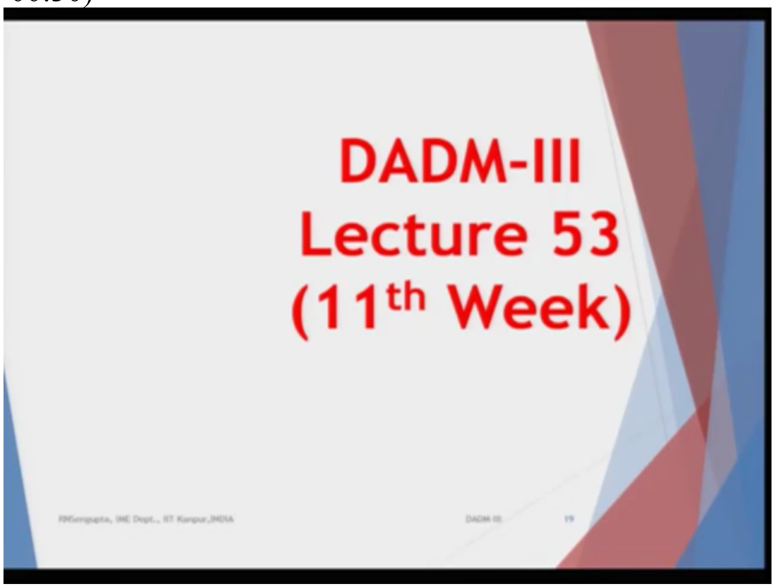
Data Analysis and Decision Making- III

Professor Raghu Nandan Sengupta

Department of Industrial and Management Engineering
Indian Institute of Technology Kanpur

Lecture-53

(Refer Slide Time: 00:30)



Welcome back my dear friends. A very good morning, good afternoon, good evening to all of you wherever you are in this part of the globe and this is the DADM 3 lecture series under NPTEL mock, which is data analysis and decision making. And as you can see from the slide is the 11th week and we are in the 53rd lecture which is the 3rd lecture in the 11th week. Now this total course duration is for 12 weeks. Contact number of hours as you all know is for 30 hours, which is basically broken down into 60 lectures.

Each lecture is for half an hour and after each week where in each week you have 5 lectures of half an hour each, you have 1 assignment. So considering that you have already completed 10 lectures you have already, 10 weeks you have already taken 10 assignments and with the finishing of the 11th week and the 12th week, you will take another 2 assignments and then go for the final examinations.

And my good name, very good name is Raghunandan Sengupta from IME department at the IIT Kanpur. So if you remember that we were discussing about the concept of reliability optimization and robust optimization with respect to the stochastic programming and the main concept which was which I mentioned was that in the concept of stochastic programming, we will consider an inherent probability distribution for the random variable or the weight decision variables which you are going to take.

And for the case robust and same would be the case for reliability one, for the robust optimization we will consider some perturbation like atom basic vibrating with along its mean value or the nominal value and the perturbations would give you the variability.

Now the concept of distribution would obviously mean that it is a parametric distribution we are aware of the parameters. And if you are aware of the parameters we will utilize their mean value and their variances to basically formulate the problem of the problems as required because the variability will come into important effect because we will be considering the concept of reliability based on which we will try to model the formulation and solve that.

Now later on at the fag endof the last lecture which was the 50 second lecture I was I basically give the example that in hypothesis testing you have a level of confidence which is known as alpha but the background of the problem is that as I mentioned in the bank managers problem, he or she wants to give the loan but there would be both alpha and beta types of risks.

Alpha being in the case we have considered that the, that is basically business loss in the sense people who could have returned the loan or denied the loan and beta loss is basically for the people who should not have been given the loan or given the loan which is the bad debt. We are trying to basically minimize both alpha and beta the same time is not possible because if you see the diagram if you remember the diagram which I drew very clearly that moving the green line which was at the level of 60 or 60 points on to the right from your side if you are considering move it on the right.

There the value of alpha would be increasing while the value of beta will be decreasing but if you basically move beta on the green line on the optimum point line based on which you will you will you will give the loan. If you will move in to the left where the value of alpha will be decreasing while the value of beta will be increasing but trying to understand the problem in both these perspectives we will see that trying to minimize the sum of alpha and beta will be the best option and you will see later on or else we would have already seen in DADM 1.

That trying to keep fix beta at a certain level and trying to basically minimize alpha would be the best policy action. You can also do the other way round like trying to basically fix alpha and trying to basically minimize beta can also be done. Now in the case of the reliability based optimization we saw that depending on decision maker who is making a decision considering any of the constraints are of the less than type or the greater than type. Say for example the variable the constants are given g1, g2, g3, g4.

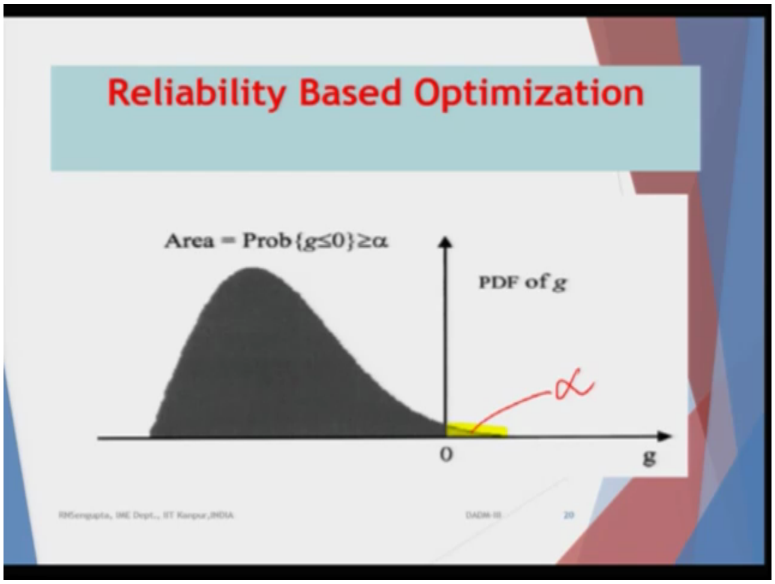
So these are the number of constraints. So considered g as a functional form less than equal to 0 and that the probability that is true is given basically given by greater than alpha where g is a functional form consisting of X which is the decision variable, D which is deterministic parameter and P is the probabilistic parameters. Now remember one thing that the value of X can also be non deterministic in the sense like say for example I am considering the prices of stocks and I am considering the speed of the car, I am considering the rainfall.

So they would basically have some distribution or a flow of fluid in a channel. They would form some distribution. So obviously when I if I have a case that if there are normally distributed then trying to basically combine them and find out the multivariate distribution which is the multi normal distribution, multivariate normal distribution will be very easy because the mean values would be calculated from the population.

And if the population is not given you can basically utilize the sample mean to basically estimate the population mean and similarly given the sample the population variance or the variance is not known you can utilize the sample variance in order to find out what is the relationship between the random variable and then utilize the normal distribution and then try to basically find out what if given the value of alpha and beta. What is alpha and beta I am going to come to that later on.

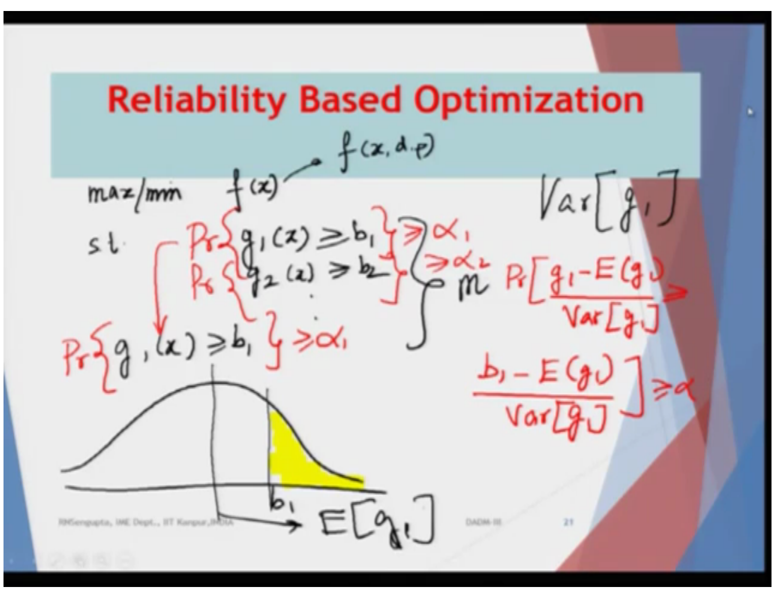
Consider for the timing it is alpha you can solve the problem. So here what I will do is that I will basically quickly go on to the go through the concept and later in maybe in the last week we will basically solve some problems. So the overall general idea is like this probability of that function g is less than equal to 0. This 0 need mean not be the 0 l. It can be any positive or negative value also.

(Refer Slide Time: 06:14)



So probability of g is less than 0 is greater than alpha and we know and we have already shown that the value of alpha is this onto the right hand side considering is less than 0. Greater than 0 will just flip the diagram on to the right and it would just be on the left hand side. So whatever I am going to consider or say would basically be applicable for the both the cases. So this is alpha. Now here I will give a pause and basically draw one diagram then explain that. So this is also we have done so you need not worry that this is a very new concept.

(Refer Slide Time: 07:53)



This is a concept which is very similar. We have done it either in this course or definitely in the course. So it will be easy basically if I create a blank slide. So this is now done. So here what I was considering. So consider the univariate case. There is only one constraint so before that let me write the general optimization problem which we have already done but I will still repeat it. So you want to maximize or minimize function f of x.

So this f of x actually would be f of x and some d and p parameters. For the timing consider d and p parameters are separate and are all deterministic such that you have one constraint g1 x, this x are vectors greater than equal to beta 1. Similarly g2x is greater than equal to beta 2. Say greater than equal to less than equal to does not matter, so you will basically have some m number of them. Now what you do? You basically add in I am just building of the story.

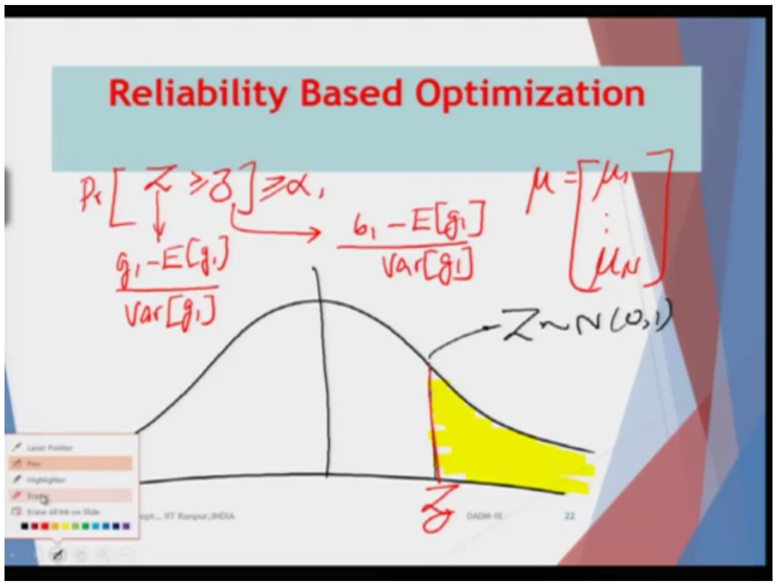
You add the slack or the surplus solve the problems using if it is a linear programming using any of the method, if it is integer programming any of the methods we have already done if it is quadratic programming any of the methods we are going to discuss and then any of the deterministic optimization problem gives you the solution so your life is done and you are happy. But what happens if you now have these. So I am changing the colors so please mark it.

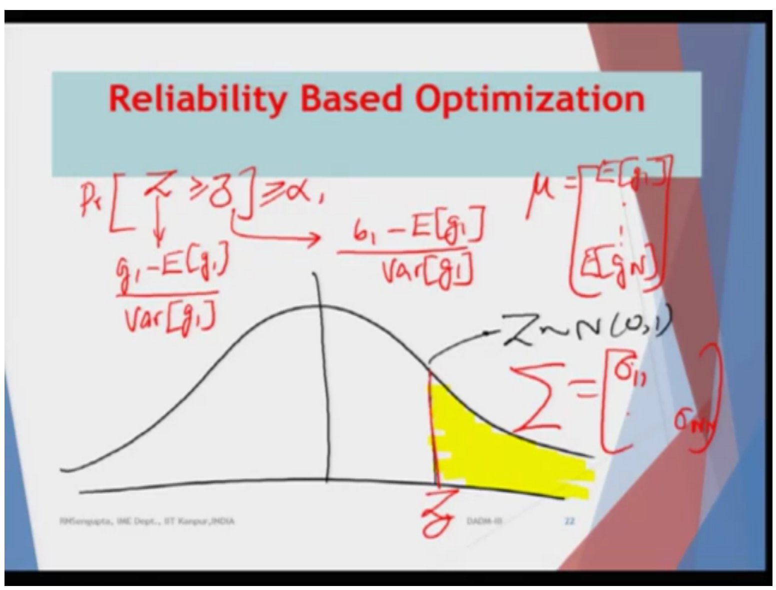
So, all are probabilities. So probability of g1x greater than equal to b1 is and that probability is given by alpha 1. Probability of g2x greater than equal to b2 greater than equal to alpha 2 is given. How do you solve it? So consider one of them and I will use the red color in order to show you how it can be solved. So I write the problem, probability g1x greater than b1 greater than equal to alpha 1. So consider this normal. So how do you proceed? So normality, this is given and this value is greater than equal to some b1.

So consider that the b1 is here, so this is the main value. So this mean value is basically the expected value of g1. And obviously you will can find out the variance. The variance will basically denote by variance of g1. So how do you convert and basically solve it? So this area is given as b1 and this is given. So this is this probability be greater than alpha. So it can be 1 minus alpha and alpha depending how you have formulated problem but I am not going to mark anything less it becomes cluttered.

But let me solve the problem using the simplistic case. So what you have is, I have used the red color so probability is true so what I do is the probability of g1. I am not going to write x dp minus the expected value of divided by the variance I am using the Var is greater than equal to. I am continuing this next space b1 minus expected value of Eg1. This is greater than input alpha. So what is this become?

(Refer Slide Time: 12:43)





So I will use I will save it keep, so it will be easier for me. So I will switch from this slide. This I have done many times but still I will basically repeat it. Few slides extra so this once I convert what it would becomes. It come to becomes probability, so now that it becomes capital Z is greater than equal to small z is equal to greater than alpha 1. So because what is capital Z? Capital Z is g1 minus E of g1 by variance of g1 and what is small z?

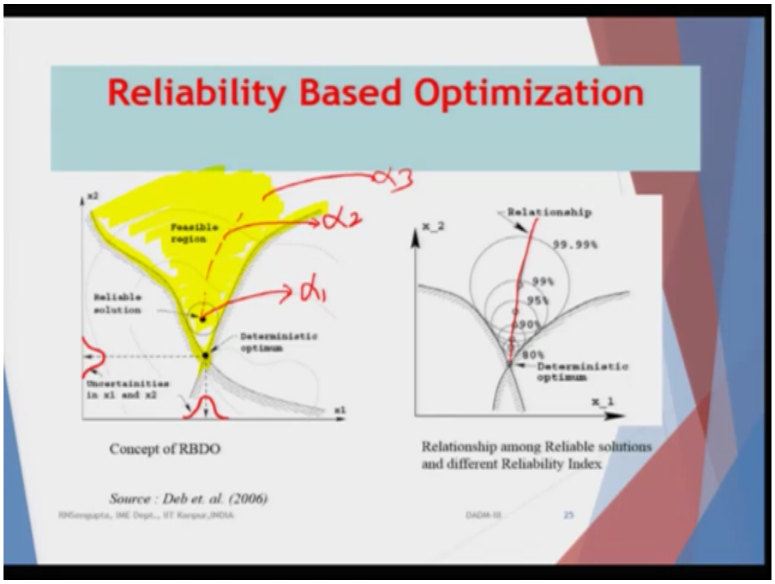
Small z is b1 minus E of g1, E means expected value, variance of g1. Now where does it how does it help? This is the standard normal deviate, this is capital is. This is the distribution of capital Z normal 0, 1 to this Z value would be given somewhere here. So let me use another color so this is Z. So this is small z I have, this value I will have from the table whatever is given because that is given alpha we can find out Z, once Z is found out you can plug in and solve the problem.

So you want to simulate basically simulate find out how many times it is true and for those sets of values where it is true where you have a feasible solution you get the answer. So you keep repeating it you will get the different type of iterative results and solve the problem. In the case if it is multivariate again you will basically formulate the problem considering these values.

So if it is a multivariate one. So in the two dimension one it will look like a the normal distribution as you have in a three dimensional one it will be hillock. And the higher dimensions obviously, it will become it will look like asymmetric one provided the variances are same everywhere. So in the case of the multivariate one you should be aware of the following the mean value of the vector for all axis. So that would be mean mu1 till mu N.

So here this mu1s are basically the expected value of g1. So technically I should be writing I am sorry I should be writing this, so this is E of g1 and the last would be E of gN provided there are N variables. And the value of the variance covariance matrix, the principal diagonals would be sigma 11 to sigma NN and off the diagram elements will be given and you can utilize the standard normal deviate in the multivariate case and solve the problem. So all the things mathematically are known and you solve it and you are happy.

(Refer Slide Time: 15:59)



So now coming back to the background, so consider you have so normal distribution why it is true I am saying that. Consider you have the two dimensional one which you are discussing. So along the x axis you have the x1 random variable along the y axis you have the x2 random variable. This picture will change remember that as we change some of the assumptions.

And considered the overall this boundary of the feasible solution where you are getting the one of the optimum solutions not one of the optimum solution is this. So this is the boundary of the feasible region. So this is for consider this is one of the constraint g1x, and this is another constraint which is g2x. And this point which is black here when I am basically highlighting is the deterministic optimum where you will solve the basic problem you get the answer.

So obviously the inner region, I will use the same color this is the feasible region. And outside you have basically the infeasible region. Now you want to be you want to basically bringing the probability concept into your model. As I said probability means that probability just we solved. So alpha 1, alpha 2 these are the levels of confidence. Now consider both are normal. So why I am saying normal and why this diagram is true look at it carefully.

If I see the distribution univariate distribution for each g1x and g2x, this is the normal distribution which I have for the first case and this is the normal distribution which I have for the second case. Now look at it very carefully, so the uncertainties are there in x1 x2 and this is basically a bivariate distribution and this is orthogonal. This one is also bivariate distribution because it has two variables x1 x2 and both are normal.

Now considered similarly I will keep repeating this with different examples so do not get bored because these examples will be coming up time and again and I will repeat it for making it much more stronger in our concept. So consider the both of them have the same standard deviation and they are orthogonal. So consider the orthogonal concept as like this a circle or normal distribution two dimension is coming from my left and another normal distribution is rising up from the bottom and both are orthogonal and they pass each other.

If you see the common area and if you look at it the top part of the combined area they would be circles. And the circles will be concentric depending on the level of alpha which have assigned for themselves. Alpha means the level or total level of confidence which I have. Smaller or larger the concentric circles with the centre as shown here where I am pointing to the pointer would basically give you the level of reliability.

So more it moves inside the feasible region that means the level of reliability is changing. So in this case the level of reliability say for example would be alpha 1 inside depending on how the locus of this is which is coming into the next diagram would give you basically the different levels of reliability which are there alpha 1, alpha 2, alpha 3. Now why it is a concentric circle because the fact is that the variances are the same.

In case if the variances are not same they would basically be either and they are orthogonal for the timing considered they would be either like looking like rugby ball placed either vertically up or kept horizontally such that the variance in which direction is high. So the variance of the one which is coming from the left if it is higher so it would be elongated rugby ball where the variance in the horizontal direction is lower in the vertical direction is larger.

And if the variants from of the further bivariate distribution which is coming from the bottom if is higher than obviously it will be a rugby ball which is kept horizontal where the horizontal variance is higher. The width the major and minor axis considering is an ellipse is more while the width or the overall variance in the vertical distance is less. Now consider that you have so this basically is a circle. Now consider that you have in the three dimensional one if the overall area you want to consider it basically be a sphere.

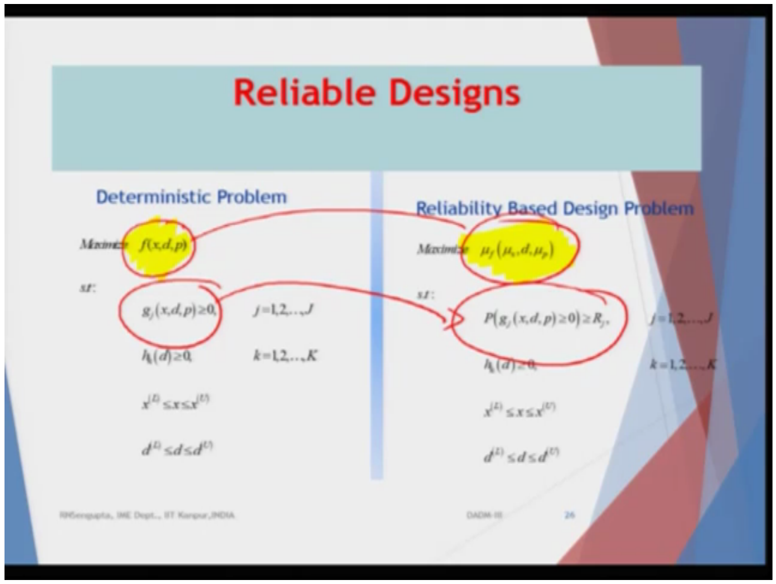
So sphere like a football. So if are considering dimensions like this in the higher dimension case it would be e hyper sphere and if you consider the variances are unequal you obviously you will have a different type off ellipsoids where the major and minor axis depending on the dimensions on the ellipsoid which you have would basically portray that what is the variance which is there which will basically have a effect on the level of reliability and the reliable solution.

Now let us move to the second diagram. So what you have is basically shows you the reliable relationship between the reliable solution and the different level of reliability index. So reliability indexes are changing the relationships also means the loci of the point based on which you will have the reliable solution will move more to inside the feasible region depending on the level of reliability which you have. So as higher and higher level reliabilities are in the case of the deterministic optima.

You will basically have a point and as you increase the level of reliability the circle would be like just like a circle expanding. Concentric circles but at the same time their centre would be moving more inside the reliability area or in the feasible area, so the locus will give you the movement of the reliable solution and the circles would give you level of reliability which is there based on which you can solve the problem. How to solve the problem I am going to come to that later.

So in the first diagram we have the so they are almost the same. Only that in the first diagram I have drawn only one level of reliability. In the second diagram I have drawn different level of reliability and how the overall loci moves.

(Refer Slide Time: 22:51)



Now consider the reliable designs and how we can formulate the reliable designs in the in order to solve the problems and this is where the earlier points which I mentioned that you solve the problem in the X dimension then basically map it in the U dimension.

Find out the MPP point again map it back, do the reverse inverse function into the x space and continue doing it till you are able to reach the optimum point depending on the (())(23:22) or process is basically how the idea works. So consider the overall problem is like this. You have you have a deterministic problem which is given as maximization of function f and there is a decision variable x which we will consider for the time being is deterministic.

Will also consider two sets of parameter for the timing considered them as deterministic which is d and p and the constants are like this. We have been able to define constraints into two sets. One set is basically j in number, j in number of constraints which is g suffix j corresponding to the fact is a function of x dp and all of them are greater than 0. So greater than 0 is just a nomenclature It could have been b1, b2, b3 till bk, bj. And the other sets of constraints are dependent on the deterministic parameters only.

Obviously x would also come which is given by the functional form h. And you have two sets of the decision variables are bounded between X lower and X maximum. So it could have been like X is greater than equal to 0. So in that case the lower bound is 0 and the maximum bound is positive in (())(24:47) they could have been integers also. So that does not stop us from solving the problem. So the number of constraints where we consider the functional form of the constraint as of g.

They are j in number and the number and the number of constraints which are based on functional form of h are capital K in numbers and this D which is the deterministic one will see later is bounded between the lower limit dL and the upper limit of dU. Now we want to basically formulate the problem in the nondeterministic sense. So what we do is that leave aside the maximization problem the objective function for the time being. Our constraints have been modified accordingly like this.

The g constraints which are there they have been if you remember have been divided into 2 sets of parametric variables. One was a deterministic set which is a vector which is d and one is the probabilistic set of parameter which is P. So hence as you basically bring that into the picture your actual the constraints which you had the first set 1st k j number have been transformed into probabilistic sense of g of the functional form being greater than alpha greater than 0 with the reliability of alpha or Rj.

So what is Rj what is alpha and what later on we will see what beta will to come to that later. The set of deterministic constants which are there which is h remains as it is. The decision variables x are also limited between max value being xU and the minimum value being xL, x and U are basically being upper and lower limit and similarly the deterministic one. What is interesting to note is that how we have been able to formulate the functional form of f.

Now if you see the problem formulation I am going to highlight it. It has been done in this way so what you do is that you have to start of the problem. So the best solution to start of the problem for iteration 1 is to find out the main values of x vectors d vectors and p vectors. And here it is what it is coming. If you see x has been replaced. So x is a vector of the decision variable x has been replaced by the vector mu x.

So the first element will be mu x1, mu x2 so on and so forth which means I am considering that distribution of x1 x2 x3 till xn replacing their values in x space from where I will start of the iteration by the corresponding expected values, d set of parameters are deterministic so I am not going to touch them. They will remain as it is when we start and the variables of parameters P which are also non deterministic will replace them with their corresponding mean values.

Now the question is that do we have the mean values? Provided we have the mean values from the theoretical results we will utilize them as the mean values for both x as well as p. In case we have not, the data is set is there will only simulate or do some bootstrapping to find out what are the mean values. So what it means that in x space we start the iteration depending on the mean values of x as a vector d remains as it is, p as a vector. Iterate it find out the first set which will be x1.

Consider x1 star or x1, map x1, I am repeating it so please follow it. So we will basically map x1 using the Rosenblatt transformation into the U space. Find out the MPP point corresponding to u1 which is there for x1, optimize it.

Find out u1 star then map back u1 star into the x space which is the second point depending on the reliability point this x space which is basically x1 star would be now optimize to x2, x2 is again mapped back to using Rosenblatt transformation to the u space and this is which is U2, U2 is basically optimized using MPP methodology to find out U2 star. Then U2 star is again mapped back to x space to find out x2 star and we continue doing this method till we reach the optimum probability. I will come in to more details in the later classes. And with this I will end the lecture and have a nice day and thank you very much.