**Data Analysis and Decision Making- III**
**Professor Raghu Nandan Sengupta**
**Department of Industrial and Management Engineering**
**Indian Institute of Technology, Kanpur**
**Lecture-54**

Good Morning, Good Afternoon, Good Evening to all of you wherever you are in this part of the globe. And this is the data analysis and decision making course 3 and of the NPTEL MOOC series and as you know this total course duration is for 12 weeks. And the number of contact hours is 30 which when broken down into lectures is 16 number considering each lecture is for half an hour. And as you can see from the slide we are in the lecture number 54 which is the 11th week.

The last but one lecture for the 11th week and each week we have five lectures of half an hour each. After each lecture sorry the week you have 1 assignment to do you and you have already completed 10 assignments you will undertake and also rule the 11th and 12th and then the final examination. And my good name is Raghunandan Sengupta from IME department at IIT Kanpur. So if you remember we were considering the concept of very basic background of stochastic programming reliability, optimisation, robust optimisation.

And there I started with by saying in stochastic programming obviously we will consider that the underlying distribution of the parameters which are unknown of the decisions variables which you want to find out are unknown. In case if they are unknown obviously will try to take the recourse of some statistical simulation techniques bootstrapping, jack knife and so on and so forth.

And then try to basically find out the best estimate using all the trying to verify the properties of the estimate with respect to the proficient parameter and complete your job and utilise those estimated parameters of the population which are the estimates from the sample and proceed. In case they are known you basically directly use them. So where are there to be used? So let us consider that in the stochastic programming and the reliabilities programming concept.

While in the robust programming concept or the robust optimisation concept we consider some perturbations like an atom vibrating about its nominal value and there the underlying distribution is not given or not known or that is not considered at all, we consider the concept of some perturbation sets and depending on the perturbation sets we basically proceed to find out some proofs of or the greater than type less than type what is that I am going to come to that later.

Type of proofs for the constraints on the objective function in the objective function obviously it can be done as a minimization or maximization to which also I will come later on. And then we proceed by the simulation concept. Now if you remember the last lecture which is the 53rd lecture I drew a lot of diagrams in order to a make you understand that if it is normally distributed univariate then trying to basically use the standard normal deviate the Z table values and the standard normal distribution values.

We can basically solve the problem depending on the level of confidence of alpha which you have. Now when if it is a multivariate normal distribution obviously we will use the concept of multivariate normal distribution, you have already the proofs then again the concept would be utilised in such a way that both the variance covariance matrix the N by N matrix considering there are N number of such decisions variables are there provided the parameters are not to be considered now.

So, you will basically considered the N by N matrix of the covariance matrix of the decision variables. The vectors of the means of the decision variables from the underlying distribution and proceed to solve the problem. Now when you want to solve the problem I mentioned that in the reliability concept you use the sequential optimisation technique and use the most probable point technique.

So what you are trying to do is that you are trying to basically take consider the probability that the function. Here I am going to consider when you one of the constants arbitrarily say for example g, and g is a function of X which is the decision variables d which is the deterministic parameters, p which is the probabilistic parameters and we will consider g as less than equal to 0, it can be greater than equal to 0 also. It can be less than equal to greater than equal to b 1, 2, 3, 4 whatever the right hand side of the equations are.

And this is true with the probability of Alpha which is the level of confidence. Then when we consider the diagram for the normal distribution case it will be on the right hand side or the left hand side for the diagram if you remember depending on how the problem formulation is been done. And when the problem formulation has been done what we do repeatedly is that solve the optimisation problem using any of the known techniques that is in the X plane or the X space and then try to basically transform that actual output which you get the first iteration into the u space.

Then optimise trying to find out the minimum variance point or most probable points, sorry most probable point and then basically do the inverse transformation of that U star point into the x space and continue doing the optimisation and then back and forth you continue till you basically able to reach the optimum solution depending on the iterative process which you are trying to aim at.

Now a question which would have occurred in your mind as we were discussing is that what is the point do we consider x naught from where we start. So you saw in the steepest decent approach the in the problem which we solved the x naught was considered as an arbitrary value. It can be chosen intelligently also. So using the same concept that you have to use intelligent starting point for your search we consider the mean values or the nominal values for the vectors x, x distribution factors which you have which are mu 1, mu2, mu3 till mu N considering mu 1 is the expected value of x1.

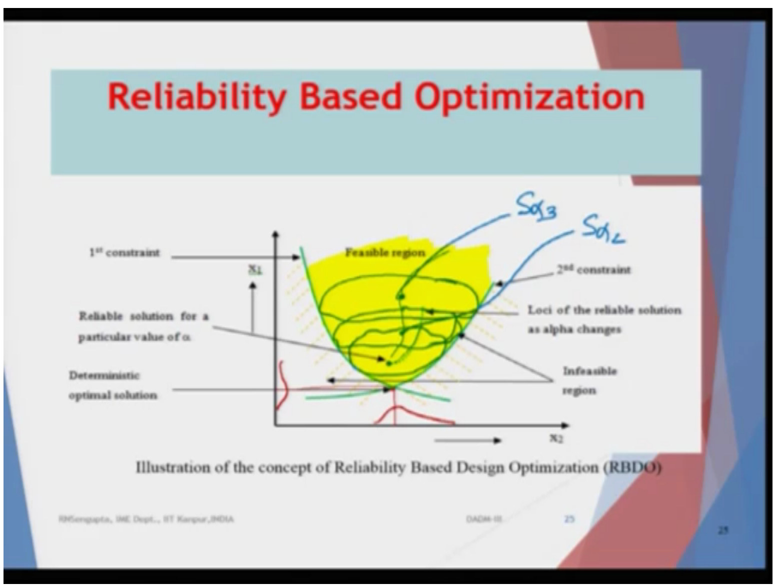
Similarly mu 2 is the expected value of x2 and so on and so forth. So you will basically consider the mean values of the axis. In a similar way we will consider the deterministic values are deterministic so they are fixed. So they would not be any mean values for them and for the parameter values which are nondeterministic p, we will also take the mean values and then basically try to basically start of the optimisation in the x space.

Now, the question is that well you have considered the normal distribution and as I showed that in the normal distribution case for the 2 dimension, if the variances are equal so hence the normal distribution if it basically arising from the x, x axis x1 and is coming from the orthogonal direction which is x2 and if the basically converge then the overall area is the circle provided the variances are both are same.

In case the variances are different you will basically have an ellipse where the major axis and the minor axis would be accordingly like this. In case if the variance of X2 is more than the major axis would be vertical. In case if the variance of x1 is more the major axis would be a horizontal line. So obviously it would mean that it is basically like a baseball in a 3 dimension case provided there are 3 random variables and also consider that the variances of two of them are equal one is more.

So in that case you will have a baseball, baseball lying horizontally on the field or vertically up on the field. Now in the case if the variances are all equal then we basically will have a sphere. And in the higher dimension case it will be a hyper sphere or how the figure is done considering they are the set of symmetric distributions on the normal distribution which we all know and based on which we can solve it quite efficiently. But now the question happens and or the question arises is that what happens if the distribution per say are not normal.

(Refer Slide Time: 08:47)




So let us consider that example here and I have drawn the similar figure but with a major change here. So the major change is what I am going to come to that later on. So consider the x axis is x2. So, you are taking x1 and x2 does not matter. Along the y axis I am measuring x1. So this and first green line which is marked as a first constraint which is example g1, the second constraint is also marked by the green line which is the second constraint considered as g2.

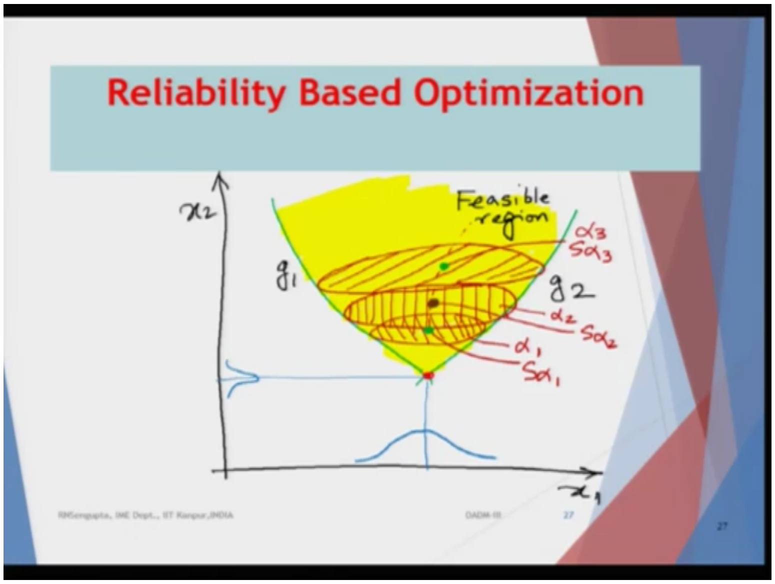
And inside region which I will mark this one which is the feasible region. It is exactly the same type of analysis which you have done in the previous lecture. Suggest highlighted this in the yellow colour in order to denote that this is the feasible region. Now in case if it is deterministic as per the norm we all know it would be one of the solutions of the which is the quantum point and let us consider the quantum point is this which is the red one which I am not going to highlight because it is already shown red.

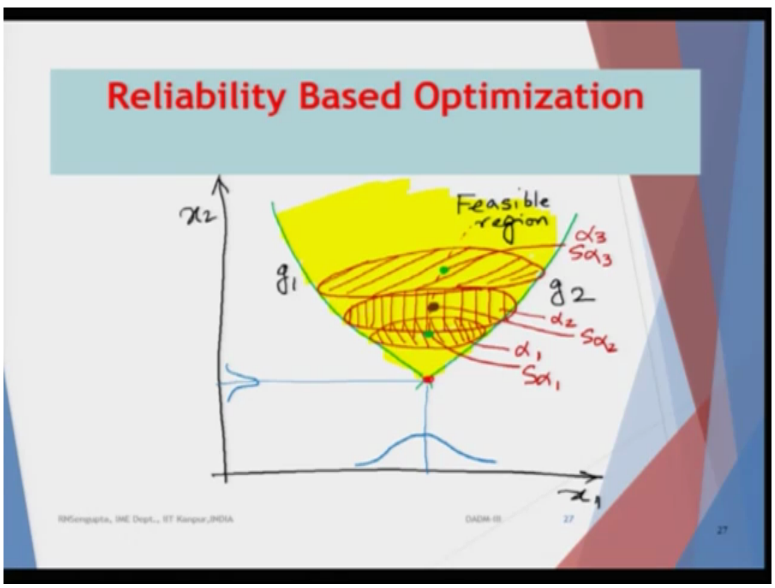
If you want I can highlight it using the red colour. So this is the deterministic solution. So if you go what is noted down here is a deterministic optimum solution which you have. Now consider as the variables are orthogonal so let us take, I will go one by one on the same picture I will draw an erase also. So please bear with me. So I will consider blue colour. So this is the axis of x to considered is a normal distribution.

This is the axis of x1 again the normal distribution. And considered the normal distribution variance is same so obviously you will have concentric circles. So this I will try my level best to draw. So as the reliability increases the line moves. So this is the loci of the solution of the deterministic case.

In case the distributions are normal but with different variances, so let me now prepare the diagrams in fresh slide. So I will basically prepare 3 slides consecutively and then come back to this. So it will be easy for us to understand. I am sure you are quite intelligent enough to understand what I want to draw but still I want to go through diagrams carefully. So in the first case I will repeat it because I thought I will draw it on the same figure but it was getting a little bit cluttered.

(Refer Slide Time: 12:19)




So you have the axis consider an x1 x2 whatever it is but I will follow x1 here. There is no change in the concept, so I will draw the deterministic this g1x and g2x line as this. This is the best solution for the deterministic one. This I consider as the normal distribution. This I consider as the normal distribution and consider the variances are same. So I will have depending the level of reliability, these are circles consider I am trying my level best plus please bear with me.

So this is the loci, so the levels of reliabilities are given as alpha 1, alpha 2, alpha 3. So what is alpha 1, alpha 2, alpha 3? I will basically mark it accordingly. So this is for alpha 1, this is for the area one is for basically alpha 2 and this is for say for example for alpha 3. So this I will say the reliability is the solution for alpha 1, this is the solution for alpha 2. Solution means the reliable solution, solution for alpha 3.

And I will mark them, I am just circling it with different colours. This is a solution for alpha 3, this is a solution I will use a different colour if possible please bear with me. This solution for alpha 2 and this is the solution for alpha 1. Now consider both the mean values the variances are the same.

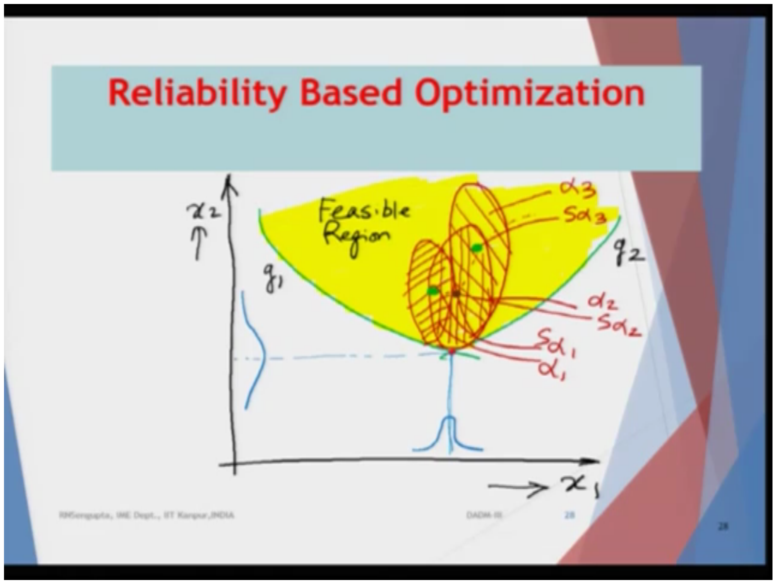
Now let me go to the variance being different one at a time. I will try to use the same colour. So this is the axis x1 x2. So let me check whether I have done x1 here and x2 yes, another thing which I did not mark. So this is the feasible region the total area. This I am mentioning as g1 this I am mentioning as g2.

Again I draw, now I will be using the colour schemes similarly so I will be switching back to the diagram also in order to make things. This is g1 this is g2 this is the deterministic solution. So consider this variance is large and this variance is small. So what you will have in this region what circles colours we used we used this one.

So it will be with different levels, so this will moving like this. So this is alpha 1 and this is S solution which is the reliable point of alpha 1. This is alpha 2 this is solution based on alpha 2. This is alpha 3 solution means the best optimum solution in the reliability case alpha 3. So alpha 1 would be this area. I am using the same nomenclature marking and the colour and same type of hash lines also.

So this is for alpha 2 and this one is basically for alpha 3 and the colouring which I did blue, violet, green. So I use this is alpha 3, this is alpha 2 the solution for alpha 2 and this is for alpha 1. And your feasible region would be this. So this is the second way of trying to consider that the variances are different.

(Refer Slide Time: 19:58)




The third way, by the time you have now understood but still I will still draw it. Here the drawings may not be exact but they should definitely give you some representation or the idea.

Sorry sorry sorry I am using a different colour. This is g1 this is g2, this is the deterministic solution. Now here the mean value is less. These are normal distribution please and forgive my drawing skills. So this has a normal distribution but the variance is large, so in this case rather than having the ellipse with the major axis horizontal the ellipse would have a major axis which is vertical. Obviously there would be some region there would be some (())(21:20) touch so I am not being very specific here.

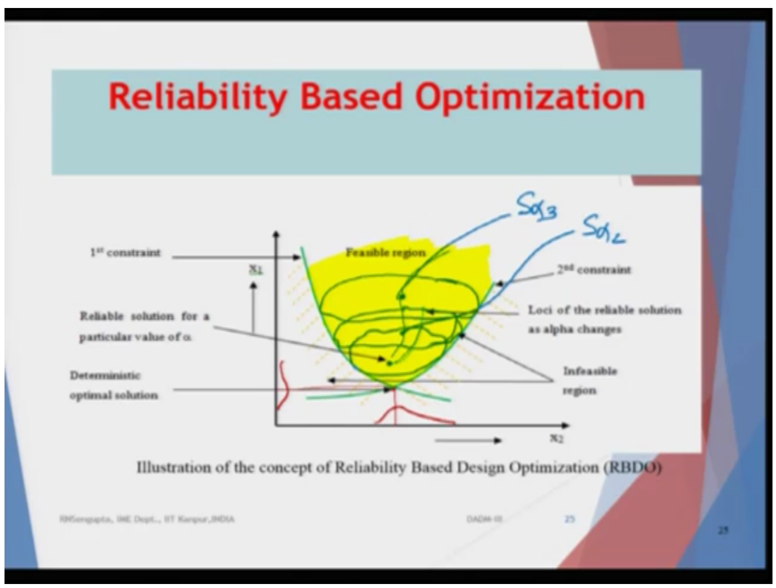
So this is so here now the values of alpha change the loci of the centre which of that ellipse or ellipse or whatever it is in high dimension that keeps changing. That means that overall path will trace out that as the level of reliability changes how the solution changes. I did not mark this. This is x1 this is x2, this would be alpha 1 for alpha 1 this one would be for alpha 2 and this one is for alpha 3. So I mark it alpha 1, S of alpha 1 and this area is alpha 1.

This second circle is with alpha 2 and the solution is S of alpha 2. Third circle third is not circle sorry ellipse is alpha 3 solution is S alpha 3. So I need to mark the solutions. What colour do I use? It is it should be blue, violet, green so sorry for that. So I will use blue one then the violet one and the green one. So you get the solution. So the first diagram and un-obviously I should mark this is the feasible region.

So it would look considering the different nomenclatures I think it should now be clear to you that for the symmetric distribution, consider the normal cases as the one which gives us the best understanding, so this is the feasible region. So this gives us in the similar way, I should mark it properly. So this is feasible region and in the similar way for the first time diagram also.

I am going a my speed is a little bit slow for this lecture because I thought I should the diagrams will come up again but there would be very concise. So I thought I would explain it here such that this clarity of the concepts are absolutely clear in all the sense that how you solve. So this is for the both the variance being equal, this both being unequal this also for both being unequal.

(Refer Slide Time: 25:48)




Let me come back to the first diagram where we left. So this is also not complete. So I will come back to this now. Consider the distributions are non-normal. So let me use the colours is it can be the left skewed or right skewed does not matter. Consider this is like this, so if these two distributions are orthogonal, they merge. Then the overall area which we see is no more a circle. So it would be a very weird area which I have tried to draw in this figure in a 2 dimensional case.

In that higher dimension case it would be a complicated surface and obviously the properties of convexity and all these things should hold such that you can find in the optimisation but they would not be any more a circle of a sphere or hyper sphere. So the reliable solution for a particular value of alpha is given and here the feasible region which is inside. So this is the feasible region.

So the yellow colour should be ischemic things clear to you, so here is alpha and area which you have as the reliability changes. So you will basically have for alpha 2, alpha 3 so this is the solution for S alpha 2, this is the solution for S alpha 3. So these loci which I have drawn would basically will like going like this. So with this I will end this lecture 54th lecture and continue more discussion about the reliability concept from the mathematical point of view and discuss few optimising problems in the subsequent class. Thank you very much and have a nice day.