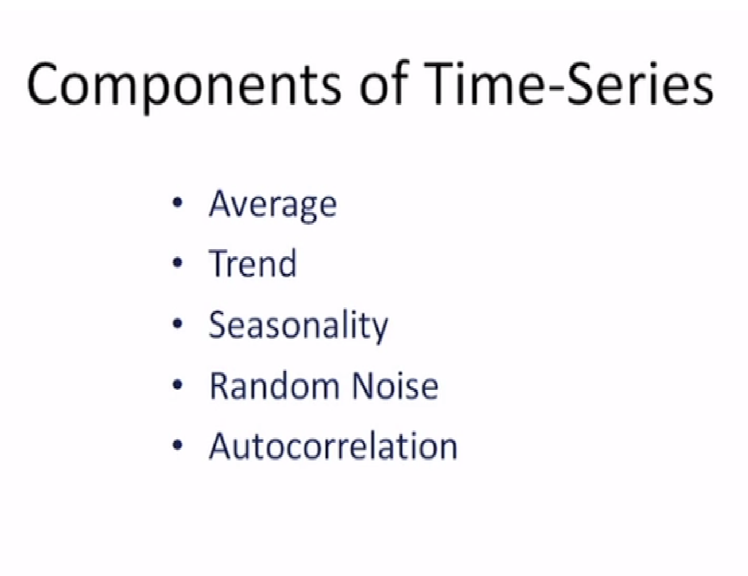
Economics, Management and Entrepreneurship
Prof. Pratap K. J. Mohapatra
Department of Industrial Engineering & Management
Indian Institute of Technology – Kharagpur

Lecture – 36
Forecasting Revisited (Contd.)

Good morning. Welcome to the 36th lecture on Economics, Management, and Entrepreneurship. In our last lecture, we had covered multiple regression method as a forecasting technique. Regression methods are usually applicable to medium-term forecasting. For short-term forecasting normally we use time series forecasting. In our last lecture we had introduced the concept of time series and we had said that a time series had got different components.

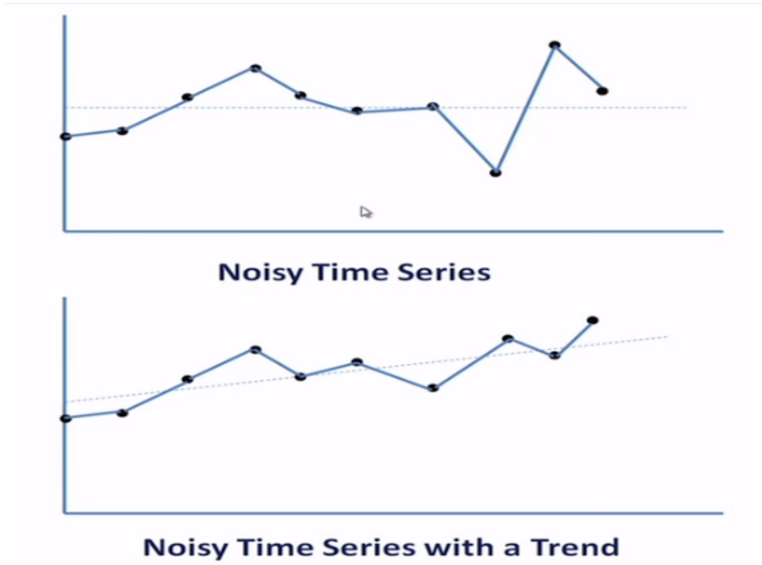
And then we had said that there are principally 3 methods: 1 the moving average method, 2 the exponential smoothing method, and 3, the Box-Jenkins method which are also known as ARMA, ARIMA methods. Let us once again look at the components of time series and the first 2 methods to moving average and exponential smoothing method.

(Refer Slide Time: 01:56)



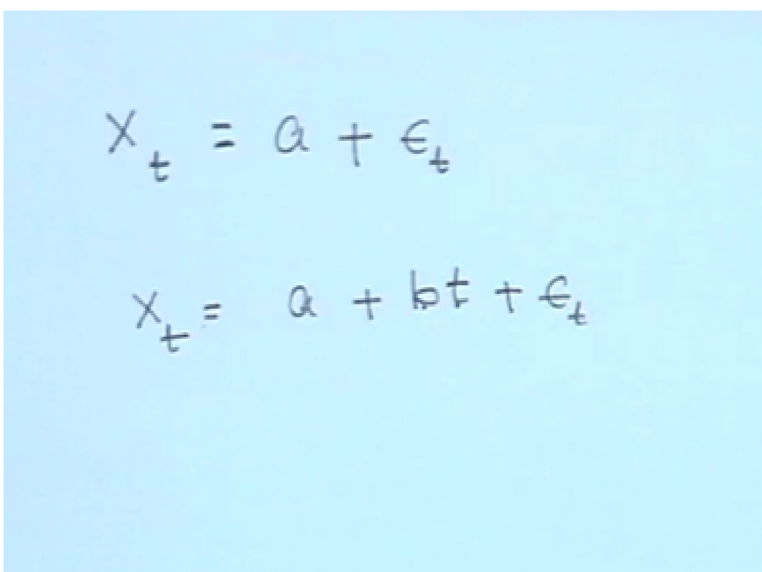
First, the components of time series and we had said that there are principally 4 components average, trend, seasonality, and random noise. There is also a 5th component called autocorrelation with the help of which one can identify whether there is trend or seasonality and also in its own right autocorrelation is a component with the help of which time series can be modeled and can be forecast.

(Refer Slide Time: 02:35)



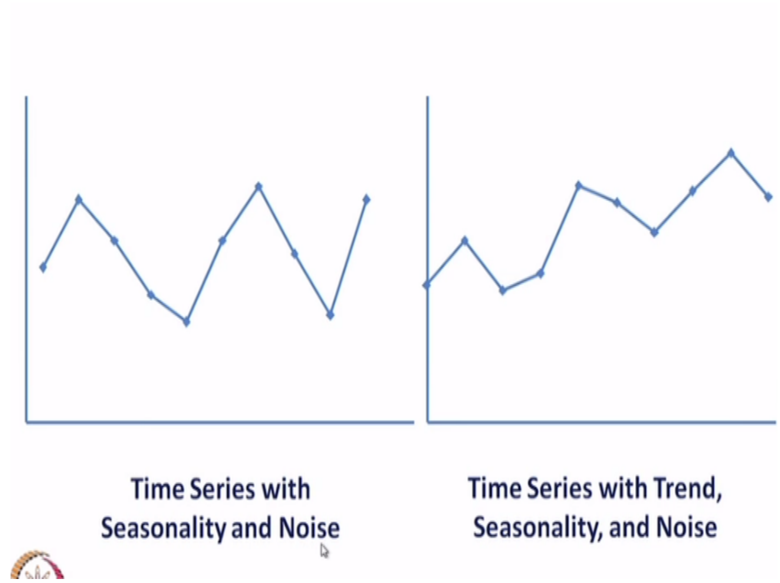
And I had told you that suppose a time series does not contain a trend or a seasonality, but contains only an average and certain fluctuations around the average, then this is equivalent to saying that Xt =. Xt the time series = a constant a + a noise term epsilon t.

(Refer Slide Time: 02:59)



That is the first diagram the noisy time series. The second diagram is noisy time series with a trend it means that we have in the second time series we have the time series Xt expressed as a constant a + a long term trend t and its coefficient b and + there is a random noise epsilon t. Thus this is the equation of a + bt and on superimposed on that is random noise therefore the actual values fluctuate around this.

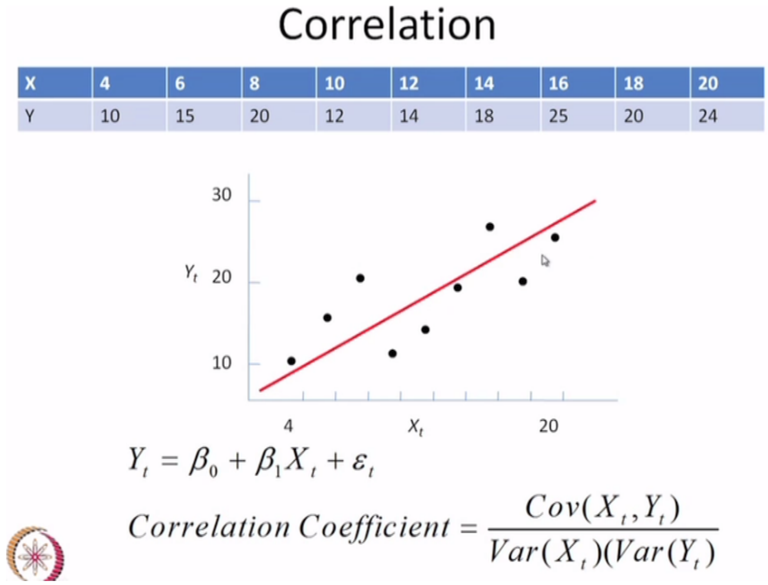
(Refer Slide Time: 04:00)



Then we have time series with seasonality and noise. So whenever there is a seasonality we will see that there is a regularity in the fluctuation it has got a constant amplitude, almost a constant amplitude and the periodicity is constant and now in this case the average is there. There is a seasonality around the average and there is certain random noise whereas here we have an average and a long term trend and on that there is a regular seasonal pattern.

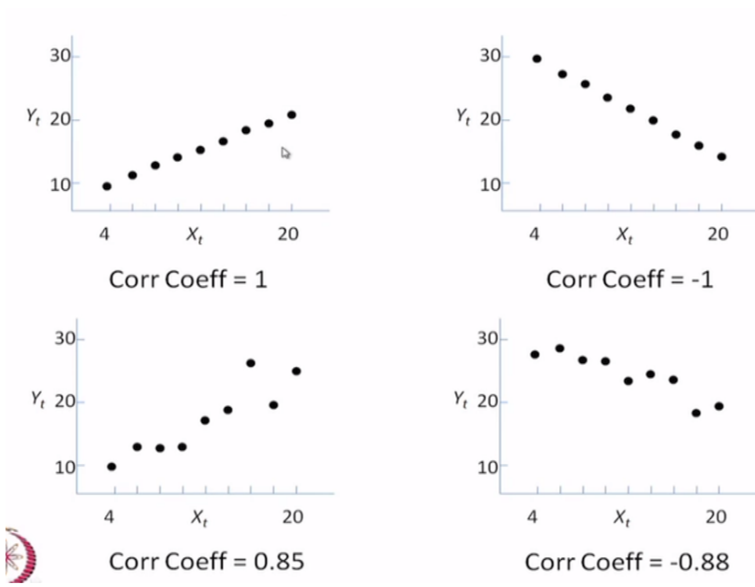
And of course there is certain amount of noise around it and there can be different other variations of these components.

(Refer Slide Time: 04:50)



Then we had also introduced the concept of correlation basically correlation is a extent of linear relationship between 2 variables x and y.

(Refer Slide Time: 05:03)

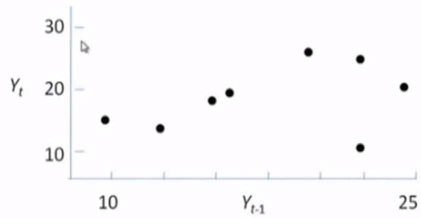


This is a case of correlation coefficient equal 1. This is a negative - 1 and these are close to 1 positive, but not 1 it is 0.85 and this is - 0.88 negative, but not exactly lying on a line.

(Refer Slide Time: 05:23)

Autocorrelation

Time, t	1	2	3	4	5	6	7	8	9
Y_t	10	15	20	12	14	18	25	20	24
Y_{t-1}		10	15	20	12	14	18	25	20



$$Correlation\,Coefficient = \frac{Cov(Y_t, Y_{t-1})}{Var(Y_t)(Var(Y_{t-1})}$$

Autocorrelation is basically defined in terms of only 1-time series, but when it is correlated with its own past values. So if we take autocorrelation between Yt and Y t - 1 we then define Yt - 1 as a 1 period lag time series so 10 and 10, 15 and 15, 20 and 20 etc. A 2 period lag time series means 10 here, 15 here, 20 here, 12 here, etc and then we find out how Yt is correlated with its 1 lag that is autocorrelation and correspondingly we can find out autocorrelation coefficient.

(Refer Slide Time: 06:12)

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

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

Next we had introduced the concept of time series smoothing. Basically whenever there is a time series Xt having certain old values, then we say a smoothed value is a weighted average of the past values. The weight is adding up to 1 and each weight > 0.

(Refer Slide Time: 06:37)

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

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

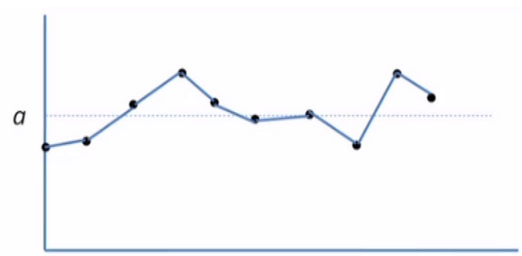
If $\{w_t\} = \{0.5, 0.2, 0.2, 0.1, 0\}$

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

This we had used in certain examples. Suppose that we have 5 values of Xt and if the constant weight Wt is given as 0.2 to each one, 0.2 is > 0 and when we add them 0.2 * 5 it becomes 1 therefore average of Xt calculated that the 10.5 = 0.2 * 5 + 0.2 * 12 etc = 11. Whereas if we change the weights still maintaining that some of the weights = 1, then the same quantity using the same values, the weights are different now 0.5 * 5 etc. The value of X5 is 10.5.

(Refer Slide Time: 07:30)

Let $X_t = a + \varepsilon_t$



Then $E[X_t] = E[a] + E[\varepsilon_t] = a + 0 = a$

For such a time series, the best forecast at time $N + 1$ is a.

The smoothed value is the best estimate of a.

Thus $F_{t+1} = \bar{X}_t$

And we say that if time series contains an average and a noise superimposed on this, then the expected value of the time series at any time is nothing but the average itself meaning the smoothed value of this. So we calculated the smoothed value of X at time t and then project that as the forecast in the next time period. So we say that the forecast at time t + 1 when we are here, we say that the forecast of the time series at this point will be the average value calculated at this point. So, when we are here, for example, we shall calculate the average value of the time series at this point and then we say that that is the forecast in the next time period.

(Refer Slide Time: 08:22)

Moving-Average Method

Consider the time series

$$\{X_t\} = X_1, X_2, \bullet\bullet\bullet, X_{t-N+1}, \bullet\bullet\bullet, X_{t-1}, X_t$$

Consider the most recent data values

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

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

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

The moving-average period is basically that we take only a constant moving-average period. Suppose we decide to take only the most recent N data points that is the moving average period. We may have many more data points, but we will take only the most recent N data points. Then the average value will be calculated in this manner. The average of that.

(Refer Slide Time: 08:50)

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:

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

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

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

And when the next data point is available, we discard the old value and add the new value and thereby we recalculate the value of X at time t + 1.

(Refer Slide Time: 09:07)

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	6			
2	8			
3	10	8		
4	3	7	8	-5
5	11	8	7	4
6	10	8	8	2
7	6	9	8	-2
8	8	8	9	-1

This was the example which we had considered in our last example before we ended the class. Suppose that we have the time series values for 8 data point, 8 time points. The values are 6, 8, 10, 3, 11, 10, 6, and 8 and suppose that we take the moving average period as 3 and calculate from here it means that when we are at time period 3, we shall average these 3 values and take that as the forecast for the next time period.

So, average value of this is 8. 6 + 8 + 10 24/3 is 8. 8 is taken as the forecast for the next time period and when we come to period 4, we shall describe this one. We shall instead include this in our consideration. So we will consider the values 8, 10, and 3 that makes it 11 + 10 21/3 is 7. So that is the moving average value calculated at time point 4. This is taken as the forecast for the next time period. Similarly we go to the next time period 5 we will consider again the 3 data points (()) (10:41) data points 11, 3, and 10 that is 24/3 is 8.

So you take this as the forecast for the next time period. Like this we continue. Now we can calculate the forecast error which is defined as the value of the time series data time t and its forecast made. For example, at period 4 we have made a forecast of 8, but the actual value is 3 therefore the forecast error is – 5. At time point 5 the forecast error is 11 - 7 is 4. Time point 6, the error is 10 - 8 is 2 and so on so forth. These are the forecast error.

(Refer Slide Time: 11:39)

Single Moving Average Applied to Data with Average and Trend

Period	X_t	Single Moving Average $\overline{X}_t$	Forecast F_t	Error $X_t - F_t$
1	2			
2	5			
3	8	5		
4	11	8	5	6
5	14	11	8	6
6	17	14	11	6
7	20	17	14	6
8	23	20	17	6

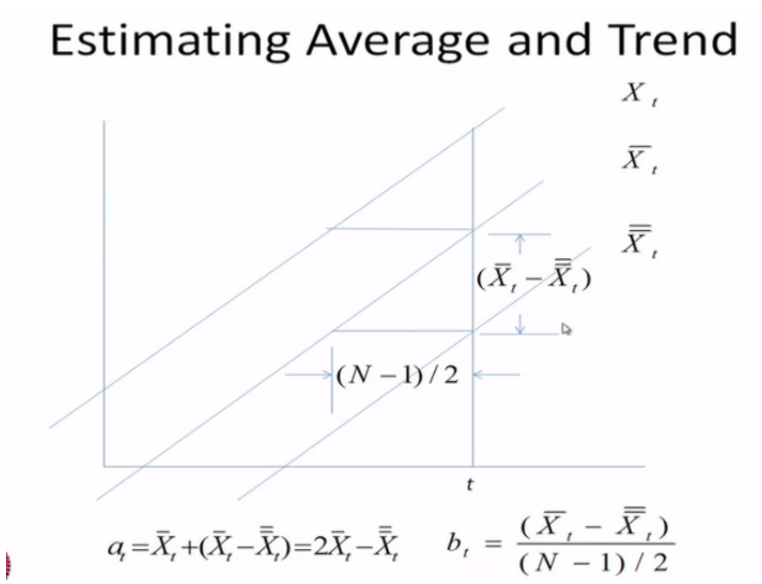
$N = 3$, Lag = 1. In general, $Lag = ½(N – 1)$

Now suppose that the time series data is all the time increasing there is an ideal case where there is no noise. It contains only an average and a trend component. You can see at time point 1, the value of x is 2. Then next period it is 5, the increment is 3. Next period it is 8, increment is still 3, 11, 14, 17, 20, and 23. So if the time series data shows a consistently rising trend, whether single moving average will be a good method we see.

Once again that we take the moving average period as 3 then the moving average value at time point 3 is 2 + 5 + 8 15/3 it is 5 and that is taken as the forecast for the period 5. Next it is 5 + 8 + 11/3 and that is 8. Likewise, we see that the single moving average value consistently lags means what was 5 is reflected in the next time period. What was 8 is reflected in the next time period. So there is always a lag of 1-time period before it can actually reflect the current value.

So we can say that if N = 3 that is the moving average period N = 3 the lag is 1. This 5 comes here after 1-time period, 8 comes here after 1-time period. So in general this lag is = N - 1/2. N in this case is 3 - 1 is 2, 2/2 is 1, so 1-time period lag and we see that the error between Xt and Ft is a constant 6. Now we can correct for this error to know or to be able to make a forecast very accurately. If you know or if you can make out what this error is we can add to X bar t to get this amount.

(Refer Slide Time: 14:14)



This is shown in this figure assuming that the value of X rises consistently in a linear fashion. Suppose that this is Xt rising in this fashion. Once we make a single moving average, once we smooth it once we know that there is a lag of sometime, which is N - 1/2. After the N - 1/2 this value is deflected as the smoothed value. We have already seen when N = 3 the value was 3 - 1/2 after 1-time period this value was obtained as the average value.

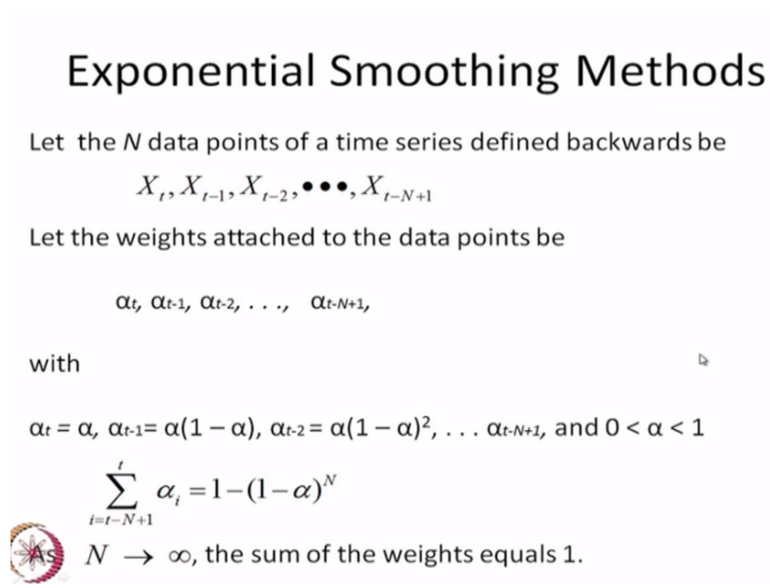
So this difference, the length of this line will be equal to N - 1/2. Now suppose that X bar t is taken it is smoothed once again we call it double smoothing and this is a single smoothing. Single smoothing of Xt is X bar t. Double smoothing of Xt is equivalent to smoothing the smoothed value of Xt meaning smoothing this. Now, if Xt rises this way, X bar t also rises in this same fashion, but it is a lagged value, lagged time series.

As we have seen in this diagram. This was 2, 5, 8, etc and this was 5, 8, 11, 14, 16 etc 1-time period lag, but it was also rising. Now if this value is once again smoothed, then this will also be rising, but it will once again show a lag and if the time period moving average period is same for the first smoothing and the second smoothing, then this distance would also be the same which is N - 1/2 and this difference is X bar t - X double bar t.

This is the value of X bar t at time t and this is the value of X double bar t at this time and that is the difference. If we know the vertical distance X bar t - X double bar t and we know the horizontal distance which is n - 1/2 then we can find the trend which is this/this. Therefore, we can find out at that means suppose that we are interested to find out this value, this will be = x bar t + this quantity and this quantity is nothing but this quantity.

So at = X bar t + X bar t - X double bar t. This quantity = this quantity. We have already calculated this. Therefore, this quantity will be = 2 X bar t - X average and the slope is = X bar t - X double bar t/N - 1/t therefore at this point suppose that we are interested to make a forecast m time period (()) (17:41) then it will be this + m * bt. If it is to the next time period, then N = 1 then the forecast for the next time period t + 1 is at calculated in this manner + bt which is this.

(Refer Slide Time: 18:09)



Let us introduce another important in fact more important than moving-average method is the exponential smoothing method because of its simplicity. Once again consider N data points of a time series, but we are writing backward defining backward meaning that the most recent value is Xt 1 period whole value is X t - 1, 2 period whole value is X t - 2, n period whole value is Xt - n + 1. We write it backward.

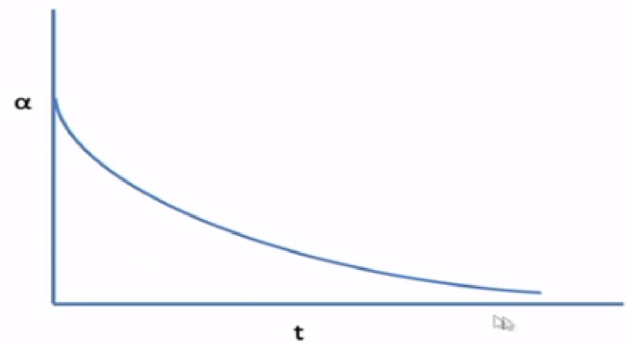
And then in order to smooth, we have to attach weights. Let the weights be alpha t, alpha t - 1, alpha t - 2, and alpha t - n + 1 and we define alpha t as = alpha, alpha t - 1 = alpha * 1 - alpha, alpha t - 2 defined as alpha * 1 - alpha square like that and so that alpha is > 0, but < 1 and as we know in order to find out the smooth value, the weights must be such that they have to be positive, but at the same time must add up to 1.

So alpha I sum will be alpha + alpha * 1 - alpha + alpha + alpha * 1 - alpha square and so on and so forth. If n tends to infinity, then this quantity is nothing but alpha/1 - 1 - alpha. This is geometric progression series with a progression ratio equals 1 - alpha which is > 0 therefore in the numerator the progression when we add it up it will be alpha/1 – 1 - alpha and that will become 1.

So if we define the weights in this fashion then the sum of the weights = 1 and each alpha is > 0, but < 1. You will see here that if alpha is < 1 then 1 - alpha is also < 1 and alpha * 1 - alpha will be < alpha. In the similar fashion alpha * 1 - alpha square will be < alpha * 1 - alpha like that it means that if we associate the weights in this fashion, then weight associated with the most recent data point is the highest.

The weight associated with the next data point 1 period old data point is < that associated with the most recent value and this is a decreasing sequence. So the weights have their values in a decreasing sequence. The highest value of the weight is given to the most recent data and the lowest value given to the oldest data.

(Refer Slide Time: 21:56)



This is shown in this fashion it is a decreasing sequence of data. Of course this is shown in a continuous manner. If t is continuous then alpha will decrease continuously in an exponential fashion and that is why the name of the technique is exponential smoothing technique, because the weights associated decrease in a negative exponential manner as if t is assumed continuous.

(Refer Slide Time: 22:53)

The exponentially weighted average is given by

$$\bar{X}_t = \alpha X_t + \alpha(1-\alpha)X_{t-1} + \alpha(1-\alpha)^2 X_{t-2} + \bullet\bullet\bullet$$

$$\bar{X}_t = \alpha X_t + (1-\alpha)\bar{X}_{t-1}$$

Whenever $X_t = a + \varepsilon_t$

the forecast for the next period is given by

$$F_{t+1} = \bar{X}_t$$

Now the exponentially weighted average is written as X bar t = alpha Xt + alpha * 1 - alpha Xt - 1 + alpha * 1 - alpha square Xt - 2 and this is nothing but alpha Xt + if you take 1 - alpha outside you get alpha * Xt - 1 + 1 - alpha * X t - 2. So that is nothing but X bar t - 1. So the exponential weighted average X bar t which is really a big expression can be certain in this fashion which is that a weight is given to the most recent value alpha and 1 - alpha is another weight given to the old average of X that was calculate 1-time period ago.

So X bar t is alpha Xt + 1 - alpha X bar t - 1. Now whenever there is only noise and average value we take the forecast for the next period as F t + 1 as = Xt itself because expected value is a and if it has no trend, no seasonality then Ft + 1 will be nothing but = the average value a X bar t. So our exponential smoothing formula is really this and when used as a forecasting method then instead of X bar t we could as well write Ft + 1. Ft + 1 = alpha Xt + 1 - alpha Ft instead of X bar t - 1 we can write Ft.

(Refer Slide Time: 24:37)

Example:

Let the time series $\{x_t, t = 1, 2, \ldots\}$ be

16, 11, 13, 12, 20, 14, 14, 16, and 18

Initialization:

$$\bar{X}_3 = (16 + 11 + 13)/3 = 13.3$$

Smoothed values are then given by:

$$\bar{X}_4 = \alpha X_4 + (1-\alpha)\bar{X}_3, \bar{X}_5 = \alpha X_5 + (1-\alpha)\bar{X}_4, \bullet\bullet\bullet$$

We can take an example to illustrate the use of exponential smoothing method. Let the time series have a value such as this and how many values we have 1, 2, 3, 4, 5, 6, 7, 8, 9 values so t = 1 to 9 and the values are like this. We can see that it has there is a noise in it looking at it we can say that there is a noise in it and there are fluctuations and we are not very sure whether there is a seasonality or not.

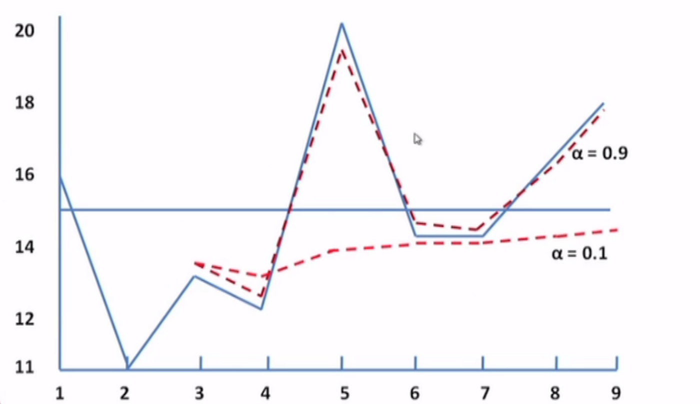
Now if you see we need in order to use this equation we need to have the first value X bar t - 1 and in this example we have only the actual time series values. We do not have any past smoothed value. So we must have to start with an initial value of X bar t. So let us we can in fact arbitrarily take 2 or 3 values, 3 initial values old values of the time series and use a moving average of the first 3 such data points to find out the initial value to use our exponential smoothing method. In this example, we have taken the moving average period as = 3 that means we have taken the first 3 data points 16, 11, and 13.

And find out an average value of that, moving average value. So 16 and 11 and 13 added/3 gives us 13.3. 13.3 is taken as X bar 3 and is used when we go the next data point, next time point 4 for the data value 12 then we take X bar 4 as = we can use our exponential smoothing formula now. We can say it is alpha * X 4 namely 12 + 1 - alpha, X bar 3 that was calculated from our moving average method.

Once we have X bar 4 we can move ahead and use that X bar 4 here to calculate X bar 5 and we can recursively use that equation to go further ahead. This is the way we proceed when we use exponential smoothing in a recursive fashion.

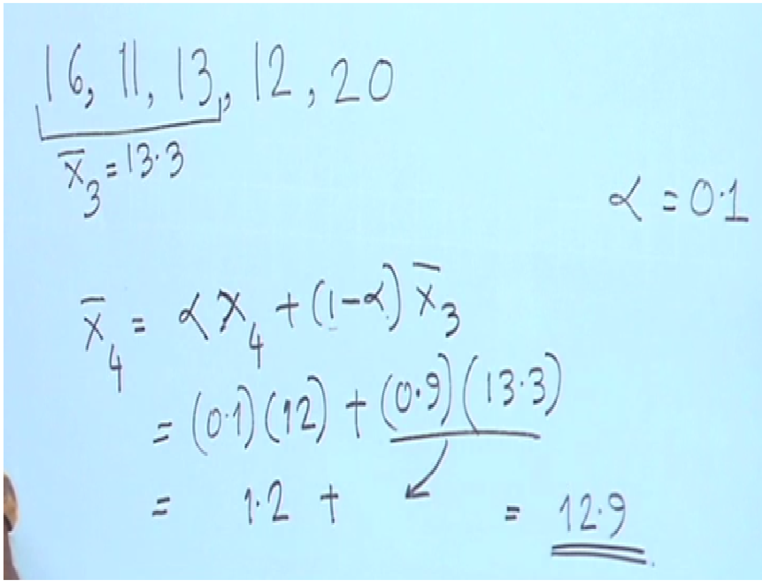
(Refer Slide Time: 27:40)

	$\bar{X}_3$	$\bar{X}_4$	$\bar{X}_5$	$\bar{X}_6$	$\bar{X}_7$	$\bar{X}_8$	$\bar{X}_9$
α = 0.1	13.30	12.90	13.61	13.65	13.68	13.91	14.32
α = 0.9	13.30	12.13	19.21	14.52	14.05	15.80	17.78



Now this is what we have tabulated in this table. First of all, recall that we have calculated X bar 3 as = 13.3 and the values of the time series data where this now X bar 3 was 13.3 and suppose that we take a value of alpha as = 0.1 then the next value would be alpha * 12. Alpha is 0.1, 0.1 * 12 is 1.2 + 1 – alpha, which is 0.9 * 13.3. Let me do it here.

(Refer Slide Time: 28:44)



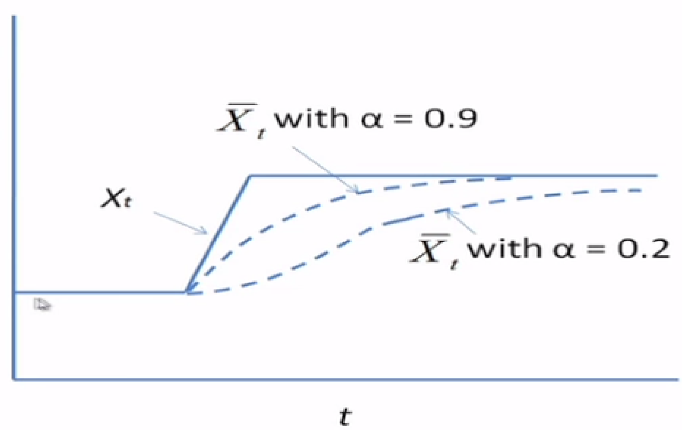
We had the data points 16, 11 now the average we have calculated here as 13.5 that is X bar 3. To calculate X bar 4, we take alpha * X4 + 1 - alpha X bar 3 and that is = suppose we take alpha = 0.1 then this is 0.1 * 12 because this is X4 12 + 1 - alpha is 0.9 * 13.3 which is X bar 3. This is = 1.2 + whatever it is coming and that is = 12.9. So this is how the value is calculated X bar 4 and now once you have X bar 4 value go to the next one you have to take 0.1 * 12.9 + 0.9 * the X5 value which is available as 20 and likewise you calculate X bar t for all the values.

Now suppose instead of taking alpha = 0.1, I take alpha = 0.9 because alpha can take a value from 0 to 1. Then the corresponding this value remains same because that was calculated on the basis of the moving average value, but this value will continue to change, but we use the same formula. Now we plot it here, this is time, and this is the data values. The firm line in blue is the actual time series data which was 16, then 11, then 13, then 12, 20, etc.

So this is the actual time series data. Smoothed values on the basis of moving average was calculated here at 10.3 and if we use 0.1 the red dotted line is this the value behaves in this fashion X bar t goes down a little bit here and then slowly rises does not consider so much of fluctuations. It does recognize that there is a rise in the average value. The average here was much less, but the average here was much higher.

It is increasing, but not a very fast rate and when we are taking alpha = 0.9 instead of 0.1 the brown red may be that dotted line is closely following the actual value. So it means a higher value of alpha gives the higher weightage to the most recent data and low weightage to the past data and therefore it has a tendency to track the change quickly. Its response is fast whereas when we take alpha = 0.1 then it is unable to track the underlying change and it smoothens the fluctuations quite a lot.

(Refer Slide Time: 33:19)



This is shown here, suppose that there is a step rise in the value of the X that is the time series data, suppose if there is no noise, no trend, no seasonality etc, then the average value remains same, but when there was a step change with a small value of alpha it increases slowly, but with the high value of alpha its response is very fast that is the advantage of taking a higher value of alpha because it keeps higher weightage to the most recent data and that is close to 1.

(Refer Slide Time: 34:06)

Low value of α

- It gives less weight to the most recent data.
- It smoothens out the noise.
- It cannot track the underlying changes very fast.

High value of α

- It gives more weight to the most recent data.
- It tracks underlying change quite fast.

So a low value of alpha gives less weight to the most recent data. It smoothens out noise, but it cannot track the underlying changes very fast whereas a high value of alpha gives more weight to the most recent data and tracks the underlying change quite fast.

(Refer Slide Time: 34:43)

Forecasting with Adaptive Smoothing Constant

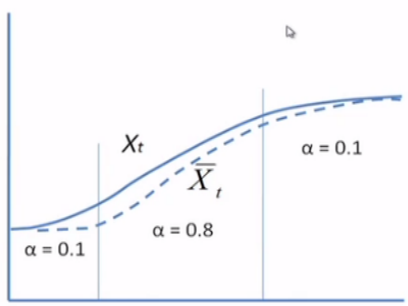
$$\alpha_{t+1} = \left| \frac{E_t}{M_t} \right|$$

Smoothed Error

$$E_t = \beta e_t + (1 - \beta) e_{t-1}$$

Smoothed Absolute Error

$$M_t = \beta |e_t| + (1 - \beta) M_{t-1}$$



Now that we know that the value of alpha can be high or low and higher value of alpha is able to track the underlying change, but the lower value of alpha can smoothen out random fluctuations. If the data does not contain any underlying any trend, but we use a higher value of alpha then it will have a tendency to even track the random noise that means the forecast would be as noisy as the original data and that is not what we want.

We would like to track only when there is an underlying change in the value of the average. We do not want to track the noise so whenever we do not foresee any change permanent change in the average, the change is noisy, the change is random fluctuation then we use a small value of alpha, but if we have strong reasons to believe that the average itself is undergoing a change then we take a higher value of alpha.

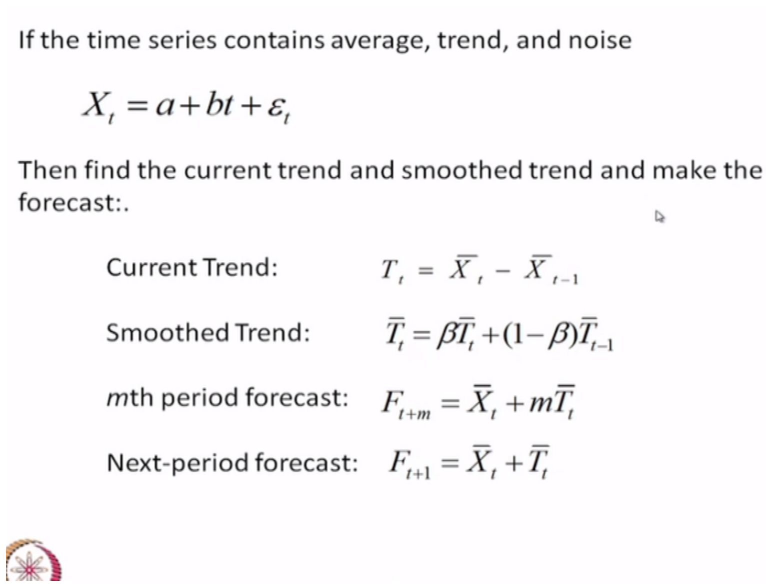
This has led to using a method where alpha is changed adoptively so that is called forecasting with adaptive smoothing constant. This alpha is called smoothing constant and with change the value of the smoothing constant we change it depending on the error forecast error. Now this is an example we visualize first of all in this diagram first let us say that in the beginning there was not much of a change in Xt so we can take a small value of alpha.

But here Xt is undergoing a change almost a permanent change is occurring, so we should be able to understand that there is an underlying change and therefore change of alpha from a low value to a higher value such as 0.8, so that it quickly tracks the change and once this is more or less stabilized there is not much of a change, it is stabilizing then we reduce the value of alpha so that the noises are not tracked so much, so this can be done by using a method which works in this fashion.

Here we first of all find out the error. Error we have already known Xt - ft that is the error. So that is Et. Et is Xt - Ft. So error is first of all smoothed giving a smoothing constant beta. Beta et + 1 - beta et - 1 exactly the similar formula to find out smoothed error. Then the smoothed absolute error we find out that is Et absolute value and find out its smoothed value we call it smoothed absolute error.

The ratio of the smoothed error to the absolute error is taken as the smoothing constant for the next time period and this is used in our calculation to make the forecast and this is how one can change alpha here for example if we continue to use alpha = 0.1 the forecasting error will be high. When the forecasting error is high alpha will be changed following this formula and that will give a higher value of alpha and at this point when forecast error will come down then alpha t - 1 this value will automatically be reduced to a value close to 0.1 or so. This is the forecasting method with smoothing constant that adopts itself with the forecast error.

(Refer Slide Time: 39:29)



Now if the time series contains average trend and noise a + bt + epsilon t then what is done? First of all we calculate the current trend. Current trend is Xt - X bar t - 1 of course divided by 1-time period. So it is X bar t - X bar t - 1. This is once again smoothed because there may be some fluctuations around the trend. Trend may not be a permanent trend.

So we smooth the current trend and use that current trend to project the future. So 1-time period projection is X bar t + just t bar t that is the forecast for the next period and X bar t is calculated following our exponential smoothing method. If it is nth period forecast, then it is m T bar t. So this is when we have a time series that contains an average a trend, bt, and the noise.

(Refer Slide Time: 40:51)

If the time series contains average, trend, and seasonality, then one has to work with de-seasonalized data by dividing the data by a seasonlity index:

$$X_t' = \frac{X_t}{I_{t-L}}$$

Overall smoothing: $\overline{X}_t = \alpha X_t' + (1-\alpha)(\overline{X}_t + \overline{T}_{t-L})$

Current Trend: $T_t = \overline{X}_t - \overline{X}_{t-1}$

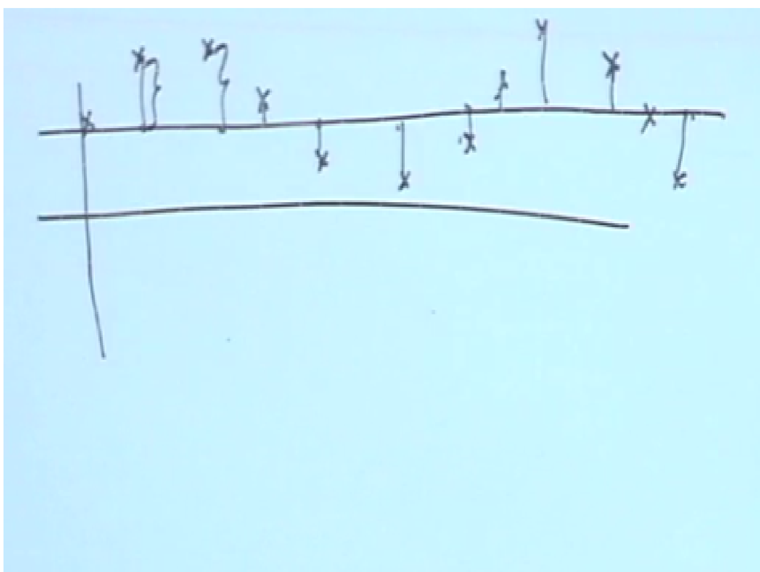
Smoothed Trend: $\overline{T}_t = \beta \overline{T}_t + (1-\beta)\overline{T}_{t-1}$

Seasonality Smoothing: $I_t = \gamma(X_t / \overline{X}_t) + (1-\gamma) I_{t-L}$

Forecast: $F_{t+m} = (\overline{X}_t + m\overline{T}_t) I_{t-L+m}$

When we have seasonal data then what we do is to first of all deseasonalize. Suppose that we have data.

(Refer Slide Time: 41:05)

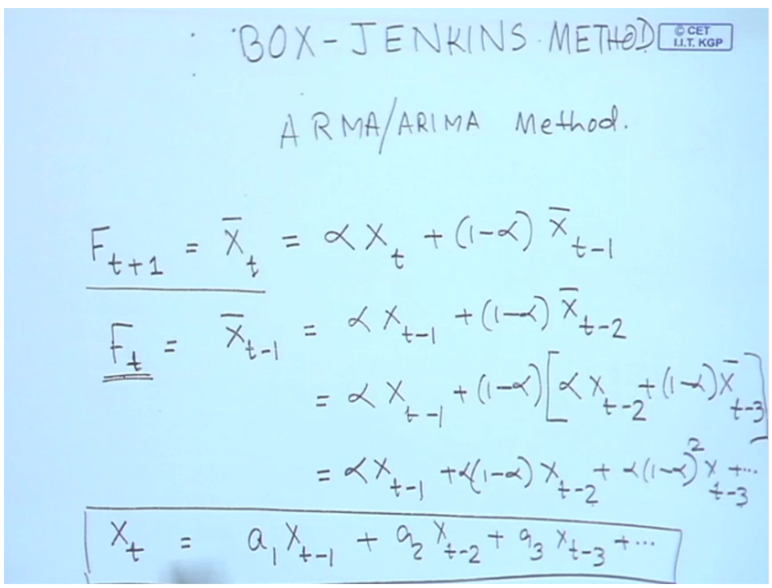


That goes up like this then what we do we find out an average value and find out to what extent the ratio of these values from the average. So they are the It values. I is the seasonality index. We calculate first of all an index I. So if we divide this by that I we get this, that is the idea. We divide this by I we get the average value. We divide this by I we get this. So first of all the values are deseasonalized so we have to find out initial value of I which is smoothed once again.

Later as you can see here using similar equation Xt - X bar t is basically IT and it is the current seasonality index. This is the L period back calculated. L is the seasonal period of seasonality. So these seasonality smoothing is done here. Overall smoothing is done here and we work with the deseasonalized data X bar X t prime. X t prime is defined in this fashion. As usual as before current trend is calculated and smoothed and I is smoothed.

The forecast finally is this smoothed value calculated in this fashion + the m period had trend added to it * the seasonality index calculated for the mth period (()) (43:08). This is how exponential smoothing is used for seasonal data. Before we close our lecture we would like to just give an introduction to a very useful, but highly sophisticated method of forecasting which is known as Box-Jenkins method or ARMA-ARIMA method. We will just give an introduction to Box Jenkins method introduced by Box and Jenkins.

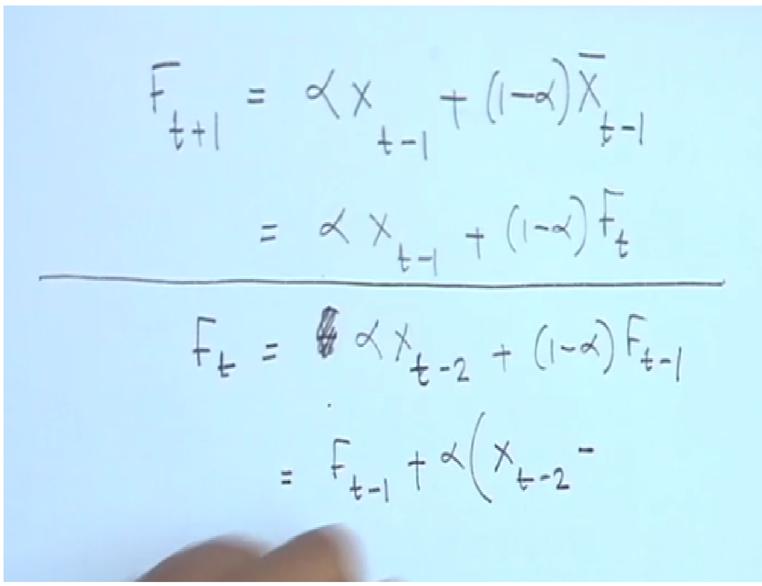
(Refer Slide Time: 43:45)



Also, known as ARMA, ARIMA method. We just give an example. Suppose, we already know that F t + 1 = X bar t that is smoothed value calculated at time period t is taken as the forecast for the next time period and that is known as alpha Xt + 1 - alpha X bar t - 1. Now we can write this. We can write Ft therefore as nothing but X bar t - 1 and we can write X bar t - 1 using it recursively as X t - 1 + 1 - alpha X bar t - 2 and this can be written as alpha X t - 1 + 1 - alpha. X bar t - 2 can be written once again recursively as X t - 2 + 1 - alpha X bar t - 3.

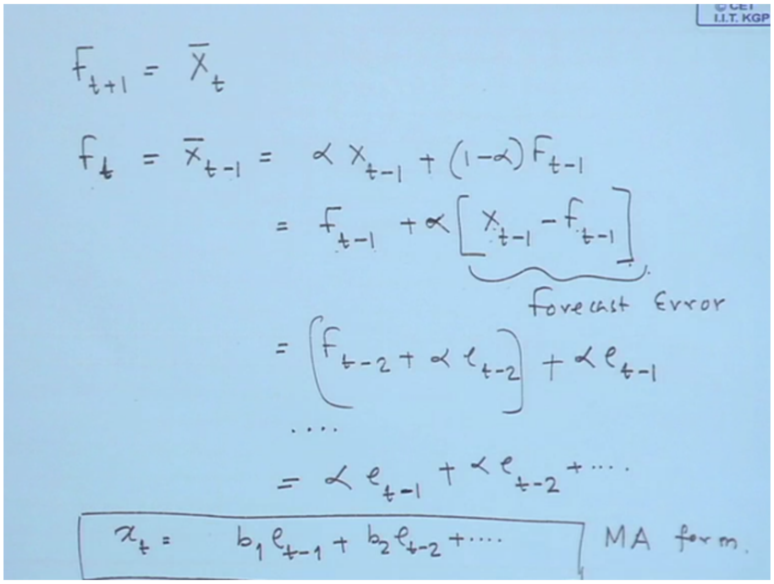
So like this if you proceed we shall get X t - 1 here + alpha * 1 - alpha X t - 2 + alpha * 1 - alpha square X t - 3 and we will proceed like this. We can therefore write you can see that forecast for the time period t is a function of the time series data X at 1 period lag, 2 period lag, 3 period lag. Therefore, there is a regression. This is a regression like equation and if we replace Ft as Xt we can say this is a1 X t - 1 + a2 X t - 2 + a3 X t - 3 so this is called an auto-regression on an autoregressive model. Keep this aside and now have the moving average form of the same thing. Now if we write.

(Refer Slide Time: 47:21)



F t + 1 as = alpha x t - 1 + 1 - alpha X bar t - 1 we can write this as alpha X t - 1 + 1 - alpha Ft forecast for the time period t is the average calculated at t - 1 therefore we can replace X bar t – 1 by Ft. This we can write as F instead of Ft + 1 we can now let us write Ft. Ft will be = alpha X t - 2 + 1 - alpha F t - 1. This we can write as 1 * F t - 1 therefore we write F t - 1 + alpha * X t - 2. I am sorry there is a mistake here. Now let us redo, there is a mistake here, I am sorry.

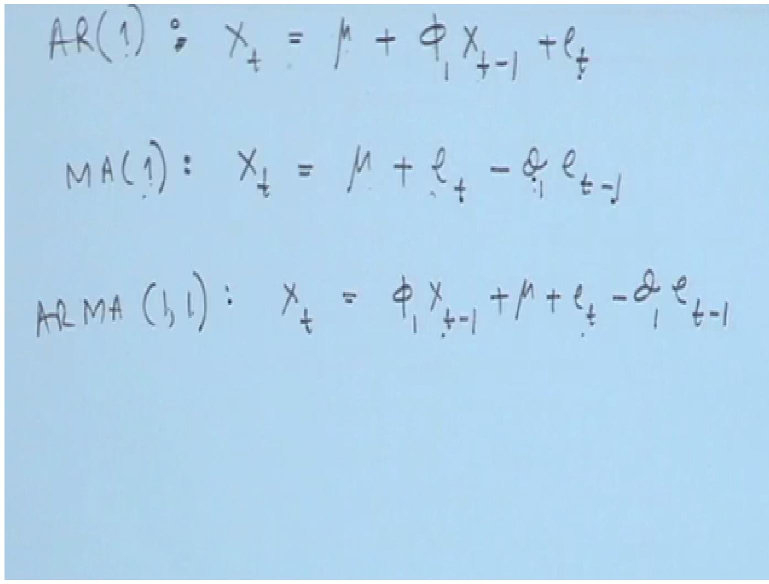
(Refer Slide Time: 49:00)



F t + 1 = we know X bar t. So F t = X bar t - 1. Now this is = alpha X t - 1 + 1 - alpha F t - 1. This is = F t - 1 + alpha Xt - 1 - Ft - 1. This is nothing but the forecast error. In fact, 1 interpretation of the exponential smoothing forecast is that forecast is the old forecast + the forecast error given some weightage alpha. Now we can expand Ft - 1 like before we can write F t - 2 + alpha e t - 2. If we call this difference as error e and this, we can write et - 1 error.
So if we proceed this way then we will land up with writing this first alpha e t - 1.

Next this alpha e t - 2 and like this. So in general we can replace Ft by Xt and we will say that Xt = b1 e t - 1 + b2 et - 2 and so on. So this is normally called a moving average form. Remember that this moving average is not the same as the moving average that we had used earlier, but unfortunately this term moving average is used in this sense. Now if we use this form which is the autoregressive form where Xt is regressed or related with its own past value.

And this is the moving average form where xt is related to the forecast error. If we combine the 2 we get what is known as ARMA method.
(Refer Slide Time: 52:20)



We will say AR1 = Xt = some constant mu + phi 1 Xt - 1 + et. Similarly, MA1 we shall say Xt as = mu + et - theta 1 et - 1 and we will say ARMA M1, 1 is Xt = phi 1 X t - 1 + mu + et - theta 1 et - 1. Basically, what we are trying to say here is that exponential smoothing method is equivalent to some sort of an autoregressive model and some sort of a moving average model of the form that I had just now developed.

AR1 is only 1 period lag value is taken that is why it is autoregressive of order 1. This is a weight phi, this is a constant mu, and this is the error term e. In the moving average form of order 1 only 1 period lagged value of the forecast error is taken which is given a weight theta and this is the current forecast error that gives the value Xt and if we use both autoregressive of order 1 and moving average of order 1 then Xt = phi 1 Xt - 1 + mu + et - theta 1 et - 1.

Now this are what is called the very elementary forms of autoregressive and moving average models used in the Box-Jenkins methodology and for short-term forecasting that considers autocorrelation this ARMA/ARIMA methods are quite important. However, time does not we do not have sufficient time to discuss these methods. We have spent nearly 4 hours on long-term forecasting which is mostly qualitative and medium-term forecasting which is usually regression.

Based and short-term forecasting that makes use of time series data in which we have introduced moving average methods and exponential smoothing methods we said that exponential smoothing methods are more convenient to use and it can be applied in various situations very easily, but the most advanced method of time series forecasting which is based on ARMA that is auto-regression or moving average concepts we just introduced we will not be able to discuss them in course of these lectures. Thank you very much.