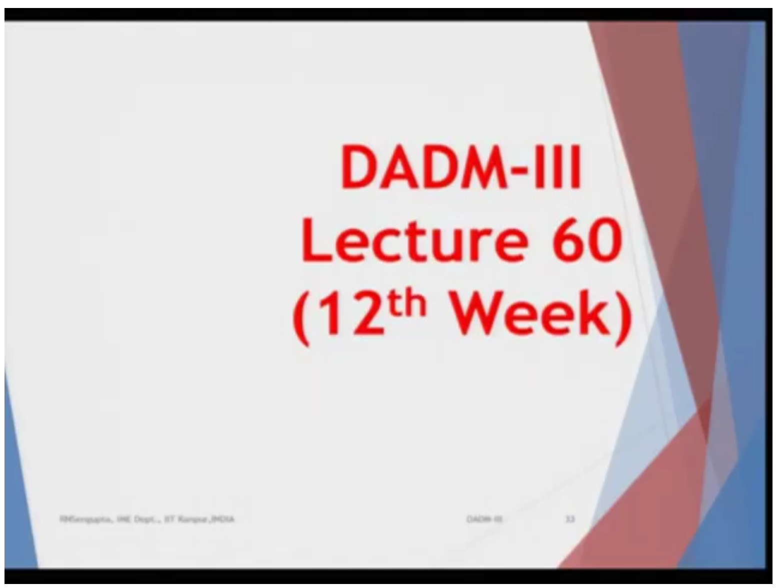
Data Analysis and Decision Making-III
Professor Raghu Nandan Sengupta
Department of Industrial & Management Engineering
Indian Institute of Technology, Kanpur
Lecture: 60
Data Analysis and Decision Making-III

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-III course under the NPTEL MOOC series. The total course duration which I basically I these few things which is basically add before starting of this lecture, in each and every lecture is for 12 weeks which is spread over sixty lectures.

(Refer Slide Time: 00:56)



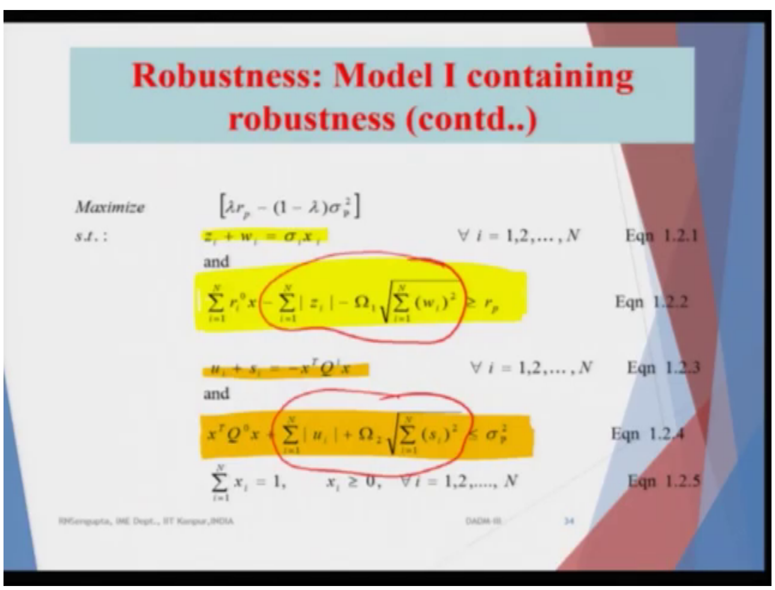
And the total contact hours is thirty because each lecture is for half an hour and in each week we have 5 lectures of half an hour each and after each week you have an assignment. So in totality have 12 assignments and as you can see from the slide we are in the last lecture for this course which is the 60th lecture in the end of the 12th week. And we were discussing about robust optimization and my good name is Raghu Nandan Sengupt form the IME department at IIT Kanpur.

So in the concept of robust optimization we considering the idea that main part is basically to model it and the concept of modeling had been mentioning (())(01:16) in its importance, and you basically consider either the ellipsoidal concept or the interval set. And the ellipsoidal concept this concept of the sets we will consider here in the area of portfolio optimization and I gave a reason for that.

We consider a perturbation based on the nominal value and the nominal value is basically we will be considering the mean value based on the prior data. I will discuss that later on. So our main steps of preceding how to solve the problem would we first discuss that deterministic part solve it, and the solving that would also be a type of precursor which you have discussed in the reliability case.

Then convert that the deterministic one in the in the probabilistic sense with the constraints being probability with the level of betas being given the level of reliability or the robustness. Then convert using the ellipsoidal set and using the nominal values perturbation over and above that, convert them into the robust counterpart, propose the theorems and then solve it using this the simulation method. So once you have the perturbation based on the less than type or a greater than type the I will only give you the essence of the models please please bear with me.

(Refer Slide Time: 02:35)



So we will basically have the same maximization problem where we have given a weightages of lambda to the return of the portfolio on 1 minus lambda to the variances and we are trying to maximize that. The perturbation sets would be based on the fact that we will consider for the less than type and greater than type corresponding to the fact that we have the returns to be greater than equal to some RP or some RP star and the variance being less than equal to some Sigma square P or Sigma square P star.

We basically formulate the produls corresponding to the fact that the ellipsoidal sets are plus type that means this wi being the weights. So they would basically be given by the concept that the perturbation sets would be more important if we consider the perturbation to be on the positive side and if I consider the perturbation set I should be basic to use a different color.

If I use the portable N sets for the variances which are the less than time obviously a minus sign come because I would be more concerned if they are going on to the left hand side for the returns and if I will be more coincident for the risk if they are going out of the right hand side. So remember this line which I drew. For t1 being greater it is good for the returns and for t2 being less on to the left hands is good for the variances that is what we want as an investor.

So once you basically do the simple model formulation for the box and ball plot and then when you consider the interval set for them you have basically the formula corresponding to the return when the probabilistic constraint is converted into the robust counterpart considering the box and ball plot is the following.

Here ri0 are the nominal values for the ith stock. This box and ball plot concept which you are considering using the L infinity norm and L2 norm would basically give you the part which is here box and ball counterpart and that would be greater than RP because it is going to the right hand side if you remember.

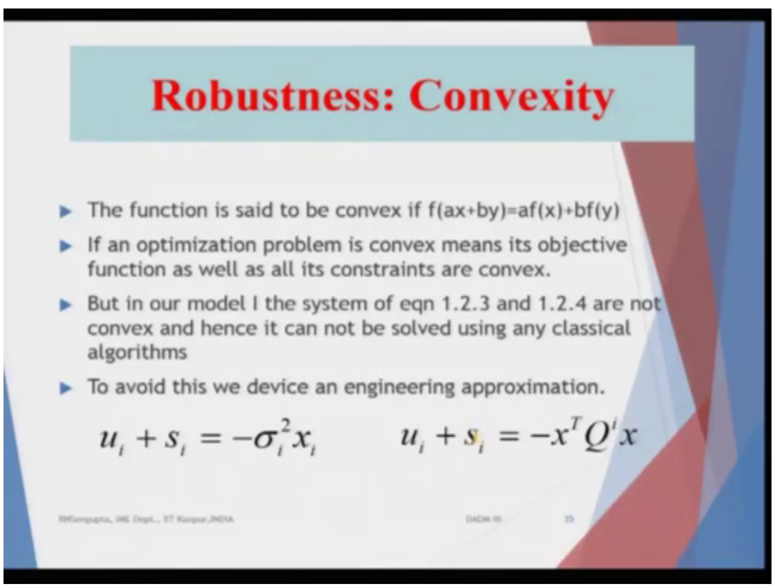
And in the case when we consider the box and ball plot for the variances is it will be given by this. So is the less than type so obviously we add the box and ball and take the intersection and this would be of the plus type because they are on the left hand side. Again Q naught Q suffix naught are the nominal values of the variance covariance matrix.

So basically we will be taking the average of the variance using the concept of bootstrap whatever it is and the of the diagonal element would also be given by the nominal values for each and every stock taken individually.

And here as you remember as you know the XS are the concept of the perturbations we are going to consider for the box and ball. The last constraint where we didn't have any perturbation was basically the sum of the weights is equal to 1, and here we are going not going to consider any short-selling. So the probabilistic first constraint is this one which I am highlighting within the yellow color has been converted into a robust counterpart and the robust counterpart for the second constraint which was related to the variance is this one which is highlighted. So objective function does not have any perturbation because it is not we do not consider robust obviously we can consider.

So once you have the model converted with probabilistic constraint into the robust counterpart you will use simulation methods and solve them and basically give the results. Now remember one thing changing the values of beta 1 and beta 2 would have a consequence on the concept of the box and ball probability levels. So they would dictate how your results would be which I will come when I consider the results accordingly.

(Refer Slide Time: 07:39)



Now we consider the concept where what is convexity and we know that we have consider convexity in more details. So I will just read it with it will just be a repetition.

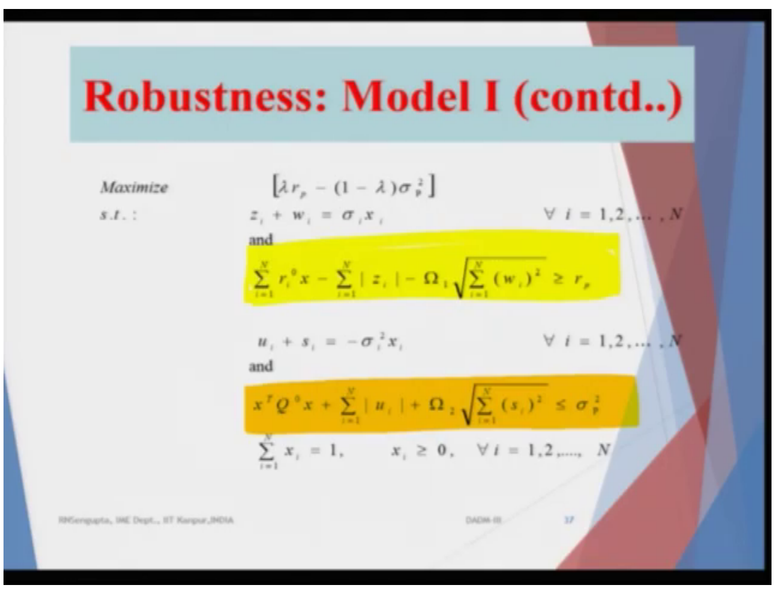
A function is said to be convex if F of lambda X plus 1 minus lambda Y here lambda and 1 minus lambda are considered an A and B is equal to or less than equal to less than equal to time and less than type or greater than equal to greater than type would basically becoming from strict convexity and convexity and strict concavity and concavity.

We have already discussed that when we are considering the concept of convex functions the semi definite, definite, semi and semi definite positive, definite positive, semi definite negatives and definite negative. So depending on the less than equal to less than or greater than equal to greater than sign. If an optimization problem is convex it means it is objective function as well as constraints are convex and we can solve it.

But in our model the system equations which is 1.2.3 and 1.2.4 which is the first and the second one and not convex hence it cannot be solved using any classical algorithm. To avoid this we devised an engineering approximation here it is. For the counterpart which when we are dealing with the variances will consider that the variances are given by Q the variance covariance matrix of size N cross N.

So here we will consider any perturbation is happening in a sense that if it is the less than type will basically add a negative counterpart if it is a greater than type will be basically at the positive counterpart and basically solve the problems accordingly.

(Refer Slide Time: 09:40)

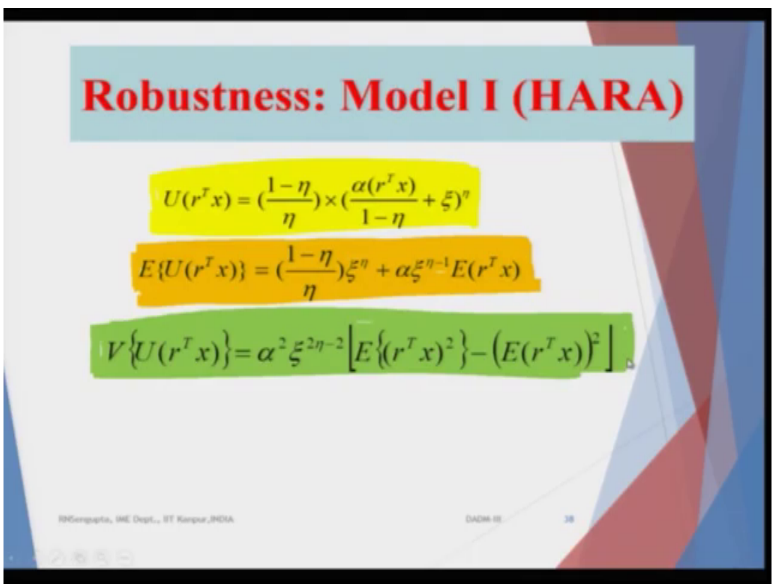


So this was the just repetition this was the model 1 and in the model 1 we had I am just repeating it maximizing lambda rP minus of 1 minus lambda Sigma square P because minus sign coming because that will be when I add the maximization becomes a minimization. Probability of the return of the portfolio is greater than equal to rP with the probability beta 1 and probability of the variance of the covariance of the portfolio being less than equal to Sigma square P has a probability of beta 2.

This beta N, beta 1 and beta 2 can be changed and that would basically have a consequence on the values of the box and ball values based on which you are going to solve the problem. Again some of the weights is equal to 1 and there is not short-selling so obviously the weights can be has to be greater than 0.

So these are again a repetition. The first constraint when converted into robust counterpart using the level of beta 1 is this. And the second constraint considering the robust counterpart when the probabilities one considering the beta 2 level of probability are converted is this. So we can solve it using simulation method and get the result.

(Refer Slide Time: 11:14)



Now we consider a model which is a little bit different in the sense that we will consider the concept of hyperbolic risk functions which is Hara and what is the story behind that? Now the story is like this so for any decision when you are trying to basically buy and this again I am going to the concept of utility for any decisions we have the investments in place and the investment gives it a utility.

Now what is important to note that if your neutral functions are quadratic then obviously we know that the returns would be normal and vice versa. This I had mentioned that not in the area of DADM-III but I have been mentioning that in quantity finance time and again.

But the fact is that the returns of the stocks of the scripts are not normal. They are EVDs, with type 1 type 2 distributions, so gamble distribution and all these things. Now how do we model it? The answer is intuitive and simple. What you consider is that, you consider I am not going to go through the proofs I will only state the sequence of how the procedure is done. We will consider the distribution to be gamble or type 1 type 2 EVDs with some parameters alpha beta gamma.

So these are shape, scale and location parameter. So how we find it I am going to come to that later on, so let us keep that aside. Consider alpha, beta, gamma are known for the EVDs. Now what we do is that we consider the utility function based on the investment as Hara and using Jacobean transformation we convert and try to find out that if the distribution of the returns are EVDs then what is the distribution of the utility function.

Once we find that utility function, we need to basically find out so utility function returns for each and every EVDs are given we find out using the Hara, using the Jacobean transformation we combine them, so what is the reason of combination? So this utility functions based on EVDs for each and every stock when they combine would give me the overall utility function for the portfolio.

So now we are intuitively assuming that EVDs is being true for the portfolio distribution I would basically have a utility which I need to find out and that what we just would do using the concept of Jacobean transformation. Once that is done they would be a functions of alpha, beta gamma. That is those parameters which we had in the EVDs. Now how would we find out the alpha, beta, gamma.

Obviously we will need to basically estimate them using the alpha hat, beta hat and gamma hat. What we will do and I will come to that later in that, given the data set for the EVDs for stock 1, stock 2 till stock N or script N will basically do the bootstrapping and find out the different parameters alpha hats, alpha hat 1, alpha hat 2, alpha hat 3 for all the scripts similarly beta 1 hat to beta N hat.

So this beta should not be confused with the level of reliability beta 1. Obviously we are using the same symbol but let us not get confused. Then we would have the gamma 1 to gamma N all the hat values the estimated values and we basically find it from the sample using bootstrapping method. Once they are found out they are put back into the utility functions for individual scripts considering EVDs and an utility function based on Hara.

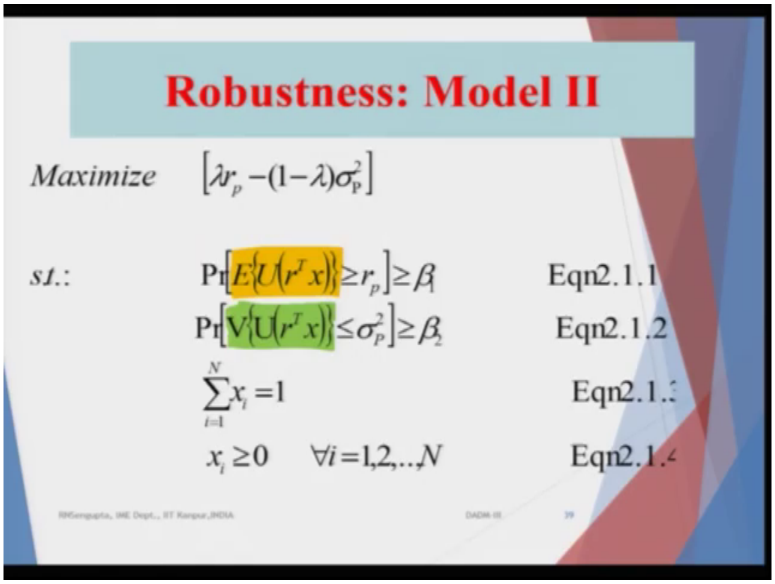
Then we will combine them to find out the combined utilities for the portfolio. Once the distribution is found out for the combined portfolio we need to find out two things one is the expected value of the portfolio considering the HARA utility function to be true and one is the variance of the of the utility for the portfolio considering HARA utility function to be true, why we need the first moment in the second moment?

The reason is that we want to basically maximize the first moment with this expected value and we want to minimize this the second moment which is the variance. So this is what we are doing once you find out the utility, the utility function is given. I will just highlight the values using different colors. So this is the utility function. And this Eta Zeta all these values are the parameters for the Hara utility function.

The expected value or the utility function is given for the portfolio is given by this. This I said I am not going to go into the proof I just gave the general idea how you solve it, and then you find out the variances. Now only point what added extra step in calculation in the various would be that you will basically have a Taylor series expansion a form which you need to basically expand in a Taylor series expansion and basically ignore the higher terms.

Because they would basically be tend to 0 because the value of return off on any particular stock is 10 to the power minus 2 so any square value cube values would be 10 to the power minus 4 or 10 to the minus 6 hence we basically ignore them. So if we consider the variance or the portfolio this is the variance of the portfolio considering the Hara utility function and based on that we will basically right to proceed and do the calculation.

(Refer Slide Time: 17:47)

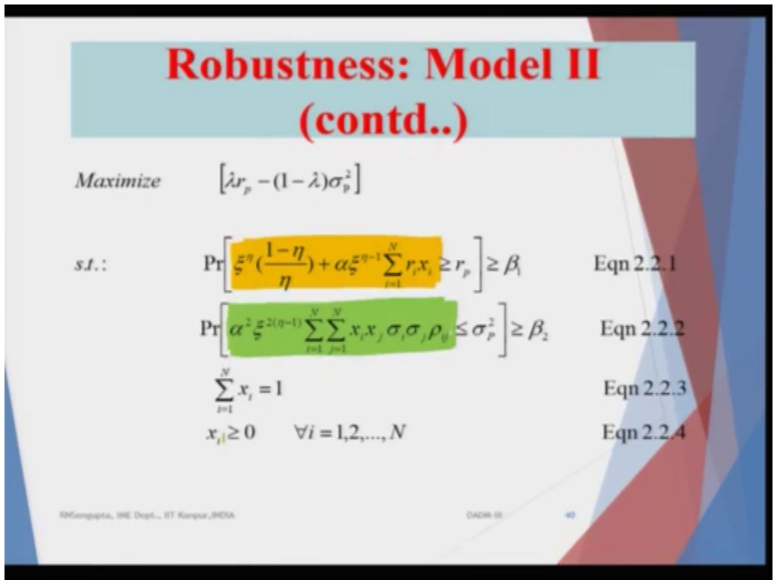


So now the problem is again the simple way. Look that at the maximization problem. Maximization problem is basically the same one tried to basically maximize us lambda into rP some value which you want to keep increasing that means we are trying to increase t1 more on to the right, and minimize t2 which is Sigma square P and basically push it on to the left. So you want to maximize lambda into rP minus 1 minus lambda into Sigma square P. P suffix is basically for the portfolio.

Now the constraints are interesting. I will use the colors accordingly the probability of the expected value of the Hara utility function for the portfolio is this formula which you have already calculated. So that has to be greater than rP and that probability is beta 1. So in the case the yellow colour it is this orange color which had been utilized if it was normal distribution would be the expected value of the portfolio the multivariate distribution be true for the portfolio if symmetric distribution was considered to be true considering the Marquis Principle and model.

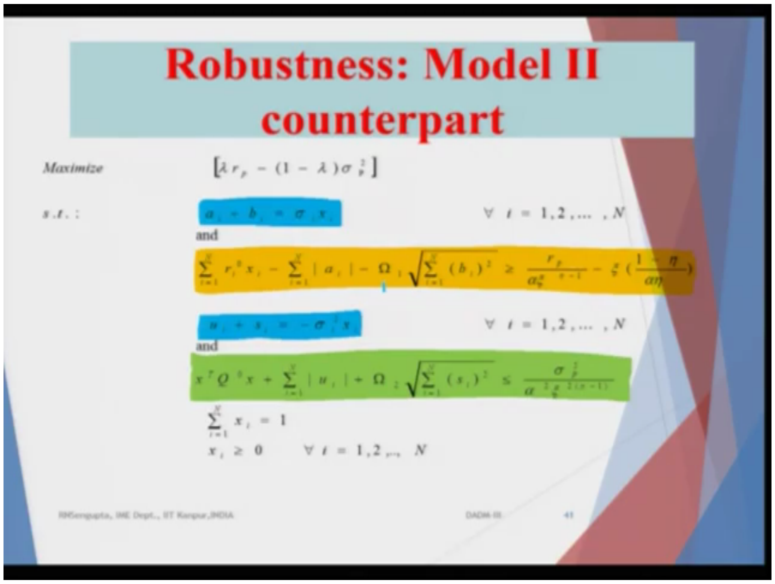
And in the case if the variance is considered using the Hara utility function for the portfolio this is the variance and again if you consider the Hara of this Marquis model to be true they would be replaced by the variance covariance matrix. So now again you will basically have a model to formulate using the probability for the return of the Hara utility function to be greater than rP with the probability of beta 1 and probability of variance of the Hara utility function for the portfolios in both the cases to be less than Sigma square P and that probabilities is greater than beta 2 again we need to basically formulate using the robust counterpart.

(Refer Slide Time: 19:52)



So we write the equations as it is. Here is the expected value of the Hara utility function. The objective function remains the same note that and the variances for the Hara utility function is this is less than equal to Sigma square P this is true with the level of beta 1 and beta 2 for the first and second constraint respectively. The summation of the weights is 1 and Xs are greater than 0 considering no short-selling is there.

(Refer Slide Time: 20:38)

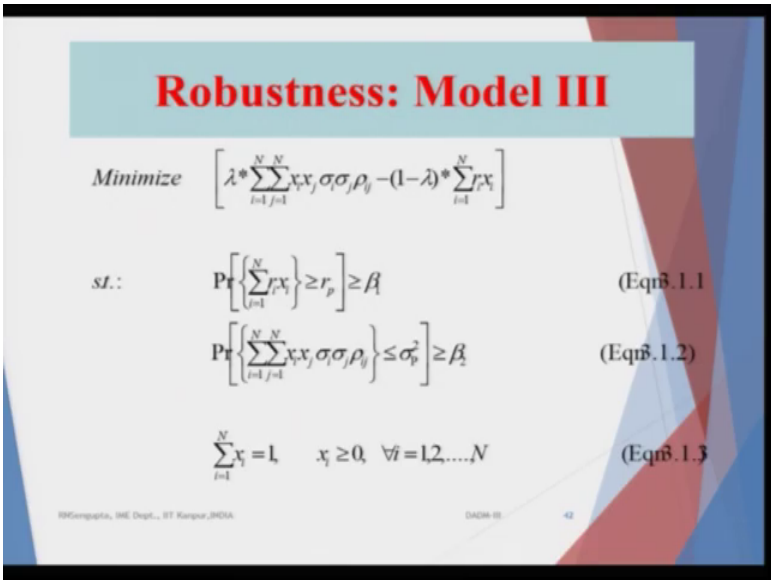


So once you basically do the robust counterpart and do the model considering the less than time and greater than time I am not going to go into the proofs. What you have are these again I will highlight. So this is the robust counterpart for the first constraint using the less-than or greater-than type and here we are using I will use another color blue. So these are the initial equations for the perturbation sets and then you use the box and ball plot.

And the box and ball plot would basically have a consequence and what is the value of capital Phi and the value we would be basically based on the level of beta 1 which we have. That is why we have written capital Phi suffix 1. The robust counterpart for the second constrain our given. So this is the robust counterpart again you see this is less than type. Obviously it should be because it is on the left hand side and the this perturbation sets what basically model using the perturbations as highlighted here.

Again the value of capital Phi suffix stood basically mean they are corresponding to the level of beta 2. So as beta-1 beta-2 change these values of capital phi 1 and capital phi 2 would change and you would give you different results. The third scans constrain summation of Xi is equal to 1 remain same Xi being greater than 0 remains same.

(Refer Slide Time: 22:54)



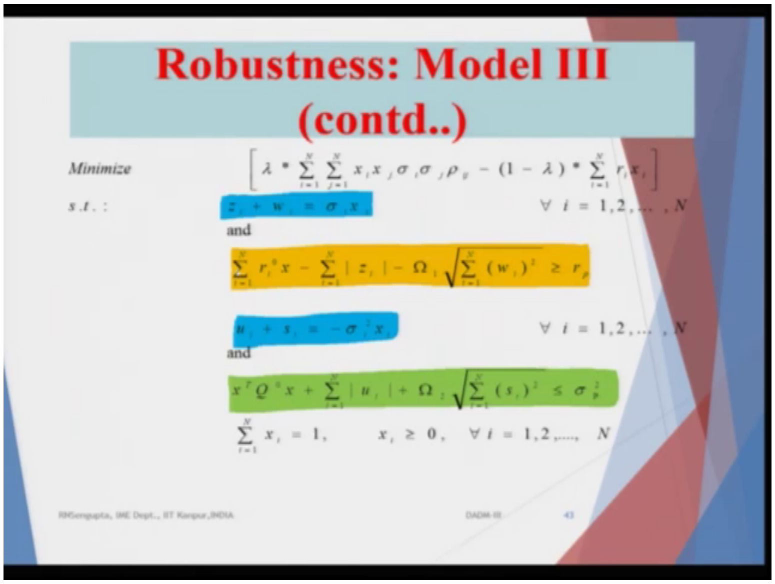
The second model is exactly the same only that we are putting the levels of lambdas on the levels of weights on the variance covariance matrix and 1 minus lambda for the return on the portfolio. Again probabilities are beta 1 beta 2 for probability of the risk return being greater than equal to rP.

Now remember one thing the change in the problem is happening here. You are trying to minimize, minimization and maximizations are not important, what is important is to note down what is being multiplied with lambda. In the initial problem you are basically multiplying our value of lambda with rP and you are trying to basically maximize my and minimize rP.

But here now trying to basically bring the value itself of the portfolios risk and return so that risk is basically given by the double summation of Xi into Xj Sigma i Sigma j Rho ij while the value of the returns are given by the summation of our ri into Xi. So in one case you are keeping the values of rP and Sigma square P as in their objective function in the next model which is in front of you, you are trying to basically keep it as the value being which is being calculated from the portfolio by itself depending on the weights which you are going to invest for each and every stock.

The third constraint which is summation of xi equal to 1 remains the same. The fourth constraint xi is being greater than equal to 0 for no short-selling remains the same and the probabilistic parts for first constraint and second contents are the same.

(Refer Slide Time: 24:40)

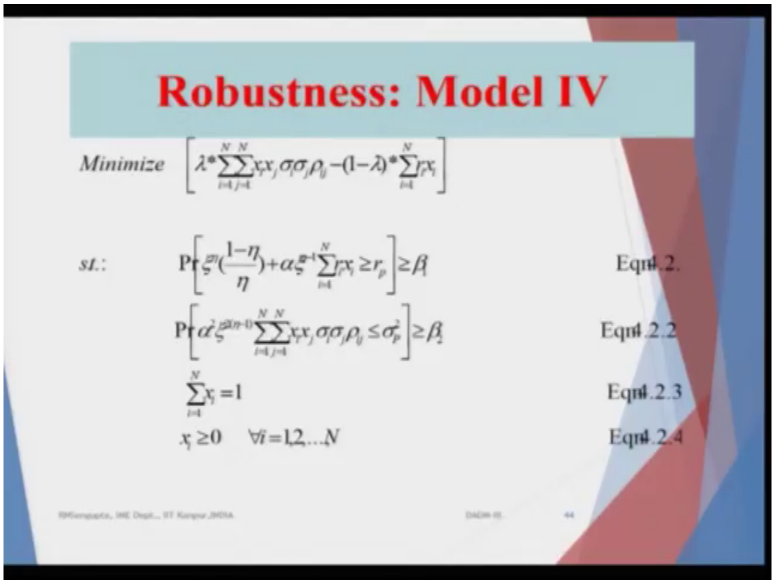


Once you basically again consider the robust counterpart considering the perturbation sets, I will highlight the perturbation sets using the blue color. So this is the perturbation sets corresponding to the first constraint which is to do with the mean value of the portfolio. This is the perturbation sets corresponding to the variance for the second constraint and that has to do with the variance only and the changed constraints depending on the robust counterpart are basically given I am using the same color. This is the robust counterpart for first constraint and this is again utilizing the concept of box and ball plot.

Again capital Phi suffix 1 basically means this is the value of Phi 1 corresponding to beta 1 and this beta 1 value changes would also have a consequence and change in capital Phi 1 would also result. And similarly when I consider the corresponding robust counterpart for the second constraint I will use the same color green. So this is the change constraint for the robust part for the second constraint.

So here capital Phi 2 suffix 2 is basically the value of the reliability corresponding to the beta 2 value and again these values I am not repeating, r1 naught and Q naught both in the other problems, r1 naught is basically the mean value of the return and Q naught is basically mean value of the variance covariance matrix. This will basically simulate it using the bootstrapping and then consider those values. And this third constraint summation of Xi is greater than 1 is equal to 1 and Xi is are basically greater than 0 remain same.

(Refer Slide Time: 26:47)



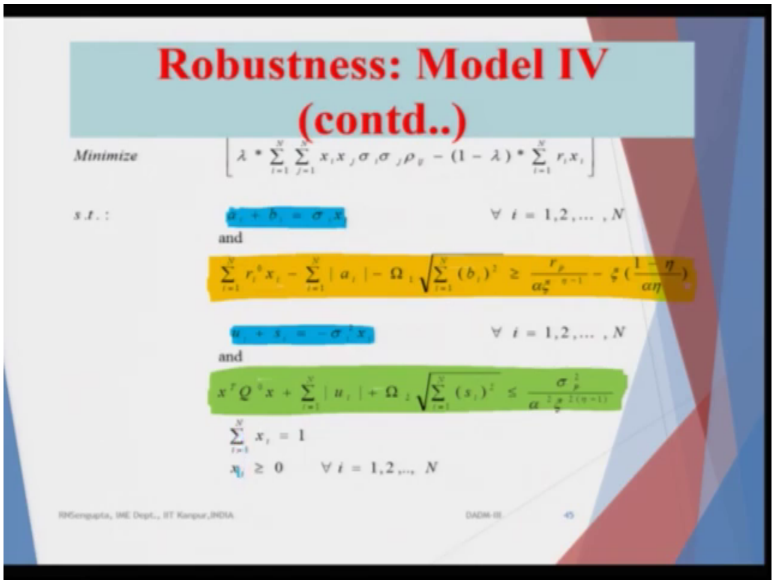
Now I consider the model where I have the minimization being true but this minimization are being considered for the case when we have the risk and return corresponding to the Hara utility function, so minimization for the objective functions remain the same as in the third model.

This is lambda into variance covariance of the portfolio minus 1 minus lambda because you are trying to minimize. Minimize or negative value would basically be trying to pull it up. Minus 1 minus lambda of the return and the probability corresponding to the first and the stress second constraint have the reliability levels of beta 1 and beta 2, but the values inside the bracket are interesting to watch.

In the first case it is basically the so called return calculation which we found out as the expected value of the Hara utility function based on the fact that we are considering the EVDs to be true for each and every stocks. So this is the part where I am just highlighting I am not going to put the color but where the pointer is this is the part greater than equal to rP is basically the value which I have for this the return of the portfolio considering the higher utility functions.

Similarly the second part is basically in the second probability is basically the part related to the variance or corresponding to the utility function from the ri utility function. The third constraint remains summation is equal to 1. The forth constraints are related to each and every investment is greater than 0.

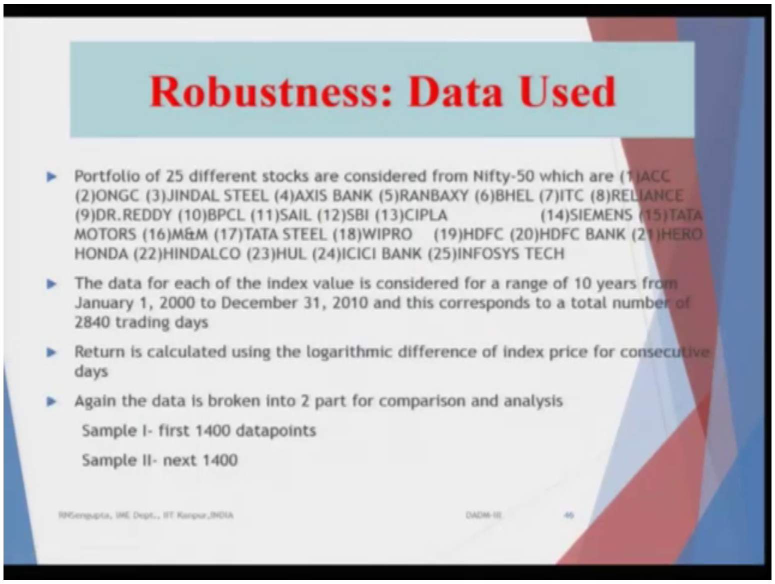
(Refer Slide Time: 28:25)



The counterparts are given I will just highlight them. This is, so this is for the first constraint depending on beta 1 whereas capital Phi 1 would basically be calculated using beta 1, and the second constrained counterpart in the reliability sense are is this. Again the less than type greater than type equality would be used.

And this capital Phi 2 would be calculated from beta 2 and these are the perturbation sets for the second case and the first case. So I will just request one thing because today being in the last class I will just extend it for by about 5-7 minutes. So please bear with me. The third constraint remains as summation is equal to 1 and the fourth constraint is Xi are greater than 0.

(Refer Slide Time: 29:32)

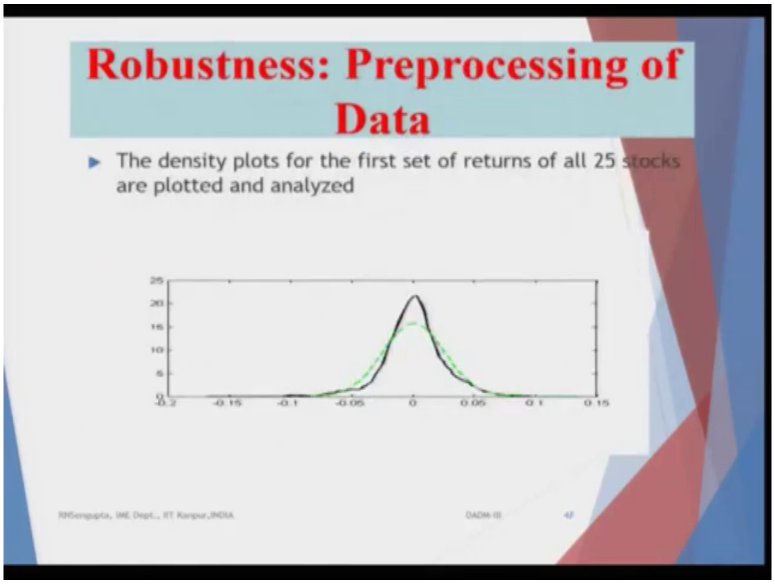


So we take the stocks off of 25 different stocks from nifty 50 starting from ACC to Infosys Tech Jindal, axis bank, Ranbaxy, BHEL, ITC, Reliance, Dr..Reddy, BPCL, SALE, SBI, Cipla, Siemens, Tata Motors, Mahindra and Mahindra, Tata Steel, Wipro, HDFC, HDFC bank, Hero Honda, Hindelco, HUL that is Hindustan Unilever limited, ICICI and Infosys Tech.

The data is taken for indices considering the 10-year range from December from January 2001 to 31st December 2010, which is 2840 days. We find out the max and the min considering the EVDs to be true. We basically find out the maximum the return, on the minimum return basically sort it out.

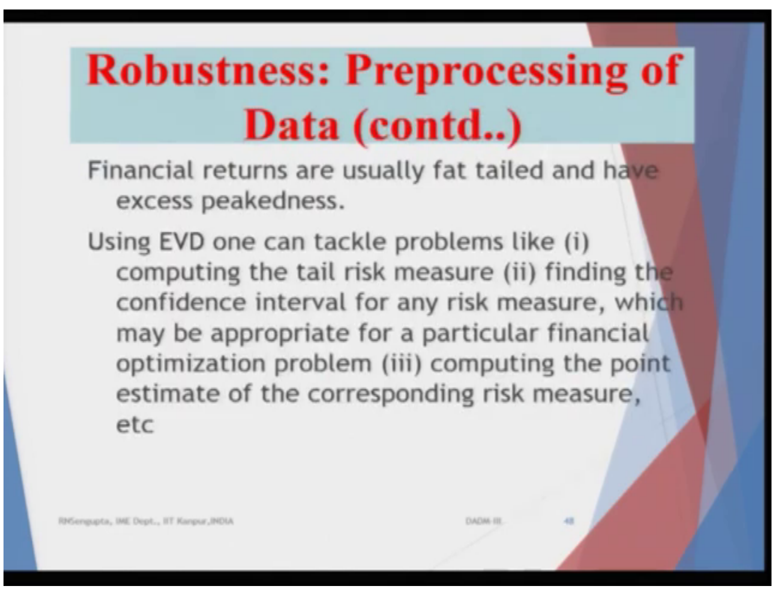
We find out two sample sets one is for in sample and out sample. The in sample basically would have about 4. So we basically divide 2040 into 2 equal sets. So we consider 1400 for the first sample and the 1400 for the second sample and for both the samples we find out the minimum and the maximum.

(Refer Slide Time: 30:45)



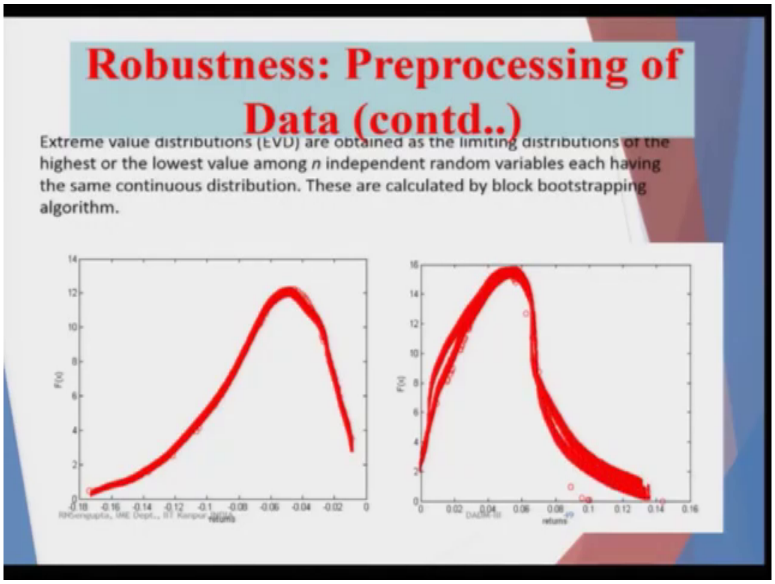
So this we do give us a density plot for all this 25 stocks. So they are extreme values depending on whether you want to take the left hand side or the right hand side of the distribution considering skewed on to the left of the right.

(Refer Slide Time: 30:59)



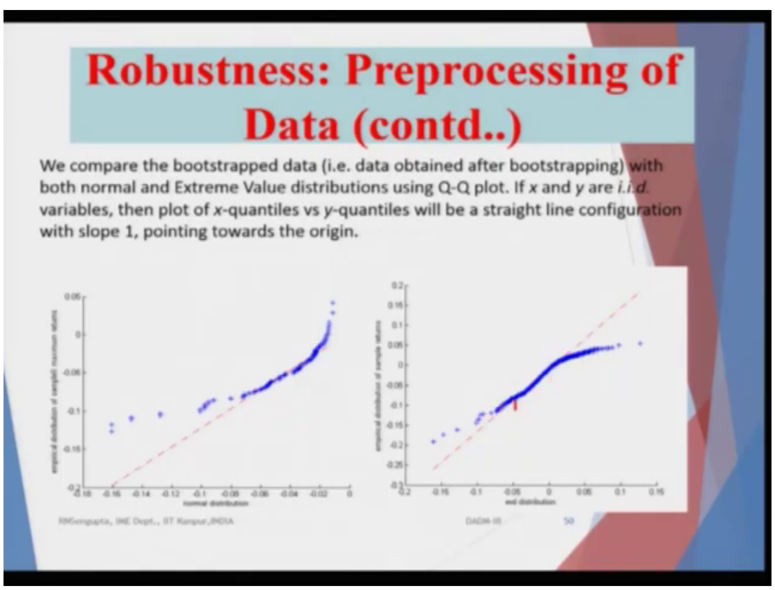
Financial returns as usually flat and have expected picnics we use the concept of EVD that has been telling you time and again.

(Refer Slide Time: 31:10)



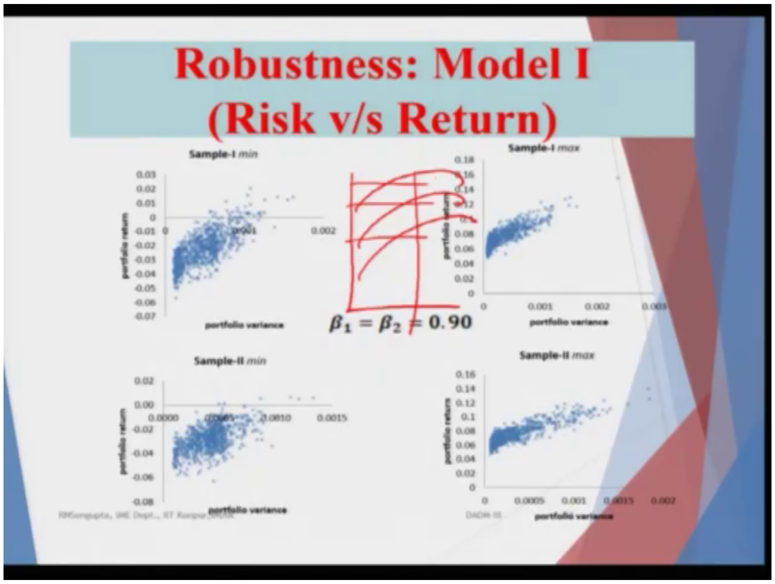
So I just give you the general results. So EVDs for the right hand side and the left hand side are basically the returns corresponding to whether the positive one. So I will just highlight using the red color. So this is the positive part which you have and this is the negative part which you have so like this so either skewed on to the left or the right.

(Refer Slide Time: 31:38)



And we compare the bootstrap results to find out whether the values are the QQ plots are related to the EVD distribution or normal distribution the normal distribution does not hold but the EVD distribution for the central region are quite visible and they agree with the fact obviously you have to do better bootstrap results and take the data accordingly.

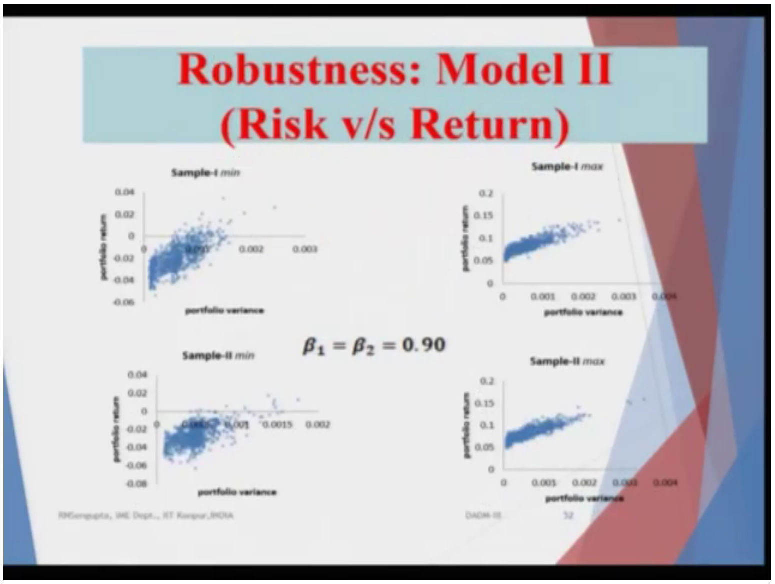
(Refer Slide Time: 31:59)



So I will just give you a few sample sets. I will only give you the highlighted points. So we take basically sample 1 min max, sample 2 min max and give the portfolio returns and the variances for model 1. So they basically formulate if you do about 1000 or 2000 simulation runs they basically agree with the fact of the mean variance theorem.

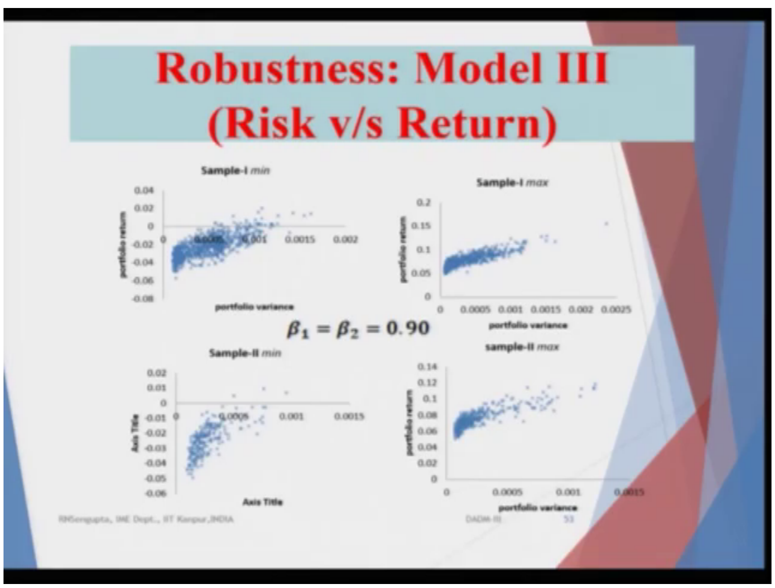
But the values what is interesting to note is that if you keep changing beta naught and beta 2 they would be conjugating on were concentrating envelopes, overlapping envelopes where you can find out to what level of beta and beta 1 or beta 2 you basically need to find out the level of risk and the return. So they would be like this so each levels of betas would give you the risk and return level for the different beta. So we have basically just giving you a beta 1 and beta 2 as 90.

(Refer Slide Time: 33:00)



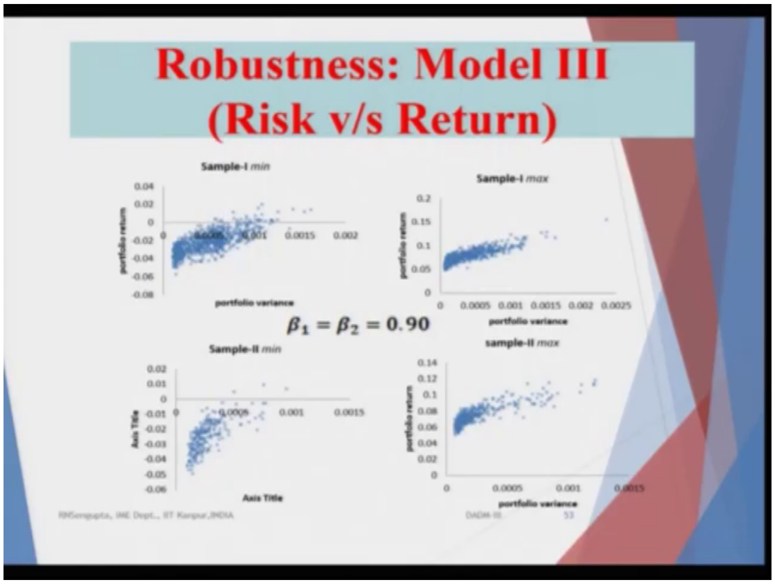
So this is for model 2, so that general structure remains the same. So this is for model 2.

(Refer Slide Time: 33:05)



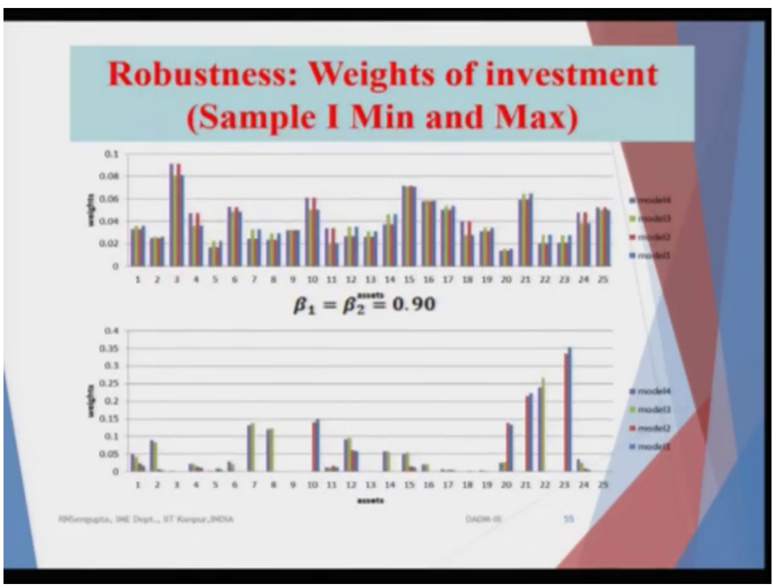
This is for model 3, so risk and return profile the frontier is same.

(Refer Slide Time: 33:14)



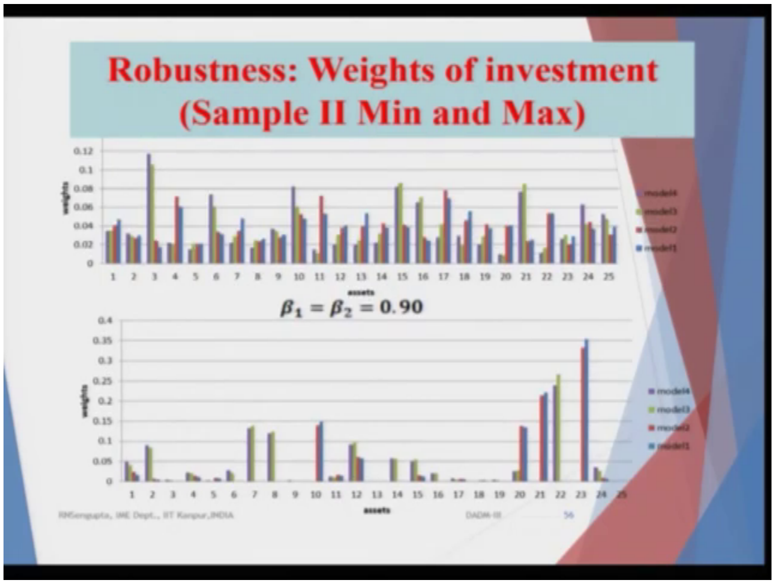
And this is for model 4.

(Refer Slide Time: 33:18)



Then we come to highlighting the sample 1 min max weights for the portfolios of the scripts. So this gives you a distribution of the weights this is nothing to do with the actual operation but they give you the results.

(Refer Slide Time: 33:34)



Then I find out the weights for the investment for sample 2 for the min and mean man max. So with this I will end this course of DADM-III and before I close the course I will take another 7-8 minutes. I want to thank all of you for your patience.

We have covered yet taken a long journey starting from concept of linear programming optimization the latest area of what are the concepts of the slacks, the surplus then the concept of the weights, why they are important. The concept of the nonlinear programming, the this concept of (())(34:20) cuts and the concept of branch-and-bound.

Then we went on the concept of reliability optimization, robust optimization and we consider a different type of problems accordingly. So obviously we not may not have been able to cover all the topics in in optimization but it basically I am sure it will give you a good feel that how you can proceed and basically try to pick up the concepts and there are many interesting areas which are still left to be explored corresponding to the fact that the DADM 3 course was basically for 30 hours.

I want to basically thank all of you for your patience and all the queries which have been there. We have tried our level best to answer them. If you are not if you have not been able to answer all the queries I apologize and I am sure that in future any of your correspondence being sent to NPTEL office or to me individually we will be able to handle and satisfy all your queries on this academic front.

I want to thank all my TA for this course. I want to thank all the staff in NPTEL office at IIT Kanpur and all the people who have been able put all their efforts in trying to edit the videos, lectures here and also IIT Madras which is the nodal agency for the NPTEL course. I am sure all of you would definitely be motivated by these NPTEL courses and take up some of you would definitely try to basically read further and we are there as a teacher, professors, tutors.

We will definitely be there to help you out and encourage students who are really willing to basically pursue this for higher level in either on a theoretical level or on the practical level. Have a nice day and I would like personally like to wish all of you the very best for a fantastic career either in academic or in professional life and I am sure you will do very well. Have a nice day and thank you very much.