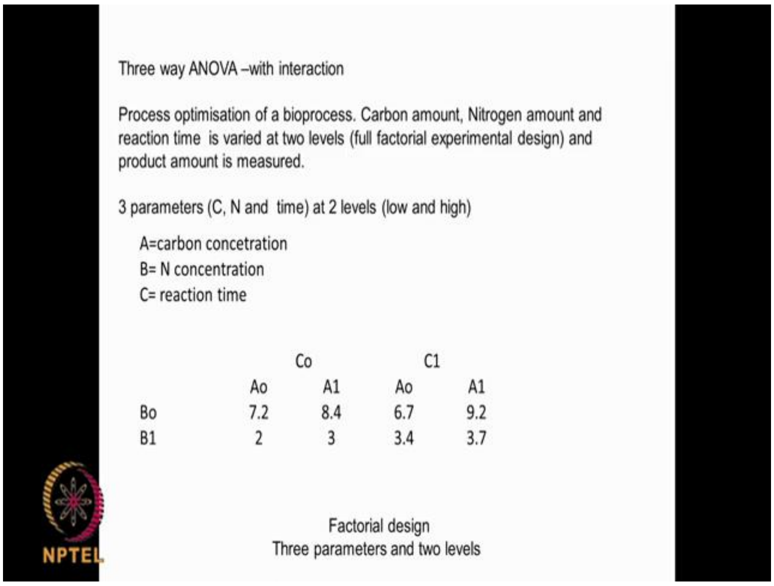
**Biostatistics and Design of Experiments**
**Prof. Mukesh Doble**
**Department of Biotechnology**
**Indian Institute of Technology, Madras**

**Lecture - 20**
**ANOVA**

Welcome to the course on Biostatistics and Design of Experiments. We will continue on ANOVA. We will look at one more problem, this is a three-way ANOVA.

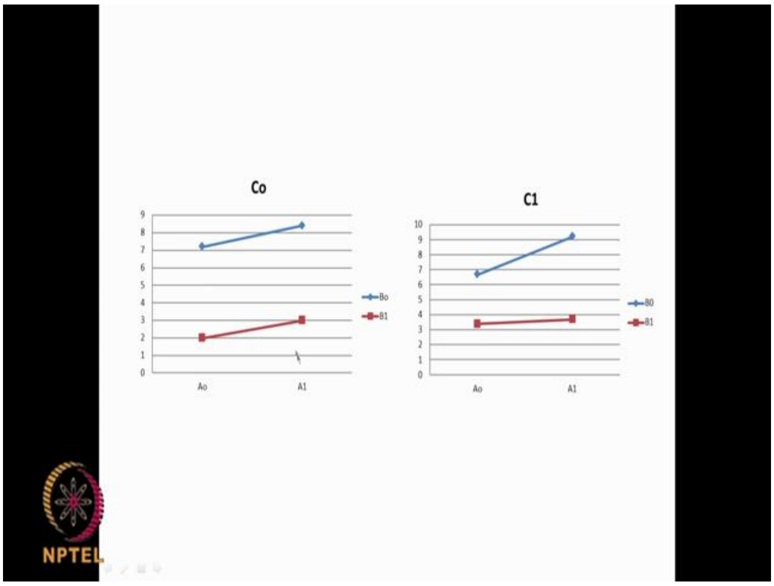
(Refer Slide Time: 00:23)




And there is a possibility of interaction also in this three way ANOVA, but the number of experiments done is not sufficient actually. Imagine, I am doing a process optimization of a bioprocess, this is very common. We do lot of process optimization where we modify carbon concentration or even the carbon source, nitrogen, then type of organism, agitation, temperature, micro nutrients, pH and so on. So, imagine a problem it is a three way ANOVA. That means, I have three independent variables or three mean effects, as we call it; the carbon amount, the nitrogen amount and the reaction time. These are the three variables or parameters or f h and modifying and I am measuring the final product yield, this is called a full factorial experimental design at 2 levels. That means, I am looking at say, carbon concentration at one level and another level; that means, it could be say 5 % carbon and this could be 8 % carbon. But I have put it in a coded form, normally when you plan the design, we always put them in a coded form.

We will talk about it later, the coding and all that. But, so at 2 levels of carbon concentration, in Ao and A1, at 2 levels of the nitrogen concentration Bo and B1; at 2 levels of reaction time Co and C1. So, if I have 3 parameters and I am changing each parameter in 2 levels; that means, 2 X 2 X 2 that is 8 experiments. So, I have done 8 experiments, I have not replicated anything. But, I have done 8 experiments and each one at 2 levels and this is called full factorial experimental design. 2 X 2 X 2; that means, $2^3$. So, I have full factorial experimental design in this particular situation. So, 3 parameters and 2 levels, A could be carbon concentration, B could be nitrogen concentration, C is the reaction time. So, I am getting some product yield here, mostly in grams per liter, this is the amount. Now, I want to look at it, look at the main effects and see whether I can even get interaction effects ,that is going to be big problematic, we will see why, as we go along.

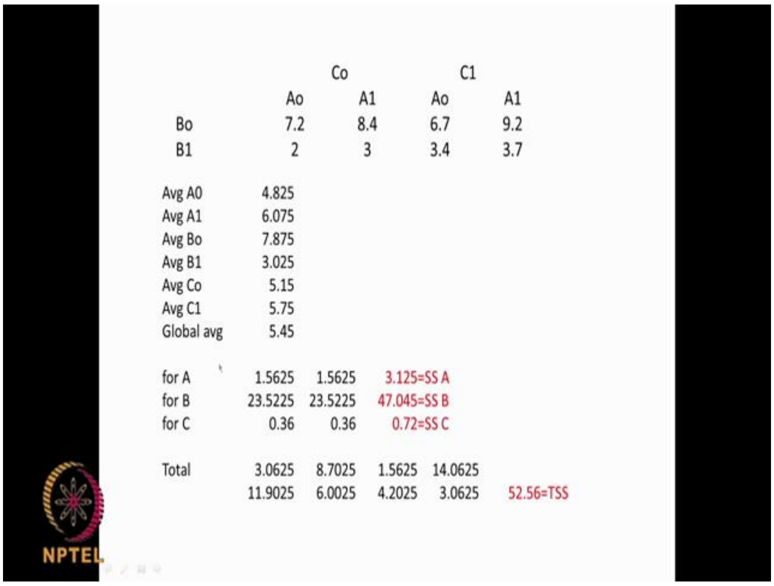
(Refer Slide Time: 02:41)




So, pictorially how do they look like? So, Ao, A1 like this; then I am changing Bo, Bo here and B1 here, at one particular concentration value of C naught that is reaction time. So, at one reaction time, Ao changes like this, Ao to A1 and this is for another B. Now at another C1, we have another picture like this actually, like this, like Ao A1 like this and this is for Ao A1 at B1 at this is at B1 at one value of B and that is at another value of B, that is Bo and B 1.

So, we have it is always good to draw up figures because sometimes figures always show as things like interaction and how their performance is pictorially. So, we have 8 experiments, so we can do an ANOVA, we can get a total sum of squares, we can get sum of squares for each of the mean effects. We can get sum of squares for A, we can get a sum of squares for B, we can get sum of squares for C. Then we can get total sum of squares also. Let us go further as usual.

(Refer Slide Time: 03:58)



| | Co | | C1 | |
|---|---|---|---|---|
| | Ao | A1 | Ao | A1 |
| Bo | 7.2 | 8.4 | 6.7 | 9.2 |
| B1 | 2 | 3 | 3.4 | 3.7 |

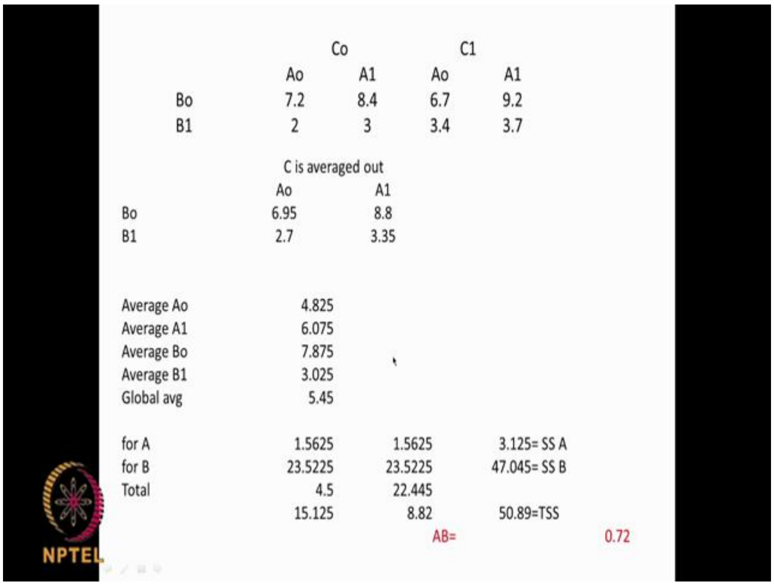
| | |
|---|---|
| Avg A0 | 4.825 |
| Avg A1 | 6.075 |
| Avg Bo | 7.875 |
| Avg B1 | 3.025 |
| Avg Co | 5.15 |
| Avg C1 | 5.75 |
| Global avg | 5.45 |

| | | | | | |
|---|---|---|---|---|---|
| for A | 1.5625 | 1.5625 | 3.125=SS A | | |
| for B | 23.5225 | 23.5225 | 47.045=SS B | | |
| for C | 0.36 | 0.36 | 0.72=SS C | | |
| Total | 3.0625 | 8.7025 | 1.5625 | 14.0625 | |
| | 11.9025 | 6.0025 | 4.2025 | 3.0625 | 52.56=TSS |

So, again I am showing you this picture; so average Ao will be this, plus this, plus this, plus this, divided by 4. Average A1 will be this, plus this, plus this, plus this divided by 4. Average Bo will be this, this, this, this added, divided by 4. Average B1 will be this, plus this, plus this, plus this, divided by 4. Average Co will be, add all these 4/ 4. Average C1 will be, add all these 4, ÷ 4. Now, the global average could be average of any one of these 2 or these 2 or these 2. I should get the same answer. So, the global average is 5.45, very simple. Now, let us look at the sum of squares for A. How do you do? you know how to do that, right? So, there will be 2, there are 2 levels for A or 2 groups for A, so you will get two sum of squares (Refer Time: 04:54)

.

(Refer Slide Time: 04:55)




So we have 2 levels for A, so how do we get this for A? What we do is, we take this 4.8 - 5.4, square it up and then multiply by 4. Why 4? We have here, 4 values for A and then how do you get this? What we do is, average of A1 is here, global average is here subtract one from another; square it up, multiply by 4.

Now, how do you do for B? Bo 7.87 - 5.45, square it up; then multiply by 4, then for the other one, 3.02 - 5.45, square it up, multiply by 4. Now, let us look for C, 5.15 - 5.45, square it up, multiply by 4 and then 5.75 - 5.45, multiply by, square it up, then multiply by 4. So, why did I say 4? Because for Ao, we have 4 terms, for A1 we have 4 terms, for Bo we have 4 terms, for B1 we have 4 terms, for Co, we have 4 terms and C1 also 4 terms. When you add each these two, we will get sum of squares for A, sum of squares for B, sum of squares for C, straight forward. Now, how do get the total sum of squares? Each of the term we subtract from 5.45, square it up, so this, $7.2 - 5.45^{2}$ and so on. So, we will get 8 terms here. We add all of them, we will end up having total sum of squares.

So, now, we can assume the error is equal to total sum of squares - SS - SS B and SS C that is one way. If I am interested in looking at interaction, imagine I am interested in looking at interactions. So, what do I do? I can look two at a time, because here I have three. So, I can look two at a time; that means, first I look at AB so, I can take an average of Co and C 1 and put it here.

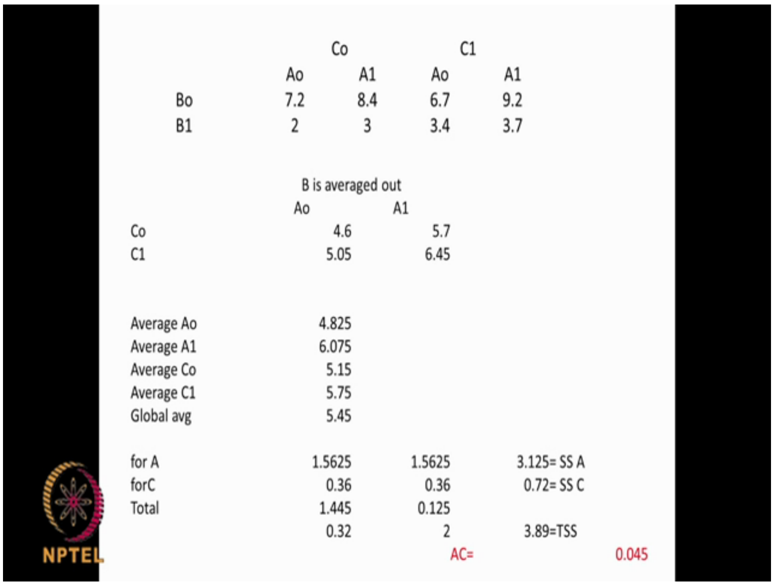
So, I can look two at a time. Let us look at again this is my, the full factorial design experiments. So, what I do? I take C is averaged out 7.2 + 6.7 / 2. 8.4 + 9.2 / 2. 2 + 3.4 / 2. 3 + 3.7 by 2. So, I have now a two effect problem, where C is averaged out, understand? So, two effect problem, so again, I can do a sum of squares, so average of Ao, will be 6, this + this / 2, for average of, sorry, average of Ao this + this / 2, average of A1 this plus this / 2 and of course if you go to the previous one, you should be getting the same 4.825 and 6.075. So, 4.825 so you should get the same thing.

So, average of Bo will be this plus this divided by 2, average of B1 this plus this divided by 2. So, we got the averages for these and then global average will be again you should 5.45, when you add up all these 4 items divided by 2. Now, we can do a sum of squares for A, sum of squares for B and then total sum of squares. For A, What do you do? $4.825 - 5.45^2$, $6.075 - 5.45^2$, and you have to multiply by 4 each term, do not forget that. Because we have here 2 terms and getting this again taken from average of 2 terms so 4. So sum of squares for A is 3.125. So, if you look at the previous one, sum of squares for A, in fact it will come out to be the same. And same thing for B, we will get 47.045. How do you get this term? This term; this minus this, square, multiply by 4, this minus this, square, multiply by 4. So, we get B here, here there is no C; because we have taken averaged out, but your total sum of squares will be different from the previous case. Total sum of squares will be different from previous case, here 52.56, but here you will get different, because the terms here are different here. How do you get that total sum of squares? There are 4 terms here, we take the $5.45 - 6.95^2$ X 2, then we take $5.45 -2.7^2$, X 2. Then we take $5.45 - 8.8^2$ X 2. Then we take $5.45 - 3.35^2$ X 2.

So, you will get 4 terms, you add all of them to get the total sum of squares. So, you can subtract the sum of squares of A, sum of squares of B from this 50.89. That will be the error or we can even say that is the interaction between A and B. So, we can call this as AB sum of squares; that is an interaction between A and B, which comes out to be 0.72. So, interestingly if you look at it, the degrees of freedom for A will be 1, because we have two levels for A. Degrees of freedom for B will be, 1; because we have 2 levels for B and the total degrees of freedom here is 3.

So, AB will have 1 degree of freedom. So, either you call it error or you call it AB. Only one you can call it, because we cannot call both AB, as well as error because, you do not have sufficient degrees of freedom here. So, we call this as AB, which is 0.72. From the previous figure; in the previous table, we got the sum of squares for A, sum of squares for B, sum of squares of C, total sum of squares and then from the next one we got sum of squares for AB here. Now same thing we can do for other two variables.

(Refer Slide Time: 11:51)



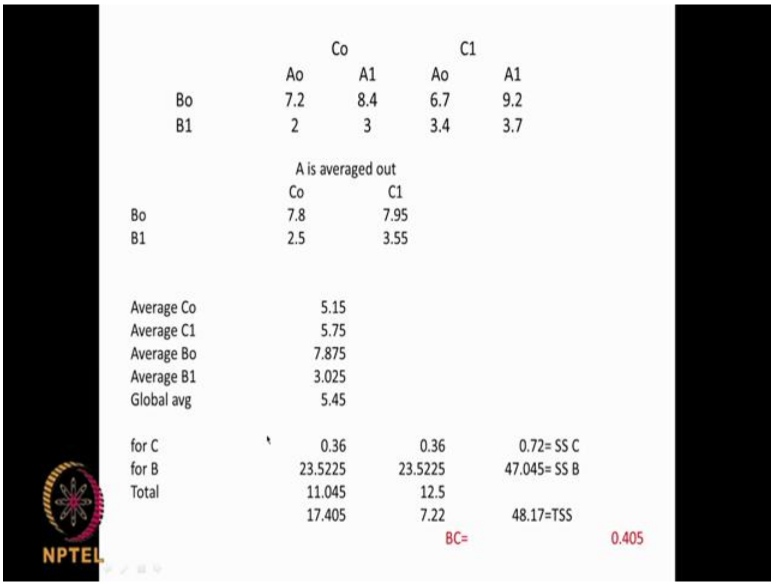

Here we can take B as averaged out, so what do we do? We will have Ao, A1 in one column and Co and C1 in other one. So, how do you? How do you do B averaging out? So, that it will come out to be 7.2 + 2 / 2, 8.4 + 3 / 2, 6.7 + 3.4 / 2, 9.2 + 3.7 ÷2. So, B is now averaged out so we have a table for Ao, A1 here and Co, C1 just like the previous one here we have Ao,A1 here and Co,C1 here. So, again we can do the same thing, we can get average of Ao, that is, this plus this by two. Average of A1, this plus this by 2. Average of A1, this plus this by 2, then average of Co, this plus this by 2, Average of C1, this plus this by 2 and global average will be average of the entire lot five point.

Then, now we can get the sum of squares, we can use this or we can take it from the original mean effect table also. But, so how do you get this particular term? We take $4.825 - 5.45^2$ X 4. How do you get this term? $6.075 - 5.45^2$ X 4. How do you get this term? $5.15 - 5.45^2$ X 4. How do you get this term? $5.75 - 4.5^2$ X 4.

So, sum of squares for A is 3.125, sum of squares for C is 0.72. Now, this will be the same as the original 0.72. You should be getting the same answer actually, 0.72, 0.72, sorry this 0.72. So, the total will come out to be different of course. So, how do we get the total? We have to subtract each one of these item do not go here, but go here. The reduced table with 2 table with 2 parameter table. So, $4.6 - 5.45^2$, $5.05 - 5.45^2$. But, please remember that, here you have to multiply by 2, because each one of this term is made up of 2. Then, $5.7 - 5.45^2 \times 2$, $6.45 - 5.45$ so when we do that and add up you will get this as your total sum of squares for this reduced table, for the reduced table.

We can do total sum of squares - sum of squares due to C, sum of squares due to A and either call it as a error or we can call it as a sum of squares for the AC. Sum of squares for AC; that is A, C interaction. And as I said before, A will have 1 degree of freedom because A has two levels, C will have 1 degree of freedom, because C has two levels, so AC will have 1 degree of freedom and there are 4 experiments so 3 degrees of freedom, you are showing here 3 degrees of freedom means if the mean, fx take 1 1, so either you can have error or you can have a interaction AC. We cannot have both, we cannot have both error as well as interaction. So, this is the interaction you see, then we go to the next one where we have BC and A is averaged out.

(Refer Slide Time: 16:11)




So, we have BC and A is averaged out, so how do you average out the A? We have this table now, we what we do is, A gets averaged out, so we combine, add these two, divide by 2, add these two, divide by 2, 2 + 3 / 2 is 2.5, 7.2 + 8.4 / 2 is this, 7.8, 6.7 + 9.2 / 2 is 7.95, 3.4 + 3.7 / 2 is 3.55. So, now, we have a reduced table with only B and C, A is not there. So, we will get an average of Co here, which is this plus this by 2. Average of C1 here this by this by 2. Average of Bo, this plus this by 2. Average of B1, this by this by 2.

The global average will be average of any two of these. Any two of these, Co C1 or Bo B1 like that. So, will get 5.45, this is same as the original value.

Now, how do you do the sum of squares for C? Simple, we have 5.15 - 5.45 $^2$ X 4. Do not forget, we have to multiply by 4. Then 5.75 -5.45 $^2$ X 4, then we have 7.87 - 5.45 $^2$ X 4 ; 3.02 - 5.45 $^2$ X 4, 4 here. Now, if you add these you will get sum of squares for corresponding to C, if you add these you will get sum of squares corresponding to B here, quite simple. Now, how do you get the total? Total is 7.8; 7.8 - 5.45, square it up and then multiply by 2. Do not forget for the total sum of squares, we are multiplying by 2 because each one of these term is made up of 2 terms right? Like, I said if you want to take this particular term, I am adding this plus this by 2, 7.2.

So, this itself is made up of 2. That is why, for total sum of squares, what do we do? 7.8 - 5.45 $^2$ X 2and so on actually 2.5 - 5.45 X 2, 7.95 - 5.45 $^2$ X 2; 3.55 - 5.45 X 2. So, we get these four terms, add up all of these, this will give you total sum of squares. So, from total sum of squares, we can some subtract sum of squares for B, sum of squares for C. That will either we call it error or we call it sum of squares due to BC. As I said, C has two levels, so DF is 1, B has two levels, so DF is 1 and the total you have 4 experiments. So, 3 degrees of freedom so you have 1 extra DF, which we can be calling it as error or we can call it sum of squares due to BC.

So, this is how we have done, we have calculated the main effects and then we have calculated the interaction effects AB, AC and BC and we also have the total sum of squares, which is given by the 52.56. So, here we have the sum of squares because of A, main effects is called sum of squares because of B, sum of squares because of C, and this is total sum of squares, so it may contain everything.

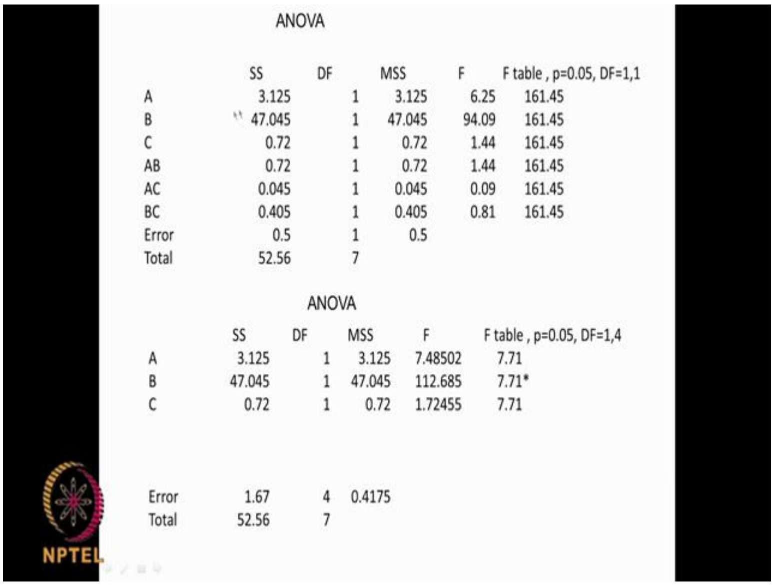
Then, we got AB sum of squares, how did you get the A B sum of squares? What did we do? We averaged the C here. So, what did we do? We just made B and A table, C is averaged out. So, 6.7 + 7.2 ÷2; 8.4 + 9.2 ÷ 2; 2 + 3.4 ÷2; 3 + 3.7 ÷2 that is what we got. So, then we did an average for Ao, average for A1, average for Bo, average for B1, global average. Then for A what did I do? 4.85 - 5.45 $^2$, but multiplied by 4; 6.075 - 5.45 $^2$ X 4. Why 4? Because each of these terms is made up of 2 and each of these is also made up of 2 from the original table. That is why we are multiplying by 4. Now, you get sum of squares for A, sum of squares for B.

Now, how do you get total sum of squares? We again, as you know this is your table so $6.95 - 5.45, 6.95 - 5.45^{2}$ but do not forget you need to multiply by 2 because this 6.95 is made up of this average of these, right? So, you get this then, $8.8 - 5.45^{2}$ X 2; $2.7 - 5.45^{2}$ X 2; $3.35 - 5.45^{2}$ X 2 that will give you total sum of squares. Do you understand? Then you can subtract TSS minus, sorry, you can subtract SS of A, SS of B from TSS to get either we call it error or we can call it interaction AB. That is what I mentioned here; it will have 1 degree of freedom, because A has 1 degree of freedom, B has 1 degree of freedom. So, same thing we can do, just like averaged out C to get a table for AB, I can average out B to get a table for AC. So, once I do that, I do the same calculation. Do not forget, when I am calculating these sum of squares for A I will multiply by 4, C also I multiply by 4, but when I am calculating the total sum of squares, I will be multiplying by 2. So, here I will get the average for Ao, average for A1, average for Co, average for C1, global average then for A, how do I get it? This minus this, square, multiply by 4. Then, how do I get total? I will take this $5.45 - 4.6^{2}$ but X 2, do not forget this.

So, I will get 4 terms which I add up to get total sum of squares. Now, when I subtract each of these effect A and C from TSS, I will get AC or I can call it error also, but I will call it AC for a time being. But, it will have 1 degree of freedom, then after that I make a table for BC, where I am averaging out A. How do I average out A? I will take this plus this, add up, divide it by 2. This plus this, add up, divide it by 2. So, so that will be column for Co and row for Bo. This plus this, add up, divide it by 2, this plus this, add up, divide it by 2 that will be column for C1. So, A is averaged out understood? So, we have the average for Co, average for C1, average for Bo, average for B1, then global average.

Now for C, when you are calculating sum of squares, what do you do? $5.45 - 5.15^{2}$ X 4. Same thing here, then B, $7.875 - 5.45^{2}$ X 4. Then for this 3.025 - 5.45 square $^{2}$ X 4. Now here, how do I do? $5.45 - 7.8^{2}$ X 2; 2.5 - 5.45, square it up, then multiply by 2 here, do not forget. When $3.55 - 5.45^{2}$ X 2, then you add up all these 4 terms together that will give you total sum of squares, so you subtract sum of squares due to effect C, sum of squares due to effect B that should give you BC. Now, we can go for our ANOVA table.

(Refer Slide Time: 25:31)



ANOVA

| | SS | DF | MSS | F | F table , p=0.05, DF=1,1 |
|---|---|---|---|---|---|
| A | 3.125 | 1 | 3.125 | 6.25 | 161.45 |
| B | 47.045 | 1 | 47.045 | 94.09 | 161.45 |
| C | 0.72 | 1 | 0.72 | 1.44 | 161.45 |
| AB | 0.72 | 1 | 0.72 | 1.44 | 161.45 |
| AC | 0.045 | 1 | 0.045 | 0.09 | 161.45 |
| BC | 0.405 | 1 | 0.405 | 0.81 | 161.45 |
| Error | 0.5 | 1 | 0.5 | | |
| Total | 52.56 | 7 | | | |

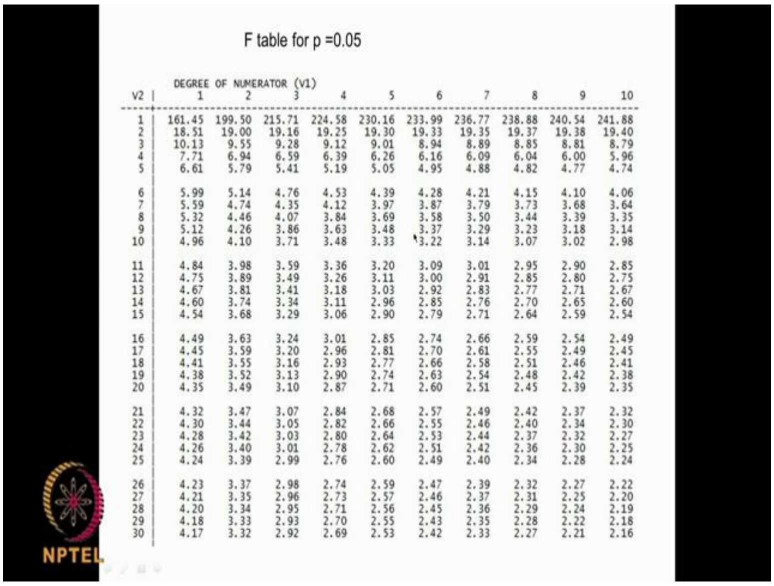
ANOVA

| | SS | DF | MSS | F | F table , p=0.05, DF=1,4 |
|---|---|---|---|---|---|
| A | 3.125 | 1 | 3.125 | 7.48502 | 7.71 |
| B | 47.045 | 1 | 47.045 | 112.685 | 7.71* |
| C | 0.72 | 1 | 0.72 | 1.72455 | 7.71 |
| Error | 1.67 | 4 | 0.4175 | | |
| Total | 52.56 | 7 | | | |



This is our ANOVA table, we have the main effects A, B, C sum of squares, sum of squares, you have written down. Now, I picked the AB, AC, BC. So, if I subtract the total from each one of them, if I subtract the total from each one of them; I should get error. So, I add up all these things, subtract from 52.56; that will give you 0.5. How did I get this 52.56? That is the original one, if you remember original table, understand?, original table that is the total sum of squares. Now, totally I have done 8 experiments 2 X 2 X 2. So, 7 degrees of freedom; A has 1 degree of freedom, B has 1, C has 1, AB has 1, AC has 1, BC has 1. So, if you add up all these and subtract from 7, you will have 1 degree of freedom. Error is coming out to be 0.5. So, how do you calculate a mean sum of square sum of squares? Divide by the DF, so you will get this. How do you calculate F? You divide 3.125 / 0.5, 47 / 0.5, 0.72 / 0.5, 0.72 / 0.5, 0.045 / 0.5, 0.45 / 0.5, 0.8. So, this is the F ratios. Now, let us go to the table for 1 degree of freedom, 1 X 1, 1 comma 1.

(Refer Slide Time: 26:54)



F table for p =0.05

| V2 | DEGREE OF NUMERATOR (V1) 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 161.45 | 199.50 | 215.71 | 224.58 | 230.16 | 233.99 | 236.77 | 238.88 | 240.54 | 241.88 |
| 2 | 18.51 | 19.00 | 19.16 | 19.25 | 19.30 | 19.33 | 19.35 | 19.37 | 19.38 | 19.40 |
| 3 | 10.13 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.89 | 8.85 | 8.81 | 8.79 |
| 4 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.09 | 6.04 | 6.00 | 5.96 |
| 5 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.77 | 4.74 |
| 6 | 5.99 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.21 | 4.15 | 4.10 | 4.06 |
| 7 | 5.59 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.79 | 3.73 | 3.68 | 3.64 |
| 8 | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.50 | 3.44 | 3.39 | 3.35 |
| 9 | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.29 | 3.23 | 3.18 | 3.14 |
| 10 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 3.02 | 2.98 |
| 11 | 4.84 | 3.98 | 3.59 | 3.36 | 3.20 | 3.09 | 3.01 | 2.95 | 2.90 | 2.85 |
| 12 | 4.75 | 3.89 | 3.49 | 3.26 | 3.11 | 3.00 | 2.91 | 2.85 | 2.80 | 2.75 |
| 13 | 4.67 | 3.81 | 3.41 | 3.18 | 3.03 | 2.92 | 2.83 | 2.77 | 2.71 | 2.67 |
| 14 | 4.60 | 3.74 | 3.34 | 3.11 | 2.96 | 2.85 | 2.76 | 2.70 | 2.65 | 2.60 |
| 15 | 4.54 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.71 | 2.64 | 2.59 | 2.54 |
| 16 | 4.49 | 3.63 | 3.24 | 3.01 | 2.85 | 2.74 | 2.66 | 2.59 | 2.54 | 2.49 |
| 17 | 4.45 | 3.59 | 3.20 | 2.96 | 2.81 | 2.70 | 2.61 | 2.55 | 2.49 | 2.45 |
| 18 | 4.41 | 3.55 | 3.16 | 2.93 | 2.77 | 2.66 | 2.58 | 2.51 | 2.46 | 2.41 |
| 19 | 4.38 | 3.52 | 3.13 | 2.90 | 2.74 | 2.63 | 2.54 | 2.48 | 2.42 | 2.38 |
| 20 | 4.35 | 3.49 | 3.10 | 2.87 | 2.71 | 2.60 | 2.51 | 2.45 | 2.39 | 2.35 |
| 21 | 4.32 | 3.47 | 3.07 | 2.84 | 2.68 | 2.57 | 2.49 | 2.42 | 2.37 | 2.32 |
| 22 | 4.30 | 3.44 | 3.05 | 2.82 | 2.66 | 2.55 | 2.46 | 2.40 | 2.34 | 2.30 |
| 23 | 4.28 | 3.42 | 3.03 | 2.80 | 2.64 | 2.53 | 2.44 | 2.37 | 2.32 | 2.27 |
| 24 | 4.26 | 3.40 | 3.01 | 2.78 | 2.62 | 2.51 | 2.42 | 2.36 | 2.30 | 2.25 |
| 25 | 4.24 | 3.39 | 2.99 | 2.76 | 2.60 | 2.49 | 2.40 | 2.34 | 2.28 | 2.24 |
| 26 | 4.23 | 3.37 | 2.98 | 2.74 | 2.59 | 2.47 | 2.39 | 2.32 | 2.27 | 2.22 |
| 27 | 4.21 | 3.35 | 2.96 | 2.73 | 2.57 | 2.46 | 2.37 | 2.31 | 2.25 | 2.20 |
| 28 | 4.20 | 3.34 | 2.95 | 2.71 | 2.56 | 2.45 | 2.36 | 2.29 | 2.24 | 2.19 |
| 29 | 4.18 | 3.33 | 2.93 | 2.70 | 2.55 | 2.43 | 2.35 | 2.28 | 2.22 | 2.18 |
| 30 | 4.17 | 3.32 | 2.92 | 2.69 | 2.53 | 2.42 | 2.33 | 2.27 | 2.21 | 2.16 |



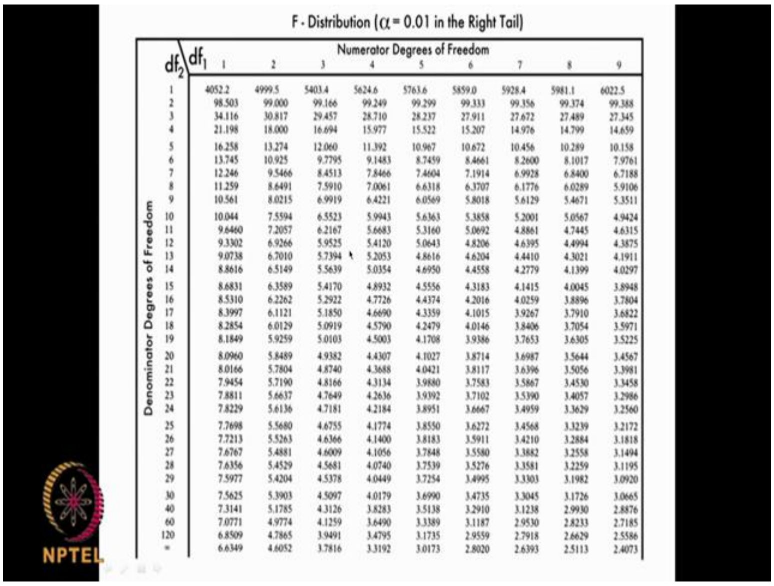
1 comma 1 for a 95 %, we get 161.45, that is a big number so, obviously, we end up saying that none of the main effects or even the interaction effects are significant. Now because the interaction once I said that interaction can even be considered as error, we can add all these into the error, all these into the error, all these into the error. So, 3.125, 47.05, 0.72, 1, 1, 1 because each of the main effects has 2 levels. Now I can add all these interactions with the error. So, the error will have 4 degrees of freedom, because total experiment is 7, each of the effects has 1, 1, 1, so 4 degrees of freedom.

When I add up all these, I will end up with the error as 1.67. So, mean sum of squares divided by one, divided by one, divided by one. Here it will be ÷4 so 0.415. So, 3.1 ÷0.4 is 7.4, 47 ÷ 0.4 is 112 and then 0.7 / 0.4 is 1.7. Now if you go to F table, the degrees of freedom is 1 comma 4. So, 1 comma 4, it comes out to be 7.71 when you do that, we see that the main effect B is significant, because F calculated is larger than the F from the table. So, B is significant that means, nitrogen seems to have a significant effect in the metabolic production actually. Now A is not really significant, because it is 7.48, but it is closer to 7.71 so you better do more experiments to see whether it is really not significant or is it statistically significant so because it is sort of closer to that. So, you understand this problem; it is quite an involved problem, we have three parameters A, B, C. These are called the main effects and then so obviously there could be interaction AB, AC, BC type of interactions are possible. So, the number of experiments are less. That means you have not done any replications so total number of experiments are 7.

If A, B, C take away one DF, we are left with only four DFs and if you want to calculate AB, AC, BC then we will have 1, 1, 1, so error will end up having only 1 degree of freedom. So, when we use that approach and calculate F values, we see none of them are significant, but then these interaction are any way not significant or very small. So, we can combine that with the error. So, what happens? DF for the error goes up to 4 and we have to add up all these with 0.5. So, this comes out to be 1.67. So, F ratios change to 7.48, 112, 1.7.

But the main thing that happens is now we are looking in the table at 1 , 4 degrees of freedom. So, the number has gone down dramatically to 7.71. Then we see that main effect B is significant, ok?, main effect B is significant, because it is 112 at 95 % or even at 99 % it is a significant, right?.

(Refer Slide Time: 30:31)



F - Distribution ($\alpha$ = 0.01 in the Right Tail)

| Denominator Degrees of Freedom $df_2$ \ Numerator Degrees of Freedom $df_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 4052.2 | 4999.5 | 5403.4 | 5624.6 | 5763.6 | 5859.0 | 5928.4 | 5981.1 | 6022.5 |
| 2 | 98.503 | 99.000 | 99.166 | 99.249 | 99.299 | 99.333 | 99.356 | 99.374 | 99.388 |
| 3 | 34.116 | 30.817 | 29.457 | 28.710 | 28.237 | 27.911 | 27.672 | 27.489 | 27.345 |
| 4 | 21.198 | 18.000 | 16.694 | 15.977 | 15.522 | 15.207 | 14.976 | 14.799 | 14.659 |
| 5 | 16.258 | 13.274 | 12.060 | 11.392 | 10.967 | 10.672 | 10.456 | 10.289 | 10.158 |
| 6 | 13.745 | 10.925 | 9.7795 | 9.1483 | 8.7459 | 8.4661 | 8.2600 | 8.1017 | 7.9761 |
| 7 | 12.246 | 9.5466 | 8.4513 | 7.8466 | 7.4604 | 7.1914 | 6.9928 | 6.8400 | 6.7188 |
| 8 | 11.259 | 8.6491 | 7.5910 | 7.0061 | 6.6318 | 6.3707 | 6.1776 | 6.0289 | 5.9106 |
| 9 | 10.561 | 8.0215 | 6.9919 | 6.4221 | 6.0569 | 5.8018 | 5.6129 | 5.4671 | 5.3511 |
| 10 | 10.044 | 7.5594 | 6.5523 | 5.9943 | 5.6363 | 5.3858 | 5.2001 | 5.0567 | 4.9424 |
| 11 | 9.6460 | 7.2057 | 6.2167 | 5.6683 | 5.3160 | 5.0692 | 4.8861 | 4.7445 | 4.6315 |
| 12 | 9.3302 | 6.9266 | 5.9525 | 5.4120 | 5.0643 | 4.8206 | 4.6395 | 4.4994 | 4.3875 |
| 13 | 9.0738 | 6.7010 | 5.7394 | 5.2053 | 4.8616 | 4.6204 | 4.4410 | 4.3021 | 4.1911 |
| 14 | 8.8616 | 6.5149 | 5.5639 | 5.0354 | 4.6950 | 4.4558 | 4.2779 | 4.1399 | 4.0297 |
| 15 | 8.6831 | 6.3589 | 5.4170 | 4.8932 | 4.5556 | 4.3183 | 4.1415 | 4.0045 | 3.8948 |
| 16 | 8.5310 | 6.2262 | 5.2922 | 4.7726 | 4.4374 | 4.2016 | 4.0259 | 3.8896 | 3.7804 |
| 17 | 8.3997 | 6.1121 | 5.1850 | 4.6690 | 4.3359 | 4.1015 | 3.9267 | 3.7910 | 3.6822 |
| 18 | 8.2854 | 6.0129 | 5.0919 | 4.5790 | 4.2479 | 4.0146 | 3.8406 | 3.7054 | 3.5971 |
| 19 | 8.1849 | 5.9259 | 5.0103 | 4.5003 | 4.1708 | 3.9386 | 3.7653 | 3.6305 | 3.5225 |
| 20 | 8.0960 | 5.8489 | 4.9382 | 4.4307 | 4.1027 | 3.8714 | 3.6987 | 3.5644 | 3.4567 |
| 21 | 8.0166 | 5.7804 | 4.8740 | 4.3688 | 4.0421 | 3.8117 | 3.6396 | 3.5056 | 3.3981 |
| 22 | 7.9454 | 5.7190 | 4.8166 | 4.3134 | 3.9880 | 3.7583 | 3.5867 | 3.4530 | 3.3458 |
| 23 | 7.8811 | 5.6637 | 4.7649 | 4.2636 | 3.9392 | 3.7102 | 3.5390 | 3.4057 | 3.2986 |
| 24 | 7.8229 | 5.6136 | 4.7181 | 4.2184 | 3.8951 | 3.6667 | 3.4959 | 3.3629 | 3.2560 |
| 25 | 7.7698 | 5.5680 | 4.6755 | 4.1774 | 3.8550 | 3.6272 | 3.4568 | 3.3239 | 3.2172 |
| 26 | 7.7213 | 5.5263 | 4.6366 | 4.1400 | 3.8183 | 3.5911 | 3.4210 | 3.2884 | 3.1818 |
| 27 | 7.6767 | 5.4881 | 4.6009 | 4.1056 | 3.7848 | 3.5580 | 3.3882 | 3.2558 | 3.1494 |
| 28 | 7.6356 | 5.4529 | 4.5681 | 4.0740 | 3.7539 | 3.5276 | 3.3581 | 3.2259 | 3.1195 |
| 29 | 7.5977 | 5.4204 | 4.5378 | 4.0449 | 3.7254 | 3.4995 | 3.3303 | 3.1982 | 3.0920 |
| 30 | 7.5625 | 5.3903 | 4.5097 | 4.0179 | 3.6990 | 3.4735 | 3.3045 | 3.1726 | 3.0665 |
| 40 | 7.3141 | 5.1785 | 4.3126 | 3.8283 | 3.5138 | 3.2910 | 3.1238 | 2.9930 | 2.8876 |
| 60 | 7.0771 | 4.9774 | 4.1259 | 3.6490 | 3.3389 | 3.1187 | 2.9530 | 2.8233 | 2.7185 |
| 120 | 6.8509 | 4.7865 | 3.9491 | 3.4795 | 3.1735 | 2.9559 | 2.7918 | 2.6629 | 2.5586 |
| ∞ | 6.6349 | 4.6052 | 3.7816 | 3.3192 | 3.0173 | 2.8020 | 2.6393 | 2.5113 | 2.4073 |



And main effect A is not significant as per this table comparison, but still 7.48 is closer to 7.71. So, you better watch out; do few more experiments to confirm whether you to accept the null hypothesis or reject the null hypothesis. So, in this problem we had possibility of interactions being measured, but then our error DF is very low only 1 degree of freedom because we have not done any replications of the experiments. If you had done lot of replications, then we might have had a better chance of seeing interaction effects. But here because we did not do much replications, either we can call all these interactions into error or we can try to put them each separately. But when we tried that it does not work out, because error ends up having only 1 degree of freedom, whereas when we combine all these together and put them as error then error gets a higher number of degrees of freedom. So, we are able to see some mean effects being significant as against to this particular ANOVA table, where none of the main effects are significant, do you understand this problem? it is very interesting, very extremely interesting problem where in a bioprocess or even in drug discovery, drug testing you will end up having interactions and if you want to see the interactions, the degrees of freedom for error goes down so dramatically that you are not able to really see clearly the effect of each of these parameters.

So, one of the take away here is error should have sufficiently large degree of freedom, it cannot be so low as 1, that is one thing. Other thing is if we have a 2 by 2 table and if you have not replicated then you can either call it interaction or you can call it error because the degrees of freedom are much less.

So, that way this problem is extremely interesting and also I taught how to reduce from a 3 by 3 type of table into a 2 by 2 table by taking average of for each variable. So, when we had A, B, C in 2 levels 8 experiments what we did? We had a, made a table of A and B alone by taking averages from C, and then later on we had A and C alone by taking averages of B. Then later on we had B and C alone by taking averages of A. So, we, every time we reduced it to 2 by 2 type of table and we looked at the interaction effects also. So, we will continue again in the next class on some other topic.

Thank you very much.

Key words: Interaction, Main effects, Replication, Error, Degree of Freedom, F value, F table, ANOVA table, Three way ANOVA, Mean sum of squares, Total sum of squares, Statistically significant, Null hypothesis