Course Name - Operations and Revenue Analytics

Professor Name - Prof. Rajat Agrawal

Department Name - Department of Management Studies

Institute Name - IIT, Roorkee

Week - 02

Lecture - 06

Welcome, friends. In our last two sessions, We discussed the use cases of predictive analytics. We primarily discussed the use case of predictive analytics in operations management, specifically in the field of forecasting. We discussed how we can apply it either in a very simple model, such as the Moving Average Method. We also discussed applying it in the form of Exponential Smoothing Methods.

We can also apply it in the form of Regression Analysis. And we only discussed regression analysis where we had one independent variable. But in practical life, you may have more than one independent variable. Two, three, or even four types of independent variables are possible. Some of you must have also heard about econometrics, where a single regression equation may not be sufficient for predicting the outcome.

We may use a combination of equations—two, three, or four—for predicting the outcome. Generally, complex regression models are easily possible, and these models may provide better estimations of future demands, better predictions, etc. But considering the scope of our discussions, we limited ourselves to just one independent decision variable. Now, since all these are simply predictions, predictions may have their own errors. When exit polls for the elections were conducted, there were different kinds of estimations about the possible results we were going to have.

And we are recording this course on June 4th, when the results are actually out, and there is a huge error in those estimations and the actual observed value. So, we need to know how we can measure the errors in our estimations, and there are so many different measures of error in our predictions. That is what we are going to discuss in this particular session—how many different ways you can express your error. The error is the same: I estimated 200 units, but the actual demand turned out to be just 150 units. So, the error is 50 units. Now, this error of 50 units can be expressed in multiple ways.

In the coming slides, we will talk about the estimation of error in the forecasting process. We will discuss one measure of that estimation in the form of average error. Another popular and very important measure is mean absolute deviation, mean squared error, mean absolute percentage error. Then we will talk about the tracking signal, and we will also cover that all these errors come from our time series analysis. It may be moving average or exponential smoothing. But we will also see some errors in regression analysis.

So, this particular session is devoted to understanding errors and their representation in multiple ways. So, as I said, the simple meaning of error is Et equals Ft minus Dt.

$$\mathbf{E_t = F_t - D_t}$$

You estimated Ft as 200, but the actual demand turned out to be 150. So, in the simplest way, I will say that the forecasting error is 50 units. So, you will say that if it is a 50-unit forecasting error, how will you represent this 50 in different ways?

One simple on a lighter side, it is in the digits, I can write it in the words also 50, but this is not the meaning I want to say. I want to say that, I can represent this 50 as a percentage error for my actual forecasted value 50 as percentage of 200 one way of representing it. 50 as percentage of actual demand another way of representing. And then, if I collect data for forecasting error for multiple periods then I will find more interesting ways of representing this forecasting error data. That is where all these different way of representing the forecasting errors are possible.

You always need to see that in how many different ways because 50 here may look very high number because 200 and 150. Now, consider another scenario. Another scenario, in that scenario I forecasted 10,000 units. Demand forecast was 10,000, but actual demand came out to be 9,950. In this case also my forecasting error is of 50 units.

So, if two of us are sitting together one manager says sir my forecasting error is 50 units, another manager also says my forecasting error is 50 units. So, for someone it may appear that both the managers have equal amount of forecasting error. But, that is not the correct representation of the data, because 50 when you are forecasting 200 and 50 when you are forecasting 10000, both these are on a very different scale. So, just 50 is not complete you need that 50 as a percentage of your forecast or percentage of your actual demand. So, that is why we require different way of representing so that you can have complete information and you can understand the forecasting error in a better way.

So, one very common way of representing the forecasting error is average error. Average error this is the formula for that. So, it is very simple that for all the periods you are calculating the forecasting errors E1, E2 up to En, if n periods are there and then taking the average of these n periods is your average error. If you take the average of the n periods that is your average error.

$$AE = (\frac{1}{N}) \sum_{t=1}^{N} E_t$$

Another is but as we have already discussed in the many of our previous sessions, particularly when we started discussions on descriptive analytics, we said that average is not a very good representation. Why? because it is quite possible, let us say I have six periods.

In one period it is plus 50, another period it is plus 100 forecasting error, another period is it is plus 75, then it is minus 100, minus 50 and minus 75. So, these are the forecasting errors of individual periods and when you take average of these, it will come out to be 0. So, if I simply have this data that average error for t is not a good forecast. You are going up to plus 100 also and you are going up to minus 100 also. So, positives are neutralizing the negatives therefore, average error is 0.

So, we need average error because it is very simple and average error should be as minimal as possible, but we cannot rely only on average error because of this reason that there are positives, there are negatives and positives and negatives will cancel out. Then, we have to resolve this particular issue, we have another very common used method of forecasting error. That is Mean Absolute Deviation MAD, which is I think most used. Now, what we are doing we are actually taking the only absolute value of the forecasting error. So, you see that with these symbols we are considering only the absolute value of our forecasting error.

So, if data is like plus 50, plus 100, plus 75, minus 100, minus 75, minus 50. So, if my average error is 0, but for the calculation of MAD, I will take the absolute values of all these periods and absolute values will be without any sign 50, plus 100, plus 75, plus 100, plus 75, plus 50 divided by 6. So, you can calculate easily that the 450 divided by 6, that is going to be the MAD. So, average error which was 0 for this data but MAD is 75.

MAD = (50 + 100 + 75 + 100 + 75 + 50)/6 = 450/6 = 75

So, you can see that actually our deviation in every period is around 75 units on an average and therefore, MAD is independent of the sign of your or the nature of your independent errors.

It is therefore, considered to be more used measure of forecasting errors. In all our predictive analytics, MAD is very very useful measure of forecasting errors. Then, another important measure of forecasting error is Mean Squared Error. It is a kind of a improvement over our absolute deviations or average deviations.

$$MSE = \left(\frac{1}{N}\right) \sum_{t=1}^{N} E_t^2$$

For example, if I have data like E1, E2, E3, E4, E5, E6. E1 is plus 2, E2 is plus 3, E3 is plus 5, E4 is minus 20, E5 is minus 4, E6 is minus 1. So, mean squared error is helping in two ways. One, because of square terms the effect of sign whether it is a positive or negative that is neutralized. So, E1 square 4 then 9 then 25 and then all of a sudden you see a kind of a outlier 400 comes then 16 and then 1. Now, when you are going to take average of these square terms, this average will significantly jump because of this one outlier 400.

Then all of a sudden you can identify that ok, I was moving at a particular rate. But, because of one outlier all of a sudden my mean square error value has changed significantly and you can see that what happened in this particular period. Why I got so high? Because, you see in relation of 2, 3, 5, 20 is not big but when you do the square then it is clearly visible. So, for having a better clarity, better visibility so that outliers can be clearly identified square terms have double parallelization and therefore, MSE is useful in identifying some kind of large deviations in your error data.

Then, another you can say measure which has become very popular in the field of our AI models. MAPE, Mean Absolute Percentage Error, this is in all our AIML based models, when we are looking that how effective is our model, whether it is giving good predictions or not. So, in predictive analytics particularly the mean absolute percentage error has become a very popular you can say criteria for deciding the suitability of a particular model. Like, if you remember in our exponential smoothing class, we discussed only one type of exponential smoothing that is simple exponential smoothing. But, if you remember we discussed nine different types of models in that class that based on the combination of trend and seasonality, whether these are multiplicative or additive, whether trend is there or not, seasonality is there or not.

Generally, by our naked eyes we cannot identify all these characteristics of our historic data. Then, you have to apply different types of possible models on your data. Now, once you have outcomes from those different models which model is giving the best result. So, that now onwards you can use that model, it is being decided by the measures of forecasting error. And for that purpose MAD and MAPE, these are the most commonly used forecasting error measures which help you in taking a decision which model is suitable not only which model is suitable, what level of parameters are suitable for our decision making?

So, both these things—because, as we know, alpha, beta, gamma—these are the smoothing constants we are using, and I said that the values may vary from 0.1 to 0.3.

Now, which value is more suitable for my particular case, for my data? All these things—selection of model and selection of parameters—can be decided with the help of MAD and MAPE. So, MAPE is basically the calculation of the average of the percentage errors, as a percentage of demand data that we are calculating. And here again, this Et is taken in the absolute manner. This Et is taken in the absolute manner—that is what we are calculating: the mean absolute percentage error. And we have some data with us, and with the help of the data, one method is giving the forecasting errors of these values—these you can say as E1, E2, E3, E4—and the other method is giving these values of E1, E2, E3, E4.

Now, method 1 has lower MSE and MAD relative to method 2 and would be preferred if either criterion was used. So, if you see method 1 and method 2, the individual values are given to you. Now, you can calculate. In fact, it is required to calculate MSE and MAD for these data. And then, based on this data, you can see which method should be used for this particular scenario when you have the data for individual periods.

Now, in organizations, because these calculations may happen over a period of time, but in real time, whether our predictive activities are happening appropriately or not. So, we create a kind of dashboard in the organization, and that dashboard is continuously tracking our actual demand, predicted demand, and the forecasting errors also. And a third kind of thing is developed over a period of time, which is known as a tracking signal, which is a relation between our demand data, the actual forecast, and the mean absolute deviation. The relation between all these things can be put together in the form of a simple thread, and on the basis of the movement of this thread, you are able to see whether our predictions are in the right direction or not. And therefore, this tracking signal has become a very common means—more used than these individual forecasting errors—because it gives you a kind of single package, which is good enough to decide whether our forecasting system is going in the right direction or not.

So, we calculate the tracking signal with the help of this particular formula: tracking signal is the running sum of forecasting error divided by the mean absolute deviation up to a particular period. So, if I have data for, let us say, 4 periods, and I have to calculate the tracking signal. So, I will calculate the tracking signal for every period and, for every period up to that point, what is the running sum of forecasting error? Because for every period, the running sum of forecasting error—let me explain with the help of data. First period, the forecast error (Et) is 50, next is 75 minus, then plus 20, then plus 40.

So, RSFE will be 50, 50 minus 75, it will become minus 25, plus 20, it will become minus 5, plus 40, it will become plus 35. So, this is the calculation of the running sum of forecasting error. Similarly, for MAD calculations for the same data: plus 50, minus 75, plus 20, plus 40. First, we need to have absolute errors, and the absolute errors I am writing like this: 50, 75, 20, and 40. Now, MAD will be 50 upon 1, 50 plus 75 by 2, 50 plus 75 plus 20 by 3, and 50 plus 75 plus 20 plus 40 by 4. These are the MAD calculations.

So, the tracking signal will become RSFE divided by MAD for respective periods. This is a detailed calculation of one example for tracking signal, and you can see that we have 1, 2, 3, 4, 5, 6 months. The forecast is given that we have a static forecast for every period, that is 1000, 1000, 1000. The actual demand data is like 950, 1070, 1100, 960, 1090, and 1050. For the respective period, the forecasting error sometimes is minus—this Et is basically this is Dt, this is Ft.

So, this Et calculation is Dt minus Ft. So, minus 50 plus 70 plus 100 minus 40 plus 90 and so on. Now, the running sum of forecasting error is minus 50 plus 20 because of the 70 addition, and because of the 100 addition, it is 120, and so on up to 220. These are the absolute values of the forecasting error: 50, 70, 100, 40, 90, 50, and these are the MAD calculations. In this column, the tracking signal is RSFE upon MAD, and in this way, the final column gives you the values of the tracking signal. Then, on a particular scale, if you can see on your screen at the bottom, we have the calculations for the tracking signal.

The middle value is 0, and then there are a few values which are positive, maybe up to plus 3, and maybe up to minus 1, minus 2, minus 3, like that. And then you see that the tracking signal starts with minus 1, then it is 0.33, 1.64. There is a slight decline here, a slight drop, but then it is continuously rising. Now, ideally, a tracking signal should fluctuate around the mean value, the zero value. But here, the tracking signal is continuously rising, which says that our predictions are slightly underestimated. Our predictions are slightly underestimated.

The actual demand is more than the predictions. So, our predictions are happening maybe on a conservative ground; the actual demand may be more, but we are predicting less. So, this kind of challenge you are able to understand from this tracking signal. The tracking signal should naturally predict that means if we are having a fluctuating tracking signal, it means randomly sometimes we are overestimating, sometimes we are underestimating, which is well accepted. But one type of trend is not a good idea. Then, all these errors, which we discussed, are generally found in our moving average smoothing systems.

But in our last session, we discussed regression analysis, and there are possibilities of errors in the regression system as well. The most common error is the standard error of estimation. If you remember, we discussed weak relations, strong relations, and if weak relations are there, Se will be high; with a strong relation, Se will be low.

$$s_e = \sqrt{\frac{\Sigma(Y - \hat{Y})^2}{n - 2}}$$

So, if you see Y and Y cap, Y is the observed value, and Y cap is the estimated value. So, with the help of your data, you can calculate this standard error of estimation, and this diagram helps you understand this concept—that this is the line which is our line of regression.

Now, if our relation is weak, if the standard error of estimation is high, there will be a lot of scattering of data around the line of regression. They may be scattered very far as well. And if we have a strong relation, most of the data will be clustered around the line of regression. You will have better predictions with your regression equation. If your standard error of estimation is very high, In fact, I can recommend that we should not use this regression equation for prediction purposes. Maybe we need to identify a new independent variable X, which can have better predictability. It is, in a way, also an indicator of the predictability of your X—whether X is a good predictor, able to predict your Y or not. So, you can see that as you move away from your regression equation—with one standard error, two standard errors, three standard errors—you are continuously increasing the scatter of your data, and if the data is more scattered, it is a weak relationship, and therefore, it also questions our selection of independent parameters. So, whatever predictive analytics we apply, there are always chances of errors—it is not a 100 percent correct estimation.

But we should also know how to handle these errors. Once you know the use of these errors, as I told you in my earlier discussions when I was talking about smoothing constants and moving averages, in that case, whether the value of the smoothing constant is correct or not, whether the model is correct or not—that is the use of error. In this particular case of regression, whether the independent variable is correct or not—that is the use of standard error calculations. So, errors may help you improve your predictability so that you can use predictive analytics in a more efficient manner. With this, we come to the end of this particular session. Thank you very much.