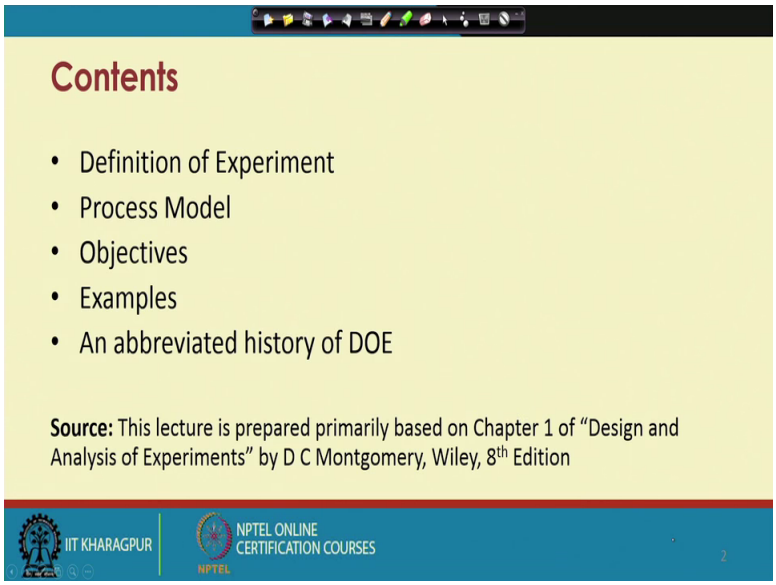
Design and Analysis of Experiments
Prof. Jhareswar Maiti
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

Lecture - 01
Introduction

Welcome to the first lectures of the MOOC course on Design and Analysis of Experiments.

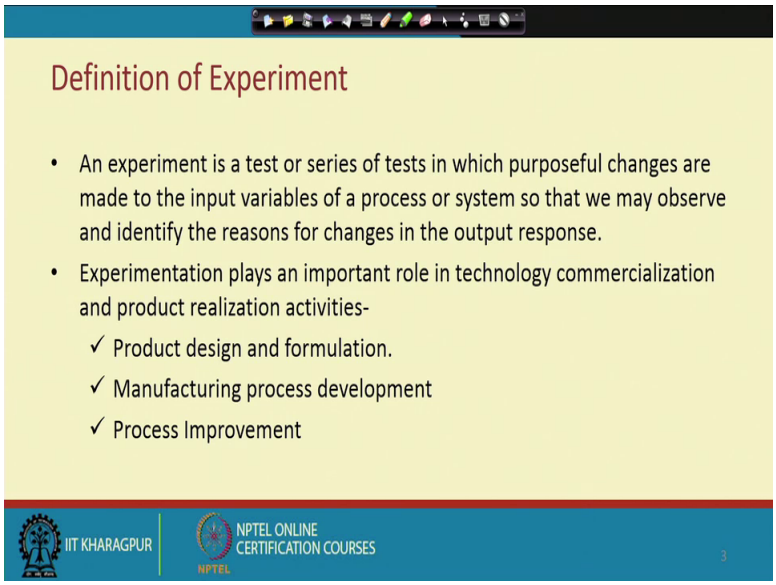
(Refer Slide Time: 00:30)



In first half an hour of time I will try to cover the following topics. First we will define experiment with certain examples, then experiments in terms of process models where we will try to link with laboratory as well as industry based experiments, we try we will find out the key terminologies related to design and analysis of experiments, what are the different objectives of conducting experiments and analyzing the data pertaining to the experiments. And in few cases case examples where doe can be used and how doe will help in achieving the objectives and then the brief history of the doe I hope that within half an hour we will be able to cover this.

The lecture material is prepared primarily based on the book entitle design and analysis of experiments by dc Montgomery which is published by Wiley and I follow the contents of chapter one and it is of the 8th addition of this book and in fact, my subsequent lectures also will be heavily drawn upon this book writ10 by Montgomery.

(Refer Slide Time: 02:41)



So, let us first know that what is experiment before giving the formal definition, let me give one example suppose you are tossing a coin then you will find out that after the after first trial you will either get head or tail, if you cannot repeat this second time again you may get head or tail and you can continue in this manner.

This is an experiment, because there are certain outcomes in this example head or tail will be the outcome and there is definite purpose of that toss also maybe you are in case of cricket you shall have seen that the toss is the first thing to happen and the part the captain who win the toss, he will he will basically choose either head or tail and after the empire or the referee a he tosses these.

Either, one of the captain will win and the purpose is that captain will be able to choose whether go for batting or fielding first. So, this thing if we think a if we are analogously we will bring it to a production system, then you will find out that there are operator and that a system is given to him or her and the production system produces something and there are certain quality of the output to be produced. And accordingly there is a set of process variables that to be controlled, and the operator what you will do he will he will fix the process variable to the desired or the given range as per the specifications and then ultimately process is wrong.

So, now you think that you have the operator has the capability as well as the he has the authority to do some kind of manipulation in the input as well as the process variables

the output. So, that the output can be changed and purposefully changed in order to see the effect of those process variables in the in the output quality then it is also an experiment. Similarly in the laboratory suppose in chemical laboratory or chemistry laboratory you will find out that suppose you are mixing to a to cover to maybe your acid and base to produce something some salt and then by changing the quantity of the material acid or base material you will find out different kind of quality of the output.

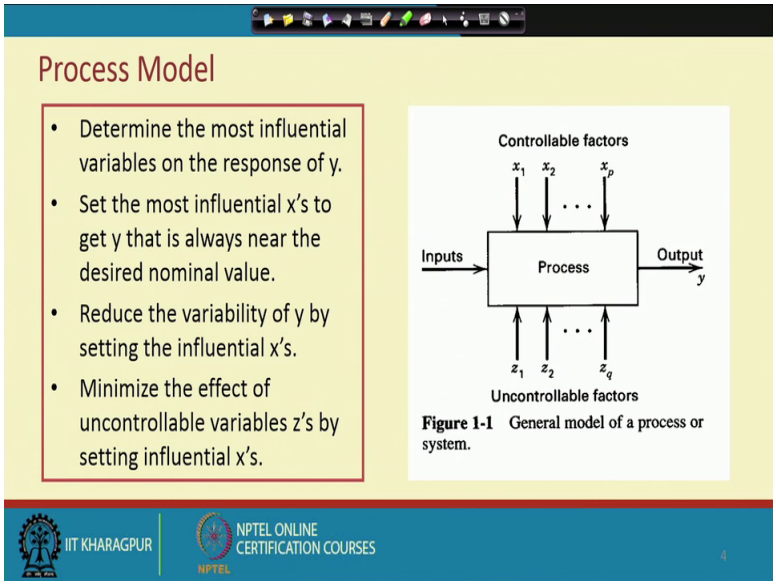
So, even you will find out in the sociology side also suppose someone is interested to object the behaviour of a certain system for example, maybe the behaviour of the traffic or maybe behaviour of some animals. So, they are also some stimulus or stimuli can be generate can be generated in such a manner that the subject here it may be the animal will perform based on the stimuli given the performance of the behaviour will be may be different for different stimuli. So, purposefully if you control the stimulus side what will happen the output also be controlled now these are also experiments. So, give in these examples. So, I and to read out the formal definition of experiment and experiment is a test or series of tests in which purpose full changes are made to the input variables of a process or system.

So, that we may observe and identify the reasons for changes in the output response. So, experimentation plays an important role in technology commercialization and product realization for example, the laboratory based experiments or prototype many typing based experiments or large scale field based experiments also these things are done in order to have good product in order to produce good product. In order to design product in such a manner that the defect during production will be minimized, similarly in case of manufacturing process development, because when if you produce something in item or component or the product as a whole it will be produced through a manufacturing process and that process design itself it also very important one because the process variables will ultimately affect the product quality.

So, how to develop a manufacturing process so that the product that will be produced will confirm to the specification, now the specification which are derived based on the customer requirements maybe through a broad customer survey or market research. Sometimes what happened you are producing something with the given process you want to improve the process maybe in terms of extinction or in terms of existing process performance improvement? So, in that case also the experiments can be conducted and

the resultant data can be analyzed in order to find out the avenues for process improvement.

(Refer Slide Time: 10:11)



So, as I told the process is a very important one now for this particular subject doe design and analysis of experiments I leaves a they so they abbreviation doe. So, in doe the definition of process is not only is not the manufacturing process alone here we will define process a more generic way more general scenes.

Here a process is something where some transfers transformation will take place on to the inputs and as a result some value added outputs will be produced. For example, I told you that in a manufacture suppose you are manufacturing some maybe steal you are manufacturing or you are manufacturing some engine wheel rim wheel rim maybe. So that means, what you require you fed a raw materials with certain other components to the process where some chemical and physicochemical a changes takes place and then out ultimately the either the wheel rim or the your molten metal in case of still make it still making this will be the output, this is one kind of manufacturing example.

Another kind of example maybe that patients going to hospital and they will be there will be a will go through a process some steps like you can say that starting from the appointment to consulting the doctor to test. Then finally, give me medication and then ultimately the patient will be cured and he or she may be admitted to hospital.

So, they are in the medication process also there are certain steps. In other wise suppose you are writing a program which will give some output suppose a prediction model you are developing. So, then the in the program the algorithm your building the there are certain steps which will be executed then only the prediction model will give you the desired output. So that means, a process means a series of steps either these are a chemical, mechanical, electrical or where your service related like health care or it is algorithm whatever may be these there are certain steps, which are taking place and these steps are taken on the inputs and some outputs are outputs are produced.

So, the process itself has certain characteristics these characteristics are named as controllable factors and uncontrollable factors. For example, if you are producing suppose who will ream and though a centrifugal casting process then the speed of the process the feed; feed rate although and the molten metal flow all those things will be the controllable factors and there will be certain factors, which may not be you may not be able to control maybe the ambient temperature during cooling after casting you will you go for cooling that may not be controllable.

So, similarly the operator who are coming from different background, and having different qualification, different training level, they may not be controlled. So, in that sense whenever you do any kind of production or any kind of service operation there are there are there are many thing many parties in board one is the process itself who is which is having 2 kinds of factors controllable uncontrollable as I given examples. So, the sole purpose of this process is to produce something or to either you perform some operation or you give some service or you produce something and the that whatever what perform a what you are performing and what you are service your are giving or whatever producing all those things those things will be tested with certain characteristics this is known as response of the process or in quality terminology known as quality variables.

The product will be of good quality, the service to be a good quality, the performance will be of the satisfactory in nature. So, now, doe is applicable or you do experiment here in in how now you manipulate the controllable factors and also sometimes inputs in such a manner that the output quality or the response here that will be at the degradable or the variability of the output will be at the at the at the desired range or you want to know the relationship between the output as well as the controllable factors or also sometimes the noise factors. That means, you do experiment in order to achieve something what is

written in the left side of these slide you may be interested to know what are the most influential X variables for Y means the contribution of which of the X variables is more on Y.

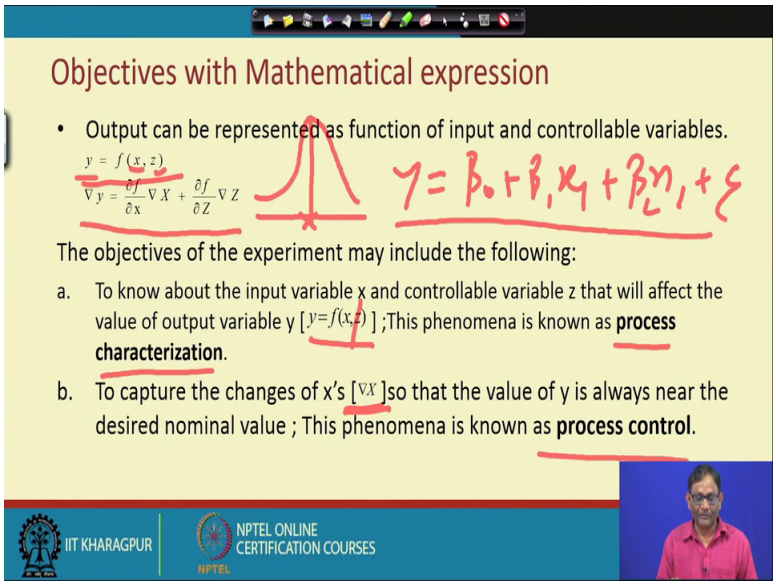
You may be interested to know where to set the X or the most influential X such that Y value the response or Y will be always at the desired level or the nominal level or for at the nominal value. Similarly you may be interested to control the X in such a manner that in such a manner that such a manner that the Y variability will be minimized or all though z the uncontrollable factors are not being controlled during experiment, because you either you are not able to control it or it cannot be control because of the cost point of view.

So, in that case is there any opportunity to manipulate the X in such a manner that the z effect on Y will be minimized. So, that is what is the purpose of experiment, what are the purpose of experiment at the first and you please understand the experiment is done on a process, what is a process a process is something which takes the inputs and converted into value added outputs and process is governed by 2 kinds of factors controllable uncontrollable factors.

So, the huge of experiment objective of experiment is to know 4 things one is determine the most influential variable that is X on the response of Y, said the most influential X in in some value or some range. So, that the Y will be always at the nominal value, other hand also you said the controllable factor X in such a manner that Y variability will be minimum. So, minimize the effect of z variable or uncontrollable variable on the output Y. So, I want to extend it little bit I want to write something here how to write anyhow. So, let us love further because there are some more important.

Student: (Refer Time: 18:49)

(Refer Slide Time: 18:50)



So, let us assume that the output variable is denoted by y x denote the x denote the controllable factors z denote the noise or uncountable factors and y is the output. The experiment is done to find out y, what is the functional relationship between y x and z. So, this can be told as y is function of x and z. So, when you do the experiment you get results of y x y and x z is uncontrollable you cannot cont manipulate so that you will not get.

So, essentially what happened sometimes you may able to manipulate z also, but that is very costly as I told you, but you for most of the time; that means, x you will control. So, if this is the relationship then from variability point of view I you can write that the variability of y is delta y and where n and it variability of f x is delta x and variability of z is delta z.

Then this kind of relationship can be written that you have seen. So, delta y will be del f by del x into delta x plus del f by del z into delta z delta y is the where the change in y delta x is the change in x and delta z is the change in z, what is del f by del x it is basically the coefficient the relationship coefficient if you change del x by delta x what will be the change in y that will be that will be finally, weighted by del f by del x similarly for del f by del z. From this equation we can we can give some terminology for of doe here, suppose you are interested to know what is this y is a function of x z or

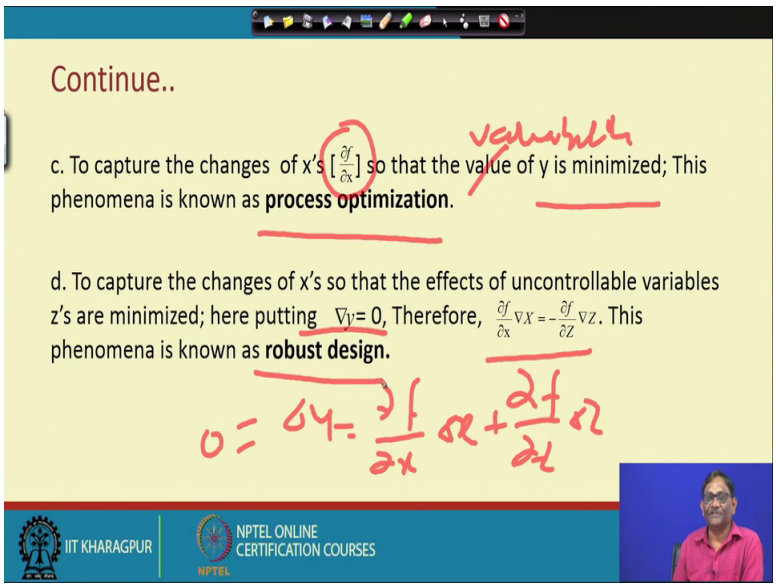
sometimes y is a function of x like y it a is a linear relation beta 0 plus beta 1 x 1 plus beta 2 x 2 plus error.

So, this is a linear relationship also you may be interested to develop this is nothing, but y is function of x. Similarly it may be non-linear, but you want to know these you want to establish these then this is once you establish this this is known as process characterization. So, what is process characterization process characterization knowing the relationship between y x and z, to know about the input variable x and controllable x and uncontrollable variable z controllable x uncontrollable z that will affect the output variable. So, y is function of x and z.

As I told you earlier most of the time this z will not be on z will not be included in the function. Second one is the delta x where to say delta x so that the y will always be at the desired level. So, what does it mean by desired level suppose I know that my y variability is like this this is a normal this curve and the mean value is here this is the desired value from (Refer Time: 22:49). So, u I say that this this is this y this is the nominal value or the desired value this value is also controlled by x or affected by x so where to put this delta x. So that Y will be always here it will not be always here.

But it will be around approximately within this. So, this is known as process characterization sorry process control y is f x z process characterization where to set delta x. So, that y will be at the nominal better that is process control.

(Refer Slide Time: 23:35)

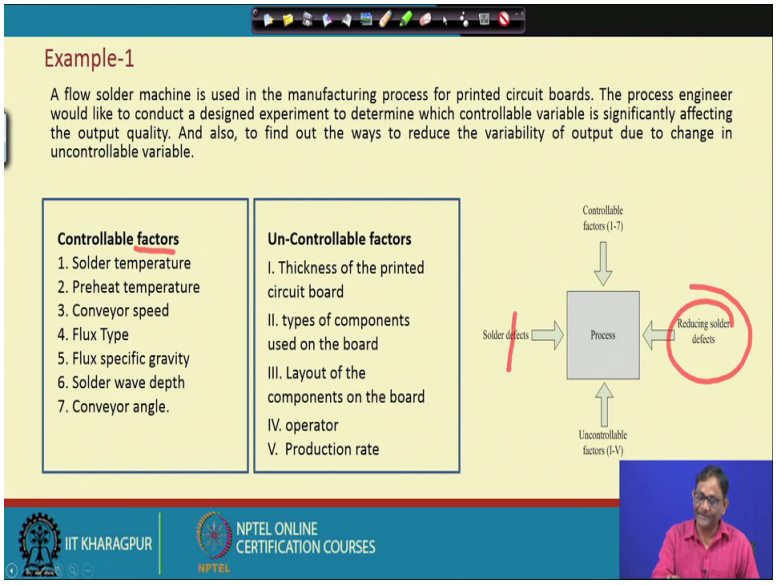


Then you will see in the another component is del f by this one del f by del x. So, to capture the changes of X that is del operate del x. So, that the value of y is minimized. So, you want to minimize the variability of y not value this is variability of y is minimized this is known as process optimization.

Another one is very important one equally that you know that z cannot be controlled during experimental that is it is uncontrolled variable, but there may be relationship between x and z. So, if you incase this relationship and accordingly you set x in such a manner, that the z effect will be nullified. So, if we put delta y equal to 0 in the equation that delta y equal to del f by del x delta x plus del f by del z delta z in this equation if put this one equal to 0, then you will get this the beauty of this equation is mean then if you change the x if you manipulate x in such a manner that that ultimately the variability of y can be further minimized and the effect of uncontrollable or noise variables that will also be nullified.

So, if not completely nullified it can be minimized so, this is known as robust design.

(Refer Slide Time: 25:29)



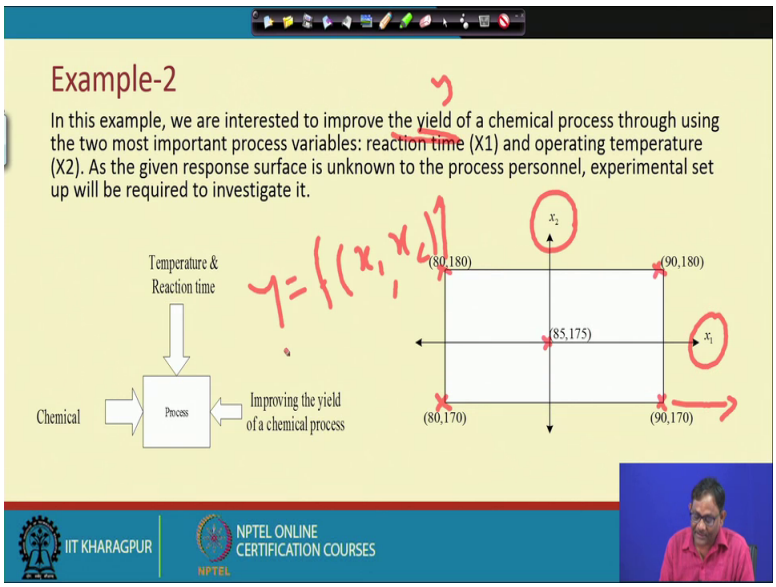
With reference to process characterization, process control, process optimization, and robust design, I will give you some of the some of the examples here. First example is the a flow solder machine is used in the manufacturing process for printed circuit boards, the process engineer would like to conduct a designed experiment to determine, which controllable variable is significantly affecting the output quality and also to find out the

ways to reduce the variability of output due to change in the uncontrollable variables. Then what is required this is a solder machine pro solder machine. So, this is the process is the this one.

Now the solder machine is the process, with a set of controllable factors we are enlisted the controllable factors, solder temperature, preheat temperature, conveyor speed, flux type, flux specific gravity, solder wave depth, conveyor angle etcetera. And it also has certain uncontrollable factors like thickness of the printed circuit board, types of component use layout of this operator production rate etcetera. So, if you want to do experiment your first thing is that you must know the system here is a solder machine and how the solder machine will work under this situation when manufacturing of process printed circuit boards and what are the inputs and what are the outputs. So, what are the controllable factors and uncontrollable factors here I have committed one mistake here that defect also I have it is written solder that it is not correct.

So, what happened the printed circuit board should have less soldering defects that is the purpose? So, once you do this then fact these are the factors, you want to control this factors, in such a manner that the defect will be reduced or minimized my second example is.

(Refer Slide Time: 27:54)



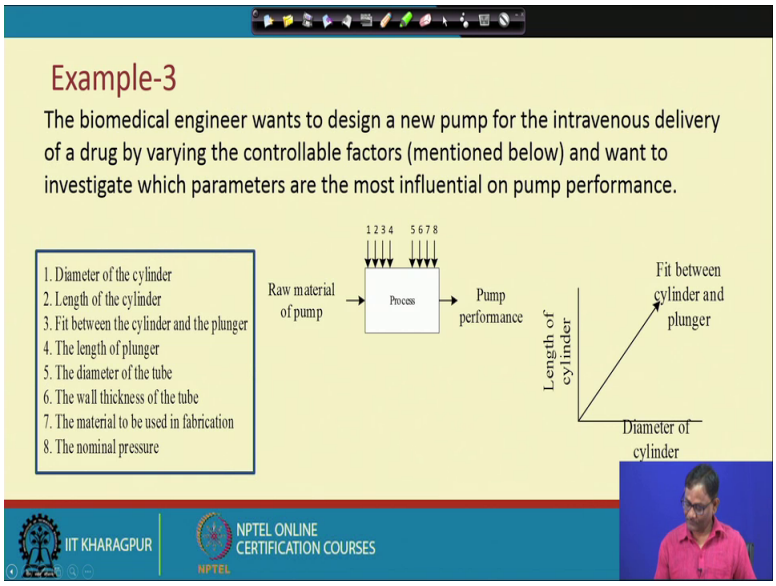
Here suppose considered a chemical process and you want to improve the yield of the chemical process and it was found that from the experts that the 2 important variables are

reaction time and operating temperature. So, you want to conduct an experiment in such a manner that you will know that how reaction time and operating temperature will affect the yield. So, in the right hand side we have shown that this axis is X 1, which is reaction time and this is X 2 which is operating temperature and the a reaction time from 88 minutes to 90 minutes and your operating temperature 170 degree Fahrenheit to 180 degree Fahrenheit this range.

Suppose you consider this range this is the of a range of operating zone that is of interest and then then you created such several settings here that first this one, second this one, third this one, 4 this one and also this is another one , but these 4 1 2 3 4 these are at the corners these are known as basically factorial points, that we will discuss later on, but further time being you understand that you will be interested to know how that process yield, how the process yield that is y that will be affected by x 1 and x 2. And you want to know what is this function x 1 and x 2 then process characterization, if you want to say no as suppose I want to I want to set x 1 x 2 in such a manner that variability of y will be minimized that will be your process control.

If you say I want to see that that the influence of x 1 and x 2 trap that that coefficient in such a manner that they should minimize then it will be your process optimization. So, you may in be interested that there are some other variables which is not considered here they may be acting as a noise variables so that their effect can be what I can say partial out. So, in that case you may go for robust design.

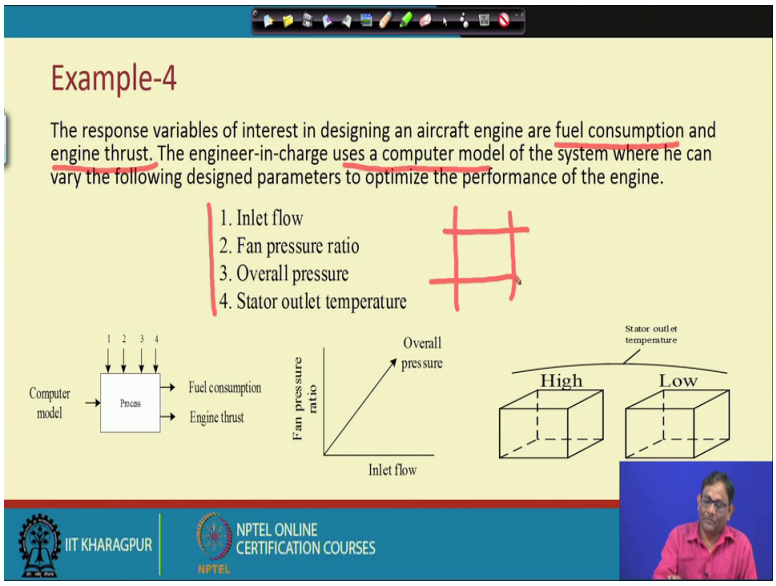
(Refer Slide Time: 30:49)



My third example is a biomedical engineer once to design a new pump for the intravenous delivery of a drug by varying the controllable factors, which are already mentioned here and want to investigate, which factors are the most influential pump performance.

So, there are 8 controllable factors identify diameter of the cylinder, length of the cylinder, fit between the cylinder and the plunger the length of the plunger, diameter of the tube, the wall thickness of the tube, the material to be used in fabrication, the nominal pressure and this pump performance basically new enter in a delivery. So, what we want basically maybe either the length of cylinder or diameter of the cylinder or fit between the see these things, they fit between fit between cylinder and plunger. So, they will be they will be manipulated to just see that how base the pump will operate.

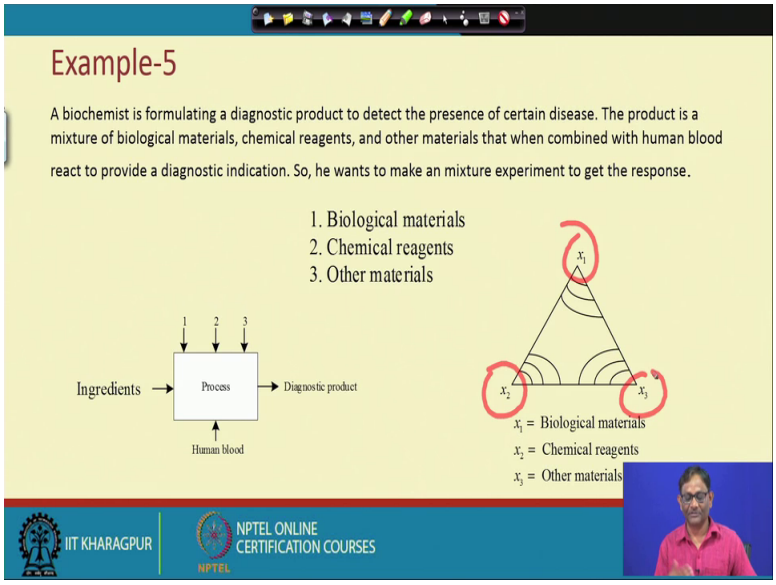
(Refer Slide Time: 32:00)



Fourth example here is maybe in the in the case of designing a aircraft engine. So, here what happened that the response and quality or the quality variable of interests are your fuel consumption and engine trust.

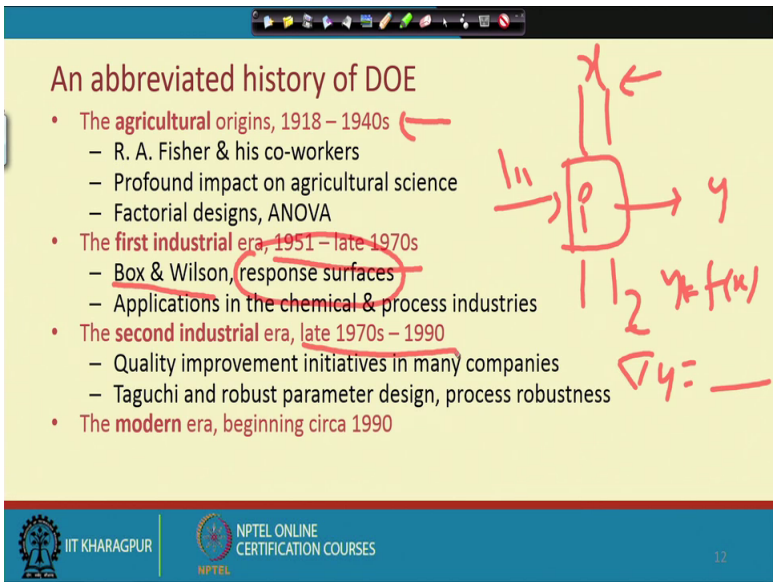
So, these are these are where controllable factors. So, you want to design an experiment in such a manner that you will be able to do those things what we have that we have already you are able to achieve the objectives like starting from process characterization to robust control robust design. So, here what happened actually the engineer is interested to use a computer model, it is not that the in the laboratory all the real productions of will be doing the experiment he will be he will be given a computer model and they are compute computer based experiments will be they will be developed. Suppose if we consider the 3 variable 3 3 factors usually this kind of cubic diagram we will use and later on I will we will extensively see that from your rectangle to a cuboid to the higher level that geometrical that geometric will be using.

(Refer Slide Time: 33:37)



Then in the example another example here suppose biochemistry is formulating a diagnostic product to detect the presence of certain disease the product is a mixture of biological materials and chemical reagents and other materials. So, your work is basically you want to see that how this 3 material will be mixed. So, that the performance of that the drug will be maximum in curing the disease. So, in that case this kind of design we will we will adopt. So, here you see that x 1 x 2 x 3 that 3 3 mix material, which will be mixed and this is known as the mixture design and there is a there is a good example or in good literature available or mixture design which is also coming under experimental design.

(Refer Slide Time: 34:37)



So, with these examples as well as what I have discussed so far. So, you must know that you must be able to or can say that define and experiment and you must be able to identify your problem in terms of the process model and when you find the process you must know that in a process there will be output variable, which are known as responsible variable or quality variables, if you are thinking in terms of quality. More generic name is responsible variable which we denote by which we denote by y that is y, there will be inputs there will be x and there will be x z input.

So, you this is what is the process. So, you must be able to from based on your system or the problem at hand you will be able to find out this unless you are finding out this one then you cannot go further. So, once you know the response variable and what is the objective you want to achieve in terms of response variable given x and z and input, then your work comes that how to how to set the controllable factors at what will be the settings of different factors, what will be the different levels of x and then how do you do the experiment and all those things will come. Primarily in this lecture I talk to you about the experiment I talk to you about process model then we say that y can be a function of x and z and we also have seen this equation delta y equal to this you have already seen.

So, your process characterization process control, process optimization and robust design those the concepts described here. So, all those things are done with the help of designed experiments, but if you see the history of designed experiment it is not that one the development it is a great history and it started here in the agricultural origin in 1918 by R.A Fisher and his co-workers primarily how to improve the yield from the agricultural land and that that era is 1918 to 1940 and there are lot of what I can this significant impact of these designed experiment on the performance, particularly the in the yield of agricultural fields.

And then that time factorial design and analysis of variance so these things where extensively used and this many concepts were developed, then after that the other pioneers or Box and Wilson actually in the in the in this period 1951 to let us 70s they worked a lot and; that means, the from agricultural region to the industrial sector it has been basically translated there and for primarily chemical and process industries are heavily benefited out of these designed experiments. And analyzing analysis of the experimental results and in box and I will come they are they are one of one of the

theoretical contribution is the response surface and that response surface, will be knowing also response surface in the in some of the lectures.

We you will also know factorial design you will be knowing analysis of variance and how to do different kind of analysis using analysis of variance that also will be known to you then a then came to this quality engineering era, which is basic primarily because of Taguchi that quality improvement initiative and Taguchi is robust design robust parameter design and he also concern given the concept of process robustness, from product robustness that time you will see that the many many other quality gurus where they are Deming, then your Juran, then your Crosby Feigenbaum, then that Taguchi and all, but they have different contribution in different era.

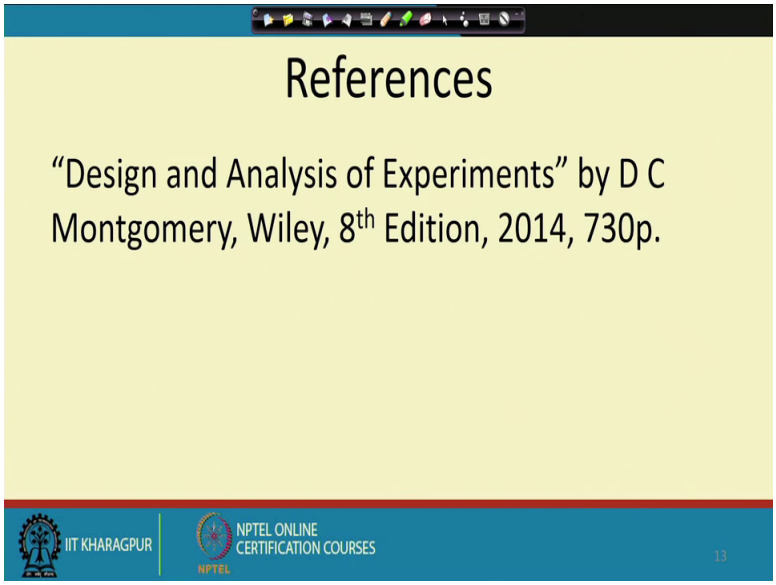
Particularly in this is a Taguchi that quality improvement all the techniques by Deming Juran Crosby Feigenbaum and Taguchi s parameter into robust parameter design mainly Taguchi kind the what their quality engineering and doe place huge role in the quality engineering domain and Taguchi table have the Taguchi given different kind of orthogonal areas and they are huge and then the Predi present they that is the modern era 1990 onwards. Now what happen that lot of data not only from experiment from other sectors also data has coming now, data analytics is a boom.

And in fact, the doe data this kind of data are also used because data coming from different sources. And we will see later on that min model, then variance model. There will be Joitn model, like mean variance model and there are lot of application of multivariate statistical models and new develop techniques that we that we will see slowly, but this subject is a this is basically a design and analysis of experiment this subject is generalized experiment. So, it will be it is a basic subject all the fundamental of experimental design and analysis will be talk taught you and the some detail calculation will be given to you different small examples will be given to you and there will be hands on hands on tutorials for you.

So, I am hopeful I am quite hopeful that you all will be will be enjoying this subject and you will be learning this, because you know the data will come design data will come from different sources also in laboratory scientists are developing in field engineers are doing it even the mangers are also experimenting ability. So, it is a huge it has huge

application and there are very good techniques also for doing this work will be taught to you.

(Refer Slide Time: 41:37)



So, with this I finish this first lecture, but at the end I must tell you that this is the book "Design and Analysis of Experiment" by Montgomery, this is an excellent book there are many good books excellent books available in the market. So, you can go through many other books also as I have heavily gained by reading this book. So, I recommend this book, but it is not that there is no other book there are many many good books available in this domain.

Thank you all, I appreciate your patience.