered the four different type of utility functions. Utility functions was quadratic utility functions were logarithmic, utility functions was exponential and utility function was power utility.

Then, based on the mathematical formulation we did we first found out the first derivative, found out the second derivative, use the first derivative and second derivative to find out a and r. Then, again we differentiate it a with respect to w and r with respect to w and found out a prime and r prime.

Now, as we found out a prime and r prime, based on that, we can basically find out what is the characteristics of the utility function, whether it has positive absolute risk aversion

property or whether it has 0 risk relative risk aversion property, which mean means that a prime can be greater than 0 can be equal to 0, can be less than 0 similarly, r prime can be greater than 0, equal to 0 and less than 0.

So, based on any of the categories we can basically club these four different type of utility functions into one of the characteristics based on which we can see qualitatively. Then we basically, took the simple example of a thought out experiment, using excel sheet and I am please request all of you to please take out your time and spend about 5 to 6 minutes in drawing to basically mark those values.

So, those arbitrary values which I have taken, you can do that is also or you can change the values accordingly, what you think is more easy for you to draw because, that I am trying to understand the diagram would be important. You basically plotted, you took some arbitrary values of w based on the utility function which was quadratic, which was exponential, which was logarithmic and which was power.

We found out the value of that utility function u. Next, which is the second and the third and the 4th column, which was basically omitted u form, we should find out u prime and r e u double prime. Based on once u prime and u double prime are found out you will basically find out a, then a prime next r and r prime.

So, when r a prime and r prime are found out, check the values what is this general functional form of those and check what is the derivation which you have got, based on the fact that the utility function is given those four utility functions based on this, we are able to find out a prime and r prime using the mathematical formula.

So, be rest assured, the way we have I have basically analyzed this problems in the excel sheet using a very the simple hypothetical example, the answers which you will get using the mathematical formula and using the excel sheet simple. Trying to basically draw the graphs and in the excel sheet would give you the same concept they would be double verified.

Then, we considered that in case we have the utility function, which is quadratic then, there is a 1 to 1 correspondence with the quadratic utility function and normality of returns. So, it means that, if utility function was a quadratic then returns would be normal and vice versa. So, this would be an if and only if condition.

Then and we showed that that considering the investments if, we basically rang this investment based on the maximization of the expected value of the utility and same if we consider that we have the returns are normal and basically try to find out the probability of the normality value, greater or less than some stipulated value. And then, we use the standard normal tables based on that, when we find out we found out that the results which we get from the case of for the case of the utilization of the values, which we have been done for the case of the utility function they would basically match.

So, in when I am saying they would match yeah, so the values which you got from the utility function using the mathematical formulation, the values which you would use u t is using the excel sheet, they would definitely match. And then, we will basically try to find out that, the utility function as I am was mentioning for the normality one and for the quadratic utility function on the normality of the details they would match because, when we rank the investment based on the quadratic utility function and then we separately rank them using normality returns, normality returns was basically for the case when we normalize the returns to normal case, there is standard normal and (Refer Time: 11:30) of the property less than or equal to or greater than equal to a stipulated value. And we found the ranking remains the same, so there is no change.

Then we consider the different axioms of probability so that means, the ranking which we do that, a is greater than b and b is greater than c then obviously, (Refer Slide Time: 11:48) the ranking which we do between a and c would also basically give us the same picture; that is a is greater than b and b is greater than c.

Then later on we also consider that if utility function is generally not given, how we are able to find out what type of utility function you are. So, for that also we took a again a is very simple example and this is the third time example is explaining there, so please try to understand. I am sure you would have what if it is there is no harm if I repeat it.

So, along the x axis, I basically draw the values of w and I take 2 arbitrary values of a and b and for a fair gamble the probabilities are half and half. And if it is an half and half and the utility function is basically linear in the sense that, m u of a is a and u of b is b which means, the utility function u w is w, then the expected value of that fair gamble would be a plus b by 2.

Then also we drew the 45 degrees line, based on the 45 degrees line based on the 45 degrees line we are able to find out the certainty value and if this certainty value falls along the on the 45 degrees line then, (Refer Time: 13:00) that the marginal rates of the utility function that is derivative first, derivative of the utility functions would be positive, but the second derivative of the utility function would be 0.

That means, constant relative and constant absolute relative risk aversion property would hold. And in case the values of the values which you plot, I am not drawing it, but I am trying to explain, I am sure you will have understood what I am what the diagram.

So, if the values of the plots based on when the expected value of the utility and the certainty values are below the 45 degrees line or above the 45 degrees line, you will basically have the concept that of double derivative being greater than 0 and in the second case the double derivative being less than equal to 0.

So, based on that you can find out whether, your risk aversion properties true or your risk loving property is true. Then, using the concept of a certainty value we found out that, if the utility function is known beforehand in the functional form, but the parameters are not known you can basically also use the certainty value to find out how the parameters of the utility functions could be derived.

Later on then we consider that apart from the expected value, what are the other different type of matrix or other different of measures or other different type of basically simple concept we can use to rank the decisions; we found out that, there was a geometric mean method mean method and geometric mean method you find out basically the if the probabilities are given and then you find out the geometric mean such that, you rank them from the maximum to the minimum.

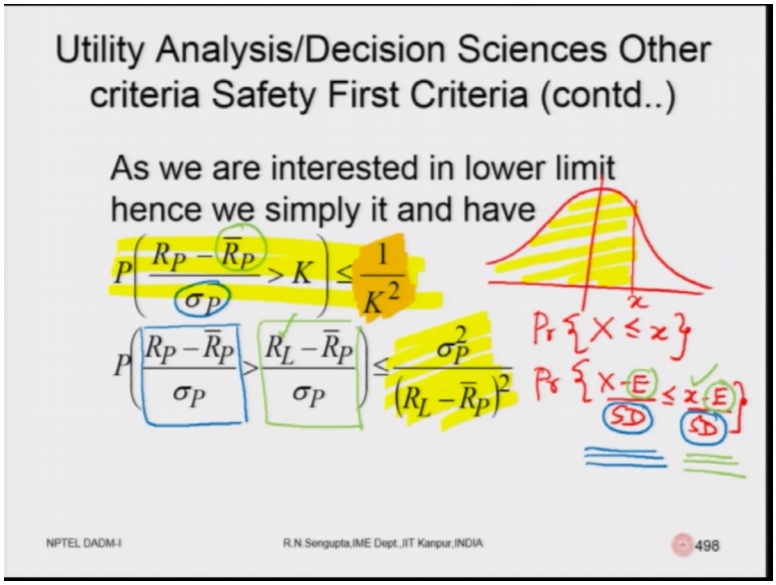
Later on we saw that that we can also use the safety first principle. And the concept that the probability of the return r p is less than equal to r l, r l is fixed value that you basically I try to decrease that overall coverage area between r p and r l, that is one method. Another method can be basically I try to push r l more onto the sides on to the right such that, the overall return on the of the so called decisions, the average decisions would be definitely on an increased scale.

And the third method was basically to push the average of that that return of the decision which is r p bar on to the right such that, the overall distribution shifts on to the right with the higher returns for each and every corresponding values of r p.

Then in the fag end of the class for the for the 49th lecture I consider the Chebyshev's inequalities, now the reason why I went a little bit slow is that, you will try to utilize that this decision concepts are very important and we will try to utilize that in some of the problems for the regression models and later on when we do the concept of different above other multivariate statistical methods and also this concept of utility theory would be repeated again when we do the DADM to course which is basically to do with different type of multi a multi criteria decision making where, many of the methods are nonparametric in nature and how we utilize the concept of utility or the pareto optimality frontiers, in order to understand that the different type of techniques which you use from the operational point of view considering the non parametric functional form are true would be utilized.

I will definitely come all to into all this concept later on, but I thought I will just mentioned it very briefly So, consider the concept of this utility analysis from the point of view of Chebyshev's inequality.

(Refer Slide Time: 16:26)



So, what we are interested is in the Chebyshev‘s inequality is the probability that R P is basically R P consider it is a random variable and R P bar is basically the expected value and similarly sigma suffix P or sigma P is the standard deviation.

So, what we are interested is if this is the distribution, let me draw the normal distribution. So, our main task is that technically we know, I am considering X only, x is the random variable is less than equal to sum X, so consider this sum x for the time being is this. So, my main idea is, I want to find out this. So, this is the is the expected value and standard deviation is S D is less than equal to x minus E by S D.

So, this E value or the expected value which we have is actually this and the value of standard division which we have is basically sigma P. So, trying to find out that expected value of the area using the value of Chebyshev‘s inequality basically, gives us a bound. So, this is the bound which I am talking about and try to utilize a different color. So, this is the bound based on we, we can say the probability of the standardized probability, so this is the standard probability I am not going to highlight it, I am just circulating it virtually.

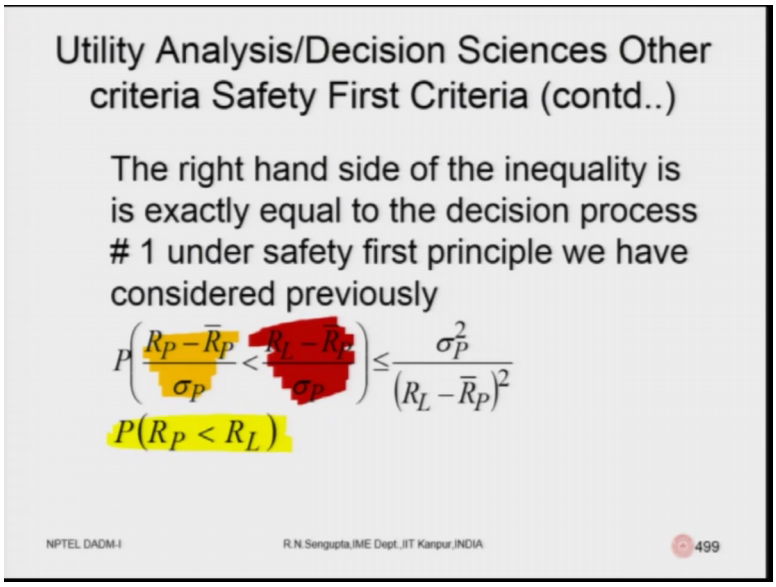
So, this standards value which is Z is greater than equal to some K value, K is a constant would be less than equal to 1 by K square; that means, I know the bounds based on which I will basically do the calculations. So, once I considering so, so this is what is going to come. So, the part we should so, this which I consider. I think I should I should use a different color, so it is easy for all of us to understand.

So, this part which I use is exactly this and the part which I use here is exactly which I found out here. So, in this case X is basically R L, which is the risk free interest rate which is generally considered for portfolio investment.

So, on definitely so, this can be considered as K, so once by k square would be equal to sigma square P which is basically the standard deviation square or the variance of the of the set of all values which are there for the portfolio divided by this is basically a sort of normalized norm where I find out the ratio of the sigma square divided by R L minus R P square. R L minus R P square are basically member R P bar is fixed R R L is fixed. So, square of that is a fixed value where, sigma square root would depending on how the overall spread of the total distribution is you can basically calculate sigma P square and then do the calculations accordingly.

So, basically this value I will highlight using the yellow color, so this would give me basically the bounds based on which I can utilize Chebyshev's inequality for to find out the bounds in the decision case.

(Refer Slide Time: 20:30)



Now, coming back to this again discussion, the right hand side of the inequality is exactly equal to the decision process number one under the safety first principle when you are trying to basically, minimize the ADR difference between R P and R L, considering R P is on the left hand side, left hand side from my side.

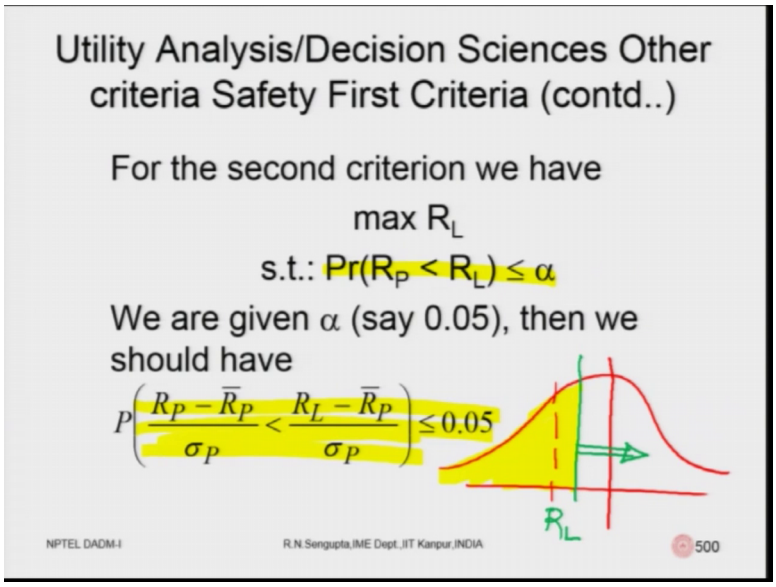
So obviously, it will be left hand side if you are looking from it, it will be right hand side from my side. Under the face safety principle we have a considered So, basically what we did was, this was the actually this is the first line so, once I normalize them, R P minus R P bar divided by sigma P, which is the left hand side here which I will highlight using a different color, in the right hand side this R L minus R P bar divided by sigma P.

So obviously, this would be thus right hand side. So, that we know it is basically equal to the ratio less than equal to the ratio sigma square divided by R L minus R P square, which is what we found out. And this is exactly what we intended to prove using the Chebyshev's inequality, as well using the safety first principle.

So, which means they are trying to minimize the area between R P and R L considering R P was in the left side left hand side, also gives out the same rule or the same concept

using the Chebyshev‘s inequality So, whichever you use you will get the same answer as such. So, Chebyshev‘s inequality basically gives us a bound, maximum value or minimum value, whichever you look at it and that is what we get from the calculations in the and the safety first principle the first criteria.

(Refer Slide Time: 22:11)



For the second criteria what we need to do is that, you have to find out the maximum of maximum value of R L. So, you will basically push R L more on to the right. So, take the average value not the average value. (Refer Time: 22:26) So, this is the distribution, this is the average value and I consider R Ls here.

So, what I am trying to do is that, I am trying to basically push R L to more to the right such that, the actual probabilities of coverages of R P less than R L; obviously, that you will still try to basically minimize in the long run but if you try to basically put R L on to the right the overall so called average value of the returns on a random scale would definitely be much higher, in case if R L is on to the left..

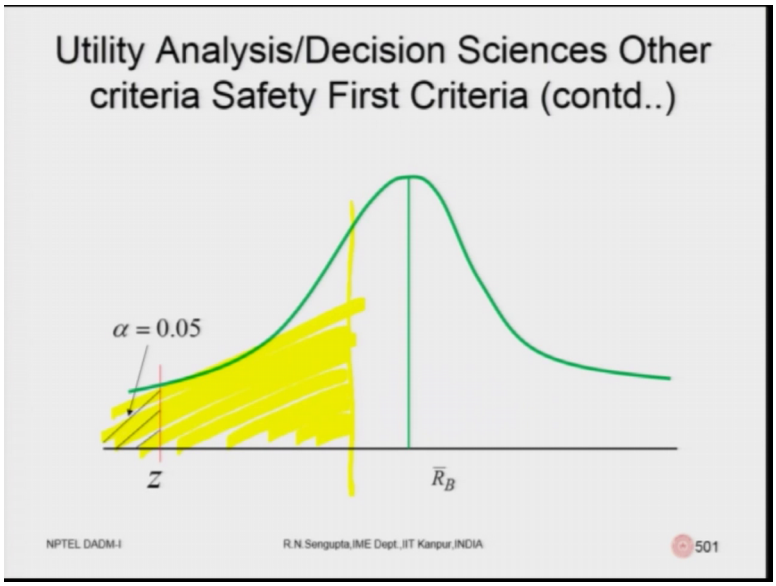
R L is basically the minimum threshold. So, obviously minimum threshold also increases; obviously, will try to push the overall distribution on to the right from my side. And we will always consider such that the probability R P minus R L which is less than alpha, so that would be true, should it be highlight. So, this is you want to put. So, this always remains fixed such that, this is less than equal to alpha.

So, if R L is moving it pulls along the green line which is (Refer Time: 23:40) if sorry sorry sorry my mistake, I know my mistake I should basically, a it is this. So, this is considered R P which is on the left hand side. So, I am trying to basically maximize R P such that, the area probabilities reduce. So, consider R I am considering an R bit value of R P, it can be anything, but technically if I consider the whole area, it should be this.

So, this is alpha So, I need basically keep fixed alpha at this level and try to basically pull R L onto the right hand side, which means, and obviously, R K fixed in the sense that I keep it less than equal to alpha, it can be alpha so, it can be alpha 1 alpha 2 alpha 3, where alpha 1 alpha 2 alpha 3 are less than alpha. So, more I push on to the right, the overall area can be decreased provided the distribution by itself is also moving. So, pulling R L on to the right, obviously, would mean that the value of the distribution as such as a whole body will move on to the right.

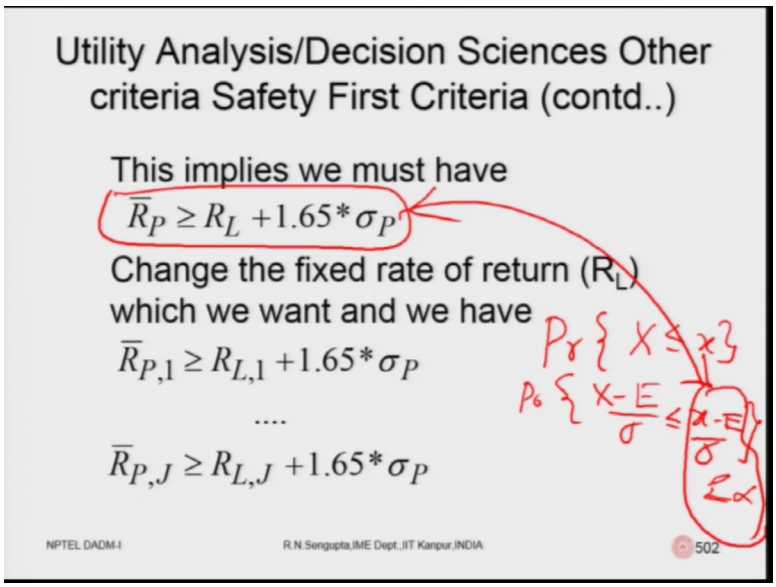
So, we are given alpha which can be say for example, 0.05 or 0.45 whatever it is and based on that once we convert that into the standard normal case it is become so, this equation which we have here exactly because the standard normal case like this and you can solve it. So, you want to maximize R L based on a on a constraint the constraint is basically a simple, stochastic constraint but we know that stochastic constraint considering normality is there it can be easily done through simple reliability optimization or simple optimization. So, I it is not a part and parcel of this course I am just mentioning.

(Refer Slide Time: 25:34)



So, in this case you have the utility analysis for the safety first principle. So, alpha is basically on to the left 0.05. So, it could have been on this side also the total area could have been 0.045, 0.4 0.45 so on and so forth, so it depends on how you basically be able to manage the problem.

(Refer Slide Time: 25:59)

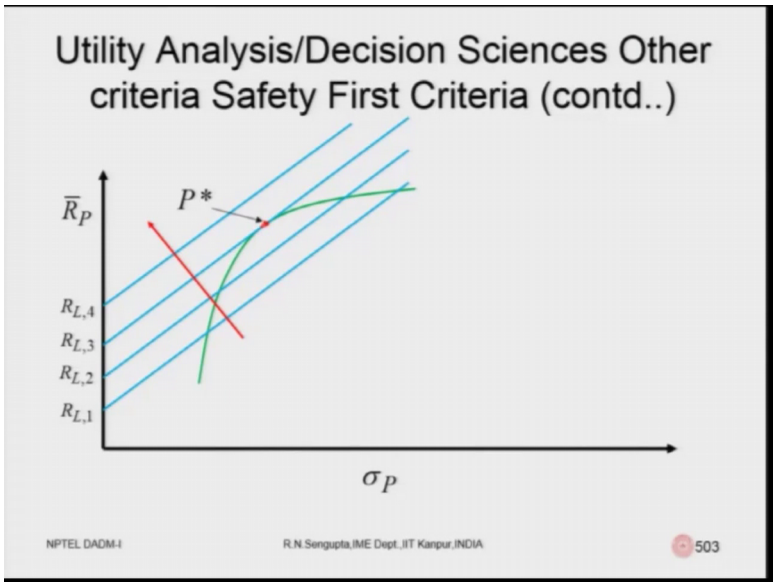


Now, considering that you are basically converting that into probability for the normalized case, so what we will do is, so if you consider the probability of X greater than less than x which is probability X minus X bar or let it me put it as.

So, let (Refer Slide Time: 26:36) be change the values accordingly. So, the E divided by sigma less than equal to x minus E by sigma. So, this is the X is equal to alpha. So, once we take this term remember and expand it so obviously, it will be some time number of times left or right of the sigma value.

So, if you put it here, so you basically, so this x and E would be considered in a such a way, so, this x is basically R L, E is R P. So, I basic and sigma is sigma P. So, how many times it is on to the right or the left would be dictated by this equation. So, just make a 1 to 1 correspondence here on the right hand side, this is the right hand side of the normalized probability. So, so change the fixed rate of R L accordingly, so we can change it and find out different values of R P and based on that you can plot it and do the calculations accordingly.

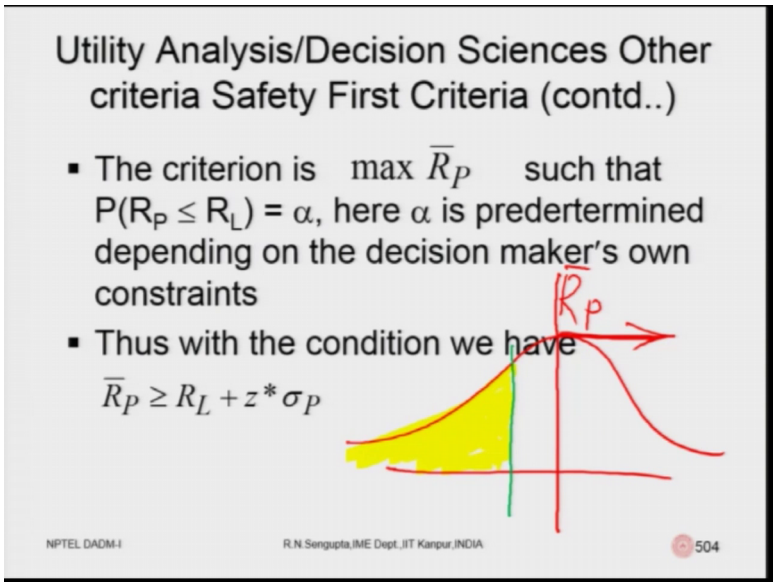
(Refer Slide Time: 27:41)



So, this is how the values would if you plot it or not on a risk return framework and the decision framework.

So, your actual value of R P star is the best optimal portfolio, but what you are doing is trying to shift it till it is orthogonal to the green line and we will take that point which will give us the best return.

(Refer Slide Time: 28:02)



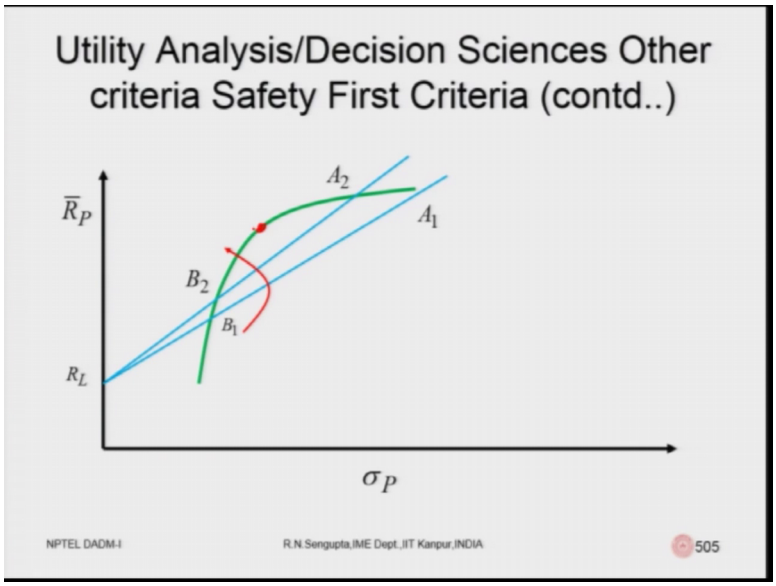
The fourth third criteria is criteria of trying to basically maximize R P which is the average referring to trying to push the average value onto the right such that, the value of the probability R P less then again the same thing R P less than equal to R L is less alpha. So, what we are trying to do is, this is the distribution, this is R I will I will draw R this value of so this is average value, which is R P bar and this is just for the information I am giving this is R L.

So, technically this is the area which is alpha, this is the second term remember here. And I am not going to highlight it, I am just hovering my electronic pen. So, what are you are trying to do is, push the average value onto the right, so as the average value moves on the right, so the whole distribution moves. So obviously, if it is equal to alpha less than alpha, obviously it can be made less than alpha, so the criteria of trying to push the values average values on to right would basically give you higher returns.

That is with the constraints, we can find out again, use the same concept convert in a standard normal, replace that R L X by R L and r p as the expected value and sigma P remains as sigma P, you can do it and then find out. So, this Z is basically the number of times it is higher than the value.

So, z is the standard normal values which you have from the table.

(Refer Slide Time: 29:39)



So, if you basically do the values of R L, now in the initial case I you saw that, I was pushing it parallel, now consider R L is fixed. So, what you do is that rotate anti clockwise, the moment it leaves that boundary it is tangent, you get the best optimum solution based on which the decision can be taken. With this I will close the 50th lecture which is completes our 10 weeks and continue discussing. Use this concept in the regression model and so on and so forth. And then again come back to the multivariate statistical methods in a simple way.

I did get some feedbacks about the factor analysis and principal component analysis those who are a little bit more theoretical, I will take that into definitely take that into considerations. Hence based on that, we have we have consciously taken a decision that use utility in a very simple sense and for the regression models also we will just give the problems and then come back to the simple analysis of how the problem is done. So, rather than doing the theory and then the problem we will basically do the problems and then come back to the theory if at all required.

So, we will keep it as simple as possible. And with I thought that I will I did mention very fleetingly about DADM 2 also, I will give a brief background of DADM 2 at the last lecture of DADM 1 such that, you can understand if you give us a very a very frank feedback that how, DADM 1 went..

So, we can basically fine tune from our end the good points and also basically try to reduce the overall the effects of some of the teaching methodology which you may have found a little bit difficult we will take that into consideration in all the seriousness. And we will expect a more of your feedbacks more in the forum and any questions you have you are most welcome to ask there and we will take care of them in the (Refer Time: 31:41) have a nice day and.

Thank you very much.