Quantitative Finance

Prof. Raghu Nandan Sengupta

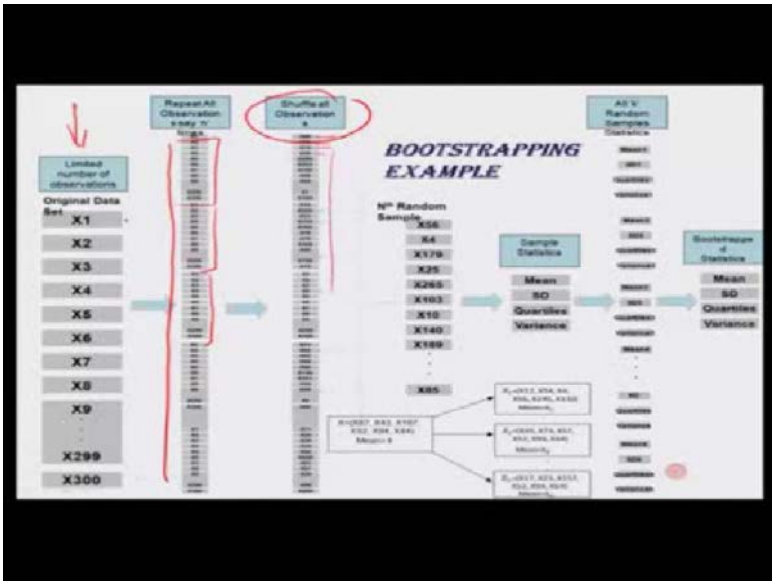
Department of Industrial and Management Engineering

Indian Institute of Technology, Kanpur

Module – 06

Lecture - 36

(Refer Slide Time: 00:18)

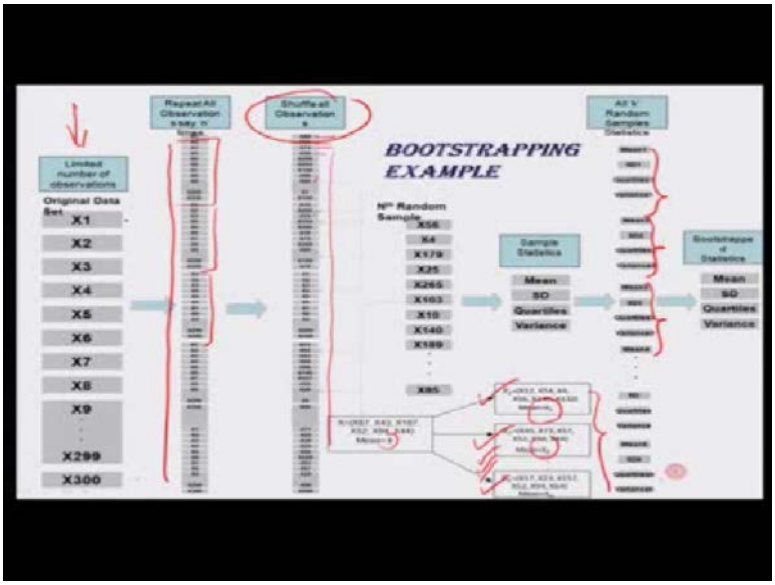


Continuing our discussion about the boot strapping, so, what we have done; we have shuffled them. Now, if you see the third column all of them are basically, jumbled up. Now, what we do is that we pick up such random k; say for example, 40 number of observation rather than picking up from the lot. So, consider 40, we are observation picking up is x 56, x 265 and so on and so forth. Now, once you have that you find out the average of that. Note it down.

Again, we are placing the jumble and again, jumble it. Again, you randomly pick up such; say for example, 40 observations. Initially, it was also 40. Now, it is also 40. This 30 can be 30 also; there is no problem. So, and pick up again, randomly from different positions. You have, they are already jumbled up. If you keep repeating it what you will have? If you say for example, do such 40 number of pickings at each go and you basically, pick up k number of them then the averages of this k number, each are totally different. Then what you need to do is that find out the averages of the average and all

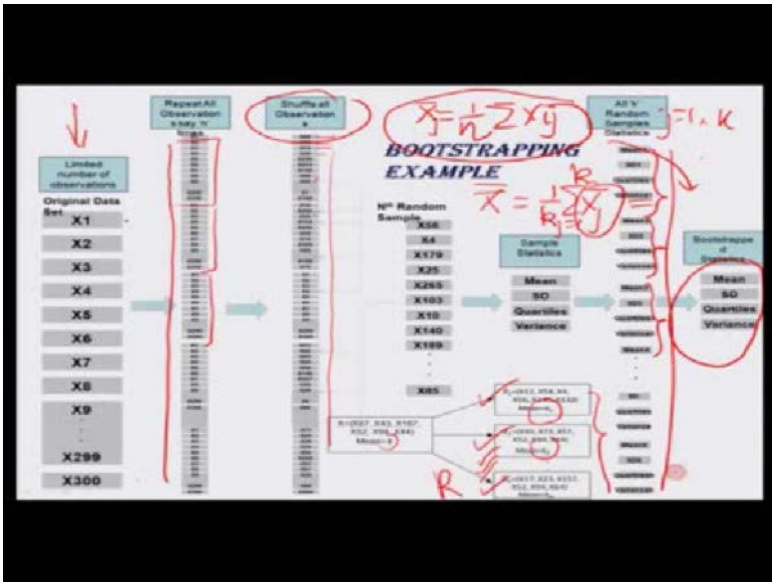
the characteristics, which are needed in order to basically, have a good study about the population, which you need to study.

(Refer Slide Time: 02:00)



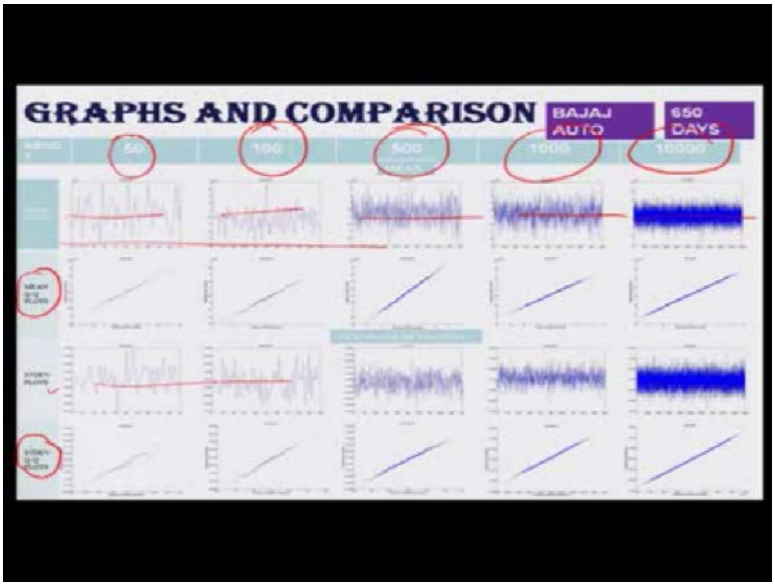
So, if you note down you have picked up x 1 for this case rather than 30 or 40. I am picking up observation 1, 2, 3, 4, 5, 6. This is just an example considering the space of trying to denote in a slide is limited; this need not be filed. So, once we find out, we find out the average from this five, and what we do is that as you keep up picking up observation, this such fives and you repeat it, say for example, 20 number of times; this is the first five; this is the second five; this is the third five; fourth five; fifth five and so on and so forth. It will continue. Once you find out these averages, these are denoted by x bar 1, x bar 2 so on and so forth, and once you have this you find out the averages in order to find out the characteristics, which are of most important to you. So, for the first one which you find out, these are the characteristics. For the second set, these are the characteristics; third set, these are the characteristics so on and so forth.

(Refer Slide Time: 02:33)



So, if you pick up k number of them, you will have such k here and finding the averages of this, you will get all the characteristics which are the mean, standard deviation, variance, courtesies so on and so forth. So, if I am talking about the mean, very simply mean would be, if I pick up n observations it will be this, and if I repeat it k number of times. So, this is, say for example, j; j is equal to 1 till k. So, what you will have is that this x bar you will calculate in the average and this average would be averaged out more in k number of observations. So, first block of say for example, n size; second block, n size; third block, n size; repeated this k number of times and then find out the averages and that will give you the actual characteristics.

(Refer Slide Time: 03:41)



So, now, let us show the actual study which we have done, considering the Bajaj Auto. This data was taken from the year, for the last three years, till 2015 July. So, once I have this, considering that the boot size that you are going to take is 50, 100, 500, 1000, 10000. You see the mean path, the average which you find out and if you keep repeating it, the average slowly smooths out. So, the overall variance which you have would be there, but it would basically be slowly smoothen out.

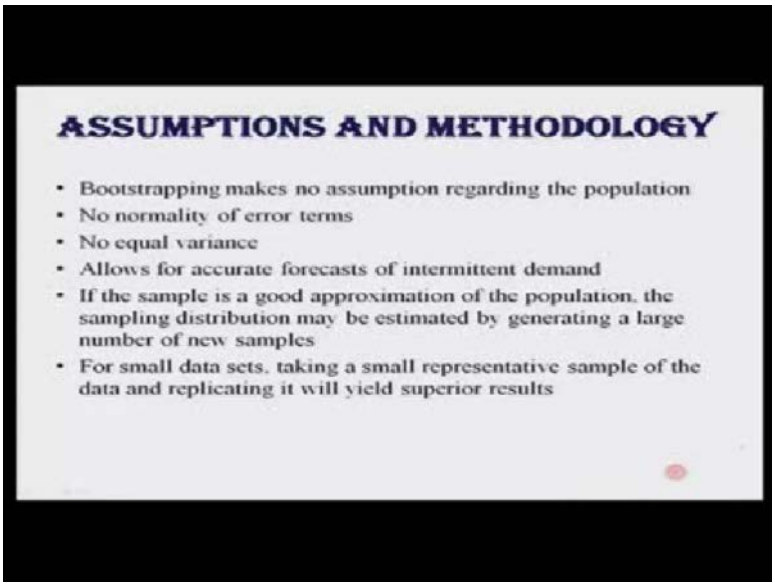
So, you want to find out what happens to the standard deviation. This is the standard deviation which you have. There is a huge amount of fluctuation from the actual fluctuation, keep us down; that is not important. What is important is to note down how does the mean qq prods and the standard deviation, qq prods vary; qq prods basically, means the variation of the actual variable which you are going to study using the quantail quantail prods. Now, if your underlying assumption is the mean, the population distribution is mean then the mean would by itself, are random sample, would be distributed by normal distribution, with certain standard deviation and certain mean, and the variance would be distributed for the samples which you pick up, would be distributed like I square.

So, if you fit the mean distribution of the sample which you are picking out with the standard normal deviate, means find out as the boot size increases, becomes exactly a straight line. Similarly, as you pick up for the variance or the sample variance, along

with the ki square distribution, initially, there is fluctuation, but in the longer run again, it is straight line, which means that it basically address to the property that from the concept of theoretical concept, which we study that the mean of the sample in the long run would basically, be exact equal to normal distribution. Then the ki square is the best fit for the sample variance that actually comes out in this case.

This is considering that you are considering for a big sample size and considering the central limit theorem to be true, but if you consider the extreme value of distribution, it need not be, but still it gives you a lot of information that if you do the boot strapping you can find out the actual characteristics of the population in a big, considering that you have only one chunk of sample and you resample it repeatedly, k number of times, where k can be a very large number.

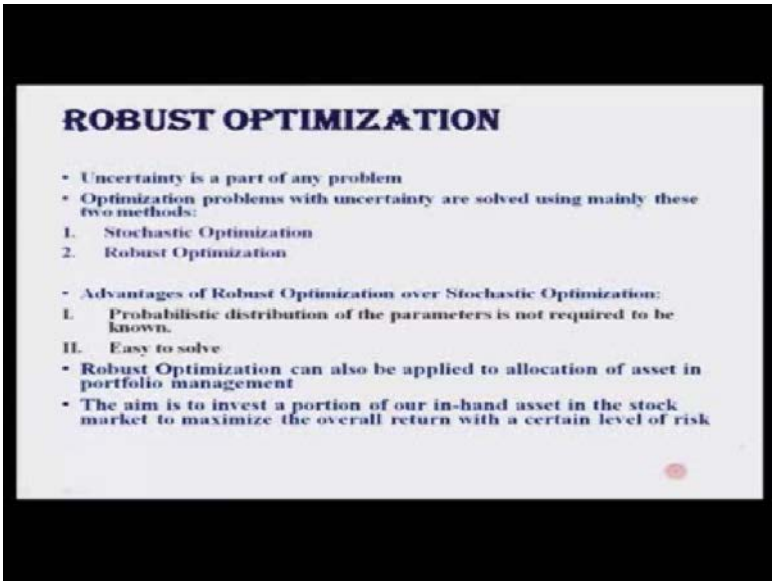
(Refer Slide Time: 06:17)



Assumptions of bootstrapping are bootstrapping makes no assumption regarding the population. No normality of error terms is considered. It can be anything no equal variance is considered; it can change, but still we are able to mix it in such a way that the variability is basically, done away with. Allows for accurate forecast for intermittent demand in between. The sample is a good approximation of the population. The sampling distribution may be estimated by generating a large number of new samples; that is what I am saying.

If you consider the samples which you have picked up at one go, actually, has lot of information of population. Then trying to repeat picking up the sample in the same manner, number of times would actually, give you the whole picture of the population if you combine all the sample information together. Because you are trying to pick up chance. These chances are picked up from the population randomly, and if you are able to pick up the chance in a nice manner then all the characteristics of the population would be coming out in the combined chance, which you found by combining the samples. For small data sets, taking small repetitive sample of the data and repeating it will yield superior results; that is what I am saying. If you do it with replacements for a large number of such re-sampling or a sub-sampling then the characteristics of the sample, considering the sample is the best proxy of the population in a very nice manner, then combining the samples would actually, give you a lot of information about the population and its characteristics.

(Refer Slide Time: 07:42)



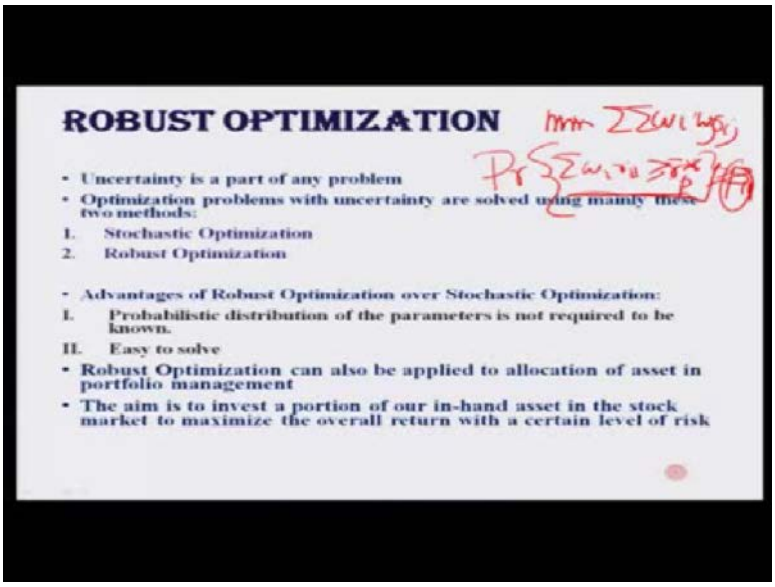
Now, with this concept background, I will just cover in a very simple manner, what you mean by robust optimization. Now, before I scribble anything, let me again go back to the Markowitz model. In Markowitz model, generally, the actual concept was given the utility, the actual characteristics of the portfolio is that you have w 1, w 2, w 3, w 4 so on and so forth; that the weights of so called n. This small line is not exactly; go to the some samples which you have talked about. Consider this n is the number of this furnished stocks which we have. So, these w 1 to w n are the different type of scripts and their

corresponding mean and the variances are given by r bar I and sigma square I, and the covariance is also known to us. So, we have either formulated the Markowitz model by trying to basically, minimize the overall portfolio risk, subject to some constraints or trying to basically, maximize the overall return.

So, if you minimize the risk or maximize the return, you are trying to basically, take two approaches, but the constraints would all be the same. Like if there is the sum of the weight should be 1 then if the weights are between 0 and 1; there is no short selling. If the weights can be negative also; there is short selling that is unboundedness on these characteristics. In hand if you note down whenever, I said I considered apriary that given the fact that observations were known; the closing price, they are fixed; they did not change. So, given the closing price we will consider them to be given and we will consider the return which is l n of p 2 by p 1 as being given and they are deterministic and based on this we proceed.

Now, the problem is these are the past data. So, the past data would not give us exact information; what is going to happen in the future. So, we will consider the past data is the sub-sample from the overall population and use this past data repeatedly, with re-sampling in order to basically, do some good bootstrapping in order to utilize our results. Now, when you are trying to utilize the results; obviously, they would be constraints, probability on the constraints.

(Refer Slide Time: 09:52)



Like I say for example, consider very simply, I want to minimize the risk sigma ij. Consider the returns are greater than equal to some r star p value, and all the constraints are there; forget about that. Now, if you looked at it carefully, what we can do is that rather than mean having the return of the portfolio being greater than some r star p, we can just put the constant of probability that this value would be greater than r star p by some probability beta 1, which means that if you consider the concept of non deterministic case then they may be some instance, where this inequality may hold and this inequality may not hold. So, what we are trying to do is that we are trying to put some probability bounds that if this; let us say for example, 90 percent and we said this inequality holds 90 percent; it means that if you keep repeating this experiment, do the simulation 100 number of times, 90 of the instances; it be true. 10 of the instances, it would not be; that means, we are slowly trying to bring some non deterministic nature in our overall problem; concept of trying to solve the optimization problem.

So, let us consider the concept of robust optimization and how it proceeds in a very simple manner. Now, the advantages of robust optimization is that we do not consider any underlying distribution for that parameter, which you are going to study. That is, it is not specific to any distribution, because if we consider some specificity of the distribution then you have to basically, consider what type of distribution it holds. Does it hold that is distribution? Does it mean that is going to be distribution or say for example, depending on n different type of problem; should be considered as a normal distribution? Should be considered it is a viable distribution? Whatever it is; that would basically, come under the purview if you study the concept of reliability based optimization or stochastic optimization, but in the robust case, we consider the underlying distribution is not taken into consideration. So, in that case your question would be what we consider.

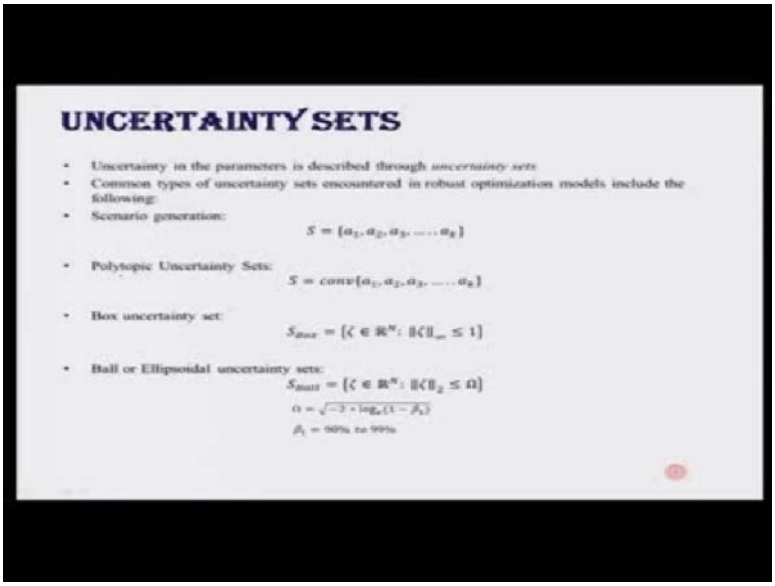
Rather than distribution, we consider a set; that is a set is basically, a small bandwidth. Like say for example, consider an atom is vibrating. So, if it is vibrating is basically, on the mean value and it is vibrating over a number of certain range on the right hand side and the left hand side. So, we will consider some sort of set, where the vibrations of the movement of the atom would give you the overall bandwidth between which, the overall stock price can fluctuate. So, if we have the bandwidth is very small; that means, the level of probability is very high; that means, we are confident that the value of the stock

price would be between a certain bandwidth. Here, the bandwidth is very large. Then the corresponding probability slowly shrinks. So, uncertainty is a way which we are trying to bring for our study, in the optimization sense, in the financial optimization sense, where we will consider the input variables are uncertain.

So, optimization can be done in stochastic optimization sense, robust optimization sense. There is reliability optimization also, we will consider as the far end of the class, because after the robust, I will again switch back to the discussions we are having and if time permits, I will definitely consider the concept of reliability based optimization, which is also very heavily used in civil engineering and mechanical engineering, but I will try to give some examples, where each has been used in financial optimization also. So, advantages of robust optimization over stochastic optimization is probabilistic distribution of the parameters are not required to be known. As I said, it is easier to solve, because once you have the actual mathematical formation, converting that into a programming concept becomes very easy.

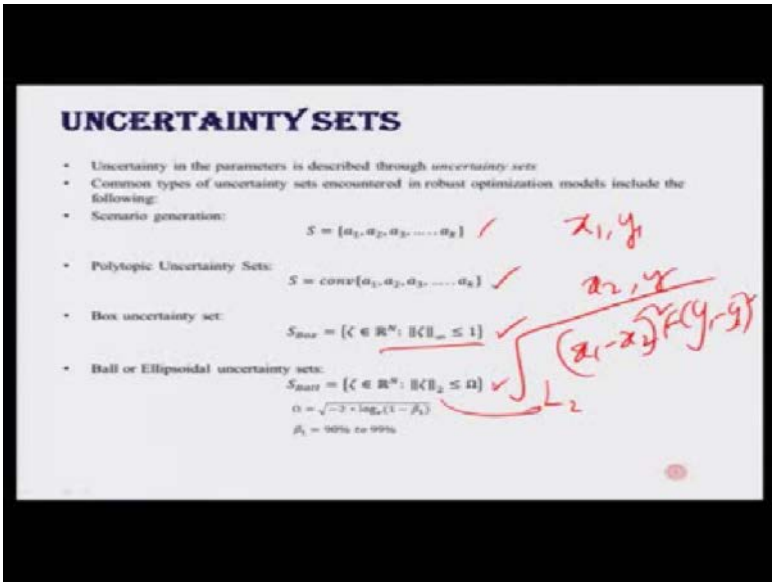
Robust optimization can also be applied to the allocation of assets in the portfolio optimization, which we will consider. The aim is to invest a proportion, out of the total money you have in your hand, in assets, in the stock market to maximize the overall return or minimize the risk or whatever will be your actual concept is, but with the fact, that the input variables are non deterministic and we would not considering this distribution to be true, and we consider the concept of set in a very general sense such that the set, will give in a way, the overall reliability or the overall robustness, which are there for an answer.

(Refer Slide Time: 14:16)



So, uncertainty sets as we know, are if we somebody studies or different time; some of them are the scenario generation; that means, we generate the scenarios. There are the Polytopic uncertainty sets depending on whatever mathematical assumption which you have. Then another is basically, the box probability and the ball probability. So, if you consider the box and the ball, there are, let me step back and there is a concept of distance measure. In distance measure, in simple mathematics is basically, the norms; l 1 norm, l 2 norm, l 3 norm; whatever you have and l 2 norm, we know is basically, the Cartesian coordinate concept which we used when we want to find out the distance between two points.

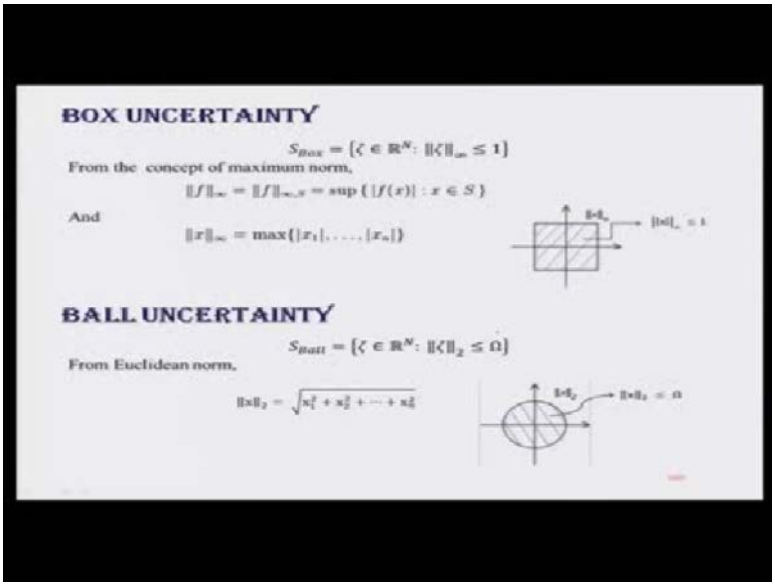
(Refer Slide Time: 15:00)



Consider the points are x 1 and y 1 in the two dimensional case, and x 2 y 2 will basically, consider x 1 minus x 2 whole square plus y 1 minus y 2 whole square, square root of that. So, this is you are taking the l 2 norm. So, this is known as the l 2 norm as is denoted and the l infinity norm is basically, when we consider the maximum of the distance and there are concept of l 1 norm, which is known as the Manhattan distance, but this is just for the information; we will only concentrate on the Cartesian coordinate of l 2 norm and the infinity norm, which is in the Robustic sense, is known as technically, the ball or the ellipsoidal uncertainty, and it is known as the box uncertainty for the infinite number.

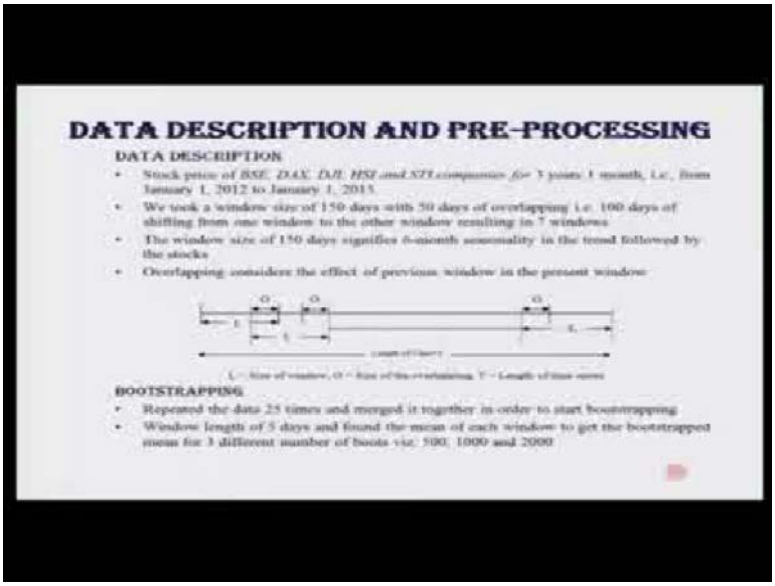
Ellipsoidal is basically, the ellipsoide which you are forming in the higher dimension. If you consider in that simple three dimension, and if all the variances are equal then it is basically, a simple football shaped and if you consider the variances are different is basically, simply like a rugby ball if you see, and as you go into higher dimension, it would basically be ellipsoide, like if you see the two dimension is basically ellipse, rather than a circle and in higher dimension, rather than a football, it is basically, a simple base, this rugby ball, which we consider, which is the American football.

(Refer Slide Time: 16:19)



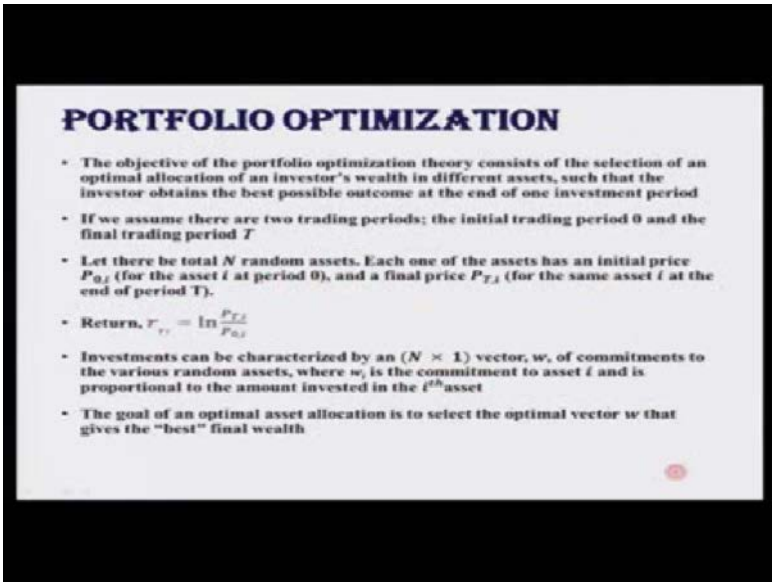
Now, in the box uncertainty, the box is denoted by as we know in a two dimension by this, in a three dimensional cube, and in the ball one if you do in a one two dimension, is a circle, and in higher dimension, is a sphere. So, what we do is that if we want to basically find out the common area between the box and the ball. So, once you are able to find out the common area between a box and a ball, it gives us the maximum reliability, based on which we can do the sliding; this is the very simple concept which we would not go into details in order to explain the mathematical concept. It is basically a union of this, and we find out the intersection such that it gives us the maximum probability of the reliability.

(Refer Slide Time: 16:59)



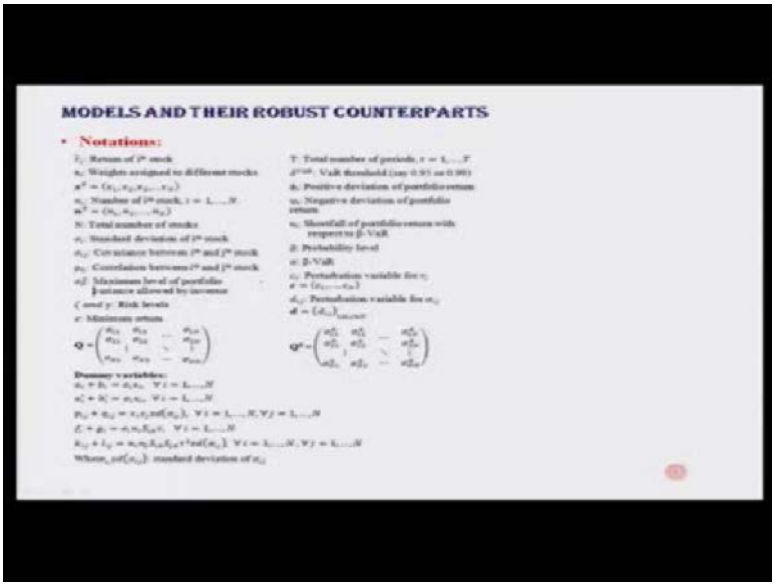
Now, the study which we do, as you remember I mentioned about three years. So, these are the data sets with the data sets we are using are the BSE in the Bombay, DAX in the Germany, Dowjans Injestial average, DJI, (Refer Time: 17:15) Hanseng which is HIS; STI is the Singapore one. Companies have been taking all the indices which are there in this index for three years and one month, starting from January 1st 2012 to January 1st 2015. We took a window size of 150 with 50 days of overlapping. They were overlapping each other, that is 100 days of shifting window. The window size of 150 signifies significant 6 months seasonality, which may be there in the stock market, and the overlapping window is considered like this; it shifts. So, repeat the, in bootstrapping, you repeated the data 25 number of times and merge them together in order to do the bootstrapping concept, if we consider. Window length of 5 days and found window length of 5 days are also being taken and we have done the bootstrapping, considering for 500, 1000, 2000 so on number of n boot.

(Refer Slide Time: 18:10)



The objective of the portfolio optimization theory consists of selection of the assets in such a way that you are able to maximize or minimize some objective function. If you assume there are two trading periods; the initial trading period is 0 and the final trading period is t. In between, it would be 1, 2, 3, 4 till t minus 1, and then the t. Let the total number of n random assets being there and the returns be calculated given by l n f pt by p 0 or pt 1 by pt 2 and so on and so forth. Investments can be characterized by n cross 1 vector, which is w 1, w 2, till w n, which are the weights, and the goal of the optimal allocation is to select the optimal vector w that gives the best final worth, whatever the concept of best is.

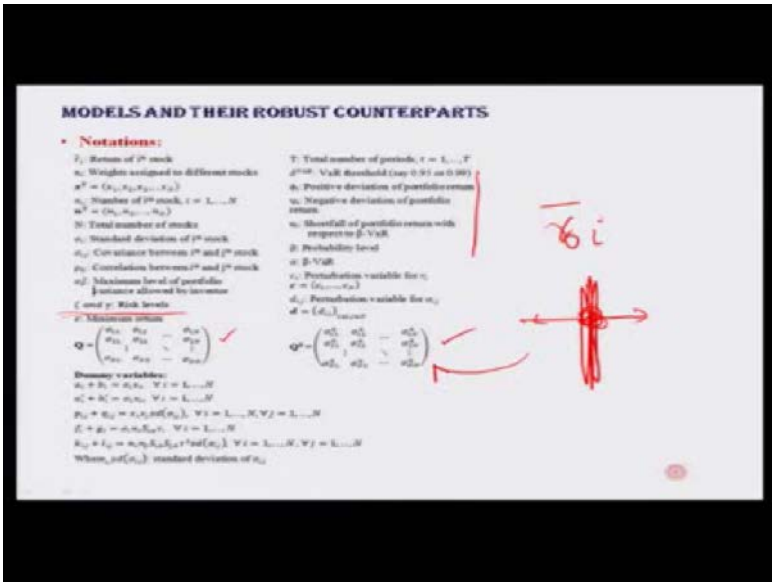
(Refer Slide Time: 18:54)



So, let us consider these notations without going to technicalities; r star is the return on the I stock; x I is the weight assigned to different stocks which is the w's; x t is the vector, x 1 to x n, depending on which time period you are considering; that time window we will consider, the number of stocks and I is the number of stocks; n t is the number of stock given on the vector; capital n, the total number of stocks; sigma I, row sigma I j, row I j are the corresponding standard deviation, covariance and the correlation coefficient. Then these are the level of risk, considering different concept of reliable robustness we consider; q is basically, the variance covariance matrix. Then you have the t as the time period.

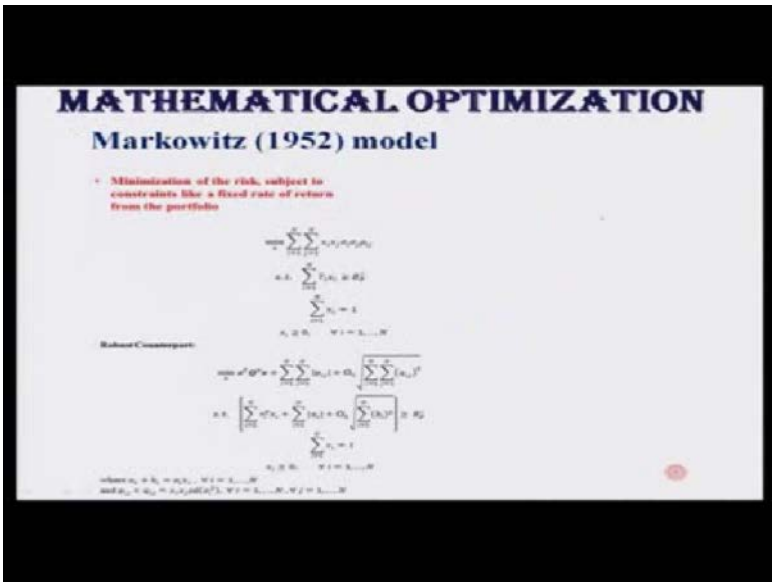
This var threshold value; what is var? We will come to that later on. Later on means another 3 or 4 lectures will be there into var and cvar and so on and so forth. Row, and this positive deviation of the portfolio, the negative deviation of the portfolio, the u th, the shortfall of the portfolio; all things would be specified depending on what you think of the level of your reliability is, and this q star or q 0 is basically, the initial deviation over on and over which, the reliability would be; that means, if you remember I mentioned the center and over which, the fluctuation is.

(Refer Slide Time: 20:15)



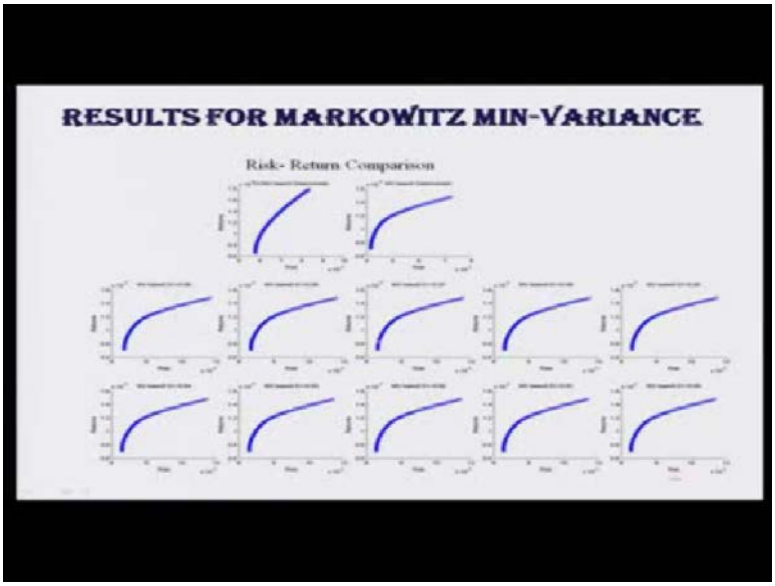
So, this is the mean value which you are considering, which is denoted, wait, is one of them is here and another one would basically, be the r bar 0 I for the stock which we are considering.

(Refer Slide Time: 20:34)



So, the first model is we consider the model, where we minimize the risk, corresponding to the return being greater than equal to r p star, and sigmas and sum of the weights greater is equal to 1, and such that stocks are is equal to 1 to n.

(Refer Slide Time: 21:00)



So, if you do the risk return comparison and do the runs for different levels of reliability, starting for the deterministic one, for the 90 percent reliability; that means, 90, 91, 92, 93, 94 and so on and so forth of the beta value which we consider for BSE, as shown accordingly. We will end the class here and if you consider later on, we will discuss the models in detail, as we discussed the runs in much more details. Later on, once we basically, finish off the discussions about different concepts which you have to cover.

Thank you.