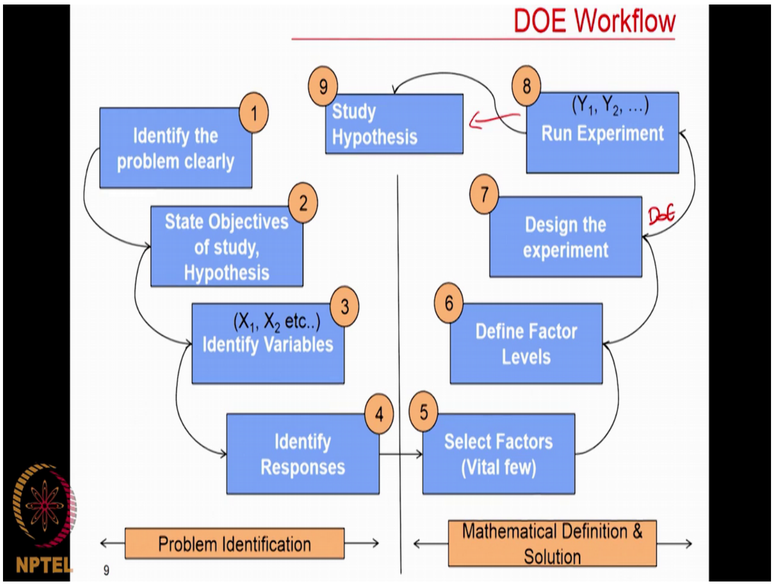
Design for Quality, Manufacturing And Assembly
Prof. Palaniappan Ramu
Department of Engineering Design
Indian Institute of Technology, Madras

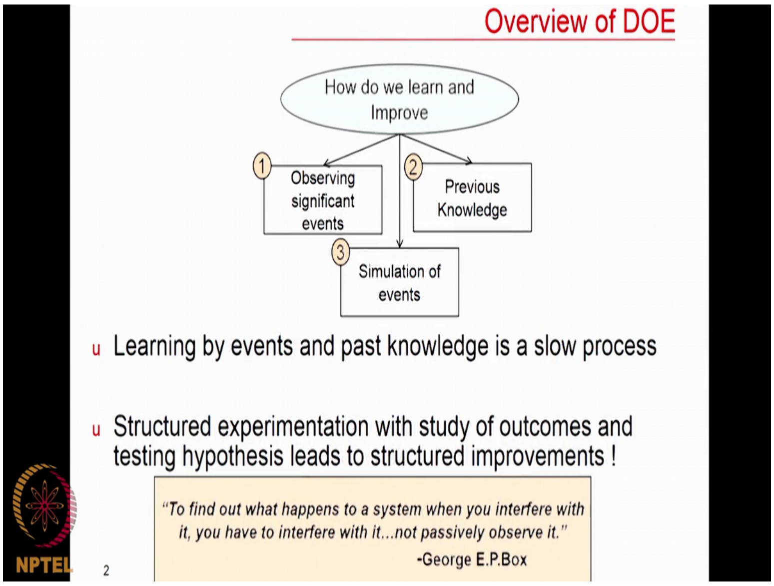
Lecture - 39
Overview of DOE Workflow

(Refer Slide Time: 00:16)



So, this is just to give you an overall idea of how to implement design of experiment and what is the basic idea of the design of experiments is also this here, ok.

(Refer Slide Time: 00:25)

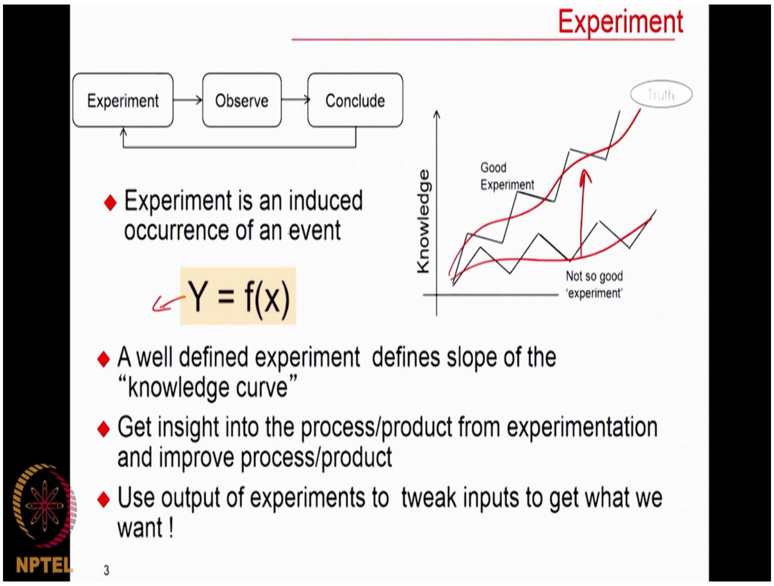


This is a general overview of a DOE. You want to learn and improve that is a basic idea in any design process. There are multiple ways in which you can do, one is by observing significant events, one is you have previous knowledge you have legacy and the third one is simulation of events, ok. In order to do 1 and 2 you have to interfere into the process ok, whereas, 3 is good because it is a computer simulation you can simulate the event, but the moment you want to simulate the event you will have to replicate these guys that becomes important. You understand what I am saying?

Let us say that you are trying to model rainfall and you have last 50 years information then you build the model. Your model will never get the 2015 rainfall situation, it will never get it because your model is based on the last 50 years in which we never saw that kind of a rain or tornado or whatever your wind speeds if you do your model will never capture the Vardah. So, what you have to do is you will have to try to simulate this significant event.

You will see it will happen in one in 100 years, but your model should have that kind of a non-linearity, you have to do that. If you can do that then you can do simulation of events, ok. What is the idea?

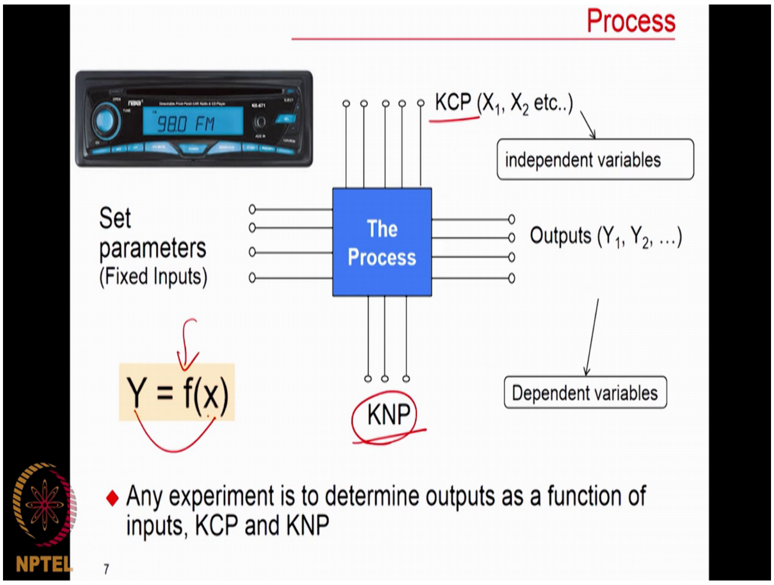
(Refer Slide Time: 01:45)



This is similar to finite element, there is a truth is here, ok. The truth part is here it is kind of a non-linear stuff I could still do a non-linear approximation or I could do a linear approximation, does not matter. The truth is not known. This is the actual stuff, but it is not known. What happens to a occupant under a specific crash condition you will never know what the truth is you cannot do an experiment with the real person inside. You will do it with the dummy and you hope that the dummy is correlated to the human being that is all, ok.

So, you do not have this information, but we hope this is closer to that and then we try to do this. This is not such a good experiment you will have to correlate the few points that you have and then you will have to adjust this guy to this guy. So, end of the day this is what we spoke about; the two things that experiment that DOE can let you do is this can give you some statistics of this and it can let you construct this f function.

(Refer Slide Time: 02:49)



The general process I did not discuss this explicitly. For instance you have something called key control parameters. You have key noise parameters. You cannot do anything about this noise; you know that these are noise you have to live with a noise. The roads are bad in India. You are a car manufacturer, the roads are not let us not use the word bad, it is not that good in India.

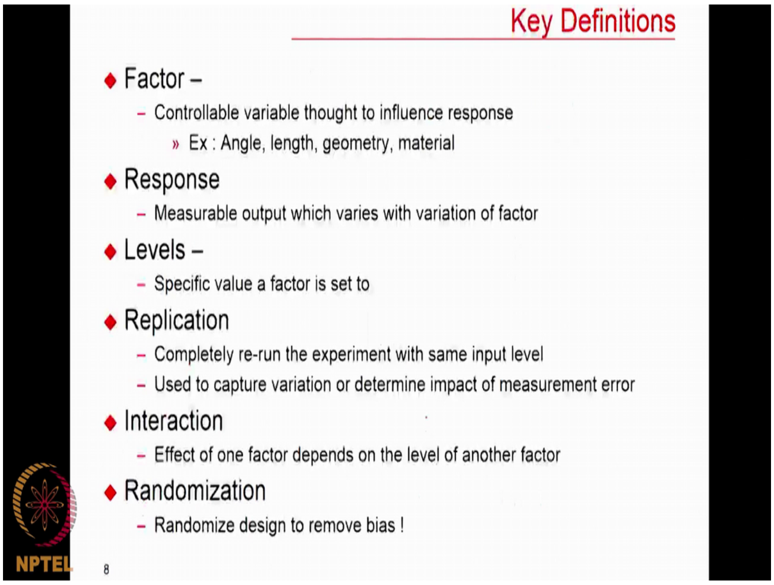
You say ok, let us wait until the roads get better then we launch our cars. Today you name the car that is there in the Indian roads you name a car, that car is there in the Indian roads anything from your Lamborghini into your Rolls Royce, goes to anything that you name it is there Ferrari. Everything is there in Indian roads need not be necessarily in Madras, it is there somewhere, ok.

So, these are all sometimes they are all like fast moving cars, sports types of cars anything, they are not designed for Indian roads ok. So, they say no let us not launch this because Indian roads are not good. And, someone buys it at few crores and let us say that it fails on day 4, because the roads were not bad were not good, that is also not acceptable right. So, this is noise parameter which is not under my control. I know these are the noise parameters, but I do not have it under control.

The way you guys attend are attentive to me is a noise parameter. All of you are not the same the way you sit, the way you nod, the way you put your head down, the way you do this, it is all noise parameters to me, but I cannot let me I will not say that let all of them be fresh and then I will start, that can never happen. So, what I do is I try to my minimise my reaction to what you are doing, I make myself insensitive to your noise that is what is robust design.

So, that three control parameters I have key noise key noise parameters I have and this X is your input this X is your input which is your independent variables and the Y is the dependent variable on your X. So, essentially you are interested in this f finding this f that relates your Y and X, that is all. And, our whole point is also to simulate these guys to understand the effect of these guys and to simulate these guys.

(Refer Slide Time: 05:13)

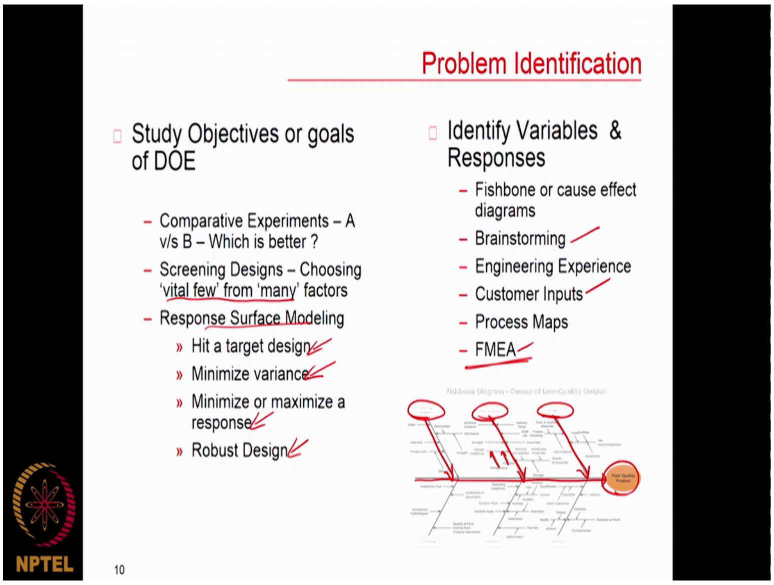


What is the DOE workflow? It starts in a design perspective, ok. So, first you will have to identify the problem clearly. You will have to study the objectives of your you will have to lock your objectives. Sometimes as you point out in the clean example, in the tile example the problem was the tile output, the variation in the tile dimensions. But, you should know what is it that you are looking at to address. You cannot go and address the non uniform temperature you will have a problem, you should know that I will identify the combination of my inputs that go in, ok. So, that becomes important you will have to state your objectives, then you will have to correspondingly identify your variables that will contribute to that objective which essentially will address the problem.

You identify the responses it is not always one it could be multiple responses, from this you select a few. You do not you do not carry all of them you select a few then within that you define your factor levels 3, 4, 2 whatever it is, Then you do your design of experiment which could be orthogonal array for us, run those experiments in some randomisation blocked complete whatever simple and then you study your hypothesis.

So, what we have discussed so far is this DOE we have done. Running the experiment we imagine that you will be able to do some kind at some combination you will be able to do, right. Then you will have to study the hypothesis. What does this mean? It could be your quality based study it could be your robustness based study ok. These are all out of scope practically for us how to choose the vital few from the trivial many and all that. Usually you use you do something called PCA, Principal Component Analysis. There is something called factor analysis you can do that and sensitivity study you can do to identify which factor or the response is sensitive to which factor the most then you select only those guys, ok.

(Refer Slide Time: 07:27)



What are the objectives or goals of DOE one please understand this is slightly different from what I told you earlier. What DOE let us you do is different from what could be the goals of DOE. So, this is an interesting stuff. So, if you look at it can let you do comparative experiments A versus B which is better? Is A better than B without getting into the physics.

There are two ways of looking at it right you build a model you understand the physics and then you tell whether A is better than B, but often times you cannot do that. The furnace example you cannot touch the example I do not know the physics I only have a product with me. So, what I do I get two – three products and then I test it and then I will say this product is better than that product. So, DOE let us you do that.

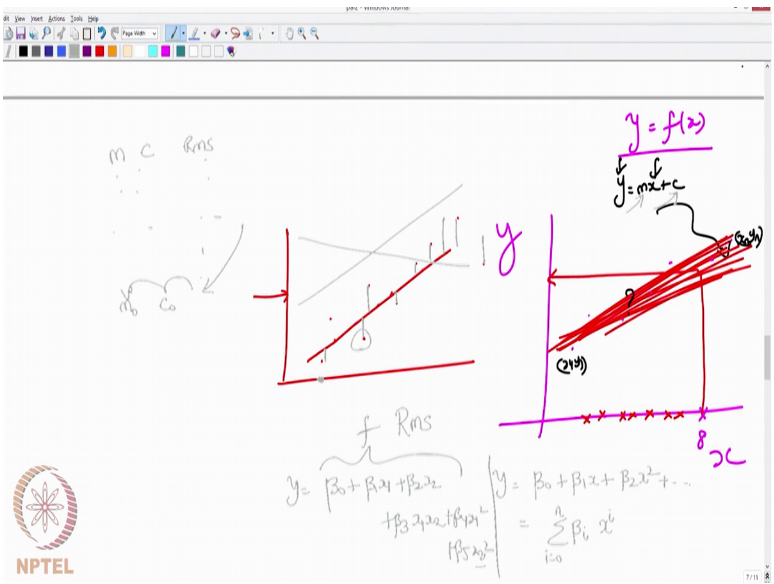
For instance you get two different bikes motorbikes and then you are reviewer in xBhp forum. You just drive it in one setup and then you say this bike is better than that bike. You cannot post a review like that; you will have to test it in different conditions, ok. You will have to ride it in a muddy road, you will have to take it up a hill, and then you will have to check on different parameters how fast it went from 0 to 100 kilometres per hour, ok, what is the fuel efficiency, on different attributes, on different levels you will have to try, ok.

You will have to try the mileage driving it early in the morning, at 1 o clock in the afternoon, in the night, and then you subject it for continues riding for like three days and then you check what was the mileage you continuously ride this guy as much as full throttle you abuse this vehicle for like three days and then you check what was the mileage. Similarly, you abuse a vehicle for three days and then you do it then only you can say that this vehicle in an overall sense was better than this vehicle or this vehicle was specifically good in these areas and this vehicle was specifically good in this areas.

So, you cannot just do one condition and say all the different attributes that a combination that I told you is what your design of experiments. So, you will have to design your experiment in such a way that it will get abused I will get in an nice combination it will go in a hill, it will go in a flood terrain, it will go into a valley whatever it is. So, all this are you are trying to simulate the different combinations. So, then I will be able to tell you a comparative study of A versus B, that is one thing.

Then, you can also conduct an DOE to identify the vital few from the many factors, ok. So, that will tell you which factor is important column wise when it says it says this factor is more is more effects more the response then you will choose only those factors, but this will be a preliminary design of experiment. Then, you can do something called the response surface modelling it is nothing, but getting the f using some regression models you can identify f.

(Refer Slide Time: 10:22)



See, if I give you a bunch of x and y, we have been talking about this for a while I will give you an idea on what this is. I did some experiment x is a dimension, y is a response for different dimensions I have data like this. I have this data I gave this data to you. Then, I say, hey I forgot can you go and find out for a dimension of 8 unit what is the response value? You have this data with you and then I am I told you can you go and find out for a data value new data value of 8, what is the response value, what will you do?

One way would be to go and conduct an experiment at this x and find the y, but is that the wise way to do it? No, the wise way is just now what this guy told he says I will look at the data and it kind of exhibits a trend. So, can I do this? If I do this then probably I can just say that this is my y and I will be able to the moment I construct this I will be able to tell this for any x that you ask me any number of x's which means I can get rid of the experiment that you are trying to do.

This is what is the f of x that we are talking about. What is our f of x in this case if it is a the red line is represented with what equation?

Student: y equal to f of x.

So, this equation is you will know y, you will know x that is what I put as dots here, correct? This is just this is x 1 y 1, this is x n y n, this was x 4 y 4. So, there is an x there is a y. So, like that I had 7 x's and y's and that is what I have plotted. The second point is how did you get this graph? You go to excel, you put this and then you set trendline, then it will give you this. But, practically how do you get it?

Student: These try and take a line such that the distance (Refer Time: 13:00) minimise the distance.

One second. The question is what he told is right the question is I can fit infinite number of lines through these through these points which one will you select? And, please understand none of these lines could pass through all the points that is not possible, when there is variability in your data. If I drew the lines like this then you can say I will draw line that passes through all the points, but I did not draw like that I drew like this. So, what will be your criteria to draw this line?

Student: Taking the mean.

Mean? Mean is here, this is the mean.

Student: Mean is standard deviation.

If this is the y, this is the mean for you.

Student: With all the data points we can calculate the mean and plot the standard deviation.

No boss, this is a mean.

Student: Yes, sir.

This is a mean, what do you do after that?

Student: Sir, standard deviation we could reduce and try to.

Standard deviation of what?

Student: All the points.

Ok, you compute the standard deviation of all the points. Then what you do?

Student: Sir, then we can reduce the which line has the minimum standard of deviation.

Standard of deviation what? You are coming to the point, but you are not at least you are not expressing if I am not standard deviation of what, ok. So, here is a point you draw this curve. So, this particular point this particular point the actual value is this red dot, but what the line gives you is this. So, what is there? There is a error similarly there is an error. So, this guy there is an error. So, this guy there is an error. So, this guy there is an error actually the line should come here.

So, there is an error for this guy. All the green vertical lines are the errors between the actual dots and the line that you have fitted. Whichever line you take from this will have this error because none of the lines is going to pass through all the points. So, now, what could be a logical criteria for you logical criteria?

Student: Least.

Sorry?

Student: Least error.

You want to choose a least error, but then you have more errors here you do not have one error. So, what you do?

Student: Average.

Average is one thing, but you.

Student: RMS.

You do an RMS. So, you should be careful what statistics to use when, right. So, you will use an RMS because you want to get one metric out of multiple values. So, RMS is what you will square your errors, you will take a mean of it, you will take a square root of it, that is your root mean square. Why you will square is because you will have errors on both direction. So, they should not go to be 0 that is all ok. So, what you will do is you will change your m and c; m is the slope and c is the intercept. So, they will control if I am going to change the c for the same slope, what will happen? This curve will become this curve, ok. For the same c if I am going to change the slope it will become like this.

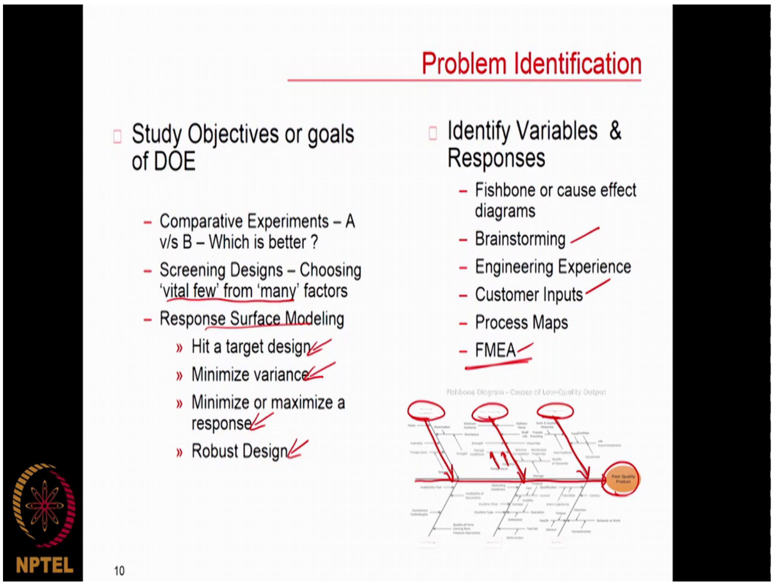
So, this m and c is your design variable. So, you are going to change them while you record your errors. So, what is it practically there is an m there is a c and there is an RMS error. So, for different m and c I will have different RMS errors. Imagine, you are you are not going to do that you will use an optimisation algorithm to do that. If you do all this what you will do is you will find go and find something where the RMS error was less and then you will choose the corresponding m and c's m optimum and c optimum. This is how this is called regression because you are trying to regress the values or the residuals, you are trying to find m and c and this can be generalized, ok.

You can go and generalize this to y equals beta naught plus beta 1 x plus beta 2 x squared you can just write it simply like this beta i x i sorry not x i sorry x raise to i, ok. So, if you say 0, i running from 0 to n. So, beta 0 x 0 is 1, then beta 1 x raise to 1 beta 2 x raise to 2 and you can also build it into an i j provided you have x 1 x 2. So, if you have x 1 x 2 you will write it as beta naught plus beta 1 x 1 plus beta 2 x 2 plus beta 3 x 1 x 2 plus beta 4 x 1 square plus beta 5 x 2 square. So, you can keep on doing this it is all fun ok, but this is not the scope of this course.

So, the idea is you can fit a response surface, this is called a line fitting, but the moment you have 2-dimensions it is not a line anymore it is going to be a surface, ok. So, this is important you need to have a function that relates your because this now, this is a function that we are talking about, this is a function f that we are talking about. This is how you get y x and then you find the relationship,.

So, this is the response surface that we are talking about. So, using that you can find a target design, you will be able to minimise the variance, you can find the minima or maxima of the response, you can do robust design, you can do reliable design, you can do optimisation, you can do design exploration there are multiple things that you can do with that provided you create a surface like this,.

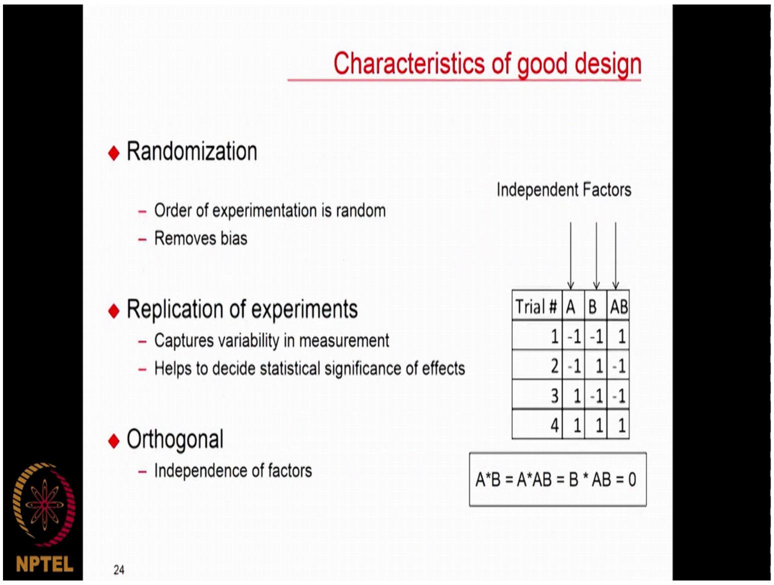
(Refer Slide Time: 19:00)



The other things is how do I identify these variables and responses. These are stuff that you might have come across in a what you call functional conceptual design, the fishbone diagram, what leads to what. Otherwise it is not a big deal, see what you do is this is called working condition. You are not able to read this, but do not worry I will just tell you what these things are. So, this is a main one and then these are the bones that lead for instance this one is called working condition, raw materials management. I will tell you what the flow of the working condition and says noise illumination humidity temperature these are the factors that, and then raw material it says moisture contents delivery time, shell life strength, storage condition all these effects, ok.

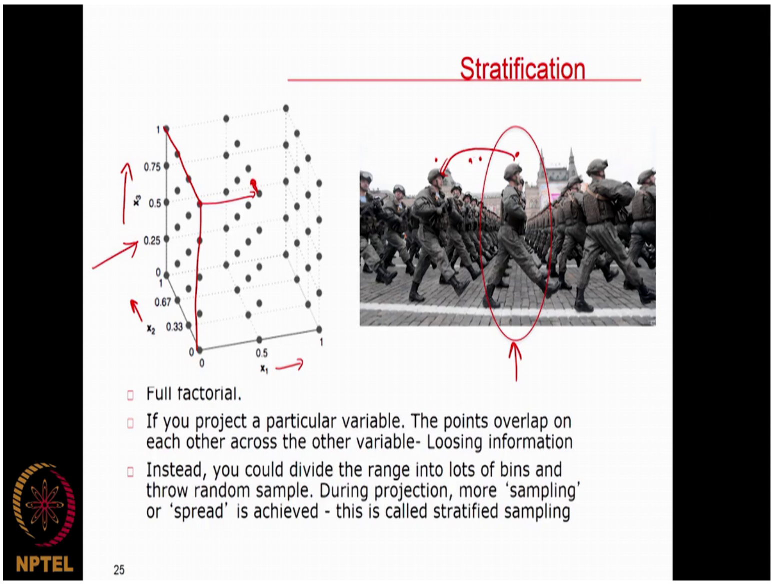
And, then you can also have multiple things from here feeding to each of those bones. Storage condition, temperature humidity influences that, ok. So, you build all these things and then all of this leads to a poor quality product. So, from this you identify which is the most important one, they are all close to the right side, ok. So, that is and then you can not a brainstorming, you can have customer input process maps, you can do a failure mode effect analysis this is one of the key things that people try to do, to identify the fuses.

(Refer Slide Time: 20:22)



So, that was just give you an idea of the general design of experiments, ok. Now, this is just one more stuff I wanted to say.

(Refer Slide Time: 20:31)



There is some limitations with respect to the orthogonal array, ok, one more from the sampling perspective itself, ok. What is it? You see this picture here, this is an orthogonal array representation. So, x 1 is in this axis, x 2 is in this axis and x 3 is in this axis. Now, how many points along x 3? 1 2 3 4 5, 1 2 3 4. So, if you take this particular plane the plane that I have drawn it is x 2 and x 3 imagine that you are standing on this side and you are you are you are just looking at this direction,. Your depth information is lost; if you are projecting x 2 and 3 your depth information is lost, your only getting the 20 point information, but how many simulations have I done? I have done 60 simulation because 20 times 20, 20 not 20 times 20, 20, 20 three times 20, 60 simulations you are doing. At 60 different points you have done the simulation.

But, out of that if I take a projection of x 2 and x 3 I am getting only 20. The idea there the picture in the right is to drive that idea if you are seeing through the this guy here this is like a orthogonal array, condition ok. So, what is happening now imagine that each of these guys are points, ok. So, similarly there is a point I keep one point, but actually there is 20 points behind that point correct, right. So, when I am seeing from here how many people I am able to see? I am able to see only one guy. So, in a projection I am able to see only one of that guy of course, if I go to the other direction I can see, but please remember I will be able to see those 20 people, but I will not be able to see the guy behind him. I will not able to see this guy I will only see this guy. So, still I am losing that information from that projection.

So, what is happening is, though you run the simulation at the 60 samples the information that you extract in a projection sense is only 20, this is one problem with orthogonal array. How can you address this problem? Very simple, if I want to see more people in this row, in this particular row, in this particular row if I want to see more people what can I do? It is simple man for a very long time the guy at the back was hidden behind this guy knowingly or unknowingly, ok. I am not able to see he is sitting.

Now, imagine that 20 people are sitting in this and I am not in a podium I am I am facing you like this I will not able to I will not be able to see the other people, but I need to see the audience reaction to teach. So, if I ask if I want to do that what should I ask you to do?

Student: (Refer Time: 23:52) move.

So, you move, but then you are going to block the other guy next guy. So, what will you do?

Student: Stand up.

Stand up?

Student: (Refer Time: 24:02).

So, let us say that I put a roof on top you; you cannot really stand up that is what you say design bound you cannot go beyond your bound right. So, there is a practically you cannot stand up. Meaning we are we are talking only about this window projection. So, you cannot get up that is a window projection that I am talking about, what can you do?

Student: (Refer Time: 24:22).

Sorry?

Student: (Refer Time: 24:23).

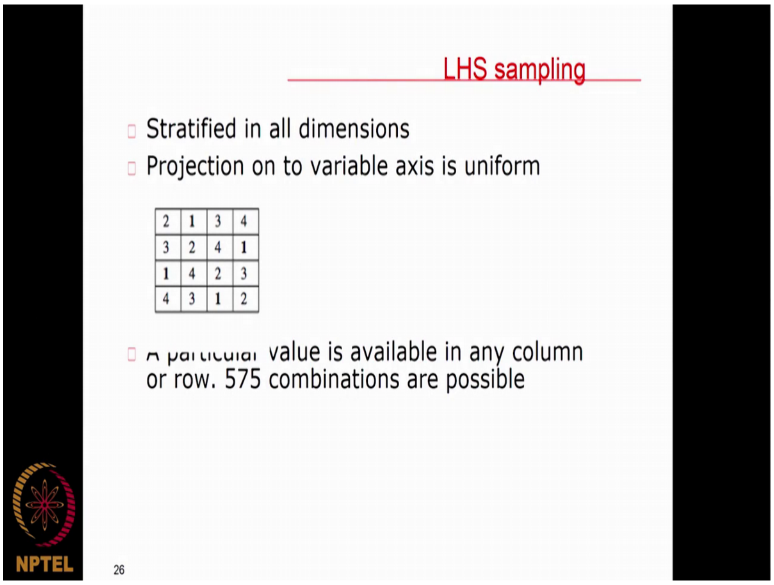
No, that is height wise think the same thing in a projection sense, correct. That is the reason why CLT and all that you have that kind of a stuff, sorry.

Student: Move and see.

So, you will have to move and see, ok. When you go to your what you call theatres, ok. So, there are two types of mix; one is along this axis also there is a mix you are not you do not sit in the same line, ok, this line is also like this and you do not sit one behind the other, bit this elevation also the one seat is between the other two seats there is a small overlap, right. So, now, if that guy has to see me also he has to just push his head a little bit. So, if I want to extract more information what I need to do here I need to push this point a little bit, because right now this guy is just behind him. So, if I want a little bit of information I will just push this guy a little bit.

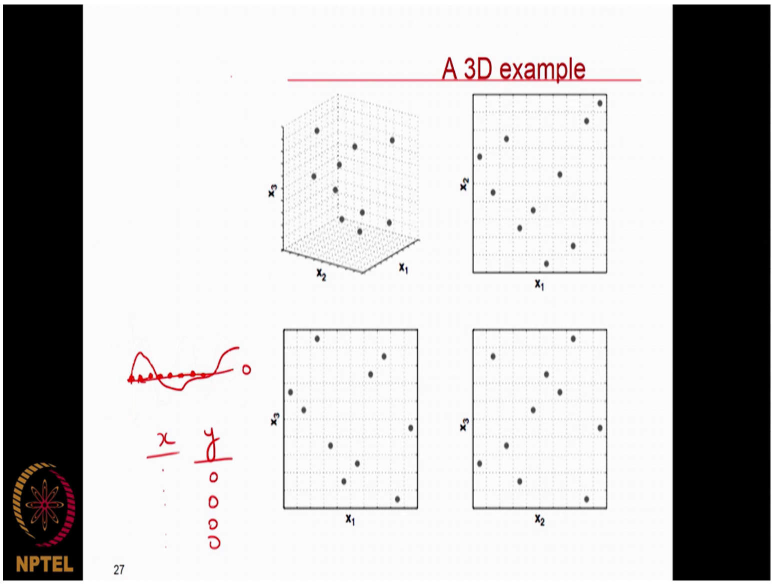
Similarly, imagine that you perturb the 60 points or the 40 points a little bit then if you take a projection you might not get 60 information you might get at least 55 to 58 information, correct. So, this movement you need not do it randomly, ok. There is statistics to do that for you and there is some quantities that you might want to preserve and do it, those are called space filling techniques and one of the widely used technique is called a Latin hypercube sampling,.

(Refer Slide Time: 26:12)



So, the idea in that is it is stratified in all the dimensions, not in a particular dimension that you talk about. And, the projection onto the variable axis is also uniform.

(Refer Slide Time: 26:18)

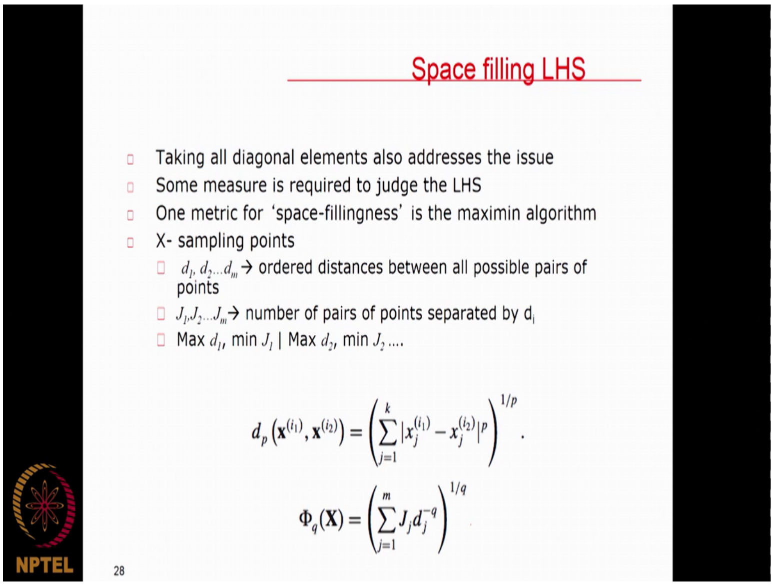


So, a very simple example of a Latin hypercube, this is only to give you the scope, we will not get into the details and all that right. So, there is a min-max algorithm which means it is minimizes the maximum distance between any two points, so that you get total exploration of the design space, ok. The basic idea between LHS in a very layman term is let us say that you have about 10 points that you can do. What I do is this I divide the design space into a 10 by 10 grid, then what I do in a random sense I allot one point to a particular column in this case it was column 1, then I immediately lock that column and that row I will not put any other point in that row and in that column. Then I put the second column in this row in a random sense then immediately I lock that row and that column.

So, when I keep doing like this is exploring because there is a 10 by 10 stuff and for sure you will get one point one row in one column, ok. So, this is all you can do with the 10 sample stuff. So, instead if you used 100 samples in this case then imagine that it will be like very small in this case right, we are almost close to. So, if you have a equidistance stuff and what we are saying is I am going to take points here, right. So, if I had an x and y for whatever x's I did right now, my y values will be 0 provided this was 0, but that is not my curve. If I going to fit this guy and this guy what will you get? You will fit like this that is not my curve. So, instead if I am just pushing this guy a little bit this side here, here, here, here and here something like that then again I might not get this curve, but I will at least get something like this, ok.

So, there is one limitation with the periodic sampling orthogonal array is kind of a periodic sampling, ok. But, there are a lot of advantages of orthogonal array that is evident from the profits and the huge success in the manufacturing industry, it is based on a very strong scientific principle and manufacturing industry has heavily benefited out of Taguchi and 6 sigma principles.

(Refer Slide Time: 28:40)



So, this is the idea behind the space filling techniques, ok. So, let us not it is, it is the max-min distance, maximize the minimum data. Did I write it as min-max? It is a max-min algorithm sorry maximize the minimum distance between any two points.