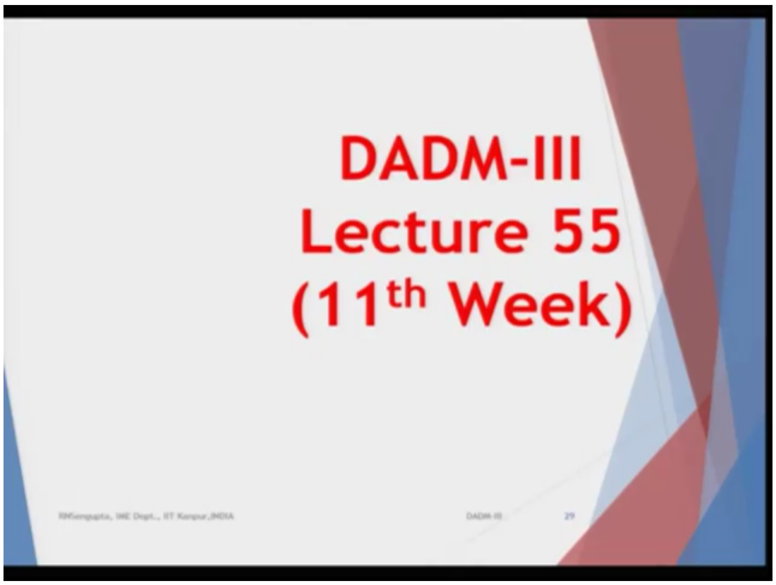
Data Analysis and Decision Making- III
Professor Raghu Nandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur
Lecture 55

A very warm welcome to all my friends in students, a very Good Morning, Good Afternoon Good Evening to all of you. And this is the DADM 3, which is the data analysis and decision making 3 course under NPTEL MOOC series and this total course duration is for 12 weeks which is 30 contact hours which when consider into the number of lectures is 16 number because each hour of the lectures are half an hour.

(Refer Slide Time: 00:43)



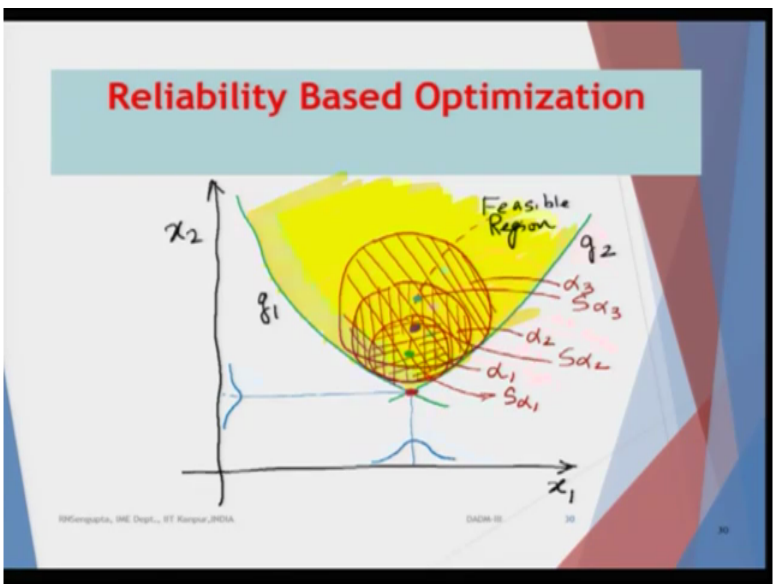
And as you can see from the slide this is lecture Number 55. The last lecture in the 11th week and each week we have 5 lectures of half an hour each and after each week you have an assignment. So with this ending of this lecture you will take the 11th assignment and then in the 12th you will finish the course 12th assignment and then you will have basically a final examination based on all the concepts which have been covered.

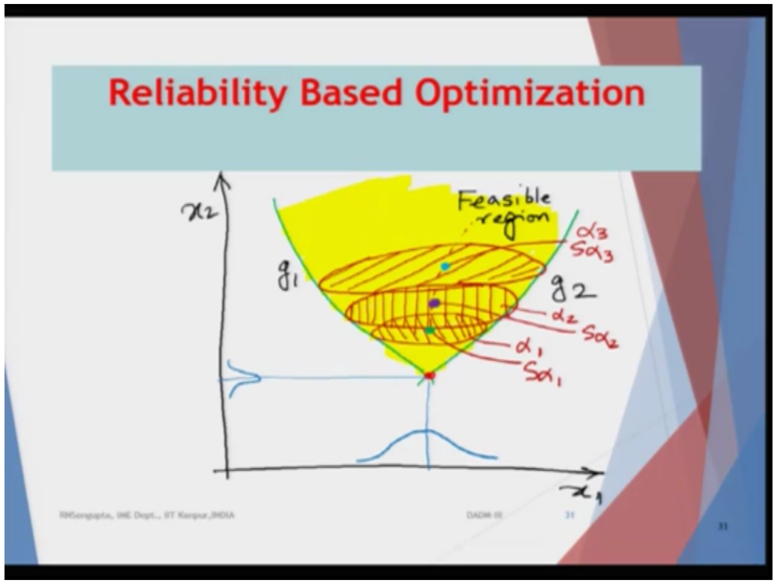
Now, if you remember in the last lecture which is the 54th lecture it was I only drew diagrams in order to make you understand. So both were the case when you had the normal distribution to be true and variance for both are equal then in the case, second case the normal distribution to be true for both the cases. I am only able to draw a 2 dimensions. In the higher dimension if I was able to draw I would have basically shown it.

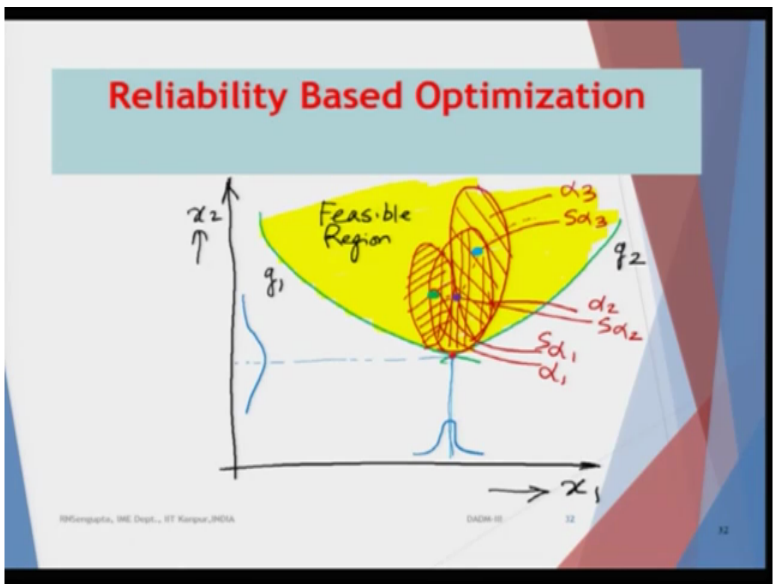
Consider the sphere would be fine but consider the variance are different for all the three normal distribution it will be difficult to draw it until and unless if I have drawn it earlier. It would have take time so I just skipped it. In the second case let me come back to the work or the discussion which was done in the last lecture which was the 54th one.

We had considered the both normal distribution and variance of x1 was more than x2. In the third case both were normal then the variance of x2 was more than x1 and later on which we will now discuss further on is basically where we consider both are non-normal. The variance would definitely be different and how can we basically think about solution techniques and how they look like.

(Refer Slide Time: 02:47)





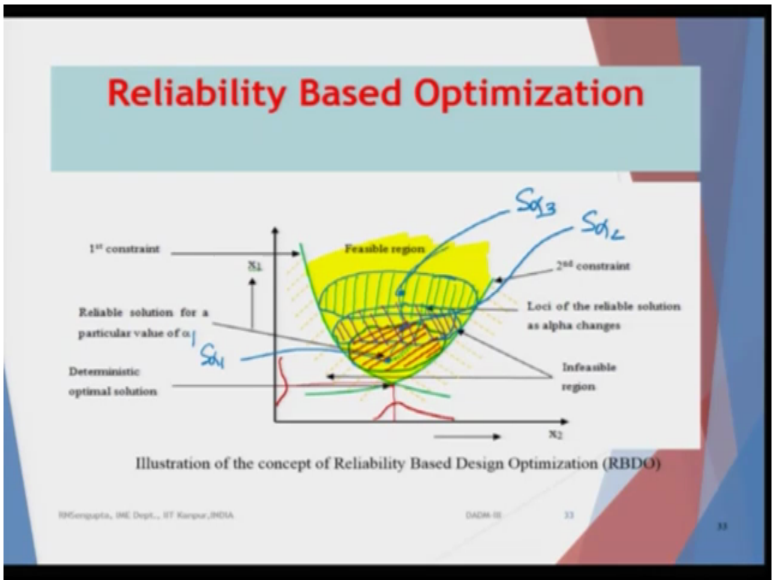


And obviously the solution techniques would be same, now we can solve it. So I should be basically copy it in the later part so it would make sense for all of us to consider. So this is the first case, both being variance being equal and alpha 1, alpha 2, alpha 3 are the consecutive values of reliability. In the second case, alpha 1, alpha 2, alpha 3 are the levels of reliability for the variances are different where x1 is more than x2.

In the third case again alpha 1, alpha 2, alpha 3 are the levels of reliability and obviously in all the 3 diagrams the solutions are given by the centroid or the centre of the gravity of these figures, the circle is the centre. Similarly, for the ellipse is also easy to find because it would be the midpoint on the major axis on the midpoint on the minor axis and this levels of reliability would dictate the solutions which are given by S by the suffix corresponding to the reliability.

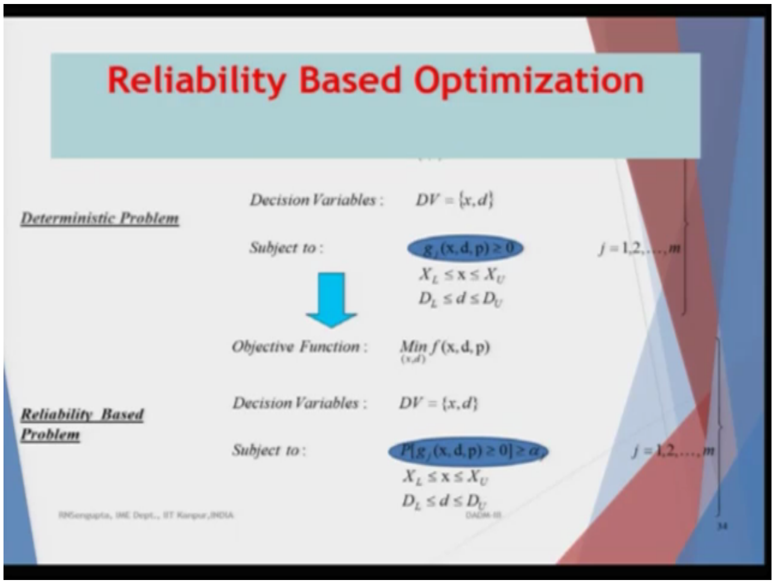
In this case the variance is x2 is more than x1 and remember these are orthogonal. In case even if they are not orthogonal the concept considering that if the normal distribution would not matter much it is not difficult to visualise.

(Refer Slide Time: 04:07)



In the fourth case, we have the distributions are normal and when we find out the centre of gravity which is given here as this one. So I should mark it with a blue colour, so this is for alpha 1. This is basically S alpha 1, this is S alpha 2, this is S alpha 3. So if I consider, this is of reliability alpha 1 so this is of reliability alpha 2 and this is of reliability alpha 3. Now here the concepts of infeasible region feasible region are exactly the same as we have been trying to draw. So this is the deterministic optimum solution or the corner point, and then of the loci is reliable solution as alpha changes (())(05:27).

(Refer Slide Time: 05:28)



Now how the overall problem formulation is being done which I have already considered but I will go into more details here. So consider the deterministic problem is given you have a decision sets variable which is x and d, x is the decision variable which would initially be considered deterministic they can change also. They can be a probabilistic also and d is the set of decision variables.

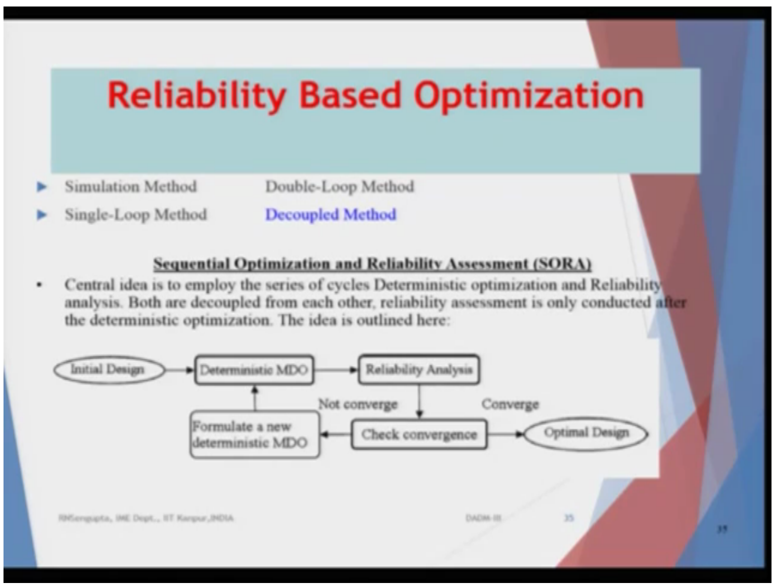
They are now the parameters. We are considering a the set of decision variables which is d. So they can be considered as parameters also depending on the problem formulation. Subject to conditions are like this. You have m number of constraints which I have already discussed g suffix j and they are deterministic in nature. So they are greater than 0.

So they can be less than 0 also. They can be greater than b j b suffix j is also. And the this concepts which you have for the x is basically bounded between x lower and x maximum XL and XU. Similarly for the deterministic variable DL and Du like X mean and X max whatever between which you have that value. So in case if X values are between 10 to 20 both inclusive continuous so x mean will be 10 and X max will be 20.

Now consider it is a minimization problem for a maximization problem whichever you try to formulate. So your decision variables would also be X and D but here X is now probabilistic. And this concept of probability would be brought down into brought into the picture through the constraints such that you have a probability of g j x d p being greater than 0 and that is given by alpha j.

So if there are say for example, capital J number of constraints, each would have basically a probability of alpha 1 to alpha k and this alpha 1 to alpha J and these values of alpha 1 to alpha J would be the corresponding levels of reliability which the decision maker has for each and every constraint. So we will basically try to take the collective reliability based on which we find out the best optimum solution. And subject to constraints are also the same X is less is between X min and X max, Similarly for D it is between D min or D Max DL and Du. L is the lower limit and u is the upper limit.

(Refer Slide Time: 08:09)



Now reliability based optimisation can be solved using different methods. One is the simulation method. I will basically go into the decoupled method, double loop and single loop methods are there. I will first consider and other method is the simulation one. So I will basically talk something about the simulation method then go in details about the decoupled method. So considered the simulation method like this.

So say for example, in the case of the normal distribution what you had. You found out probability of X capital X is greater than x. I am not going to write because I have been writing this equation time in again so let me explain. Probability of capital X weather in the univariate case or multivariate case the concept remains the same. Probability of X is greater than say small x, say for example greater than alpha and you know that expected value of x and you know the variance of x.

So it would be probability of x minus expected value of x. The whole thing divided by the variance which now becomes capital Z and that is greater than equal to X which now is x minus expected value of capital X. Small x is basically the or realised value which will be realised later on depending on how the simulation is being done divided by variance of X. So you will basically have probability of capital Z greater than equal to small z that is equal to alpha.

So you choose the value of alpha, find out its small z and you choose the value of small z you find out the value of alpha. So in this case normal distribution is there either in the multivariate or the univariate case life is done. But what happens if there non normal? So what you do is that you take you want to find out that how many times the distribution, random variable X is greater than equal to small x which has been given.

And that is the value of is probability is given. So what you do is that you simulate 1000 values and considering that you are only simulating one at a time not taking the sample statistic of the sample estimate. You simulate one at a time rank them lowest to the highest or highest to the lowest whatever it is. Now what is alpha? Alpha is basically the area on the right hand side as for the diagrams which you have been seeing.

So you want to basically consider 1 minus alpha. 1 minus alpha considered is 95 percent. So you will keep adding so for each of these random variables generated from the minimum to the maximum there is a corresponding value of probability also. So X and consider they are nomenclature as x1, x2 till x100. Considering that generated 100 values, so for each of these axis you have the probability also.

So probability of capital X, is equal to X1, Probability of capital X is equal to X 2 so on and so forth this given. You keep adding the probabilities one at a time and find out the cumulative probability. The cumulative value where it becomes exactly equal to 1 minus alpha or some value near that 1 minus alpha which is 95 percent that value of X star is reported, so you keep repeating do it 100 times 200 times and each time you will have a different value of x star you will try to basic.

So if you have done it once, considering that you have the whole population then you would have found out the best value of x and that you would have reported and basically completed your optimisation problem. So you want to basically find out at that case where the distribution would be true for that value of alpha and that considering those values you will basically solve your problem to find out the reliable solution.

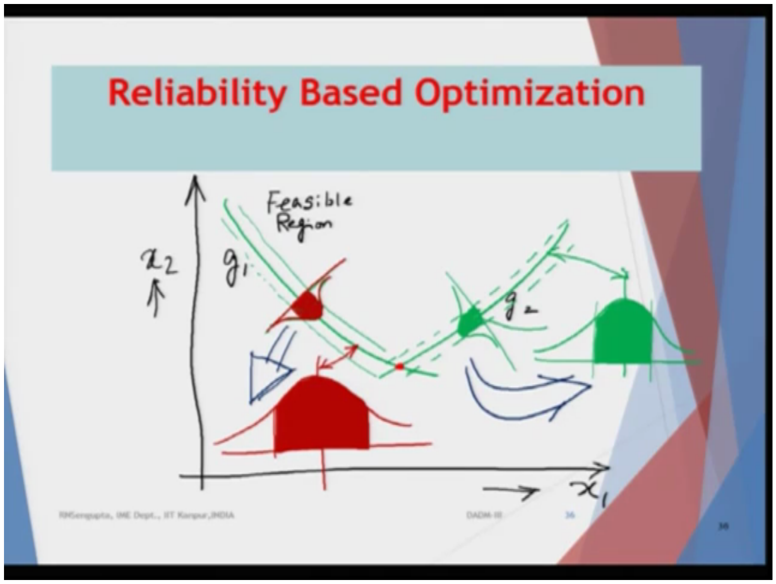
But in the case you do not have a whole set of population values you take a sample generated, say for example, hundred number of times with each set of samples having another hundred. And find out X star in each case and you will be basically depending on what type of distribution it is. If it is an exponential distribution, normal distribution then considering if is extreme value distribution you will basically take the mean value of the or some the minimum of that value or the maximum of the value and then basically report it.

As the best x based on which you will do the simulation. Now if you remember I am not changing the topic I am coming back to 1 concept which you have done long time back considering the whole course. In the first 2 weeks or 3 weeks in the course we have done something to do with this sensitive analysis and if you remember in the sensitive analysis what we did we basically changed the right hand side by 1 unit decreased or increased or basically changed the parameter of this decision variables by 1 unit.

So that will give that if the cost or the profit of x1, x2, x3 basically changes by 1 unit how does your optimum value solution change, Which means that in the case if it was you are trying to change the right hand side which is the b value by 1 unit that would mean that overall constraint is shifting by 1 unit parallely. Now what happens in the case when your distribution is normal, that concept also needs to be understood.

So I will try to basically draw that. Again I am saying these diagrams will be coming up again in the later parts so hence things would become much clearer. So I will repeat it please bear with me so but these things would definitely be clear.

(Refer Slide Time: 14:22)



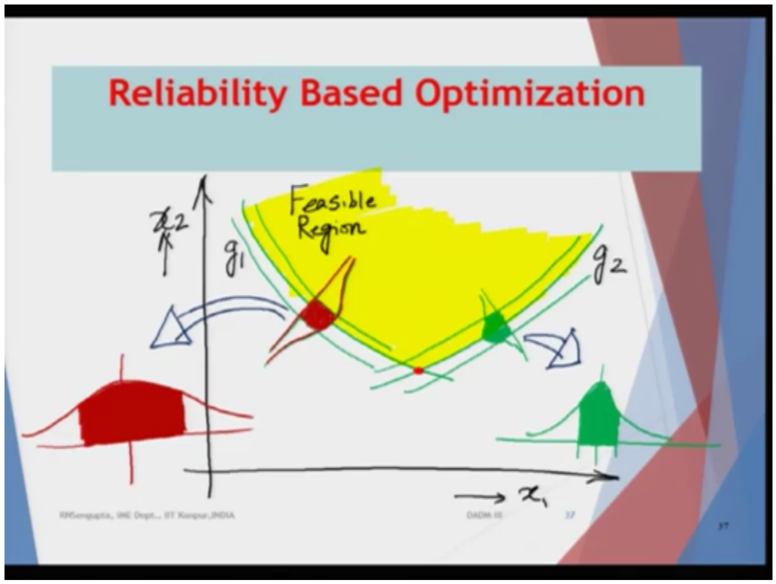
So first step draw the diagram and come here. I will take wait, 4 different blank slides to draw it. The diagrams would be almost the same but I will now highlight at other particular point this is the deterministic solution. Feasible region would be somewhere inside, I will just write it. If required I will mark it but what for the point of time being I will just mark it.

So this is g1, this is g2. Now what is important is that for perturbations or the sensitive analysis what we do is the that constraint shifts. So consider the constraints shifting is both are normal and both have the same variance. The first diagram, so it will be with say for example perturbations are happening at 90 percent reliability which means if I take the considered this is the normal distribution which we have very interestingly the mean value would be the bold line.

And this area which you have and also outside region because this will be in feasible region. That will give you the level of confidence the overall breath. So it is like this if I draw it this line is this line and the overall region if I am talking about, so this was what was talking about. In the similar way variance was there and consider the variances are the same so this is the region, so again if I so this diagram actually is if I am permitted so this is similarly if I draw it here.

I will use, I should be using at different colours so it is easy for us to highlight. This is how it is coming so these values and this value is this one, this mean value the line. So I utilise this concept so these are the level of confidence. So if you are talking about the level of confidence among the circle is drawn. So that will highlight the level of confidence which I am trying to highlight here.

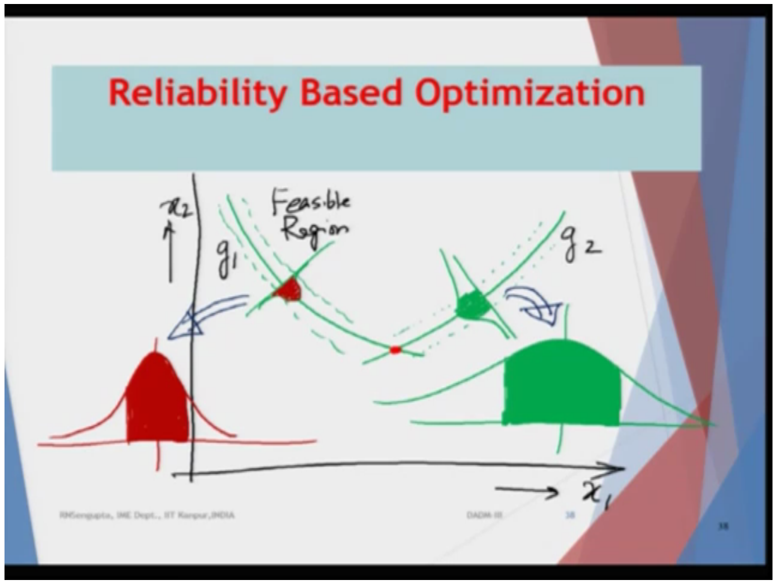
(Refer Slide Time: 19:32)



In the case so this diagram did I draw? Yes I have drawn x1 x2. X1 x2 is given. Now this is normal but one of these variances is high so let me just draw it. This is much higher variance you should be using the so other line, I am sorry. And in this case the normal distribution. Let me highlight this feasible region, it would look nice. Here I am trying to basically mark it because you know which is the actual constraint.

And when I draw it, the diagrams would be like this. So this is the level of probability, so Remember this level of reliability changing depending on the level of confidence which is the decision maker wants to make problem formulation. This is much flatter, spread out. You have the feasible region in the case when the variances are different.

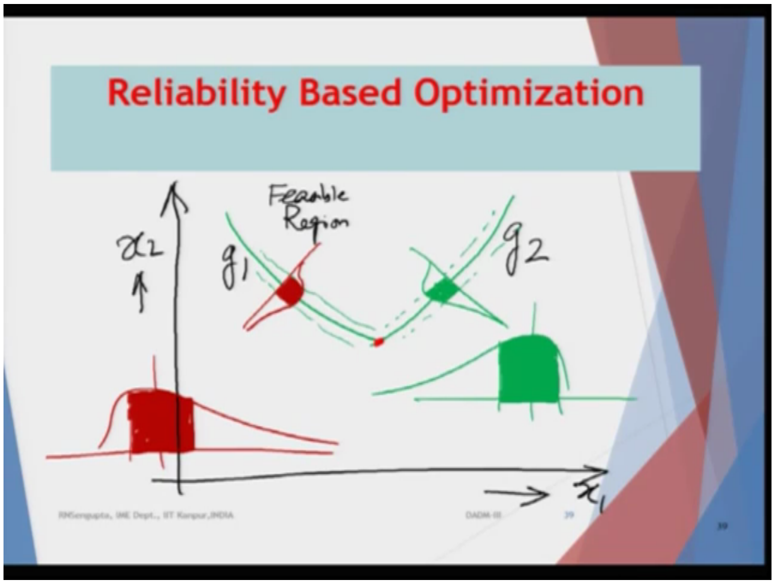
(Refer Slide Time: 23:56)



In the case x1, x2 1, 2 the second boundary g1, g2 feasible. I am just repeating the diagram but trying to highlight.

So consider this distribution, in the other diagram you should have dotted the boundaries but I did not do that and this is much more spread over. This one much more narrow this portion is low. So I highlight level of reliability here depending on the level of reliability which is there. And for this case so I draw the diagram this much steeper, this is much flatter, this one is the value and then highlight the bounded regions. You are the probability is given and can finally find it. But similarly for the non-normal distribution it will get complicated but you can find it out.

(Refer Slide Time: 27:20)



So let me draw it again also for the non-normal case. These 2 classes are going a little bit slow but considering that these concepts are important, I thought I will go quite slowly. I will have g1 here, g2 here. This is g1 this is g2 this is the solution in one case, so it will be like this.

Now the variance would be important but more important is the distribution per say. I use the, it is possible like this there is nobody stops us. So in this region this is the area (())(28:45). I consider one case is like this the highlighted areas you want to find out the level of reliability. So with this I will close the 55th lecture which is the end of the 11th week and consider more of once this concept is clear. In the higher dimension obviously all the repeated things will be repeated so we will be considering the solution methodology from the (())(30:05) and a most probable point techniques, have a nice day and thank you very much.