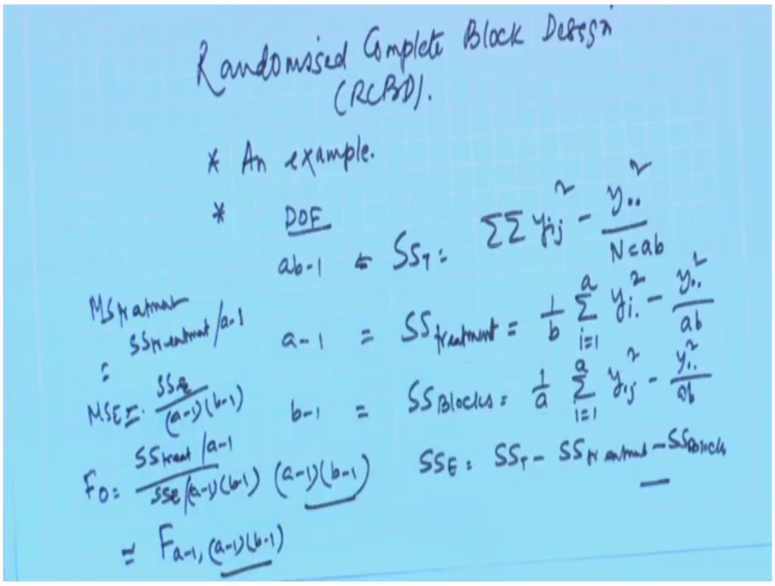
Design and Analysis of Experiments

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

Lecture – 26
Randomized Complete Block Design (RCBD): Estimation of Model Parameters

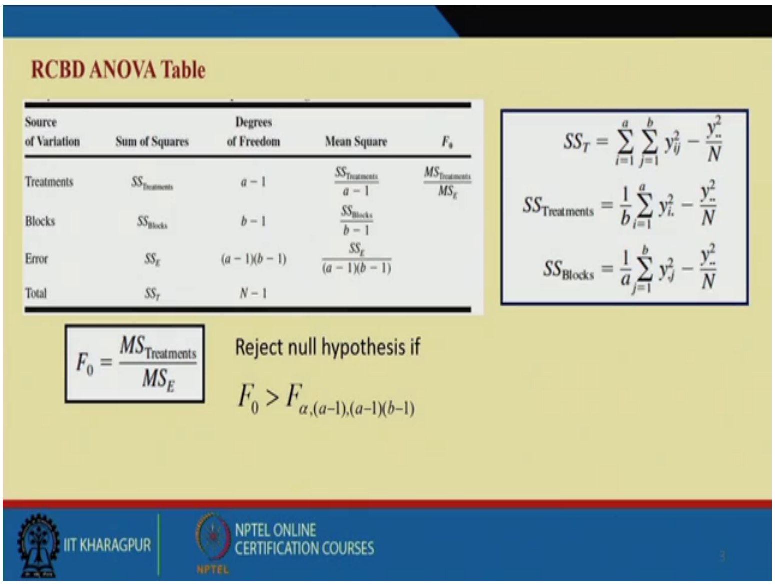
Welcome, we will continue Randomized Complete Block Design, RCBD.

(Refer Slide Time: 00:22)



In this lecture, I will first show you one example, and then how the parameters are estimated including SS T, SS E, SS treatment and SS blocks.

(Refer Slide Time: 00:54)

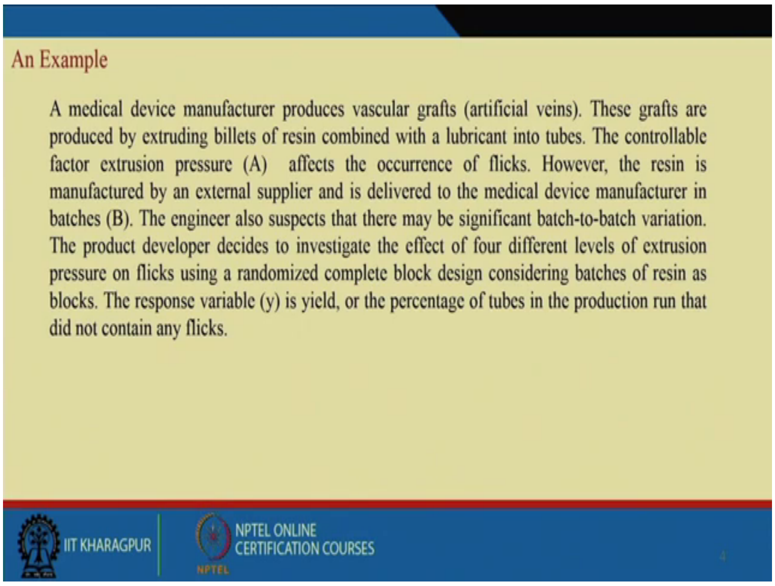


In last lecture, you have seen the ANOVA table for RCBD. We have seen that there are 3 sources of variation with one effect treatment, that is, treatment block error and then this is what is the things and the right hand sides S T equal to SS T equal to you got sum double Sum y i j square minus y double dot square by N, N equal to basically ab. Then U bar we found out SS treatment this we found out that 1 by b sum total i equal to 1 to a y i dot square minus y double dot square by ab.

Similarly, SS block we got, 1 by a sum total i equal to 1 to a y dot j square minus y double dot square by ab. And, SS E will be SS T minus SS treatment minus SS blocks and degree of freedom for this is ab minus 1 for this one is a minus 1 this is DOF, Degree Of Freedom, and this is b minus 1 and this one, a minus 1 b minus 1, we have seen this.

Also, you have seen that that MS treatment, we calculate this is SS treatment by degree of freedom and MS error is SS error by a minus 1 and b minus 1 degree of freedom, then you have seen that we have computed F 0 which is SS treatment by degree of freedom by SS error by degree of freedom; that means, MS treatment by MS error and that follows F distribution with a minus 1, a minus 1 into b minus 1 degrees of freedom.

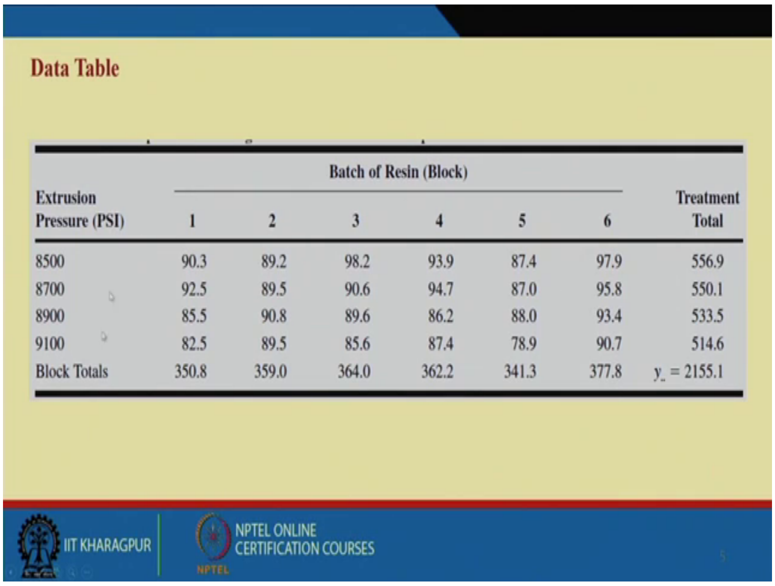
(Refer Slide Time: 03:23)



So, we will see 1 example some kind of tutorial. Here, that example is a medical device manufacturer produces vascular grafts, this is known artificial veins. These grafts are produced by excluding billet us of resin combined with a lubricant into tubes. Controllable factor is extrusion pressure that affects the occurrence of flicks, which is basically defect; however, the resin is manufactured by an external supplier and is delivered to the medical device manufacturer in batches so; that means, different batches of resin is used to manufacture this. And, these resin is supplied in blocks, in batches and the quality of that batches are beyond your control because it is coming from the supplier side that is what is the situation.

The engineers suspect that there may be significant batch to batch variation for the resin supplied. The product developer decides to investigate the effect of 4 different type of extrusion pressure on flicks, using randomized complete block design considering batches of resin as blocks. The response variable is yield that it means that is the good minus the defective products that sends and or the percentage of tubes in the production run that did not contain any defects means any flicks.

(Refer Slide Time: 05:00)

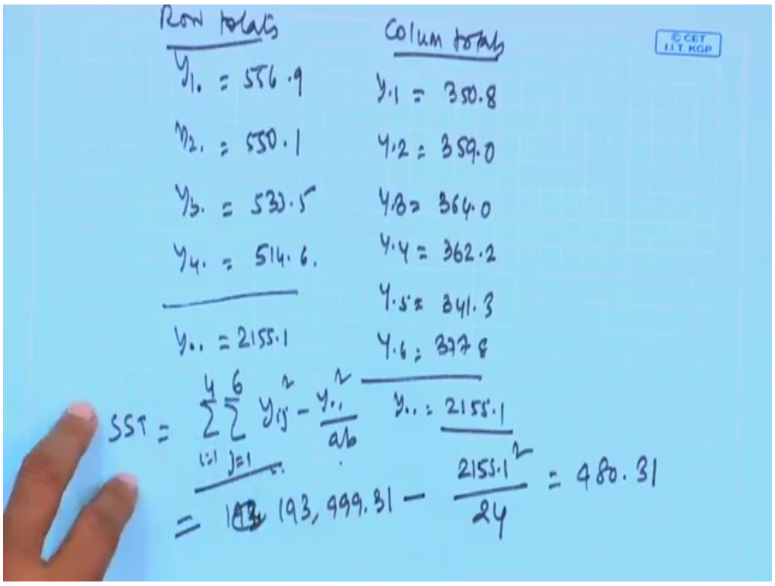


Data Table

Extrusion Pressure (PSI)	Batch of Resin (Block) 1	2	3	4	5	6	Treatment Total
8500	90.3	89.2	98.2	93.9	87.4	97.9	556.9
8700	92.5	89.5	90.6	94.7	87.0	95.8	550.1
8900	85.5	90.8	89.6	86.2	88.0	93.4	533.5
9100	82.5	89.5	85.6	87.4	78.9	90.7	514.6
Block Totals	350.8	359.0	364.0	362.2	341.3	377.8	$y_{..}$ = 2155.1

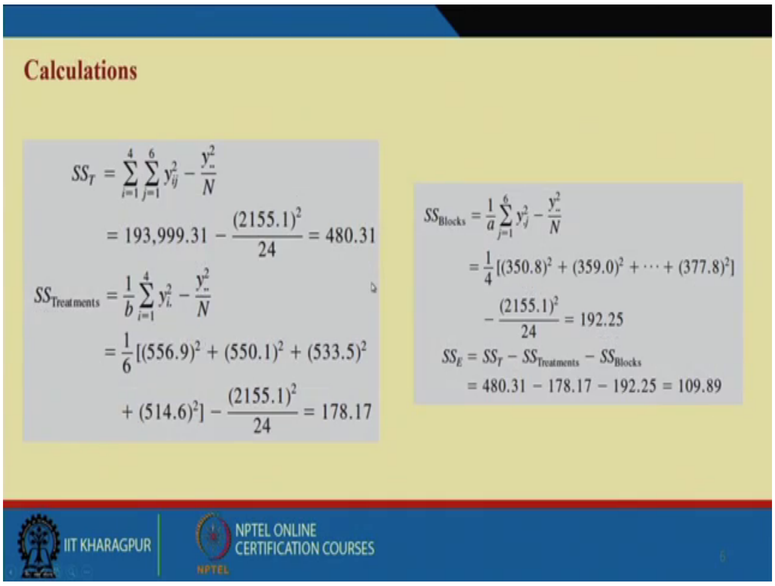
So, the data once you get that mean there are extrusion pressure is the controllable factor. It is having 4 different levels and there are 6 different blocks and you have done the experiment under each block. All the level of extrusion pressures experiment is conducted and this is what is the data on the yield value, y.

(Refer Slide Time: 05:37)



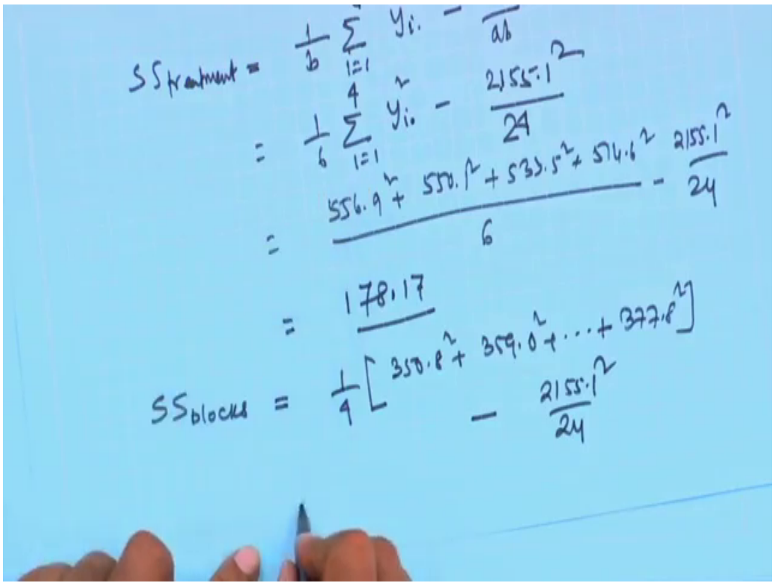
And, then you see that right hand side this is the total, so, y 1 dot is total means y 1 dot is 556.9. Similarly, y 2 dot is 550.1and y 3 dot is 533.5 and y 4 dot is 514.6, these are all row totals. Similarly, if you go by column totals you will be getting y dot 1 equal to 350.8, y dot 2 equal to 359.0, y dot 3 equal to 364.0, y dot 4 equal to 362.2, y dot 5 equal to 341.3, y dot 6 equal to 377.8 and some of these will be y double dot which is basically 2155.1 which is also some of these equal to 2155.1.

(Refer Slide Time: 06:51)



Then, you will calculate SS total. What is the formula SS total, double sum then y ij square minus y dot dot square by ab. I here i equal to j equal to 1 to 6, i equal to 1 to 4 and all those values if you combine you will ultimately this value means the square of all those 6 into 4, 24 observations, when you sum up you will get the value 193999.31 minus y double dot is 2155.1 square divided by ab is 24. So, this quantity is 480.31.

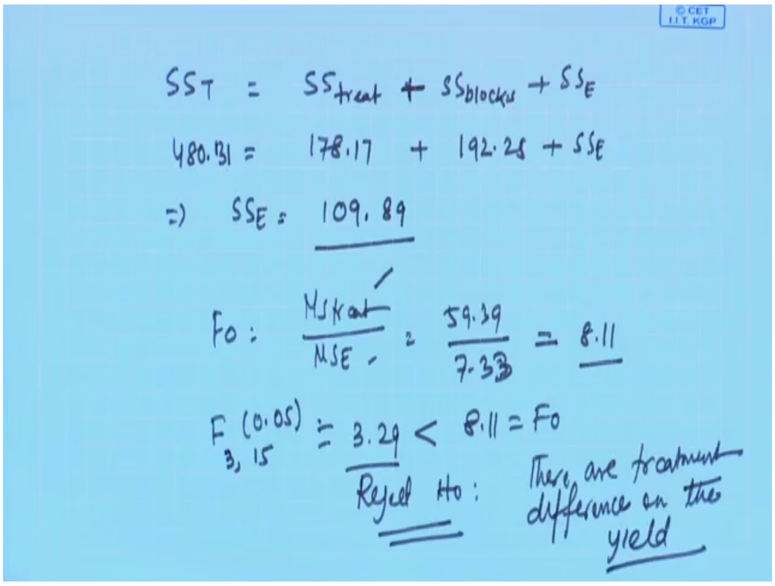
(Refer Slide Time: 07:54)



Similarly, if you want to calculate SS treatment what is this formula? The formula is 1 by b, sum total i equal to 1 to 4 here or 1 to a, y i dot square minus y double dot square by ab. In this case, 1 by 6 sum of i equal to 1 to 4, y i dot square minus this quantity 2155.1 square divided by 24.

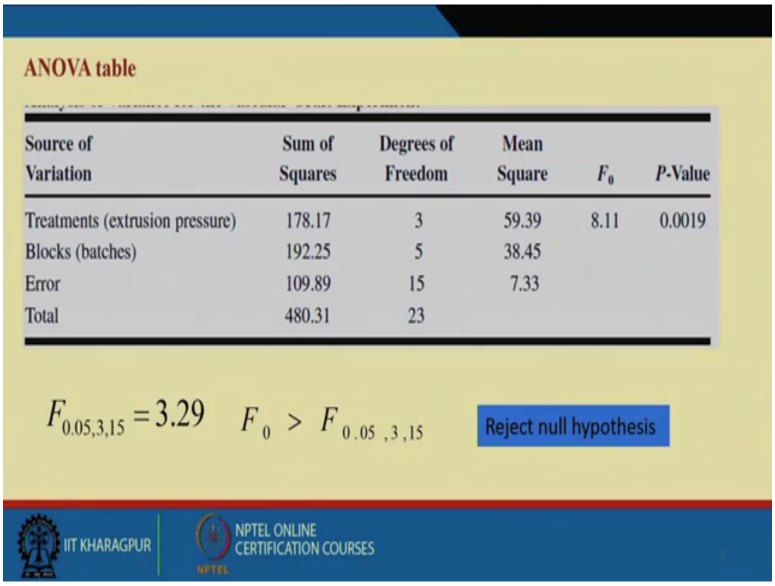
What you do here, how many totals are there? Row totals, 4 row totals are there. This is basically 556.9 square, 550.1 square plus 533.5 square plus 514.6 square divided by 6 minus 2155.1 square divided by 24 and then this quantity will become ultimately it will become 178.17. In the same manner, if you calculate SS block this will be 1 by 4, sum total of that mean the your column square that is 350.8 square plus 359.0 square, like this the 6th one is 377.8 square minus 2155.1 square by 24.

(Refer Slide Time: 09:42)



SS total is 31. SS treatment is 178.17; SS block is 192.22 then, what is SS E? From here you will get SS E equal to 109.89.

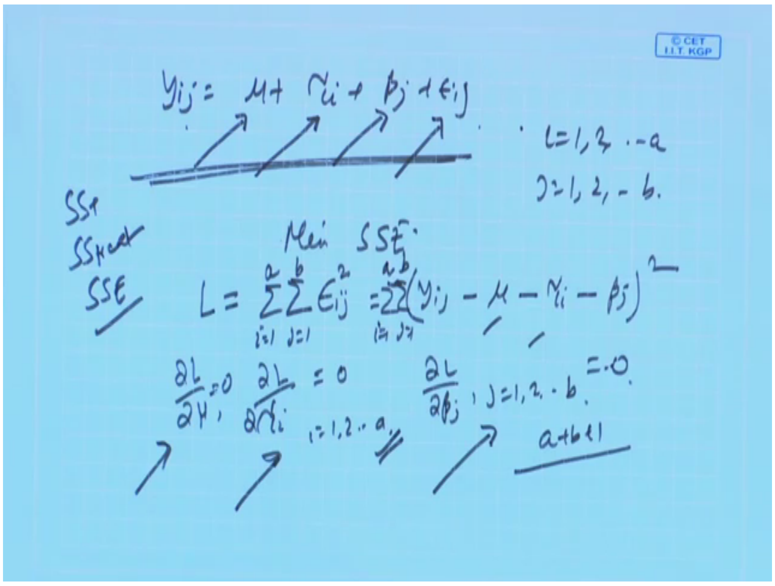
(Refer Slide Time: 10:07)



You are in a position to compute the ANOVA table. Treatment extrusion pressure effect sum square 17 this, block this, error this, total this. The degrees of freedom for treatment level minus 1, 4 minus 1, 3. Then 6 batches 6 minus 1, 5 and error is 5 into 3, 15, a minus 1 into b minus 1 and there are 24 observations then 24 minus 23. Then you are getting mean square, now F 0 is 59.39 divided by 7.33 that mean means F 0 is MS treatment by MS E 59.39 by 7.33 which will be 8.11.

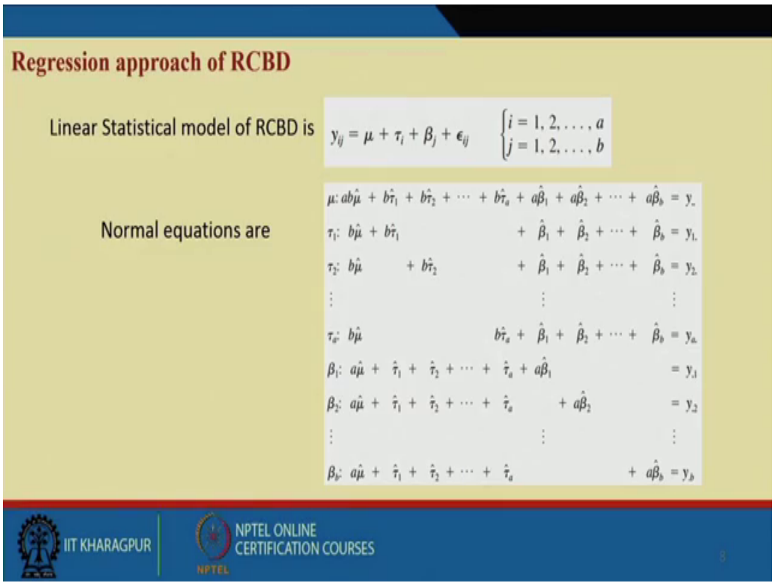
So, F 0 is MS treatment divided by MS error which is 59.39 by 7.33, this equal to 8.11. So, the if we consider 0.05 is the threshold, then it is basically 3 and 15 because this degree of 3, 15 this value is 3.29, now this value is much less than 8.11which is equal to F 0. So, reject H 0. There are treatment differences on the yield. Now, this is what so far we have computed.

(Refer Slide Time: 12:06)



Now, we have I have given you the model y ij equal to mu plus tau i plus beta j plus epsilon ij, what how to what is the estimate of this? What is the estimate of tau i? What is the estimate of beta j? What are the error estimates and all those things. In addition, that how the; that means, the computation we have shown partitioning, but from regression approach also we can find out SS T, SS treatment, then SS error. First I will see that how the parameters can be estimated using regression approach.

(Refer Slide Time: 12:48)

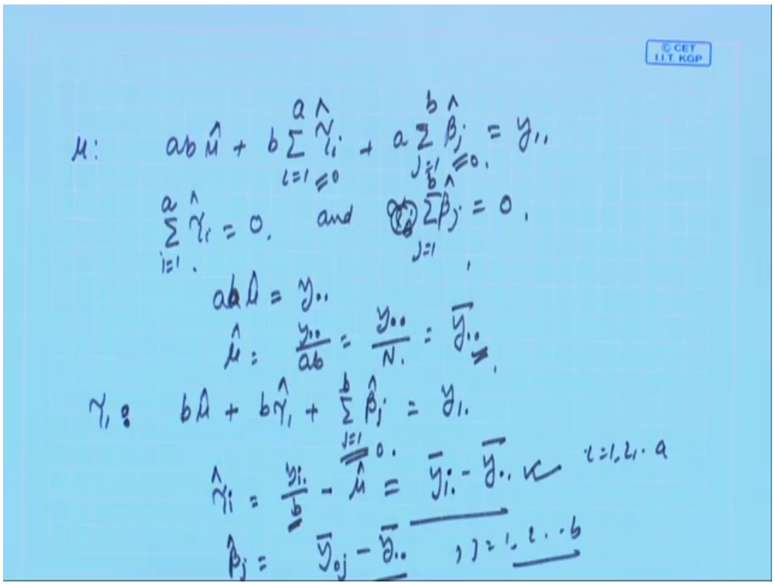


Our model is y ij equal to this and i equal to 1, 2, a and j equal to 1, 2, b. In regression approach what we do, we want to minimize SSE, that mean in this equation epsilon ij equal to y ij minus mu minus tau i minus beta j. If I take square, this will be square if I take sum, double sum, j equal to 1 to b and i equal to 1 to a, then here also double sum j equal to 1 to b, i equal to 1 to a, suppose this is the function L, which we want to minimize.

Then what you do you know the parameters, so, del L by del mu, del L by del tau i, i equal to 1 to a and del L by del beta j, j equal to 1, 2, b. If you differentiate and put them equal to 0 here 1 equation, here, a number of equations, here b number of equation. So, ab plus b plus 1 number of equation you will find out and then the equation will be like this you see.

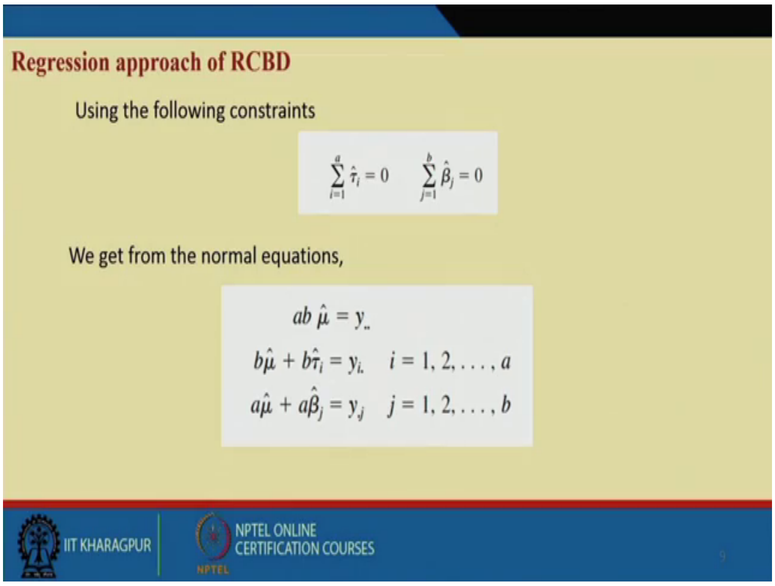
All normal equation 1 normal equation related to mu will be like this. The left hand side these are the estimated value and the sample; that means, the levels for blocks and levels for treatments it will be the grand total. Then we have a number of tau i, tau 1, tau 2, tau a. You see that everywhere that mu, mu is coming because of this b mu will coming for tau 1. The tau 1 quantity is coming and beta j all will be coming beta 1 to beta b and this will become y 1 dot that is first row total, tau 2 the left hand side accordingly changing, second total tau a single this and then all the column totals.

(Refer Slide Time: 15:21)



If you see that in the first equation ab mu cap plus if you take b here this will be, suppose, I consider the first equation with reference to mu the equation is ab mu cap plus if I take b common, then it is basically i equal to 1 to a tau i and plus if I take a common it will be j equal to 1 to b beta j. This cap these are all estimated value, this becomes y dot dot total.

(Refer Slide Time: 15:52)

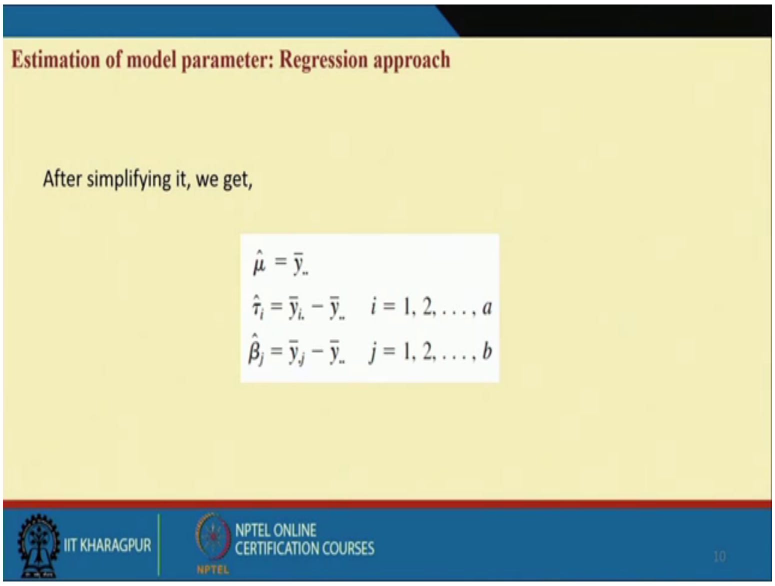


Now, we have 2 constraints because of the design, because of this design tau i, i equal to 1 to a equal to 0 and tau b sorry beta j, j equal to 1 to b cap this also become 0. If we put this become 0 and this also become 0 from this then you are getting ab mu cap equal to y dot dot. That mean mu cap is y double dot by ab which is nothing, but y double dot by N which is your grand average. What we earlier you were basically reproducing the same thing using regression approach.

Now, if you see the second equation here for tau 1, you see it is basically tau 1 related equation is b mu cap plus b tau 1 cap, then plus sum of bj cap j equal to 1 to b, this equal to y 1 dot. What about this? This is 0. If, that means, you are getting tau 1 cap equal to y 1 dot by b minus mu cap, where mu cap is y bar and this is y 1 bar. That mean y 1 dot bar minus y double dot bar. This is the row average minus grand average, that way we have seen. If I put here i, it will be i, it will be i. In the same manner we will be get beta j cap, we will get y bar dot j minus y bar dot dot.

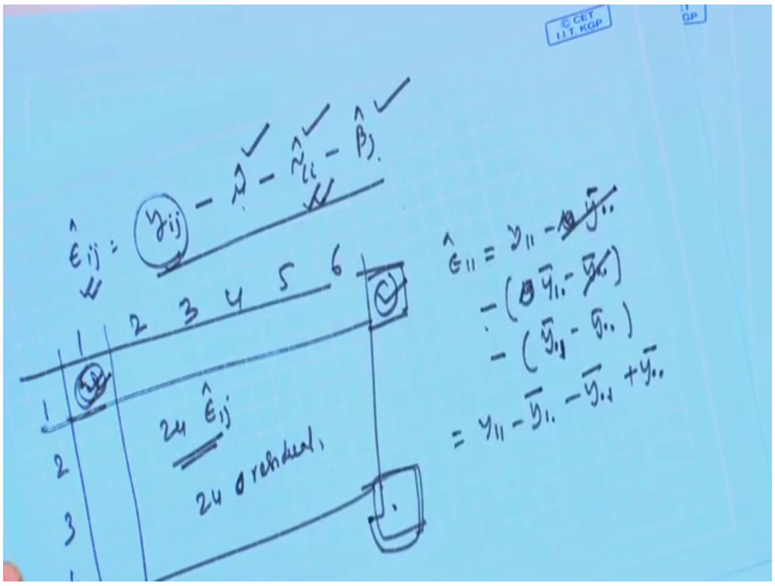
This is what you see the equations are coming. Tau i is this, i equal to 1 to a for beta j, j equal to 1 to b and if through regression approach you are also getting same equations.

(Refer Slide Time: 18:27)



Now, from the data you can find out all those things estimated values tau i, tau beta j and other things.

(Refer Slide Time: 18:33)

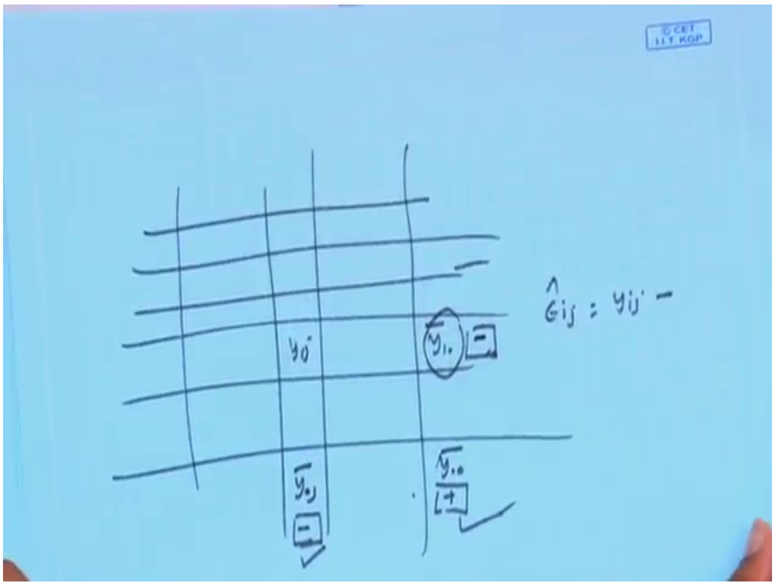


And, once beta j and beta all those things are estimated then you will you are in a position to estimate epsilon ij which is for every y ij will be subtracted by mu cap will be subtracted by tau i cap will be subtracted by beta j cap.

In this example, what happen we have 4 extrusion pressure and 6 numbers of blocks. We have 24 observations. We will be getting 24 epsilon ij cap which is known as 24 residuals because, you have already computed this, you have computed this and you know this value. That mean every cell value will be subtracted by the corresponding the grand mean then the row effect and the column effect, column effect row affect, this.

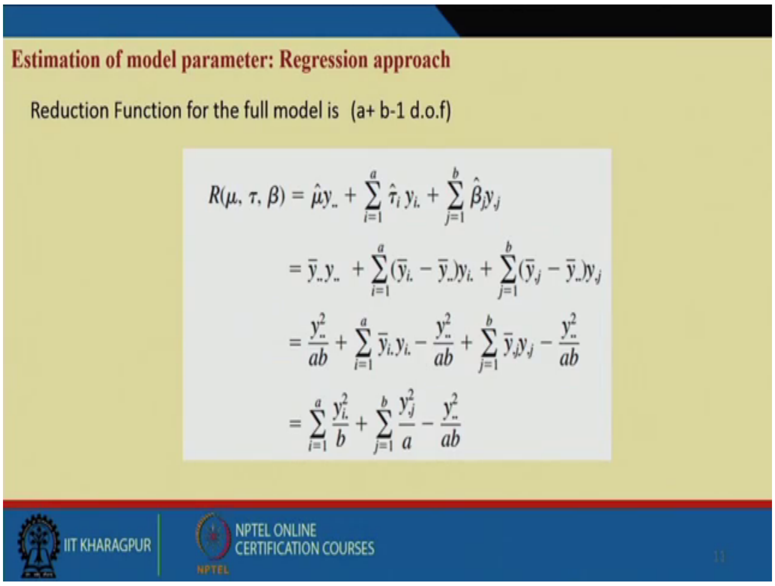
Now, you can do that row effect is what, that mean if I want to know what is the epsilon 1 1 here then this will be y 1 1 which is already there minus mu cap is y double dot bar. So, minus tau i is nothing, but y 1 dot bar minus y dot bar. Minus y dot 1 bar minus y dot bar this is what for beta j. Then this is y 1 1. This quantity minus plus, this will be cancel out. Now, this is minus y 1 dot bar minus y dot 1 bar plus y dot dot bar. Having this quantity will become error; residual will be computed y ij this value then minus this value minus this value.

(Refer Slide Time: 20:40)



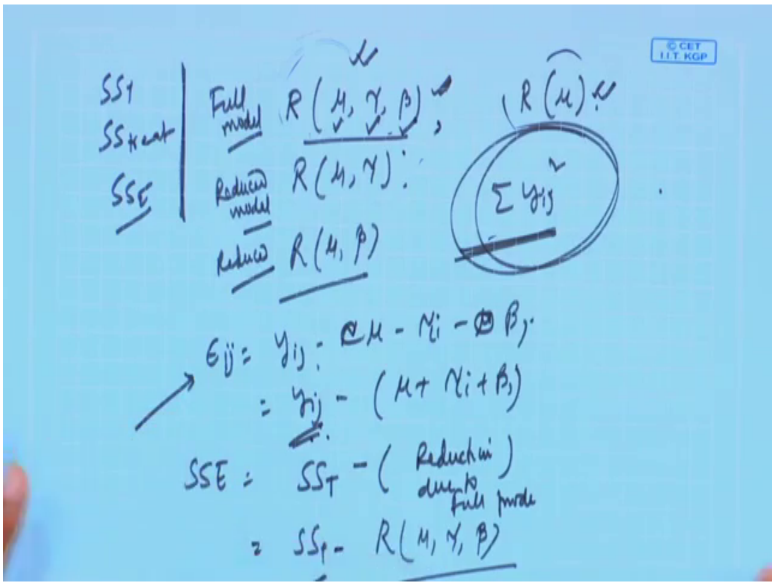
I can write like this, a general one I can write. Suppose, here it is y ij suppose and this is your y i dot bar this 1 is y j dot bar and this is y double dot bar then epsilon ij cap is this one this will be minus, this also will become minus, this plus. y ij minus this minus this minus this plus this. So, that sense all residuals will be computed this is what the computation part form is, it is easy one. In assignment, you will be doing it and then you will learn more.

(Refer Slide Time: 21:30)



Now, I will show you the actually we have computed SS T, SS treatment and SS error.

(Refer Slide Time: 21:37)



Here, I will show you the concept of from regression concept how it is coming. We will use the concept called that full model and reduced model concept. What is the variability that can be captured by full model; the full model means all parameters are considered.

Suppose, if we do not consider the treatment and blocks and all those things then only mu is there that is a reduced model. We may consider that mu and tau this is another reduced model, so, this is full model, this is reduced model and another one will be mu beta another reduced model.

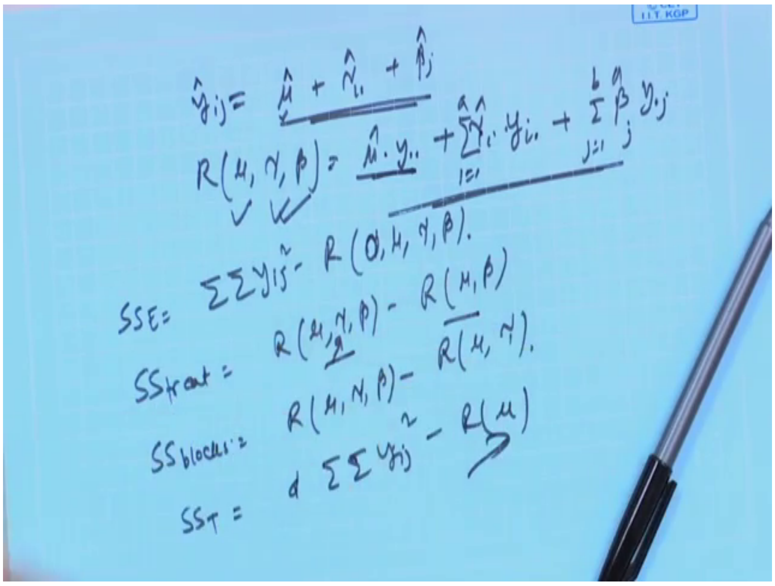
Now, if you do not consider any model then we, will see that ultimately we thought of that sum of y ij square this is what is the total some observation square is there and then if we use a model then there sum square will be from the full model will be with respect to this, with respect to this, with respect to this and if we subtract the row sum square minus the reduced model that what sum square that is having then this difference will give you the error part; like we know that epsilon ij equal to will be y ij minus mu minus tau i minus beta j. That mean y ij minus mu plus tau i plus beta j.

What I am saying that the sum square here is SS E, this sum square here will give a some kind of SS T and this one will basically that reduction due to full model reduction part due to full model. This can be denoted like this. This is basically this will be equal to that one R mu tau and beta.

Here, I am saying that with reference to y ij only not y ij minus y bar or mu. That sense it is there, but if you subtract this mean this one, then you will get the other kind that formula that what we have discussed earlier, some y ij square minus y dot dot square by N will come because you will consider this. This part we have discussed earlier also, but approach is this.

Then what happened, this is what the full model reduction in reduction of variability is.

(Refer Slide Time: 25:05)

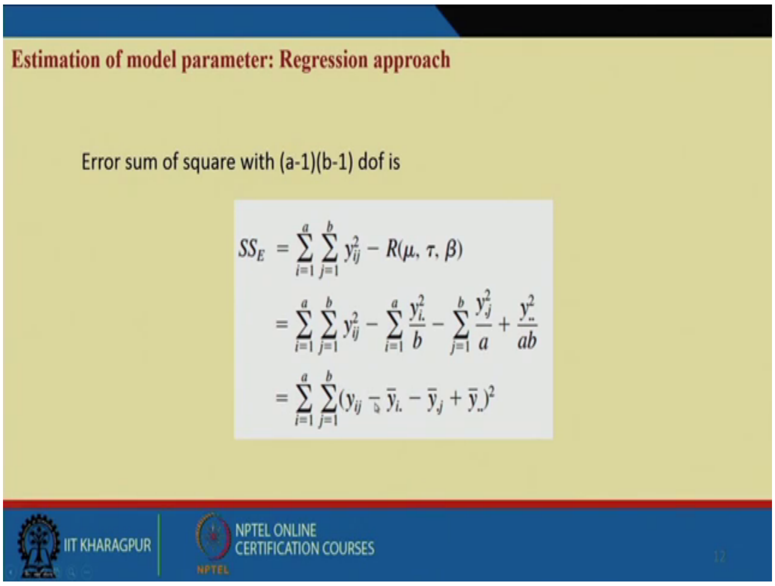


What we are doing, we are using the model using the model estimate like y ij you are using it is mu cap plus suppose, y ij cap I one then this is is equal to mu cap plus tau i cap plus beta j cap, this is what is coming.

And, then we are when we are writing like this R mu tau and beta we are multiplying these mu cap with the total for that for that parameter. Similarly, how many tau i we have? We have some tau i means we have tau i, i equal to 1 to a that you multiplied by the corresponding total and then plus beta j cap with corresponding total j equal to 1 to b. This will give you the reduction in variability.

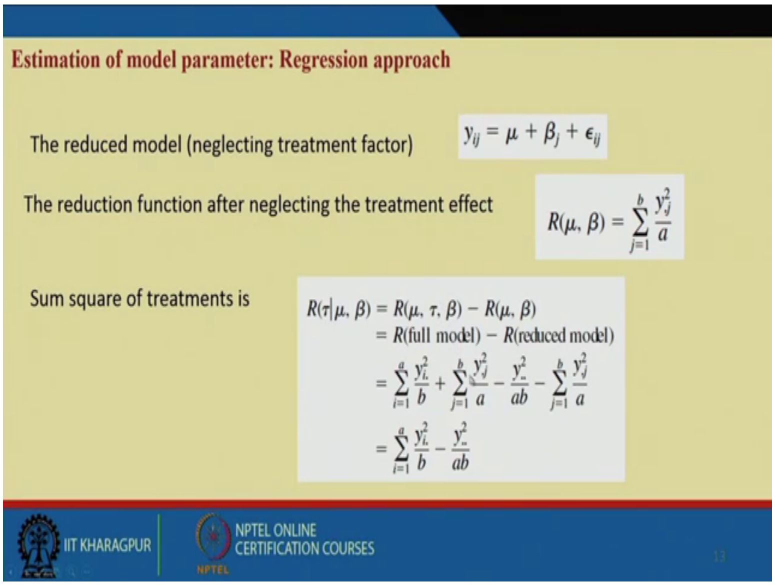
Now, see the slides. The estimate of mu cap is y bar this 1 is y i dot bar minus this and beta j is y dot j bar minus this. If you do algebraic manipulation all those things you are finally, getting this one that mean i equal to 1 to a this, j equal to this.

(Refer Slide Time: 26:39)



Now, if you subtract this quantity from the sum square of the observations that mean this minus this sum square of the observations minus sum square because of the model, the difference will be SS E. You are getting this equation and actually earlier also you have seen this equation earlier also you have seen this equation.

(Refer Slide Time: 27:17)



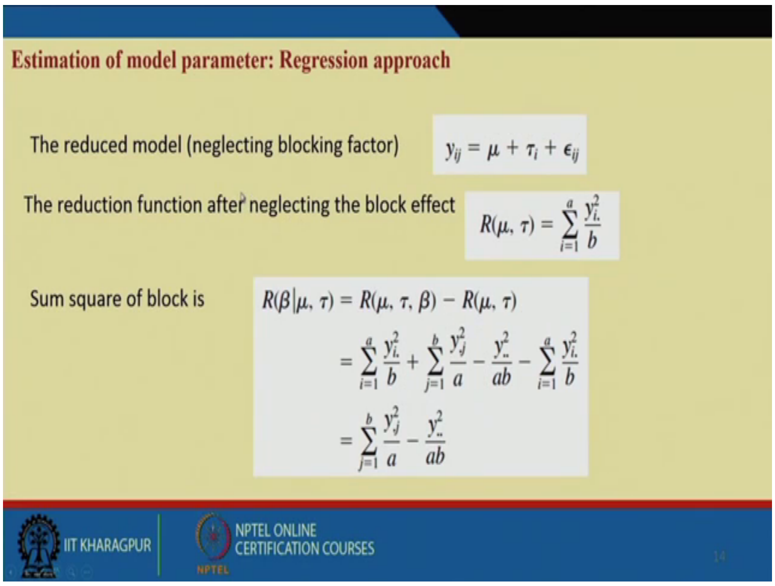
Now, if I want to calculate that what is SS treatment, then what we will do, we will simply find out, sorry. We want to find out then we will find out that this is the full model now here what happen we will find out that if we keep mu and beta instead of tau then this is the reduce model without tau and the corresponding reduction of variability will be this.

Then what is the contribution of the treatment? The treatment combination will be the full model minus the reduced model. This full model treatment combination you have seen earlier this is the full model treatment, full model reduction is this and then minus this is this quantity this is nothing, but this and this is what we have seen that when you are computing SS treatment we are writing y ij square by b minus y dot dot square by n, n equal to ab.

So, that mean if you want to know SS E then what happen you require to know some square of the observation minus R of tau, mu, tau and beta. If you want to know SS treatment then it is nothing, but R tau, beta minus R mu and beta because this portion is eliminated from here. Similarly, if you want to know treatment blocks then you find out R mu, tau beta R mu minus R mu tau. If you want to know the SS T basis then you write down that what is the y ij square total minus R mu how much is coming from mu anyhow.

So all those things from this regression concept approach from regression approach we can estimate. You see what happen here reduce model is this you are getting the sum square treatment.

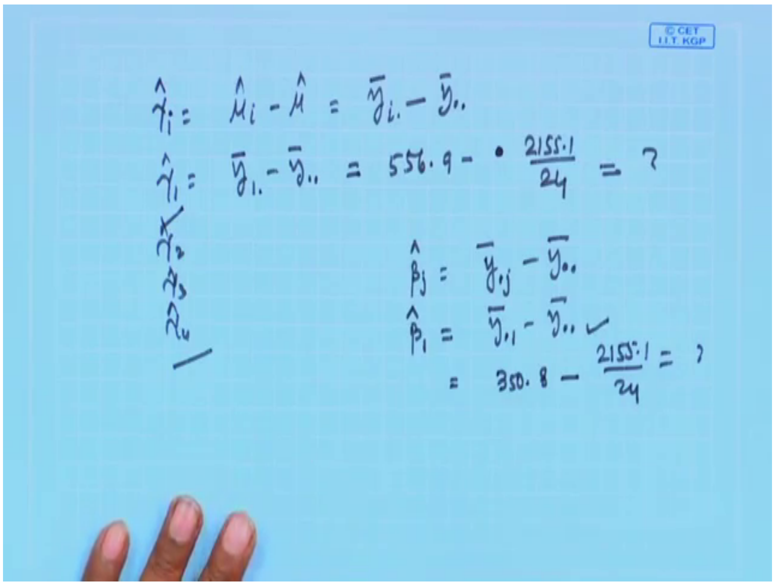
(Refer Slide Time: 29:30).



If I want to get the sum square blocks your model is without block what is the thing and then what is the reduction in variability without block is this and then this minus this, is this. Similarly, you will get SS T also.

I hope that you have understood that mean the estimation of parameter in this particular case RCBD and I will go back to the numeric example just to give you some idea how the numerically it will be computed.

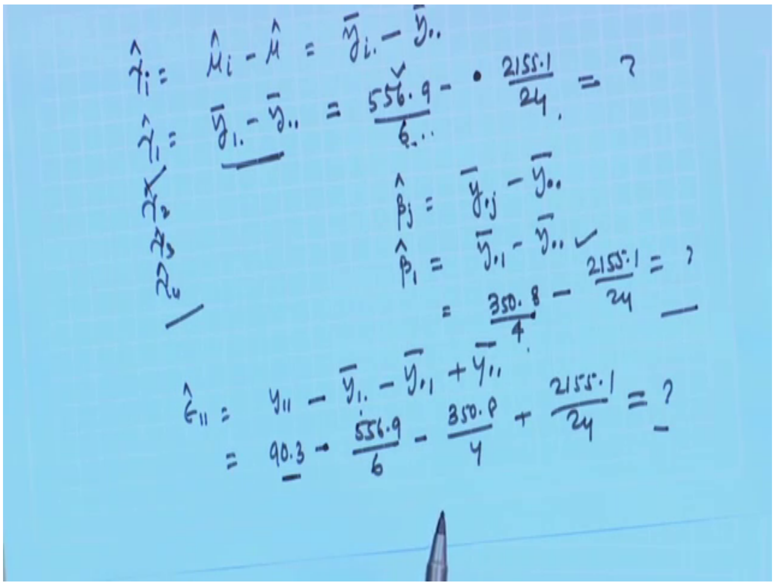
(Refer Slide Time: 30:32)



Here, you see suppose you want to know what is you will be able to compute tau 1 cap. So, I know tau i cap is mu i cap minus mu cap which is nothing, but y i dot bar minus y double dot bar. If I say what is tau 1 cap then, you will say y 1 dot bar minus y dot dot bar this is nothing, but 556.9 minus, what is the grand mean, grand average? Just let me check, we have not computed here, but it will be 2155.1 by 24, this will give you value of tau 1. Similarly, you will get value of tau 2, tau 3, tau 4.

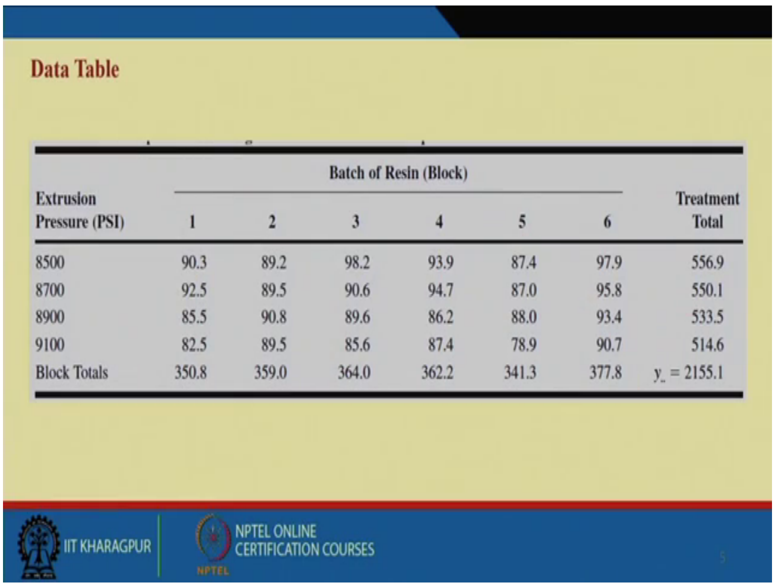
If you want to know beta j cap then this is nothing, but y dot j bar minus y double dot bar. For example, I want to know the beta 1 cap although though although beta 1 cap then it is y dot 1 bar minus y double dot bar which is y 1 dot bar is 350.8 minus y double dot bar is 2155.1by 24, find out this value. Same manner other beta will be able to be.

(Refer Slide Time: 32:20)



Suppose, you want to know the error component E 1 1 bar for the particular cell then this will be y 1 1 minus y 1 dot bar minus y dot 1 bar plus y dot dot bar. y 1 1 is 90.3 minus y 1 dot bar is. Here one mistake we have done 556 is the total divided by 6. Similarly, here this is basically the total divided 4; this is total value not average value. These are the average value. This by 6 minus this, this by 6 minus this, 90.33 minus y 1 bar, y 1 dot bar is 556.9 divided by 6 minus the column average 350.8 by 4 plus 2155.1 by 24 this quantity find out later will be the error.

(Refer Slide Time: 33:35).



Extrusion Pressure (PSI)	Batch of Resin (Block) 1	2	3	4	5	6	Treatment Total
8500	90.3	89.2	98.2	93.9	87.4	97.9	556.9
8700	92.5	89.5	90.6	94.7	87.0	95.8	550.1
8900	85.5	90.8	89.6	86.2	88.0	93.4	533.5
9100	82.5	89.5	85.6	87.4	78.9	90.7	514.6
Block Totals	350.8	359.0	364.0	362.2	341.3	377.8	$y_{..} = 2155.1$

So, using this table go to slide, using this table you will be able to find out if you compute average here across rows, across columns; then grand average, then just subtract the row average with grand average, you will be getting the level effect treatment effects for different levels. If you do it for the blocks also that effect you will be getting.

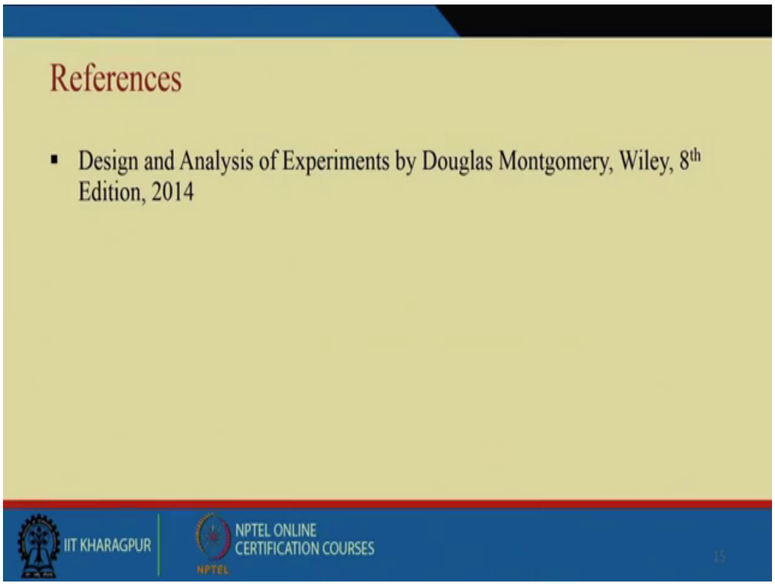
Now, when you are interested to find out the errors using this model is this that cell value minus this corresponding row average, minus corresponding column average plus the grand average. In this way the all the parameters can be estimated. 1 is the parameter estimated, the next issue is that the (Refer Time: 34:15) the errors are residuals are estimated. You have to go for residual analysis for what purpose for test of assumptions.

What are the tests of assumption? First one is normality, second one is your homoscedasticity, and third one is independence, uncorrelated errors. All those things the way we have discussed in one way analysis of variance, in the same manner you can do it. You will go for normal probability plot, control quantity control plot of the errors, you will see that whether the errors are normal distributed or not.

Second 1 homoscedasticity predicted value versus residual values plot. If there is the funnelling right or funnelling left that kind of effect is there then the heterogeneous variability. That is the another issue and for independence you have to know the order of experimentation. X axis you write down the order of experimentation from 1 to that N capital N, y axis you give the error values residual values and you see that whether they are random or not. If they are random then that is independent, otherwise you can go for go for Durbin-Watson test for errors, that is, taking one leg at a time you can find out DW, Durbin-Watson and also the correlation coefficient there and from there you see that whether the correlation coefficient give large or not and accordingly, you will be able to do the residual.

For RCBD residual analysis we will not do further, this is the same way in CRD whatever residual analysis we have done using one way ANOVA the same way you will be doing and once you have residuals calculated, residual analysis event.

(Refer Slide Time: 36:39)



The thing are basically taken from Montgomery book chapter – 4.

Thank you very much.