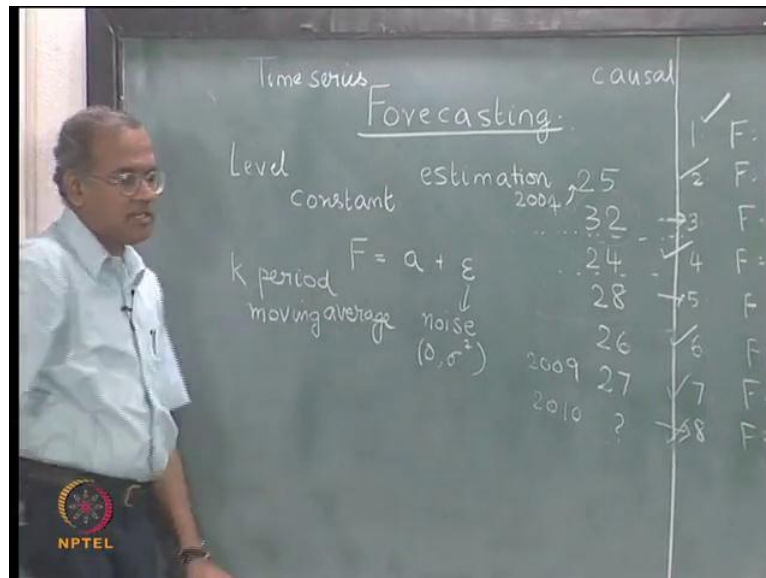


Operations and Supply Chain Management
Prof. G. Srinivasan
Department of Management Studies
Indian Institute of Technology, Madras

Lecture - 2

Forecasting - Time Series Models - Simple Exponential Smoothing

(Refer Slide Time: 00:17)

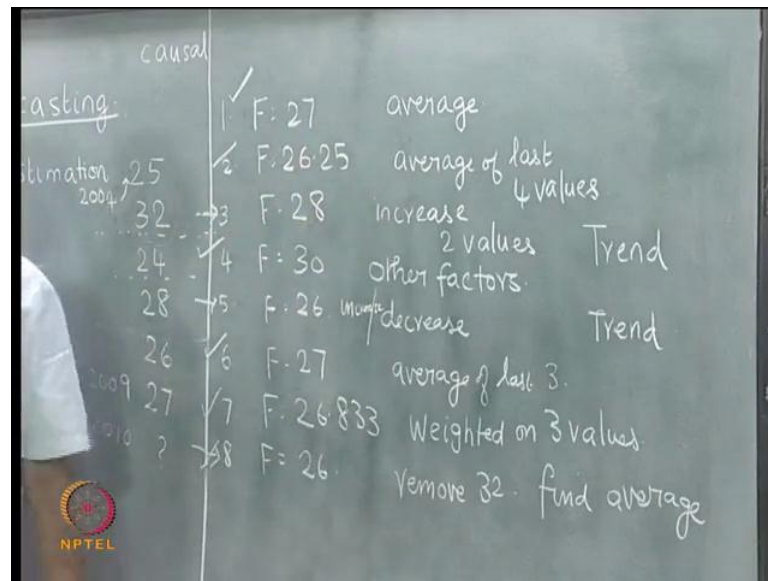


Today we begin our discussion on Forecasting; forecasting essentially deals with estimation of future demand for a product. For example, they could have demand for a particular product; let us say in the last 6 years as 25, 32, 24, 28, 26 and 27. The product could be anything, the product could be an automobile, a product could be phones, a product could be machinery and so on and these numbers could be in multiples of 1000 or multiples of lakhs and so on. For the sake of illustration we will take these 6 numbers, we can assume that this is the sale in 2009, and this is the sale in 2004.

So, the data for 6 years starting from 2004 to 2009 are available, and let us say for the sake of illustration that these are the values. Now, we want to estimate the demand for the year 2010 say and which I have indicated by a question mark, this demand estimation is extremely important because this demand estimation is a starting point of production planning, as we saw in the last lecture. Only when we have the forecasted value of the demand, we will be able to make both capacitive planning and production planning.

So, the whole area of production planning begins with the forecast, so we will spend time on learning forecasting models. So, we will start with these 6 values as values of demand and we will try to forecast for the next period, now whenever I pose such a question and give 6 values of demand, and ask a class of students to make their own forecast. I normally get several answers that the students give, and let me write down some of them on this board, and then try to explain them one by one and lead to the various basics and principles of forecasting.

(Refer Slide Time: 02:39)



So, let us consider these 6 values of forecast and let us say these have been given by 6 different people. Now, when I asked the students the basis or the reason behind which these numbers were given, I got a variety of answers, so I will write down these answers and then we will slowly see the relevance of each one of them, in the context of forecast. So, the first F is equal to 27 forecast equal to 27, comes from simple average, so one of the students believes that a simple average of these 6 numbers, adequately represents the forecast.

And we add all this and divide it by 6, we get f is equal to 27, another student said that he was interested only in the last four values, and this student felt that the first 2 values were little old and therefore, he is not keen on considering these 2 values. He believes that the last four values adequately represent what is happening and the value and the average of

the last four values turns out to be 26.25. A third person said that the forecast could be 28.

And the third person reasoned out this way by saying that there is a increase and then a decrease, and then an increase and then a decrease and then an increase. And so the estimate after the decrease happens to be the third student who said 28, started by saying that the student looked at the last 2 values, which happened to be 26 and 27. And the student believed that considering only the last 2 values, this seems to show an increase and therefore, the student gave forecast as 28 which is an increase based on the last 2 values.

Another person looked at these values and said that the forecast could be 30, if we assume that or if we consider that this data corresponds to sale of automobile. And this person believes that the sale of automobile is related to the population in that area and is also related to the buying power of the population. And the person believes that due to increase in buying power of the population, the demand is expected to go up from its last values. And therefore, the person estimated that 30 is a very good forecast for the year 2010.

So, this is based on other factors, now this person who gave forecast equal to 26 believed that there was an increase and a decrease. So, it shows an increase and then a decrease, and then an increase and then a decrease, now there is an increase and the person expects a decrease, and believes that the value 26 comes from decrease and increase, increase and decrease. Now, this person who gave forecast equal to 27 took the last 3 values instead of the last 4 and said $28 + 26 + 27$, so this is the average of the last 3.

The person, who gave 26.833, also considered the last 3 values, but believed that since this value is the oldest among the 3 this is older, and this is the most recent value. The person gave weights and said, this is going to have a weight age of 1, here there will be a weight age of 2, and here there will be a weight age of 3 which is the closest. So, the forecast would be $1 \times 28 + 2 \times 26 + 3 \times 27$ divided by 6 which gave 26.833, so it is a weighted average on the last 3 values.

The last person who gave value equal to 26 felt that if you look at all these values 32 seems to be an outlier, and there could have been some assignable reason for this 32 which looks like a spike. So, the person took away this 32 and then found the average for

rest of them, so this is remove 32 and find the average, now let us look at each of these 8 answers, and more importantly the reasons associated with these answers, and try and learn some aspects of [forecasting](#).

Now, this is centred around the average which is a measure of central tendency of this data. This is also centred around the average even though smaller or less amount of data and more recent data has been considered. This has a slightly different assumption that this is based on the increase that happens here, here the person has considered other factors. Here also the person has considered an increase and decrease, this is based on the average, this is also based on some kind of a weighted average, and this removes the out layer.

So, let us say out of these 8 different reasons given, 5 of the reasons are centred around the average in some form or the other. But, if we look at this data and before we start making any forecasting, we need to understand whether this data actually represents a constant or what is called a level or whether there is some kind of a trend that has happened. So, if we look at these values in some sense we do not see a clear increasing trend from this to this we also do not see any kind of a decreasing trend.

So, I am going to assume that this does not show any kind of a trend and therefore, I am going to start developing models, only for level or what I call as a constant model, which in some sense has been validated by 8 out of the 8 people saying that it is centred around the average. So, essentially we are trying to find out a constant or a number which adequately represents all these numbers, under the assumption that there is no increasing or decreasing trend.

So, which simply means we are trying to fit something of the form F is equal to a , where a is the constant or the level that we are trying to find out. But, then we also understand that if this were to represent a constant, then why do we have these variations in these values, and not the same constant that comes as a forecast. The reason behind is there is a certain inherent variation or inherent variability in the system, which we call as some epsilon. And this epsilon is the noise or inherent variation or variability in the system.

And this noise is expected to have parameters 0, σ^2 , which means this noise has a mean of 0, which also means that these inherent variation or statistical fluctuations will cancel out as time passes. And these variations are small because their variance

sigma square is also assumed to be a very small number. So, if we make the assumption that this represents a constant, and we are trying to estimate that constant here, now we will try to find out various ways of estimating the constant.

So, the simplest and the best way to estimate the constant is by looking at the simple average, which the first student actually mentioned as the forecast. So, a simple average is the simplest and the easiest way to give a forecast for a set of data, which shows level or which represents a constant. But, when we take the simple average we also assume a few things, number one is we are considering all the data points, and we are not leaving out any of the data points.

Second is we give equal weightage to all the data points, so under these two important assumptions that we consider all the data points, and then we give equal weightage to all the data points, we get simple average to represent the forecast of this. So, if we make a decision based on the simple average, we would say that the forecast for 2010 is 27. It also means that the forecast for 2011 right now is 27, because if we add a 27 and then take the average we will get the same 27.

So, the forecast made here for every period is the same and that will be 27, now the moment we assume that we are going to give equal weightage to all the data and consider all the data, then the question asks whether thus data which is 6 years old is equally important as this data which is 1 year old. So, we need to answer that question and that makes us move towards what is called a weighted average, so we can give some weights to each one of these, and then compute a weighted average.

For example, if we give weights w_1 , w_2 , w_3 , w_4 , w_5 and w_6 then the weighted average will be 25 into w_1 plus 32 into w_2 plus 24 into w_3 plus 28 into w_4 plus 26 into w_5 plus 27 into w_6 divided by w_1 plus w_2 plus w_3 plus w_4 plus w_5 plus w_6 . And as long as the w 's are greater than or equal to 0, we would get an answer which is within the range of 26 and 32, it is also a customary to define the w such that, sum of the w 's is 1. So, that the denominator is 1 and you do not divide it by the sum of the weights.

So, a weighted average is as good as a simple average, it has the additional advantage that we consider all the points as in simple average. But, we are able to capture the relationship that more recent pieces of data, add more value to the forecast and therefore, they are given higher weights or there is a possibility to give higher weights, to more

recent data. So, that the effect of the more recent data can be captured into the forecast, the obvious disadvantage of a weighted forecast is, how do we determine the weights.

One person may want to give weights 1, 2, 3, 4, 5, 6 another person may want to give weights 1, 1, 2, 2, 3, 3 and so on. So, for different weights we have different values of the weighted average, and it is not perhaps possible to find out a consistent set of weights or a value that comes out of a given set of weights, as acceptable to a group of people. But, if a group of people sit together and say that, this in our opinion are the weights w_1, w_2 to w_6 , then the weighted average can be a slightly better estimator than the simple average.

Then we move to the next idea that was shown here, where the person had taken the average of the last 4 values. So, there are 6 values, but then somebody had decided that the first 2 values being very old are not going to be as important as the last 4 and therefore, the person has taken the last 4 values and computed the average. So, the average of these 4 become 26.25, so when we leave out certain data points, and consider only the last k data points, then we develop what is called a k period moving average, it is called a k period moving average.

So, the 4 period moving average gives us a forecast of 26.25, a 3 period moving average taken here gives us a forecast of 27 which is the average of 28, 26 and 27. So, moving averages essentially are made or decisions based on moving averages, are essentially made by assuming that the last k periods, are more relevant and the others are left out.

So, a k period moving average can be thought of as giving weights of 0 a 4 period moving average can be thought of as giving weights 0, 0, 1, 1, 1, 1. So, every moving average can be thought of as a weighted average with weights given corresponding, a 3 period moving average can be thought of as giving weights 0, 0, 0, 1, 1, 1 and it becomes a weighted average.

While moving average is able to capture, the principle that more recent points are important, the older points do not contribute much, then there are two issues that are associated here, one is do we essentially ignore all of that, after all do not they add a small amount of values, which perhaps could be modelled or could be taken into account by giving slightly smaller weights.

But, then as we keep increasing the k for example, a 4 period moving average gave 26.25 a 6 period moving average for this data would mean that you are taking all the 6 data points and n period moving average if there are n data points will be the simple average etcetera. So, a moving average is related to simple average, it is also related to weighted average, so this is another form of average, which is the moving average a 4 period moving average is shown here, a 3 period moving average is shown here.

The next thing that we can look at is, so if we have a simple average and a weighted average, can we have something like a weighted moving average. So, a weighted moving average is shown here with 26.833, which was obtained by considering the last 3 points, but giving differential weights of 1, 2 and 3. The oldest has a lesser weight of 1, the middle one has a weight of 2, and this has a weight of 3, and the weighted 3 period weighted moving average gave us 26.83.

So, we have seen four different forms of average, the simple arithmetic average, the weighted average, the moving average, and the weighted moving average. All four of them are measures of central tendency; all four of them would give us values which is well within the range. Now, I also need to explain the reasoning behind these other three, and also we need to try and position the other three reasons, in the broad context of forecasting. Now, when we look at these values and when I introduce these values, I said these would represent the demand of a particular product.

And the only thing that we have right now here is that, how the demand behaves with respect to time. So, the behaviour of demand with respect to time has been captured in this data, and all data where the demand is given with respect to time, come under the category of time series models or time series forecasting. So, this is an example of time series data, where the behaviour of the demand with respect to time is provided.

Now, when we start looking at this F is equal to 30, now here the person, who said F is equal to 30, assumes that there is a relationship, there is some other factor which is influencing the forecast. And that factor could be the buying power of the people and so on, models where other factors influence the forecast are called causal models, where there is a cause and effect relationship between the forecasted variable, which is the independent, which is the dependent variable, and the causal variable which becomes the independent variable.

So, even though this is a reasonable value the thinking process behind this value comes from the assumption, that there are other factors. And that leads to other type of forecasting models. Right now in this lecture, we are restricting ourselves to time series forecasting. And much later we will see a couple of models for causal forecasting, so that is how one can explain this other factors, now there are two other answers which need an explanation.

Now, both of them are centred around the belief that this data is not stable and in fact this data shows a certain increase or decrease, which essentially means that in both these instances the person believes that there is a trend. And here also there is a trend, the rest of them based on averages do not believe that there is a trend. So, once again time series models have different classifications, now the first ones are called only level models or constant models, which for example, was the assumption when somebody made this forecast, and somebody made this forecast, and somebody made this forecast this forecast.

Whereas, when someone gave this value which is F is equal to 28 which was based on an increase and F is equal to 26 which was also based on increase and decrease. There is an assumption here that the data is exhibiting a trend, a noticeable increase or decrease which cannot be attributed to randomness or noise. So, there are trend models which we will see a little later, after we address this particular model called the models for data that exhibit constant.

So, the last point that requires requires certain explanation is this, now this person said if I look at these I find that this is an outlier. So, I am removing this and then I am trying to take the average, so the general assumption is that this person believes that this data represents a constant or level data, person assumes that is there is no trend or anything. Except that the person believes that this number seems to be slightly high, and perhaps may not be adequately justified by the assumption that it is a noise.

So, in such cases it is also a customary to remove certain pieces of data, after the analyst is fully convinced that this may not actually contribute or it might take away. So, in such cases you remove a piece of data, and then you find the average, so these represent different kinds of thinking, but largely centred around average, which represents that the

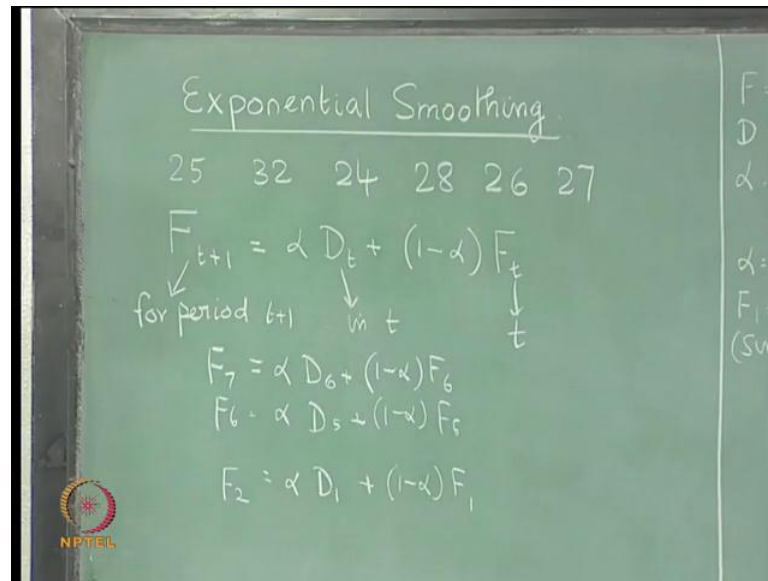
assumption behind here is that it represents a constant or a level model, and whatever variation we have is attributed to noise or statistical fluctuation.

And we are trying to forecast F is equal to a plus epsilon where epsilon represents the noise or the variability, and it which has the mean of 0 and it has a variance of sigma square. Other things such as trend models, and other causal models we will see as we move along, now let us look at one more thing, let us look at this pieces of data once again. Now, if we look at the four models that we have seen, the arithmetic average, the weighted average, the moving average, and the weighted moving average all of them have advantages and perhaps a little bit of limitation.

The arithmetic average gives equal weights to all the points, which we may kind of question because it is only customary to believe that more recent data should contain more weight, and are more important. The weighted average has the disadvantage of not being able to arrive at some kind of a consensus weight or an acceptable weight, moving averages are very nice, but moving average tends to ignore or give 0 weight to earlier points and so on.

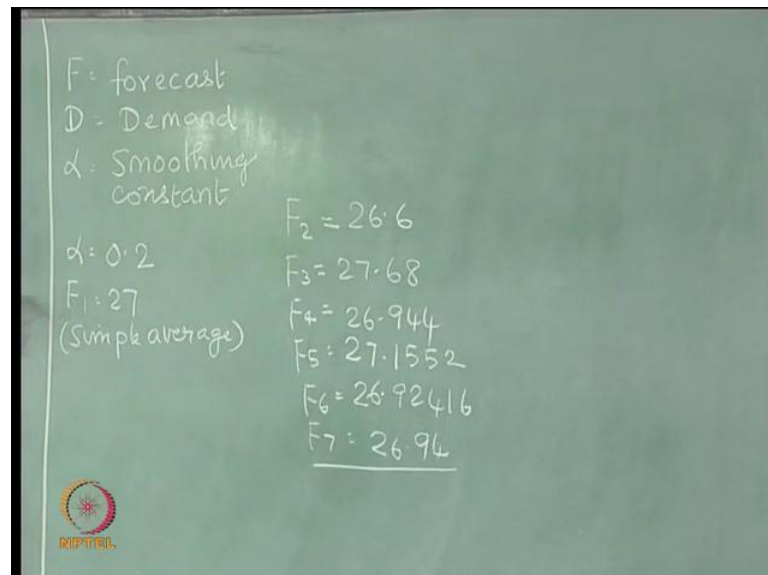
Any weighted model has the limitation of getting the right kind of weights, so it now boils down to another question is there a model where, we actually consider all these points. We give progressively increasing weights to more recent data, and the weights seem to come logically and are acceptable to everybody, now such a model is called the exponential smoothing model.

(Refer Slide Time: 28:31)



So, let me write the exponential smoothing model here, it is called the exponential smoothing model. Now, let us look at the 6 pieces of data 25, 32, 24, 28, 26, and 27, now the exponential smoothing equation comes from. So, F_{t+1} is equal to αD_t plus $(1-\alpha) F_t$.

(Refer Slide Time: 29:39)



I am going to follow this notation for sometime, where F represents the forecast and D represents the demand, alpha is called as smoothing constant. Now, F_{t+1} is the forecast for period t plus 1, so F_{t+1} is forecast for period t plus 1, D_t is the demand

in period t and F is the forecast for period t . Now, in this notation this is D_1, D_2, D_3, D_4, D_5 and D_6 , D_6 is the most recent demand D_1 is the oldest demand, since we know the demand for 6 periods up to D_6 , we are interested in forecasting the demand for the seventh period. So, we are interested in finding out F_7 .

Once again let me repeat that I am using this notation F_{t+1} is equal to $\alpha D_{t+1} + 1 - \alpha F_t$ I am defining F_{t+1} as forecast for period $t+1$. There are other notations possible right now it is a bit too early to describe other notations, but as we move along at a suitable time, let me compare this with other notations and explain the significance of this notation, but let us proceed with this notation. So, we are interested in F_7 which is the forecast for period 7, because we know the demand up to period 6.

So, based on this F_7 is equal to $\alpha D_6 + 1 - \alpha F_6$, now we can write F_6 as $\alpha D_5 + 1 - \alpha F_5$. So, F_6 will now be written as $\alpha D_5 + 1 - \alpha F_5$ and proceeding further we can write F_2 is equal to $\alpha D_1 + 1 - \alpha F_1$. Now, we know D_1 right now we do not know α and we do not know F_1 , now if we assume let us assume α equal to 0.2 α as called a smoothing constant, and α is a number between 0 and 1.

So, let us assume α equal to 0.2 we also do not know F_1 , so if we know F_1 or if we assume a value for F_1 , then we can use this to compute F_2 . And then we can substitute F_2 somewhere here and find F_3 , similarly use F_3 to find F_4 , F_4 to find F_5 and then once F_5 is known you can find F_6 , and then once F_6 is known we can find F_7 . So, let us simply do these calculations first let the value of F_7 , and then try to explain a few things.

So, let us assume that F_1 is 27 which is the simple average, so if we use F_1 is 27, D_1 is 25, α is 0.2 and calculate we will get F_2 . So, F_2 becomes 26.6 and F_3 is $\alpha D_2 + 1 - \alpha F_2$, which is $0.2 \times 32 + 0.8 \times 26.6$ and this on calculation will give us F_3 equal to 27.68, now F_4 is equal to $\alpha D_3 + 1 - \alpha F_3$, F_4 is equal to $\alpha D_3 + 1 - \alpha F_3$.

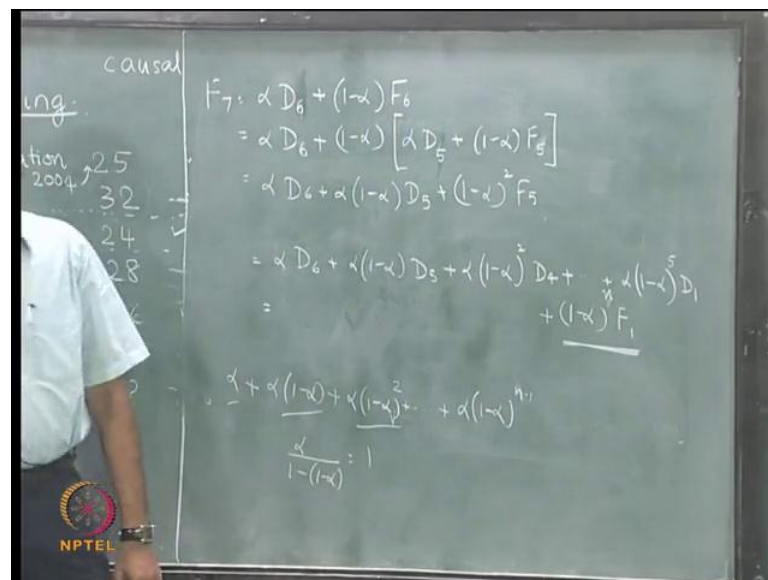
So, F_4 is $0.2 \times 24 + 0.8 \times 27.68$ and F_4 will become 26.944, F_5 is $\alpha D_4 + 1 - \alpha F_4$. So, it is $0.2 \times 28 + 0.8 \times 26.944$ and F_5 will become 27.1552 F_6 is $0.2 \times D_6 + 1 - \alpha F_5$, so it is $0.2 \times 26 + 0.8 \times 27.1552$ which would give 26.92416, and once F_6 is

calculated F_7 is $\alpha D_6 + 1 - \alpha F_6$, 0.2 into 27 plus 0.8 into 26.92416 , so F_7 will become 26.94 .

So, F_7 is the forecast for period 7 based on this exponential smoothing model, we have calculated the forecast for period 7 as 26.94 , this 26.94 is calculated based on two assumed values, one is α equal to 0.2 and the other is F_1 is equal to 27 . Now, this 26.94 if we start comparing with these, it is very close to the simple average of 27 and it is slightly far away from a weighted average of 26.833 , if we consider a weighted moving average on the last three values.

Now, what are the advantages of exponential smoothing and what are the areas we need to address in exponential smoothing? So, first let us address perhaps the advantages of exponential smoothing, now when we set out to do exponential smoothing we said that we want a model where we use all the values. Now, we have used all the values we have used D_1, D_2, D_3, D_4, D_5 , and D_6 we have used all of them, and we also said that we should give progressively increasing weights to more recent data. Now, let us see how we give the progressively increasing weight to more recent data, now in order to do that.

(Refer Slide Time: 37:55)



So, let us write F_7 is equal to $\alpha D_6 + 1 - \alpha F_6$. So, $\alpha D_6 + 1 - \alpha$ into $\alpha D_5 + 1 - \alpha F_5$, this on simplification will give $\alpha D_6 + \alpha$ into $1 - \alpha$ D_5 plus $1 - \alpha$ square F_5 .

Now, F_5 will be written as $\alpha D_4 + (1 - \alpha) F_4$, and if we continue to do this we will get $\alpha D_6 + (1 - \alpha) D_5 + (1 - \alpha)^2 D_4 + \dots + (1 - \alpha)^5 D_1 + (1 - \alpha)^6 F_1$.

So, this is how this will expand to this if we keep inductively substituting the values for this F . Now, if we look at this expression what we have calculated as 26.94 here, can also be calculated in a single pass by using $D_1, D_2, D_3, D_4, D_5, D_6$ from these values or from three values, using α equal to 0.2 and using F_1 is equal to 27. Now, if we start looking at these terms, each term uses the demand D and is multiplied by a number, here it is α it is $(1 - \alpha)$ etcetera.

So, this can be treated as a weight associated with this demand, now I already mentioned that when we took α equal to 0.2, I already mentioned that α is called the smoothing constant, and α is a value between 0 and 1. So, $(1 - \alpha)$ is smaller than 1, so α here $(1 - \alpha)$ is smaller than this weight, this is smaller than this weight and so on. So, we have progressively decreasing weights from the more recent data value which is 27 to the past.

So, 27 has a higher weight which is α which itself is less than 1, 26 has a weight smaller than α , 28 has a weight smaller than this and so on. So, if we leave out the last term the exponential smoothing is a weighted average, it is a weighted average of all the demand points, with progressively decreasing weights as we move past or as we move backwards in time. But, we also know that if we have a weighted average then we have to divide by the sum of the weights.

So, let us find out the sum of the weights the sum of the weights are $\alpha + (1 - \alpha) + (1 - \alpha)^2 + (1 - \alpha)^3 + (1 - \alpha)^4 + (1 - \alpha)^5$. Now, if you look at these sums this is one term, this is one term, this is one term and so on, now we know that each term is multiplied by a $(1 - \alpha)$. So, it is a finite geometric series, some of terms have a finite geometric series is a into r to the power $n - 1$ by $r - 1$, if r is greater than 1 and a into $(1 - r)$ to the power n by $(1 - r)$ if r is less than 1.

Now, if we make an assumption that n is large here we have n equal to 6 we have 6 demand points. So, if we make an assumption that n is large, then this will be α ,

alpha into $1 - \alpha$, α into $1 - \alpha^2$ and so on, this will become n or $n - 1$ as a case may be. But, if n tends to infinity we get an infinite geometric progression, with first term equal to α and common ratio equal to $1 - \alpha$.

So, sum of an infinite geometric progression is $\frac{\alpha}{1 - \alpha}$, so that will be α by $1 - \alpha$ which is 1 . So, as the number of terms increases this will converge to 1 , and even within a very a reasonably small number of terms like 6 terms or 7 terms or 8 terms, whatever be the value of α the sum will be very close to 1 . And therefore, this is like a weighted average, where the denominator which is a sum of the weights is 1 and therefore, we do not divide.

But, then we also have a term like this $(1 - \alpha)^n$, and if n is large this term is $(1 - \alpha)^n$. Now, as n is very large $(1 - \alpha)^n$ tends to 0 , if α is smaller than 1 and n is large $(1 - \alpha)^n$ tends to 0 and this can be ignored. So, the exponential smoothing is like a weighted average, which is what we wanted, we wanted to use all the terms we are using all the demand terms.

We want progressively decreasing weights as we move to the past or progressively increasing weights for more recent data, which is what is happening here. Some of the weights adds up to 1 , the last term tends to 0 , so it is like a weighted average that we have, and what is also interesting that given a value of α we get progressively decreasing weights, which in some sense can be accepted by everybody provided everybody accepts the value of α . So, over a given value of α we seem to be getting a very nice and consistent set of weights which also add up to one.

So, exponential smoothing meets all other advantages or all the things that we expect from it, which is use all data points give progressively decreasing weights to older data, and have a consistently acceptable set of weights. So, which is what exponential smoothing does, plus a fact that this weight is very small $(1 - \alpha)^n$ is very small, even for a reasonable value of n . So, the F_1 equal to 27 actually does not impact the final answer significantly.

Because, the weight associated with this F_1 is practically 0 therefore, the initial value of forecast can be taken as 27 which is this simple average, sometimes people would take the last of the values as F_1 which also happens to be 27 in this case. But then even if we

take 26.5 and 26.35 or whatever as long as it represents adequately another measure of central tendency or average the effect of this F 1 is not very high. So, in a way we can say that F 1 equal to simple average can be accepted, but we have to spend a little more time on this alpha. Now, first of all alpha should satisfy two conditions, alpha is greater than 0 and alpha is less than 1. So, 0 less than or equal to alpha less than 1 and why are we looking at this?

(Refer Slide Time: 47:37)

Exponential Smoothing

25 32 24 28 26 27

$$F_{t+1} = \alpha D_t + (1-\alpha) F_t$$

\swarrow for period $t+1$ \downarrow in t \downarrow t

$$F_7 = \alpha D_6 + (1-\alpha) F_6$$

$$F_6 = \alpha D_5 + (1-\alpha) F_5$$

$$F_2 = \alpha D_1 + (1-\alpha) F_1$$

F - forecast
D - Demand
 α - Smoothing constant
 $\alpha = 0.2$
 $F_1 = 27$
(Simple average)

NPTEL

Now, if we write or if we come to the most basic equations which is given here from this equation we understand that the forecast for period t plus 1 is a sum of two things, which is a weight associated to the current demand and a weight associated to the current forecast. And because it is a sum of two terms a weighted sum of two terms, we do not want a negative to come here we want both of them to contribute, and we do not want a negative term to come here.

Therefore, we will not have alpha greater than 1, the moment alpha becomes greater than 1 then we realise that this term starts contributing in a negative manner. Similarly, the moment alpha is less than 0, then this term gives a negative contribution, since we do not want the negative contribution from either of the terms, we have alpha between 0 and 1, now, having kind of agreed on the fact that alpha has to be between 0 and 1.

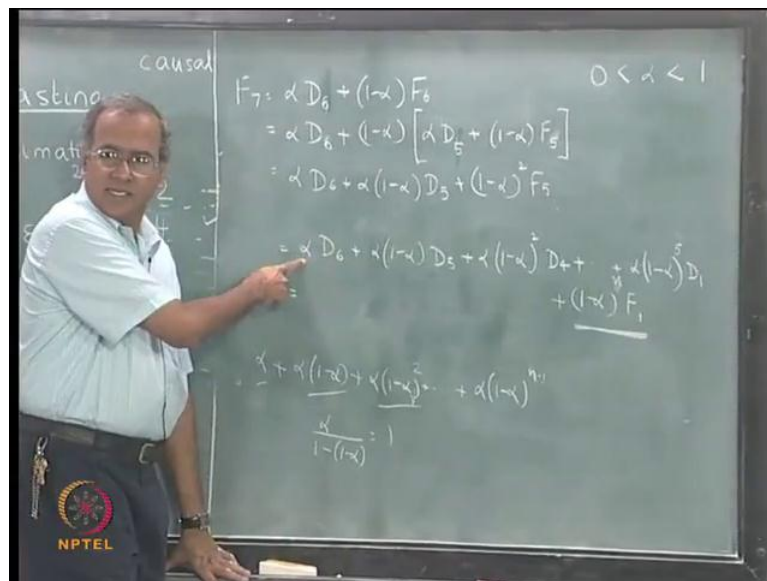
Now, what should be the value of alpha should alpha be small or should alpha be large, now in this example we have taken alpha is equal to 0.2. So, we have taken alpha as a

smaller value, and not as a larger value, now if we take alpha as a smaller value and start substituting here, we will immediately realize that the weight or the contribution of the demand is actually if D_t and F_t are comparable then the contribution of the demand seems to be less than the contribution of the forecast.

If alpha is greater than 0.5, then the contribution of the demand will be more than the contribution of the forecast. There it actually brings us to another question which says, should not the contribution of demand be more than that of the forecast, which means should alpha be closer to 1 or should alpha be closer to 0. If alpha is small and 0.2 then; obviously, the contribution of forecast is higher, but if we look at it very carefully we would want the contribution of forecast to be higher.

The whole forecasting comes because this demand is not extremely stable and it exhibits a certain amount of noise, just for the sake of argument if all these numbers were 27, then we do not forecasting and we would say that the next value is 27. So, the demand is very stable and does not show noise or variation within, in such cases we can have a larger alpha. So, long as there is variation and noise it is advisable to have a smaller value of alpha, which means here, we realize that the weight associated with the demand in this term can be less and the forecast will have a higher weight.

(Refer Slide Time: 50:51)



There is another way of explaining it if alpha is large here, then what will happen is this F_7 will be guided by fewer terms which will contribute to F_7 . For example, if alpha is

0.8 now this is 0.8 this is 0.16, so two terms are contributing to a weight of 0.96, so they have almost contributed to F_7 . Whereas, if alpha is 0.2 then this is 0.2 and this is 0.16, so two terms are contributing to a weight of 0.36.

So, smaller the alpha more terms will start contributing to F_7 , smaller the alpha in some sense the weights even though there are progressively decreasing weights, this will contribute a little more to it smaller the alpha, larger the alpha these the weight associated with this become very, very small. So, if we want all of them to contribute reasonably well alpha has to be slightly smaller, so this is how the exponential smoothing model works and is dependent on the alpha and F_1 .

And exponential smoothing provides us with a very good model to forecast the level, where your able to use all the data, we are able to give progressively decreasing weights to the data as we move past or progressively increasing weights to more recent data. So, with this we have seen a little bit of forecasting for level or constant data, in time series how to do forecasting when the data exhibits trend we will see in the next lecture.