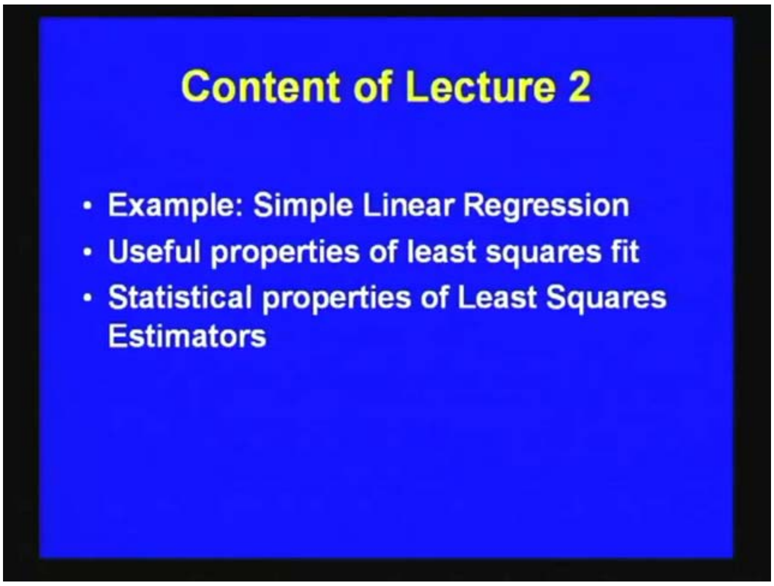
Regression Analysis
Prof. Soumen Maity
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 2
Simple Linear Regression

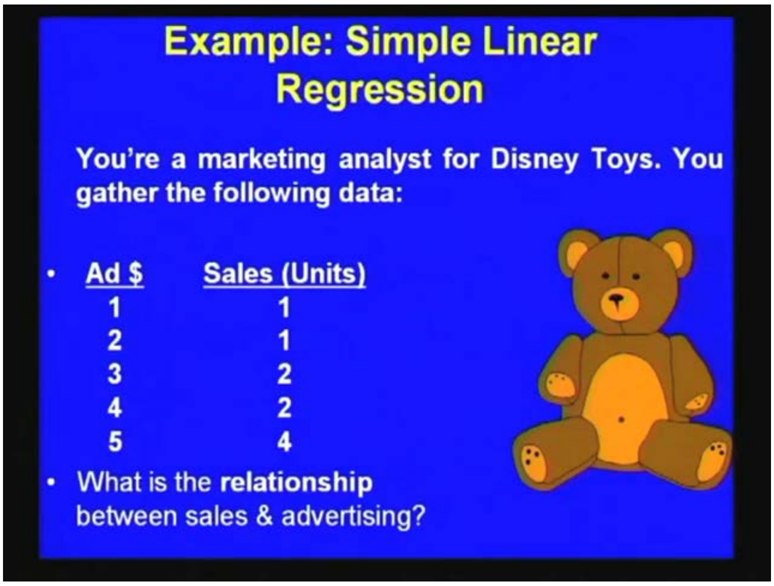
Hi, this is my second lecture in module one and on simple linear regression. In the first lecture, we have introduced simple linear regression model and they have learnt how to estimate the regression coefficients using least square technique.

(Refer Slide Time: 01:25)



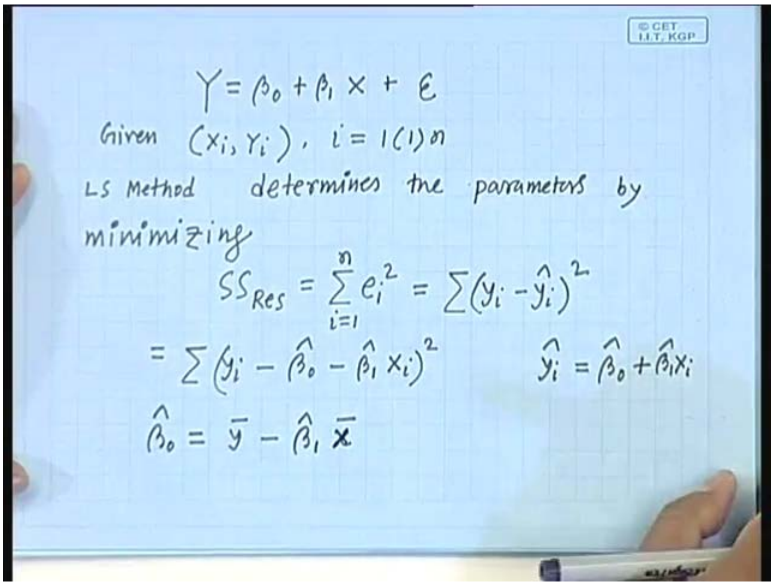
Here is the content of today's lecture first will give one example on simple linear regression. And then we talk about useful properties of least square fit and then the statistical property of least square estimator.

(Refer Slide Time: 01:53)



Well let me just recall the simple linear regression model. The general form of simple linear regression model is a Y equal to beta naught plus beta 1 X plus epsilon.

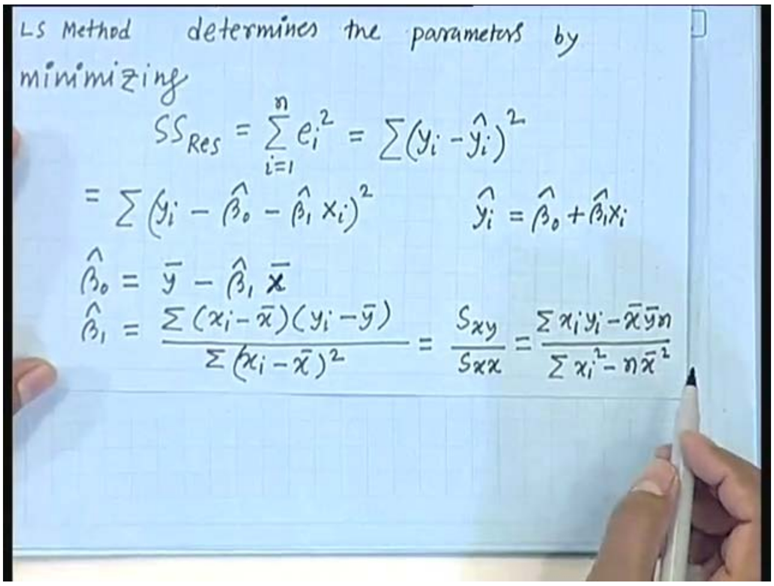
(Refer Slide Time: 02:08)



Were Y is the response variable X is the regressor variable and epsilon is the error term and beta naught is intercept and beta 1 is slope and we call beta naught into beta 1there regression coefficients. Now, given a set of observations say X i Y i for i equal to 1 to n we learned how to estimate the regression coefficients beta naught and beta 1 using least square technique. So, least squares method determines the parameters means the regression coefficients beta naught and beta 1 by minimizing residual sum of squares SS residual. Which is equal to e i square i equal to 1 to n, which is basically equal to the difference between the observed response value and the estimated response value y i hat.

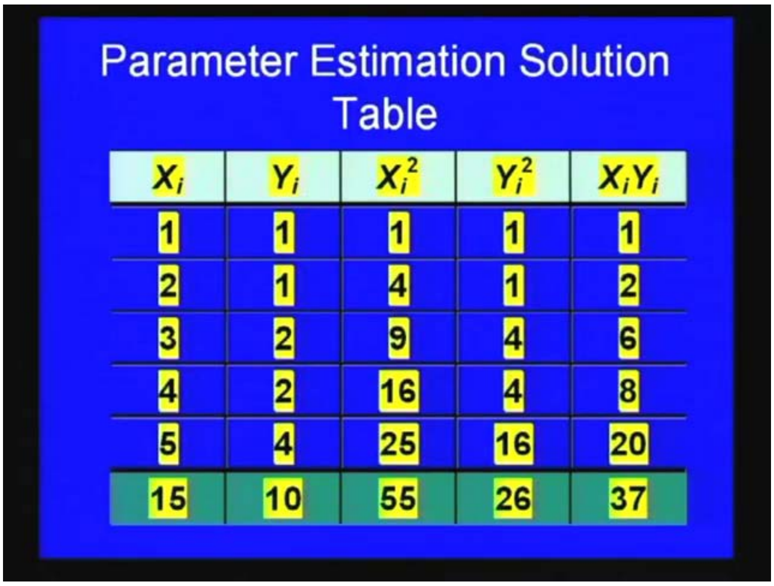
So, this is the residual and y i hat this is y i minus t i hat is the is equal to beta naught hat plus beta 1 hat X. So, this is beta naught hat minus beta 1 hat X i whole square. Because the fitted equation is y i hat equal to beta naught hat plus beta 1 hat X i. So, we learned that beta naught I mean this SS residual or residual sum of squares is minimum. When beta naught hat is equal to Y bar minus beta 1 hat X bar.

(Refer Slide Time: 05:53)



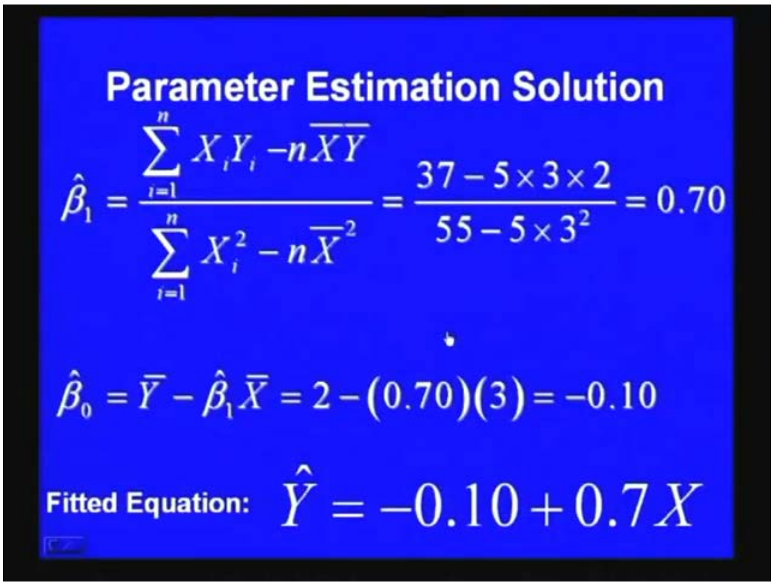
And beta 1 hat is equal to summation x i minus x bar y i minus y bar by summation x i minus x bar whole square. Well this quantity is also denoted by the symbol S x y by S x x and this one can also be written in the form summation x i y i minus x bar y bar n times by summation x i square minus n into x bar square. This is not difficult to observe just a simple algebra shows that this quantity is equal to this quantity. So, let us move to the disney toy example this is the cost on advertising and this is the sales amount, so these are the x i values and this are the y i values. And even to fit a straight-line model to this data this table shows.

(Refer Slide Time: 08:02)



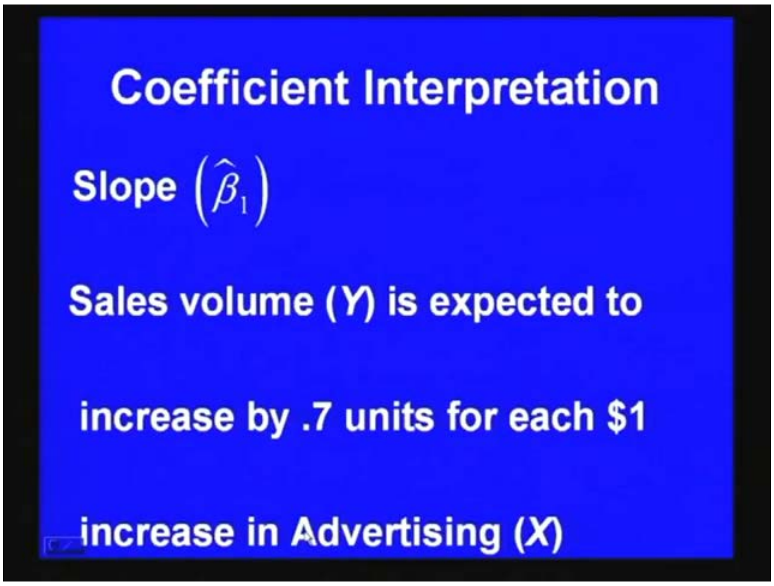
Summation X i is equal to 15 summation, Y i is equal to 10 summation, X i square is equal to 55 summation, Y i square is equal to 26 and summation X i Y i is equal to 37.

(Refer Slide Time: 08:31)



And then we can compute the value of beta 1 hat. So, beta 1 hat is equal to summation X i Y i minus n times X bar into Y bar by summation X i square minus n times X bar square which is equal to 0.7. And similarly beta naught hat is equal to Y bar minus beta 1 hat X bar which is equal to minus 0.1. Well so the fitted equation is Y hat equal to beta naught which is equal to minus 0.1 plus 0.7 into X. So, this is the fitted equation for the given observations, now what is the interpretation of this regression coefficient beta 1. So, it says that that the expected value of the response variable or.

(Refer Slide Time: 10:03)

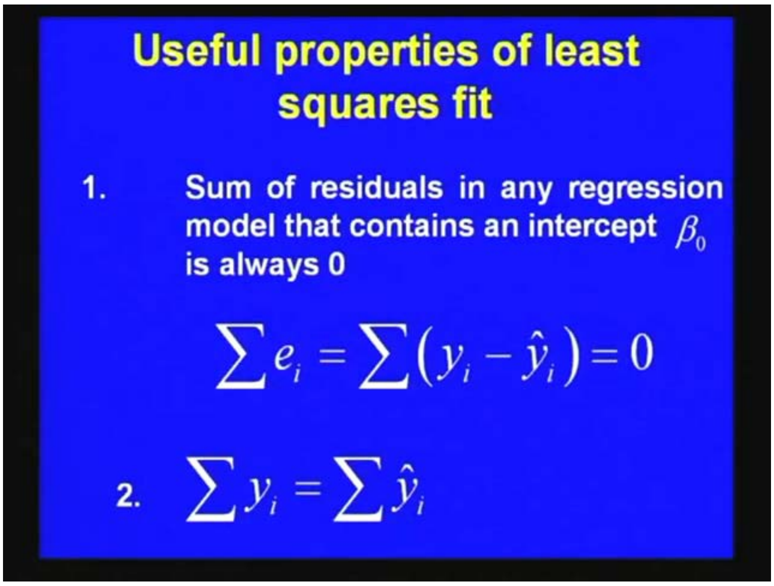


The expected sales amount is increased by 0.7 units for each 1 dollar increase in advertising and the interpretation of the coefficient beta naught hat is that.

(Refer Slide Time: 10:17)

What is the average sales amount when X is equal to 0 beta naught is equal to Y hat, when x is equal to 0. So, beta naught beta naught hat this basically gives some idea about the average value of sales amount when the advertising cost is equal to 0. So I mean its very difficult explain why it is so because see we can expect some sales amount without advertising also. But here it is negative anyway, so the beta naught hat is the average sales volume which is equal to minus 0.1 when the advertising is equal to 0.

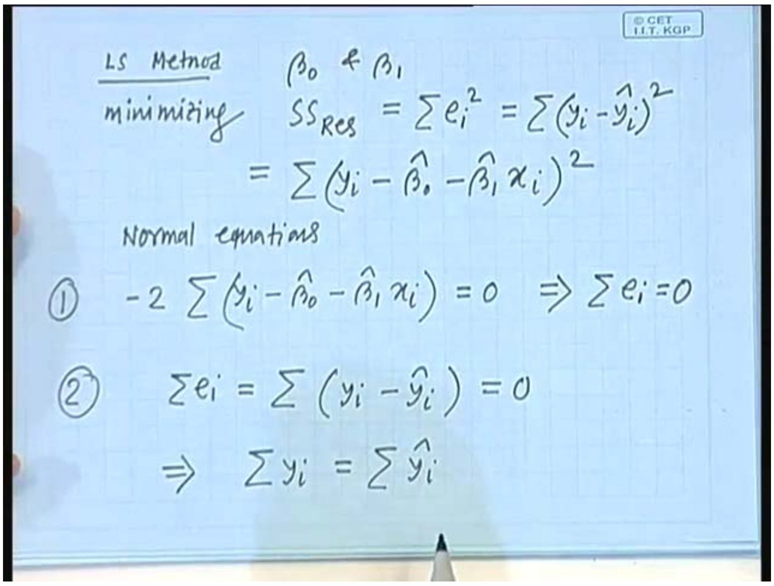
(Refer Slide Time: 11:24)



Well next we move to the useful properties of the least square fit. It says that we know what is residual is e i is the difference between the observed response value and the predicted or the estimated response value. So, it says that the sum of residuals in any regression model that contains a intercept beta naught is always 0. I am going to prove this one, so this summation e i is equal to 0.

And the second property says that the sum of the observed value is equal to the sum of the fitted values y i hat. So I mean this second property is consequence of the first property it is obtained from this one only. Let me prove the first property which says that the some of residuals in any regression model that contains an intercept beta naught is always equal to 0. So, what it to do is that we have to recall the.

(Refer Slide Time: 13:00)



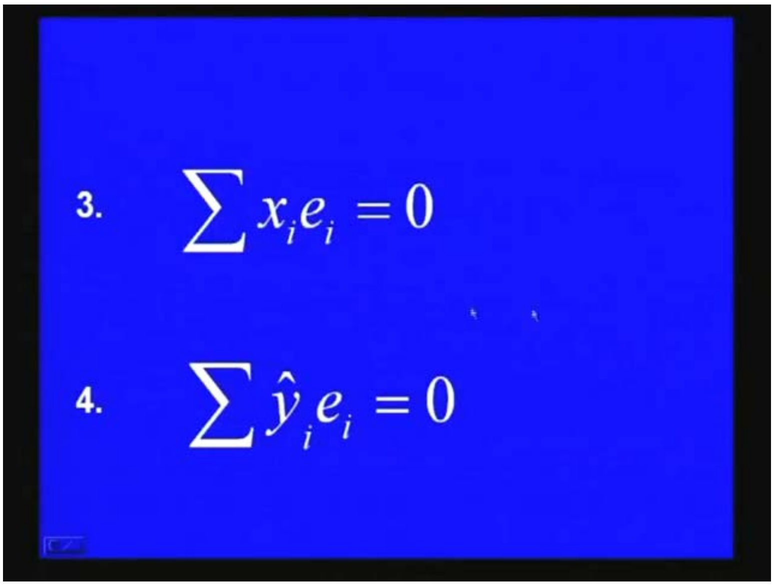
L S method least squares method least squares method determines the parameter beta naught and beta 1 by minimizing by minimizing SS residual that this residual sum of square. Which is equal to summation e i square and this is basically equal to summation y i minus y i hat square. And this is equal to summation y i minus beta naught hat minus beta 1 hat x i square well.

So, least squares method or L S method determines the parameter beta naught hat and beta 1 hat by minimizing this quantity. So, what we do is that we differentiate the residual sum of square with respect to beta naught hat and we equate that with 0. We did the first normally equations the normally equations differentiate SS residual with beta naught hat that gives one normally equation.

And again we differentiate residual sum of squares with respect to beta 1 hat that gives another normally equation the first normally equation. We differentiate SS residual with respect to 0, sorry with respect to beta naught hat that gives summation y i minus beta naught hat minus beta 1 hat x i 2 times of this and one negative sign here. So, and we equate this 2 equal to 0 this is the first normally equation and which is nothing but see.

This is nothing but y i minus y i hat, so this one is nothing but summation e i equal to 0. To the first normally equation says that the sum of the residual is equal to 0. And this residual sum of residual e i is nothing but summation. So, this is the first property the second property says that well this is residual some of residual is equal to y i minus y i hat this is the residual ith residual. And we know this is equal to 0 from the first property, which implies that summation y i is equal to summation y i hat. So, this one is the consequence of the previous property or the first property it says that the sum of observed values equal to is equal to the sum of the fitted value well. So, next will move to the third property.

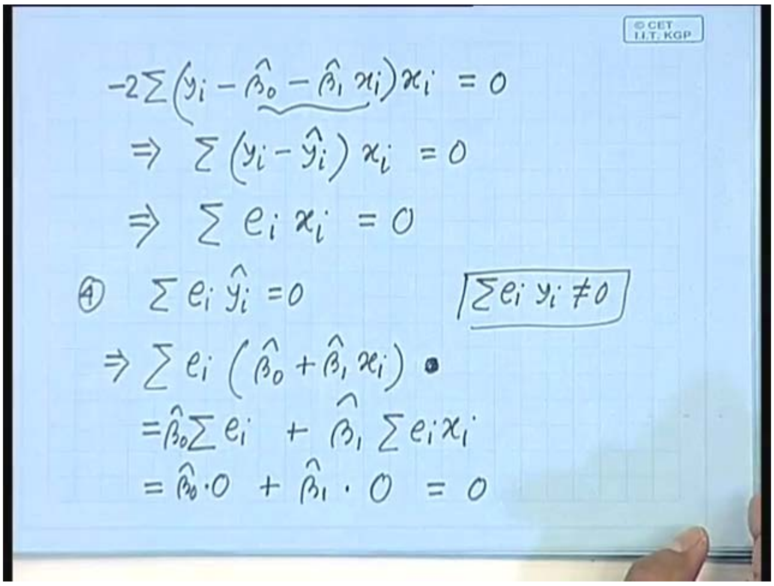
(Refer Slide Time: 16:52)



It says that summation x i e i equal to 0, that means sum of residuals waited by the corresponding value of the regressor variable is equal to 0. So, sum of residual e i waited by the corresponding value of the regression variable this is the value of the ith regression variable this is equal to 0. And the fourth property says that summation y i hat e i is equal to 0. That means the sum of residuals waited by the corresponding fitted value of the response variable is equal to 0, well let me prove this two properties.

Property three and property four well by differentiating this residual sum of square with respect to beta naught hat. We got this normally equation this is the first normally equation, now again we differentiate this normally equation sorry this some of residual with respect to beta 1 hat.

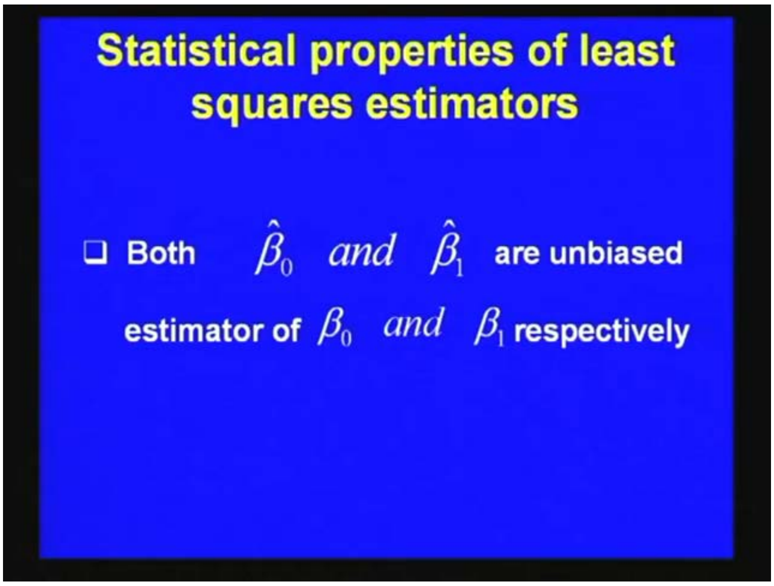
(Refer Slide Time: 18:28)



And that gives summation y i minus beta naught hat minus beta 1 hat x i and we are differentiating with respect to beta 1 hat. So, this quantity will be multiplied by minus 2 into x i, so minus 2 into x i which is equal to 0. This is the second normal equation and this one is nothing but summation y i. And this part is y i hat into x i equal to 0, which implies summation e i x i equal to 0 which is the third property. Now, the fourth property is summation e i y i hat this is equal to 0. But, summation e i y i is not equal to 0 you should note that how to prove that summation e i y i hat equal to 0.

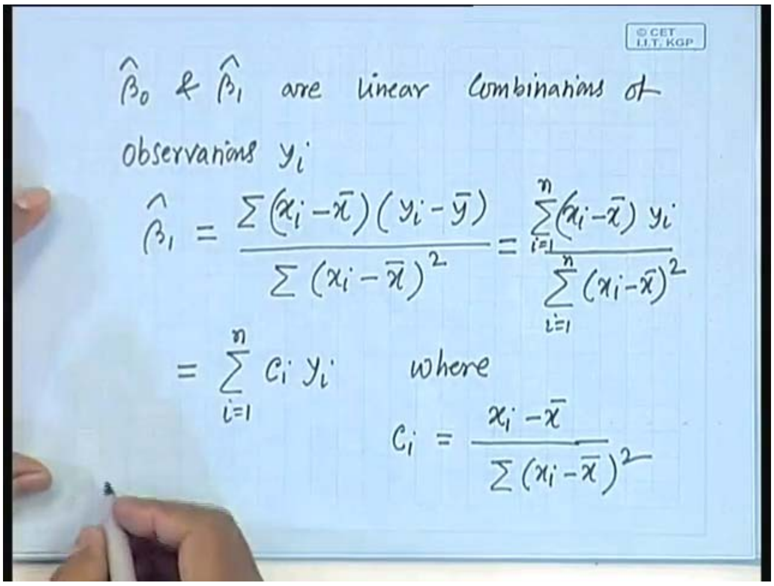
This is again consequence of the first property and the third property well this can be written as summation e i what is a y i hat y i hat is equal to beta naught hat plus beta 1 hat x i which is equal to 0. Which is equal to which is equal to summation beta naught hat e i plus beta 1 hat summation e i x i. Now, see summation e i from the first property this is equal to 0. So, this is equal to 0 plus the third property says that summation e i x i equal to 0. So, this is basically beta naught hat into 0 plus beta 1 hat into 0. So, this is equal to 0, so you prove that summation e i y i hat equal to 0. Ok, so these are the some properties of least square fit and will be using them in the future.

(Refer Slide Time: 21:30)



Next move to the statistical properties of least square estimators. So, we have estimated the regression coefficient beta naught and beta 1 using least square estimators. And we are going to prove that both beta naught hat and beta 1 hat are unbiased estimator of beta naught an beta 1 respectively well let me prove that.

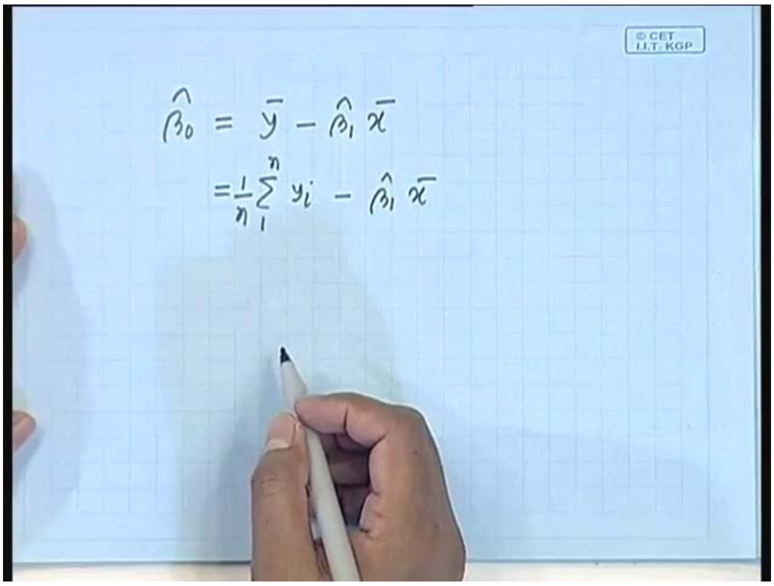
(Refer Slide Time: 22:20)



First I will prove that beta naught hat and beta 1 hat are linear combinations of observations y i, that means we want to say that they are linear estimator. Ok for example, considered beta 1 hat , so what is beta 1 hat beta 1 hat is equal to summation x i minus x bar into y i minus y bar by summation x i minus x bar whole square. Now, I said that this is linear combination of the response variables y i s other observation y i s. This can be written as this can be written as summation x i minus x bar into y i only by summation x i minus x bar whole square this is it i mean this two quantity are same.

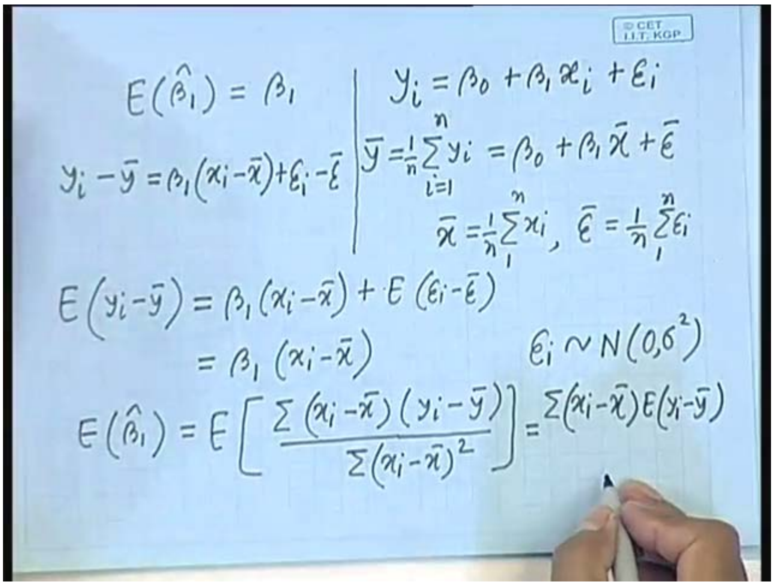
One can very easily prove that summation x i minus x bar into y bar that quantity is equal to 0. This is not difficult to prove that now this is from i equal to 1 to n and i equal to 1 to n. Now, this can be written as summation C i y i were C i is equal to here i equal to 1 to n were C i is equal to x i minus x bar by summation x i minus x bar whole square. So, I prove that beta 1 hat is a linear combination of the observations y i. Similarly, one can prove that beta naught hat is also easily what is beta naught hat beta naught hat.

(Refer Slide Time: 25:38)



Beta naught hat is equal to y bar minus beta 1 hat x bar. So, this is also a linear combination of the observation. Because this is nothing but this is nothing but summation y i from i equal to 1 to n 1 to n and minus beta 1 hat x bar. We already prove that beta 1 hat is a linear combination of the observation y i and the first term is also a linear combination of y i. So, the whole thing is a linear combinations of the observations y i. So, this just prove that the estimator we got they are linear estimator they are linear in y i next let me prove that the estimator beta naught and beta 1 they are.

(Refer Slide Time: 26:55)

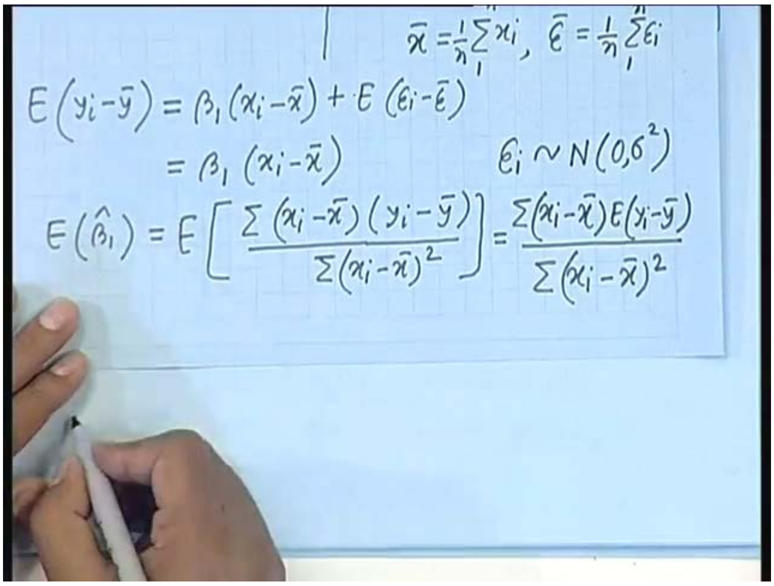


Unbiased estimator that is we are going to prove that expectation of beta 1 hat is equal to beta 1. So, if this is true then we call beta 1 is an unbiased estimator. Well let me start from here the simple linear regression model is y i is equal to beta naught plus beta 1 x i plus epsilon i. Now, y bar y bar which is equal to summation y i i equal to 1 to n. This is equal to beta naught plus beta 1 x i will be replaced by x bar and epsilon i will be replaced by epsilon bar. So, were x bar were x bar is of course, equal to summation x i 1 to n and epsilon bar is also equal to summation epsilon y summation epsilon y 1 to n.

Then from here y i minus y bar is equal to y i minus y bar is equal to beta 1 into x i minus x bar plus epsilon i minus epsilon bar. So, the expected value of the expected value of y i minus y bar is equal to. See one thing you should observe that you should always remember that y is a random variable. The response variable y is a random variable, but x is not a random variable it is its controlled variable. So, for given i this is just a constant, so expected value of y i minus y bar is equal to beta 1 x i minus x bar plus expectation of epsilon i minus epsilon bar.

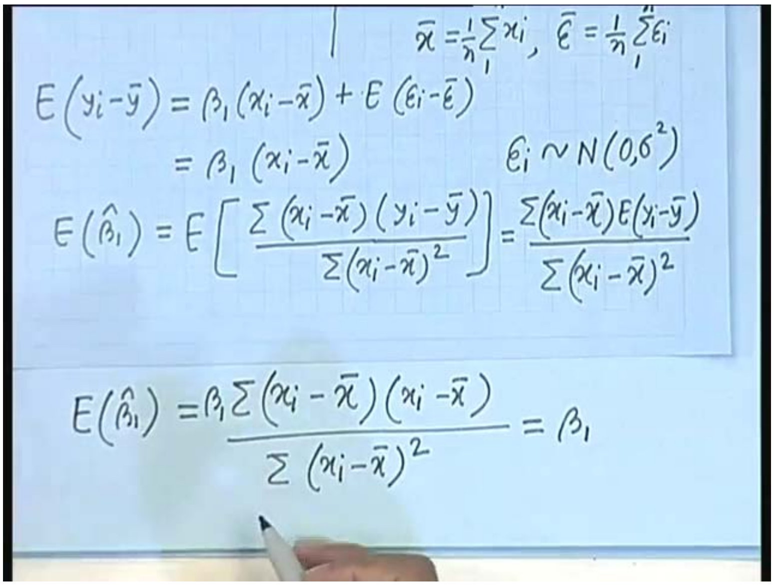
Now, epsilon i is the random error which follows which is a random variable and this follows we are assumed that epsilon i follows normal 0 sigma square. So, the expected value of e i is equal to 0 and similarly the expected value of epsilon bar is also equal to 0. The expected value of y i minus y bar this term is going to be 0 is equal to beta 1 into x i minus x bar. Now, our aim is to prove that beta 1 hat is an unbiased estimated of beta 1, so expectation of beta 1 hat is equal to expectation of what is beta 1 hat beta 1 hat is equal to summation x i minus x bar y i minus y bar by summation x i minus x bar whole square. So, this one is nothing but summation x i minus x bar into expectation of y i minus y bar by summation x i minus x bar whole square.

(Refer Slide Time: 32:37)



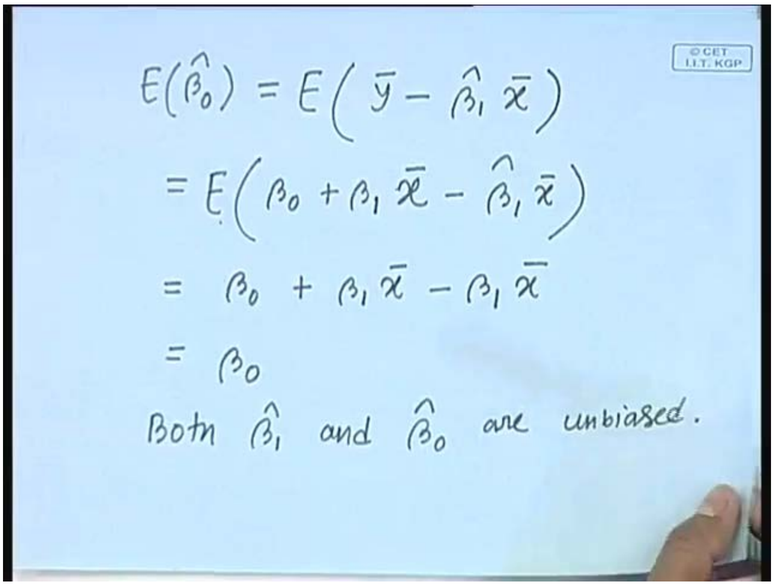
Now note that we proved that expectation of y i minus y bar is equal to beta 1 into x i minus x bar, so expectation of beta 1 hat is equal to summation x i minus x bar.

(Refer Slide Time: 33:04)



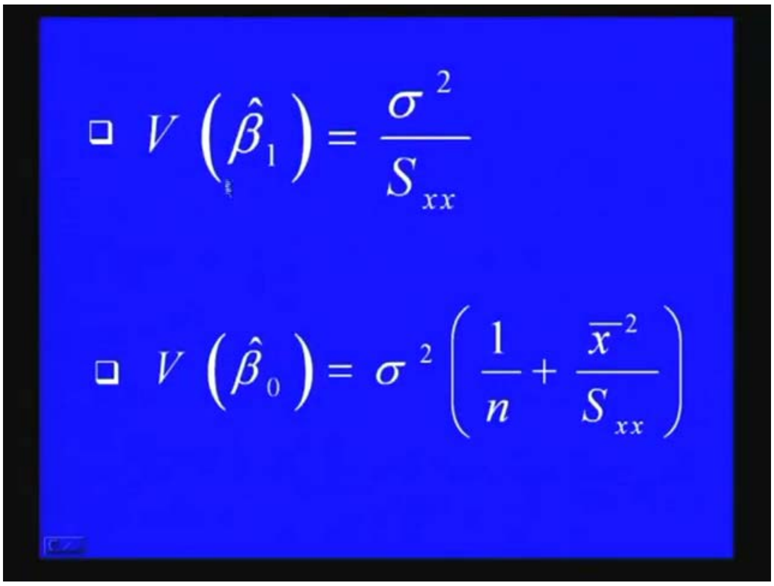
And this quantity is equal to beta 1 beta 1 x i minus x bar by summation x i minus x bar whole square. So, this is nothing but beta 1 only because this one is nothing but x i minus x bar whole square. So, he proved that beta 1 is an beta 1 hat is an unbiased estimated of beta 1 similarly, next we prove that beta 1 hat.

(Refer Slide Time: 33:53)



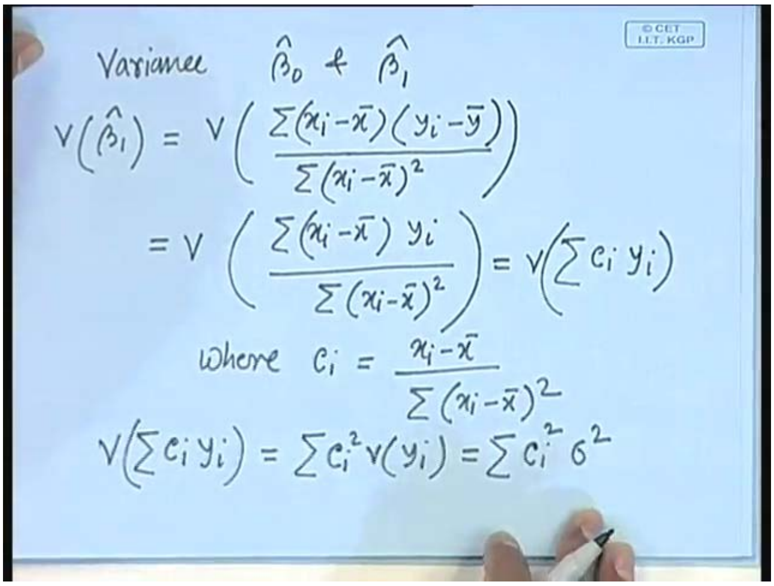
Sorry beta naught hat is also unbiased that means we are going to prove that expectation expected value of beta naught hat is equal to beta naught, this is equal to expectation of y bar minus beta 1 hat x bar. So, what is y bar y bar equal to we proved in the previous light y bar is nothing but beta naught plus beta 1 x bar minus beta 1 hat x bar. So, this is equal to expectation of beta naught is beta naught expectation of beta 1 x bar is equal to beta 1 x bar. Now, the expectation of beta 1 hat just now we proved that expectation of beta naught beta 1 hat is equal to beta 1 hat sorry beta 1. So, this is equal to beta 1 x bar which is equal to beta naught, so we prove that both beta 1 hat and beta naught hat are unbiased.

(Refer Slide Time: 36:01)



Next, will talk about the variance of the variance of beta 1 hat and beta naught hat the variance of beta 1 hat is equal to sigma square by S x x and variance of beta naught hat is equal to sigma square into 1 by n plus x bar square by S x x. Anyway I mean we need to know how to how to derive these things well.

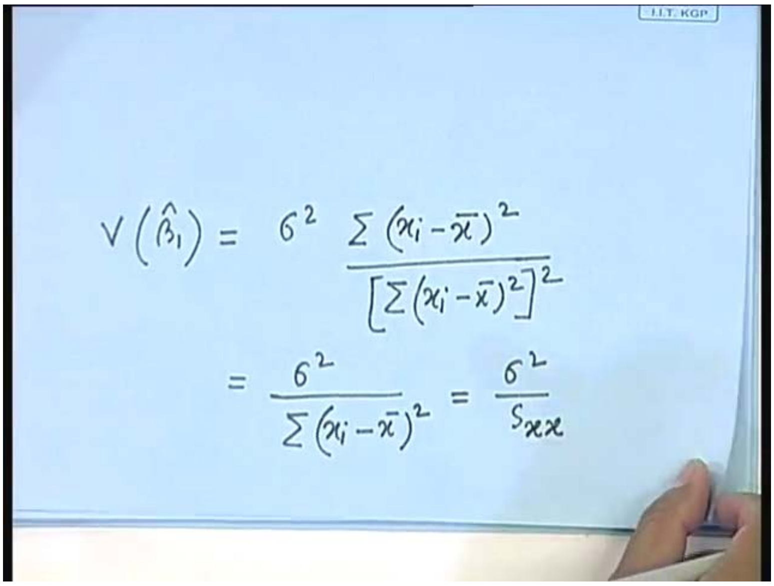
(Refer Slide Time: 36:36)



So, next we talk about the variance of variance of beta naught hat and beta 1 hat. So, the variance of beta 1 hat is equal to variance of what is beta 1 hat beta 1 hat is equal to summation x i minus x bar into y i minus y bar by summation x i minus x bar whole square. Well now this can be written as the variance of summation x i minus x bar into y i by summation x i minus x bar whole square. And this is nothing but see I before also I proved that this estimator is a linear combination of the observation y i.

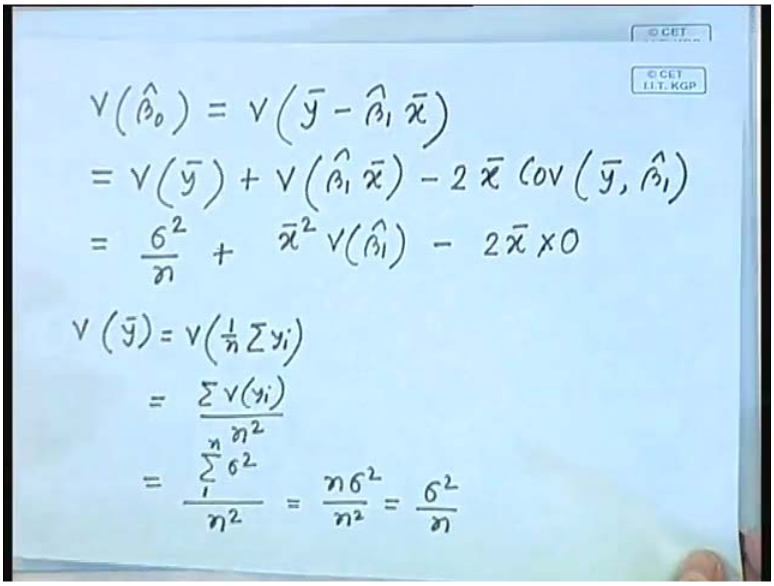
So, the variance of this quantity is and since y i y i is are independent you know well this is equal to variance of summation C i y i, were C i were C i is equal to x i minus x bar by summation x I minus x bar. Now, we know that the variance of the variance of summation C i y i is equal to summation sum of variances C i square variance of y i. This is true because because y i s are y i s are independent and we know that variance of y i is equal to sigma square this is equal to summation c i square sigma square. Now, what is summation C i square summation C i square we know this is this is C i. So, variance of beta 1 hat is equal to sigma square into summation C i square which is equal to summation x i minus x bar C i square.

(Refer Slide Time: 40:36)



So, summation x i minus x bar whole square by summation x i minus x bar whole square and the square of the whole thing. So, just replacing C i square by its value and this becomes sigma square by summation x i minus x bar whole square, which is equal to sigma square by S x x by notation. So, he proved that variance of beta 1 hat is equal to sigma square by S x x.

(Refer Slide Time: 41:54)

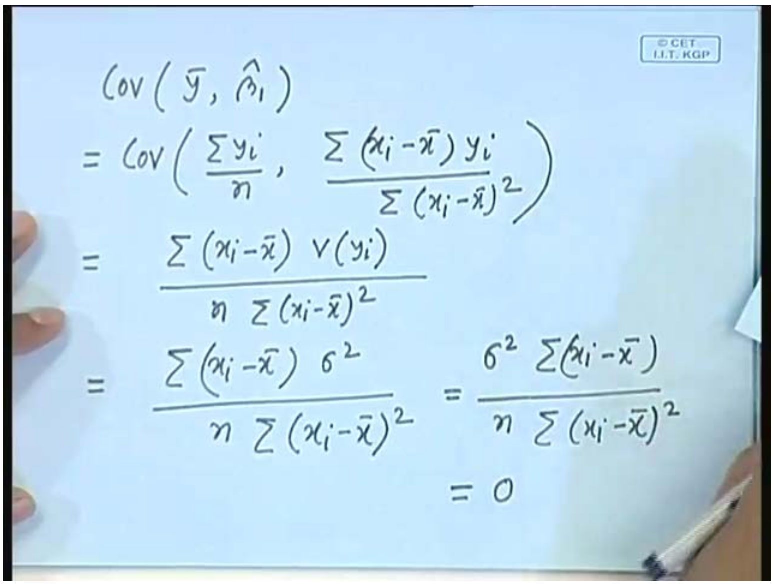


So, next we talk about the variance of beta naught hat variance of beta naught hat is equal to variance of y bar minus beta 1 hat x bar right. Now, this variance can be written as variance of y bar plus the variance of beta 1 hat x bar minus 2 times x bar the covariance of y bar and beta 1 hat right.

Now, you know the variance of y bar i hope you know it it is equal to its not difficult to prove that this is equal to sigma square by n. Let me prove this variance of y bar is equal to the variance of 1 by n summation y i know y i is an independent. That is assumption we made at the beginning because e i is an independent. So, epsilon i s are independent y i s are also an independent the variance of this quantity is equal to summation of variance of y i by n square right and variance of y i is equal to sigma square.

So, summation sigma square n times 1 to n y n square, so this is basically n sigma square by n square. This is equal to sigma square by n the variance of y bar is equal to sigma square by n, we know the variance of beta 1 hat and this is the constant quantity. So, the variance of this quantity is x bar square into the variance of beta 1 hat which is equal to which is equal to sigma square by S x x just we proved. Now, what about this covariance this is going to be this co-variance is going to be 0 this is 2 x bar into 0 into 0, but we need to prove this one. Ok the covariance between this can be proved that the covariance is equal to 0.

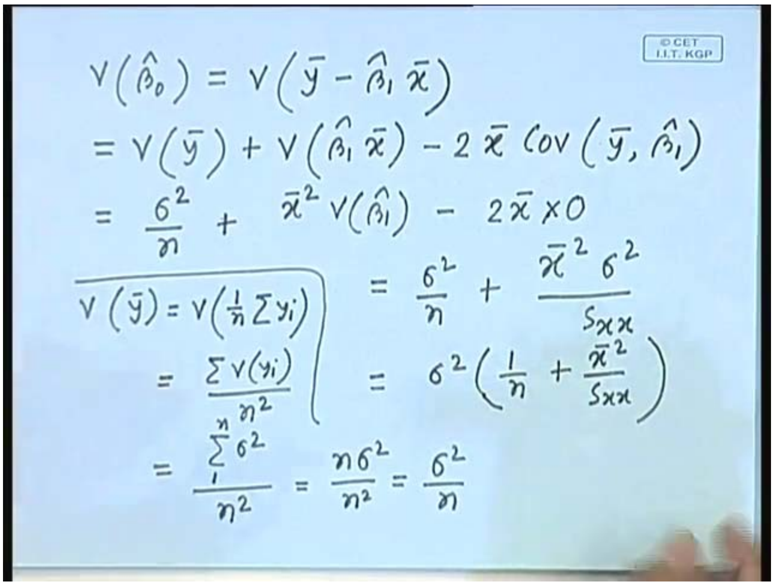
(Refer Slide Time: 45:07)



The covariance between y bar and beta 1 hat is equal to the covariance between what is y bar y bar is equal to summation y i by n. And beta 1 hat is equal to summation x i minus x bar into y i by summation x i minus x bar whole square right. Now, this covariance is see y i are independent, so the covariance between y i and y j is equal to 0 when i is not equal to j. So, this is nothing but the summation of x i minus x bar the covariance between y i and y i. So, which is nothing but the variance of y i right by summation x i minus x bar whole square and 1 n here. This is equal to this is equal to summation x i minus x bar variance of y i is equal to sigma square by n into summation x i minus x bar whole square. And this quantity see summation x i minus x bar is equal to 0.

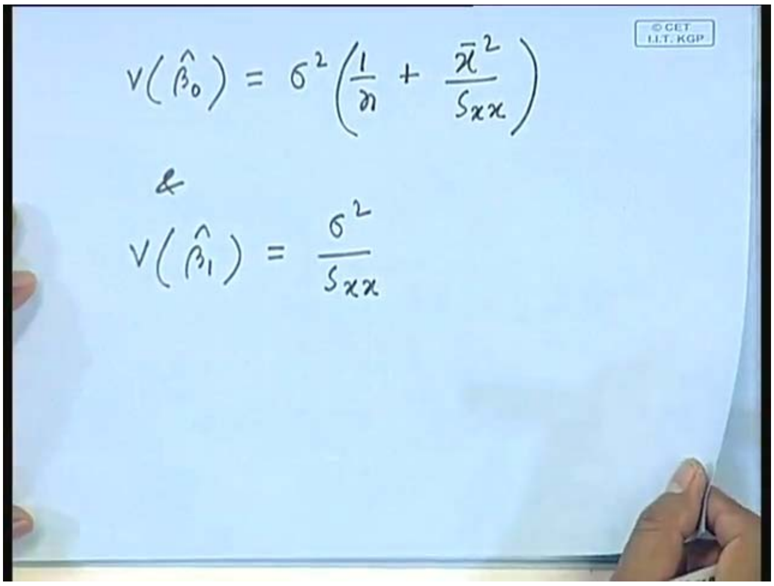
That is why the numerator is 0, so this is equal to sigma square into summation x i minus x bar this quantity is 0 always by n into summation x i minus x bar whole square. So, this is nothing this is equal to 0, because of the fact that summation x i minus x bar is always equal to 0. So, he proved that so he proved that covariance is equal to 0.

(Refer Slide Time: 47:53)



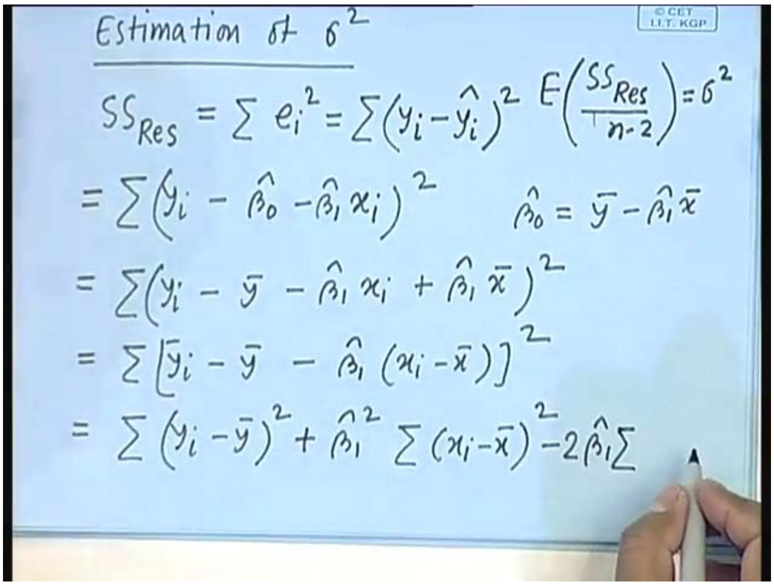
That means the variance of beta naught hat is equal to sigma square by n plus x bar square into variance of beta 1 hat which is equal to sigma square by S x x which we just proved. This is sigma square by into 1 by n plus x bar square by S x x, so we found . So, we proved that both the beta naught and beta 1 they are unbiased estimated sigma square.

(Refer Slide Time: 48:48)



And also we proved that the variance of beta naught hat is equal to sigma square 1 by n plus x bar square by S x x and the variance of beta 1 hat is equal to sigma square by S x x. Now, see both the variance formula the variance for beta naught hat and beta 1 hat both the involved a sigma square. So, but we do not know what is the value of sigma square, so sigma square must be replaced by its estimators. So, you need to estimate the value of sigma square, so next we talk about the estimation of sigma square, how to how to estimate sigma square.

(Refer Slide Time: 50:06)

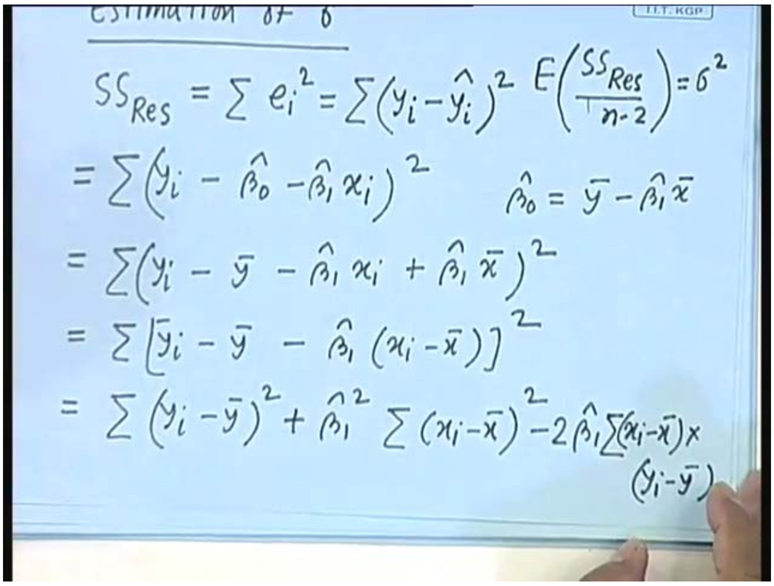


Estimation of sigma square well the estimation of sigma square is obtained from the residual sum of square SS residual. And finally will be proving that expected value of SS residual by n minus 2. This thing is equal to sigma square that means the sum of residual sorry residual sum of square by n minus 2 is unbiased estimated of sigma square. And will be we can compute the value of residual sum of square given a set of observations and we know n. So, using this formula we can we can estimate the value of the population variance sigma square. So, well we need to prove this one SS residual is equal to summation e i square. That we know and which is equal to the difference the ith residual is the difference between ith observation and the estimated value y i at.

So, this is equal to this is equal to square here this is equal to summation y i minus what is y i at is equal to beta naught hat minus beta 1 hat x i whole square. So, what we do here is that we know beta naught hat his equal to just trying to find a convenient form for SS residual. So, beta naught hat is equal to y bar minus beta 1 hat x bar right. So, will just plug this value here this is equal to summation y i minus y bar minus beta 1 hat x i plus beta 1 hat x bar whole square.

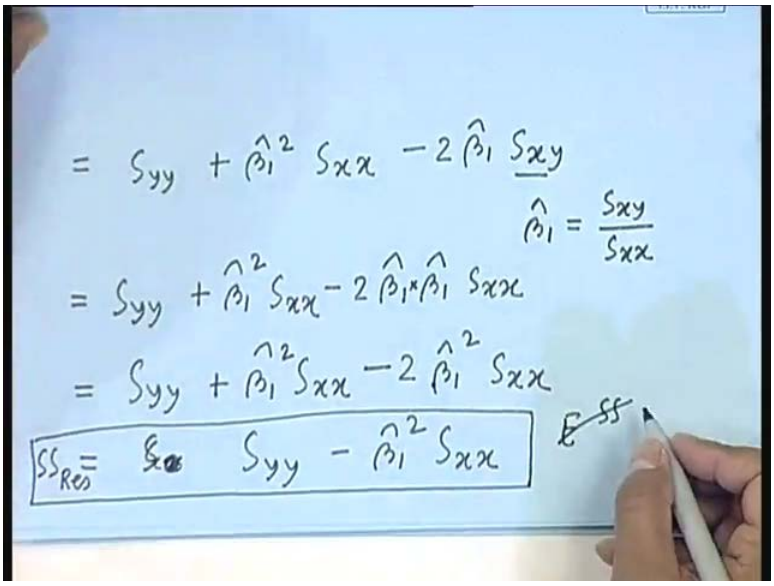
This is equal to summation y i minus y bar minus beta 1 hat x i minus x bar whole square, now this can be written as summation y i minus y bar whole square minus. Let me write plus beta 1 hat square summation x i minus x bar whole square minus 2 times beta 1 hat summation x i minus x bar into y i minus y bar right.

(Refer Slide Time: 54:23)



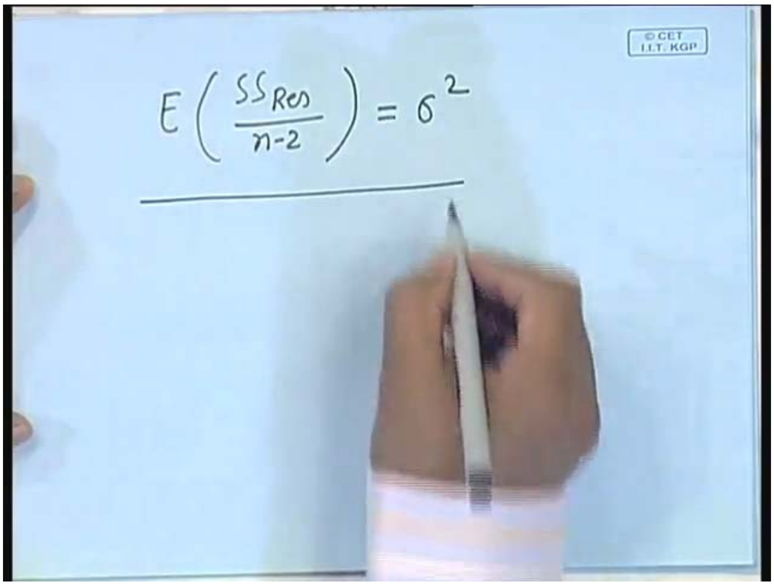
So, this can be written in the form S y y plus beta 1 hat square is x x minus 2 beta 1 hat is x y, just notation.

(Refer Slide Time: 54:55)



Now, see we know that beta 1 hat is equal to S x y by S x x right so what I will do is that I will replace this S x S x y by beta 1 hat into S x x. So, this is equal to S y y minus beta 1 hat is S x x minus 2 beta 1 hat. Now, replace this one by beta 1 hat S x x into this becomes S y y plus beta 1 hat S x x minus two times. This is square here square here beta 1 hat square S x x, so which is equal to S x sorry S y y minus beta 1 hat S x x. This is a convenient form of SS residual and we are going to use this one in the next class. To prove that because you need to find the expected value of this one to prove.

(Refer Slide Time: 56:45)



The expectation of SS to prove that expectation of SS residual by n minus 2 equal to sigma square well, so will continue in the in the next class.

Thank you.