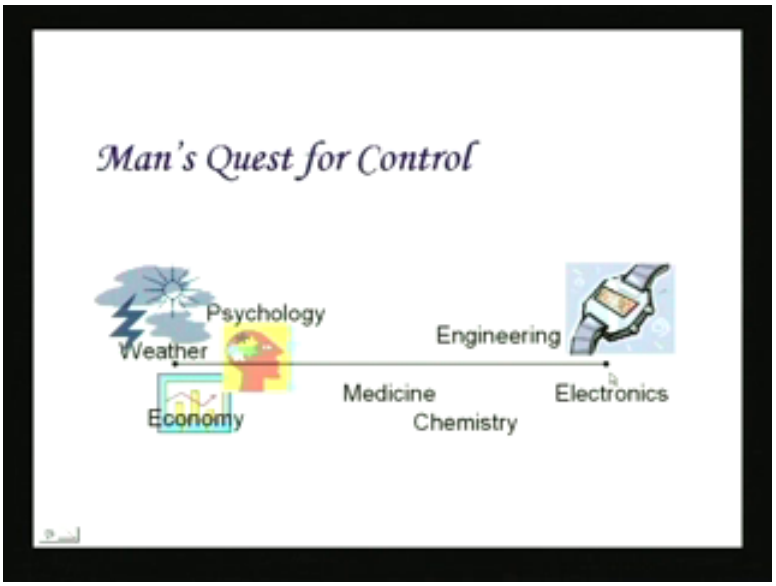
Six Sigma
Prof. Dr. T. P. Bagchi
Department of Management
Indian Institute Of Technology, Kharagpur

Module No. # 01
Lecture No. # 28
Factor Effect Calculations

Good afternoon, we resume the lecture today again on the series of lecture on Six Sigma. The theme that, we are discussing presently are design of experiments, this is process that is used by many times, when theoretical knowledge is not sufficient advanced it is not advanced yet to the point, when you can write down the equations that describe the behaviour of the system.

(Refer Slide Time: 00:47)



Just you can remind you we are coming from recall this picture, that I have showed you long time back. It had the the theme again here was the quest of mankind for control of various systems around us. We have on one hand systems such as the weather system or economy or psychology. These are systems where we cannot even list the causes that produce the effect; we can only observe the effect for most part.

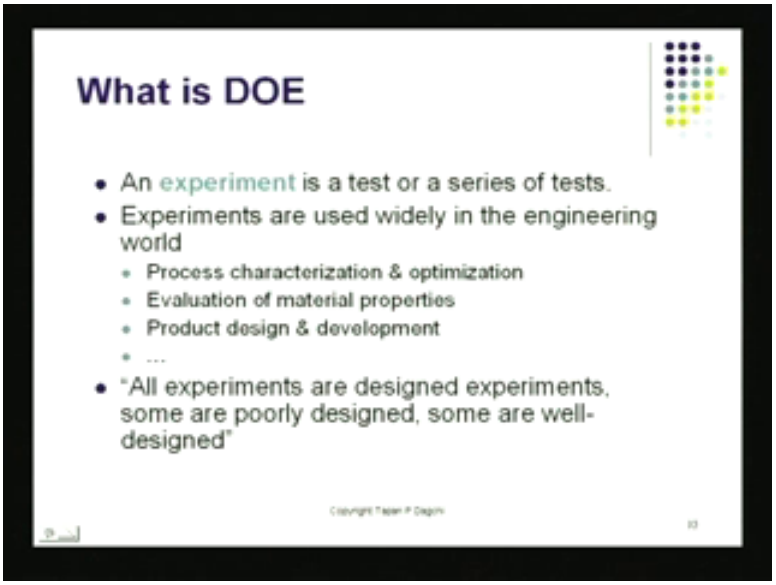
At the other extreme, we got the digital watch for example, where theory has progressed to the point, when we can write down the exact equations that basically describe the functioning of my digital watch. Here is my example, this the theoretical knowledge has

advanced to the point here, where I can write down the exact equations, that describe the behaviour of this watch; in between of course, are all these other systems there is medicine, there is chemistry, there is physics, there is engineering and so on, so forth.

Here of course, we are certainly not at the stage, when we can write all the equations but, we are trying to use experiments to move to the right, to slowly move to the right there was a point of course, when even electronics was experimental. The theory got built much later as we picked up on more and more better and better understanding of what is going on in the real system.

So, we started at the left extreme, which is like a system like weather or psychology or economy, when we could not even list the factors, we got to the point when we replaced list out the factors, we could draw the cause and effect diagram and then we used techniques various experimental techniques, empirical techniques try to sort out develop knowledge that would then, gradually move us toward equations, that would help us design systems, that are predictable behaviour, that we can predict.

(Refer Slide Time: 02:42)



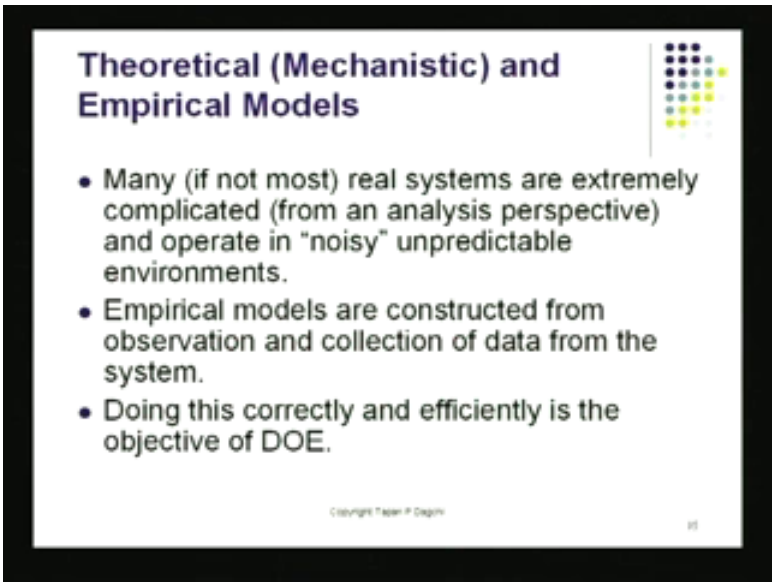
In today's context of course, we are going to be discussing this technique called DOE, Design Of Experiments and in particular I am going to be focussing today on the calculation procedures, that are involved in DOE. When I conduct DOE, you collect data. You collect data in a controlled manner, now what do you do with data our goal is basically, is try to understand, what are the effects, what are the different effects of

different factors that are impacting the system, that is like something you would like to know is that effect significant, when I compare that to noise and also if there is a interaction between the effect of two factors two or three factors that might be effecting this.

So, basically an experiment is a test of a series or series of test and basically these are used widely in a engineering and you know natural sciences and also physical sciences, what we really trying to do is we are trying to characterise the process and we are trying to achieve optimization of (( )) chart; we might be valuating material properties of things or might be designing a new product for example, or conducting some sort of developments.

And in fact, it turns out that, all experiment, they are designed basically in such a way that, some experiments turn out to be well designed, where as some other turn out to be poorly designed; the result is this with some experiments you can calculate you can go on further and you can convert the data into information that, you are able to do with certain types of experiments these are well designed experiments. There are other experiments where this is not possible, that actually indicates some problem with the design of these experiments that is the manner in which you manipulate the system and you collected some response data.

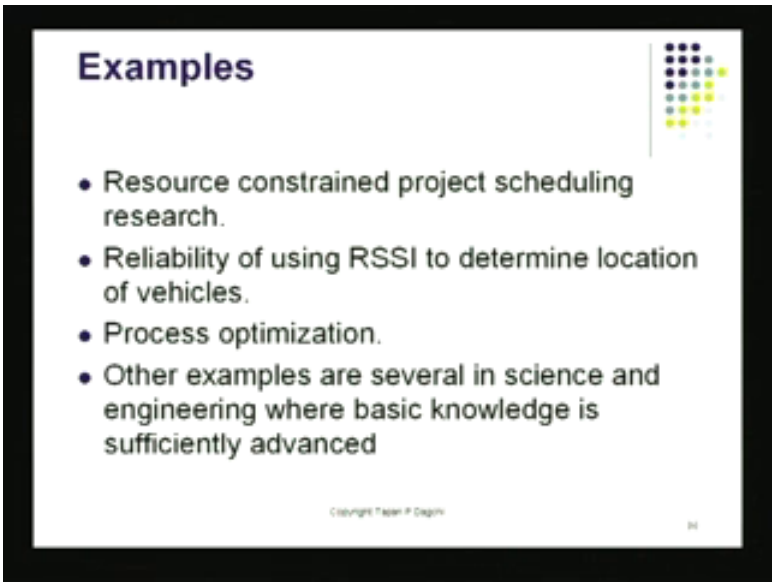
(Refer Slide Time: 04:20)



Let us see how we do it, the theoretical mechanistic models, these are empirical, these are models, which are based on theory then of course, you got empirical models, there are like I said before there are some systems, that we understand very well, we understand there functioning's very well and we can write down their equations as for thermodynamics or electronic circuits circuits or queuing equations and so on so forth. We can describe the behaviour of these systems in an exact theoretical manner that we are able to do, then the other systems, which are really complicated and these also involve a lot of noise and for such systems to for us to understand, how they behave we have to conduct experiments.

And the kind of experiments we would like to conduct these experiments are going to be the most effective ones or the ones that i planned by DOE design of experiments, so in fact, if you are trying to approach data gathering or knowledge acquisition through experiments, the best way to go there is to use the DOE path.

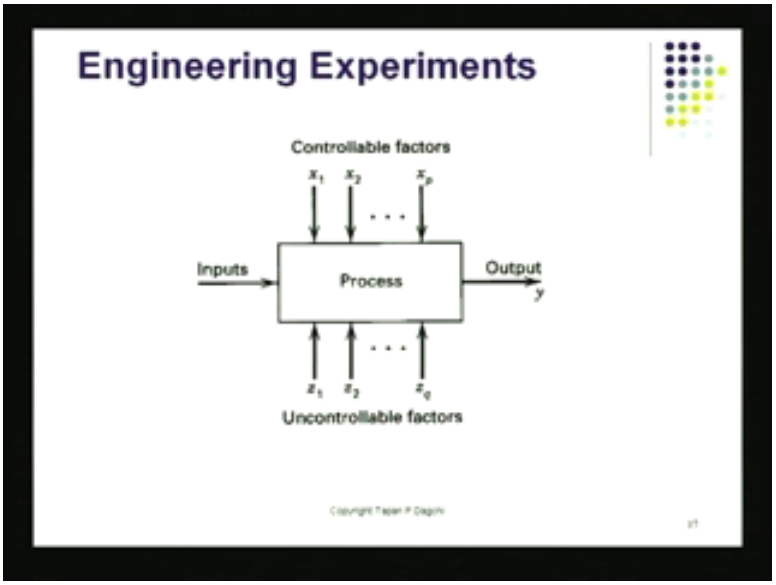
(Refer Slide Time: 05:25)



What are some examples? There are some examples, where you can use theoretical models for example, project scheduling constraint by resources; you can construct mathematical programming in model for it. This is something we could probably do, when you are looking at the RSSI system, this is the system that helps us locate the determine the location of vehicles; this is also possible with a theoretical model. Many times process optimization can be done by applying calculus or research methods, there

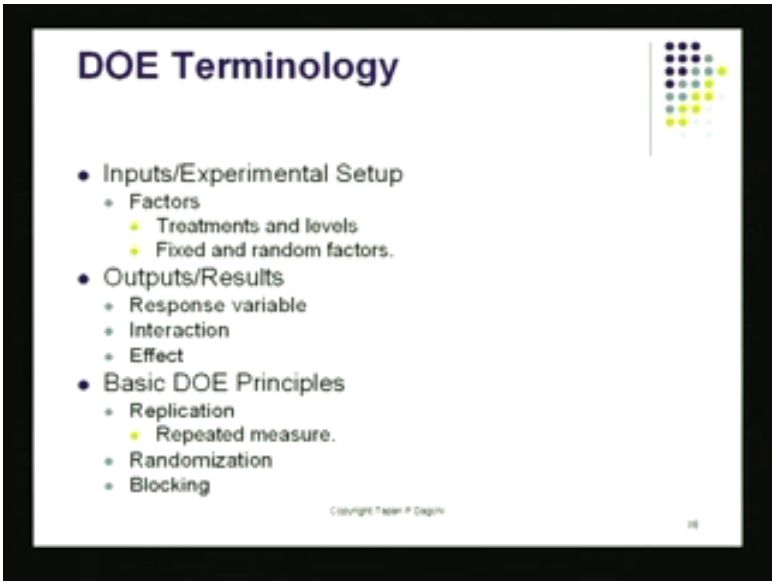
again theoretical model is available and you can actually go ahead and do it. And of course, you know all these areas science is advanced to the point, science and engineering both are advanced to the point, when we can write down the basic equations that describe the behaviour of the system.

(Refer Slide Time: 06:15)



This of course, is not possible with all systems and example is engineering systems, many engineering systems are such we can perhaps list the controllable factors and we find there are lot of uncontrollable factors, that cannot really enter our equations; then there are some inputs that the user or the driver of the of the of the of the basically, if person interested in leading the process operate or perform he provides some inputs, then we of course, we observe the output, these are these are typically what most engineering systems comprise.

(Refer Slide Time: 06:52)



What of course, we have to do is, we have to see, how this particular system behaves without really having that understanding not be able to control the output of the system. And the system by itself would be just standing alone and we would not really have much use of it. What we would like to be able to do is, we would like to be able to set up an experimental setup, where we control the inputs in a manner where we understand what factors are, what treatment levels are these are the settings, that we these different factors are set then of course, which factors are fixed and which factors are randomly fluctuating when the experiment is conducted.

Then of course, the output is what we are interested in so they could be something that, we call the response variables would also like to see as the output of the total empirical study process into action between factor affect this is something we would also like to be able to understand and also would like to definitely understand the effect of the main factors the prime actors in the process the the different main factors such as temperature, pressure, concentration, these are the factors that vary in the system and the these are the ones that is required some sort of adjustment, in order for us to optimize the system and get the performance come in a manner, come out in a manner that is optimum for us.

There are certain principles we apply when we use DOE for example, replication is a process that is used whenever, we are replicating the same thing trial again and again several times. This is because no matter, what system you are looking at like the system

that, I showed you there are some factors, which are controllable there are many factors that are uncontrollable.

Now, the ones that are uncontrollable those I am not setting and therefore, I do not know when they are on and when they are off or to what degree they have changed during the experimental and so on, that is really not known because, these factors are uncontrollable and many are actually unknown. In a situation like this the output is going to be noisy what would like to be able to do is, would like to fix the settings of the control factors at a certain design setting and then what would like to repeat the trial several times, so that the on the average the effect of the uncontrollable factors would get neutralised by the averaging effect this is something that would like to be able to do.

So, that is why we use replication we also sometimes use this principle called randomisation which is like, let us say you are running an experiment you are running a chemical process and this experiment is being conducted in an at in a in a situation in an environment, where the environment in changes from morning to lunch time to evening and so on.

Now, in the morning of course, temperatures are cool, so you may run the experiments you may be run experiments in the morning and if you are not controlling the temperature externally then of course, the experiment is going to be subjected to this morning and the temperature which is going to be cool, then around 2 o clock which is like about just pass lunch time perhaps temperature heats up and the weather gets warm and so on, so forth.

If you run your trials in that point perhaps the response would be different because, now the external environment which is uncontrollable factors has changed from being cool to warm, then the evening of course, again temperature is cool down and again if you runs an experiments at lunch time, when temperature temperature is ambient temperature is warm and then you run it again in the evening that will be different if if you are trying to experiment if you are trying to run an experiments in the these three settings for example, there should not be an order there should not be this this sort of things should be randomise.

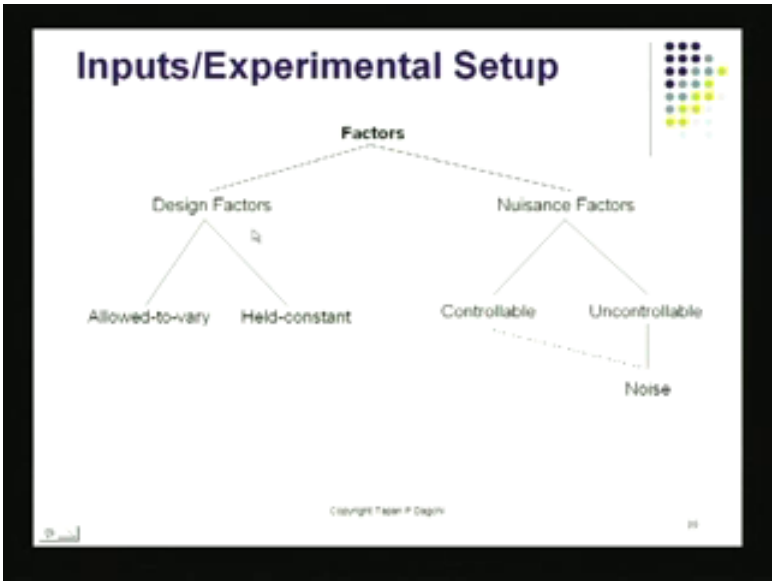
So, if you look at experimental trials you should randomise their conduct the the timing of their conducting between morning, lunch time and evening and one should not really

have for example, certain experiments run only in the morning and certain and process settings change but, those experiments are run only at 2 o clock then, again you run the same trials which some other settings at the process or the process factors in the evening if you do this there will be some effect of the environment also but, if the environmental environmental sequencing if it is randomised then of course, the effect of such external factor is going to be then more less neutralised, so this is like this the idea of randomising.

Then there are certain other principles also utilised in conducting experiments these are called blocking, and blocking is a way when there may be a factor that might be impacting your process but, you are interested in neutralising that effect so you you in fact construct blocks and this was actually this procedure was invented when Ronald Fisher was trying basically to try to come up with schemes or methods or identify the effect of various various factors such as fertilizer, water, sunshine, soil condition and so on, and seed quality on yield from farms there he used blocks, because he could not really take the same piece of land and run all these different trials and different sequences he did that by dividing up the total land into blocks and he then randomised the allocation of the experimental treatments those he randomised over this, so this is called blocking.

Blocking, basically does away with the with the any effect that might be there because of that other factor, which is not part of my experiment but, it is something that could also impact my my outcome of the experiment that is why blocking is there.

(Refer Slide Time: 12:16)



Now, let us take a look at different types of setups that could utilize in experiments, now as I as i mentioned to you experiments are done by controlling certain factors which are which we call the experimental factors, these are also the control factors, we change their settings at will we we follow a certain scheme in changing the settings of these factors.

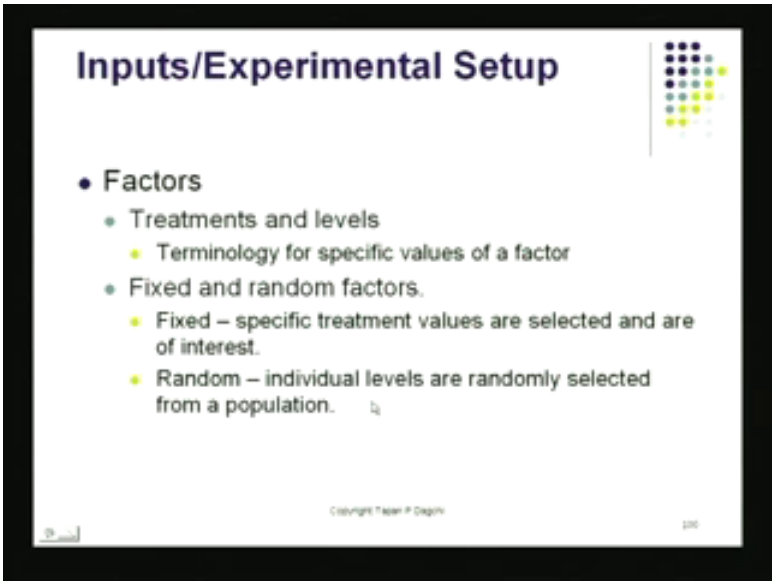
So, there are going to be obviously many different factors and if you do a little cause and effect diagram you will find that you know the main stream of the of the fishbone that goes horizontally then you got some factors that are affecting these (()) basically, coming and these are going to be the design factors these are the once you will be manipulating then the whole lot of other factors which are like random factors or noise factors, nuisance factor these are also going to be impacting your your experiment. So, basically the total process would be under the control of part your design variables and part these noise variables, they are both they are both both of these sets they are going to be impacting this thing.

So, this is something you got to remember when we design a experiment when we design the scheme or the plan to conduct experiments we got to be mindful of the fact that the design variables we are going to control and the basically the noise factors nature is going to control and what we have to do is we have to somehow neutralise the effect of noise or have a data analysis scheme, that can still help me get get basically, reliable results while looking at the output in a special way so that I neutralise the effect of the

noise variables this is something that will be done when we get into this phase of data analysis.

Now, design factors they may be allowed to vary randomly or they may be having constant or they may be allowed to vary according to some plan, that also is possible so either these may vary randomly or they may be varying according to a particular matrix or they may be held constant, that is our choice that is basically the experimental plan. Noise factors may be controllable or may be un controllable but, all of these are called basically noise factors and they are the ones, they are factors of course, they are impacting the system we are not really interested in looking at the effect of noise that we are not interested rather would like to find out whatever the effect it is of the design factors, so that we can optimise the process this is this is what we would like to be able to do.

(Refer Slide Time: 14:42)



So, what are we talking about, when we talk about design factors we talk of treatments and levels, treatments are for example, they are the different settings at which, I could conduct the experiment, there is some process factor which is going to be varying and I will be able to vary them this process factor for example, if it is temperature I could set temperature at 10 degree Celsius, 20 degree Celsius, 30 degree Celsius, 40 degree Celsius. So, there are four treatments now, 10, 20, 30, 40 four settings possible for temperature for pressure, I could have to two settings, I could have 10 P S I and 20 P S I

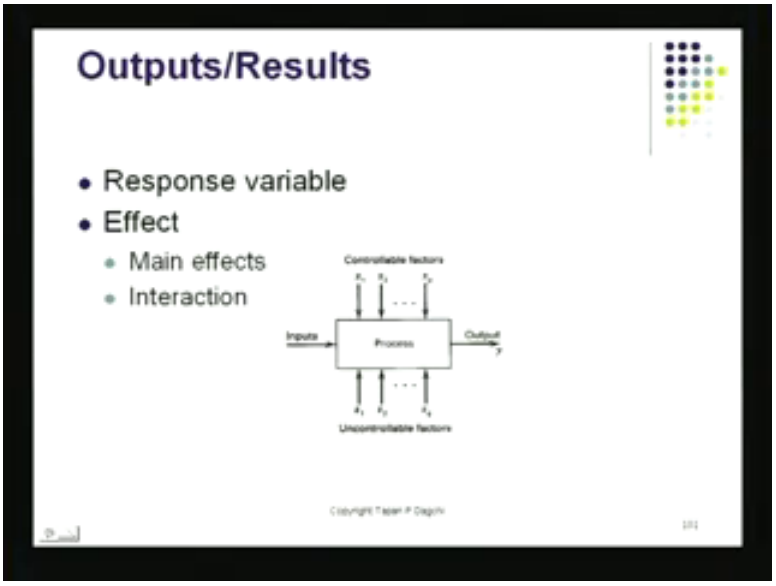
those are going to be my two settings I have got two treatment levels for pressure and I have got four treatment levels for temperature.

When I conduct the experiment of course, I combine the two in the full factorial situation I will be combining every pair I will be basically running the experiments under every pair possible with a pressure setting and a temperature setting, this is what I will be doing. So, it will be two times four that will be the number of experiment, that I will be running that will be like eight trails but, there are of course, this is the full factorial design there are other designs, we do not require to run all eight experiments and of course, they do not also produce the same amount of information the full factorial experiment will always produce the maximum amount of information not only about the main factors but, also about interaction between the factors factor effect.

The partial factor design when we choose not to run eight trails but, perhaps only four trails, these are special designs a special matrices that you that you use to guide your experiment and these these matrices are such, they require fewer trials but, they are clearly not as revealing as the full factorial experiments we will see that in a couple of minutes. And of course, when the factors are discussed when experiment factors are discussed there are situations, when I can fix the settings of these different factors, the factor levels at certain fixed levels these are going to be fixed factor experiments, then of course, its also possible that I run my experiment by letting the experimental factor vary randomly and we look at the impact.

The settings are chosen randomly in the random, randomly chosen experimental experiments fact to fact (( )) and if the factor levels are randomly selected from a population the effect is also going to be variable and the data analysis methods are going to be different from the one, when I have got fixed levels. In the fixed level case I am going to be show you both of these in fact I will be focussing mostly on the fixed level experiments, because most practical experiments they are of the fixed nature, fixed treatment level time but, there are situations when we might use randomly set selected settings of these different process factors that also we can do.

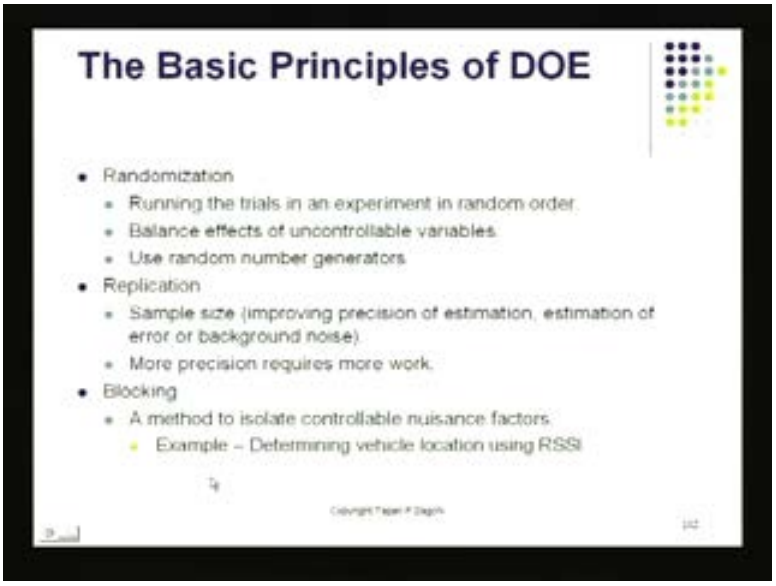
(Refer Slide Time: 17:44)



Now, what about the response variable the response variable is something that you measure and it is quite possible you probably measure more than one response variable because they are running the experiment, if you can observe multiple things coming out of the same trials go ahead and do it, because you have to reveal more out of the experiments as far as effects are concerned as far as factor effects on those responses are concerned there are two types in which, we have a lot of interest the first is the main effect and these are the effect of the individual factors such as pressure temperature concentration and so on.

Then of course, there are other effects which are called interaction effect these are perhaps the interaction between pressure between the effect of pressure and also the effect of temperature perhaps there is some interaction between the two and this is what brings in complexity when you look at the response; when you look at the total response of a process many times the total response is the composite effect of main effects, which could be as many main factors, that you have as many experimental factors you have then some complex interaction contributions also that might be there, we will do some calculations; we will try to see for the particular system chosen are main effects present there and what is their extent and our interaction effects present there and what is there extent we will we will see that as we get into it.

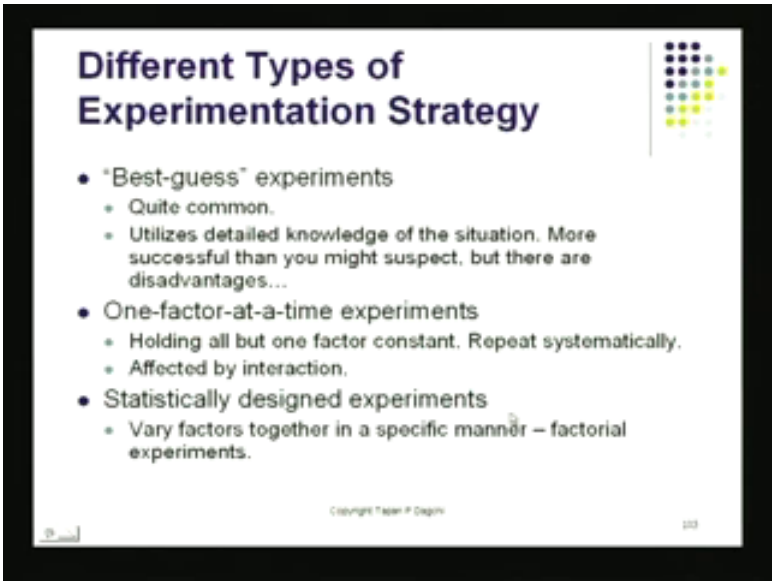
(Refer Slide Time: 19:06)



So, the basic principles just recap again we have the idea of randomising, which is like running trials in an experiment in a random order, so that the factors that you are not controlling they are not able to affect the the the experiments, they trials in a systematic way and of course, we many times we would also like to help like to utilise the effect of the the help the random numbers this is probably used this of used quite often to try to basically sequence the trials you have got.

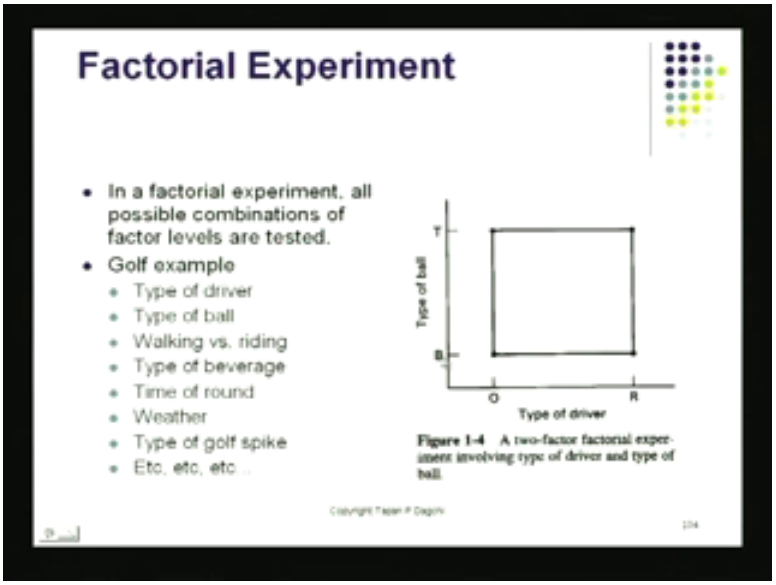
So you may run the first trial first, then you may run the sixth trial, then you may run the fourth trial and then you run the third trial and so on, you are randomising the sequence should be running your trials this is being done primarily like I said before in order to neutralise the effect of any factor that you are not controlling but, you want to fool nature. Replication is something that we also would be interested in when, we are conducting experiment and basically, the more you replicate the truer you are going to be to to closer to the true true response; because you are looking it them with with replications you end up finding the average response that should be looking at then, of course, the the there is the effect of blocking which is to try to isolate controllable noise factors and should able to block them, when you neutralise basically the effect of the these of nuisance factors and you are able to basically basically cancel out the effect of which on on the real process itself that also is is quite possible.

(Refer Slide Time: 20:34)



What sort of strategy is people use the design of experiments perhaps of course, is the most sophisticated method statistically designed experiments then of course, there are best guess (( )) best guess experiments I will give you some examples of this, then there are experiments that are like one factor at a time and of course, the best experiments are these statistical design multi factor experiments.

(Refer Slide Time: 21:00)



Let us try to take a look at what these are the factorial experiment and I am going to be giving you giving you one example and this is example comes from the game of golf and

when you are playing golf your score of course, your idea is to try to your score low hit the ball as few times as possible to complete all the eighteen holes complete walking through the full course and you complete sinking of ball in each of the eighteen holes, the fewest number of fewest number of strikes that you make in doing so, that becomes your score and the and the and the person who is able to do this in the shortest number of trials he becomes the winner in the golf game.

What are (( )) on the variables? What are the different factors that might be affecting your score? The type of the driver you know the driver is that hockey like stick, it is got that head there at the end, then you got the long stick there and you bash the ball you bang the ball, you out you do whatever you are doing that with the help of this driver various types of drivers are available and these are heavy, they are round they have they got grass face they got composites and all kinds of things are there then of course, the length of the bar that can also change there are various types of drivers available and they have different names also.

So even if you are looking at only the driver the real the real stick that you use to hit the ball straight of the tea to get your first shot the first shot is usually very long shot and like to get a driver that actually hits accurately and also hits fast and far this is what would like to be able to do so there will be some choices there there may be a special driver or there may be regular driver.

So, there are two type two types of possibility, let us say we are trying to optimise the choice of driver the type of ball we are going to be using or we going to be walking around when we play the game of golf or we going to be riding a cart what kind of beverage we will be drinking, when we basically have you know the need to drink something we can either be drinking beer or could be drinking water or perhaps he could drink nothing, that is also is possible the time of round which could be morning, afternoon, evening whenever it is what is the weather like that might also affect your score and the kind of spike you are using and so on, so forth.

There are many different factors that can impact your total score total golf score lets simply conducting an experiment and we are going to be doing a very simple experiment and we are saying we are saying we got a choice of two drivers, two type of drivers one is the ordinary driver and the other is the special driver. So I will I will just going to be

calling them type O driver and type R driver these are the two types of drivers we have, two type of sticks we have to bat for ball, as far as the ball is concerned there is the regular ball and this is called the bollate Italian ball or there is the T type ball this is a special ball which is marketed by some particular super markets super markets, so we have got two types of balls also that are available.

So what are the two choices I have two, I basically two drivers with which I can play with which I can hit and of course, I have got two types of balls, that I can use in my place, so two times there are four possibilities then there are four possibilities you see then that this becomes this actually gives me two times to four different ways in which i could play this game and those are shown here if you look at the matrix here, these are these two four these these two by two.

And therefore, four possibilities, the first one is going to be type of ball, the type of type of drive driver and type B ball, so, this is both O type driver and B type ball and this one is the R type driver and again B type ball this is like another possibility then of course, O type driver and T type ball and of course, this one is the last one last combination possible which is like T type ball and R type driver.

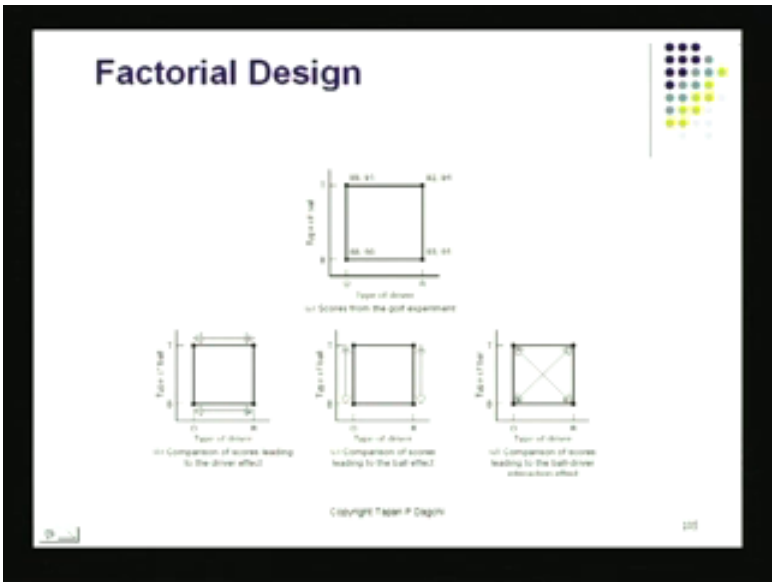
So I have got O B and I have got R B, I have got O T and I have got R T these are the four different ways I could play this game and borrowing you know something that we mentioned earlier we should not just play one round of game with this four different combinations, we should play multiple trials because there may be these other factors that I am not including in my game

So let us actually do this let us play multiple multiple replicates of the same game so we play the game play the same game multiple number of times let us how we play it. Perhaps if you have enough money four times each, so you will be playing the B O combination four times we will playing the B R combination four times, we were playing the T O combination four times, we were playing the T R combination four times after that of course, we gather some data and we record the data from this we are going to working out the effect of the driver the main effect of the driver this is something we need to be working out.

So, you find out if going from O to R if this reduces our score it could reduce or it could increase we do not know that here and of course, also when it comes to changing the ball

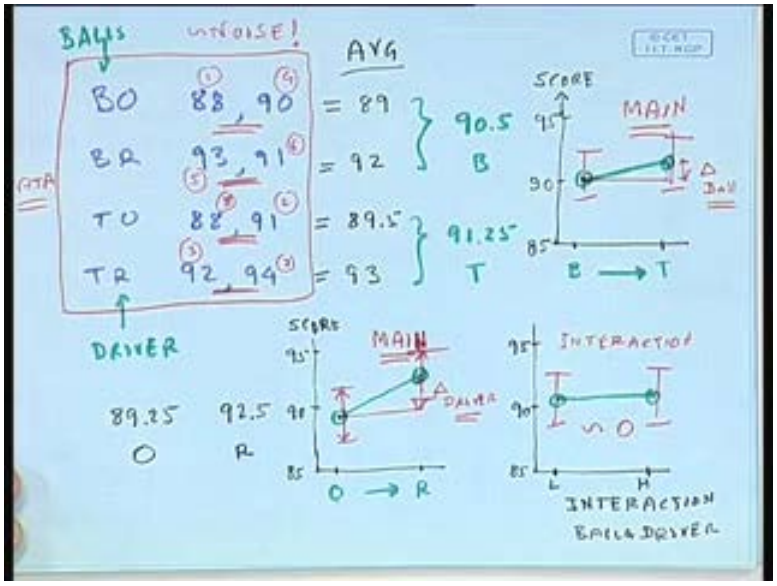
type in using ball type B versus ball type T what happens to the main effect of it I mean do I do i score? Do I do i end up with the smaller number when I play with ball type T, I suppose to play with ball type B that is also, something we would like to find out.

(Refer Slide Time: 26:49)



What we now do is we will take a look at the results of this game let us say I have done that and if this yes of course, I show you two trials only I show you the effect of two trials, so look at the top this is where they will look at the top diagram there this is what I have recorded my two rounds basically I said earlier that would be playing four rounds at each setting, let us say I I did not have enough money, so I i was, I have enough money to be able to play four rounds once and then four rounds again this was I was able to play I have recorded my scores here.

(Refer Slide Time: 27:37)



So, per I am going to be writing them down so i have here for example, i have for the B O combination, for the B O combination B O combination my scores are going to be 88 and 90, for the B R combination my scores are 93 and 91 and for the T O combination my scores are 88 and 91 and for my basically T R combination my scores are going to be 92 and 94. These are this is the data that I calculated data that I recorded.

Now of course, I have to do my calculation because, I have to find the effect of for example, you know type B ball, type T ball I have to find the effect of using the O driver as posed to you using the R driver we have to find that, basically what we would like to know is would like to know is there an effect due to change my my change in the driver with which I hit the ball is there is there an impact on my score; if I change the driver from the O type driver from being, O type driver to a R type driver this is something this is one of the things that, I would like to find out the other thing I also like to find out is, when I use ball type B versus ball type T is there an impact on my average score this is like something I would like to be able to work out.

I have already got my experimental data which is shown here this this is the experimental data that I have gathered I have drawn the two experiments and I have done it in a way when I have randomised the trials. So perhaps my randomise trial was if I just put down the sequence here, perhaps I played this one first then, I played this one, then I played this one, then I played this one, then I played this one, then I played this one, then I

played this one and I am left with this one, that (Refer Slide Time: 29:44) that is what I played the last as (( )) are concerned I have randomised them.

So, hopefully the effect of all the other factors is going to be averaged out because of my randomisation now, what we would like to be able to do is we would like to find the effect of going from type O driver to type R driver. Let us see how we find that out, what we have to do is I again go back to the picture here and I show you what I do is I take an average I take an average here, first of all I have to find the average of 88 and 90, that is something that would be the average score there, I do the same averaging here, I do the same averaging here, I do the same averaging here (Refer Slide Time: 30:32).

This of course, is not going to be terribly difficult this is going to be 89, this is going to be 92, this is going to be 93 and this is going to be, basically 89.5 those are going to be my averages.

So, let me write down my averages and my averages are going to be and I am just going to be put down the averages here, average score is going to be here is going to be 89 it is going to be 92 it is going to be 89.5 and its going to be 93 these are the averages the utility of these averages is this is neutralise then the effect of the factors that I did not control when I was playing my game I put all of those as factors that are other than the driver and the ball this is what I did from this of course, we are suppose to find the effect of the changing balls or changing the driver

Let us see how we do that what I do is, I find the average of two quantities and those two quantities are going to be the following i average out these two and I average out these two, let us try to do that let us try to do that precisely I will be using a calculator, so I will just put that here and I have got essentially 88 plus 90 divided by two that is 89 that is already there so that is confirm.

Now, let me average out this one these two I have 89 plus 92 divided by 2 that is 90.5, so that that number is going to be 90.5, so I have 90.5 and what is common between these two B b is common so this is going to be the effect of B I am going to give, I am going to tell you a little bit more about this in a couple of minute, that is also average out these two quantities I have got 89.5, 89.5 plus 93 divided by 2 that is 91.25 and when I am doing this what is common between these two is T.

Now, remember my column here this column was for balls and this column here was for driver and what I have done here is I found really the average effect of playing with ball type B and the average score of playing with ball type T, so I could quite easily at this point produce a plot, now what would that plot would be like I could just draw the plot here I first draw the axis and start at 95 and then I will have or let us make it 85 it is quite low 90, 95.

Let us lets put on this side the two ball types and those are going to be my B and T b and t what do I see now, I see the score of B to be 90.5 that is what here somewhere, score of T is going to be 91.25 and that is going to be somewhere here, if I join them there is a slight rise as you can see there is slight rise from 90 to 91.25 that slight rise is to, is because of the change that I made in going from ball type B to ball type T, this is the this is the effect, in fact if you look at the the number carefully you will find this delta this delta this delta is due to ball change, that is the effect of basically changing your ball.

Let us do the same calculation using the driver, so for that what we have to do is we have to average something different let us see what we average, we write down we write down here the average of type O first so that, is going to be I put down here the average of o type first that is going to be 89 plus 98.5 and I will work out the average of it that is going to be 89.25.

So I can directly write here, the effect of O type driver is 89.25, O type driver just like I have B type ball giving me 90.5 for O type driver I end up with score of 89.25 only the average, let us look at the other one and the other one is going to be R type driver and that is going to be the basically the average of 92 and 93 that is going to be 92.5, that is going to be the effect of R type driver see here again the change that took place when I change my driver from O to R is going to be an increase in score and this I can again write I i can plot it again, like I did my plotting before I have to pot like this, I will have some quantities there start 85, 90, 95 this always is this side be y axis is my golf score.

So, it is going to be score here, score is going to be on this side and what do I plot on this side let us try that let us try to write that I have 89.25 that is going to be, lets first put down what what is that we are plotting we will plot type A, type O is here and type R is here and these are what these are driver types O type driver and R type driver or O the score or O the score is 89.25 that is the average score, that is I am here and for R the

average score is 92.5 that is going to be somewhere up here, if I join them that shows me an increase and again as you did before let us show this effect this also main effect this is also a delta that is due to driver this this this change is due to the driver changing from O to R this is the change in score this is the change in the score

So both this this is the main effect I could just write it main, this is also main effect main effect always is the main effect always is the effect of a single factor main here and main here, so I have got a main effect calculated there and the main effect calculated there now, what about the interaction interaction becomes a little little more complicated to evaluate and for that what we do is with average of two quantities should average of 88 90, 92, 94 which is like 89 and 93 if you average these two out I can actually get interaction factor I will say the and I am just going to be plotting that directly. So let me just my scale here first I am going to be plotting this time the interaction between the two's I am going to be my scale here, interaction low and high and I just call it interaction and this interaction is going to be now between ball and driver low and high.

Let us say how I calculated that all and of course, my scale is marked at 85, 90, 95 that is my score, let us try to see if you could get my interaction plots done I will just make sure the reading is I have the low level is going to be the middle point so that, is going to be the average of 92 and 89.5 that is going to be very close to if you really look at it I could I could just use the calculator do that 92 plus 89.5 divided by 2 that is 90.75 ninety point seven five is the low level that is what here somewhere and let us do the other one, that is like the between 89 and 93 and that actually 91.

So I am going to be putting that right about here 91, notice here it looks like this almost no interaction this is the interaction, once so let us see at least give it a title this is interaction what have we done. Let us try to summarise what we have done we ran some experiments and our experiments for run using in this particular case different types of balls two different types of balls and two different types of drivers.

We ran those trials and we ended up the data here what is the data just try to put that within a red sort of zone my experimental data right here, that is what my data is, this is my experimental data and if I just shift up in little bit you will be able to see I wrote data there when I have my data there the only thing I have to do is I have to remember what

settings are obtained those things what are these two numbers, these are two replicates of experiments run that is like my game game played under the same setting.

So I was using ball type B and driver type O and play two rounds and the average turn out to be 89, I ran this other trail this is like the other game i played another game I played ball type B driver type R and the average turn out to be 92 then I used ball type T and driver type O and the the result turn out to be 89.5 and of course, ball type T and driver type R gave me an average of 93. Once I have these because of particular structure of this matrix here I am able to find my main effects, the two main effects one is of course, for this is going to be for ball type and this is going to be for the driver, so I have got the effect of the ball and the effect of the driver both I have got pulled out of the same data and of course, not just this I have also got the interaction found out.

So in fact, it turns out this system has very little interaction very little interaction this is close to zero that means there is no interaction between these two that means if I was trying to lower if I try to reduce my golf score, I could manipulate this variable independent of what I do with this variable I could I could bring this setting to b and I could bring this setting to O, so if I have to minimise my trials if I have to minimise my trials I would be best of playing with the ball that is of type B and that is and and using the club type that is going to be your O.

In fact you do see that this course are the lowest here the average scores are the lowest with this so this has vary files what we just concluded while looking at the graph this method of doing your calculations its quite straight forward and it can be done if you follow particular matrix structure and that matrix structure is shown here.

In fact I show the same matrix structure if you look at the diagram let us comeback to the diagram here, again here is the diagram I have a two by two situation and I would be running like if you assume this to be minus and this to be plus and this setting to be minus and this setting to be plus, I have minus minus at this setting, minus for this minus for this variable, two variables minus, minus this sign is going to be plus, minus plus minus is here, here it is going to be minus plus and this this setting is going be plus plus. So, I have drawn basically matrix which has got minus minus, minus plus, plus minus and plus plus this is a matrix structure, this is a matrix discipline basically, this is called

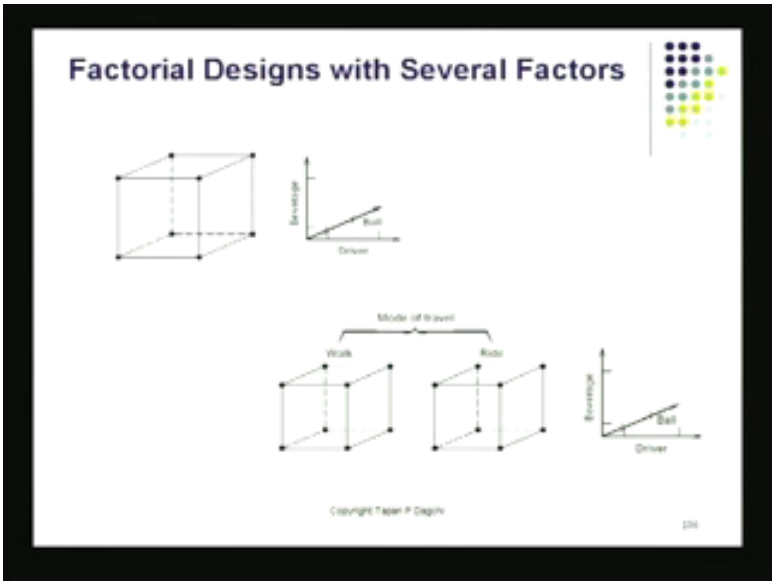
the full factorial plan when I using two factors at two levels each if you cross them you end up with four combinations and that is exactly what we covered in doing these thing.

Now, this is this approach is if the noise variables do not have too much of an affect will you say noise had no effect what do you think, well look at these numbers look at these numbers if the other factors which I am now ignoring if the other factors no effect out of had 89 write here, throw out if noise had no effect out of 92 throw out if I, if the if the other variables are no effect out of it 89.5 in both of these trials and out of had 93 at both of these trials but, that has not happened.

So, the variation that you see here is really due to noise, this is something very important first to realize that actually noise has disturbed the system and many times this noise is so high, that we are not able to draw these diagrams, because in place of just one value we will end up with a variation there like this, a variation like this, a variation like this, a variation like this, a variation like this, a variation like this (Refer Slide Time: 46:57).

And when you got background variation that is varying like between this number and this number of, this number and this number (Refer Slide Time: 47:06), then to be able to tell the difference whether there is a statistical difference between this number and this number I have to do an over I have to do analysis on variance just like a special mathematical technique which will be in shortly will begin it in this session I will complete it on next session, that is something will be doing if the background noise is too high. Now we could ignore that because these variation these variations here they are not too big there are not too big so we got away by doing this this is something we got to remember.

(Refer Slide Time: 48:04)



Lets see what we produce we are proceed next it is possible that you do not have enough money to be able play all four plus four plus four eight round is very possible that you would like to probably cover another factor if possible but, you do not have enough money to be able to do that now I had i had four plus four eight rounds run here. Suppose I had the same money available could I somehow could I somehow study another factor now what happens if you just bringing another factor, suppose you just bringing another factor what happen, suppose I wanted to know experiment with three factors we actually wanted to while we are playing the game we wanted to vary three experimental variables one would be the driver the other would be ball type I used the third would be the beverage that I drink.

If I just doing this if I had three variables I will end up with eight different possibilities I will be changing driver this way, I will be changing ball this way, I will be changing my drink this way and I end up with eight different possibilities eight different combinations so instead of the previous one which had only four trials now I end up with eight trials but, the advantage is, now I can study not only the main effect of three variables that is ball type driver type and the drink I use when I am playing golf but, also they are interactions those also I can figure out.

If I run these eight trials so its an advantage this is actually a factorial design in effective you can keep adding like this if you are willing to conduct more trials if you conduct,

willing to conduct more matrix of experiments you can get a lot of information, you can get the information about the main effects and also you get effect on the of the of the interaction it is a both types of both types of effects you can figure out if you are willing to run full factor experiment.

Now, suppose i have the (( )) fourth factor and that factor was going to be the mode of travel I use when I play my golf for example, I wanted to play golf not just by walking around, because it is a lot of distance probably three, four kilo meters I have to walk perhaps (( )) see more in some golf courses on some golf courses what I would like to when do is i perhaps like use that cart, the electric cart that take me from you know hole a to hole one, hole two, to hole three, to hole four and so on, so forth.

This is where they have little, these little flags there and you basically drive from point a to point b to point c and so on, so forth. It basically ride this cart that we could probably save some energy if you look up more rounds and probably do it quicker also all those things are possible if instead of walking using this cart.

Now, that might give you some rest also right so the expectation is if I want rise my performance which is like if found lower my golf scores perhaps I should ride in a cart, how do I find out if that has any effect well I could run, now I already have 8 trials, I could run 8 trials while I am walking and I could run the same 8 trials or similar 8 trials, when I am riding a cart, now I got 8 plus 8 16 experiments and of course, 16 are not going to be enough, I will also have replicate all these because there are these noise factors that I ignored.

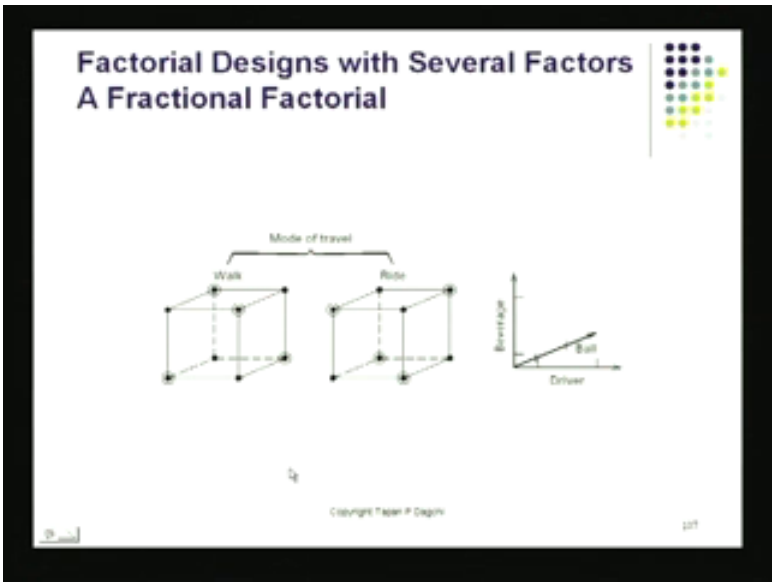
So I will have to probably now run 16 times 2 may by 32 rounds of golf that is not going to be all done in 1 day by the way you may have to come back on different days and then of course, you will have to also think about randomise in the sequence in in which you play your golf golf care unless you do that, which is going to be messi situation this is something where we would like to make sure that you use a proper matrix use a proper matrix to plan your trials you use these sequencing correctly.

So that, you know the effect of these random factors they do not meet into bias your trials you do not like play your games when your fresh some of the games when you are fresh and some of the other one's like for example, if you if you look at this matrix you

should not play all the, all you know all four games here four or eight games if you replicate in the morning and then all these four games in the evening.

If you do that then the effect of drinks is also going to be what we call basically it is going to be confirmed at with the effect of temperature variation from morning to evening, so you would like to run some of the experiments some of those you know I got the eight points there some of those you might like to play randomly in the morning some of those play, some of the you might like to play in the evening, if you do that then of course, you will end up with the completely randomise full factorial trial, then you come out there your your effect of noise is going to be low and you could really work out the these calculations without too much trouble at all.

(Refer Slide Time: 52:41)

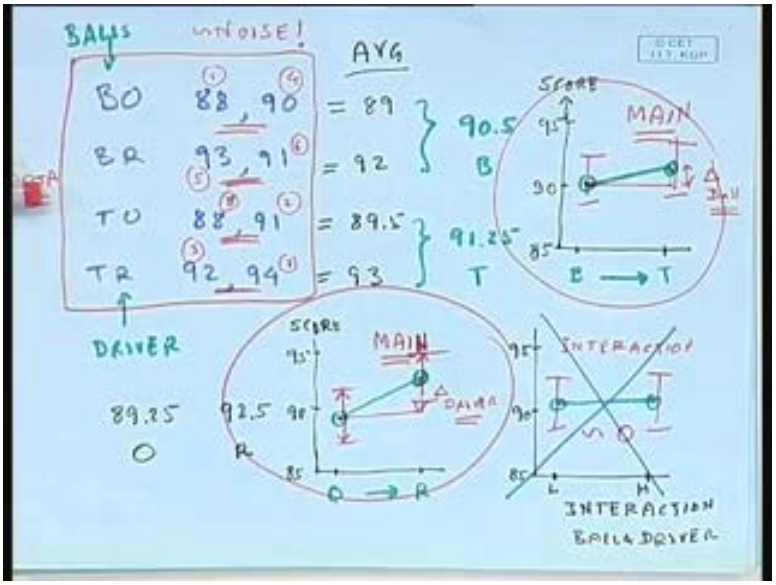


Now, suppose you did not have enough money and you still wanted to do this you will have to give up something, you still wanted to find the effect of balls, you wanted find the effect of drivers, you wanted to find the effect drinks and also you want to find the effect of riding a cart (( )) but, if you did not have enough money so you could not really run in this case eight plus eight sixteen trials you could not do in that case you could actually run what we call partial factorial design.

So instead of running the full factorial design you could run partial factorial design and here some example, I show on the screen here three variables are here, so I have go balls, I have got private types and I have got beverage, those three are here and then across on

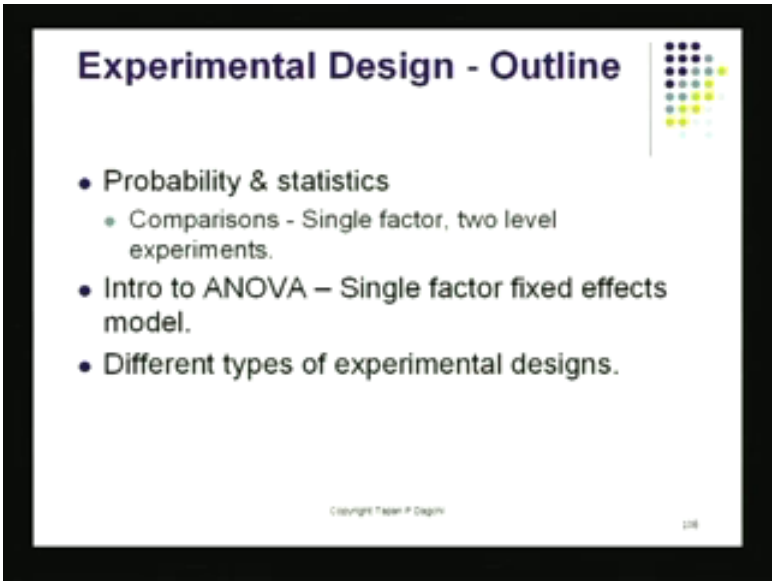
this other side I have here, I am walking here, I am riding the cart but, notice here I do not run all the all the eight experiments I select to be run this corner, this corner, this corner and this corner, after this when I run the other four corners that corner, this corner, this corner, this corner, (Refer Slide Time: 53:45) in doing, so I have really cut down my total numbers of experiments some sixteen to eight, so this is not a full factorial design this is called a fractional factorial design or a partial factorial design.

(Refer Slide Time: 54:02)



This can again lead to the calculation of the main effects, so the main effects that we saw here, I could do this experiment I could do these calculations using a what we call a fractional factorial design, I could do this one also using fractional factorial design but, I am sorry we could not do this experiment, we could not do these calculations using fractional factorial design.

(Refer Slide Time: 54:30)



So what you have to remember is we have to know the objectives, we have to know the objectives of doing the experimental work if you do that of course, then we got no real problem there and for more complicated situation something that I will be discussing in the next session is to try to utilize the technique of anova which can be done when the background is pretty noisy and just be your replication is not enough for you to complete your study I will continue with the discussion of anova in the next session. Thank you very much.