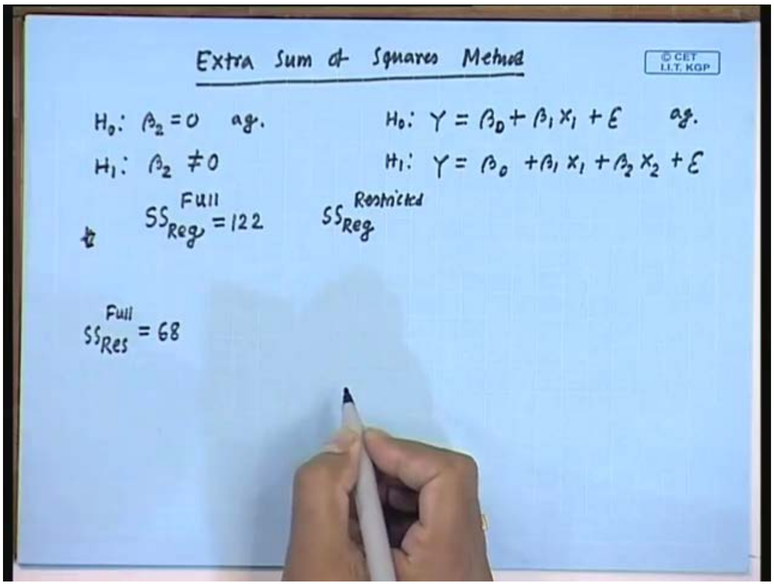
Regression Analysis
Prof. Soumen Maity
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 9
Multiple Linear Regression (Contd.)

Hi, this is 4th lecture on multiple linear regression, and know the last lecture we have considered one example on Multiple Linear Regression, there the model multiple linear regression model has been fitted. And the significance and the usefulness of the multiple linear regression model has been a tested using the global test, and also using the partial test, we have the tested, the significance of say for example, X 1 in the presence of X 2. Basically, the test statistic, we have used there is t statistic, well so this can be done also using the extra sum of square technique. So, I am going to talk you know, about how to test the significance of one regressor variable, in the presence of other regressor variable, using extra sum of squares technique well.

(Refer Slide Time: 02:23)

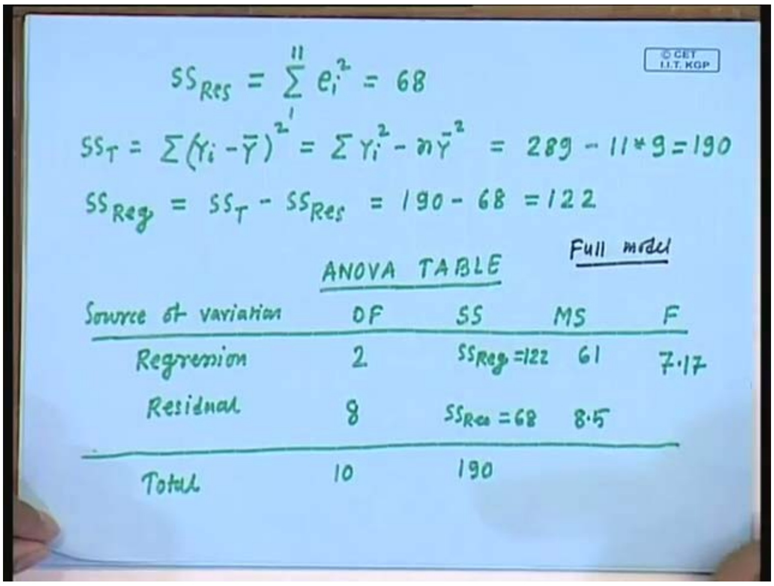


Well, so this one is you know, this one is usually used, to test for several parameters being 0, but here I am using this technique to test simple hypothesis like H naught beta 2 equal to 0, against the alternative hypothesis H 1, we says that beta 2 not equal to 0. So, this one, we can be I mean, we already we have tested this hypothesis, using the t statistic, but now we will be using you know extra sum of square technique to test this hypothesis.

Well so this hypothesis can also be written as you know H naught non hypothesis Y equal to beta 1 plus beta naught plus beta 1 X 1 plus epsilon against the alternative hypothesis, that is Y equal to beta naught plus beta 1 X 1 plus beta 2 X 2 plus epsilon. So, whether it is enough to consider, this model or it is necessary to consider the full model, so this is basically the full model and the non hypothesis says that, it is to go for the restricted model well.

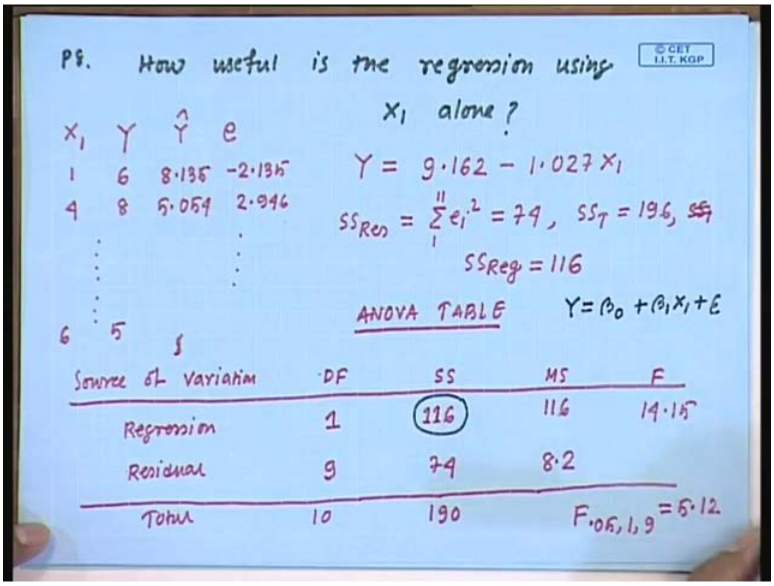
So, what we do in the extra sum of square technique is that, we compute S S regression both for the model like, for first we compute the S S regression, for the full model and also, we compute the value of the S S regression for the restricted model. So, here is the restricted model, so what is the S S regression value, when there is only one regressor, in the model and what is the S S regression value, when there are 2 regressor variables in the model.

(Refer Slide Time: 05:52)



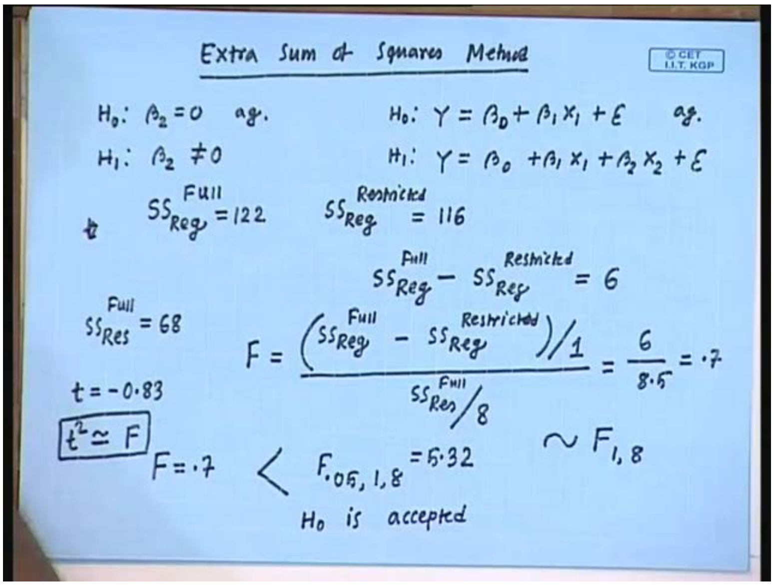
So, for these 2 values I will recall my previous class, you know this one is my ANOVA table in my previous class, for the full model, this one is for the full model, please refer my last class. So, here the S S regression value is a 122 and the S S residual value is 68, so I will copy these 2 values, S S regression is equal to 122 and also you know, S S residual for the full model is equal to 68.

(Refer Slide Time: 06:47)



Now, I will refer my previous lecture, for the restricted model, so here is the restricted model, I mean we have fitted this model, using only one regressor and here is the ANOVA table for my restricted model, that model here is called to is beta, Y equal to beta naught plus beta 1 X 1 plus epsilon and here is the S S regression 116 well.

(Refer Slide Time: 07:27)



So, S S regression for the restricted model is equal to 116, now the difference you know, we know that S S regression increases as the number of regressor variables increases. So, this is the S S regressor for the full model and this is the S S regression, for the restricted model, now the difference S S regression minus for the full model and S S regression for the restricted model is equal to 6.

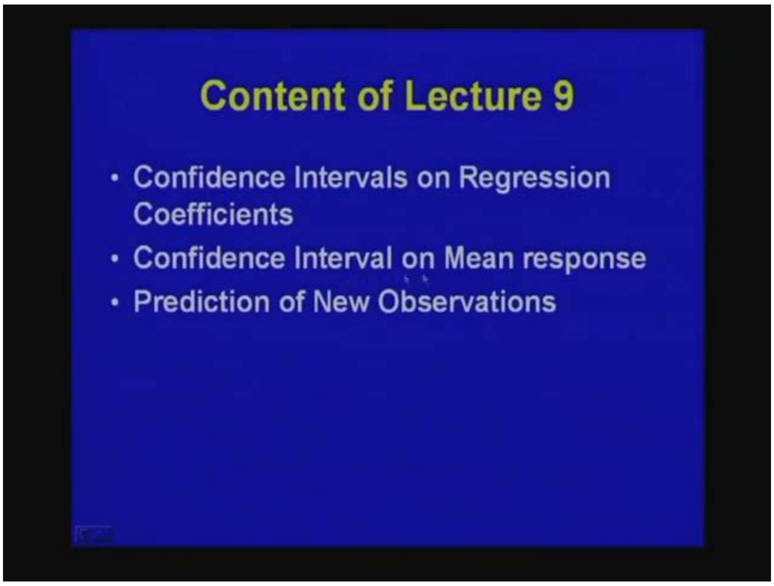
So, this one is you know, this one is called the extra sum of square, due to beta 2 or you can say that, this is the extra sum of square, some of square means, it is extra the regression sum of square, due to the regressor X 2. Because, this one this S S regression is involving both X 1 and X 2 and this S S regression is involving only X 1, so the difference is the S S regression, due to X 2. Now, the 8 statistic for this extra sum of square technique is you know, S S regression for the full model minus S S regression for the restricted model by the degree of freedom.

Here the degree of freedom of this one is only one, because here beta 2 is not a affecter, it is just only one regression coefficient and it is corresponds to X 2, so F is this by S S residual for the full model by the degree of freedom. The degree of freedom is equal to 8, if you can you can refer my previous class, well so this is equal to 6 by this is nothing but M S residual, which is equal to 8.5, which is equal to 0.7 and we know that this, if statistic follows F distribution with degree of freedom 1 and 8.

So, what we do is that, we compare the observed value of F, which is equal to 0.7 with the tabulated value F 0.0518, the value of this one is 5.32. So, the observed value is less than the tabulated value; that means, the conclusion is that, the hypothesis H naught is accepted. So, the meaning of this one is that, you know that, we accept the hypothesis that beta 2 equal to 0, the meaning of this one is the regressor variable X 2 is not significant, in the presence of X 1 in the model.

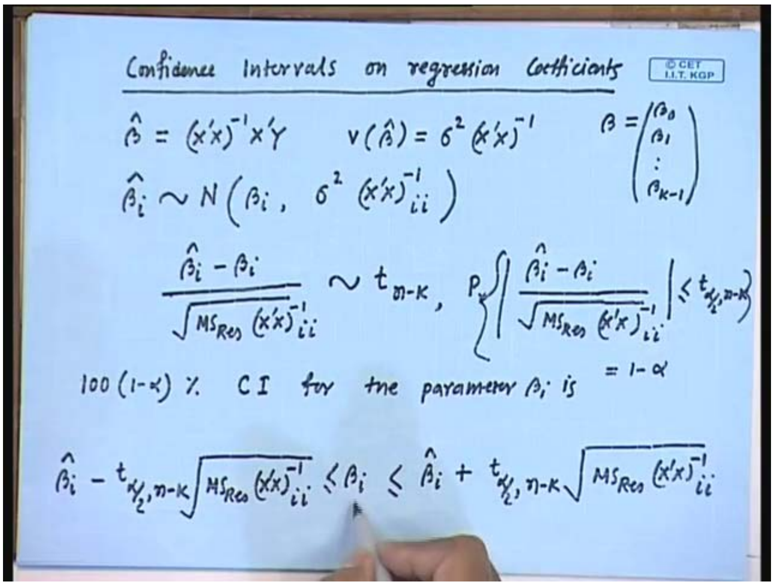
So, basically we got the same result, I mean we concluded the same thing in using that t test also, so this is another way to you know, this is another way to do the same testing, this is you know using the extra sum of square technique well. Also just I want to mentioned here is that the, if you use the t statistic then the t statistic value is equal to minus 0.83. So, here you can check that, t square value is equal to is almost, you know same equal to F and this is you know in general, this is true. So, whether you go, for the t test or you use the extra sum of square method, to test this hypothesis, you will be getting the same result, well next basically the content of today’s lecture is you know will be talking about, confidence interval on regression coefficients.

(Refer Slide Time: 13:27)



Also the confidence interval on mean response and ones the model has been fitted you know, it is very important you know, one important issue is to predict prediction of new observation, for a given value of regressor variable well.

(Refer Slide Time: 13:59)



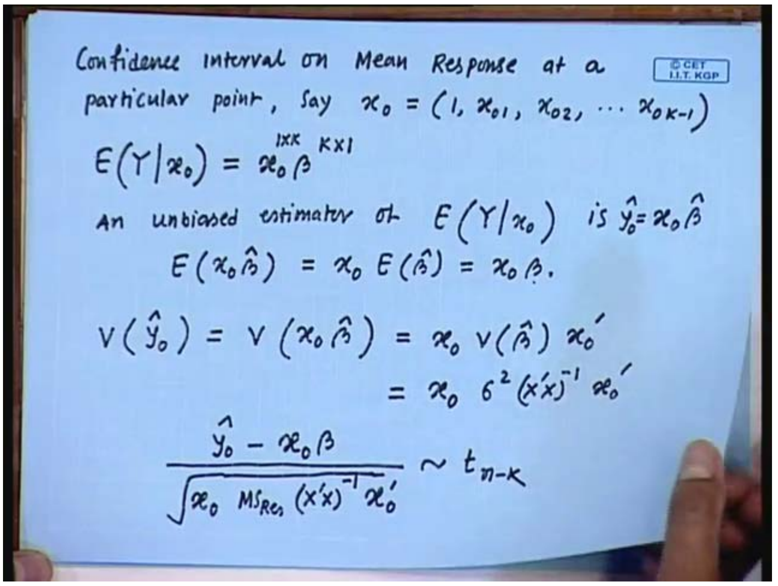
So, next we talk about, confidence intervals on regression coefficients, so here the regression coefficient is beta, which is vector beta naught beta 1 up to beta K minus 1 right. And you want to find you know the confidence interval for beta i for any i well, what we know is that to find the confidence interval, first you need to have the point estimator of beta, we know that, beta hat is equal to X prime X inverse X prime Y, this is an unbiased estimator of beta.

So, basically this one is n is a point estimator and also, we know that, the variance of beta hat is equal to sigma square X prime X inverse, now from here, we can say that beta i hat. So, I am talking about, the i th regression coefficient, this follows normal distribution with mean beta i and the variance sigma square X prime X inverse i i. So, this one is you know, this notation, we have used several time, this is the i i th element in X prime X inverse.

Now, from here, I can write beta i hat minus beta i root over, I am replacing this is sigma square by M S residual X prime X inverse i i, this random variable follows t distribution with degree of freedom n minus K, because we have K minus 1 regressor well. Then obviously, you know, we can say that, beta i hat minus beta i by this quantity M S residual X prime X inverse i i element of this the absolute value of this one is less than or equal to t alpha by 2 n minus K, this has probability equal to 1 minus alpha.

So, if you choose you know alpha equal to 0.05, this random variable is a absolute value, this random variable is less than t, alpha by 2 n minus K is 0.95. So, from here, we get 100 into 1 minus alpha percent confidence interval for the parameter b i is you know, you can get b i is in between b i hat plus t alpha by 2 n minus K, M S residual X prime X inverse i i and similarly the lower bound is beta i hat minus t alpha by 2 n minus K into M S residual X prime X inverse i i. So, this is the 95 percent confidence interval for the i th regressor point alpha is equal to 0.05, similarly you can get similar confidence interval, for the other regressor coefficient from for the other coefficient also.

(Refer Slide Time: 20:12)



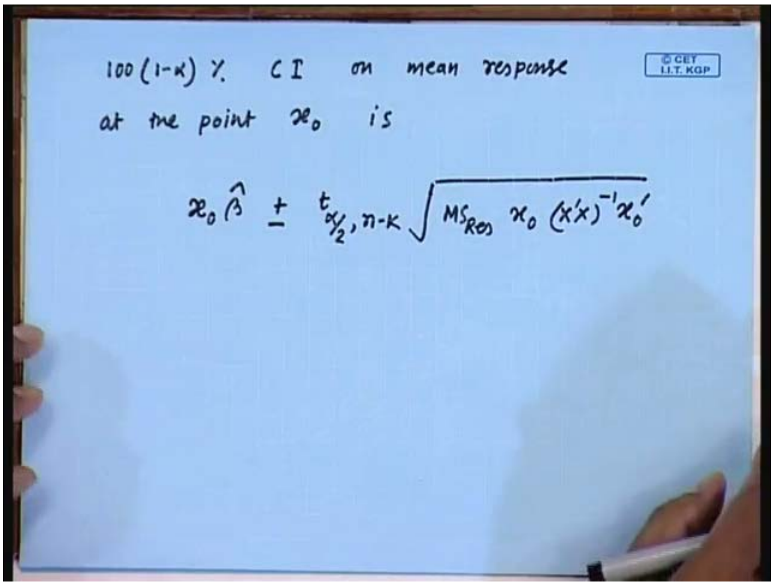
So, next will be talking about, confidence interval, on mean response at a particular point say x naught is equal to 1, x naught 1, x naught 2, x naught K minus 1 well, so what we want is that, we want the expected response value at this point. So, these are the this is the value of first regressor, second regressor and the K th regressor, so at this point, we are looking for the expected response value, so what we want to estimate first is that, we want to estimate the expected, so if you need we looking for the confidence interval for this expected value or mean response at the point x naught.

So, the usual technique, you know to find the confidence interval for this one, first you have to look for the point estimation of this one well, what is this quantity, this quantity is nothing but or this is nothing but x naught beta, so x naught is a 1 cross K vector and beta is a K cross 1 vector right. Well we know that an unbiased estimator of this expected response at the point X naught is for the unbiased estimator of this one is nothing but X naught beta hat is X naught beta hat and we call it say y hat y naught hat.

So, this one is an unbiased estimator of this quantity, because beta hat is an unbiased estimator of beta, so you can prove that, you know expected value of x naught beta hat is equal to x naught expectation of beta hat, which is equal to x naught beta. And the variance of next, we compute the variance of the unbiased estimator y naught hat, which is equal to the variance of x naught beta hat right. Now, this variance is equal to x naught the variance of beta hat into x naught prime well.

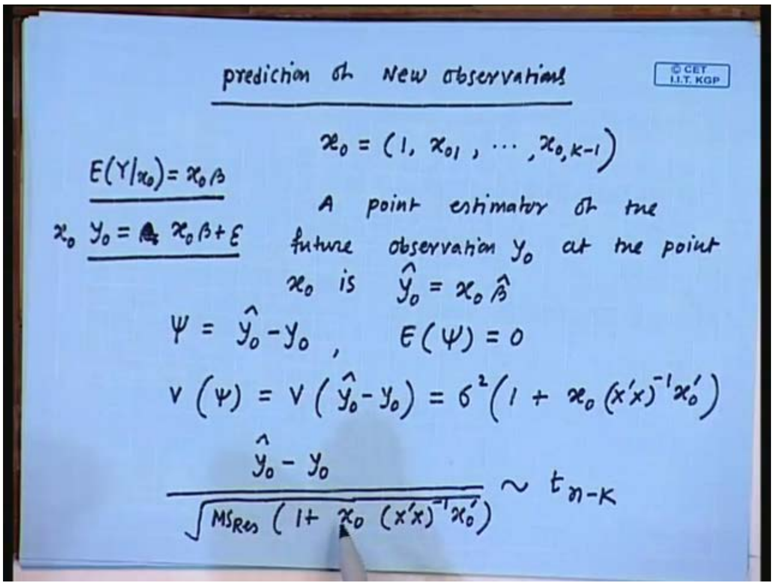
So, this one is nothing but we know the variance of beta hat is equal to sigma square x prime x, x inverse x naught prime well, so from here, I can say see my unbiased estimator y naught hat has expectation this and variance this. So, I can say that, y naught hat minus x naught beta, you know this by x naught M S residual, I am just replacing sigma square by M S residual x prime x inverse x naught prime, this quantity for this random variable, it follows t distribution with degree of freedom n minus K right. So, from here you know using this, I can give now 100 into 1 minus alpha percent confidence interval for the expected response, this is the expected response. So, i can get the confidence interval for this expected response now.

(Refer Slide Time: 26:33)



So, therefore, you know hundred into 1 minus alpha percent confidence interval on mean response at the point x naught is x naught beta hat, which is nothing but y naught hat plus t alpha by 2 n minus K into M S residual x naught X prime X inverse x naught prime. So, this is the upper bound of the interval and the lower bound is obtained by just replacing the plus sign by the by minus, so this one is basically, the confidence interval for the expected response at the point x equal to x naught well.

(Refer Slide Time: 28:10)



So, next we will be talking about, a prediction of new observation, well so this is a very important aspect, because you know, once you have the fitted model that, fitted model, if it is a significant 1 then that model can be used the regressor model can be to predict new observations, corresponds to particular value of the response variance well. So, here what you want is that, we want to predict the value new observation at the point at say x naught equal to 1 x naught 1 x naught K minus 1 well.

So, this one is bit for simple linear regression also, we had the same sort of problem, the difference between the expected response and the new observation at the point x naught is you know expected response at the point x naught is nothing but x naught beta. But, the observation here, what we looking for is that you know, we are trying to predict the future observation, y naught at the point x equal to x naught, I mean when x naught means, it is for the given values of the regressor variable well.

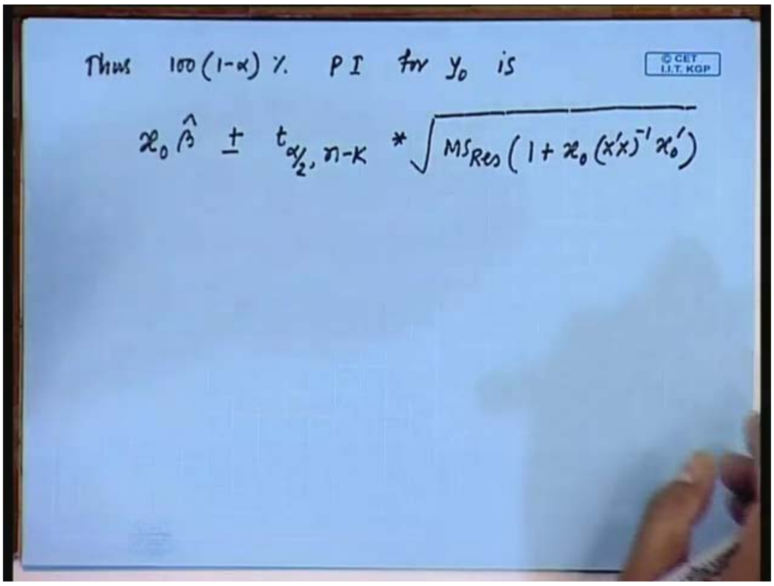
So, here we are trying to estimate y naught, this y naught is according to the model, this y naught is nothing but x naught beta plus epsilon, so in the previous case, we try to estimate this expected response and here, we are trying to predict the value of this one difference. But, the starting point will be same will start with the point estimation or point estimator of this one, so a point estimator of the future observation y naught at the point x naught is again, we call it y naught hat, which is equal to x naught beta hat.

Well so you see the difference here, you know you are starting with the same point estimator, so this point estimator, we have used to estimate the expected response well, but now, this point estimator is not an unbiased estimator of this thing, but because we have the excellent term here. So, that is why, we define new random variable chi, which is equal to y naught hat, the same strategy has the simple linear regression model, y naught hat minus y naught and the expected value of this random variable is equal to 0.

This is not difficult to check, because expectation of these is equal to 0 and the variance of this one, you can check that variance of this one is variance of chi is equal to the variance of y naught hat minus y naught, which is equal to sigma square 1 plus x naught x prime x inverse x naught prime. Let me explain see the first term is sigma square, which is basically the variance of y naught and the second term is the variance of y naught hat, you know why it is.

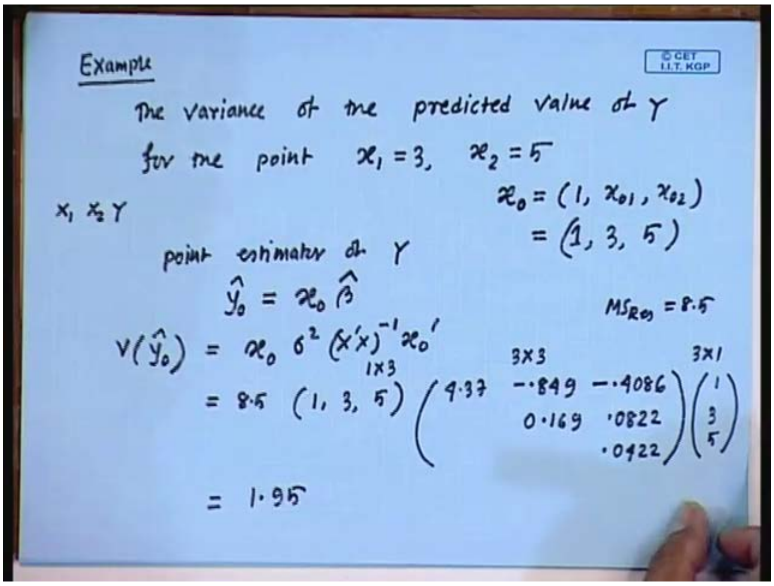
So, and these 2 are independent, because this is a new observation and this y naught hat, you know it consist of the previous observations the given observations y 1 y 2 y n and this one is an independent observation future observation right. So, from here, we can say that, y naught hat minus y naught by M S residual 1 plus x naught x prime x inverse x naught prime, so I just replace sigma square by M S residual. So, this random variable follows t distribution with degree of freedom n minus K right and from here from the distribution of this random variable, we can get the prediction interval, we call it prediction, the same thing confidence interval or prediction interval for y naught.

(Refer Slide Time: 35:15)



So, thus 100 into 1 minus alpha percent prediction interval, for y naught is equal to x naught beta hat, this is nothing but y naught hat plus minus t alpha by 2 n minus K into M S residual 1 plus x naught x prime x inverse x naught prime right. Well so this is how we get the prediction interval, for future observation, now just give one example to illustrate, this confidential interval or prediction interval, let me consider the same example.

(Refer Slide Time: 36:39)



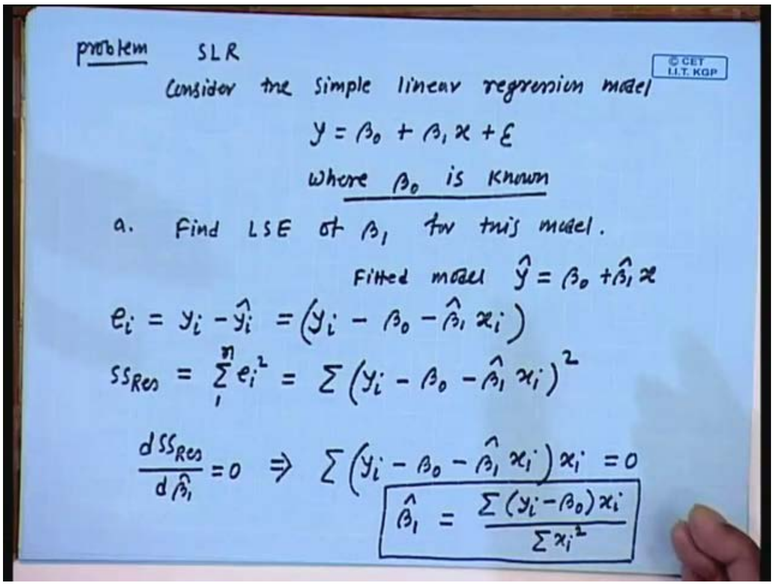
As in my last class and here I want to considered the problem that, you know, you find the variance of the predicted value of Y, for the point x 1 equal to 3 and x 2 equal to 5. So, if you can recall my last example, there we had 2 regressors x 1 x 2 and Y, now what this problem says you that know, you find the variance of predicted value of y for when x 1 equal to 3 and x 2 equal to 5, so here my x naught is what I said x naught is you know, 1 x naught 1 x naught 2, so this one is nothing but 1 3 5, this is my x naught right.

Now, what is the point estimation of estimator of y, the point estimator of y call it y naught hat, which is nothing but x naught beta hat and what we want here is that, we want variance of this point estimator. So, the variance of y naught hat is equal to x naught the variance of beta hat, which is equal to sigma square X prime X inverse x naught prime right. Now, to estimate you know, we cannot compute the variance, because sigma square is not known, so you have to replace this sigma square by M S residual, so sigma square is equal to 8.5, because M S residual value is equal to 8.5.

Now, my x naught is you know, this is a scalar quantities, now my x naught is 1 3 5 and you know that x prime x is x prime x inverse, which is equal to 4.37 minus 0.849 please refer my last lecture, minus 0.4086 0.16 9.082 2.0422 into 1 3 5. May just I am giving this example, so that you know, just to illustrated just to illustrate, whatever the theory, I just explain. And from here, you know it is not now not difficult to check that, this is you know this is 1 cross 3, this one is 3 cross 3 and 3 cross 1, so ultimately a value of this one is going to be 1.95.

So, this is the estimated variance of the predicted value of Y at this point you know, once you have this variance, you can find the confidence interval or prediction interval very easily. So, that is all regarding the confidence interval or prediction interval in multiple linear regression, next I have some time, so I want to solve one problem from simple linear regression, this will I hope this will help you, to understand, you know more on one degree of freedom, how to calculate the degree of freedom. So, constitute this problem from the simple linear regression.

(Refer Slide Time: 42:23)



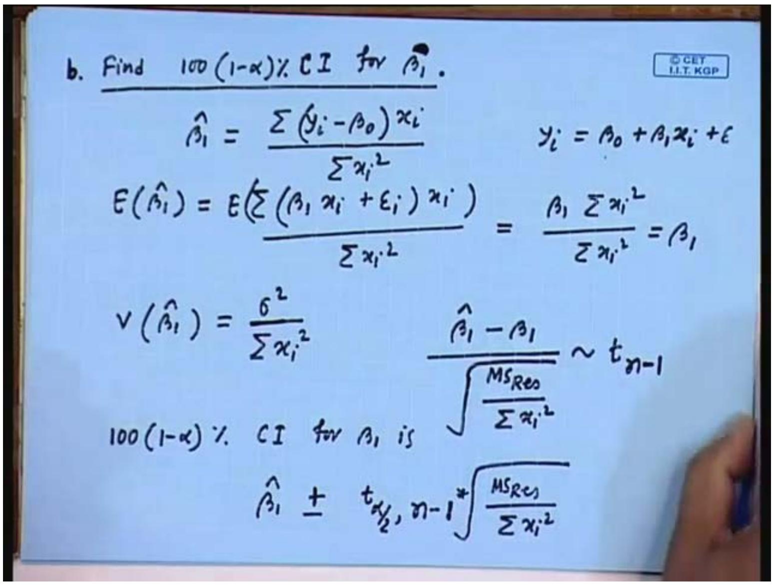
In the simple linear regression, well so the model is y equal to beta naught plus beta 1 x plus epsilon as usual, here we say that beta naught is known, so this is the only difference. So, you really you know do not need to estimate beta naught, because you know beta naught value is given, so only thing you need to do is that you know, you need to estimate the value of the slope, that is you know beta 1. So, the first problem, it says that find list of square estimator of beta 1, for this model well you know.

So, you need to understand that opposite that, you know to estimate word to find the list square estimator of beta 1, you need to I mean that will be obtained by minimizing the S S residual. So, you assume that the fitted model is you know, y equal to beta naught plus beta 1 hat x plus it is x, so suppose your fitted model is y hat, which is equal to beta naught plus beta 1 hat x. So, I am not putting see, I did not put beta 1 had because, beta naught hat, because beta naught is known well.

So, my e i is equal to y i minus y i hat, so here it is y i minus beta naught minus beta 1 hat x i, this is my i th residual, now S S residual sum of square, which is equal to some e i square 1 to n, this is going to be y i minus beta naught plus beta 1 hat x i square well. So, the list square estimate of beta 1 can be obtained by minimizing the S S residual and here, you have only one unknown parameter, so you just differentiate S S residual with the expect to beta 1 hat and this equal to 0 will give you the normal equation.

So, the normal equation is y i minus beta naught plus beta 1 hat x i, which is basically e i into x i equal to 0, so here you will get only one normal equation, because there is only one unknown parameter and solving this normal equation, we will get the estimator of beta 1 hat. So, from here, we get beta 1 hat is equal to summation y i minus beta naught x i by summation x i square right. So, this is the list square estimator of beta 1 hat, so I mean, what I want to say is that you know, you here you do not need to differentiate with respect to beta naught, because beta naught is known, so you do not need to estimate that.

(Refer Slide Time: 48:00)



So, the next problem no find problem b, it says find hundred into 1 minus alpha confidence interval for beta 1 hat, so to find this one, we need to confidence interval for beta 1, it is not beta 1 hat find 100 into 1 minus alpha percent confidence interval for beta 1. So, you know that an unbiased, we do not know whether it, so will start from the list square estimator of beta 1 that is beta 1 hat, which is equal to summation y i minus beta naught x i by summation x i square.

Well so this is a point estimator of beta 1, now will check, what is the expected value of this beta 1 hat, beta 1 hat is equal to you put, what is y i, y i is equal to beta naught plus beta 1 x i plus epsilon. So, if you put this will be replaced by beta 1 x i plus epsilon i into x i by summation x i square and the expected value of this one, so here you can check that you know, the second term is 0, because expected value epsilon i is equal to 0. So, this is going to be beta 1 into summation x i square by summation x i square.

So, it is an unbiased estimator of beta 1 and also you can check that the variance of beta 1 hat is equal to sigma square by summation x i square right. So, from these 2, I can write you know, beta 1 hat minus beta 1 by if I replaced this sigma square by M S residual by M S residual some over x i square is this follows t distribution with degree of freedom, so the here is the here, I want to discuss little bit it is degree of freedom n minus 1.

Because there are degree of freedom of S S residual is n minus 1 by S S residual is summation e i square i equal to 1 to n and you have the freedom of choosing n minus 1 e i and the last one the n th one i has to be choose an in such a way, there is such that, it satisfy the constant that e i x i equal to 0. So, here you have to note that, you know there is only one constant on e i.

So, that is why you are losing, the degree of freedom by 1 and thus the degree of freedom residual is equal to n minus 1, it is not n minus 2, in case of you know usual for the simple linear regression model, when both beta naught and beta 1 are unknown the degree of freedom, for this one is n minus 2. Here since beta naught is no is beta naught is known, we are defined setting with the respect to only beta 1 and we get one normal equation, which is nothing but e i x i equal to 0. So, there is only one constrain on residuals. So, that is why the degree of freedom is n minus 1. So, now, I can write 100 into 1 minus alpha percent confidence interval, for beta 1 is beta 1 hat, I write just plus minus t alpha by 2 n minus 1, M S residual by summation x i square. So, that is all for today.

Thank you very much.