Production and Operation Management
Professor. Rajat Agarwal
Department of Management Studies
Indian Institute of Technology, Roorkee
Lecture 14
Forecasting errors

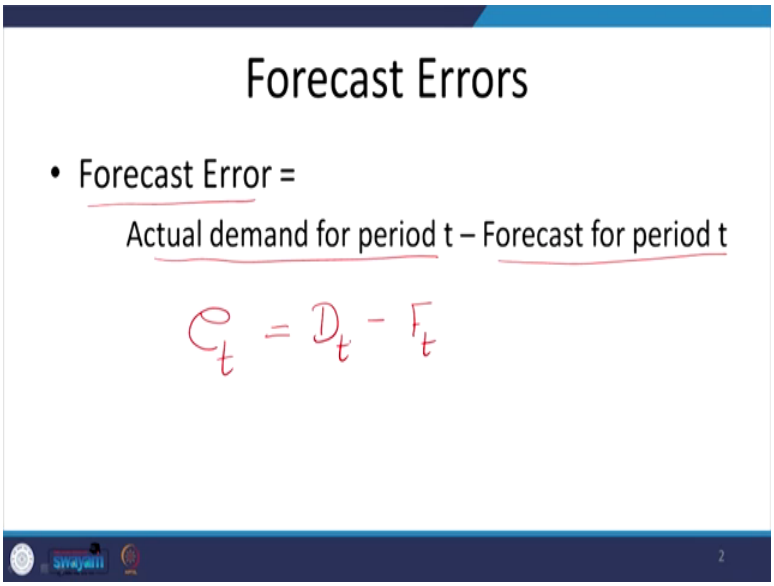
Welcome friends, now we are moving into the fourteenth session of this course of Production and Operation Management. In last many sessions we were continuously talking about one thing, that whether our forecasting model is accurate or not. Because forecasting is estimation of future demand. It is not the actual demand.

So, there will always be some difference between our estimations and the actual the demands. But a good forecasting model is one where this difference is minimum. Because if difference is too much, or if difference is large, therefore either you will be having more inventories or you will be running out of the stock. Either of these two things will happen.

If you have overestimated, if you have overestimated, and in that case you will go with extra inventory at the end of the period. If you have underestimated then you will run out of the stock and some of your customers may go empty handed and their customer satisfaction level will go down, your service level will go down.

So, that will also affect the perception of customers towards your company towards, retail outlet, whether you will be able to provide the service or not. So overestimation and underestimation both are dangerous for the organization. Therefore accuracy of forecast is very important. That if you accurately forecast your demand, it will not be hundred percent accurate. But if you are close to that actual demand then probably you can maximize your profit you can minimize the customer dissatisfaction and you can also become competitive in this very tightly available resource world.

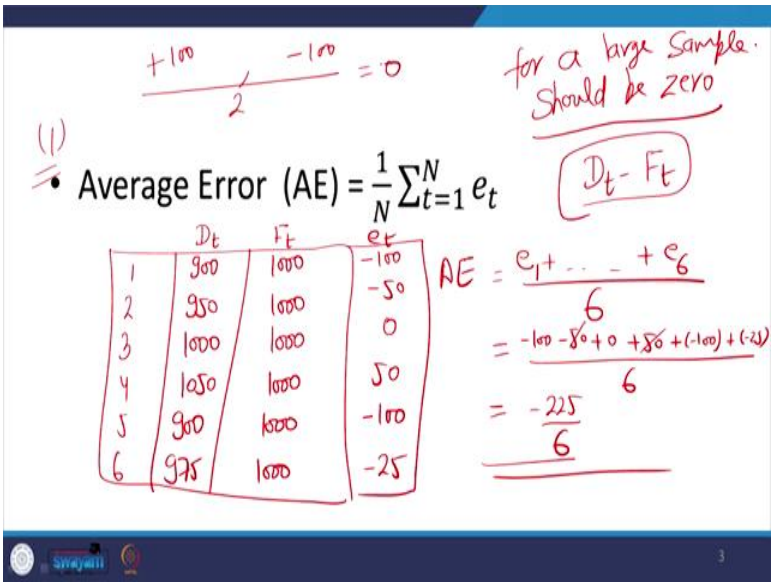
(Refer Slide Time: 02:44)



So, therefore the issue of forecasting error is very important. Now first we understand what is the basic meaning of forecasting error. So, the basically forecasting error is simply defined as the difference of actual demand for period t and the forecast for period t. That means the forecasting error for a particular period t is the difference of actual demand and the forecasted value.

So, in the formula way, we can write that forecasting error for period t, is the difference of actual demand D represents the actual demand F represents the forecast. So Dt minus Ft is the forecasting error for a particular period. But alone, calculation of this forecasting error is not going to serve my purpose and for that purpose we have developed different types of measures, different types of computations out of this forecasting error and these different measures of forecasting error will help us in taking good number of decisions. So, let us see what are those measures of forecasting errors.

(Refer Slide Time: 03:56)



The first measure of the forecasting error is average error. Now average error is the simplest measure of forecasting error. That is you have this data for past many periods 1, 2, 3,4,5,6. In each of these periods, the actual demand and forecasted demands are available. Actual demand was 900, 950, 1000, 1050 and then again 900 and 975 and for all these periods, you are forecasting 1000 like this. So, for each period you will calculate your et. You will calculate et for each period, so Dt minus Ft it is minus 100, it is minus 50, it is 0, it is 50, it is minus 100, it is minus 25.

So, the average error actually is the average of all these individual average, individual forecasting error. So, that is the formula of average error that you are taking summation of et from t equals to 1 to t equals to 9. So, you can write this formula in this particular case, e1 to e6 divided by 6. That is the formula for average error in this particular case and therefore minus 100, minus 50, plus 0, plus 50, plus minus 100 and plus minus 25 divided by 6.

That will become the forecasting error. So, this is going to be cancelled, so this is minus 225 by 6. That is the forecasting error. Now, this is forecasting error for this small sample of six periods. Normally it is desirable, normally it is desirable that when we have a large sample, when we have a large sample, forecasting error in terms of average error should be 0. Average error should be 0 for a large sample.

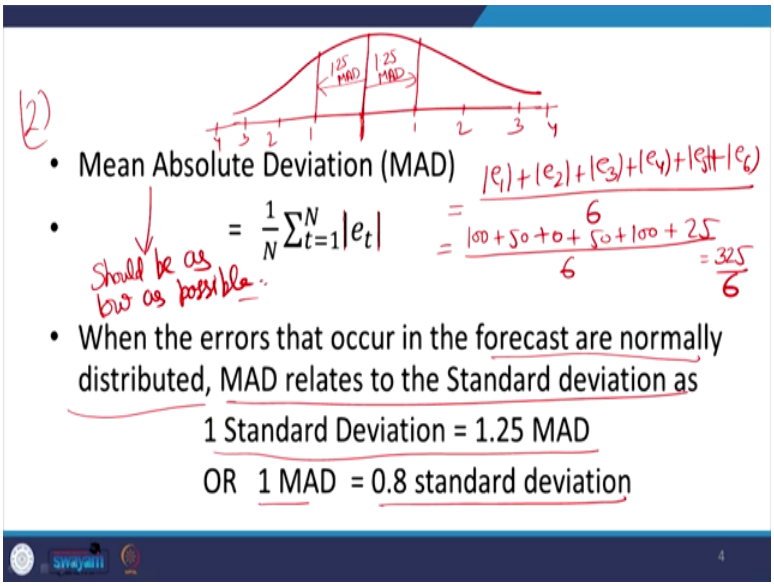
Because sometimes forecasting error will be positive, sometimes forecasting error will be negative. So, positives will cancel out the negatives. So, for a large sample, forecasting error should be 0. If forecasting error for a large sample is not 0, if it is either positive or negative like in this particular case it is minus 225 by 6.

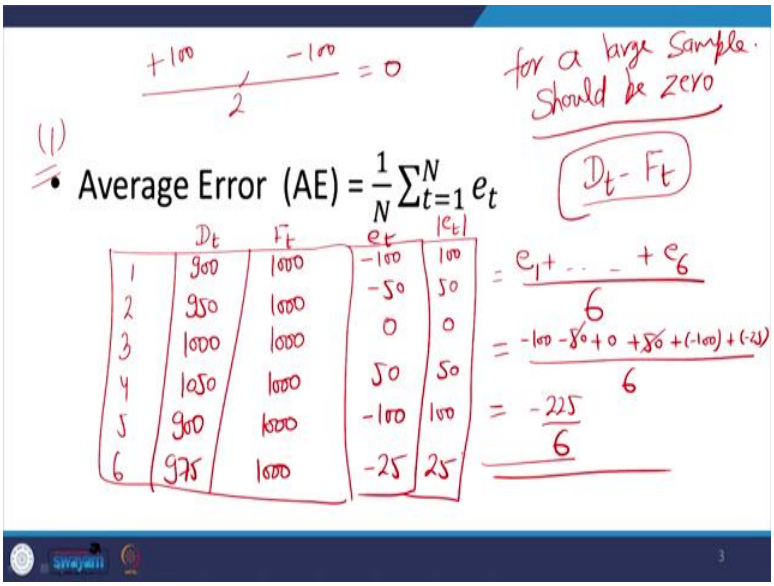
So, it is saying it is saying what is that forecasting error is Dt minus Ft. So, when forecasting error is coming in negative, it means we are continuously overestimating, we are all the time overestimating, in first period you estimated 100 unit extra in second you are estimated 50 units extra then in fifth also you estimated hundred extra, in fourth you estimated 25 extra. So, there is lot of over estimation.

And this overestimation will result in extra inventory at the end of the six periods. So, therefore this is not desirable. For a good forecasting model average error should be 0. So, this is one measure of, and we also know from the classes of our statistics that average a very good measure. Because sometimes the forecasting error in one period is plus100, another period it is minus 100.

So, if you take the average it will be 0. Though for individual periods, forecasting errors are very high, plus 100, minus 100. But when we take the average, it becomes 0. So when I am just the average error and it is coming 0, I may be happy that my model is very good. But for individual periods the errors are very, very high. So, average error is not sufficient condition for checking the suitability of the model.

(Refer Slide Time: 08:47)\





	D_t	F_t	e_t	$\lvert e_t \rvert$
1	900	1000	-100	100
2	950	1000	-50	50
3	1000	1000	0	0
4	1050	1000	50	50
5	900	1000	-100	100
6	975	1000	-25	25

So for that purpose, we go to the second measure of forecasting error and that second measure of forecasting error is mean absolute deviation MAD. This is a very popular measure of forecasting error. Mean absolute deviation is a very popular measure of forecasting error, here you see that we have put these signs of absolute values.

So, you take only absolute values of et. Now going back to the previous data, here if I take the value of the absolute et, so this becomes 100, this becomes 50 this is 0, this becomes 50, this becomes 100, this becomes 25. So, here for the calculation purpose I will only take the values

which are without sign. These values are without sign and therefore MAD will never be 0 MAD cannot be 0, because I am not considering their sign and therefore it is 100, 50, 50, 100, 25, 100, 50, 0, plus 50 plus 100 plus 25 divided by 6.

So, this 100, 50, 100, 50, 300, 325 by 6. That is the value of MAD. So, MAD cannot be 0. Because you are taking the values of individual forecasting error without the applicable sign, because you are taking their modulus values. Now, this is very useful tool for determining the suitability of your model. We want the values of MAD should be as low as possible. It cannot be 0 but if MAD values are low, it is a very good indicator it is a very good indicator that my forecasting model is performing well.

If MAD values are increasing, then there is a matter of concern and we need to see whether my correct, whether this same model is suitable for the changing data. So, because sometime the characteristics of your data may change intermediately and therefore it is possible that you need to change your model intermediately. Initially you started with alpha equals to 0.1, beta equals to 0.2. But maybe after one year or two year, you need to change the values of alpha and beta. So, all those things can be easily determined whether you require some intervention or not with the help of values of MAD.

So well MAD should be as low as possible. It is also very important to know or to understand, that the forecasting errors, if you see the previous table, so we have data for 6 periods minus 100, minus 50, 0, 50, minus 100 and minus 25. So, for a large sample, so for a large sample the values of forecasting error are normally normal distributed. So, the forecast are normally distributed, the errors that occur in the forecast, these are normally distributed and then you can have a relation between MAD and standard deviation.

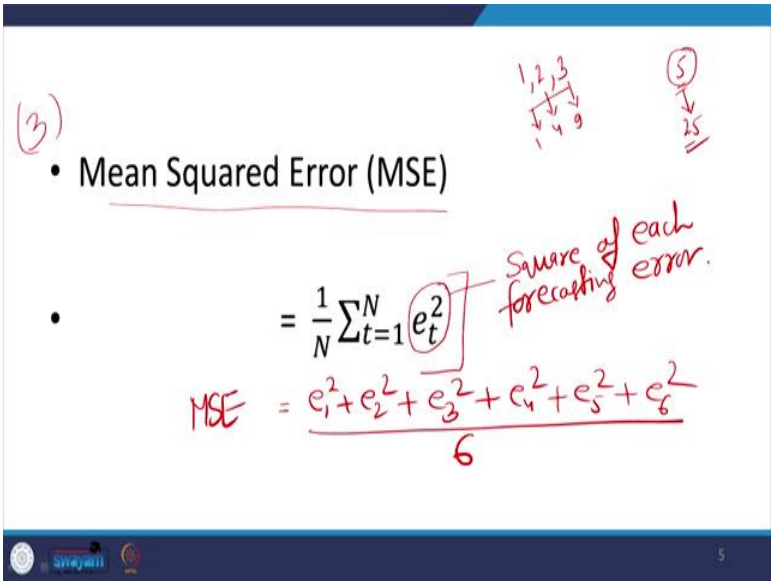
One standard deviation is equal to 1.25 MAD or vice versa 1 MAD is equals to 0.8 standard deviation. Now if you know the standard deviation curve, this normal distribution curve in that you can see that these are 1 standard deviation, this is 2 standard deviation, this is 3 standard deviation and then you can have 4 standard deviation also. So, now you can know that if you are within 1 standard deviation on the either side of the mean value, this is covering your 1.25 1.25 MAD on the either side of this mean value. So, this is 1.25 MAD, this is 1.25 MAD.

So, this way you can also understand that how the MAD value is determining, if MAD value is less. So, your most of the forecast are close to this mean value and if your MAD value is more that means you have a wide range of forecasting errors. So, that is not so desirable that sometimes it is minus 100 and sometimes it may be plus100 also.

So, we want, because forecasting is estimation. So errors are always there. But the limit of error should not be too much. There should be a very small range, within which forecasting errors are acceptable. So you can decide, you can decide for your organization that the forecasting error should be plus minus 2, standard deviation or plus minus 3 standard deviation plus minus 4 standard deviation.

So, depending upon organization to organization, depending upon the type of product you are going to have, if a high unit value product is there, so you may not expect high standard deviation in the forecasting error. But if it is a low unit value product, maybe you can go for 3 or 4 standard deviation type of forecasting error. So, depending upon the type of product the unit value of product the longevity, the perishability all these are aspects affect to choose a particular limits of standard deviation in the forecasting error which you can bear.

(Refer Slide Time: 15:33)



Now the third type of measure of forecasting error is mean squared error is mean squared error. That is the third type of error and mean squared error this is you can say to some extent is an

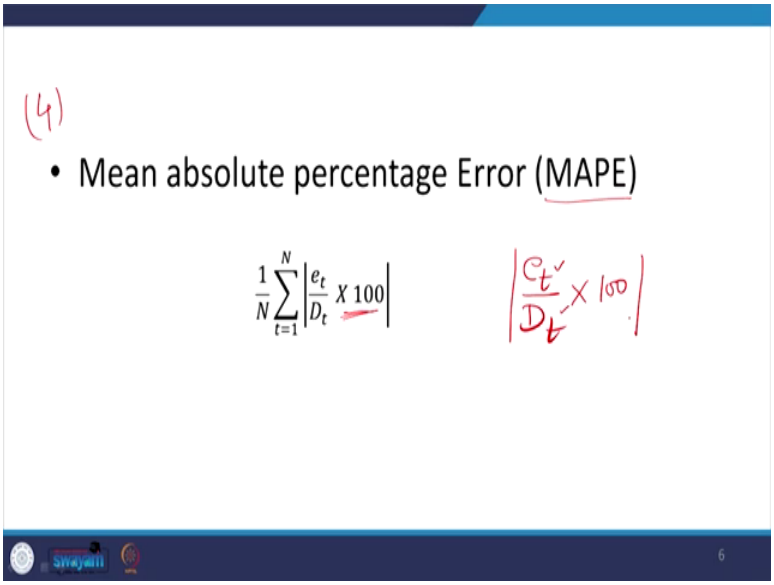
extension of mean absolute deviation. This is an extension of mean absolute deviation and in this case what we are doing we are doing the square of each forecasting error and then we are taking the average.

So, like the example which we are taking e1 square plus e2 square plus e3 square e4 square e5 square plus e6 square divided by 6. That is my mean squared error. So, we are taking and what is the purpose of taking this square, because when we know that the value is 5 let us say, the forecasting error is 5 for some period, for other periods it is 1, 2 or 3. For one period it is 5, now the square of 1 is 1, square of 2 is 4, square of 3 is 9. But square of 5 is 25.

All of a sudden the impact of slight increase in forecasting error is too much when we are doing the square. So the, these errors are identified very uniquely, very separately because of this squared impact. So, mean squared impact gives you direct visibility of a particular period, that why in this particular period we have more forecasting error.

So, that is the purpose of mean squared error. That is also a very common way of penalizing the higher forecasting errors.

(Refer Slide Time: 17:45)



Now another forecasting error measure is mean absolute percentage error, which is also known as abbreviated as MAPE. This is mean absolute percentage error. Now in this case we are determining that what is the percentage deviation, what is the percentage deviation of forecasting

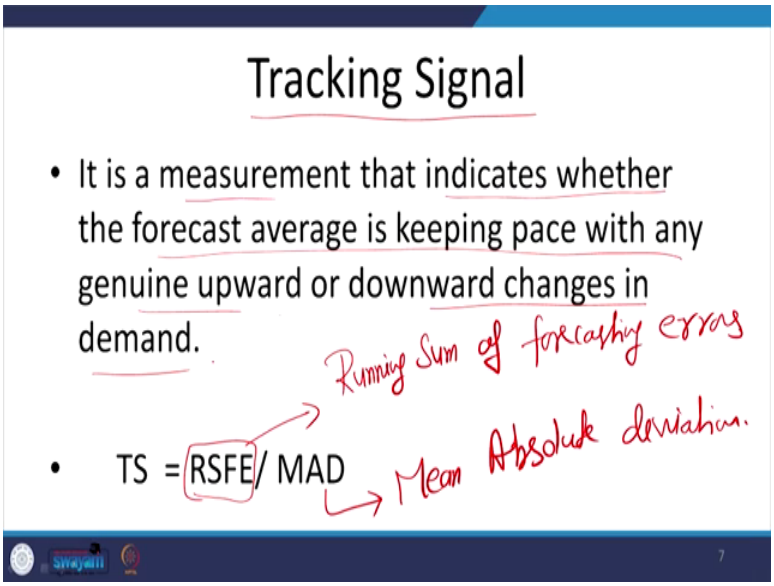
error from actual demand? So percentage error is because we are calculating the percentage or therefore it is multiplied with 100. Otherwise this is et upon DT which tells you that how much is the deviation of forecasting error with actual demand and then we multiply this deviation with 100, it becomes into the percentage.

And sometimes the forecasting error is in positive. Sometimes forecasting error is in negative. So in this particular case, we are taking the mod values of these percentage deviations. Because we want to see that how much overall over forecasting error has deviated 10 percent, 20 percent whether it is in the overestimation size or in the underestimation size. Because there is always a chance if you do not take the modulus value over estimations of 1 period will neutralize the underestimation of another period.

But we have discussed in the beginning of this section that both over and under estimations have their own demerits. So, you cannot create a situation where one cancels out other. So, we require both these things to have a better understanding about forecasting error and therefore in 3 out of 4 we have taken either square or absolute values only in case of average error we have considered the sign of the forecasting error.

Otherwise it is MAD on MAPE in both these cases we have taken the mod values and in case of mean squared error, since we are doing the square so that automatically takes care of the sign. Minus if you are taking the square of minus 5 it is going to be plus 25. So, in that also we have taken only positive sign. So, because we want to see in absolute terms, what is our forecasting error. So, this is another important type of forecasting error measure, MAPE that how my forecast error is percentage of actual demand. So, these are the four important types of forecasting errors.

(Refer Slide Time: 20:44)

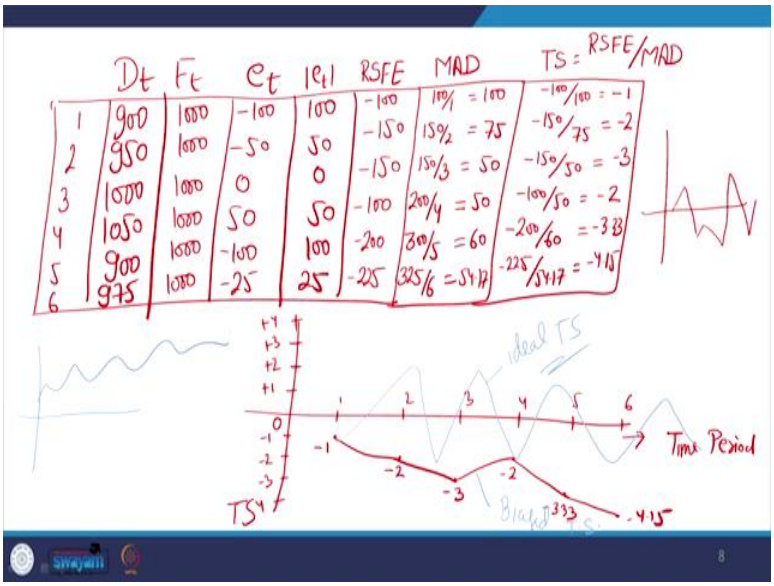


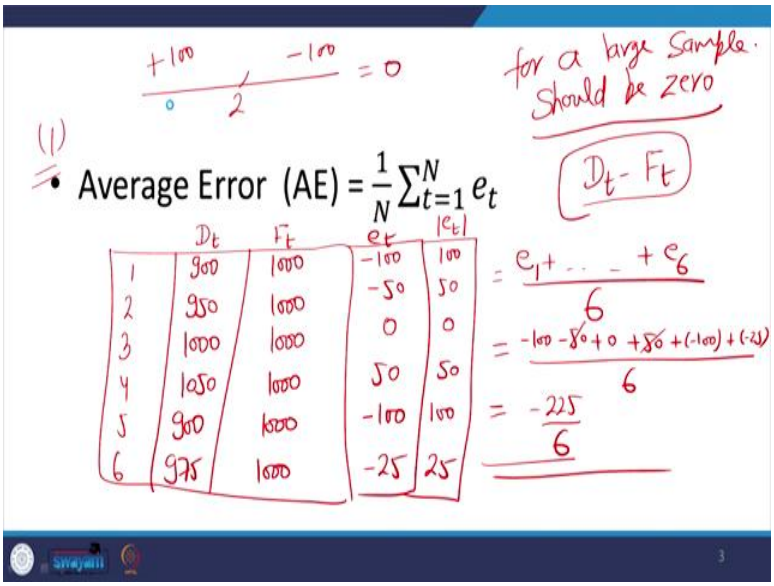
Now there is one more important thing in this continuation and that is known as tracking signal. Because we need some kind of mechanism, where just by visual display you can understand whether my forecasting whether my modelling is producing the appropriate results or not and that is very much possible with the help of the tracking signal.

Now what is this tracking signal? It is in fact a combination of two types of forecasting errors and what is this, these are RSFE which is running some of forecasting errors and MAD we have already discussed just now is, mean absolute deviation. So, these are the two terms which are going to be helping us in determining the value of tracking signal.

Now what does it say that it is a measurement that indicates whether the forecast average is keeping pace with any genuine upward or downward changes in the demand. So, whether my forecast is showing any genuine trend of upward or downward. We want our tracking signal should be fluctuating one.

(Refer Slide Time: 22:22)





I will show you what I want to say and for that purpose let us consider some data and with the help of that data we will prepare this tracking signal table, let us say we take the same data. These are the six periods, if let us quickly see 900, 950, 1000, 900, 950, 1000, 1050, 900, 975. These are the Dts and then we have the forecast for these periods.

So the forecasts are 1000s and the et will be DT minus Ft that is minus 100, minus 50, 0, 50, minus 100 and minus 25. Now you calculate the et that is 100, 50, 0, 50, 100 and 25. Now let us see the calculation of some of the things first is RSFE.

Running sum of forecasting error. So for the first period this is minus 100 then minus 100, minus 50, it becomes minus 150 minus 150 plus 0 , it becomes minus 150 minus 150 plus 50 it becomes minus 100 minus 100 minus 100, minus 200 minus 200 minus 25 it becomes minus 225. The second thing is the calculation of MAD. MAD is the calculation of average of summation of modulus errors.

So for first period, it is 100 by 1 becomes 100. For the second period 100 plus 50, 150 divided by 2 it becomes 75. For third period, it again remains 150 divided 3 it is 50. For the fourth period it becomes 200 divided by 4 it becomes 50. For the fifth period, 200 plus 100 it becomes 300 divided by 5 comes to be 60 and for the sixth period 300 plus 25, it is 325 by 6. So, it is coming to be 325 divided by 6. So it is coming 54.17.

Now we can calculate tracking signal which is RSFE divided by MAD. Now for the first period the RSFE is minus 100 and MAD is 100. So tracking signal becomes minus 1. For the second period it is minus 150 divided by 75 minus 2 minus 150 divided by 50, minus 3, minus 100 divided by 50 minus 2, minus 200 divided by 60. So, this comes to be 200 divided by 60 3.33 minus 3.3 3. Then minus 225 divided by 54.17. So this is 225 divided by 54.17 comes to be minus 4.5.

Now these are the values of tracking signal. Now we can plot this tracking signal on a piece of paper also and for that purpose you can take the values of tracking signals on y axis and on x axis you can take different time periods. This is the 0 value of tracking signal, this is positive values plus 1, plus 2, plus 3, plus 4 and here it is minus 1, minus 2, minus 3, minus 4 like that.

Now we have started plotting for period 1, period 2, period 3, period 4, period 5, period 6. Now for period 1 it is minus 1, for period 2 it is minus 2, for period 3 it is minus 3. Then it is minus 2 then it is minus 3.33 and then it is minus 4.15.

So, you join this, so by joining these dots you get this value of this shape of tracking signal which is all the time below the baseline. This is all the time in the negative zone. We want a tracking signal which should be fluctuating, sometimes it would be above the baseline and sometime it should be below the baseline and this type of fluctuating tracking signal where

randomness is there. Fluctuating tracking signal means the tracking signal showing the randomness. It is not biased.

Now this tracking signal which we got now it is all the time in negative. So, it is showing a kind of biasness. That you are continuously over stocking, you are doing more forecasts than required and therefore every time your forecast your tracking signal value is in negative. Though there are some fluctuations but those fluctuations are even from minus 3 to minus 2 and then again minus 3.33.

So, these fluctuations are happening, but these fluctuations are limited only to the negative part of your graph. We want fluctuation, so that sometime our fluctuations contain the curve to this type of phenomena, so that this is some kind of Ideal tracking signal and our tracking signal which we are using right now, this is a biased tracking signal. So, biased tracking signal is a good indicator.

So any type of trend if your tracking signal is continuously moving upwards in the positive that is also not desirable. It means you are continuously underestimating, that means you are a very, very you can say conservative type of person and you do not want to stock enough. So, you are not taking any rest and this may bring down the service level of your organization.

So, if tracking signal is only in positive that is also not desirable. If it is only negative it is also not desirable. If it is only in negative that is also not desirable. So, it is required that it should fluctuate around this base value and the fluctuations are there around base value. We can say that our model is not having any kind of bias. It is free from bias and that is the important requirement of a good forecasting model.

If the model is biased, if the model is biased then you will be either continuously over stocking or continuously understocking. So we require, you can say two very important things, one we need to have plot of MAD that will give us the suitability of the model and we also need to have tracking signal form our model which tells us whether the model is Biased or not biased.

So, by combining these different measures of forecasting errors, we can find out the suitability of a particular forecasting model for our historical model data. With this we come to end of this discussion of forecasting errors. Where we discussed different types of measures of forecasting

errors and we also discussed tracking signal and how do we use this information for identifying a good forecasting model for our purpose. Thank you very much.