Data Analysis and Decision Making – I
Prof. Raghu Nandan Sengupta
Department of Industrial & Management Engineering
Indian Institute of Technology, Kanpur

Lecture – 25

Very good morning, good afternoon, good evening to all of you my dear students, welcome to this Data Analysis and Decision Making I course under the NPTEL MOOC series. And we are in the 25th lecture, and this course as you know is for 12 weeks total number of hours is 30 hours, total number of lectures would be 60 because each week we have 5 lectures and 5 each lecture being of half an hour. So, 25th means we are going to end the 5th week.

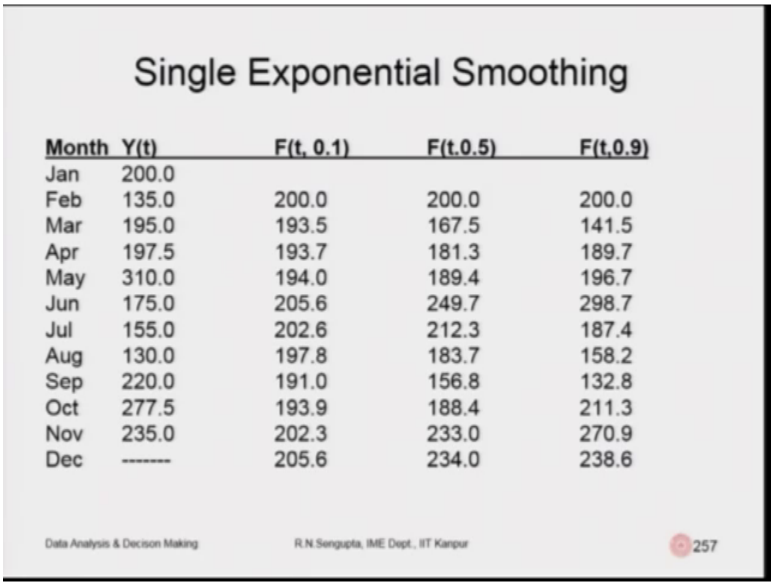
Now after the multiple linear regression conceptually which I had discussed, I basically went into different type of forecasting techniques. I will discuss the forecasting techniques (Refer Time: 01:01) see a few problems, then again come back to the concept of regression with models and solutions. And, then go into the analysis of variance and multivariate distributions and all these things.

So, if you remember we were discussing a problem where the numbers of passengers was given, starting for few years total number of passengers in the airline sector and you want to basically predict forecast. And we will consider that given the datas which are basically actual values, which are y 1 y 2 y 3 y 4 till y n you want to predict technically for y hat, which will be y hat for n plus 1 which will be given by forecasted value or predicted value of f n plus 1 and you will basically continue accordingly.

Now according to the formula for the exponential smoothening we had taken that we will put some weights on the actual value for the last time plus the predicted values of the last time. So, considering that I want to predict for February 2018 I will put some proportional weights for January actual value and January predicted value.

Now; obviously, the weights which had discussed should be equal to 1 and also the main idea was that the weight should be such that if I want to. So, the weights are basically if their some is 1, one would be alpha another would be 1 minus alpha so; obviously, we want I need to find out alpha. So, what we will do? We will basically minimize the sum of the squares of the errors, errors are basically the difference between the actual value and the predicted value square that sum it up, and differentiate with partially differentiate. I am using the word partial differentiation because if there are more than one parameters here it is only one which is alpha. So, you will differentiate that with respect to alpha put it to 0 phi naught alpha hat and then proceed accordingly just like as we have done for multiple linear regression.

(Refer Slide Time: 03:25)



Single Exponential Smoothing

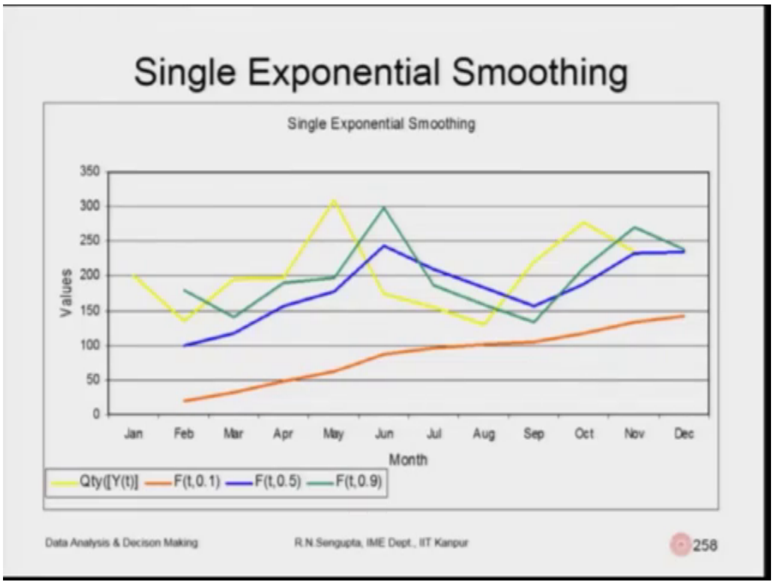
Month	Y(t)	F(t, 0.1)	F(t,0.5)	F(t,0.9)
Jan	200.0			
Feb	135.0	200.0	200.0	200.0
Mar	195.0	193.5	167.5	141.5
Apr	197.5	193.7	181.3	189.7
May	310.0	194.0	189.4	196.7
Jun	175.0	205.6	249.7	298.7
Jul	155.0	202.6	212.3	187.4
Aug	130.0	197.8	183.7	158.2
Sep	220.0	191.0	156.8	132.8
Oct	277.5	193.9	188.4	211.3
Nov	235.0	202.3	233.0	270.9
Dec	-------	205.6	234.0	238.6

Data Analysis & Decision Making R.N.Sengupta, IME Dept., IIT Kanpur 257

So, the month wise data points are given I am only taking a snapshot. So, January February March April are the months given in the first column, in the second column we have the Y t, Y means the actual values they start from 200 and go to November it is 235. And, what I have done in the corresponding third fourth fifth column is, I basically written or calculated the forecasted value for the month of February to December, for the month of February to December, for the month of February to December in the third fourth column, corresponding to the fact that the alpha values are changing for the third column which is 0.1, which is; that means, you want to give up weightages of 10 percent. Similarly for of the second last column the weight is 0.5; that means, you are putting equal weightages 50 percent and the third column and the last column basically has a alpha value as 0.9, which means you are trying to put 90 percent weightages.

Now these weightages is that that how much confidence do you have on the values of the past data whether it is forecasted or predicted based on the we are going to basically consider.

(Refer Slide Time: 04:43)



Considering the values of alpha as 10 percent 50 percent 90 percent, which are basically given by the red line, blue line, green line and the yellow one is basically the actual quantity which you want to predict. So, if you find out the overall value for 10 percent it only gives the trend. So, the prediction is bad; that means, how you want to forecast if you find out 50 percent and then basically look into 90 percent; obviously, 90 percent value would be the best predictor as it happening, because you are putting more weights on the last value, but they would be a lag. Now the lag is coming because you are going one step backward; that means, standing at t you are trying to basically utilize the data from t minus 1 then try to predict.

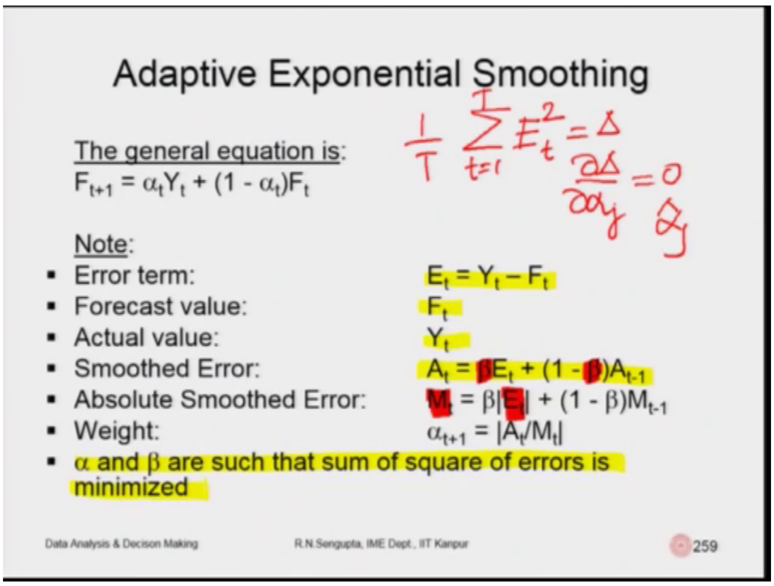
Now, this explanation smoothly can be expanded further on like rather than only going 1 time backwards, you can go 2 time backwards in the sense that it can be standing and in February I want to basically forecast or predict. So, I want to find out F suffix February using the predicted value and the actual value for January; that means, 1 month back and also predicted and forecasted value for December; that means, 2 months back.

So, in this case; obviously, the weights would be alpha 1 for the predicted value for January alpha 2 for the actual value of January alpha 3 for the predicted value of December and 1 minus sum of alpha 1 plus alpha 2 plus alpha 3 basically would be the other case for December. Now obviously, why I use the word 1 minus in the bracket alpha 1 plus alpha 2 plus alpha 3 was due to the fact that the sum of the weights is equal to 1; 0.1.

Point number 2 obviously, you will try to say that yes there are 3 variables alpha 1 alpha 2 alpha 3, again you will utilize the same concept find out the difference between the predicted value and the actual value square that that value would basically be the errors squared the error, find the sum differentiate partially with respect to alpha 1 put it to 0. Differentiate the sum of the squares with respect to alpha 2 put it to 0 differentiate the sum of the squares with respect to alpha 3 put it to 0 solve them and find out alpha 1 hat alpha 2 hat alpha 3 hat and you can basically proceed accordingly.

Now, the question which you may be asking yourself intuitively is that, well I am taking alpha 1 alpha 2 alpha 3 or whatever alpha values are there, and I am able to predict and I am able to basically find out the sum of the squares try to minimize it with respect to the parameters and continue with the job. But the question is that is it necessary that alpha 1 alpha 2 alpha 3 till whatever values of alpha we take, are they fixed with respect to time? That means, is it necessary that the weights also which we give to the predicted and the actual values also get updated as we proceed further on with respect to time? The practical answer is yes it should be considered and we will consider that in the halt linear and halt winters methods later on.

(Refer Slide Time: 08:03)



So, consider the adaptive exponential smoothing process, which I am going to now discuss. Now in the adaptive smoothing method the general equation is all again the same, but here I will consider that there is an trend which I am going to consider. So, the general equation is that I want to forecast for t plus 1; that means, I want to focus for February utilizing the information from January.

So, F t plus 1 suffix is equal to alpha t, which is now see here I am putting a suffix t because this values of alpha 1 or alpha 2 or alpha 3 whatever they are there and whichever values we take how many such alphas we take, they are dependent on time that as we proceed the values of alpha would basically either increase or decrease. So, I take alpha t Y t minus into Y t plus 1 minus alpha t into F t. So, Y t would basically be the actual value for time period t, F t would be the forecast or the predicted value for time period t.

So; that means, I am putting a weightages if alpha is say for example, 20 percent I am going to put 20 percent weightages for a Y t and 1minus 20 percent which basically would be 80 percent would be the weightages on the forecasted value for time period t. Now actually if you remember I did mention errors square of the errors and all this things. So, the errors are this is the error. So, technically I want to basically find out the sum of the error t is equal to 1 how many such errors are there, square that up divided by t.

So, that considered as delta differentiate delta with respect to alpha, and consider the j is the suffix for different alphas. So, basically you put to 0 and then find out the alpha j hats now further on they are being updated. So, let us first go through a nomenclature; the nomenclatures are as mentioned F t is the forecasted value with the suffix t, Y t is basically the actual value with suffix t. Now, there would be two different terms which will be coming one to smoothen out the errors and find out the absolute smooth error.

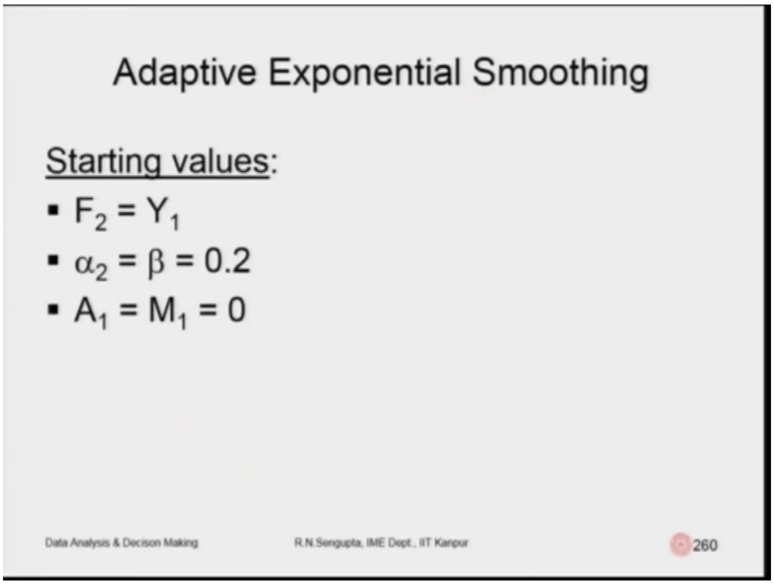
So, smoothing means that if there is a trend that the values are increasing. So, basically increase it slowly and bring it to the average value; that means, it is basically acting like a so, called not damping effect, but basically trying to basically bring down that fluctuation to an average value. So; that means, you know the trend as well as seasonality if they are there.

So, A t is the smooth error. So, this is given and here for the first time we find out there is another constant which is beta, which is basically the smoothing constant and smoothing variable which you will also try to find out. Based on the fact that are main idea is to find out the sum of the squares and differentiate with respect to the parameters, because here also beta is a parameter and put it to 0 and find it out. And the absolute smoothing error would basically be the values of the smooth error which you have, that would basically be added or subtracted depending on how you are trying to proceed to find out the error.

So, this value of E t we know it is an error. So, now this smoothing constant beta would either increase or decrease depending on what the value of beta is. If you note down this value of beta can also change and in the later problems, now we will consider them to be fixed. And, this smooth error M t minus 1 for the last time would basically be weighted by the factor of 1 minus beta and that would basically give a general trend how the overall error would basically behave based on which we will proceed. And obviously, the fact that the smooth error which is A t, and the value of the absolute smooth error which is M t they if they are 1. So; obviously, it would mean the value of alpha is 1 if they are increasing or decreasing; obviously, they would have an effect and how the alpha value is created.

So, again I will basically highlight the value and alpha beta are such that the sum of the squares of the errors is minimized and we will basically come to the problems accordingly.

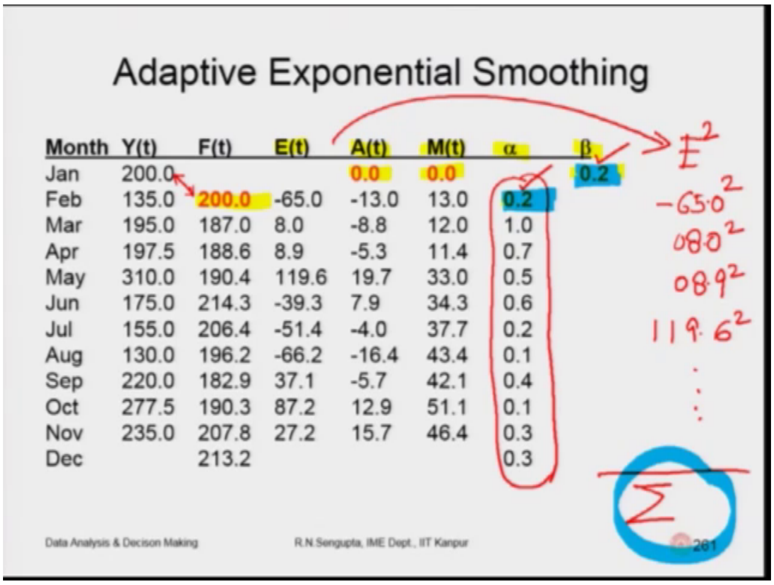
(Refer Slide Time: 13:01)



Now consider these values of we are taking arbitrary values and we will proceed and then discuss that how they can be done intuitively. So, consider A 1 and M 1 which is the absolute and the smoothening errors are 0 initially when we start at time period 1 it can be for time period 0 also. So, we proceed for time period 1 2 3 4 and so on and so forth. So, we considered A 1and M 1 as 0, and we are for our problem considering the weights of initial alpha as 20 percent beta as 20 percent.

So, beta would be fixed as 20 percent, but alpha may change depending on the values of the ratios of A to M; A t divided by M t whatever we had in the last slide. And we are considering that if it is the beginning then the forecasted value for the next time period would be the actual value for this time period; that means, f suffix t would be equal to y suffix t minus 1. So, if you are trying to basically start at the actual is considered a and m for time period 1 as 0 we will consider that the corresponding to f for the next time period which is 2, we will consider that equal to y 1 and then proceed accordingly.

(Refer Slide Time: 14:31)



So, now, I have again written down the values on the as described just few minutes back in one on the example. We consider the months along the first column, the actual values of Y t along the second column, and then what I do is basically I consider the difference between I first I and have a third column where I need to find out the predicted of the forecasted. And the corresponding fourth fifth sixth column, and then the corresponding one would basically be the error, which is here. Then the absolute smoothing error, the relative one the alpha value, but alpha value is basically alpha suffix t because this is changing and beta.

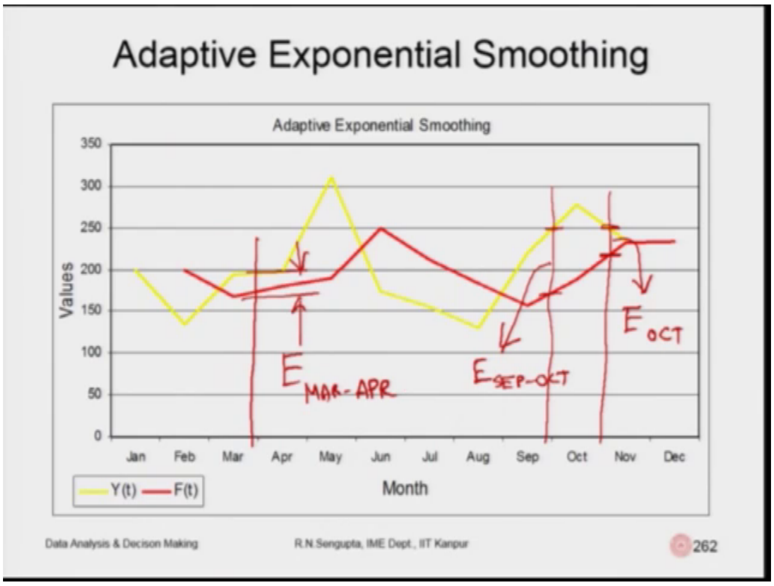
So, now you see I am just mentioning for the time being (Refer Time: 15:25) alpha values are changing. Now consider I am taking A 1 and M 1 as 0 0 as mentioned. So, let me highlight it I consider alpha value of 0.2, beta value as 0.2, this is arbitrarily I am taking and I am putting f 2 as equal to Y 1. So, this is the case. So, which means this value goes here.

Now, I proceed in the same formula. So, if you consider the same formulas utilize them accordingly, you find out the errors are here as given based on that you find on the forecasted value the A value and the M value. What is interesting is that now you see the alpha values are changing within the initial example for the exponential smoothing method where they were not adaptive, the values of alpha was fixed, but here the alpha values are changing. And what you will do is that considering a starting value of 20 percent 20 percent for alpha and beta you will basically find out the sum of the squares of the errors. So, this actual column which is not there I am just putting it here.

So, basically you will find out E square corresponding to the fact that you are taking alpha and beta as 20, and you will basically find out this value minus 65.0 whole square, then 08.0 whole square, then 08.9 whole square, then 119.6 whole square continue adding find out the sum. So, you note it down.

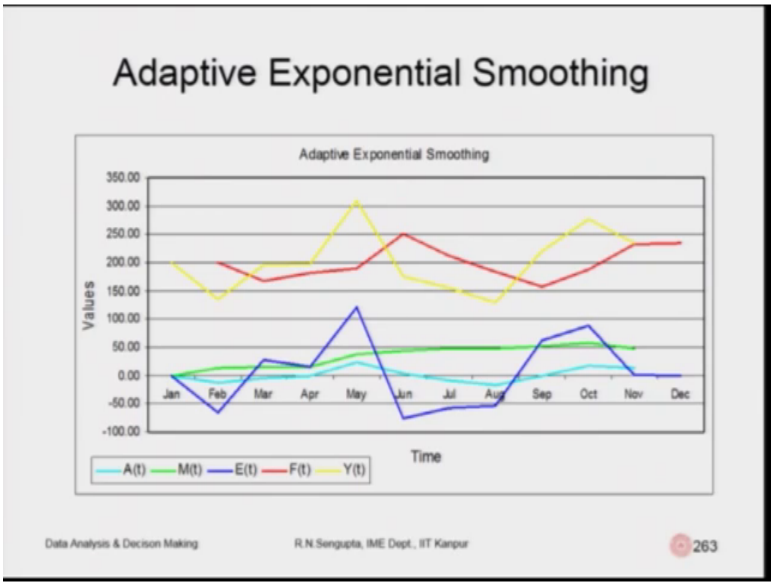
Now, what you do in a simple excel sheet now keep changing these values, keep changing this 0.2 here and some value here to take arbitrary values. And then check for which values this error is coming out to be the least and obviously, that would be the value of alpha and beta corresponding to the problem, that your main emphasis is to minimize the error which is basically sum of the squares of errors.

(Refer Slide Time: 17:45)



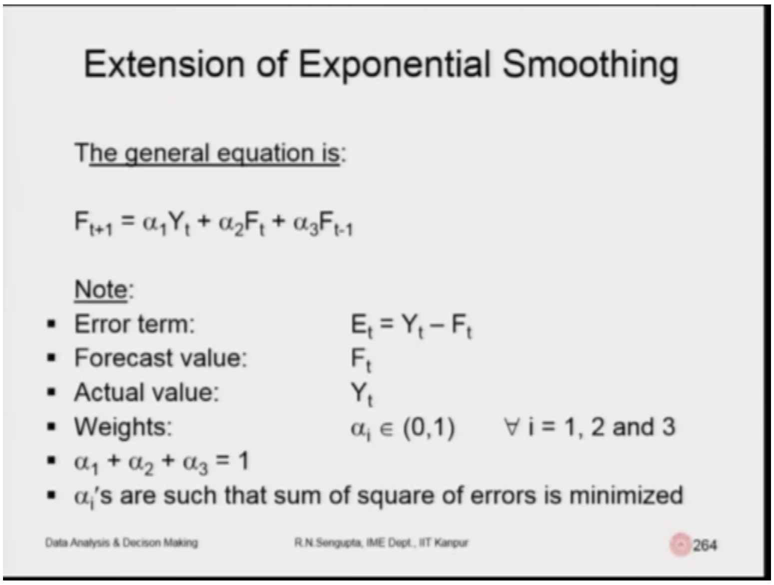
So, considering some values of 20 percent 20 percent for alpha and beta, I have Y t which is the yellow line, the red one is basically the forecasted line and if you find out if you plot try to plot it, the difference which you have which you would have let me consider a color say for example, dark red. So, for September for October whatever you do, you will find the differences here. So, these are the errors. So, this technically would be the error for between September to October, this would be the error for October. Similarly if I take month of April so, this would be the error of March to April, you are basically squaring them and basically trying to find out the sum of this, it does not minimize.

(Refer Slide Time: 19:01)



For the adaptive exponential smoothing, I now basically write the values of A and M corresponding the t values and then also I draw the errors which is the blue line. So; obviously, in the long run the expected value for the error or the average value of the error should be 0 and; obviously, we will consider that if the values of the variances of the errors are not changing, we can find out the variances of the errors also correspondingly.

(Refer Slide Time: 19:41)



Now, we will go into the extension of the exponential smoothing method. So, the general equation is we will try to find out the forecasted value for t plus 1 time period, based on the fact that you want to basically have put weights. And these weights are alpha 1 alpha 2 alpha 3 alpha 4 whatever you consider, but; obviously, remember the fact the two main notions continue, sum of the weights is 1 and the square of the errors should definitely be minimized in order to adhere to the fact that differentiating the sum of the squares of the errors with respect to all parameters individually partially differentiating; putting it to 0 finding out those hat values for the parameters would be the best way in trying to basically find out the alpha 1 alpha 2 alpha 3 hats whatever it is.

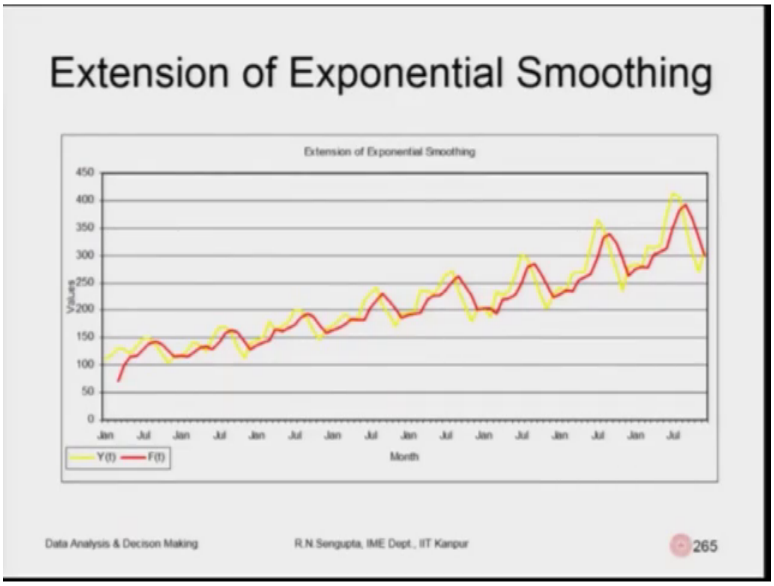
Now, here we are putting consider alpha 1 is the weight. So, we have putting alpha 1 weights on the actual value for time period t, and we are trying to also put alpha 2 and alpha 3 weights on the forecasted value for the t-th time period and the t-th minus 1 time period; that means, we are quite certain that the forecasted value have a huge amount of significance for predicting or forecasting for the further on time.

So, again in this same case this is not the adaptive one, this is a simple exponential smoothing, but we are going into a little bit towards the past in more details. In the sense if in the initial case we take only one time period, now in this case we are taking 2 time periods and it can be increased to 3 time periods 4 time periods and all this case can be considered, but; obviously, the fact remains sum of those parameters are 1 and;

obviously, sum of the squares of the errors then differentiate it with respect to the parameters if you put it to 0, then is the best way how you can find out the parameters themselves.

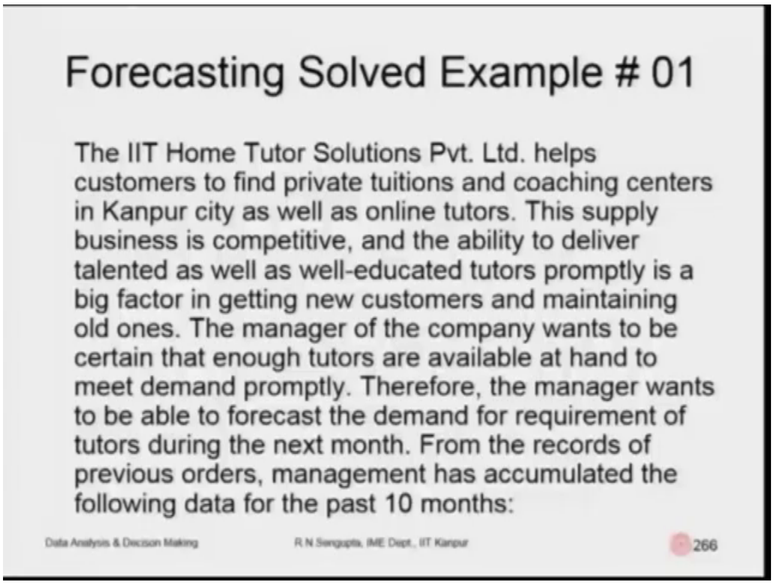
So, the error term is basically the difference between Y t and F t, and the forecasted value as usual is F suffix t for whichever time period. The actual value is given by Y suffix t which is that time period; and here alpha 1, alpha 2, alpha 3 are the errors and the sum is 1 as I have already mentioned, and here alphas are such that the sum of the squares of the areas is minimized. So; obviously, you can continue this problem in the context of adaptive explanation smoothing, also where you have now alpha 1 suffix t alpha 2 suffix t alpha 3 suffix t and beta would also be there.

(Refer Slide Time: 22:25)



So, for the extension this exponential smoothening which I do for the same data set. The data is the same with different values of alpha 1 alpha 2 alpha 3 I am basically brought it. So, the yellow one is the actual value, red one is the predicted value and I take different combinations of alpha 1 alpha 2 alpha 3 keeping in it mind that the sum is 1, then I can find a good fit such that the overall way how the actual values fluctuating is very well mimicked by the forecasted values. So, let us consider 2 or 3 simple problems.

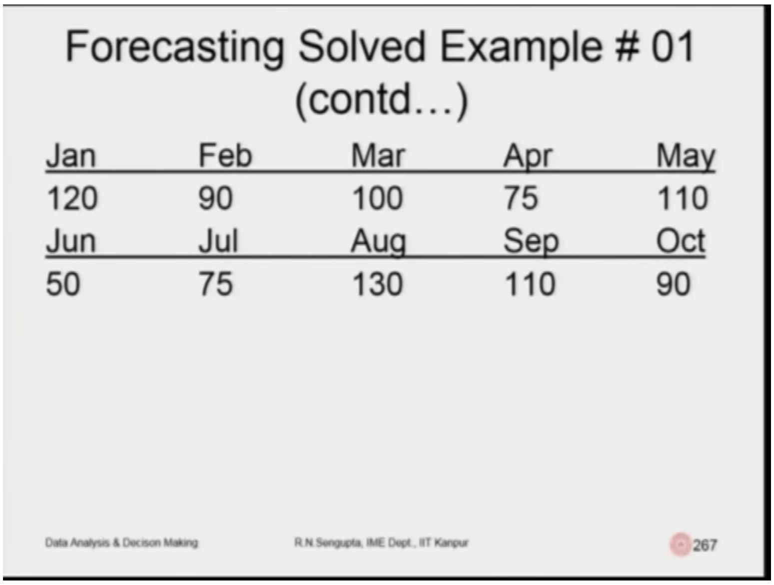
(Refer Slide Time: 23:01)



So, the problem is like this. The IIT Home Tutor Solution Private Limited, helps customers to find private tuition and coaching centers in Kanpur city as well as and for online tutors. This supply business is competitive and the ability to deliver talented as well as well educated tutors promptly is a big factor in getting good new customers and maintaining the old ones. The manager of the company wants to or the owner of the company wants to be certain that enough tutors are available at hand to meet the demand.

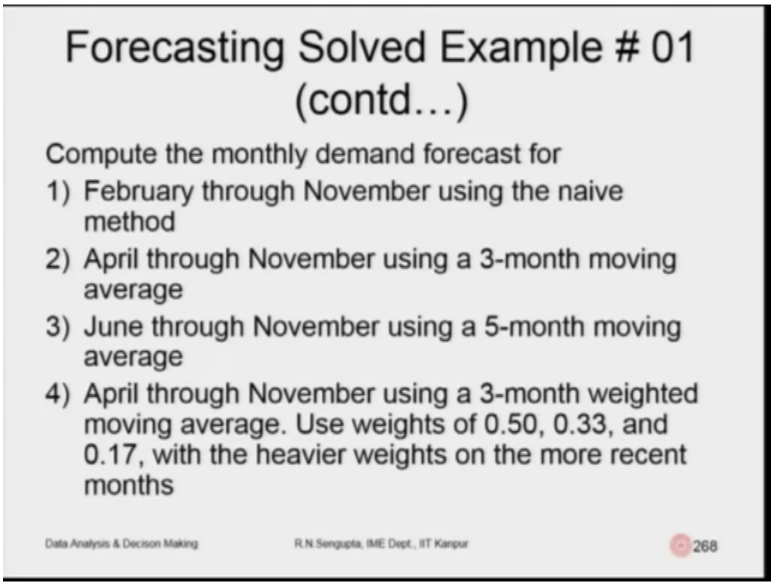
So, tutors maybe for English, for physics for Maths, for Chemistry whatever it is. Therefore, the manager wants to or the owner wants to be able to forecast the demand for requirement of tutors during the next month or in the coming future. From the records of the previous orders the management or the manager has accumulated the following data for the coming past 10 months. So, I am trying to consider the data as given, it is like this

(Refer Slide Time: 23:59)



It will be from starting from January to October the numbers are basically 120 90 100 75 110 50 75 130 110 and 90. So, based on I want to basically solve (Refer Time: 24:18) problems and what are the questions I am going to state it now.

(Refer Slide Time: 24:21)

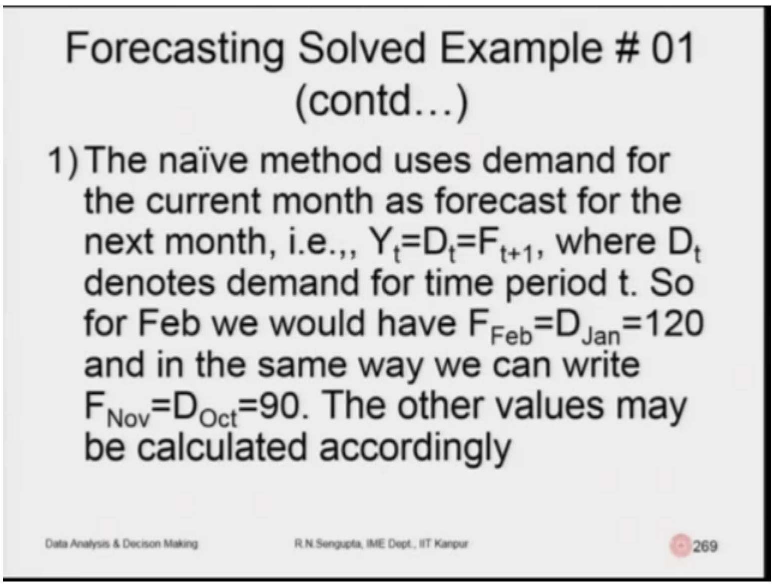


So, questions are as follows compute the monthly demand forecast based on four simple methods; number 1 February through November using the very simple naive method; that means, what is actually to be the case you will basically use that simple without any much calculation very simple conceptual ideas based on which you will proceed. This I have not discussed, but it is very simple to understand you are basically going very intuitively. The second part of the problem is using April through November, you want to basically use a 3 month moving average; the third one is June through November using the 5 month moving average.

Now, see April and June are the cases because in the first case is 3 months second case is 5 months so; obviously, in the amount of information we utilize would be much more in the third problem hence you can only start predicting from June. The fourth part is April through November using a 3 month weighted moving average use weights of as given as 50 percent 33 percent and 17 percent with the heavier weights on the more recent months and lower weights later on. And, if you notice here very interestingly and rightly so, the sum of the weights which is 50 percent 33 percent and 17 percent should be 1.

So, I am considering them as arbitrary values, but; obviously, keeping in mind that the sum is 1, but if you go one step forward; obviously, your question would be yes I will try to basically make the sum as 1, but does it ensure that the sum of the squares are the minimum. So; obviously, that we have to find out; and what you can do is that, there are different type of mathematical methods, iterative methods based on which you can find out this alpha 1, alpha 2, alpha 3 which is 50 percent as given in this problem 33 percent and 17 percent. So, let us do this problem in a simple way.

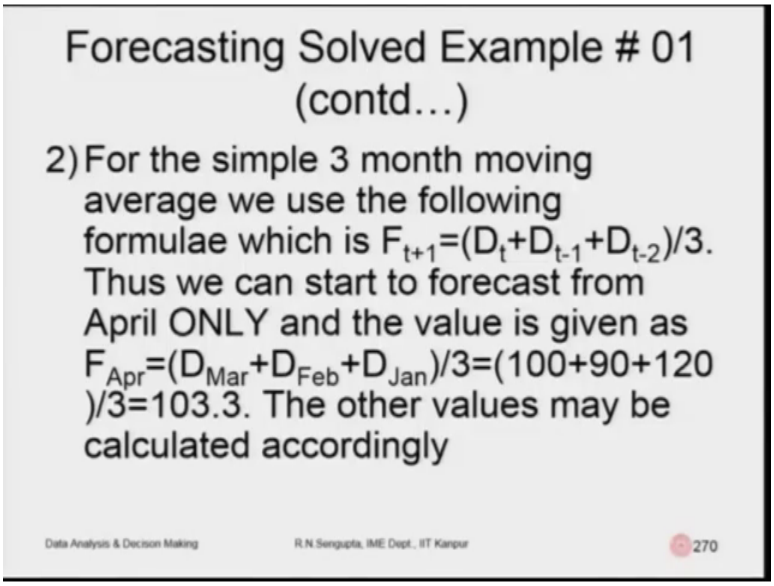
(Refer Slide Time: 26:23)



Now, the first method the naive method uses the demand for the current month and the forecast for the next month. So, whatever the demand is would be the actual forecast for the next month so; that means, F t would be equal to Y t minus 1. So, F forecast for month 2 would be with the actual demand for month 1 and we will proceed accordingly. Now, so hence it will be given that Y t is equal to F t plus 1 and in place of Y we are basically writing the D which is the demand for period t.

So, for February we would have February would basically be the demand of January which is 120, and in the same way we can continue writing where for November forecast would be the actual demand for October. Then the forecast for December would be the actual demand on November and so on and so forth. For in this case the forecast for November is equal to the demand of October which is 90, the other values may be calculated accordingly. So, this name one is that very simply the demand of today would basically be the forecast for tomorrow one time period.

(Refer Slide Time: 27:31)



Next one for the simple 3 month moving average we are using the simple formula as this. So, it will be basically be we add them up and divide by 3. Now, if you remember we are trying to basically put actually the weights as one third. So, the forecasted value if it is you are trying to find out for t plus 1, then I am trying to utilize the demand for t, demand for t minus 1, demand for t minus 2 and the weightage which I am going to give equal weightages would basically be one third, one third, one third which is 33.33

percent, sum is 1. Thus we can start only to forecast from April only because, April will use the month of January February March and I have written in this way. So, hence the forecasted value of April would be the sum of all these 3 months, which is 100 plus 90 plus 120 divided by 3 which comes out to 103.

Similarly when I go to the month of May, so the May value would basically be equal to the demand of 3 months, which will be February March April add them up divide by 3. If I go to say for example, month of December so, December would basically be we are trying to forecast for December. So, the last 3 months would be considered accordingly. So, it will be September October November add them up divide by 3.

So, with this I will end this 25th lecture and continue more discussions about the forecasting and different type of adaptive methods, and I am sure it will clear many of the doubts you may have. And, then slowly once I finish I will definitely come to the multiple linear regressions later on. Have a nice day.

Thank you very much.