Economics, Management and Entrepreneurship
Prof. Pratap K. J. Mohapatra
Department of Industrial Engineering & Management
Indian Institute of Technology – Kharagpur

Lecture – 35
Forecasting Revisited (Contd.)

Good morning. Welcome to the 35th lecture on Economics, Management, and Entrepreneurship. In our last lecture, we revisited forecasting methods. We elaborately discussed regression methods. In that, we had developed a vector matrix representation of the dependent variable y with a set of independent variables x1 through xk and we said that it is possible to estimate the values of the regression parameters beta0, beta1, beta2 etc that minimize the squared errors, error between the estimated values of the dependent variable and the actual values.

We also said towards the end of the lecture, that there are a few statistics that indicate the adequacy of the regression of the dependent variable y on the independent variables x1 through xk. We shall to start with today discuss an example to illustrate the use of the regression method and thereafter we shall discuss 2 methods of time series forecasting that are useful for short-term projection of an independent variable. Let us take off the example first to illustrate the use of the regression method.

(Refer Slide Time: 02:35)

Example:

Sale of a product is assumed to depend on price and average personal income. The values of these variables are given below for five cities.

Observation No	Sale (Thousand Rs/year)	Price (Rs/unit)	Income (Lakh Rs/year/person)
1	10	2	20
2	12	4	30
3	20	3	40
4	16	6	50
5	25	4	60

Fit a regression line. Find the R^2 and the *t* statistics.

If the price and income are predicted to be 5 Rs/unit and 80 Lakh Rs/year/person), then what is the most likely sale of the product?

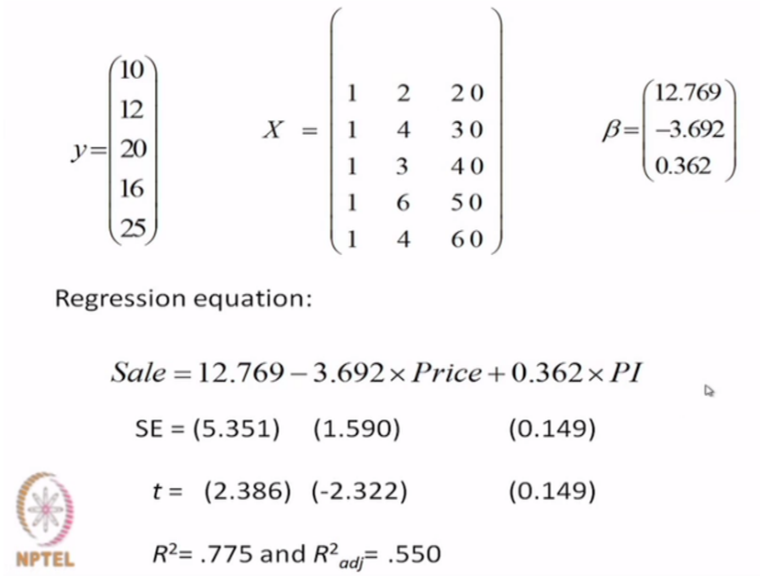
Here is the example. Sale of a product is assumed to depend on price and average personal income. The values of these variables are given below for 5 cities that is, 5 observations are made in different cities. city 1 the sale was 10,000 rupees per year, where the price of the product sold there was only 2 rupees per unit whereas the average personal income in that city was 20 Lakh rupees per year per person. City number 2 sale was 12,000 rupees per year.

Price of the product selling there was rupees 4 and the average personal income was 30 Lakh rupees per year per person. So like that in 5 different cities 5 values of sale and the prevailing unit price of the product in each of these cities and the average personal income of a person per year in terms of Lakh rupees per year per person there also found out. So this is the table of observations.

Here we have assumed sale to be a function of price and income that means price and income are independent variables are assumed independent and sale depends on these 2 independent variables therefore sale is the dependent variable or explain variable and price and income are considered as explanatory variables. We are required to fit a regression line that is regression sale on price and income that means find out the regression coefficients of the regression between sale and price and income.

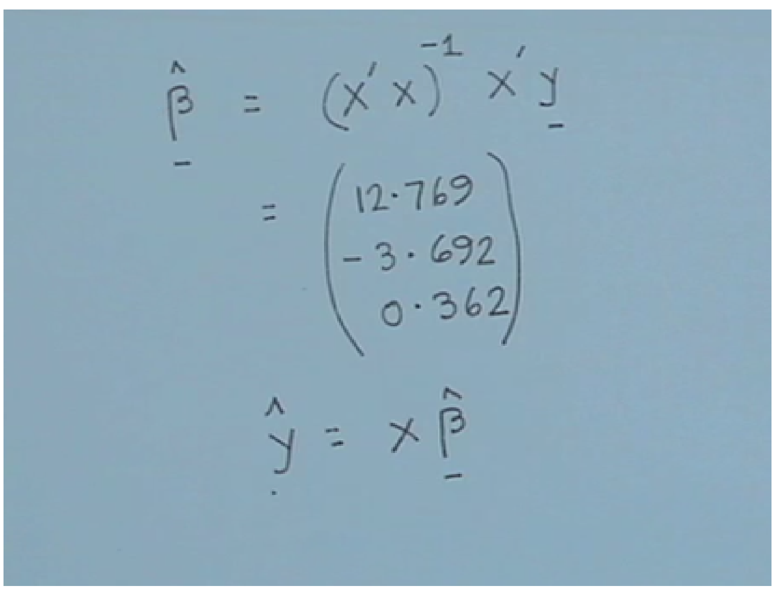
Find also the R square and t statistics. Now if the price and income are predicted to be rupees 5 per unit and 80 Lakh rupees per year per person, then what is the most likely sale of the product. This is the question. The first part of the question deals with estimating a linear relationship between sale and price and income. The second part of the question is after we have established a relationship between these variables. Use that relationship to project the value of sale of the product when price and income values are predicted to be 5 and 80 respectively.

(Refer Slide Time: 06:07)



From the given data the observed value of the dependent variable, the sale is y. sale is basically y. The matrix of independent variables contained the first column is 1, 1, 1, 1, 1, but the other 2 columns are the price figures and the personal income figures.

(Refer Slide Time: 06:48)



Then we use our formula for beta which is x transposed x whole inverse x transposed y. Now that we have the values of y and x are given here. We can use them here and find out the value. The value for beta is a vector that comes as 12.769 - 3.692 and 0.362. Now these values are given here in this vector. Hence the regression equation is y hat = x beta hat and from there we get sale = 12.769 - 3.692 * price + 0.362 * personal income.

So that the beta values are given beta 1, this is beta 0 12.769 - beta 1 * price + beta 2 * personal income. Now these results can also be obtained as I said by using any of the standards of software packages, one of the most well known package is what is known as SPSS. SPSS is one of the oldest software packages, statistical packages. There are nowadays a large number of statistical packages however now these results have been obtained by using SPSS and the value of beta has been obtained as this.

Hence the regression equation is obtained as this that is beta 0 + beta1x1 + beta2 x2. In this case, beta1 is - 3.692 and beta2 is 0.362 PI. Now from this equation the interpretation is that if there is a unit change in price whereas these are held constant. The other variables like personal income remains as it is, but price increases by 1 and the sale we will fall by - 3.692. The reverse is also true if price falls, personal income remaining same then the value of sale would rise by 3.692.

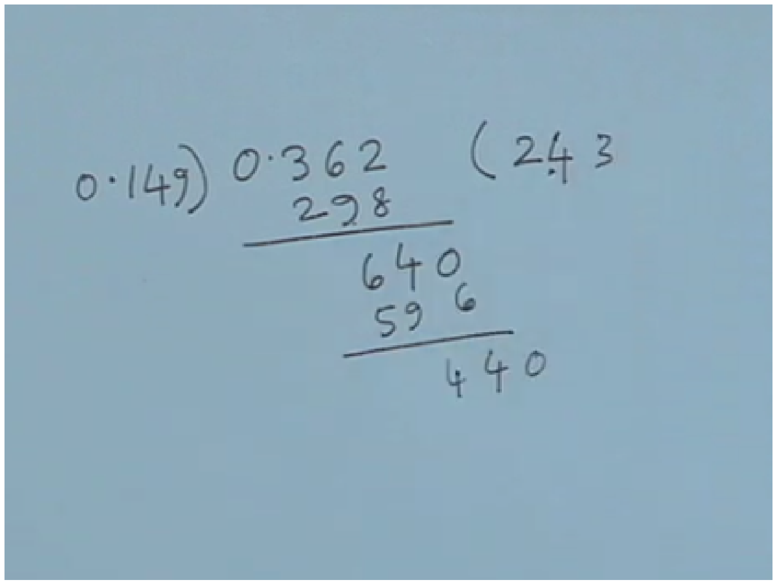
So this negative sign indicates that there is a reverse (()) (10:10) between sale and price. If price increases sale decreases and the amount by which cell will decrease, then price is given an increment of 1 is 3.692 when all other variables remain constant at that (()) (10:30) values. We can give a similar interpretation for 0.362. It says that price held constant at its (()) (10:41) value. If personal income rises by 1, then sale would raise by 0.362 thousand rupees per year.

So there is a positive relationship between sale and PI and the magnitude of change in sale for a unit change in PI given price is constant remaining constant is given by the coefficients. So these are the interpretations of the regression coefficients which are in this case this and this. From our own observation also we know that if price increases in most of the product sale would fall and if personal income raises sale would raise thus intuitively this relationship is adequate.

However, there are statistical relationships also. Statistical estimates can be made of standard error of these estimates. Beta0 the average value is 12.769 and its standard error calculated by this software package is 5.351 and the t statistics is this divided by this that comes to that is beta0/its standard error that comes to 2.386. Similarly, the standard error in making the estimate of the values of beta 1 is 1.592 giving a t value of the ratio of the 2 which is - 2.322.

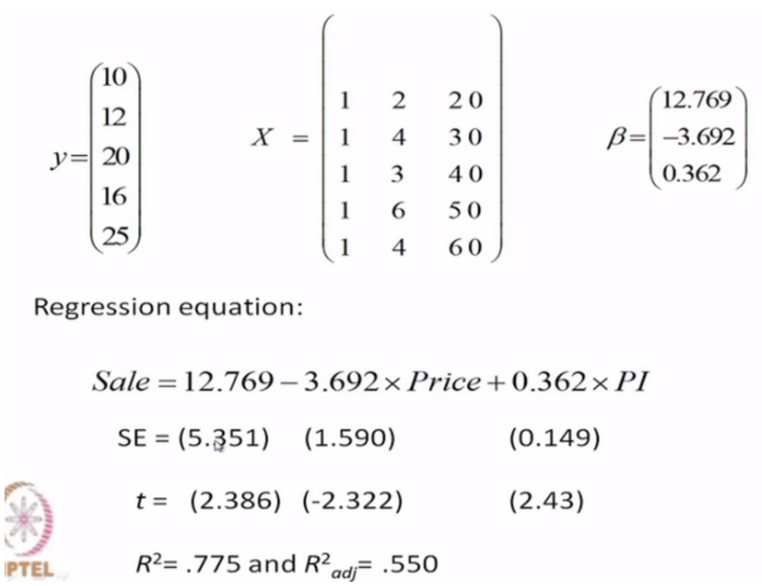
And the average value is 0.362 that is the estimate. The standard error of the estimate is 0.149 and the t value this might be little erroneous I think it is because 0.362/0.149 would be little different may be 2 point something, and not 0.149 this may please be corrected cannot calculate it right now, but it appears that there is an error and I have to wait for the values I have forgotten the values it is not coming.

(Refer Slide Time: 13:51)



This value will be 0.362 division 0.149 so this will be close to 2.43 so let me make this correction. This will be close to 2.43. It has not taken this value, but anyway this value is going to be 2.43 it is not taking. Now it has taken.

(Refer Slide Time: 15:58)



The t value is 2.43. Now basically this you can see that this standard error being less than the actual values, the t values are higher and this means that the beta values are not significantly different from 0. These are the interpretation of the standard error values and the t values. If the t values would have been less it would have indicated that the values of this are close to 0 and if they are close to 0, it means suppose that the t value for pi was close to 0.

It would mean that this coefficient is also close to 0. It means that the personal income is not an explanatory variable it does not explain sale, but because it is high it means that this also is different from 0. This means that personal income does have an influence on sale. Now I also introduced another statistical or R square statistics. In this case the software SPSS get a value of R square as 0.775 and the adjusted R square value came us 0.550.

We would rely more on adjusted value of R square 0.55 means that 55% of the variation of the actual values from its average was explained by our regression method. 55% is not very good, it only means that there are other factors that rather play and that influence the sales. Our initial assumption that price and personal income alone are the main factors that govern the variation in sale perhaps needs to be relieved into and there may be a few other factors that are governing the variation of sale.

The other factors could be the competitive products for example a price of competitive products. This could be another explanatory variable and if we now that we know that the R square are adjusted value is just 0.55 it now tells us that we should look for such other explanatory variable and if it is at the price of a competitive product is playing a role then find its value and take also that therefore we then have 3 explanatory variables.

The price of this product, price of the competitive product and on top of that the personal income. So it is a very 3 explanatory variables which are explaining the variation of sale, but assume that this is our estimated relationship between sale and price and personal income which is reproduced here.

(Refer Slide Time: 19:33)

$$Sale = 12.769 - 3.692 \times Price + 0.362 \times PI$$

The predicted price and income are 5 Rs/unit and 80 Lakh Rs/year/person) respectively.

The most likely sale of the product is given as

$$Sale = 12.769 - 3.692 \times 5 + 0.362 \times 80$$

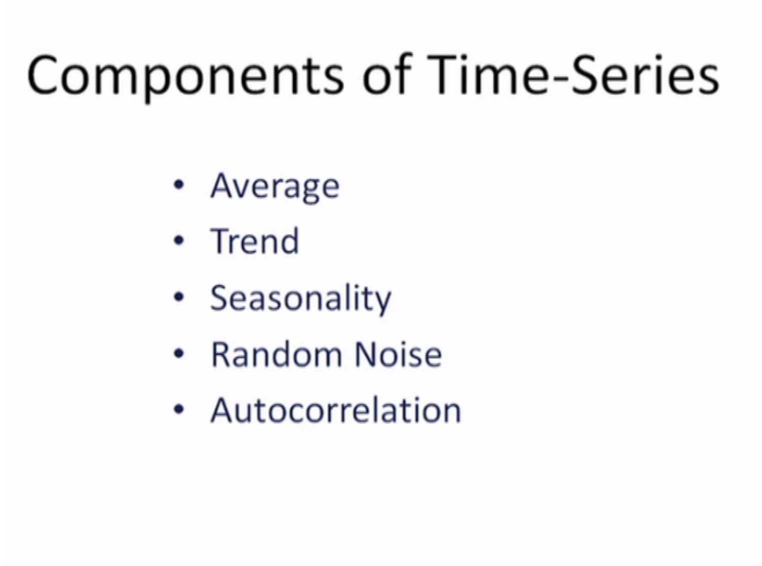
= 23.269 thousand Rs/year

= 23,269 thousand Rs/year

The second part of the question says that suppose the predicted price is 5 rupees per unit and personal income rises to 80 Lakh rupees per year per person that what is the likely sale. What we need to do is just to put the value of this predicted values price = 5, pi = 80 and that gives us a value of 23.269 and they are thousand rupees per year and therefore this is 20, 269 and not this thousand again, this is a mistake. This is 23.269 thus this should be just 23. I am sorry there is a problem here. It is not responding very fast. I must be careful. Now that it has come.

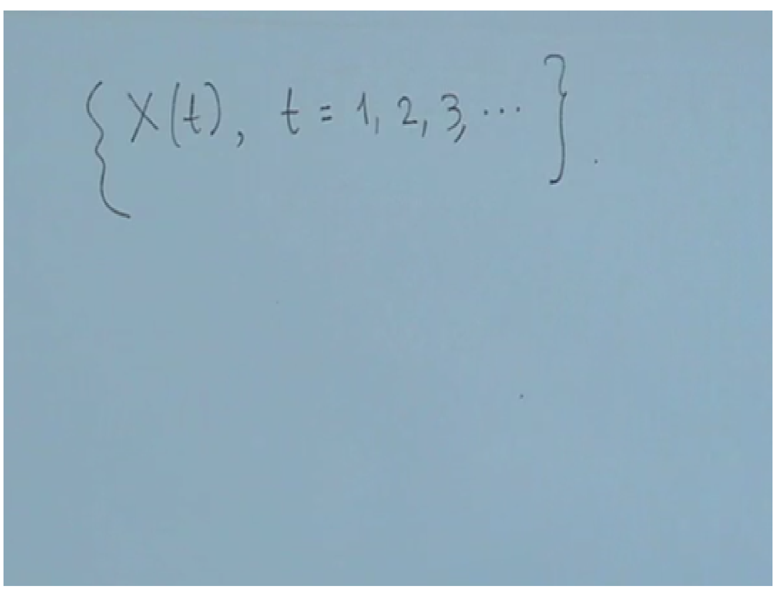
So we have through this example understood how to formulate a model how to estimate the coefficients of a regression model and then how to interpret the values of the regression coefficients and their signs that is + and - and how to interpret R square and t statistics. Now these are the very fundamental requirements for any regression analysis. Regression methods and econometric methods are very, very well developed and one has to study much more in detail to get to know the other (()) (22:08) of this interesting and very useful method.

(Refer Slide Time: 22:16)



Now we go to time-series forecasting method. Time-series forecasting methods if you recall are useful for short-term forecasting. Short-term forecasting is needed for production planning, for inventory control, for working capital management and such other things. So we should first of all understand the components of time series. Please recall that time series means that we have observed a particular variable x.

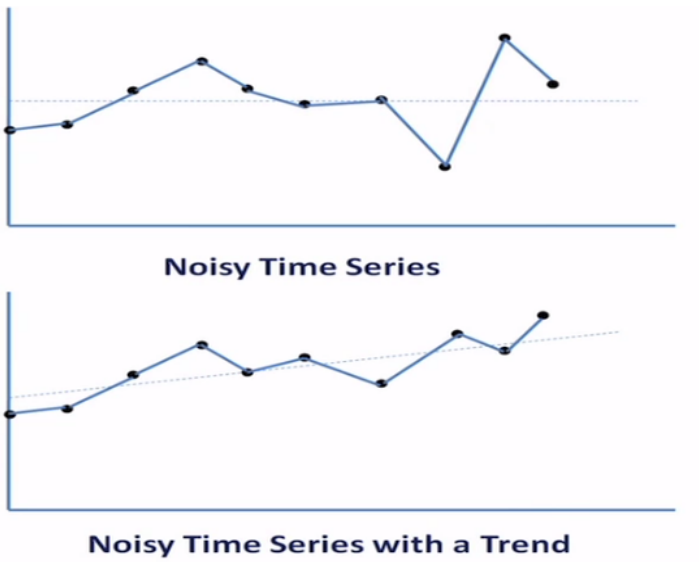
(Refer Slide Time: 23:01)



For various values of t, this is called a time series. Observed values of x are different points of time. Now normally a time series we will have 4 components although I have written 5, the first 4 components it will have an average, it will have a trend, a seasonality, and a random noise. These are all of them are independent of each other.

Autocorrelation is another component of time series with the help of which one can find out whether there is a trend, whether there is a seasonality and so on and so forth. Let us first of all study the meaning of these components. Before, we do that just 1 second. There is a lot of problem here. I am sorry this whole thing is giving a problem.

(Refer Slide Time: 25:38)

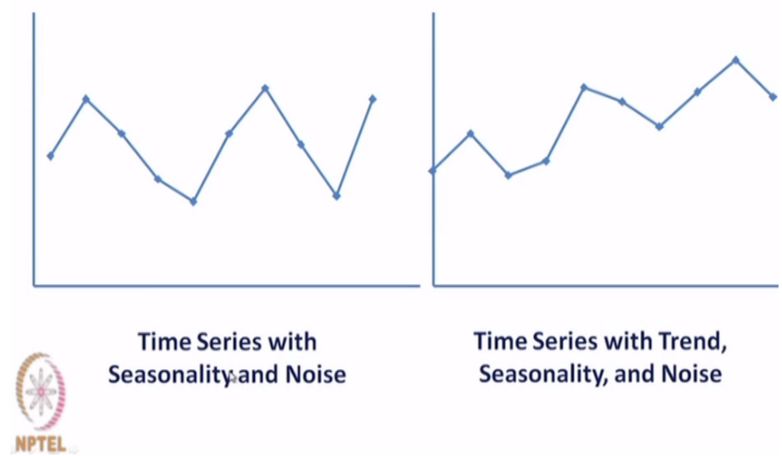


Yes, now I told you that there are 4 components of a time series. 1 is the average, the other is the trend, 3 is seasonality, and 4th is random noise. There is also a 5th component which is autocorrelation with the help of with which you can understand whether there is a trend or seasonality or what. Now let us see this is a noisy time series. It means that this series has got noise associated with it, but of course it has an average value such as this.

So this has an average value. This is an average value of this time series. This is XT and the X axis is time T. The Y axis is XT, the value of the observed variable X at time t are these values indicated by the field dots and we can say that it has a noise that is superimposed on this average value. The noise sometimes takes less value, sometimes takes more value, but on an average it has a value 0. Now come to this particular time series.

In this case we can see that there is a long term trend overall above the average. The average is somewhere here, but there is a trend or if we take this as the intercept a, then there is a trend here and on top of that there is a noise associated on this. So this is noisy time series with a trend. Now look at this.

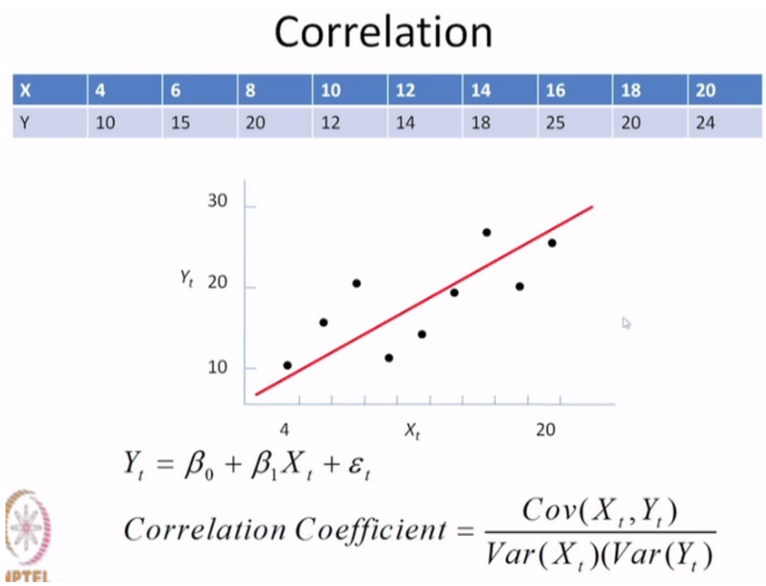
(Refer Slide Time: 27:38)



This time series you will see that there is an average which is here and now it looks like there is a seasonality here. There is a regular pattern up and down up swing and down swing, up swing and down swing in regular interval. This whenever there is a regular fluctuation with regularity we usually called it a seasonality with a length of periodicity remaining constant. So here we do not have a trend we have instead an average, a seasonality, and there are small changes, and that is because of the noise present.

The 4th one contains all the 4 components such as the average, the long term trend, there is a seasonality over the long term trend, and there are various fluctuations because they are not actually the same amplitude so that is difference in amplitude is because of the noise that is present. So these 4 examples illustrate how a random or how a particular time series can be expressed in the form of its main 4 components.

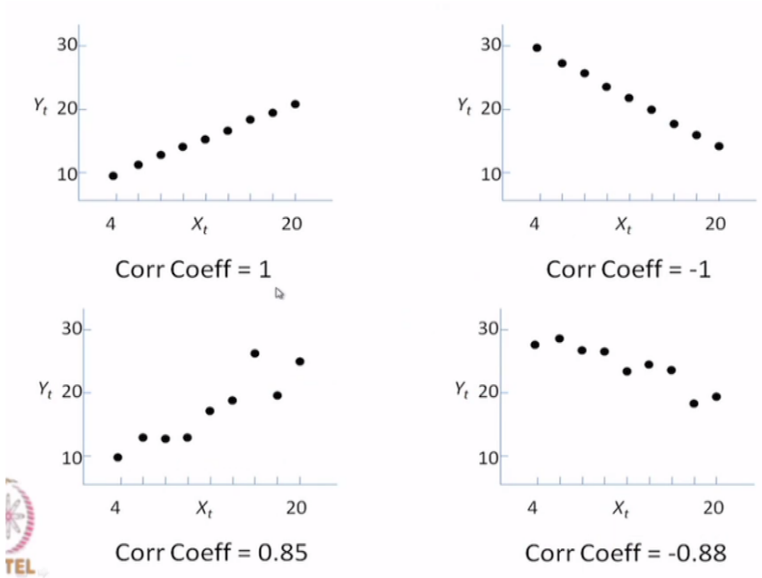
(Refer Slide Time: 29:03)



We also introduced a component called autocorrelation. To understand autocorrelation, we have to fist understand the meaning of correlation. As I told in the last lecture correlation can be expression to indicate linear relationship between 2 variables, let us say the variables are x and y. If at a particular time we find the value of x and y then and suppose that we have in this case.

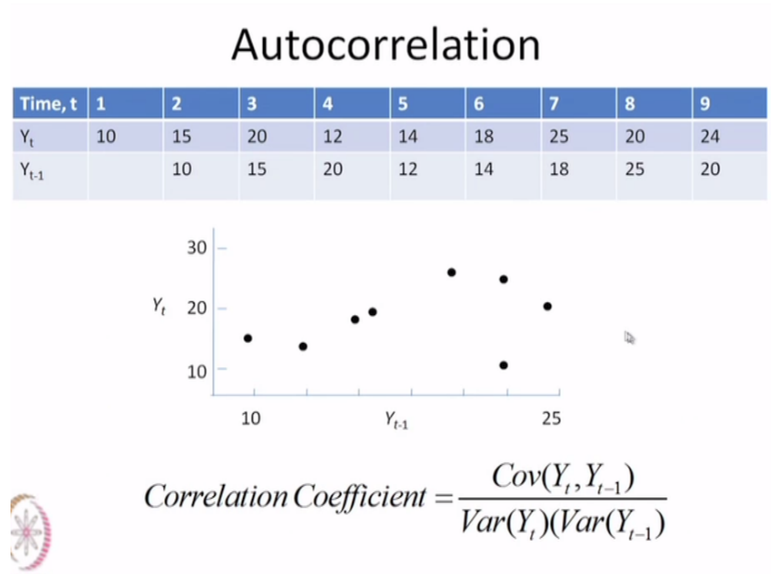
We have taken 1, 2, 3, 4, 5, 6, 7, 8, 9 data points. For each data point, we have x and y so this is the plot of x and y for the 9 sets of observations and we find that this is on the rise. So this is the probably the regression relationship between y and x, but the correlation coefficient is positive. This is expressed here in this manner.

(Refer Slide Time: 30:21)



If the line almost exactly in a straight line then they are 1 or close to 1 with a positive slope if the line in a straight line with a negative slope then the correlation coefficient is - 1 and if there is a slight deviation from linearity, but the trend is very high, the departure from the straight line is not very high, the correlation coefficient is 0.85. In this case the correlation coefficient is - 0.88 so negative and close to 1. So this is the meaning of correlation coefficient between 2 different variables, but we are dealing with only 1.

(Refer Slide Time: 31:10)



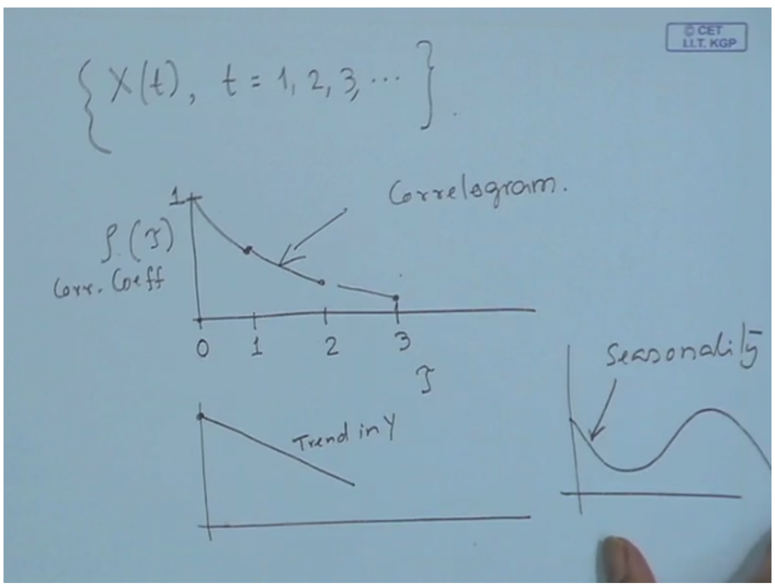
Now this is another example that suppose the values of we are not taking about autocorrelation. We have understood what is correlation. Autocorrelation is always discussed in only 1-time series. Let us say the time series is Yt and the value for 9 readings are 10, 15, 20, 12, 14, 18, 25, 20, and 24. What we do? We take one period lag Y values. We call it Y t - 1 that means at time 2 we take Y1 2 - 1 is 1 what was Y1 was 10. So Y t - 1 is 10.

We write it here. At t = 2 we write the value of Y t -1 that is 10. At t = 3, we take the value of Y2 that was 15, so we take write 15 here. 20 is written here, 12 is written here like that 14 here, 18 here, 25 here, and 20 here. Now this is like 1-time series y and this is another time series which is basically a lagged variable by 1 period lag. If we instead consider y t - 2 then when t = 3 10 will come.

So 10 will come, 15 will come here, 20 will come here, 12 will come here, so we shall have another time series that is basically Y t lagged by 2 time periods and similarly we can have Yt lag by 3, by 4, by 5, and 6 periods. Now suppose we consider only 1 period lag time series that is Y t- 1 and then take we can now find correlation between these 2-time series. So this is the type of correlation that I have plotted here for time period when y = 10, the value of Yt is 15.

Then Y t - 1 is 15, the value is 20. So like that they have been plotted. Now this may have a correlation of something like 0.5, positive trend, but not exactly linear relationship. So it is something like 0.5. So that gives the value of the autocorrelation between Yt and 1 period lag Y as 0.5 and similar such things we can find between Yt and Yt - 2, Y t - 3, Yt against Yt - 4 so on and so forth.

(Refer Slide Time: 34:14)



Now that we have the correlation coefficients of we call it normally rho as the correlation coefficient and that is the function of lag called as lag as tau. So if we take tau here when this value is 0 it is 1, it is correlated with itself. When it is 1 it will be less still 2, it will be still less. So there will be some sort of a curve here. This curve is called correlogram. Now this curve may take different shapes. If this curve takes a shape such as this, it means that there is a trend in Y. If it takes a shape such as this, it means there is a seasonality. So this is the meaning of autocorrelation.

(Refer Slide Time: 35:54)

Smoothing of Time Series

Given a time series

$$X_t, t = 1, 2, \bullet\bullet\bullet, N$$

$$\bar{X}_N = \sum_{i=1}^{N} w_i X_i$$

is a smoothed (average) value of the time series computed at time point *N* if

$$w_i > 0, \sum_{i=1}^{N} w_i = 1$$

Now let us go to our topic of time series forecasting. So given a time series Xt, t = 1 through N we can find out its smoothed value by taking weighted average of each of the pass data when the weights are all > 0 and they add up to 1. So X bar N is the weighted average of X.

(Refer Slide Time: 36:24)

If $\{X_t\} = \{5, 12, 8, 20, 10\}$ and $\{w_t\} = 0.2$ for all *t*

$$\bar{X}_5 = (0.2)(5 + 12 + 8 + 20 + 10) = (0.2)(55) = 11$$

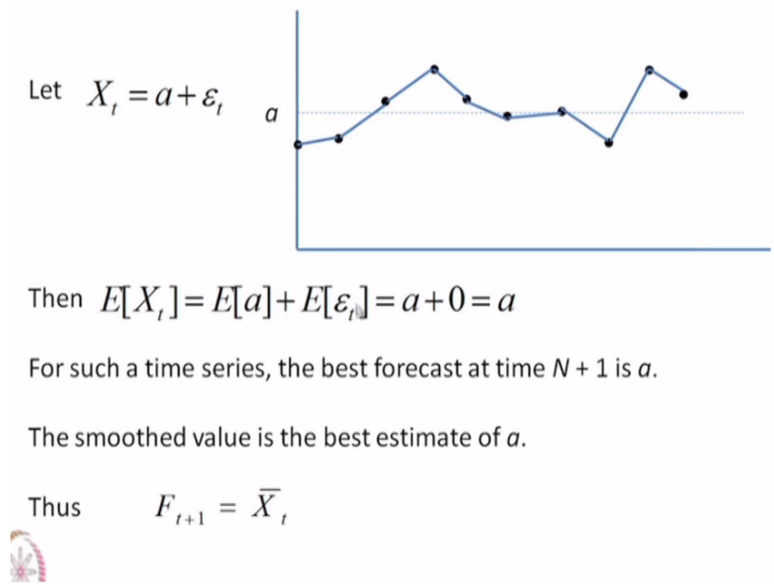
If $\{w_t\} = \{0.5, 0.2, 0.2, 0.1, 0\}$

$$\bar{X}_5 = (0.5)(5) + (0.2)(12) + (0.2)(8) + (0.1)(20) + (0)(10)$$
$$= 10.5$$

For example, if Xt is observed to have values 5, 12, 8, 20, and 10 and suppose that we take all its weights each of them is given a weight of 0.2 then 0.2 * 5 + 0.2 * 12 + 0.2 * 8 and (()) (36:42) it gives us a value of 11. Thus 11 is the average value of Xt, but suppose the weights are different. Suppose that we give highest weight to 5 that is 0.5, 0.2 to 12, 0.2 to 8, 0.1 to 20, and 0 to 10.

We consider each of them is higher than > 0 and they add up to 1 and 0.5 + 0.2 + 0.2 + 0.1 + 0 it gives 1. So these are also weights. The smoothed value of X now becomes < 11. It becomes 10.5. So there are different ways to calculate smoothed value or average value of a time series by giving different weights. The only 2 conditions are each weight must be > = 0 and they must add up to 1.

(Refer Slide Time: 37:54)

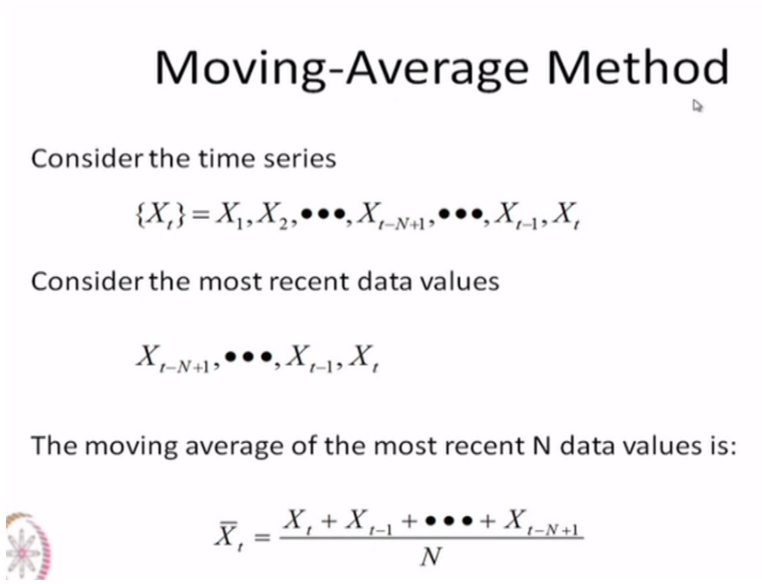


Now the first time series, this time series if you recall is having a average value a and superimposed on that is a random noise (()) (38:07) call this epsilon t. Random noises are normally assumed to follow normal distribution with mean 0 and a constant variance that means it can go deviate from 0 by certain amount. So the expected value of X is nothing but expected value of a + the expected value of epsilon and this is being 0 it is = expected value of a constant a which is a itself.

Therefore, for such a time series the best forecast at time N + 1 is a. So if by plotting this suppose that this is N we have plotted this and we have found that the best model for this time series is Xt = a + epsilon t that is it has an average and a noise then the best forecast is average itself. a is its best forecast. The smoothed value is the best estimate of a. Now by smoothing we are basically trying to find out what is the value of a.

Thus the forecast is time period N + 1 or t + 1, the next period forecast is nothing, but the average value of x calculated at this point of time. This is thus the forecast very simple.

(Refer Slide Time: 39:38)



Now we introduced moving average method. Consider the time series Xt = X1, X2 etc up to t - N + 1 and then X t -1 and then finally Xt. So we have given the present value as t and the past values as this. Now consider the most recent n values, it starts from here this is the most recent Xt, the present value and last N data. Last N data points come to this. So up to this we will not consider the previous values. So up to this we are considering.

So if you are considering only these values we have with us X t - N + 1 etc up to Xt. Write it reversely it will be Xt + X t -1 the previous value and then the last N value which is Xt - N + 1. What is its average? By giving a weight of 1/N to each one of this it will come to Xt bar. So this is the average value of the most recent N data values.

(Refer Slide Time: 41:10)

After one time period the most recent N data values are

$$X_{t-N+2}, \bullet\bullet\bullet, X_{t-1}, X_t, X_{t+1}$$

The moving average of the most recent N data values is:

$$\overline{X}_{t+1} = \frac{X_{t+1} + X_t + \bullet\bullet\bullet + X_{t-N+2}}{N}$$

$$\overline{X}_{t+1} = \frac{X_{t+1} + (X_t + \bullet\bullet\bullet + X_{t-N+2} + X_{t-N+1}) - X_{t-N+1}}{N}$$

$$\overline{X}_{t+1} = \overline{X}_t + \frac{X_{t+1} - X_{t-N+1}}{N}$$

Now suppose we have we go forward by 1-time period. So we will then have another time series data Xt + 1 and suppose that we still rely on only N data points that means we shall delete the oldest data X t - N + 1 which was there earlier. We shall delete that and we have added the most recent data 1 period. Now the time has advanced by 1 period, it becomes Xt + 1. So we are dealing with another set of data, the oldest data being deleted and the newest data being added.

Therefore, the moving average. So we can now freshly calculate an average based on the fresh set of data we call it a moving average. Calculated as t + first time period and that will consider the most recent data 1 period old data and N period old data/N and we can continue these process. However, one does not have to calculate this X bar t + 1 is = we can now say X t + 1 up to this and - this.

So you will see that this is nothing but the past average plus the one X new data that is added - the old data that was deleted/N. That means you have to remember only the last average value and the 1 that you are deleting and the 1 that you are retaining or adding. If you know these 3 data points we can calculate the new value of X bar t + 1.

(Refer Slide Time: 43:14)

If the time series consists of an average and noise only

$$X_t = a + \varepsilon_t$$

then the moving average is the best forecast for the next period.

If the time series consists of an average, a trend, and noise only

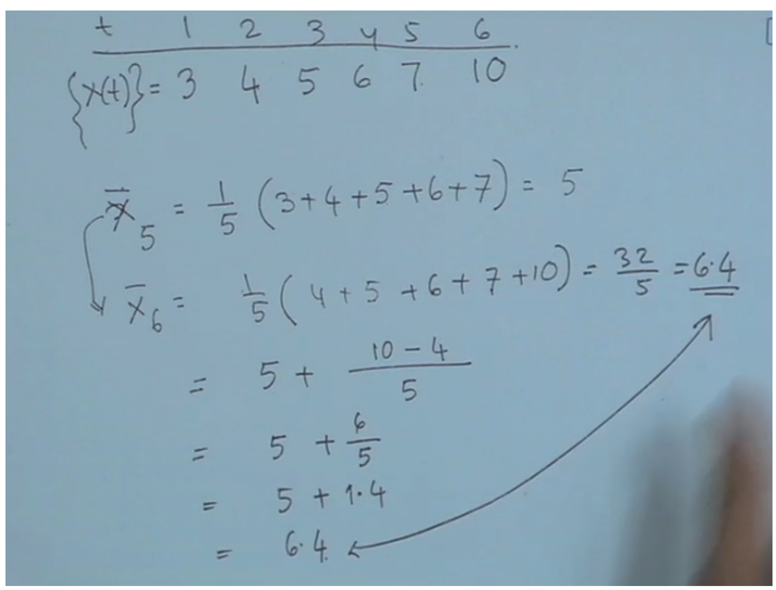
$$X_t = a + bt + \varepsilon_t$$

Hence, m-period forecast is

$$F_{t+m} = a_t + b_t m$$

Now I will give a small example to illustrate this point.

(Refer Slide Time: 43:16)



Suppose that we have 5 data points 3, 4, 5, 6, 7 like that so these are Xt values. We are interested to calculate this is time 1, 2, 3, 4, 5. So we calculate X bar 5 = 1/5 3 + 4 + 5 + 6 + 7 whatever that comes to this is = 5. Now that suppose the next value is 10 then X bar 6 we calculate as 1/5 4 + 5 + 6 + 7 + 10 which is = 32/5 = 6.4 that is incidentally = the last average which is 5 + what has been added is 10 and what is subtracted was 4.

Let us see whether it is coming that way. 6/5 which is = 5 + 1.4 which is 6.4 so this is matching. Now we have seen that the average is being moved by 1-time period that is why the name moving average. Now if the time series consists of an average and noise only it is expressed in this manner then the moving average is the best forecast for the next period that means you find out the X bar P and that is the best forecast.

If the time series consists of an average, a trend, and a noise then it is usually expressed in this fashion. Xt = a + bt. this is the trend rising with time t and superimposed by a random noise. So the m- period forecast will be F tm = at + btm. M-period forecast. If 1 period forecast it is just a + b, 2 period forecast it is a + 2b, a + 3b etc. We write at here because as new time period as time advances we estimate the value of a by taking another moving average value. That is at. That why moving average value changes at changes. We also will make an estimate of the trend b how we do it we are showing it here.

(Refer Slide Time: 46:46)

Single Moving Average Applied to Data Containing Average and Noise Only.

Period	X_t	Single Moving Average $\bar{X}_t$	Forecast F_t	Error $X_t - F_t$
1	7			
2	8			
3	10	5		
4	3	7	5	-2
5	11	8	7	4
6	10	8	8	2
7	6	9	8	-2
8	8	8	9	-1

Now look at this case. Suppose that Xt is given by 7, 8, 10, 3, 11, 10, 6 and 8. Observe that time periods 1, 2, 3, 4, 5, 6, 7, 8. Let us assume that we have used N = 3. That means based on these values we calculate these value. Just one second I think there is a problem here.

(Refer Slide Time: 51:42)

Single Moving Average Applied to Data Containing Average and Noise Only.

Period	X_t	Single Moving Average $\overline{X}_t$	Forecast F_t	Error $X_t - F_t$
1	6			
2	8			
3	10	8		
4	3	7	8	-1
5	11	8	7	4
6	10	8	8	2
7	6	9	8	-2
8	8	8	9	-1

To illustrate this problem this approach let us say that we have 8 period data for a time series Xt. the values are 6, 8, 10, 3, 11, 10, 6, and 8 and let us assume that we have taken only n = 3. So if you take n = 3 then 6 + 8 + 10 that is 24/3 is 8 that is the moving average calculated at this time point and when we go to the next time period we will not consider 6, but consider 3 in addition so we shall only consider 8, 10, and 3 so if we take that then the moving average of this time series at period 4 = 8 + 10 is 18 + 3 which is 21/3 which is 7.

Similarly, when time period proceeds or advances by 1 more period then we will consider only 10, 3, and 11. So this is 13 + 11 is 24/3 is 8 and next 1 is 3 + 11 + 10 is 24/3 which is 8. Next 11 + 10 + 6 which is 27/3 it is 9 like that we can calculate the single moving average X bar t and what we have said that suppose we assume that the forecast made for next period is the moving average calculate today.

Then forecast for the next period when we have calculated the moving average as 8, forecast for the next period is taken as 8 that means we have initially assumed that this time series contains only an average and a noise and no trend. In that case, the best forecast is the moving average value for the next period. So forecast for the next period is the moving average value made today so 8 is appearing here, 7 here, 8 here, 8 here, 9 here.

Now here we find that there is a forecast error. Now this forecast error will also be different. This forecast is Xt - Ft in this case Xt is 3 and the forecast is 8, therefore the value is 3 - 8 which is - 5. The next forecast is 11 - 7 the error is 4, 10 - 8 the error is 2, 6 - 8 the error is - 2 and like that. So this is a forecast and the forecast error is Xt - Ft which is this.

So we are trying to say that whenever we are trying to make a forecast with moving average there will be a forecast error and that this is the way to use the moving average method and we will try to see how to use various alternatives to single moving average method. Thank you for today.