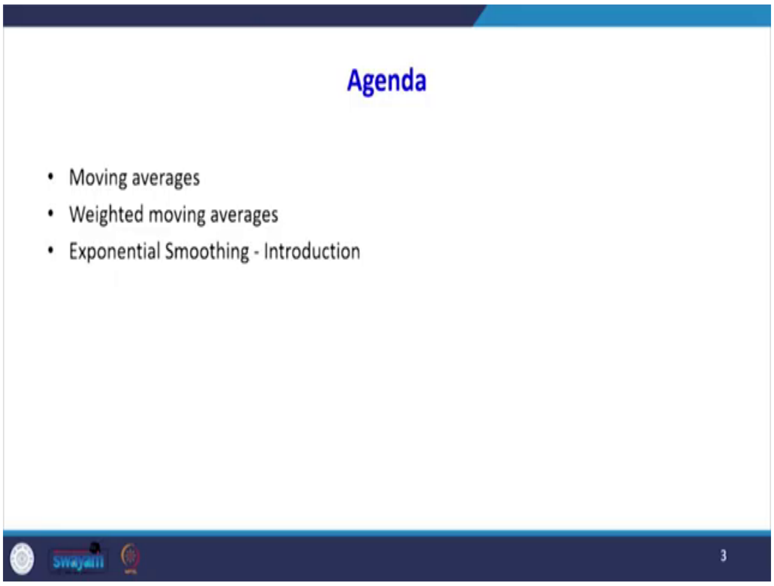
Decision Making with Spreadsheet

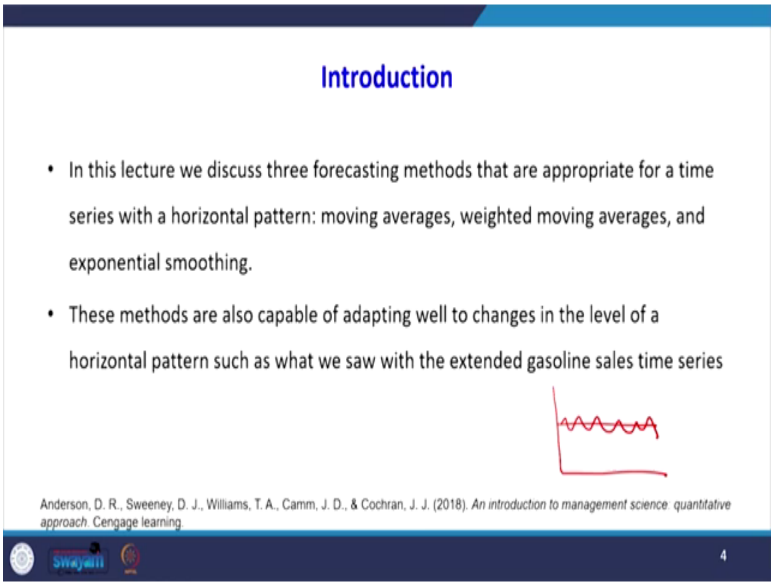
Prof. Ramesh Anbanandam
Department of Management Studies
Indian Institute of Technology – Roorkee

Lecture – 58
Time Series Analysis and Forecasting – III

So, dear students, in the previous lecture, I discussed different types of forecasting errors. In this lecture, I am going to discuss three forecasting techniques. One is the moving average method. The second one is the weighted moving averages method, and the third one is exponential smoothing. For the exponential smoothing method, I am going to cover only the introduction. In the next lecture, I will cover in detail about exponential smoothing.

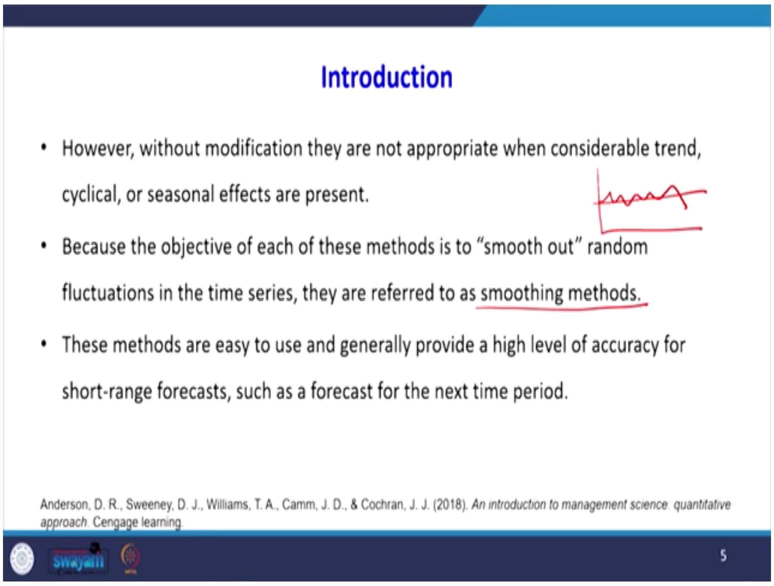


So, the agenda for this lecture is moving averages, weighted moving averages, and exponential smoothing.



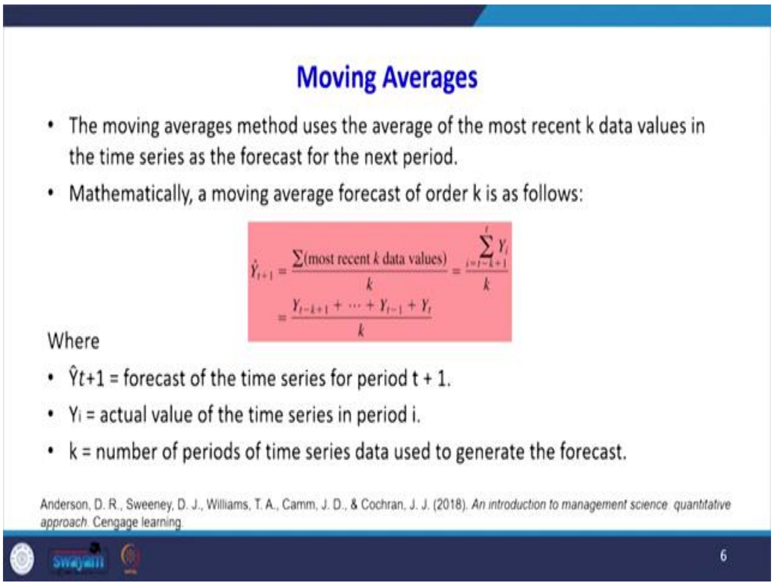
So, in this lecture, we discuss the three forecasting methods that are appropriate for time series with horizontal patterns. Those are moving averages, weighted, moving averages, and exponential smoothing. These methods are also capable of adapting well to changes in the level of a horizontal pattern, such as what we saw with the extended gasoline sales time series. So, whenever the data follows a level there is a horizontal pattern is there.

The horizontal pattern need not be exactly a straight line; it may be like this. Then, this moving average method, weighted moving average, and exponential smoothing method are more suitable.



However, without modification, they are not appropriate when considerable trend, cyclical, or seasonal effects are present. See another drawback of this moving average and exponential smoothing method is if there is any trend, so, this method cannot be used as it is. We need to have certain modifications. Because the objective of each of these methods is to smooth out random fluctuation in the time series, they are referred to as smoothing methods.

Actually here, these methods are helping to smooth out the random fluctuations. If there is any trend, this method is not suitable. These methods are also easy to use and generally provide a high level of accuracy for short-range forecasts, such as forecasts for the next period.



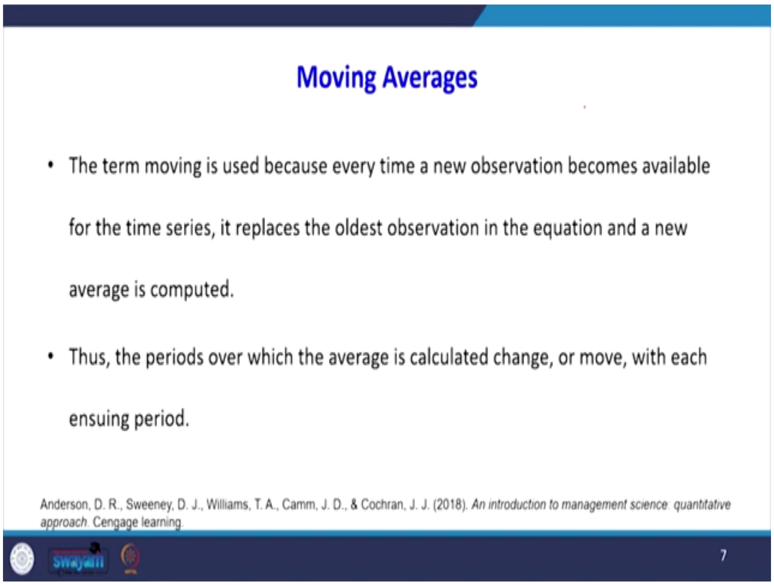
So, first, we will discuss moving averages. The moving averages method uses the average of the most recent k data values in the time series as the forecast for the next period. Mathematically, a moving average forecast of order k is as follows.

$$\hat{Y}_{t+1} = \frac{\sum(\text{most recent } k \text{ data values})}{k} = \frac{\sum_{i=t-k+1}^{t} Y_i}{k} = \frac{Y_{t-k+1} + \cdots + Y_{t-1} + Y_t}{k}$$

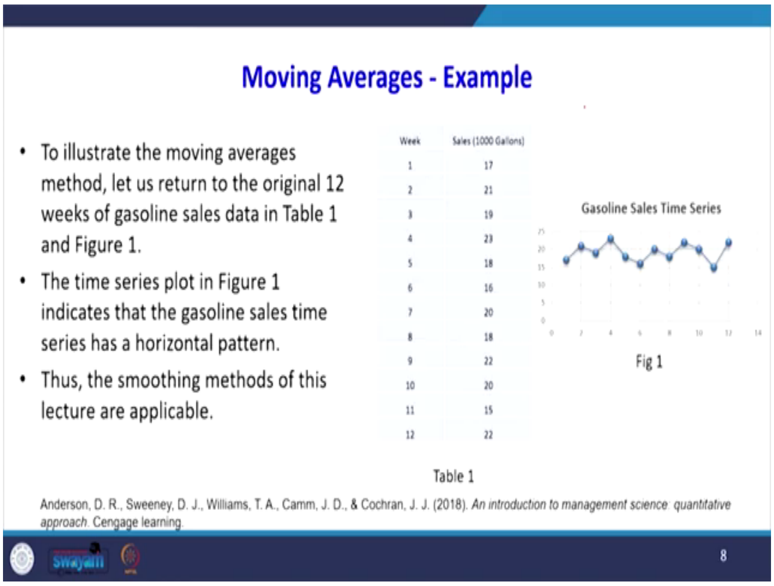
Where

- $\hat{Y}_{t+1}$ = forecast of the time series for period t + 1.
- Y_i = actual value of the time series in period i.
- k = number of periods of time series data used to generate the forecast.

So, how is this formula exactly used? I will explain, with the help of an example in the coming slides.

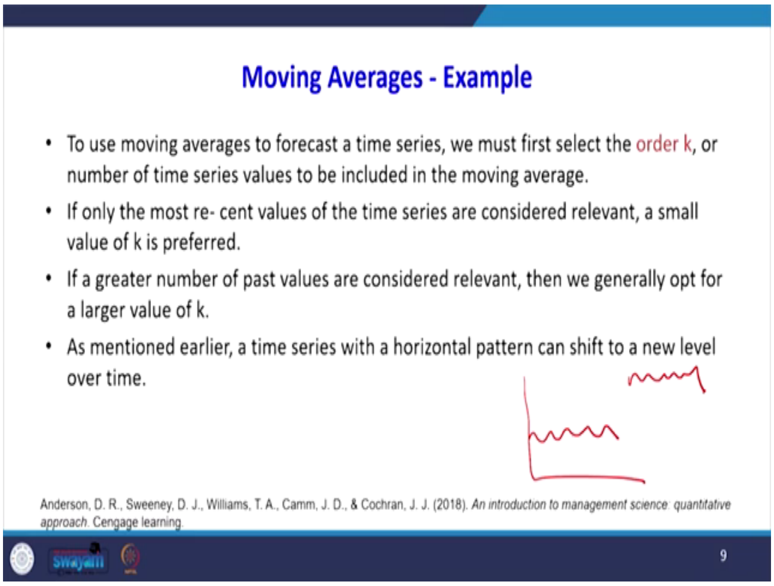


The term moving is used because every time a new observation becomes available for the time series, it replaces the oldest observation in the equation and a new average is computed. So, the average is moving. That is why it is called the moving average method. Thus, the period over which the average is calculated changes or moves with each ensuing period.



So, with the help of an example, I will explain the concept of moving averages. To illustrate the moving method. Let us return to our original 12 weeks of gasoline sales data, which I have explained in the previous slide, the data also given in this lecture in this presentation. So, the time series plot for the given data indicates that the gasoline sales time series has a horizontal pattern.

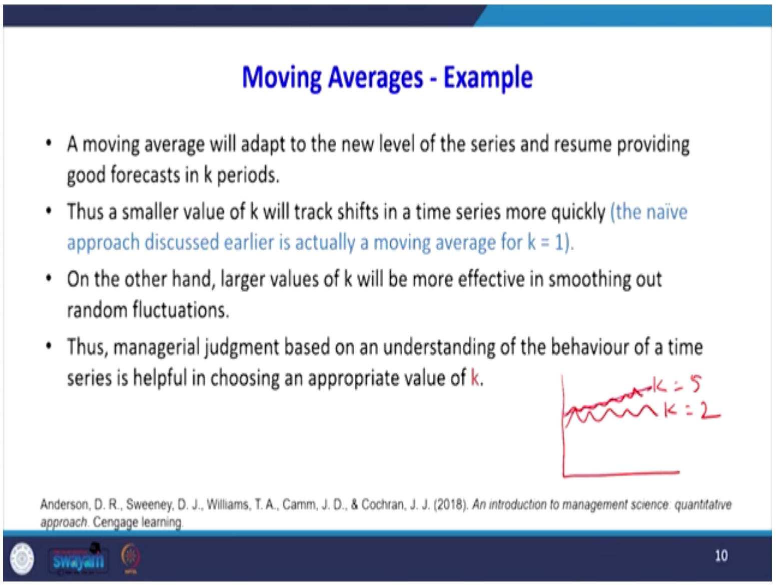
You see that when I plot to see that there is a horizontal pattern there. There are slight fluctuations, but the total effect is only a horizontal pattern. Thus, smoothing methods of these moving averages are more suitable for this kind of data.



To use moving averages to forecast a time series, we must first select the order k or the number of time series values to be included in the moving averages, which is the meaning of k. In our formula also, the k is the number of time series values to be included in the moving average. If I say k = 3, the last 3 values will be used for finding forecasting for the next period.

If only the most recent values of the series are considered relevant, a small value of k is preferred. If you find in the given data set you want to consider only the most recent values, then we have to choose. The value of k is very small. If the greater number of past values is considered relevant, then we generally opt for a larger value of k. As mentioned earlier, the time series with the horizontal pattern can shift to a new level over time.

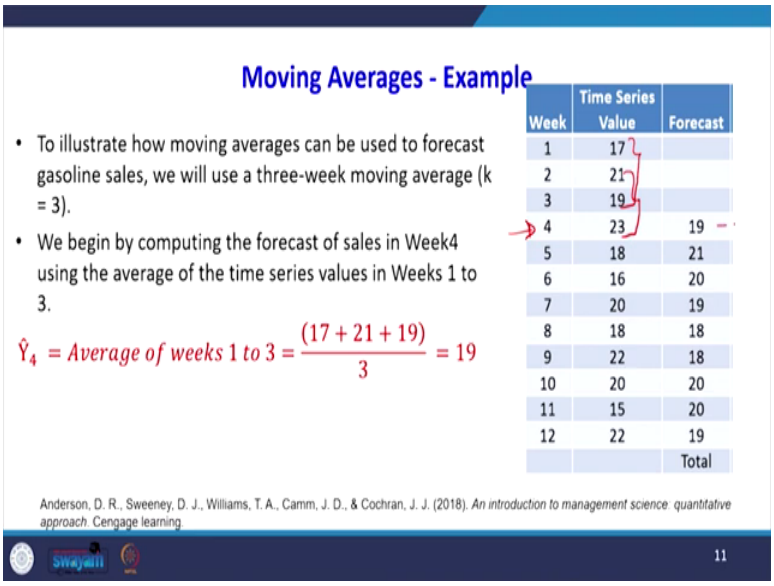
In that case, the moving average is not suitable. What is the meaning of that? Suppose the level is like this: if there is a slight shift, then this moving average method cannot be used as there are some modifications required.



A moving average will adapt to the new level of the series and resume providing good forecasts in k periods. Thus, a smaller value of k will track shifts in the time series more quickly. The naïve approach discussed earlier is actually a moving average for the k = 1. Remember the first presentation of the forecasting techniques. I have explained the naïve approach. That is nothing but the moving average where k = 1.

On the other hand, larger values of k will be more effective in smoothing out random fluctuations. When you increase the value of k it will become smoother. That is why choosing the right k is most important. Thus, the managerial judgment, based on an understanding of the behavior of your time series, is helpful in choosing the appropriate value of k. If the k is higher, for example, you plot it suppose this is k = 5.

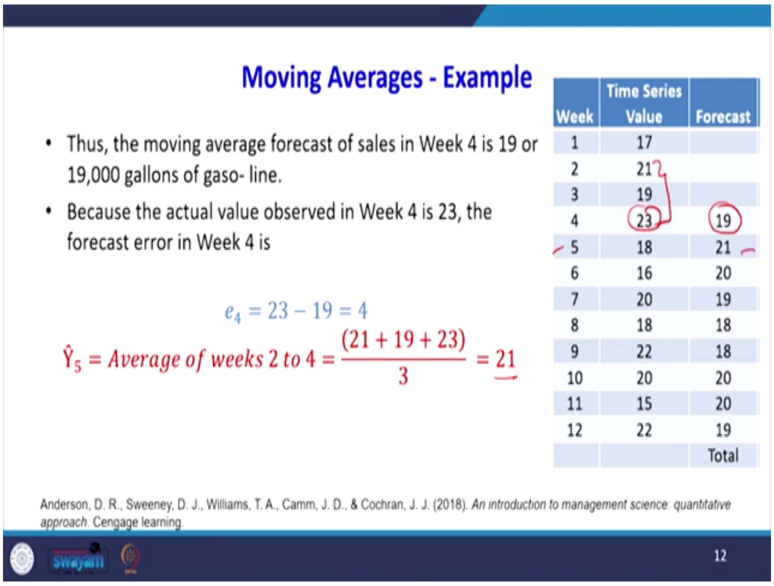
If you choose the smaller value of k. For example, k = 2 like this. So, what will happen when you choose the higher value of k is that it will become flat, and the line look flat.



To illustrate how moving averages can be used to forecast gasoline sales. We will use a 3-week moving average period. So, what is the meaning of the 3 weeks? Look at the table which is given on the right-hand side first; you will take 3 values, then find the average that will be the forecast for the 4th period. The last three periods will be the forecasted value for the fifth period. So, like this, we will continue.

So, I will explain exactly what is happening in this methodology. We begin by computing the forecast for sales in week 4. Suppose I want to forecast that I have the actual value of 23. So, Y4, the average of weeks 1 to 3, what will happen? 17, 21 and 19. So, when you find by three because it is a 3-period moving average, it is 19. So, this 19 is the forecasted value for week 4.

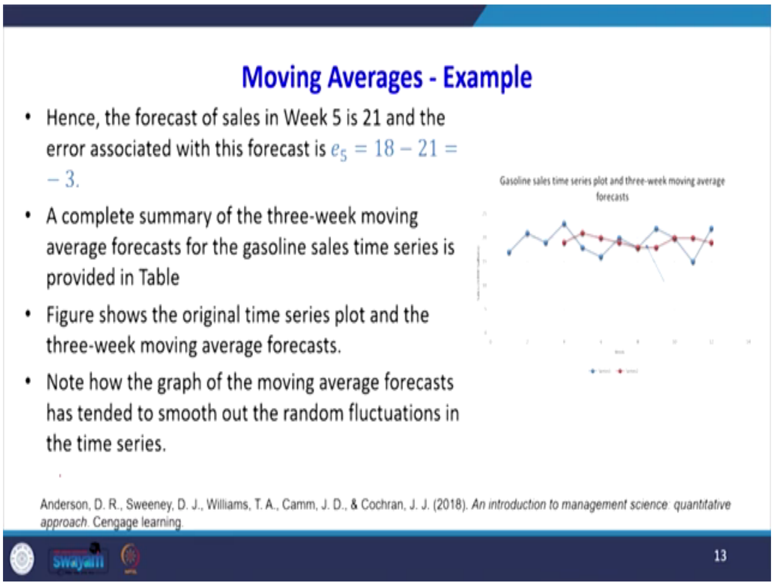
$$\hat{Y}_4 = Average\ of\ weeks\ 1\ to\ 3 = \frac{(17 + 21 + 19)}{3} = 19$$



Thus, the moving average forecast in sales week 4 is 19 are 19000 gallons of gasoline. Because the actual value observed in week 4 is 23. The error in week 4 is that we know what the error is. Actual minus predicted. So, the actual is 23, and the prediction is 19. The difference is 4. Similarly, if I want to forecast for the fifth week, so the last 3, that is 21, 29, 23 upon 3. That will provide the 21. So, this 21 is the week 5 forecast.

$$e_4 = 23 - 19 = 4$$

$$\hat{Y}_5 = Average\ of\ weeks\ 2\ to\ 4 = \frac{(21 + 19 + 23)}{3} = 21$$



Hence, the forecast of sales in week 5 is 21, and the error associated with this forecast is actual minus predicted, 18 – 21. That is –3.

A complete summary of the 3-week moving average forecast for the gasoline sales time series is provided in the table. A complete summary of the 3-week moving average forecast for the gasoline sales time series is provided in the table, which is available in the next slide. Now, look at the right-hand side.

This figure shows the original time series plot and 3-week moving average forecast. Note that the red line shows how the graph of the moving average forecast has tended to smooth out the random fluctuation in the time series. When you increase instead of k = 3 or k = 4, it will become smoothen.

Moving Averages - Example

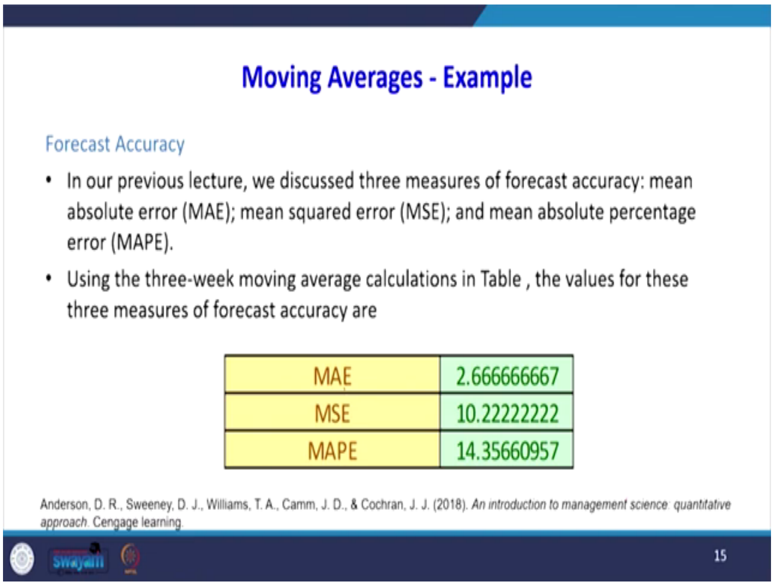
Week	Time Series Value	Forecast	Forecast error	Absolute Value Of Forecast error	Squared Forecast error	Percentage error	Absolute Value Of percentage error
1	17						
2	21		0	0	0	0	0
3	19		0	0	0	0	0
4	23	19	4	4	16	17.39130435	17.39130435
5	18	21	-3	3	9	-16.66666667	16.66666667
6	16	20	-4	4	16	-25	25
7	20	19	1	1	1	5	5
8	18	18	0	0	0	0	0
9	22	18	4	4	16	18.18181818	18.18181818
10	20	20	0	0	0	0	0
11	15	20	-5	5	25	-33.33333333	33.33333333
12	22	19	3	3	9	13.63636364	13.63636364
		Total	0	24	92	-20.79051383	129.2094862

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., & Cochran, J. J. (2018). *An introduction to management science: quantitative approach*. Cengage learning.

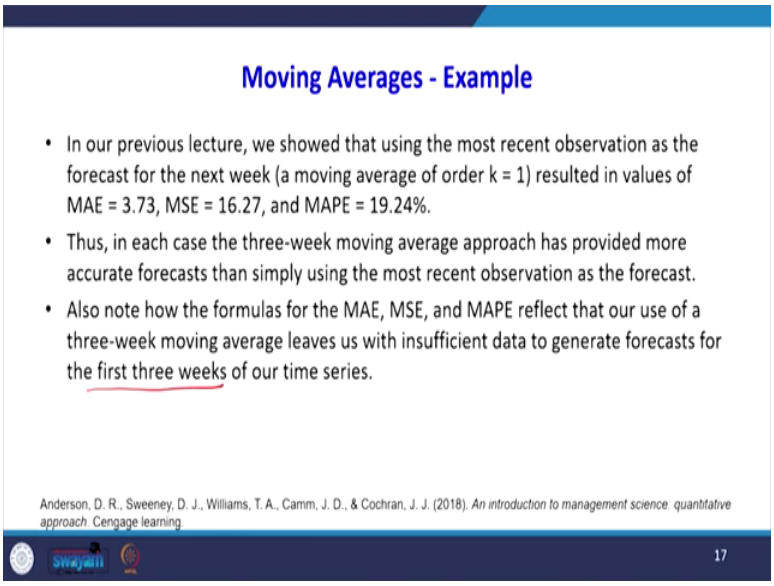
14

Now, this is the complete table, which shows the 3-period moving average. So, we have actual values there. Here, there is a forecast is there. The error is the absolute value of the forecast error. Then the squared forecast error, then the percentage error. What is the percentage error? Error, for example, is 4 upon 23. That is called percentage error, then the absolute value of percentage error. So, these values will be used.

For example, this is the total absolute value of forecast error when you divide 24 by 9 because these values only have the error. So, we will get a mean absolute error. Then, when you divide this 92 by 11, you will get a mean squared error. Then, when you divide by 129, you will get an absolute percentage error.



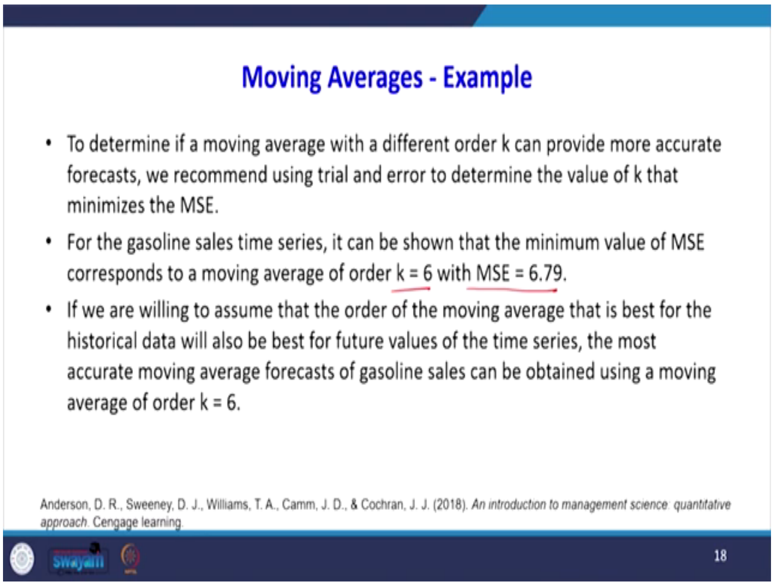
So, now for the 3-period moving averages, let us discuss forecast accuracy. In the previous lectures. We discussed 3 measures of forecast accuracy: mean absolute error, mean squared error, and mean absolute percentage error. So, using the 3-week moving average calculation table, the values of these 3 measures of forecast accuracy are for mean absolute error, which is 2.66, mean squared error of 10.22, and mean absolute percentage error of 14.35.



In our previous lecture, we showed that using the most recent observation as the forecast for the next week is a moving average of k = 1. That resulted in a value of mean absolute error of 3.73, mean squared error of 16.27, and mean absolute percentage error of 19.24%. Thus, in each case of the 3-week, the moving average approach has provided a more accurate forecast than simply using the most recent observation as the forecast.

So, when we compare it to a naïve method where k = 1 when you increase k = 3. So, the accuracy of our forecasting method has increased. Also, note how the formulas for the mean absolute error, mean squared error and mean absolute percentage error reflect That the use of a 3-week moving average leaves us with insufficient data to generate a forecast for the first 3 weeks because the first 3 weeks, we are not able to find the forecast.

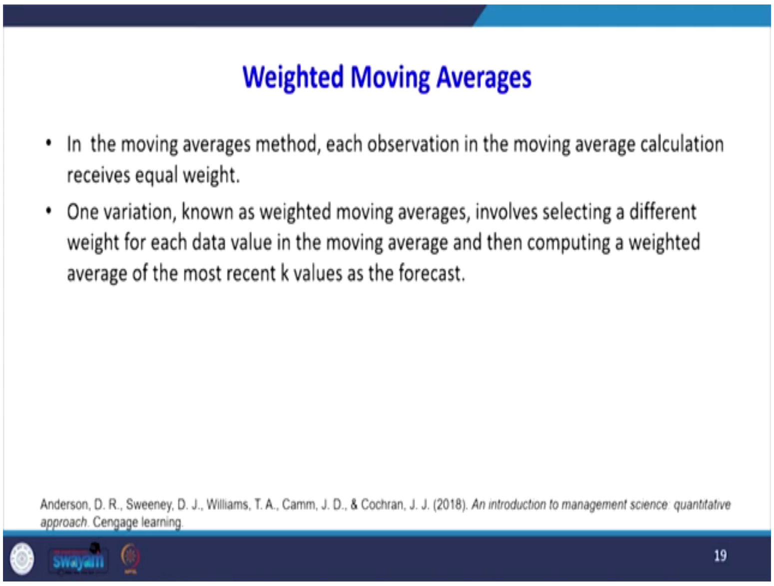
So, the first, 3 out of 12 data sets first 3 data sets will not be considered for finding the mean. Only 9 datasets will be considered.



To determine if your moving average with the different order k can provide a more accurate forecast. We recommend using trial and error to determine the level of k that minimizes the mean squared error. Many times, this will be an obvious doubt: what should be the value of k? So, what do we have to do? For the same problem, you have to use different k values for each k value. You have to find out the mean squared error for what value of k you have a lesser mean squared error.

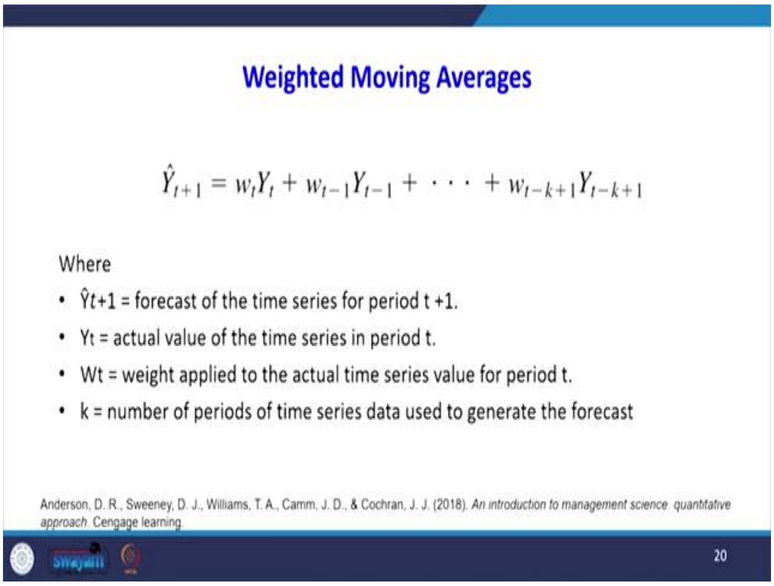
So k is the most appropriate for the gasoline sales time series. It can be shown that the minimum value of the mean squared error corresponds to a moving average of order k = 6. So, when we have done so, for k = 3, suppose you put k = 6. So we will get a more accurate value. If you are willing to assume that the order of the moving average that is best for the historical data will also be the best for future time series values.

The most accurate moving average forecast for gasoline sales can be obtained using a moving average of order k = 6. Here, it is an assumption if you want to forecast. So, we have seen a pattern of 12 weeks. In the future weeks also, if you assume that the same type of pattern is going to be followed, then this moving average method is a most accurate method.



Next, we will go for weighted moving averages, in the moving average method. Each observation in the moving average calculation receives equal weight. That was one of the disadvantages of this moving average. To overcome that, we are going to provide different weights for our past data. So, one variation when compared to the simple moving average method is weighted moving averages.

That involves selecting a different weight for each data value in the moving average and computing weighted averages of the most recent k values as your forecast.



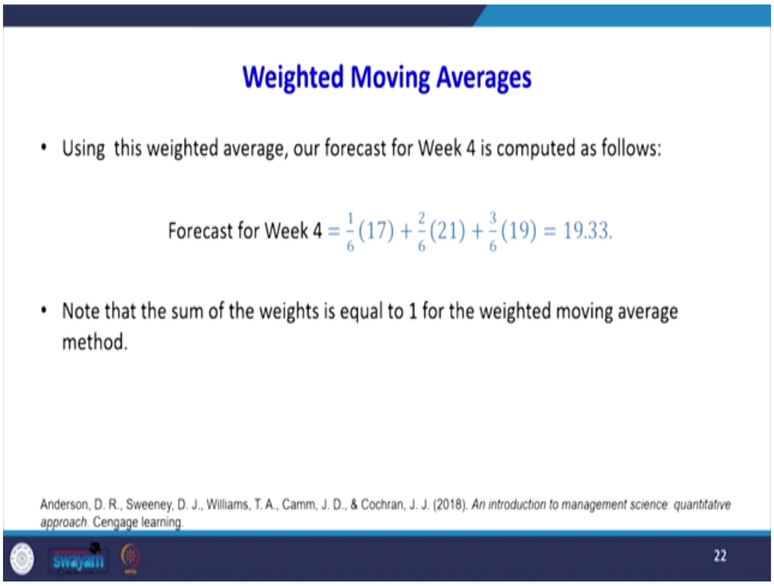
So, the formula is here: we are giving different weights.

$$\hat{Y}_{t+1} = w_t Y_t + w_{t-1} Y_{t-1} + \cdot \cdot \cdot + w_{t-k+1} Y_{t-k+1}$$

Where

- $\hat{Y}_{t+1}$ = forecast of the time series for period t +1.
- Y_t = actual value of the time series in period t.
- W_t = weight applied to the actual time series value for period t.
- k = number of periods of time series data used to generate the forecast

Weighted moving averages, generally the most recent observations receive the largest weight, and the weight decreases with the relative age of data values. Suppose we are doing 3 period moving average method. The most recent value is 19. So, the 19th that is week 3 data will give the weightage 3 by 6. The second weak data will get 2 upon 6. This is one upon 6 that has given some values but the recent value will have the higher weightage.



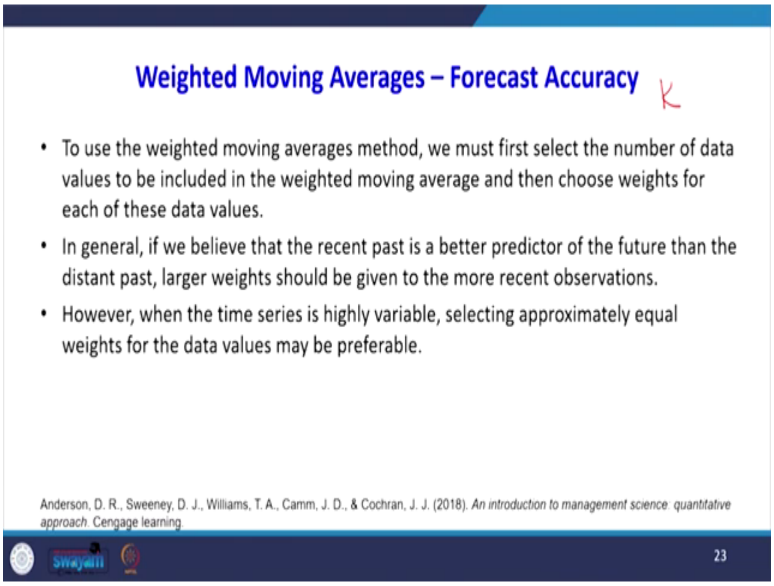
Using the weighted average the forecast for week 4:

- Using this weighted average, our forecast for Week 4 is computed as follows:

Forecast for Week 4;

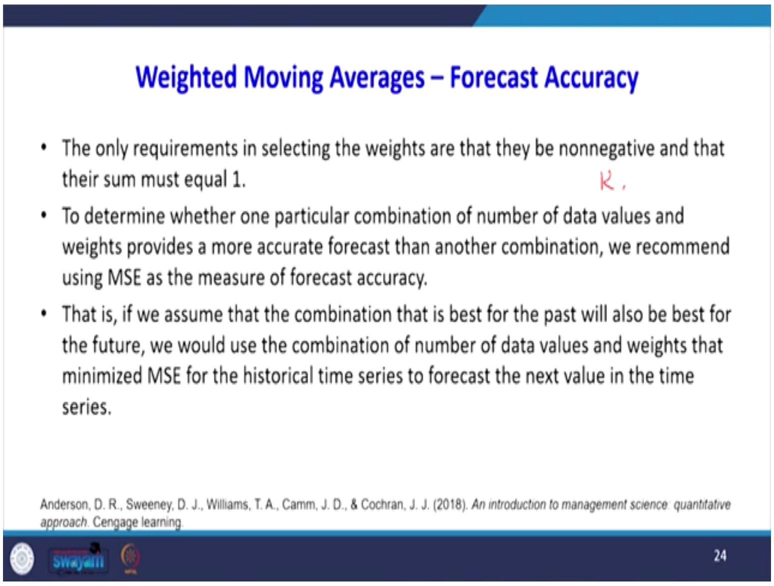
$$= \frac{1}{6}(17) + \frac{2}{6}(21) + \frac{3}{6}(19) = 19.33.$$

Note that the sum of the weight is equal to 1 for the weighted moving average method.



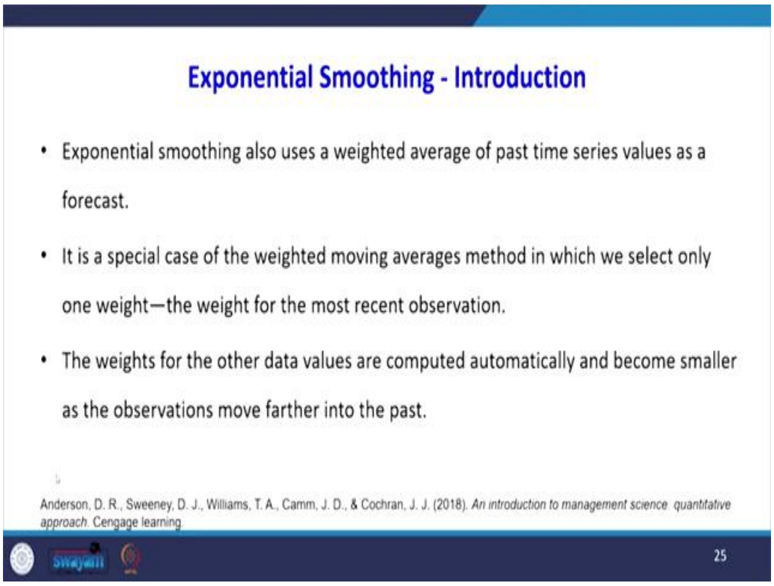
How to find the accuracy for your weighted moving averages: to use the weighted moving averages method, we must first select the number of data values to be included in the weighted moving averages and then choose the weight for each of these data values. So, we have to find out to k then you have to decide how much weight has to be given for each data set.

In general, if we believe that the recent past is a better predictor of the future, then the distant past larger value should be given to the more recent observations. In our example problem also, we have done the same thing; according to the recent data, we have given the higher weightage. However, when the time series is highly variable, selecting approximately equal weight for the data values may be preferable.



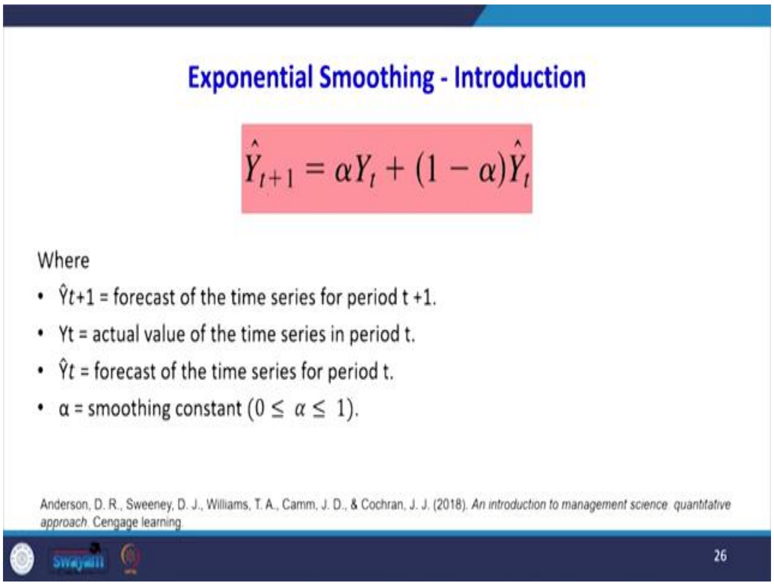
The only requirements in selecting the weight are that they may be nonnegative and their sum must be equal to 1. To determine whether one particular combination of a number of data values will provide a more accurate forecast than another combination, we recommend, as usual, using mean squared error as the measure of forecasting accuracy. So, we can change the values of k, and you can change the weightage for different sets of k and weights.

Then, you have to find out the mean squared error. So, for what value of k and the weight weightage, when the mean squared error is lesser, that is a more accurate method. That is, if we assume that the combination that is best for the past will also be the best for the future. We should use the combination of a number of data values and the weight that minimized mean square error for the historical time series to forecast the next value
in the time series.



The third method which I am going to discuss in this lecture is exponential smoothing. Exponential smoothing also uses a weighted average of past time series values as a forecast. It is a special case of the weighted moving average method in which we select only one weight. That is the weight for the most recent observations. The weight for the other data values is computed automatically.

And become smaller as the observations move farther into the past. So, we are going to consider it alpha. Another name is the smoothing constant.



So, the model for exponential smoothing is

$$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$$

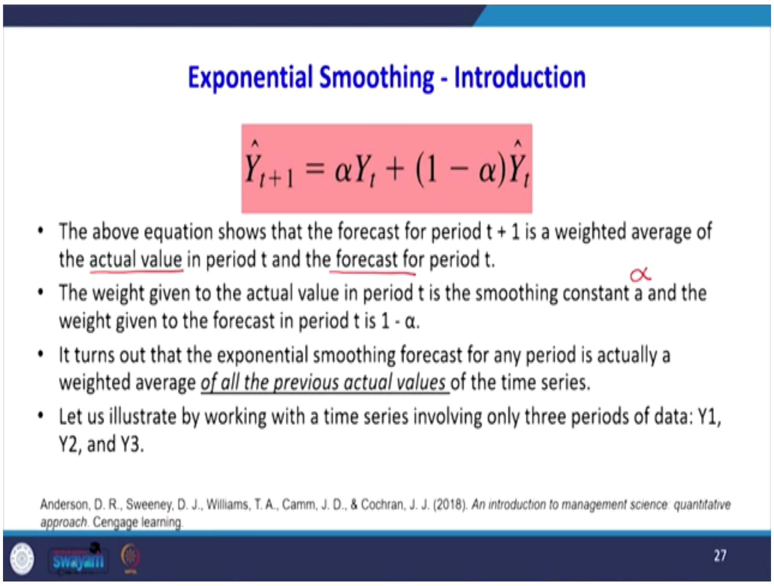
Where,

- $\hat{Y}_{t+1}$ = forecast of the time series for period t +1.
- Y_t = actual value of the time series in period t.
- $\hat{Y}_t$ = forecast of the time series for period t.
- α = smoothing constant

$(0 \leq \alpha \leq 1)$

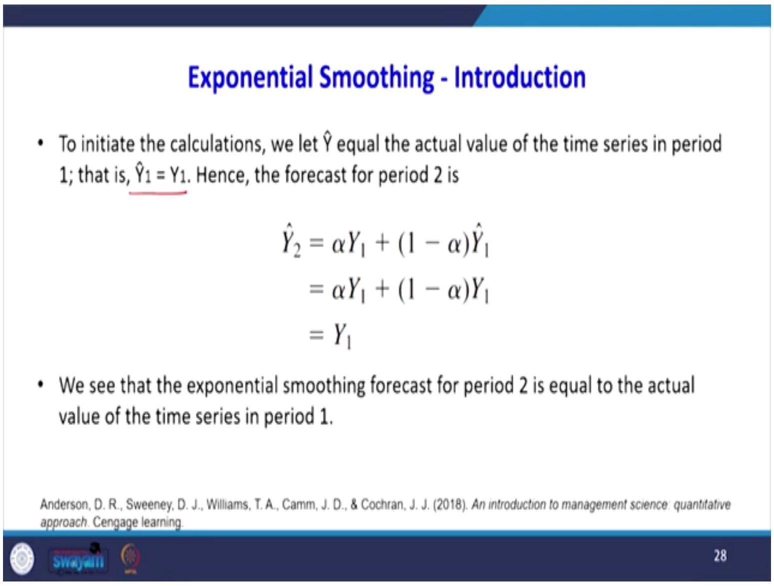
.

The value of alpha should be between 0 and 1.



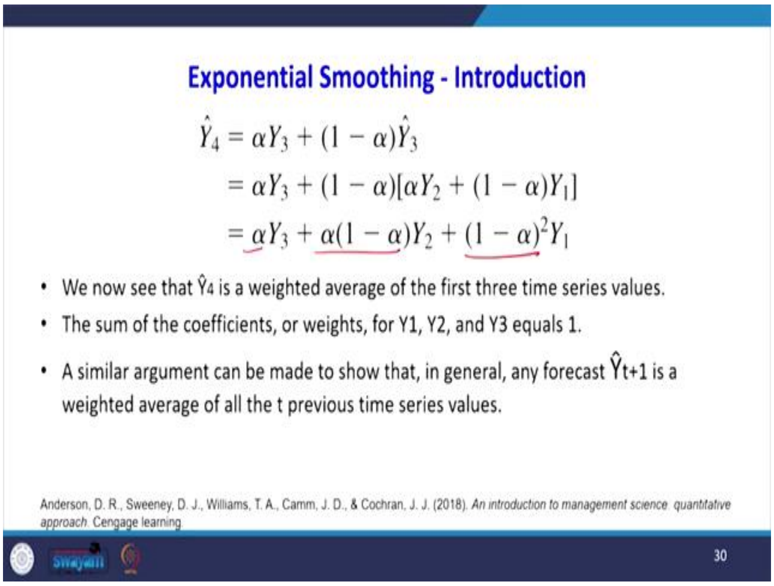
The above equation shows that the forecast for the period (t + 1) is a weighted average of the actual values in period t and the forecast for period t. So, we are taking 2, considering one is the actual value, which has the weight of alpha, and the forecast value, which has the weight of (1 – α). The weight, given to the actual value in period t, is the smoothing constant alpha, and the weight given to the forecast in period t is (1 – α).

It turns out that the exponential smoothing forecast for any period is a weighted average of all of the previous actual values of the time series. In the coming slides, I will explain how the value of alpha decreases. Let us illustrate by working with the time series involving only 3 periods of data: Y1, Y2, and Y3.



To initiate the calculation, we let Ŷ hat equal the actual value of the time series in period 1. So, we assume that the forecasted value for Y period 1 is the actual value. Hence, the forecast error for period 2 is:

$$\begin{aligned}\hat{Y}_2 &= \alpha Y_1 + (1 - \alpha)\hat{Y}_1 \\ &= \alpha Y_1 + (1 - \alpha)Y_1 \\ &= Y_1\end{aligned}$$



The forecast for period 3 is

$$\hat{Y}_3 = \alpha Y_2 + (1 - \alpha)\hat{Y}_2 = \alpha Y_2 + (1 - \alpha)Y_1$$

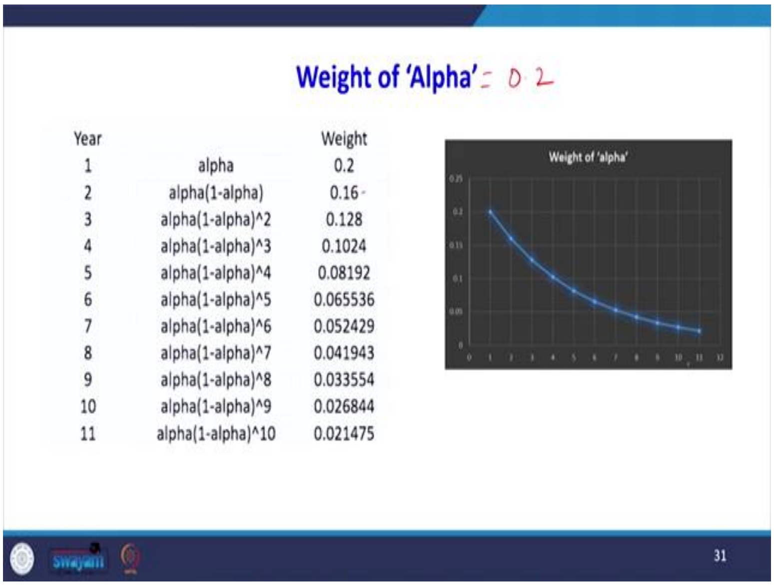
Finally, by substituting this expression for Ŷ3 with the expression for Ŷ4, we obtain

$$\begin{aligned}\hat{Y}_4 &= \alpha Y_3 + (1 - \alpha)\hat{Y}_3 \\ &= \alpha Y_3 + (1 - \alpha)[\alpha Y_2 + (1 - \alpha)Y_1] \\ &= \alpha Y_3 + \alpha(1 - \alpha)Y_2 + (1 - \alpha)^2 Y_1\end{aligned}$$

We now see that Ŷ4 is a weighted average of the first three-time series values.

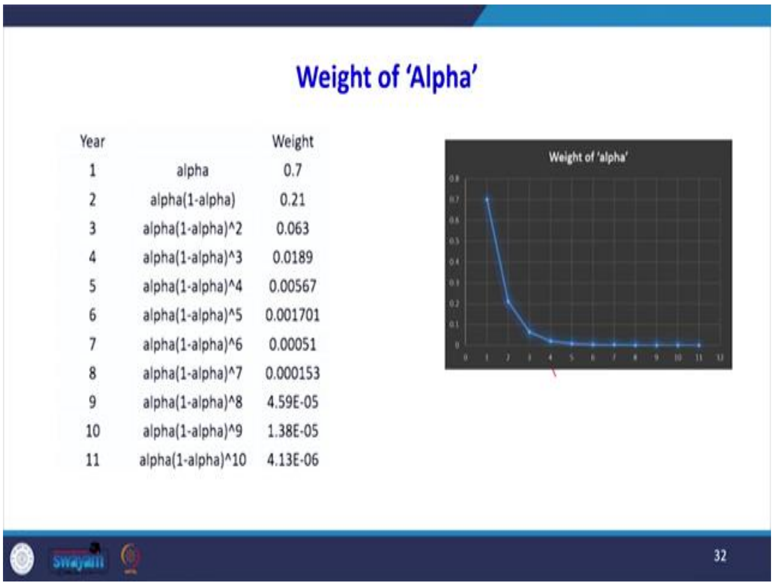
The sum of the coefficients, or weights, for Y1, Y2, and Y3 equals 1.

A similar argument can be made to show that, in general, any forecast Ŷt+1 is a weighted average of all the t previous time series values.



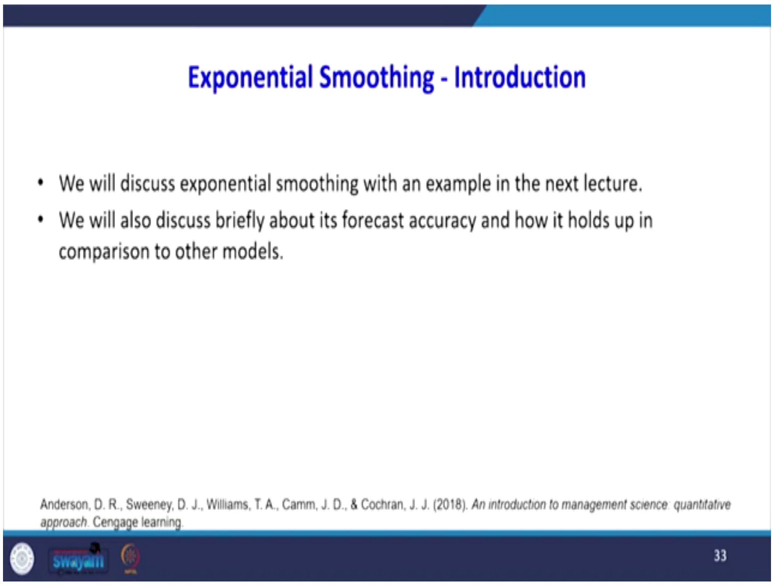
Year		Weight
1	alpha	0.2
2	alpha(1-alpha)	0.16
3	alpha(1-alpha)^2	0.128
4	alpha(1-alpha)^3	0.1024
5	alpha(1-alpha)^4	0.08192
6	alpha(1-alpha)^5	0.065536
7	alpha(1-alpha)^6	0.052429
8	alpha(1-alpha)^7	0.041943
9	alpha(1-alpha)^8	0.033554
10	alpha(1-alpha)^9	0.026844
11	alpha(1-alpha)^10	0.021475

Here, I am going to explain the behavior of the alpha smoothing constant by looking at the picture on the right-hand side. Suppose I have taken alpha value = 0.2. So, what will happen in year 1 is 0.2; year 2, 0.16, and year 3 see that the value of alpha decreases exponentially. That is why this method is called exponential smoothing. And you see, when we take the smaller value of alpha, it is considering all 11 values years of data. So, what is happening 8th year, 9th year, 10th year, the weightage is very less.



But you see the example, suppose I take a higher value of alpha, so what will happen? It is taken see for first 4, previous data set; it has taken more values. Beyond that, for example, fifth year, 6th year, almost the weight is 0. So, this is how the exponential curve or the weightage for alpha decreases for the smaller value of alpha and the larger value of alpha. If you choose a smaller value, the weight is spread across all the data, all the years.

If you take the higher value of alpha, most of the weights are consumed only for the 2 or 3 previous years. Beyond that, it is almost the weights are 0.



We will discuss exponential smoothing with an example in the next lecture. We will also briefly discuss its forecast accuracy and how it holds up in comparison to other methods. Dear students, in this lecture, I discussed three forecasting techniques. One is the moving average method. The second one is the weighted moving average method. In the third one, I will give an introduction to exponential smoothing.

Then, I explained the behavior of the smoothing constant. In the next lecture, we will continue with the example problem using exponential smoothing, and also, I will explain how to use regression analysis for trend forecasting. Thank you.