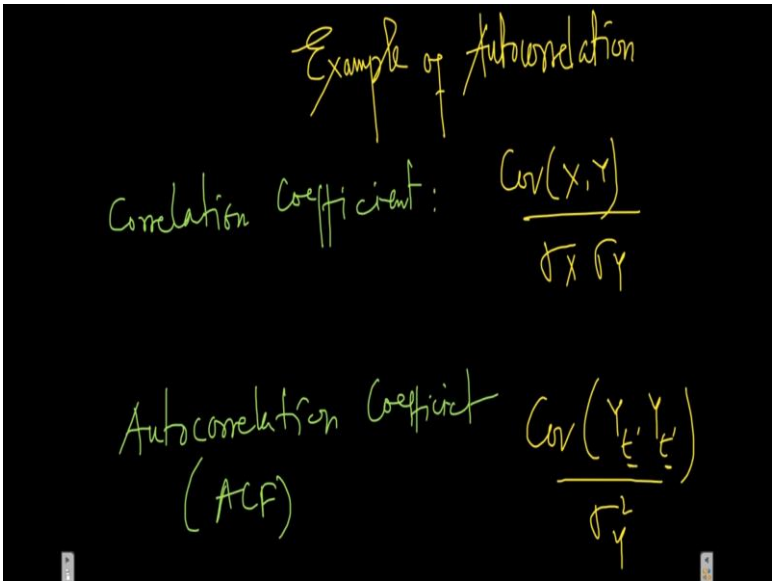
Applied Econometrics
Prof. Tutan Ahmed
Vinod Gupta School of Management
Indian Institute of Technology – Kharagpur

Lecture – 71
Autocorrelation (Contd.)

Hello and welcome back to the lecture on applied econometrics. So, in this lecture we are going to talk about autocorrelation.

(Refer Slide Time: 00:31)



And in this lecture we are going to give an example of autocorrelation. Now, before we actually start with an example, let us first start to understand what is the difference between correlation coefficient and autocorrelation coefficient? So, correlation coefficient we are already familiar with. And we will see autocorrelation coefficient is just the same representation but with a little difference correlation coefficient and autocorrelation coefficient.

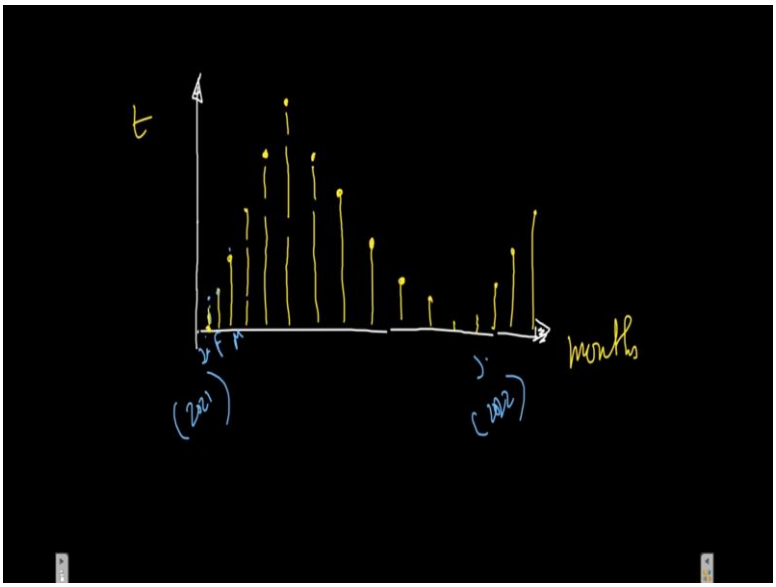
Sometimes we know autocorrelation coefficient as autocorrelation function is represented as ACF. It is commonly known ACF. Now, we know we are familiar with you know how to estimate correlation coefficient and we write let us say for 2 variables X and Y. We write covariance X Y on the numerator and in the denominator, we have sigma X and sigma Y.

Now, similarly autocorrelation coefficient is just the same but here we have. Remember when we are talking about autocorrelation, it is always correlation to itself. So, it has to be

correlation of the same variable with itself but for different time period. So, if I have X t, or let us maintain the consistency let we have Y t. So, it has to be with Y t prime, let us say. And then in the denominator since we are talking about the same variable, t y.

So, then it is going to be sigma square Y, let us say. Here my variable under consideration is Y. So, it has to be sigma square Y. So, that is basically the autocorrelation coefficient. And it depends, the value will depend between the difference that we take t and t prime. So, you just come to that.

(Refer Slide Time: 02:46)



Now, let us actually try to understand intuitively, to sort of understand what is autocorrelation? Let us actually draw something. And let us say, I want to understand how the temperature is actually wearing or across the different months. And here I have my, let us say, it is not a really great straight line. Let me draw it again. Yeah, much better.

And let us say, I have ah let us actually draw does not matter a straight line. Now, let us so, in my Y axis, I have temperature. And here I have months. So, let us say in the month of January, I will start and here let us say temperature is almost at 0, let us say is more than 0. But, it is close to 0. I am talking about India. Let us say in the month of February, I will have a little more temperature.

In the month of March, I will have a little more temperature. In the month of April, I am going to have a little more temperature. In the month of May, I am going to have quite a lot of temperature. It is going to be the summer, hot big summer. June let us say, it is even more

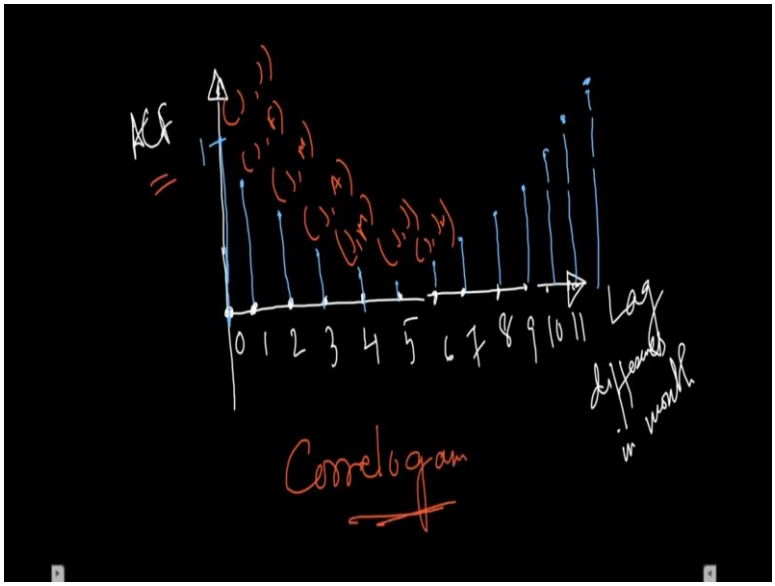
hot where from July onward because the monsoon and all the temperatures started declining, July.

It is August a lot of rainfall and temperatures declining. September let us say, it is further down. October, it is down. November is down. December is close to 0. And January is again that and February is again that. March and this sort of cycle repeating March, April. So, this is let me use a different colour. This is my January, February, March. And let us say, this is January 2021. And this is our January 2022.

Now, see the pattern here. Of course there is a pattern. Now, I want to understand how I sort of measure the correlation coefficient just I have seen for the understand the autocorrelation part. Now, if I take let us say correlation coefficient between Jan and Feb. So, the temperatures are pretty close actually. So, what I get is the correlation coefficient is going to be a little high.

But, if I take Jan and March, let me actually Jan and March. So, it is going to give me a little less value on the correlation coefficient simply because the temperatures are a little more different.

(Refer Slide Time: 05:46)



Not as close as, it was for Jan and Feb and in fact let me actually draw this. I will have a different plot going on parallel. So, if I have, let us say this is the, if this axis is representing, let us say ACF, autocorrelation function or the correlation coefficient for the measurement of autocorrelation. And this axis actually represent the lag.

Lag, I will just talk about lag and let us say this is the differences in month, differences in month. So, if I have Jan and Jan, if I measure Jan and Jan so, there is no difference. If I measure Jan and Feb, there is a one difference. So, lag is 1. If there is a difference between Jan and March so, the difference is 2 months, so, lag is 2. If there is a difference, you know, if you are measuring between Jan and April so, then there is a difference of 3.

So, that way and then Jan and May, it is the difference of 4, Jan and June is difference of 5, Jan and July is 6, Jan and August is 7, September is 8, October is 9, November is 10, December is 11. Now, if I so, this is what is lag. So, depending on when you are measuring whether using this formula. When you are taking this Y t and by the way, this is Y t, this is not, this is not anything else, Y t is just a comma and Y t prime.

So, the t prime, basically you see the difference between t and t prime. So, if it is a difference of one which is, if the month is Jan and Feb then the difference is 1. So, in that case, t prime is going to be t + 1 or, if it is going to be let us say April then t prime is going to be t + 3. So, we just trying to see this is what is called lag basically the time difference between these 2 points.

And we are trying to see what will happen to the correlation coefficient because of the lag. Now, if I take let us say, Jan with Jan so, basically the same month, same data, same everything. So, then I will get a value of what? A value of 1, my correlation coefficient is going to be 1, For the very fact that I am basically taking correlation between 2 variables, the 2 same observations so, it is basically going to give me a value of 1.

But when it goes Feb, I am going to have a little less correlation coefficient. When it is March I am going to have even less. April it is even less, May even less, June even less. And July, let us say July temperature is let us say increasing a little bit. August let us, it is increasing a little bit sorry, decreasing a little bit September is decreasing a little bit. So, again it is coming close, October is you know decreasing even more, November is decreasing even more, December it is decreasing even more.

So, the value is almost close to what we had here. And then again let us say, Jan 2022, I have the same value almost the same value. So, it will be close to 1. So, it is basically so, this is for

the month of Jan 2021. So, I took the correlation coefficient. So, here let me use a different colour. So, Jan, Jan or J J, J F, J Jan March. So, Jan remains constant. Jan April, Jan May, Jan June, Jan July.

So, this way, we can actually sort of see the impact of lag on the ACA for the correlation that we get for the, data for the observations in different time periods. And this plot this is actually a very important plot and this plot is actually known as correlogram. That is basically, the idea of how we see a time series data and how lag plays an important role when we look at time series data.

So this, we will end this lecture and in the next lecture we are actually going to see what happens because of the autocorrelation. Consequence of autocorrelation. And then going forward, we will see how to actually address the problem of autocorrelation. Thank you.