Course Name - Operations and Revenue Analytics

Professor Name - Prof. Rajat Agrawal

Department Name - Department of Management Studies

Institute Name - IIT, Roorkee

Week - 02

Lecture - 08

Welcome friends. So, in our previous session, we discussed that how predictive analytics is helping us in solving the situations, where we have more complexity in our data. We introduced two cases, one where we have trend and one where we have trend and seasonal factors also in our historic data. Now, in that case, how are we going to solve our situations? That are we going to discuss in this particular class.

Here, if you see we are going to discuss primarily the case of trend and seasonal variations. We discussed in our previous session that you know these types of models under the name of winter model, where we have both trend and seasonal components. Now, when we have trend and seasonal components in our data like for your refreshment, I am saying that this type of models are possible, when you have both trend and seasonality in your data. Now, in this particular case, if you have trend and seasonality, our answer will come with this equation. The systematic component of demand which is going to be our forecast will come with level plus trend into seasonal factor and seasonal factor which I am assuming in this equation is a multiplicative factor.

Systematic component of demand = (level + trend) * seasonal factor

So, I will calculate for my next forecast Ft1, I will calculate the current level strength LT plus Tt and I will multiply it, please remember with the seasonal factor of the respective time for which I want to get the forecast, that is T1.

$$\mathbf{F_{t,1} = (L_T + T_t) \times I_{t,1}}$$

Please remember this very important thing and if I want to calculate the forecast for n period, it will be possible to do if I do not have that much of data, I can do with the help of this current data, where this is Lt plus nTt into Itn.

$$\mathbf{F_{t,n} = (L_T + nT_t) \times I_{t,n}}$$

This is the big mistake many of us do when we are using the winter model for our forecasting purpose. That you have to use the more most commonly what we do let me tell you what majority of us do that is LT plus Tt multiplied by It.

$$\mathbf{F_{t,1} = (L_T + T_t) \times I_t}$$

This is what most of us will do, you are using current trend current base and current seasonal index that is what you are going to do most of the time, but this is wrong this is correct.

So, please ensure that you use the appropriate seasonal factor. The use of an appropriate seasonal factor is required in the Winters model or in cases where seasonality is present in your demand data. Now, how are we going to do this? This is how it is detailed here: Ft1 is the forecast for the next period that I am going to calculate.

LT plus Tt multiplied by St plus 1—please remember what I just explained. And if I want to forecast for L periods ahead, LT plus LTt—this is T plus L—this is again a mistake even by me. It should be LT plus L that I have to calculate, and I have to use the seasonal index of the L period ahead.

$$\mathbf{F_{t+1} = (L_T + T_t)\, S_{t+1} \text{ and } F_{t+l} = (L_T + lT_t)\, S_{t+l}}$$

Now, with this in mind, I have to regularly update my base, my trend, and my seasonal factors. And for calculating these, I am using three smoothing constants: alpha, beta, and gamma. First, I will calculate the initial values of the base, the initial value of the trend, and the seasonal factors S1 to SP, if there are n number of different types of seasonal factors. Because your data may have seasonality at a particular interval—most likely monthly, in some cases half-yearly, and in others quarterly. So, depending on how many seasons there are—how many periods are in a particular cycle—the seasonality will be calculated accordingly.

Now, first—and I hope you remember my equation—for the calculation of St, Tt, or LT, alpha a plus 1 minus alpha b, this equation is the most generic one.

$$\mathbf{S_t, T_t, L_t = \alpha A + (1\text{-}\alpha)B}$$

So, you have to replace alpha, a, and b as the case may be—whether you are calculating the seasonal index, the trend, or the level values—as your a, b, and alpha will change. For example, if I am calculating the level value, it is alpha Dt plus 1 minus alpha St minus 1, LT minus 1.

$$L_{t+1} = \alpha(D_{t+1}/S_{t+1}) + (1 - \alpha)(L_t + T_t)$$

If you remember the original equation when we discussed the simple exponential smoothing method, your LT1 was alpha Dt plus 1 minus alpha LT.

$$\mathbf{L_{t+1} = \alpha D_{t+1} + (1\text{-}\alpha)L_t}$$

Now, if this was A here, this was B here.

In the current circumstances, this will be alpha Dt plus 1, St plus 1. Dt plus 1 upon St plus 1 is basically de-seasonalized demand. This is the de-seasonalized demand of the current period and this is LT plus Tt which is B here. Similarly, when I am calculating the updated trend it is Tt plus 1, LT plus 1 minus Lt, that is the current trend minus the existing trend and then here you see this T plus P, plus, 1 using my current data.

$$\mathbf{T_{t+1} = \beta(L_{t+1} - L_t) + (1\text{-}\beta)T_t}$$

Let us say, if I have monthly data and I am currently in the June.

So, if I want to calculate Lt plus 1, let us say it is L June, Tt plus 1 it is June and St plus 1, St plus P, plus 1 is S June 2025. All these are June 2024, this is also June 2024, but this P means number of periods. So, in fact, in my current calculation, in my current calculation, this St plus P, plus 1 has no use. In the current calculation, this St plus P, plus

1 will have no use. This St plus P, plus 1 will be used, when I will move to year 2025, when I will be requiring this value, when I am in the month of June 2025.

$$S_{t+p+1} = \gamma(D_{t+1}/L_{t+1}) + (1-\gamma)S_{t+1}$$

At that time, when I will be looking for forecast for July 2025, sorry when I am in the month of May 2025 and I want to forecast for June 2025 this updated seasonal index will be used at that time. So, we are doing these calculations for a future period not for immediate period. So, L and T will be calculated like this and third the updation in my seasonal index will be done with the help of gamma a plus b and this is the seasonal index Dt plus 1 upon base value plus the existing seasonal factor. Now, all these three alpha, beta, gamma between 0 to 1 and as we have continuously saying popular values of these coefficients are between 0.05 to 0.30. These are the popular values of alpha beta gamma and we use a lot of forecasting error measures for deciding what are the correct values of alpha beta gamma.

Now, this was the data we were discussing in our previous session also, and with the help of this data. You see, here our seasonality is coming in the quarter. So, you have four periods in a year, and then again the data is going to be revised. So, Year 1 has data of only three quarters; the first quarter data is missing here. For Year 2, we have data of all four quarters, and for Year 3 also, we have data of all four quarters, and for Year 4, we have data of the first quarter.

So, in fact, you can easily understand our forecast requirement is for Year 4, Quarter 2. Now, before I go to the next slide, I would like to ask you, and I request all of you that before I write, you should also try to think: what should be the equation for getting the forecast for Year 4, Quarter 2? The Year 4, Quarter 2 equation will be: you have to calculate the L for Year 4, Quarter 1, plus the trend for Year 4, Quarter 1, which will be multiplied by the seasonal factor for Year 4, Quarter 2.

$$F_{Y4Q2} = (L_{Y4Q1} + T_{Y4Q1}) \times S_{Y4Q2}$$

This is the forecast for Year 4, Quarter 2. So, my duty now is to calculate all these three terms: level, trend, and seasonal factor for Year 4, Quarter 1; Year 4, Quarter 1; and Year 4, Quarter 2, respectively.

To start this process, let us first understand how we are going to do this. So, initially, we are going to have these initial estimations of S1, S2, S3, S4. Now, how are these S1, S2, S3, S4 estimations coming? For that purpose, you can take the average of all the demand. And you need to divide the demand of a particular period by the average demand, and that will give you the S1, S2, S3, S4—sorry, we have only three quarters of data given here.

So, you have S1, S2, S3, and then you may use these S1, S2, S3 for continuously updating your data using this particular equation for updating the seasonal factor. And with this, now to show you how these calculations flow, we would like to show you an example in the Excel format. So that you can easily understand how these calculations can flow in an automated manner. So, now we will go to the Excel sheet, where we will see how all this data flows in a systematic manner. So, let me stop this and go to the Excel sheet.

So, here, if we see the data available to us, it is slightly longer so that we can understand the calculations better. We have given you the data for 5 years, and every year is divided into 4 quarters: 1, 2, 3, 4, like that. Here, I am requesting you initially to focus only on the first 3 columns. CDE, where we have quarterly data available to you. Now, using this quarterly data, we are first required to calculate the seasonal index.

Now, how we have calculated the seasonal index is given in the calculation of the seasonal index. Now, for quarter 1, quarter 2, quarter 3, and quarter 4, the data given to us is 20, 250, 3514, 4120. If I use the data of quarter 1 and these. So, the average demand of these quarters is given to us and taken here, which is the average of, if you see in the formula bar, E7, E11, E15, E19, and E23, what I have encircled just now. And that average is coming to 4205, 4205. Similarly, we have done the calculations for the average demand for quarter 2, quarter 3, and quarter 4, which you can see in this particular column where the average demand is listed.

Now, after getting this average demand, in fact, please do not see this column. Let us come back to column number F, where we have the demand of a particular period divided by the average demand for that period. So, as you see, 2250 divided by 4205. If you see the formula which we have inserted in column F for cell number F7, that is directly the value we have transferred from column T here. So, if you see the value in T7, 0.83, if you see this 0.83, this is S7—that is, this value divided by S7 to S10.

So, S7 to S10, that is the average of these four, and the value we have here, 4205, is like this. If I explain this to you, it is 4205 divided by the average of 4205 to 9542. So, in the same way, we have done all the other calculations: 2945 divided by the same denominator, 3517 divided by the same denominator, and 9542 divided by the same denominator. So, you can see that because of the seasonal impact, the demand of quarter 4 increases almost double the average demand of all the quarters, but in quarters 1, 2, and 3, the demand is relatively low—it goes very low in quarter 2 and jumps significantly in quarter 4. So, these are the initial seasonal indices, these are the initial seasonal indices which we have put here.

Now, our focus is required on this part. Here, our seasonal indices, which we have directly taken from column T, are 0.83, 0.58, 0.70, and 1.89 or 1.90—for that purpose, rounding off has been done. Now, initial levels and initial trend are given to us. Let me go to cell number M6. If I go to M6, this is C56, and C56 is derived from my regression calculations. So, let me tell you that to get the initial values—initial L naught and T naught—we already discussed in our previous session that whatever data you have, if you put that into a regression equation.

So, this intercept on the y-axis is L0, and the slope of this line is our T0. So, we have used the data of, if I go to the upside. So, the initial data of my demand and period, if I use this historic data, this gives me the scope of creating a de-seasonalized demand. Because, when I am talking of 2250, 1737, 2412, etc., there is a seasonal component of 0.83, 0.58, 0.70, 1.89. So, with this seasonal component, this regression equation is not possible or will not be a good idea.

So, what I need to do is de-seasonalize my given demand, and for that purpose, if I go to column I, here what we are doing is dividing the demand, whatever is the current demand, with the seasonal factor corresponding to that particular period. For example, this calculation 2703 is coming by 2250 divided by 0.83, because 0.83 is the corresponding seasonal factor. The next is coming 2980, which is 1730 divided by 0.58, and so on we are going for developing my de-seasonalized initial demands, de-seasonalized initial demands. And then, using this de-seasonalized demand with periods, if I go to the regression equation, I get the coefficients of my regression equation where the y-intercept is coming 2644. So, that is my L naught, and the T naught, that is the slope of this line, that is the trend, which is 229.

So, we got these values of L naught and T naught to start our process, and that is what we have written here directly as our level and trend. So, this becomes L naught, T naught, and our initial seasonal factors are also given to us. Now, what we are going to do is our forecast, therefore, in this column, you may remember that it will be L plus T multiplied by the seasonal factor of the corresponding period. So, in fact, this 2391, I request all of you to please calculate on your own. This 2391 will come if I calculate the F1, this is L naught plus T naught into seasonal factor 1, that is the calculation for getting 2391.

And similarly as I go ahead, as I go ahead you see that when I am calculating 1772, I am using the 2822 as my base, 219 as my trend and the initial seasonal factor which is already available with me 0.58. And now, when I was here at the first factor 0.83, so, I will update this 0.83 also and that updated value of 0.83 incidentally it is 0.83 only but, this updated value is coming here, this 0.58 updated value is coming here. Since, the gamma is 0.1, so the impact is not so high. 0.70 becomes 0.71 and 1.90 becomes 1.87. So, the initial seasonal factor for quarter 4 was 1.90 and after year 5 the seasonal factor for quarter 4 has become 1.85. So, in this way, there is some changes happening in the seasonal factor which is going to be more, if I can show you that if I change the value of gamma from 0.1 to let us say 0.20 just to demonstrate.

And then you see that the values of our seasonal factor which was initially 0.83 in the quarter, in the year 5 quarter 1 this has become 0.87 some visible change you can see that from 0.83 to 0.87. So, the gamma is not so dynamic beta or you can say tread is slightly more dynamic and level is even more dynamic. And generally, generally you will say though it is not a rule generally, you will see that in alpha, beta, gamma, alpha is generally more, beta is less than alpha and gamma is further less than the gamma than the beta. So, here with this kind of a static development you can very easily increase your model and then you see that some more interesting things in this table, because you can see now various things in a glance. Now, when we are doing this using the Winters method, our forecasting errors are coming for each period as 6.29 to may be it is going up to 44 percent in some period which is relatively very very high forecasting errors.

And then, after all these calculations, you can calculate your MAPE, mean absolute percentage error, which has become very popular in the field of analytics these days. It is coming close to around 11.81 percent. It is around 11.81 percent. If, I see some other methods also, if I can see the column G and H in this table, where we say naive that naive means I am not doing anything whatever was the demand of the previous period, I am considering the same demand for my next period. I do not consider any kind of seasonality, I do not consider any kind of trend, I have no idea, no idea of predictive analytics for me the predictive analytics is that whatever is the current demand the same demand will be in the next period.

There is a change; whatever change will happen in the next period will be the one for the further next period, and on the basis of that, you see that our forecasting errors are so high for every period because this data is not suitable for this kind of methodology. However, this naive method can also work. You can say that it is a kind of very high simplification—the highest degree of simplification of our moving average method—where we are taking the moving average of just one period. You can say that this naive method is taking the moving average of a single period. On the other hand, you have the regression method or Winters method, which are more or less equally powerful. In the regression equation, we are developing y equals a plus bx, where x is the period and y is the forecast.

So, as the period changes, we are having a new forecast. But in this case also, I see that whenever there is a seasonal factor, the forecasting error will be very high—like in the fourth, eighth, or twelfth period—you see very high forecasting errors happening. The overall mean absolute percentage error is more than 10 percent. Then, you have this Winters method, which we discussed in detail in this class. Here also, the forecasting errors are more or less comparable to the regression method. In this class, we discussed the method of three constants using multiplicative trends.

I request you to please try the same data with the additive trend. Think of the additive trend, and when you do, the equation for determining the forecast will also change. I also request that, in this particular case, we are doing the calculation with a particular set of alpha, beta, and gamma values, where we are taking alpha as 0.3, beta as 0.2, and gamma as 0.2. I request all of you to play with the values of alpha, beta, and gamma and see how your forecasting errors and MAPE change if alpha, beta, and gamma are adjusted. We have inbuilt the formula of MAPE also in this table. If I can go to the MAPE, that is Q28, you can see the formula there.

So, this is the average of Q7 to Q26. So, that is how we have calculated the mean absolute percentage error—the absolute percentage error for every column, every value in the column Q—and the average of all these error values gives us our mean absolute percentage error. So, that is how predictive analytics can be used to improve our estimations for the future, and these forecasting models will be the base inputs for various other decision-making processes, such as inventory management, material planning, facility planning, etc., which are other important decision areas in our operations management discussions. With this, we come to the end of this particular session. Thank you very much.