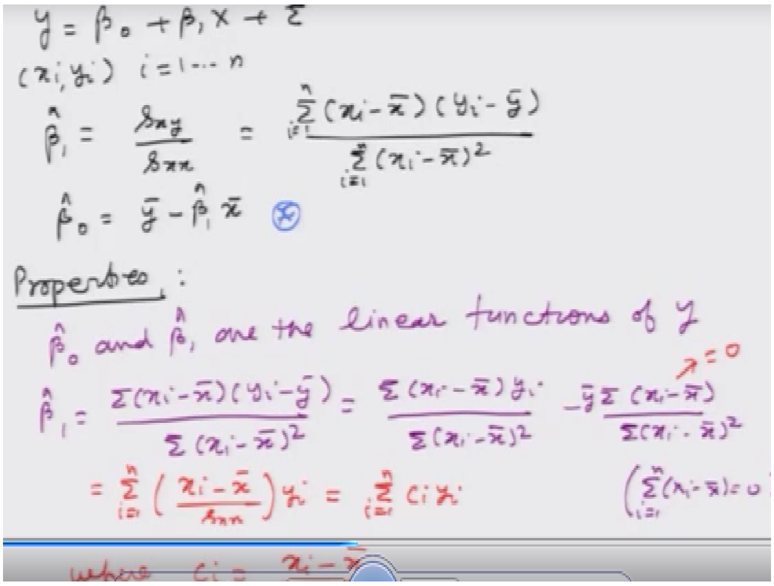
Regression Analysis and Forecasting
Prof. Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology – Kanpur

Lecture -05
Estimation of Parameters In Simple Linear Regression Model (continued)

Welcome to lecture number 5 in this lecture we are going to continue with our earlier lecture number 4, in earlier lecture we had obtain the parameter estimates of beta0 and beta1 keep in mind that we are three parameter beat0, beta1 and sigma square out of them sigma square is still remaining, but before that we would try to investigate the statistical properties of beta0 hat and beta1 hat.

The reason is a follows the parameter beta0 and beta1 they are the values in the entire population and we do not know, we have try to estimate them on the basis of a given sample of data. So how to ensure that the values which are being obtain using beta0 hat and beta1 hat they are good are bad values, so we try to expose them to statistical properties and we try to see whether they satisfy some statistical properties or not, so if you try to recall our model was

(Refer Slide Time: 01:29)



Was y=beta0+beta1 x+ epsilon, and we had obtain the observation xi, yi, i goes from one to n and we had obtain beta1 had as a sxy upon sxx and beta0 hat Was obtain as ybar - beta hat 1x bar, just for you remembrance this sxy was summation xi- xbar yi- ybar upon sxx which is summation xi - xbar whole square and all this summations are going from 1 to n.

Now let us you try to investigate the properties of beta0 hat and beta1 hat, one thing we can see her that both beta0 hat and beta1 hat are the linear functions of y what is this mean and how I can show it? So let me try to take it here beta1 hat, beta1 hat here is given by summation see here xi - xbar yi - ybar upon summation xi - xbar whole square

So this can I write down here summation say xi - xbar times yi divided by submission xi - xbar whole square - submission xi - xbar and times ybar. So ybar is independent of I, so I can write down over here and this becomes summation xi - xbar whole square, so we know that a summation xi - xbar is always = 0 i goes from 1 to n, this terms becomes 0.

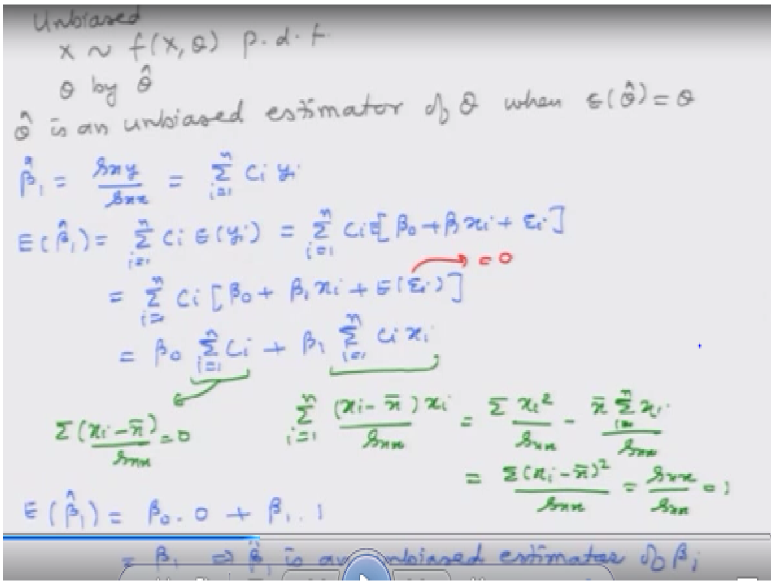
Now you can see her that this now only here a function of yi and so I can write down here it's a summation i goes from one to n, say xi - xbar upon say here sxx times yi or I can write down here i goes from 1 to n say ci yi where c i is xi - xbar upon sxx So I can is show that this beta1 hat is a simply a linear function of yi, 1 question cross of that why I am emphasizing on the linearity of the estimator, you see in this model we have epsilon as a random variable and because of that y also becomes a random variable.

Later on we will see that when we try to find out the expectations or variance and another statistical properties over beta0 hat and beta1 hat then doing mathematical manipulation over a linear function is easier than doing it on say any other nonlinear function, that is why if my estimators are linear the algebra and further properties become simpler to derive so that is the advantage of having a linear estimator.

Similarly I can also show here that beta0 hat is also a linear function now y which is not difficult to show and this can and this can also been shown say directly from this equation you see ybar is nothing but a linear function of y1, y2 , yn and beta1 hat, we already have shown that because the linear function and xbar is a fixed value that is the constant value that is non-stochastic value

Again the linear combination of the linear combination Again a linear function, so it is not difficult to prove that beta0 hat and beta one hat both are the linear functions of y.

(Refer Slide Time: 06:20)



The next property what we try to investigate here is about the unbiased character of beta0 hat and beta1 hat, just for the sake of review I can explain you what is the property of unbiasedness. So if I say if I have here a random variable say x which is following a probability density function f x theta and when we try to estimate theta by some estimator theta hat.

So we say that theta hat is an unbiased estimator of theta when expected value of theta hat = to theta. Right so now if I my question is I want to see whether beta one hat or beta0 hat is an unbiased estimator or not then essentially we have to see whether expected value of beta1 hat = beta1, and expected value of beta0 hat = beta0, So let us try to attempt on it, now we have consider this beta one hat as a sxy over sxx.

And early slide I had shown you that this can be written as summation i goes to now 1 to n, c i, yi. So when I tried to take its expectation, expected value of beta 1 hat, This becomes i goes now 1 to n, c i expatiated value of yi, now the question is what is expatiated value of yi? So let me try to solve it further i goes now 1 to n, says c i and expected value of yi is nothing but your beta0+ beta1 see here xi+epsilon i.

So this becomes i goes from here 1 to n c i and this becomes here beta0, because expected value of beta0 is Same as beta0 because beta0 is a constant that is a fixed value and this becomes beta1 expected value of xi which is again xi because we have assumed that xi is non-stochastic and then it becomes expected value of epsilon i and expected of epsilon i we have assumed that this is going to be 0.

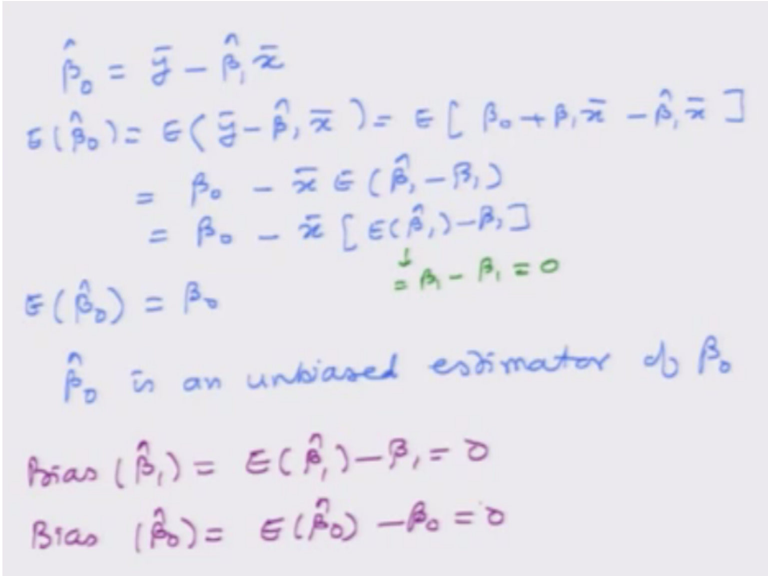
So now if you try to open it this becomes beta0 submission i goes from to n, c i + beta1 time submission i goes from one to n say c i, xi, what is the value of a summation c i? So summation c i if you try to find out here this comes out to be summation xi - xbar over sxx, so this going to be 0, so the submission of c i this is actually 0 Let us now try to obtain the value of summation c i xi.

So this can be return as i goes from one to n xi - xbar upon sxx times xi, so this can be written as summation xi square over sxx - summation i goes from one to n, xi over sxx. So submission xi is nothing but n time xbar so this quantity can be return as a summation xi - xbar whole square over sxx and his quantity is again sxx upon sxx which = 1 So I can write down here that expected value of beta 1 hat.

This is = now here beta0 into 0 + beta1 into 1, so this = nothing to but beta1, so this simplify that beta1 hat is an unbiased estimator of beta1. So now we have reason to be happy because the estimator that we have obtained on the basis of principle of least square by using the level observation this terms out to be an unbiased estimator.

Next we try to find out whether beta0 hat that had that we had obtain has ybar – beta1 hat xbar whether this is an unbiased estimator or not so we try to find out its expectations,

(Refer Slide Time :11:24)



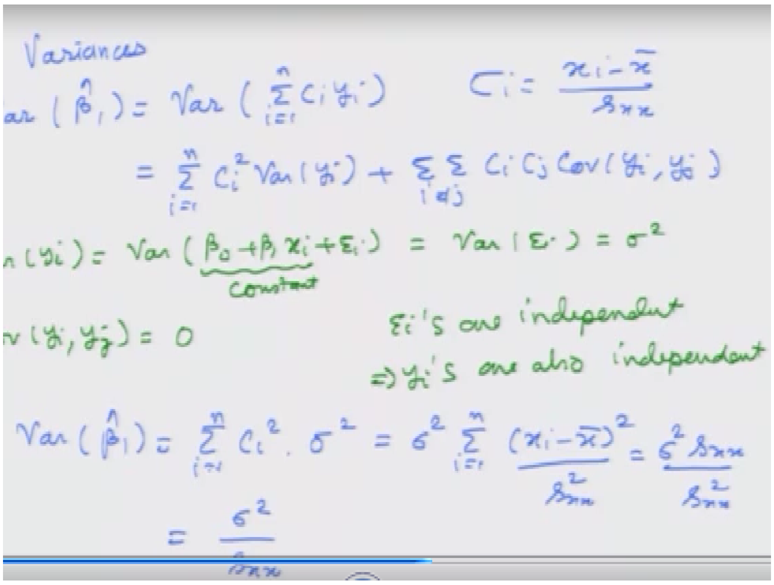
Expected value of beta0 hat that is nothing but your expected value of ybar – beta1 hat xbar, and so if you try to rewrite it ybar is nothing but beta0, + beta1 xbar – beta1 hat xbar, so this comes out to be again beta0 - xbar expected value of beta1 hat – beta1 because we had assume that xi are non-stochastic so xbar is again a fix quantity.

So we can write down here beta0 and this xbar and this quantity is nothing but your expected value of beta1 hat - beta one and this we have shown that this is an nothing but beta1, so this become beta1 – beta1, this = 0 and so I can write down that expected value of beta0 hat is beta0, so this beta0 hat is also an unbiased estimator of beta0.

So now I can define the basis of beta1 hat as expected value of beta1 hat – beta1, which =0 and bias of beta0 hat that is expected value of beta0 hat – beta0 that = again 0, so beta1 and say beta1 hat and beta0 hat they comes out to been unbiased estimator .So next property is a variance, well we have obtain a value of beta0 and beat one using beta0 hat

And beta one hat and we also have seen that both of them are unbiased estimator, now our next step is that we would you like to know what about their variability, so we try to find out the variances.

(Refer Slide Time: 13:45)



Once we have obtain the value of beta0 and beta one sing beta0 hat and beta one hat and after that we have establish that both of them are going to be unbiased estimator, so the next question is what about their variabilities? So let us try to first find out the variance of beta one hat. To make it more simple, I would try to express beta one as summation i goes one to n, c i, yi where your c i was xi - xbar upon sxx.

So you can see here this I can write down here summation i goes from one to n, c i square variance of yi and + double submission i is not equal to j c i, c j and covariance between yi and y j. Now I need to find out these terms, so let me try to find out thing, this variance of yi first, this is separately, so variance of yi is nothing but variance beta0 + beta one say this xi + epsilon i.

Since I have assume that beta0 and beta one are non-stochastic and they are fixed values and xi is also a fixed it's not random, so this entire part becomes constant, so this various expression becomes same as the variance of epsilon i, and we have assumed that the variance of epsilon i is sigma square So the variance of yi also become sigma square, this is very important result, later on you will see that using this result we will try to establish

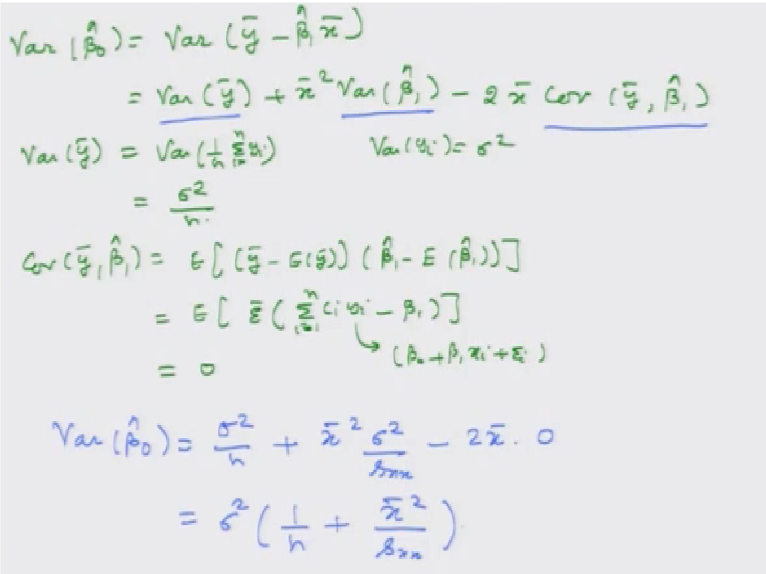
And we will try to infer many statistical properties, what is the advantage? You had assume that epsilon one ,epsilon two ,epsilon n are the random errors, they are unknown to us that you cannot know what are their values, but we assume that they have got variant sigma square. Now in case of a real data set you only get a value of xi and yi you will never have a value of epsilon I, but your model depends on sigma square, 0

so now incase if you want to infer about sigma square on the basis of a xi and yi, this result will help us at that place, so variance of yi and variance epsilon i they are the same. Now the next question is what is the covariance yi and y j? So this I can write this is 0 why because we have assume that epsilon i are independent, so once epsilon I 's are independent, so yi's are also independent .

So I can write down the covariant between yi and y j is 0, so this variance of beta1 hat, this becomes simply summation i goes from 1 to n, c I square times here sigma square, so let us try to simplify this expression further, So I can write down sigma square submission i goes from 1 to n, this is xi - xbar whole square upon sxx square .So this sigma square sxx upon sxx square so this is nothing but your sigma square upon on sxx.

So this is the variance of beta1 hat. The next question is that, well this is depending on sigma square, sigma square is the population variant again that is unknown to us, so how to know it on the basis given sample of data.

(Refer Slide Time: 18:27)



Right so that we will try to address in the coming slides, but before that before that we try to find now the variance of beta0 hat. So variance of beta0 hat is nothing but variance of ybar –

beta1 hat xbar, so this i can write down here using the simple statistical rules variance of ybar + xbar square variance of beta1 hat - twice of xbar covariance between bar hat and beta1 hat.

So now let us try to find out these expressions one by one, so what about here variance of ybar variance of ybar is we have assume that each observation yi has got a variance sigma square, so this becomes variance of one over n summation i goes from one to ny i, and this equals sigma square by n. What about this covariance of ybar and say beta1 hat, this can be retained as expected value of ybar -.

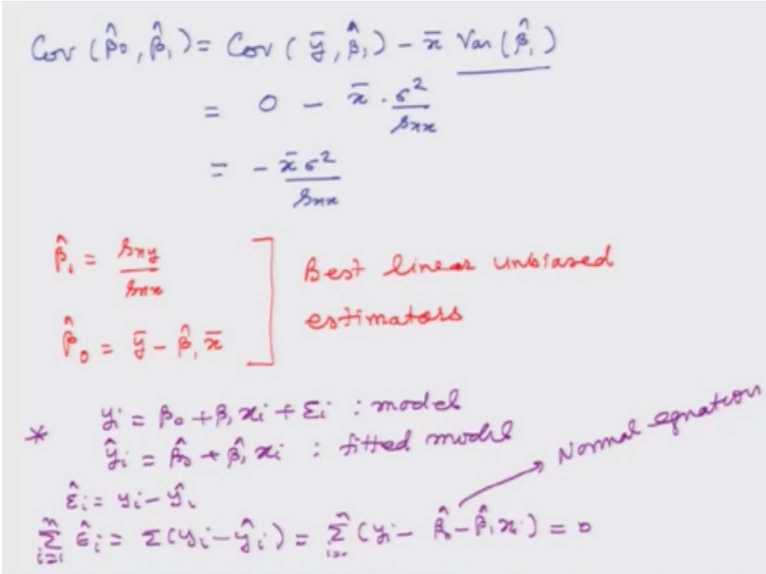
Expected value of ybar and beta1 hat - expected value of beta1 hat. Right, so if you try to expand it this becomes ybar - expected value of ybar is epsilon bar and this beta1 hat is I goes from one to n summation, c I yi and expected value beta1 hat we already had obtain beta one.

So again if you try to substitute here the value of beta0 + beta one say xi + epsilon i and if you try to simply open it and then take expectation within the property that expected value of epsilon = 0, it will turn out to be 0. I can write down now here that variance of beta0 hat is variance of ybar from here, this is sigma square by n + xbar square and variance of beta one hat that we already had obtain.

Right, so this I can write down here as sigma square over sxx and - this quantity that is twice of xbar into 0, so this comes out sigma x square 1 over n + xbar square over sxx. So this the variance of beta0 hat but again you will see that is the depending on the quantity sigma x square and again sigma square is the population various that is unknown to us in practice.

So how to cement that will be our next question, but before that we have obtain the mean and variance of beta0 hat and beta1 hat, so let us try to find out also the.

(Refer Slide time: 21:53)



Covariance between beta0 hat and beta one hat that will help us in establishing some more statistical properties later on, so if you try to see this can be retain us covariance between ybar, beta1 hat - xbar variance beta1 hat, that will be covariance between beta1 hat and beta1 hat, which is nothing but your variance of beta1 hat. This expression we already had obtain in the earlier slide.

This comes out to be 0 and this we already had obtain that variance of beta one this is nothing but your sigma x square over sxx so you can see here very clearly that this - xbar sigma square upon sxx and this quantity is not going to be 0 in general. And that is expected, so don't assume that what we have obtain beta0 hat and beta1 hat they are independent no, under usual circumstance they are not going to be independent unless and until either xbar becomes 0 or sigma square becomes 0.

Right, so there is another thing that we have obtain the value of beta1 has beta one hat which is sxy upon sxx, and we obtain the value of the beta0 hat as a ybar – beta1 hat xbar. Actually I can show that they are the best linear unbiased estimator, which means this beat one hat is the best linear unbiased estimator of beta1, and beta0 hat is the best linear unbiased estimator of beta0.

But after this we would like to investigate some properties of the model based on the values of beat 0 hat and beta1 hat, so one important property that can be observed from here is the following. Right we had got the model which was beta0 + beta one xi + epsilon i this a was our model and we had obtain the fitted model like this, and based on that we had obtain the fitted values for a given value of x n and xi.

So when I try to define my residual as a yi- yi hat let us try to see what happens to submission epsilon I hat, so when I try to find out the summation of epsilon I hat here this becomes submission yi- yi hat right, and this nothing but was a summation i goes from one to ny I - beta0 hat – beta1 hat xi, Now what is the value, I going to write it this = 0.

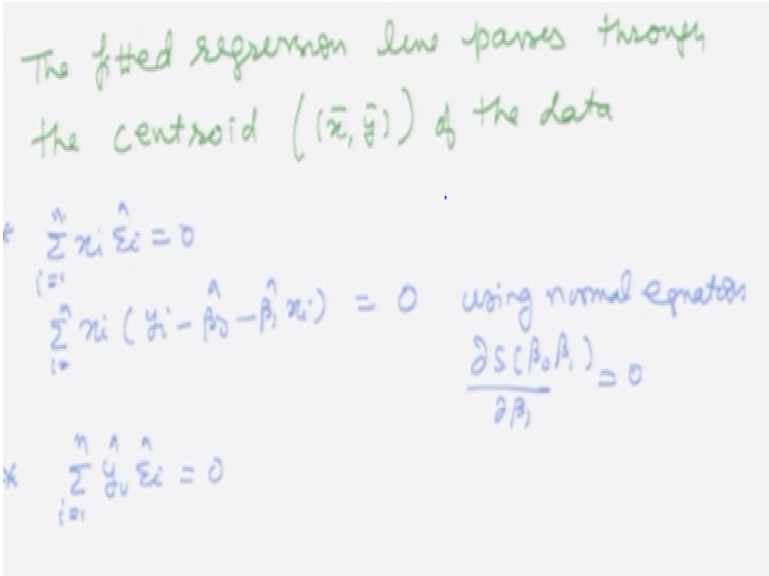
But how? It does not look like right but if you try to recall when we had obtain the normal equation after differentiate the function s beta0 beta1 hat, so from that normal equation that you hat obtain here right, you can see her that this summation yi- beta0 – beta1 xi = 0, and now what are you beat0 and beta1 they are the solution of these two equation, this equation and this equation.

So I can also write that yi- beta0 hat – beta1 hat xi = 0 and similarly I can also write yi- beta0 hat – beta1 xi times xi = 0, so these are your 2 normal equation they have been obtain at the time obtaining the least square estimate, so now using this equation I can write down here directly because this is coming from the normal equation.

So you can see here what I am trying to say in this result is that the sum of residuals is 0. Now here you have to be a little bit cautious when I make an assumption that expected value of epsilon i=0 then some time people get confused with the same result, but you have to keep in mind that when I say expected values epsilon i=0 that means I am saying that the average of the random errors in the entire population is 0.

But here I am trying to say that we have got a small sample, we have got n data sets, based on that we have obtained the residuals. Now I am raying to show that the some of those residuals is also 0 and in turn the sample mean of residual will also be 0, sometimes people get confused that what is the different between saying epsilon bar or the sample mean of epsilon is 0 or the population mean of epsilon is 0.So here you have to be little bit careful.

(Refer Slide Time: 28:20)



Next result which I would like to inform you is that the fitted regression line pauses through the centroid, centroid is what here in our case the centroid is xbar ny bar of the data that you can actually verify if you try to substitute the value of xbar ybar you can easily verify it. Next result which, I want to show you is that submission xi epsilon I hat =0 right i goes from on to n, so now exactly using the same technique.

Which I used in demonstrating the earlier result, I can also do the same thing here, this xi, this yi- beta0 hat beta1 hat x I. Now how I can right this value 0, so if you come here again on this light using this result, the result based on the second normal equation this is simply saying that submission yi- beta0 hat – beta1 hat xi times xi=0.

So without doing her much I can write down this value to be 0 using normal equation, right which normal equation that beta0 beta1 with respective beta1=0 Right similarly it is not difficult to prove that i goes from 1 to ny i hat epsilon i hat = 0, so these are some simple properties that we will use later on in establishing some more statistical properties of estimators.

So now we have consider the estimation of beta0 and say beta1 and we have investigate there statistical properties, one thing is still left, how to estimate the sigma square that is a parameter that is still remaining for us. Now we stop here in this lecture and in the next lecture we will try to estimate the remaining parameter sigma square till then good bye.