The following minutes to cut

00:00-12:18

12:48-13:20

22:20-22:59

24:19-26:00

26:39-26:52

28:34-29:05

29:23-30:21

31:09-31:20

36:15-37:02

45:30-45:57

46:45-46-56

50:50-50:51

51:58-52:02

52:25-52:30

52:47-54:28

So exponential smoothing, what is exponential smoothing? Exponential smoothing is a moving average, a wider moving average. What is moving average?

Say we have data, 100 period, 200 period, something.

We have 100, 200, 1000 numbers or periods. Someone ask us "What is your forecast for the next period?" I say "I will add up all the numbers I have and average it, that is my forecast for the next period. So I have been very conservative here, I try to satisfy everybody and find if all the demand are correct what's the average for the next period's demand?

That is one extreme. On the other extreme we just may take the last and say "I only rely on the newest piece of data and my forecast for the next period'

Moving average is a mediation, a sort of compromise between these 2 extreme points. Instead of relying on the last period, I rely on the last two periods. Instead of two periods I rely on three. That's three period moving average.

The larger the number of periods the smoother the data and the less the impact of the last period. This is 100%.

Do not forget, we don't want you to memorize this. When you go to class I may throw some data to you, some facts. We may transform those data into information. Then we make a little more play with those information and transfer those information to knowledge. We may play with that knowledge and transfer our understanding.

We may play with that then. The more prepared you are in class, the more opportunity we have to go from memorizing things to applying them. That is the whole idea.

So in an end period moving average I try to find an appropriate place in the continuum between relying on all pieces of data and relying on the last piece.

My measure of effectiveness is MAD. Having said that, now I can think about something else. I may say my forecast for the next period, which is period T+1 is equal to my forecast for the previous period plus a fraction of the difference between what my forecast was and what I observe. [writing on board]

So this is what my forecast was [On overhead.] This is what I observed in reality [On overhead.]

If I'm going to take all of that difference, 100% of that difference, what will I have? I will have Ft+(At−Ft)

What is left is At. If I set alpha to 1 period moving average .. if I put *α* =0then 0 is multiplied by this. No matter what the reality is I won't make any change. My forecast for this period is exactly what it was for this period.

I am moving from somewhere which 100% relies on last period, and something which does not rely at all on the last period. The question is where is the best place to stay? How much should I give attention to the demand of the last period and it's derivation from my forecast. So *α* is defined as a variable between 0 and 1. If it's 1 my only focus is on the last period. If it's 0 I have no respect for the last period. I don't incorporate any new knowledge into my forecast.

If it is 0.5 that means last period my forecast was 100, the actual demand came out 120, which is 100 units more, there's a 20 unit difference. I don't take 100% of that and incorporate it. I take 50%. Okay? That is how exponential smoothing works.

Therefore exponential smoothing forecast for the next period equals ... [On overhead.] Any questions?

I can simplify this equation, I can multiply *α* for Atand I can also multiply this [On overhead.] Then I can put it together. [writing on board.]

This is another formula for exponential smoothing.

Exponential smoothing can be forecast as [writing on board]

When alpha goes up, this part goes up, this part goes down. This means I pay more attention to recent periods. If alpha goes down I pay more attention to this. As I explained my lecture, the exponential smoothing is a moving average where all pieces of data are taken into account. However it's weighted. The weight of all elements are low? Okay.

[On overhead.]

Again we put numbers into formula. [ reading from the overhead ] Any questions?

Okay. [Teacher reading: Question 3]

We know the actual for this period is ... 16000. So I can put this actual and forecast into formula ... sorry, what? This is our forecast for period 4, 16000. *α* =0.4.

So I have the forecast for period 4, forecast is given, alpha equals 0.4, I put it into the formula and I get 17600Okay?

Now I have actual and forecast for period 5, then I compute the forecast for the next year. [On overhead.]

Professor: Again, same formula. [Teacher reading: Question 4.] Question put this into the formula to compute.

Professor: The previous forecast was 66, 5 units larger than the actual demand. So this is exponential smoothing formula. The previous forecast was 66, which was 5 units larger than the actual demand. This was 66 and it was 5 units greater than ... 5 units larger than actual demand. So actual - forecast = -5right?

Professor: The previous forecast of 66 turned out to be 5 units larger. [ reading from the overhead ]

Our forecast for next period is 65. Therefore we do 65=66+-5

*α* =0.2

Any questions? Any questions?

Female Student: Do we always keep it like that? With the negative?

Professor: It can be positive, it could be negative. What it says here, the previous forecast ... is Ft. The previous actual is At

Previous forecast was 66, which was 5 units larger than actual. If that is difficult you can assume this ... forecast was 5 units larger than actual.

If forecast was 5 units larger than actual, 66. Therefore forecast was that.

So we use those 5 units to compute 61, but we don't need to do that.

If we subtract them, this is 5 units more than this, so a negative sign. Right? Then we have one equation, one unknown. *α* =0.2

[Teacher reading: Question 6]

Forecast error is A−F. So previous forecast is F. So a percentage of A−F and define it as exponential smoothing. Exponential smoothing is new forecast equal to old forecast + alpha times actual.

[Teacher reading: Question 6]

Is which? It is D. Any objection?? Next question?

When you use large numbers you give more values to recent period observation. If you give small that means you give less value to the most recent observation and more value to all before this, right?

Exponential smoothing, we want one period moving average. [ reading from the overhead ] that is what we want. Therefore what we want is this to become 0 and this to become 1. If I set *α* =1 then 1− *α* =0. Right? Any questions?

What is the answer is that? Question 9.

This is a straight line, no matter if forecast is 0 or 7000, I stay as I was before. [ reading from the overhead ] This part becomes 0. [On overhead.]

Okay?

Now we have 7 periods of data and we want to use 5 period moving average, exponential smoothing and compute MAD and compare. Usually when we do forecasting, that's what you will do in the game is you draw the data to see if you observe a pattern. Right? Then we try to find via an appropriate method, which minimizes this.

This is our data [On overhead.] We don't see any specific pattern. 7 pieces of data. We use 5 period moving average. We will not be able to compute 5 period moving average until period 5.

SOM 306

Monday, February 17, 2014

Typewell; Dave Strong

Sirob Malkonian, Pamela Torres

**Last 35 minutes of Exponential Smoothing**

Now if I use 5 period moving average to calculate the next forecast I do this [ reading from the board ] Right?

If I'm going to forecast for whatever I can using 5 period moving average, the first time I can compute 5 period moving average is 5. [On overhead.] It's assumed as forecast for period 6. 5 period moving average of 7 forecast for 8.

In 5 period moving average I only have two pieces of data to help me calculate MAD. In no other place can I compare actual with forecast.

In this period I have my forecast. Here I have forecast and actual. Here I have forecast and actual. Here I have forecast but no actual.

2 period isn't enough for a reliable MAD. But that is all forecast using 5 period moving average.

Now we go for exponential smoothing. [ reading from the overhead ]

We have just taken actual of February and assumed it as forecast for March. In exponential smoothing, for the first forecast we can use several techniques. In real life we have 50 periods of data. If you use exponential smoothing one way is to assume forecast for 2 is the period 1.

No one can criticize you while you use this one. Personally, when I use exponential smoothing I compute average of all data and then build up. For first exponential smoothing forecast we need something. As I said in my lecture you may assume actual of period 1 as forecast for period 2.

Or you may just put numbers you think are more reliable for forecast for period 1 or 2. Usually we use actual of period 1 as forecast of period 2. Which is what we do here [On overhead.]

[ reading from the overhead ]

Male Student: Any reason for moving average of 1 and 2?

Professor: There is no logic to do that. But if you do no one can criticize you. You may use all, or 1 and 2. But that's not common. What is common is 1 as 2 or average of everything as 1.

Here we are in period 2, actual is 18. Forecast is 19. [ reading from the overhead ]

This gives us 18.8.As soon as I have forecast for period 3 I can compute for all.

So here we have exponential smoothing [On overhead.] I was able to have forecast and actual for 6 periods. The same is for naive technique [On overhead.]

[ reading from the overhead ]

We cannot take one for 3 periods, the next for 4th period. They must have same number of periods and same periods.

Here we have actual, naive method, exponential smoothing. [On overhead.]

This is the difference between actual and forecast. [ reading from the overhead ]

The method with lower MAD is preferred. In this case we stay with exponential smoothing.

What is the answer to this question? [On overhead.]

What about this one? [On overhead.] True or false?

Male Student: False.

Professor: False, because when alpha is large we give more value to the most recent data. Any questions? Age of data in exponential smoothing is 1/α

What is age of data in moving average?

Professor: The last piece of data is one period old in moving age. The oldest piece is *n* period old.

You have A and D, what is your average? 1+4÷2

In moving average the newest data is one period old, the oldest is *n* period old.

Why we can add these together and divide by 2? Because the change is uniform. The first period is 1 period old, the second is 2. The last period is 1 period. The oldest is *n* period old. If changes like this - for example if the period before was 8 period old, then the last period was *n* period old we can't add the last and first together.

When changes are smooth we go from large to small in a smooth way. Otherwise we add them all together and divide by the number.

For moving average the last piece of data is 1 period old, then 2, then 3 etc.

If you want to find out what exponential smoothing performs like in moving average it is easy. *n* +1÷2is age of data in moving average. It equals 1− *α*

[On overhead.]

Therefore *α* =0.5 this is like this [On overhead.] We know moving average like that has age of data 1/α=1/0.5=2

So here age of data is 2 period.

N+1/2=2

N+1=4

N=3

The conclusion is this, exponential smoothing with *α* =0.5 is somehow similar to a moving average with N=3

If *α* =0.2, age of data in this exponential smoothing ... is 1/α=1/0.2=5. So age of data is 5 periods old.

To see this exponential smoothing equivalent with what moving average? [writing on board.]

N+1/2=5N=10−1=9

So that's age of data in exponential smoothing and moving average.

Age of data in exponential smoothing follows a complicated mathematical equation, but we just accept the results.

We combine this to get N+1/2

Where the average age of the data in both techniques.

What is the advantage of 0.5? What makes 0.5 a specific number in exponential smoothing?

Female Student: You give the same weight.

Professor: What is the meaning of adding them by 0.5? What is the meaning of multiplying 10 by 0.5 and adding it to 20 multiplied by 0.5?

Female Student: Average.

Professor: Average. When it is 0.5 I can take actual and forecast, add them together and divide by 2.

[ reading from the overhead ]

So I can just add them together for the summation. Right? So 0.5 times this + 0.5 times this ... [ reading from the overhead ]

What is the forecast for period 4? Here's 300 for period 1. I put as forecast for period 2.

[ reading from the overhead ]

Therefore my forecast for the next period is what? Any questions? Any questions?

Answer this question quickly. [On overhead.]

[Students working.]

What is the answer? We are talking about 6 period moving average. I am period 100, I want forecast for 101. I want to use 6 period moving average. The most recent piece of data is 100. The oldest piece of data is 6.

[ reading from the overhead ]

In a 6 period moving average, average age of data is 3.5 periods old.

[ reading from the overhead ]

So *α* =0.29is equivalent to 6 periods moving average. They are the same, right?

Okay.

We talked about exponential smoothing and moving average. Finding if 2 period is better, 5 period moving average is better, 7 period moving average is better. Our measure was average of summation of the absolute value of the difference between actual and forecast.

So take actual, take forecast, find the difference. Drop the negative sign, add them together, divide by number of observation. I select that number of periods which leads to minimal MAD.

So they find the square. The square is another way to drop a negative sign. -5squared becomes 25. If it's 5and you square it, it's 25.

So if forecast is more than actual or actual is more than forecast it doesn't matter, both are bad.

In different cases we assign different weights to positive or negative, but we don't go there in this course. In this course we assume overestimate is as bad as underestimate.

In regression they use square instead of absolute value, why? Mathematically dealing with squares is easier than absolute value.

So one reason for going there was - the main reason for replacing absolute value with squares was ease of mathematical manipulation.

But we know if we square then the method tries to avoid large differences between actual and forecast. If the difference is 10, when we square it it becomes 100. So one other feature of square is it gives more weight to larger.

In general, list squared tries to draw a line across a set of dots, such that a square of the difference of dots and line is minimized. We will go through this on Wednesday.

Thank you.

[End of class.]

Dr. Asef these are the minutes you need to cut out from the lecture:

* 45:30-45:57 minutes.
* 46:45-46-56 minutes
* 50:50-50:51 minutes
* 51:58-52:02 minutes
* 52:25-52:30 minutes
* 52:47-54:28 minutes