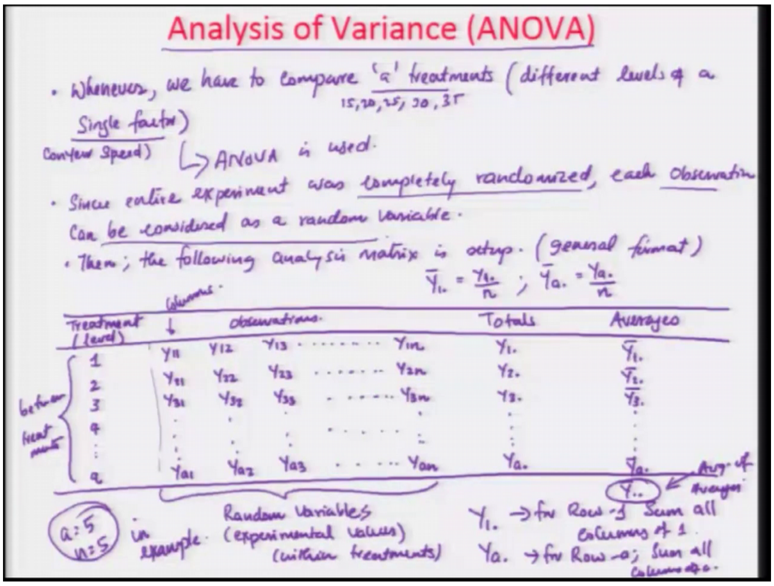
Advanced Green Manufacturing Systems
Prof. Deepu Philip
Dr. Amandeep Singh Oberoi
Department of Industrial & Management Engineering
Department of Mechanical Engineering
Indian Institute of Technology, Kanpur

Lecture – 27
ANOVA – Part 2

Good afternoon students, welcome to the course of Advanced Green Manufacturing Systems. And today we are going to discuss new technique or a special technique called ANOVA; Analysis of Variance, which is in continuation with the experimental design concepts that we are covering in this course. And I am Dr. Deepu Philip and I am from IIT Kanpur.

(Refer Slide Time: 00:41)



So, then we get into what we called as the analysis of variance ok. And what we going to talk about how and how do we do the mechanics of the all thing is what we are going to talk about. So, we are saying whenever we have to compare, whenever we have to compare a treatments a treatments ok which is different levels of a single factor different levels of a single factor.

The single factor here is what is a single factor? Single factor is a conveyer speed ok, the levels a treatments these are 15, 20, 25, 30, 35 say (Refer Time: 01:35) which is the speed of the conveyor right. ANOVA is used we use ANOVA here ok. Then since entire experiment was completely randomized, each observation can be each observation, each observation can be considered as a random variable considered as a random variable ok. So, because you completely randomize the experiment we are comparing each observation can be compared as a random variable ok.

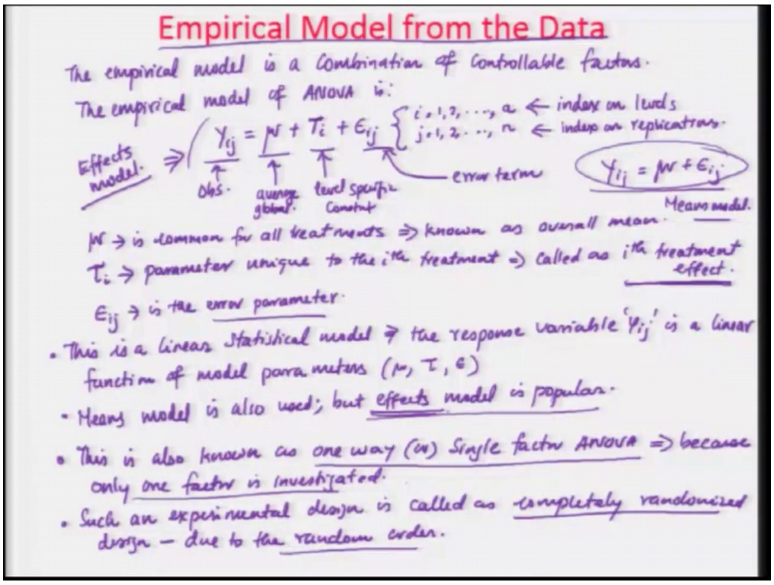
Then the following analysis matrix is setup ok. So, we set up the next matrix. So, what we do is by this is the general setup general format. So, it is usually a table matrix. So, you have a treatment here treatment is also the level of the factor right that I told you it this 1 2 3 4 all the way up to a, ok. Then you have is your observations. The observations include Y 1 1 Y 1 2 Y 1 3 etcetera all the way up to Y 1 n. You have Y 2 1 Y 2 2 Y 2 3 all the way up to Y 2 n then Y 3 1 Y 3 2 Y 3 3 all the way up to Y 3 n ok.

Then you have similarly each one of these values are there ok. Y a 1 Y a 2 Y a 3 all the way to Y a n right. So, these are your what you call as random variables or these are the experimental values experimental values whatever the observations you get of Y as part of the experiment; then you do something called as totals. So, totals is Y we denote it as Y 1 dot, Y 2 dot, Y 3 dot all the way to Y a dot. So, what we are saying is for this particular row Y 1 this row sum all the; so, Y 1 dot means for row 1 sum all columns ok.

So, you are ended up summing up all the columns for the Y first column. Same way Y a dot means for row a, sum all columns of a of 1. So, we are just summing all the columns for that. So, these are these are your columns alright. So, once you have then what you do is the next one is called as the averages ok. So, it is Y 1 dot bar, Y 2 dot bar, Y 3 dot bar all the way to Y a dot bar. So, here how do you calculate? Y 1 dot bar is equal to Y 1 dot divided by whatever is the n total n. So, Y 1 by Y 1 dot by n will give you the bar; Y a dot bar is equal to Y a dot divided by n, whatever the number of observations.

So, in this particular case in our case n is equal to, now our example a is equal to 5 n is equal to 5 in example alright. So, this is the system that we have set up for the analysis of variance ok. And most almost all of our analysis actually falls into this particular ambit ok. And I am giving you the fundamental, so that when we solve the problem you will have an clear idea of how what we are trying to do right.

(Refer Slide Time: 06:54)



So, from there once we set up this data and all those kind of things, then our next aim was I told you earlier is that we have to build an empirical model from the data.

So, the empirical model, the empirical model is a combination of controllable factors. So, in this case in ANOVA the empirical model of ANOVA, model of ANOVA is Y i j whatever is the value of Y i j is equal to mu plus T i tau i it is called as tau i plus E error epsilon i j ok where we are saying i equal to 1 2 all the way to n j equal to 1 2 all the way to n; sorry not n this is Y is equal to 1 2 all the way to a ok. So, this is i is the index on levels j is the index on replications ok. So, what we are saying here is the Y i j so this is your observation ok, is mu is a average then there is a the level specific constant; and here is the error term ok.

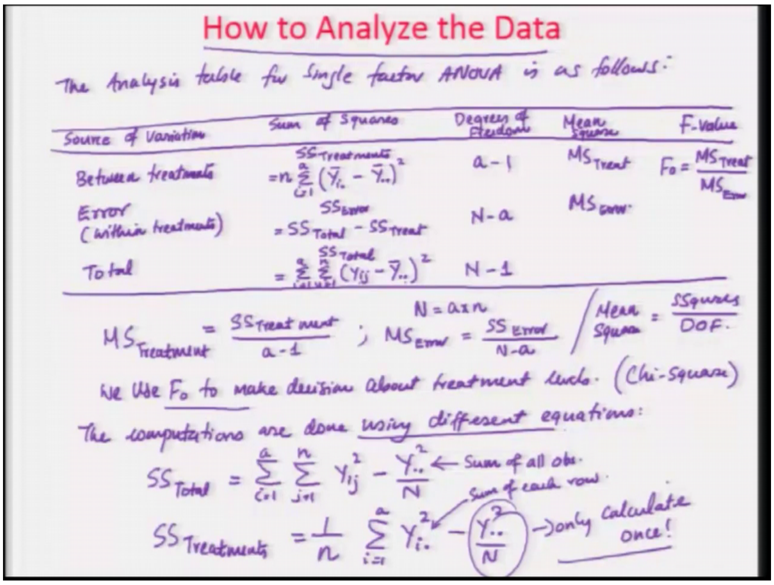
This is the global average you can think about right; so, this kind of a model. So, what we are saying is mu is common, for all treatments known as, overall mean or overall mean is kind of an average overall mean tau i tau i ok, we call it as a parameter unique to the ith treatment ith treatment is called as ith treatment effect ok. And that the last one is epsilon i j is the error parameter ok. Because of this you have the treatment effect other than this. This model is called as this is the effects model, we call this as the effects model because there is effect involve. If the model was Y i j equal to mu plus e i j this is what it was then this would have been called as the means model that means there is only the mean here is an effect ok.

A treatment specific the level specific constant is used which is called as a ith treatment effect and then that is the error on this ok. So, then you need to notice few things, this is a linear statistical model linear statistical model. So, implies what? What does this mean? The response variable, the response variable Y i j this is your response variable Y i j is a linear combination, is a linear function is a linear function of model parameters what are the model parameters mu, tau, E epsilon right means model is also used; but effects model is popular.

People want to know which effect which of this ith treatment effect is of more importance to us, and that is one of the reasons why effects model is more important rather the means model does not tell you, which is the effect that you should be known as ok. This is also known as one way or single factor ANOVA and because why do we do this because only one factor is investigated, factor is investigated you only investigate one factor in this regard right. Then such an experimental design, such an experimental design is called as completely randomized design completely randomized design ok, due to the random order; because we did the experiment in the random order alright.

So, as I told you this is the means model this is the effects model right. And I told you that the mean means model is used, but the effects model is more popular. And this kind of an experiment what we did this kind of a data collection this kind of a thing is called as the one way or single factor ANOVA because we only looking at one factor here investigating multi level multiple levels of a single factor. The factor here is the speed of the conveyor belt and this kind of an experimental design where each one of these values even you use each one of these values it is called as because each one them are random variables ok, random variables purely because this is a completed random experiment and you done it in a random order in this regard fine.

(Refer Slide Time: 14:27)



With this now the question is you are set up the whole thing you know what the model is then the next question is you have set up the entire experiment then collected the data then how do you analyze the data, what is the way to analyze the data? The analysis table, for single factor ANOVA is as follows. I follow this analysis table purely because of the fact that it actually helps you to do things in a very systematic manner.

So, we have a matrix here again for analysis ok. So, it is the first thing we will call it as source of variation where are we getting the sources of variation. And our sources of variations are one is between treatments. And which are those between treatments? These are your between treatments ok. Then it is error, error is within treatments ok. So, where is the within treatments? This is the within treatments ok. So, the columns gives you the within treatments ok then that the last one is your total alright. So, you have a matrix like this then you have something called as sum of squares.

So, the sum of squares of this is known as S S sum of squares of treatments. It is calculated by given by the equation n times; sigma i is equal to 1 to a Y i dot bar minus Y bar dot dot whole square. So, Y bar dot dot will be what ok? So, Y bar dot dot will be the average of these averages this will be your Y dot dot bar ok, so, the average of averages ok. So, that this is the Y dot dot bar. So, you some more there all the values are taken and take the grant average right. Then this is called as the S S error sum of treatments of error sum of squares of error. So, that is given by the equation S S total minus S S treatment.

So, the error is calculated by the equation sum of squares of total minus sum of squares of treatments.

Treatment I given you the equation. So, then what is the sum of squares of total ok? So the S S total is called as is given by the equation sigma i is equal to 1 to a, sigma j is equal to 1 to n; Y i j individual observation minus Y dot dot bar the whole square. So, in this case what you doing is the previous case you are only looking at the individual mean ok. So, this is the row mean, the mean of the each row this is the total mean this is individual observation this is the total mean right. And the third column is what we call as the degrees of freedom ok. So, degrees of freedom for between treatments is given as a minus 1 there are a treatments as pretty much what it is.

The degrees of freedom total is N minus 1; big N is equal to a times little n right. Then this is N minus a. So, we usually calculate between the treatments degrees of freedom, total degrees of freedom; the difference of them will actually give you whatever is the degrees of freedom for that right. Then comes the mean square right, the mean square is given by M S treat M S error ok. So, mean square treatment is given by the equation S S treatment by a minus 1 ok; then M S error. So how do you calculate the M S error? M S error again is the equal to S S error sum of squares of error divided by N minus a; is degrees of freedom.

So, the mean square is mean square is sum squares by degrees of freedom whatever the degrees of freedom that is how you actually calculate the mean square in any case right. Then the last one we need to do here is the F value, whatever is the F value. So, the Fo the observed value is MS treatment, the mean square treatment divided by mean square error. This ratio when you take this and do the ratio then you get a value called at F value right. So, then this F value is we use F0 to make decisions about treatment levels right. So, we use this value F 0 it is a critical value that we use it as, it comes from a distribution called chi square distribution ok.

Please read more about it by yourself, ok. Then by the way the computations, the computations are done using different equations ok. We do not use this equations for computations we use different equations for computation. So, the sum of squares of total S S t total is given by the equation sigma i is equal to 1 to a sigma j is equal to 1 to n Y i j square minus Y dot dot square over N; this is the computation. So, this Y dot dot is th,e this is the sum of all the sum of all observations and you square that right. And Y dot dot square is square each individual observation and then sum them up right.

Then sum of squares of treatments the computational equation is 1 over n multiplied by sigma i equal to 1 to a Y i dot square minus Y dot dot square over N. So, this is the sum of each row ok, the sum of each row, you only need to calculate this only calculate once. So, once you calculate this then you can use it in both equations. Here also it is a same exact thing that happens. So, this equation is what we actually use in our calculation when we do the manual calculations for the same. Thank you very much for your patient hearing and wish you all the best.

Thank you.