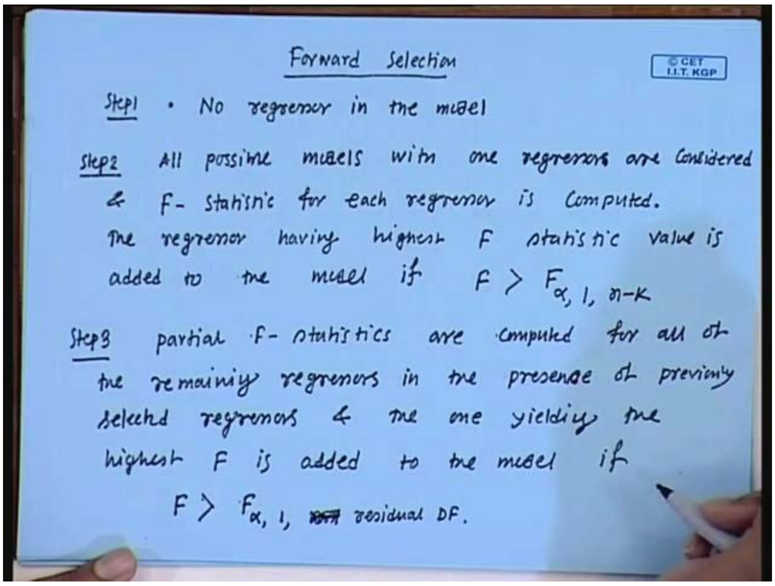
Regression Analysis
Prof. Soumen Maity
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 13
Selecting the Best Regression Model (Contd.)

Hi, this is my forth lecture on Selecting Best Regression Model, and this one is basically you know continuation of the previous class. In the last class we talked about backward elimination, and today we are going to talk about forward selection and stepwise selection.

(Refer Slide Time: 01:28)



Forward Selection

© CET I.I.T. KGP

Step1 • No regressor in the model

Step2 All possible models with one regressors are considered & F- statistic for each regressor is computed.
The regressor having highest F statistic value is added to the model if $F > F_{\alpha, 1, n-K}$

Step3 partial F- statistics are computed for all of the remaining regressors in the presence of previously selected regressors & the one yielding the highest F is added to the model if
$F > F_{\alpha, 1, \text{residual DF}}$.

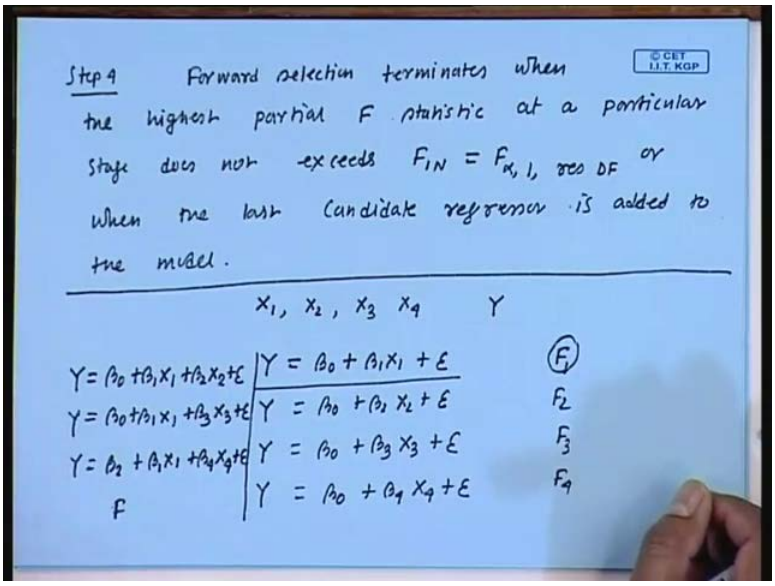
First I will talk about forward selection, so the basic motivation behind the forward selection is that you know we start with no regressor in the model, and then every step we add the most relevant regressor to the model. So, every step we keep on adding one regressor to the model and there is some stopping criteria, we will talk about those thing now.

Well, the first step is start with no regressor in the model and then step two all possible models with one regressor are considered and F static for each regressor is computed also. If there are k minus 1 regressor in the model you have to compute k minus 1, F statistic and the regressor having highest F statistic value is added to the model provided if the value the highest F statistic is written then F alpha 1 n minus k.

So, this is basically no residual degree of freedom, something in the last class I told you know, I told that this is era degree of freedom, so era degree of freedom and residual degree of freedom are the same. Well, I will illustrate using some example later on let me just write down the algorithm first, and then step 3 partial F statistics are computed for all of the remaining regressors in the presence of previously selected regressors.

And the one yielding the highest F is added to the model provided, if a course F is greater than the tabulated a value; that means, F alpha 1 n minus k. I should not say minus k maybe, I should write you know one residual degree of freedom, here it might be, no here also I should write residual degree of freedom.

(Refer Slide Time: 07:30)



Step 4 Forward selection terminates when the highest partial F statistic at a particular stage does not exceeds $F_{IN} = F_{\alpha, 1, \text{res DF}}$ or when the last candidate regressor is added to the model.

$X_1, X_2, X_3\ X_4 \qquad Y$

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$ | $Y = \beta_0 + \beta_1 X_1 + \varepsilon$ $\quad \textcircled{F_1}$

$Y = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \varepsilon$ | $Y = \beta_0 + \beta_1 X_2 + \varepsilon$ $\quad F_2$

$Y = \beta_2 + \beta_1 X_1 + \beta_4 X_4 + \varepsilon$ | $Y = \beta_0 + \beta_3 X_3 + \varepsilon$ $\quad F_3$

F | $Y = \beta_0 + \beta_4 X_4 + \varepsilon$ $\quad F_4$

Well, and the stopping criteria is that the step 4, the forward selection terminates, when the highest partial F statistic at a particular stage does not exceed, F I N, so F I N means the tabulated value of F. So, basically this one is F alpha 1 and the residual degree of freedom or when the last candidate regressors is added to the model, so this is the algorithm. And now, what I want to do is that I want illustrate this algorithm using one example, let me just give the over view of this one.

Initially, stating with no regressor in the model, and then all possible model with one regressor are considered; that means, suppose my problem has 4 regressors X 1, X 2, X 3, X 4, and one response variable Y. So, what I will do is, I will in the first step I will consider the F statistic value for the model Y equal to beta naught plus beta 1 X 1 plus epsilon, Y equal to beta naught plus beta 2 X 2 plus epsilon, Y equal to beta naught plus beta 3 X 3 plus epsilon, Y equal to beta naught plus beta 4 X 4 plus epsilon.

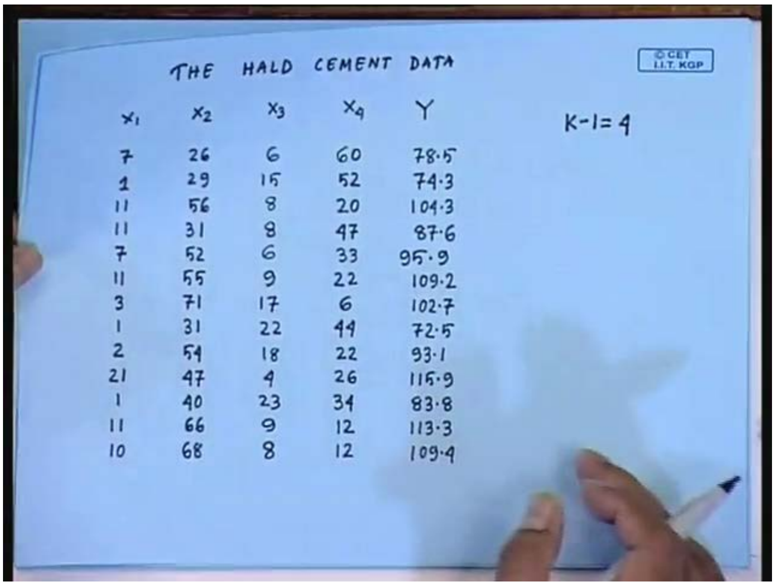
So, these are the regressor model with a involving one regressor variable, so I will compute the F statistic for each of these model and the model of the regressor having the highest F, having highest F means the associated if say F 1 is with this model this is F 2, F 3, F 4. So, F 1 is highest means associated random associated regressor X 1 is significant to explain the variability, these are most significant among four regressors to explain the variability in Y.

And once you select you know the suppose X 1 is selected and then the next step is that know keeping X 1 in the model Y equal to beta naught plus beta 1 X 1 is there in the model. Now, we seek for the next basic regressor in the models, so we will try with the model like beta 1 X 1 plus beta 2 X 2 plus epsilon, and will try with the model Y equal to beta naught plus beta 1 X 1 is fixed, and then beta 3 X 3 plus epsilon, and Y equal to beta 2 plus beta 1 X 1 plus beta 4 X 4 plus epsilon.

So, we compare in the presence of X 1 we will see which one is the based, whether X 2 is based in the presence of X 1, or X 3 based in the presence X 1 or X 4 is based in the presence of X 1. So, and depending on that can be evaluated by computing the partial F value and highest partial F value the model which is having the highest partial F value will be the associated regressor will be included in the model. So, this is the and we keep on doing these thing and there is some stopping criteria, this says the stopping criteria. So, this is that you know forward selection terminates when the highest partial F statistic at a particular stage does not exceeds the threshold value.

That means, once you have if the partial F value here is not greater than the threshold value; that means, X 1 is enough for the model you do not need to include any other regressor in the model, so we will stop there I mean the final will be this one. Well, let me explain in detail the algorithm, so this is a outline of the algorithm now I will be considering the HALD cement data again.

(Refer Slide Time: 13:43)



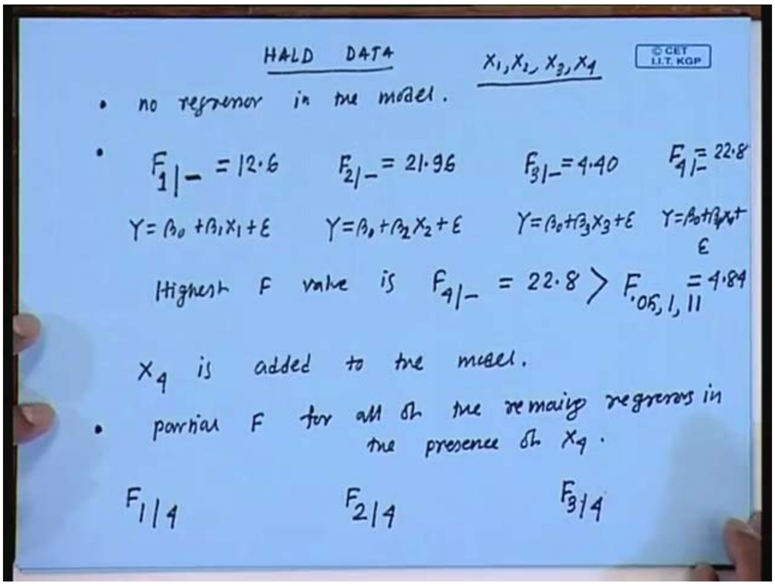
THE HALD CEMENT DATA

x_1	x_2	x_3	x_4	Y
7	26	6	60	78·5
1	29	15	52	74·3
11	56	8	20	104·3
11	31	8	47	87·6
7	52	6	33	95·9
11	55	9	22	109·2
3	71	17	6	102·7
1	31	22	44	72·5
2	54	18	22	93·1
21	47	4	26	115·9
1	40	23	34	83·8
11	66	9	12	113·3
10	68	8	12	109·4

$K-1=4$

So, you know it has the 4 regressors, and one response variable Y well now I will illustrate the forward selection technique using the HALD cement data.

(Refer Slide Time: 14:04)



HALD DATA X_1, X_2, X_3, X_4

- no regressor in the model.
- $F_{1|-} = 12{\cdot}6$ $F_{2|-} = 21{\cdot}96$ $F_{3|-} = 4{\cdot}40$ $F_{4|-} = 22{\cdot}8$

$Y = \beta_0 + \beta_1 X_1 + \epsilon$ $Y = \beta_0 + \beta_2 X_2 + \epsilon$ $Y = \beta_0 + \beta_3 X_3 + \epsilon$ $Y = \beta_0 + \beta_4 X_4 + \epsilon$

Highest F value is $F_{4|-} = 22{\cdot}8 > F_{{\cdot}05,1,11} = 4{\cdot}84$

X_4 is added to the model.

- partial F for all of the remaining regressors in the presence of X_4.

$F_{1|4}$ $F_{2|4}$ $F_{3|4}$

So, what I do is that initially there is no regressor in the model, and then I will compute for this data there are 4 regressors X 1, X 2, X 3 and X 4. Now, I will compute F value for the model Y equal to beta naught plus beta 1 X 1 plus epsilon, the F value associated with this model. I will call it F 1 using the previous notation I will say F 1 nothing; that means, this is the F statistics, but in terms of the partial F value and this is the meaning of this one you know this is the partial F value for the regressor, X 1 in the presence of no other regressor in the model, that is why this dash we compute this value.

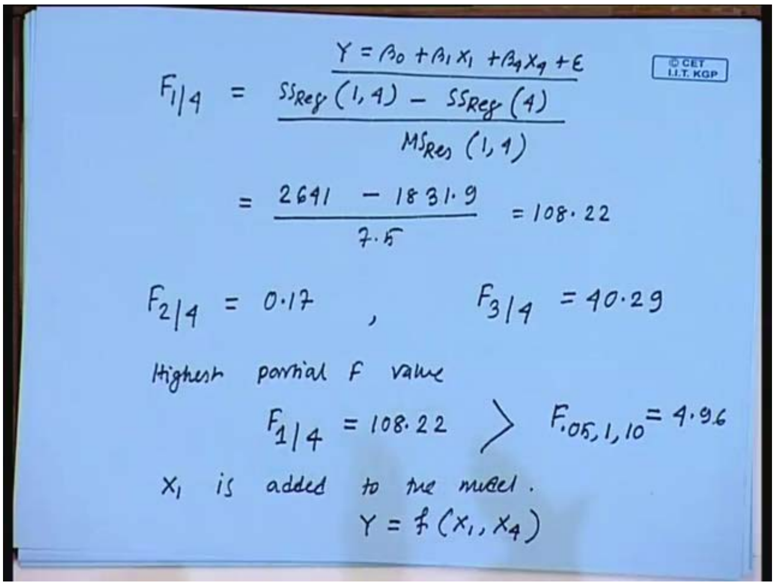
Similarly, considered the second simple linear equation model that is beta Y is equal to beta naught plus beta 2, X 2 plus epsilon the associated F value is F 2 in the presence of nothing. So, we will evaluate this model also Y equal to beta naught plus beta 3 X 3 plus epsilon, so F 3 in the presence of nothing and Y equal to beta naught plus beta 4 X 4 plus epsilon, so I will compute F 4 in the presence of nothing.

Well so how to get these value you know, you just I have this inverted blue for this model see, this is fitted equation for the HALD cement data you know the same thing you compute once you have the fitted value you compute the S S residual, you know S S T this is the S S T and then you have the S S residual and then you have the F statistic. So, this is how you have to find the a S statistic value or this model and in the forward selection algorithm, we denote this F by F 1 the presence of we call if partial F basically you know the global F, but we call it F 1 in the presence of no other regressor model, so F 1 value is 12.6.

So, it is a 12.6, similarly you fit this model you will get the value is equal to 21.96, this value is equal to 4.40, and this value is equal to 22.8. Now of course, the highest F value is F 4, which is equal to 22.8, so F 4 is, so X 4 is most significant to explain the variability in Y. So, X 4 and this value is also you know this is greater than F, if you compute F 0 5 1 and the residual degree of freedom here you have only 1 2 unknown in the model, so 13 minus 2 that is 11. So, this is the residual degree of freedom and this one has value 44.84, so that observed value is greater than the tabulated value, so X 4 is added to the model.

Now, the next thing is that we compute the partial F statistic for all of the remaining regressor in the presence of X 1, remaining regressor means see X 4 is already in the model, so in the presence of X 4. So, X 4 is already in the model, so what we do is that we will compute F 1 in the presence of X 4, we will compare the partial F statistic associated with X 2 that is F 2 in the presence of X 4 in the model. And also we compute F 3 in the presence of X 4 in the model, well I hope that you know how to compute this value for example, say F 1 given 4, that means you have to consider the model Y equal to beta naught plus beta 1 X 1 plus beta 4 X 4 plus epsilon, you feed this model.

(Refer Slide Time: 20:48)



$$F_{1|4} = \frac{SS_{Reg}(1,4) - SS_{Reg}(4)}{MS_{Res}(1,4)}$$

$$Y = \beta_0 + \beta_1 x_1 + \beta_4 x_4 + \epsilon$$

$$= \frac{2641 - 1831.9}{7.5} = 108.22$$

$F_{2|4} = 0.17$, $F_{3|4} = 40.29$

Highest partial F value

$$F_{1|4} = 108.22 > F_{.05,1,10} = 4.96$$

x_1 is added to the model.

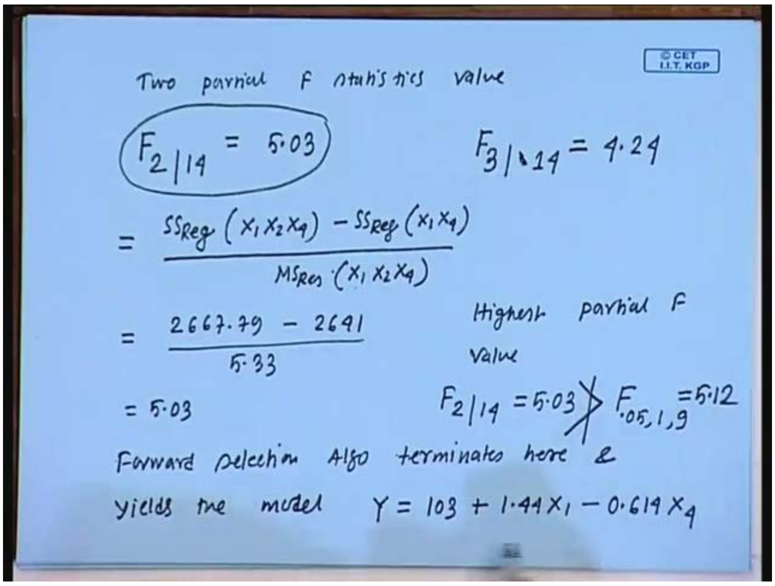
$$Y = f(x_1, x_4)$$

And then S S regression for the full model that is 1 4; that means, X 1 X 4 minus S S regression, X 4 by MS residual for the full model means you know 1 and 4 well. So, this can be this you can check that value is equal to 2641 minus 1831.9 by 7.5 and this is going to be 108.22. And similarly, you can check that if 2 in the presence of X 4 in the model is equal to 0.17 and if 3 in the presence of 4 X 4 in the model is equal to 40.29.

So; that means, if the highest F statistic value is equal to 1 0, so the highest partial F value is F 1 in the presence of 4, which is equal to 108.22. And you know now we will check the tabulated value F 0.05 1 and he residual degree of freedom here, now you have three unknown in the model, so 13 minus 3 that is 10 that is residual degree of freedom. And this 1 is equal to this value is equal to 4.96 and of course, this one is larger than then 4.96, so in the presence of 4.

So, the meaning of F 1 4 F value is 108.22 which is larger than the tabulated value; that means, in the presence of first we added X 4 in the model, now in the presence of X 4 one is significant. So, next we will add, so X 1 is now added to the model, now my model you know it consist of 2, we have selected two random 2 regresors, so Y is now function of X 1 and X 4. Now, what we do is that you know see X 1 and X 4 have been added or selected, now in the presence of X 1 X 4 what we will see whether X 2 is significant in the presence of X 1 X 4 or whether X 3 is significant in the presence of X 1 and X 4 in the model.

(Refer Slide Time: 24:58)



Two partial F statistics value

$\boxed{F_{2|14} = 5.03}$ $F_{3|14} = 4.24$

$$= \frac{SS_{Reg}(X_1 X_2 X_4) - SS_{Reg}(X_1 X_4)}{MS_{Res}(X_1 X_2 X_4)}$$

$$= \frac{2667.79 - 2641}{5.33}$$

Highest partial F value

$= 5.03$ $F_{2|14} = 5.03 \not> F_{.05,1,9} = 5.12$

Forward selection Also terminates here & yields the model $Y = 103 + 1.44 X_1 - 0.614 X_4$

So, if none of them are significant then we stop here, so this is what we have to do, now we will check this partial value 2 partial F statistics value, what we will do is that we will check the significance of X 2 in the presence of X 1 and X 4 in the model. So, we will compute these value also we will compute F 3, the significance of X 3 in the presence of X 1 and X 4 in the model.

So, will compute these two value you know how to compute you know I do not want to repeat again, you can check that this value is equal to 5.0,3 and this value is equal to 4.24. Still, maybe just I will write down this one see this is equal to S S regression, for the full model that is X 1 X 2 X 4 minus S S regression for the model involving X 1 and X 4 by MS residual for the model involving X 1 X 2 X 4.

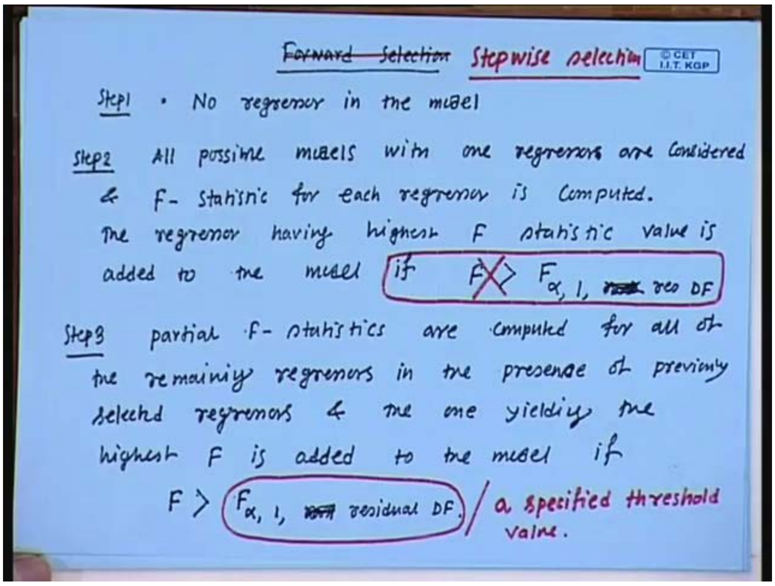
So, this one you can check this 2667.79 minus 2641 by 5.33 and this is equal to 5.03, so the highest partial F value is this one, now whether this will be included. So, whether next X 2 will be included in the model, that depends on the tabulated value of the F or the threshold value, so the highest partial F value is F 2 in the presence of X 1 and X 4 in the model which is equal to 5.03.

Now, what is the tabulated value F 0.05 1 and the residual degree of freedom, so I am talking about 3 regressors in the model full model for computing the ms residual. So, there will be 4 unknown; that means, the degree of freedom is 10 minus 4 that is equal to

9, so this value is 5.12 see now this one is not greater than that tabulated value, so we cannot include X 2 in the model at this moment.

We have no X 1 and X 4 in the model and X 1 and X 4, in the model if X 2 has the highest partial F value, but that value is less than the tabulated value, so that means, X 2 is not significant in the presence of X 1 and X 4 in the model. So, we cannot include or add X 2 in the model, so you have to we have to stop here, so this is the stopping criteria right well. So, the forward selection algorithm terminates here and yields the model output of the forward selection algorithm is Y equal to 103 plus 1.44, so I am giving the fitted model. Basically, what I want to see say by this one is I want to say that the final model involve X 1 and X 4, ye final model involve or ye involve X 1 and X 4 well, so this is the output of forward selection and next we move for stepwise selection.

(Refer Slide Time: 30:44)



~~Forward Selection~~ Stepwise selection

Step1 • No regressor in the model

Step2 All possible models with one regressors are considered & F- statistic for each regressor is computed. The regressor having highest F statistic value is added to the model $\boxed{\text{if} \quad F > F_{\alpha, 1, \text{res DF}}}$

Step3 partial F- statistics are computed for all of the remaining regressors in the presence of previously selected regressors & the one yielding the highest F is added to the model if

$F > \boxed{F_{\alpha, 1, \text{residual DF}}}$ / a specified threshold value.

Stepwise selection, is it is basically you know combination of forward selection and the backward elimination, here also we start with the no regressior in the model. And let me explain the model you know this is very similar to the forward selection model, what I will do is that I just revise this forward selection model. Now, I am talking about say stepwise selection, I will modify this algorithm, here also we start with no regressor in the model.

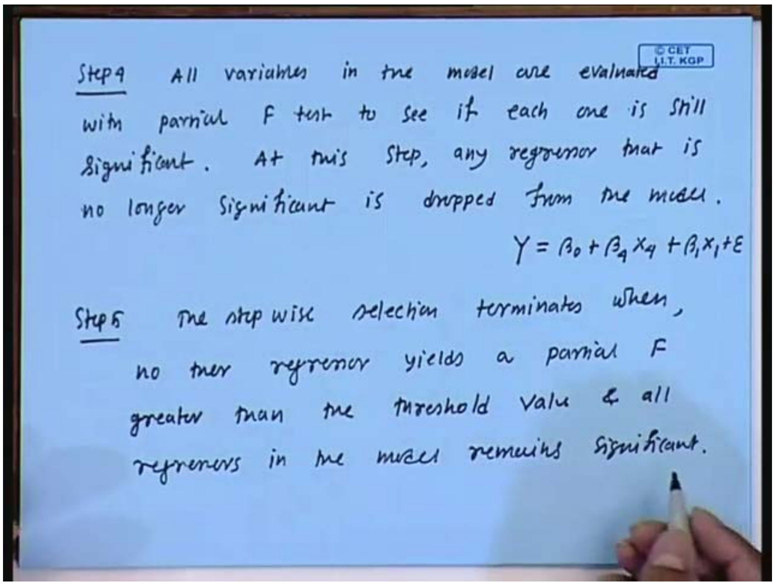
And then all possible models with one regressor are considered and F statistic value is computed, and the regressor having the highest F statistic value is added to the model.

So, there is no for stepwise selection there is no condition I mean we do not need to check the tabulated value of F for the first regressor in the model, so you do not need to check this condition, so the highest F statistic value is added to the model.

So, if you recall last example X 4 has the highest F statistic value and that had been included in the model, now the next step is the partial F statistics are computed for all of the remaining regressors in the presence of previously selected regressors. That means, let me recall the previous example X 4 was included first, and then you compute the partial F statistic for X 1 in the presence of X 2, partial F statistic for X 2 in the presence of X 1 and X 3 in the presence X 4.

And the one yielding the highest F is added to the model if the F is greater than the tabulated value, here you know maybe sometime instead of this tabulated value we can also consider specified threshold value. So, this value might be say for example, 5, so instead of finding the tabulated value every time we specify some threshold value here. If the observed F is greater than the specified threshold value then add the regressive variable in the model, well the first 2 steps are same now in step 4 we have change here.

(Refer Slide Time: 35:08)



Step 4 All variables in the model are evaluated with partial F test to see if each one is still significant. At this step, any regressor that is no longer significant is dropped from the model.

$$Y = \beta_0 + \beta_4 x_4 + \beta_1 x_1 + \varepsilon$$

Step 5 The step wise selection terminates when, no other regressor yields a partial F greater than the threshold value & all regressors in the model remains significant.

In step 4 what we do is that we check for a possible exit, what we do is that all variables in the model are evaluated with partial F test to see if each one is still significant. And at this step any regressor, that is no longer significant is dropped from the model, well I know no whether it is clear to you, but let me recall the previous example. So, in the first step we included X 4 the regressor X 4 in the previous example, and then in the second step we included X 1.

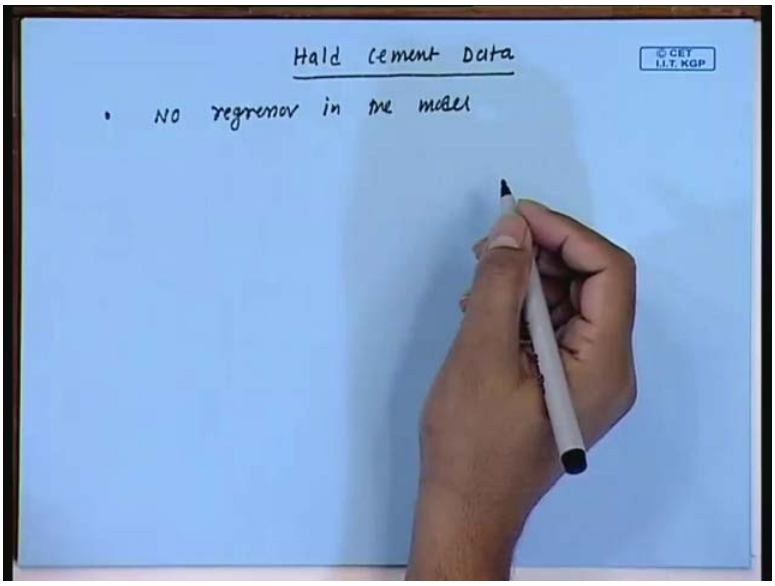
So, here it says that once we have this model say Y equal to beta naught plus first inclusion was beta 4 X 4, and then we added beta 1 X 1 in the model, so it says that see beta 4 was or X 4 was significant alone, when we consider one regressor model. Now, X 1 is significant in the presence of X 4 in the second step X 1 has been included in the model, now the question is no see of course, the X 1 is the last added regression in this model, so X 1 is of course, significant in the presence of X 4.

Now, we need to check whether X 4 is significant in the presence of X 1, so this is the difference, if it might be the case you know X 4 was significant alone, so at the first step what at this moment in the presence of X 1 X 4 might not be significant. So, we need to we need to, that is why it says that all variable in the models are evaluated with partial F test to see, if any if each one is still significant, so and at this step any regressor that is no longer significant is dropped from the model.

So, in the presence of X 1 if X 4 is not significant then we will drop X 4 from the model, so this is the difference, so this is called you know we check for the possible exit at step 4. And the next step is the stopping criteria, step 5 it says that the stepwise selection terminates, when no other regressor is partial F greater than the threshold value and all regressors in the model remains significant.

So, the first, so this includes two condition if there is no other regressor, which has significant partial F value in the presence of the regressors in the model. Then we cannot include any more regressor in the model, and at the same time the regressors, which are present in the model, if they remain significant then you do not need to remove one some regression from the model also. So, then we I mean then the stepwise selection algorithm terminates, well I want to explain or illustrate this stepwise selection algorithm using the HALD cement data again, so here we have a table.

(Refer Slide Time: 41:33)

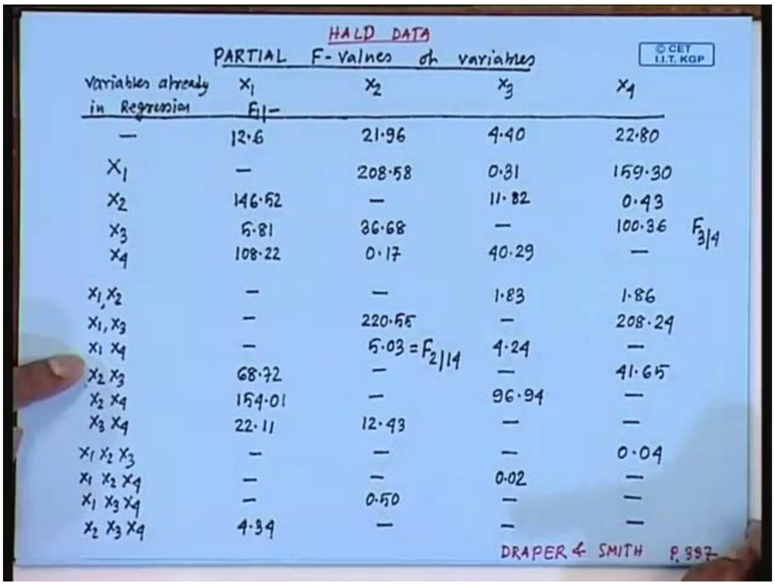


Hald cement Data

- No regressor in the model

So, we will be considering the HALD cement data, so the same technique, first you know no regressor in the model, now let me introduce one table I have one table.

(Refer Slide Time: 42:14)



HALD DATA

PARTIAL F-Values of variables

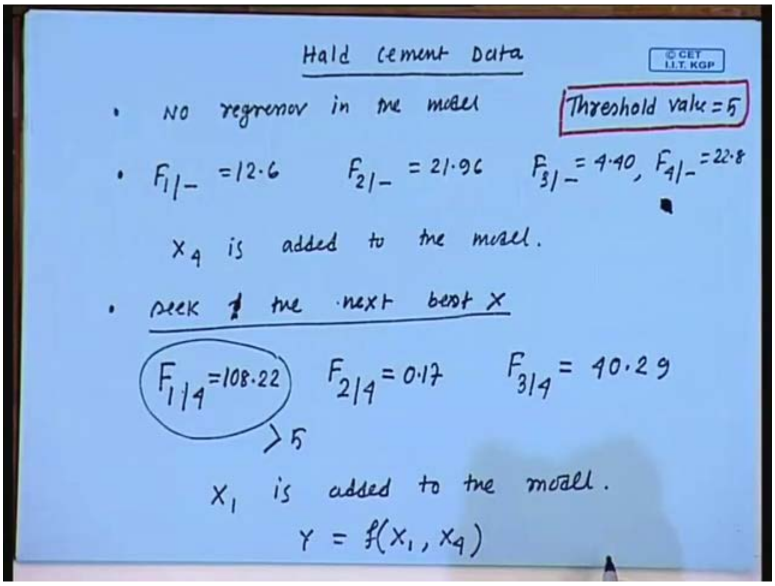
Variables already in Regression	x_1 $F_{1\|-}$	x_2	x_3	x_4
—	12.6	21.96	4.40	22.80
x_1	—	208.58	0.31	159.30
x_2	146.52	—	11.82	0.43
x_3	5.81	36.68	—	100.36 $F_{3\|4}$
x_4	108.22	0.17	40.29	—
x_1, x_2	—	—	1.83	1.86
x_1, x_3	—	220.55	—	208.24
$x_1\ x_4$	—	$5.03 = F_{2\|14}$	4.24	—
$x_2\ x_3$	68.72	—	—	41.65
$x_2\ x_4$	154.01	—	96.94	—
$x_3\ x_4$	22.11	12.43	—	—
$x_1\ x_2\ x_3$	—	—	—	0.04
$x_1\ x_2\ x_4$	—	—	0.02	—
$x_1\ x_3\ x_4$	—	0.50	—	—
$x_2\ x_3\ x_4$	4.34	—	—	—

DRAPER & SMITH P. 337

This is from the Drapper and Smith book page number 337, so let me explain the significance of these values, it says that the partial a values for variable X 1. So, this 12.6 is the F value for the model Y equal to beta naught plus beta 1 X 1 cursive sigh, and there is no other regressor already present in the in the model. So, X 1 is the only regressor in the model, so this one is basically what we used the notation F 1 in the presence of nothing, and this one is F 2 in the presence of nothing F 3 in the presence of nothing F 4 in the presence of nothing, so if you can recall this value 22.8.

Now, the meaning of this value is this one is basically F 2 in the presence of X 1 in the model, so this is the partial statistic value associated with X 2 in the presence of X 1 in the model. Let me consider these value, so this F value is the partial F statistic value associated with X 4 in the presence of X 3 in the model, so this one is basically F 3 4, so this show I have all the partial F statistic maybe I will say again this is a known figure 5.03. So, this one is the partial F statistic for the random for the regressor X 2 in the presence of X 1 and X 4 in the model, so this one is nothing but F 2 in the presence of 1 and 4 in the model, so this is I mean we have all the partial F values.

(Refer Slide Time: 44:54)



Hald cement Data

- No regressor in the model

Threshold value = 5

- $F_{1|-} = 12.6$ $F_{2|-} = 21.96$ $F_{3|-} = 4.40$, $F_{4|-} = 22.8$

X_4 is added to the model.

- seek the next best X

$\boxed{F_{1|4} = 108.22}$ $F_{2|4} = 0.17$ $F_{3|4} = 40.29$

> 5

X_1 is added to the model.

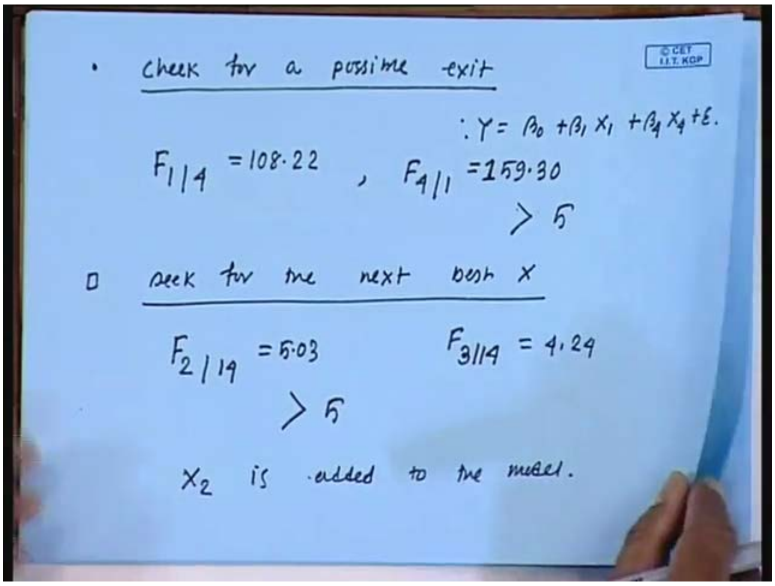
$Y = f(X_1, X_4)$

Now, I illustrate the stepwise selection, so initially there is no regressor in the model then the same like you know for the forward selection. We have the value F 1, which is equal to 12.6, F 2 in the presence of nothing which is equal to 21.96, F 3 in the presence of nothing, which is equal to 4.40, and F 4 in the presence of nothing, which is equal to 22.8. So, this one is the highest and also this is greater than, so it would not check any threshold or tabulate we do not compare it with the with the tabulated value, so since the F statistic associated with X 4 is highest, so X 4 is added to the model.

Now, next what we do is that we seek the next best X, so X 4 is already there in the model, so which one is the next best that can be added to the model. So, for that what we do is that we check the partial value for F 1 in the presence of X 4 in the model, we check the partial F value for regressor X 2, in the presence of X 4 in the model F 3 4.

And from this table from this table X 4 in the presence of X 4, so this one is F 1 4, this is F 2 4, this is F 3 4, so F 1 4 is equal to 108.22, F 2 4 is 0.17, and F 3 4 is 40.29 for the highest partial F value is this one. And here you know we fix the threshold value equal to 5, I know here we do not want to check the tabulated value, every time and since this is greater than the threshold value 5, X 2 sorry X 1 is added to the model. Now, at this moment we have X 1 and X 4 in the model, so Y is in terms of X 1 and X 4 at this moment, now what we do is that we do not this is the difference here.

(Refer Slide Time: 48:35)



- Check for a possible exit

$\therefore Y = \beta_0 + \beta_1 X_1 + \beta_4 X_4 + \varepsilon.$

$F_{1|4} = 108.22$, $F_{4|1} = 159.30$

> 5

□ seek for the next best X

$F_{2|14} = 5.03$ $F_{3|14} = 4.29$

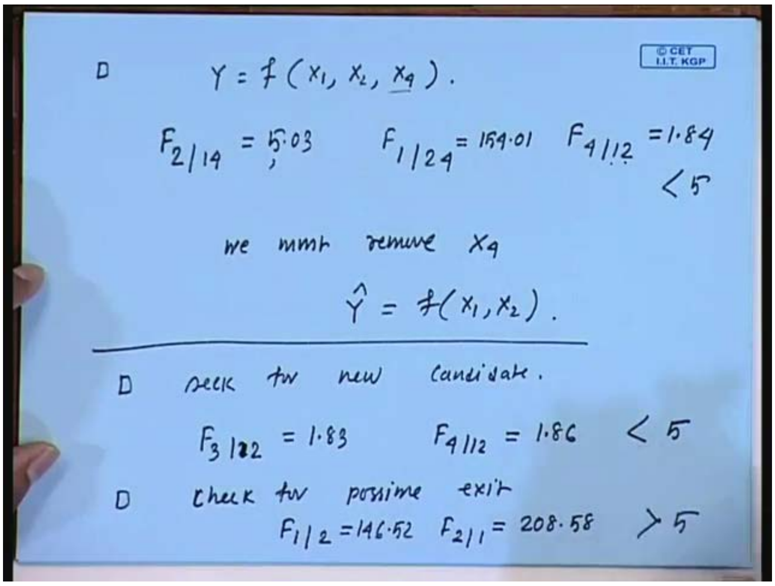
> 5

X_2 is added to the model.

Now, in the next step we check for a possible exit, so we have X 1 and X 4 in the model, so the model is Y equal to beta naught plus beta 1 X 1 plus beta 2 sorry, beta 4 X 4 plus epsilon. And now, what we do is that see of course, X 1 is significant in the presence of X 4 because the F value is F 1 4, when we compute this F 1 4 which is equal to 108.22, and the other one also we check F 4 in the presence of one which is equal to 159.03, see this is the last added relation. So, of course, this is you know, this one is significant in the presence of X 4 that is why it has been included I mean in the previous step. Now, the question is whether X 4 is significant in the presence of X 1, that we need to check and here, you will see that x the partial F value associated with X 4 in the presence of X 1 is also, it seems to be significant, because this is greater than the threshold value 5.

So, we do not drop any regressor from this model, so next step is that we seek for the next best regressor X, so I have X 1 and X 4 in the model, and both X 1 is significant in the presence of X 4 and also X 4 is significant in the presence of X 1. Now, we check the partial F statistic value with a 2 can be included in this model, so F 2 1 4 and also we check F 3 1 4. So, here is my table, so 1 4 is already there in the model, so F 2 1 4 is 5.03 and F 3 1 4 is 4.24 and this one is the highest partial value and also this one is greater than the threshold value 5, so X 2 is added to the model.

(Refer Slide Time: 52:27)



D $Y = f(x_1, x_2, \underline{x_4})$.

$F_{2|14} = 5.03$ $F_{1|24} = 154.01$ $F_{4|12} = 1.84$ < 5

we must remove X_4

$\hat{Y} = f(x_1, x_2)$.

D seek for new candidate.

$F_{3|12} = 1.83$ $F_{4|12} = 1.86$ < 5

D check for possible exit

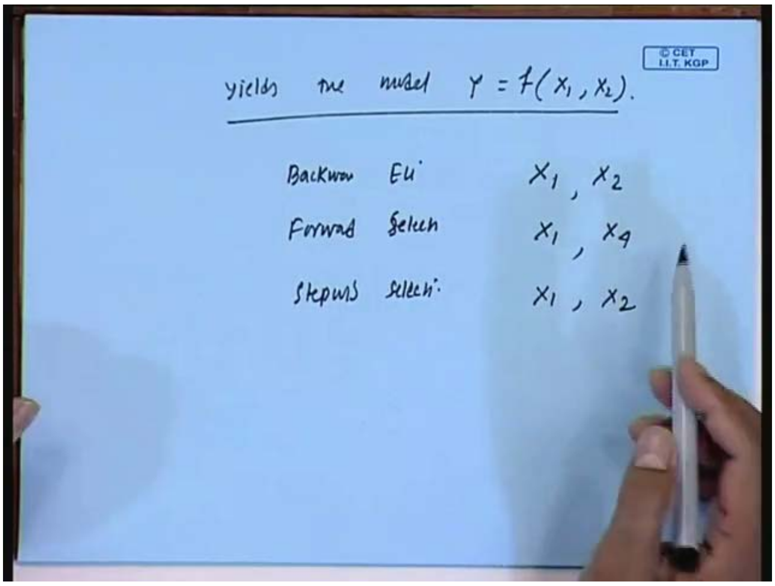
$F_{1|2} = 146.52$ $F_{2|1} = 208.58$ > 5

Now, I have the model with in my model we have 3 regresors, Y equal to F in terms of X 1, X 2, X 4, now we need to check you know X, whether the newly inserted. Of course, F we compute the partial value F 2 1 4, we compute F 1 2 4 we compute F 4 1 2, so this is to check you know this one is 5.03, this one is 154.01 and this one is equal to 1.84, that you can check from this table. And now, see before X 4 was significant, now in the presence of S 1 S 4 was significant, that is what we proved in the last slide you know see in the presence of X 1, X 4 was significant.

Now, in the presence of X 1 and X 2 X 4 is not significant, this is less than 5, but in the presence of this 2 random variable this one is significant, in the presence of these 2 regressors X 2 is significant. So, this two will be there in the model, but we need to remove we must remove X 4 from the model, so now, model is Y equal to F in terms of X 1, X 2, and again we know we check with the any candidate can be included we seek for new candidate. So, for that we compute F 3 1 2 and we compute F 4 1 2, these values are 1.83 and this is equal to 1.86 you can check from my table and both are less than 5.

So, cannot include anymore regressor in the existing model, which involves X 1 and X 2 now we need to check that whether X 1 is significant in presence of X 2 I mean we look for possible exit. So, we check for F 1 given 2 and F 2 given 1, this one is equal to 146.52 and this one is equal to 208.58 both are greater than 5, so you know we cannot remove anymore regressor from the model, and also this is that you cannot include anymore regressor to the model.

(Refer Slide Time: 56:10)



yields the model $Y = f(X_1, X_2)$.

Backwar Eli·	X_1, X_2
Forward Selecn	X_1, X_4
Stepwis Selecn·	X_1, X_2

So, the stepwise selection terminates and yields the model Y equal to F in terms of X 1 and X 2, so this is the result of stepwise selection. Now, just I want to mention that you know different algorithm, I mean the result of the different selection algorithms are different you know, if you can recall the backward elimination the result was X 1, X 2, final model was involving X 1, X 2. Now, the forward selection finally, selected regressors are X 1 and X 4 and for this stepwise selection, the final output is X 1 and X 2. So, that is you know the result are not (Refer Time: 57:25) that is what I want to say, and that is all about you know the model selection.

Thank you very much.