Management and Operations 340: Exponential Smoothing Forecasting Methods

[Chuck Munson]: Hello, this is Chuck Munson. In this clip today we're going to talk about forecasting, in particular exponential smoothing models. The book does a nice job in chapter four of covering a good introduction on forecasting methods. Forecasting is extremely important because most of the rest of the book, the decisions, such as how much inventory to have or how to do scheduling or planning, are all based on demand forecasts, so you might use the right inventory model but if you're basing it on bad forecast you're going to make bad decisions, so this is really the introductory- one of the first things you do as an operation, as a manager before you start making these decisions, so two of the more popular and more important models that are in that chapter are exponential smoothing models. We'll first look at single exponential smoothing and then include a trend component, which is double exponential smoothing.

[On Screen]

Exponential Smoothing Method

Form of weighted moving average

- Weights decline exponentially
- Most recent data weighted most

Requires smoothing constant (α)

- Ranges from 0 to 1
- Subjectively chosen

Involves little record keeping of old data

[Chuck Munson]: So let's get started with single smoothing, and the idea is that this is actually a form of weighted moving average that has a special construction of the weights. If we draw a little chart here and the x-axis is just the age of the date, so recent over here, old over here, and the y-axis, the weights of the- applied to each whole demand. If we draw those out it goes in an exponentially decreasing function, so that's why they call it exponential smoothing, so what it means is all the old data that you've ever had is still partially captured in your new forecast. The weight may be very, very small, but it's always in there unlike, say, a 3 period moving average forecast where that 4th period and older are all gone, you never use that information again. So it requires what's called a smoothing constant, which we call alpha, and this ranges from 0 to 1. Typically alpha is .1, .2, or .3, although it could be anything you want. The higher the alpha the more weight you're putting on recent data, so if current trends are more indicative of the future, you may want to have a high alpha. If current trends may not be indicative then a lower alpha puts more weight on that older data. This is subjectively chosen, there's no magical formula for what alpha to use, but once you've decided to use it the map is pretty easy-going forward. One nice advantage of exponential smoothing is that it involves little record keeping of old data, so while all the old data is in the forecast because it is a recursive formula, you only need to know what happened last period. So there's very little record keeping involved.

[On Screen]

Exponential Smoothing Equation

F_t = forecast value for period *t*

A_t = demand at period *t*

α = smoothing constant ($0 \leq \alpha \leq 1$)

Then the forecast for period *t* is:

$F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1})$

Thus, you only need to look at this period's forecast and actual values to compute the forecast for the next period.

[Chuck Munson]: Okay, so let's look at that formula next. We'll let F_t be the forecast value for period *t*. A_t is the actual demand in that period, and then αis our smoothing constant. The formula is presented here and it simple means this: F_t is this period's forecast. F_{t-1} then, $_{t-1,}$ is one period before this period, so that's last period's forecast. A, remember, is actual demand, so A_{t-1} must be what happened last period, so that's last period's actual, and so inside the parenthesis here 18 minus 1 minus F_{t-1}, that must be the error that we had last period, so error in last period's forecast. So what is the formula doing? It's basing the next forecast on the last one you made and it's adjusting it up or down. If last period your forecast wasn't high enough and your demand was higher than you thought it would be, then this number in parenthesis is positive and you're going to use αpercent higher in that direction next time. If the forecast was too high, then the number in the parenthesis is negative, and your forecast will be lower than it was last time.

So let me show you a quick example of that, and, again, what's nice about the method is all you need to know to make this period's forecast is what happened last period. So let's draw a little graph here. And in our example, we'll say αis 10% and suppose last period we thought demand would be 100, so F_{t-1} is 100. Turns out it was actually higher, 120, that's usually good news, that means more sales, however, it wasn't a very good forecast. We were 20 units too short, so next period we're going to go 10% of that distance in that direction, so if we were 20 too short, 10% of 20 is 2, so next period F_t is going to 102. And it just goes that way period after period, and you keep adjusting your forecast up or down by some percentage of the error that you just had, okay?

[On Screen]

Simple (Single) Exponential Smoothing

Level only; no trend of seasonality

<u>Update Equation</u>

$F_1=F_{t-1} + \alpha(A_{t-1} - F_{t-1})$

<u>Example</u>

The initial forecast (F_1) is assumed to be 1000 units. Demand in period 1 was 900 units and in period 2 was 1200 units. Using simple exponential smoothing with α = 0.3, what is the forecast for period 3?

F_1= 1000, A_1 = 900, A_2 = 1200, α = 0.3

F_2 = 1000 + 0.3(900-1000) = 1000 + (-30) = 970

F_3 = 970 + 0.3(1200-970) = 970 +69 = 1039

[Chuck Munson]: So let's look at an example next. So this is single smoothing. We don't have trend or seasonality. Here once again is the equation, and here is our little example. The initial forecast F_1 is assumed to be 1000 units. You have to initiate- initialize this method in some way. You have to have an initial forecast. It could be your demand or maybe you had a different way to make a forecast back then. You have to start it and then go forward in time, and then once you've caught up to now, going forward is just one equation each period.

So we'll say in period 1 we thought demand would be 1000 units. It actually was 900 and the next period was 1200. We're going to use single exponential smoothing with an αvalue of .3; what is the forecast for period 3? So here I've got what we know: forecast we made in period 1, the two actual demands, and the αthat we're using, so for the forecast for the 2^{nd} period you would take the forecast for the one before, so that's 1000 right there. And then we're going to adjust that by α times the error last period, so you take the actual from last period, that 900, and subtract the forecast. Well we were off by 100 units, so we thought it would be 100 more than it actually was, so 30% of 100 is 30 and because we were higher than actual, it'll be a negative number, so our new forecast is going to be 30 less than the old one, and that's 970 units for period 2. And period 3 is the same idea: you take the forecast from period 2, that's the one we just made, and then adjust it by 30% of the error. The 1200 is the actual in period 2. The 970, again, is the forecast that we had made for period 2. This time we weren't high enough. Demand was higher than we thought it would be. So 30% of that difference is 69, and our final forecast then for period 3 is 1,039. And because this method assumes no trend, no seasonality, without any more information, we're just going to assume it's 1,039 for the 4^{th} period, 1,039 for the 5^{th} period, etc. because, again, we assume there is no trend, no growth going on or reduction, okay? So that's how you do single smoothing.

Before we look at double smoothing I wanted to explore briefly the effect of αbecause it can make a big difference in the forecast values that you have.

[On Screen]

Forecasting Effect of the Smoothing Constant α

The exponential smoothing formula can also be written as:

$F_1 = \alpha A_{t-1} + \alpha(1-\alpha)A_{t-2} + \alpha(1-\alpha)^2 A_{t-3} + \ldots$

Weight of Old Demand

If α is…	Prior Period α	2 Periods ago α(1- α)	3 Periods ago $\alpha(1-\alpha)^2$
10%			
50%			
90%			

[Chuck Munson]: If you do some algebra, you can rewrite the equation this way, and if you stare at it long enough, you'll see this is why I said this is a form of weighted moving average because we're adding in each case an old demand times some number, and the number keeps gets smaller because α is a fraction, 1- α is a fraction, so α(1- α) is smaller than α, and $\alpha(1-\alpha)^2$ is smaller than α(1- α), etc. So these are the weights, and they keep getting smaller and smaller, and if you plotted them on a graph, you would get the exact exponential function that I showed you at the beginning. So that's what's happening with α, but the size you choose will affect the results quite a bit. If we use a 10% α, that means we're applying 10% of our forecast based on what just happened in the previous period, and then 90% of that is 9%, so that's how much we apply to the period before that. 90% of that is 8.1%, and so on. So a 10% α actually has a lot of weight going quite a ways back in time. If you use a much higher α, say 50%, half your forecast is based on what just happened, and another 25% is based on what happened 2 periods ago, another 12.5% 3 periods ago. And if you go to an extreme, 90% α means 90% of your forecast is based on what just happened and another 9 on the period before that, so 99% of your forecast is based on what happened in just the last 2 periods. You will only do this if you think there's a big change in demand and there's no longer a reason to look at very old demand. It's another 0.9% after that, okay?

So as I said there's no magic formula for α. A good thing you can do is put these formulas into an excel spreadsheet and then try it on old data with different α's, and if one of the α's gives you a less- a lower forecast error, that could be a good α to use for the future. But just because it worked in the past doesn't mean it will be good for the future, and in particular, pay attention to what's going on. If there is a lot of recent change in data, maybe open up a new market or your products getting more popular or less popular, then you probably want to either start over in time or try a higher α so it puts more weight on that recent data.

On the other hand, a big demand that just happened can really mess up your forecast if it's an unusual event. So an example might be, say, a McDonald's restaurant. When I was in high school and we played basketball we would travel two to four hours away for some high school games, so we would drive into McDonalds at one in the morning with a bus load of hungry basketball players and cheerleaders, and, you know, if they weren't ready for us, this was an enormous amount of business at one o' clock in the morning, so that's why we called ahead, but the point is that was unusual. We weren't going to come back there for another year, and so if they think what happened last night is going to happen again tonight, they're probably going to be way off, so that's an example of why you don't want a very big α because a random event might really mess you up in terms of your forecast for the future.

[On Screen]

Double Exponential Smoothing (Exponential Smoothing with Trend Adjustment)

Level and trend; no seasonality

<u>New Notation</u>

F_t = smoothed forecast average for period *t*

T_t = smoothed trend for period *t*

B = smoothing constant for the trend ($0 \leq \beta \leq 1$)

FIT_t = forecast for period *t* (including trend)

Forecast Equations

$FIT_t = F_t + T_t$

$FIT_{t+n} = FIT_t + nT_t$, for *n* periods in the future

Update Equations

$F_t = \alpha A_{t-1} + (1- \alpha)(F_{t-1} + T_{t-1})$

$T_t = \beta(F_t - F_{t-1}) + (1-\beta)T_{t-1},$

Okay, let's next talk about double exponential smoothing, and this includes trend. The problem- if you have growth in your demand, it's growing or its going down over time, if you use single smoothing, you'll always lag behind that, and the lower your α value the more you're going to lag, and so that's problematic because many products do see a trend either up or down in sales, and so double smoothing tries to capture that trend up front, so you're estimating not only the base level of demand but the growth, and so when you make a forecast you include both of those pieces in there, so we call that double smoothing, this is also known as Holt's Method. Your book calls it exponential smoothing with trend adjustment. Our notation changes a little bit: we have F_t again, but now it's what they call a smoothed forecast average for period *t*, so this is kind of like a base level forecast not including growth. T_t is the growth, the trend, and this can be positive or negative. If it's negative that means you expect demand to go down. If it's positive, you expect demand to go up. Since it's double smoothing we have another Greek letter here: this is β. βis also between 0 and 1 and it's the same idea. You choose a β that you think is right and it's meant to put some weight on the old trend estimate and some weight on what just happened. Your actual forecast then for period *t*, we call FIT forecast including the trend, and that's an automate thing we're looking for, so don't be confused. F_t and single smoothing is the forecast and double smoothing, that's only part of the forecast. So forecast and double smoothing is the F_t plus the trend term, so it's this kind of base level plus or minus the growth that you think you'll see, and if you expect to see growth, of course, and you expect to see that in the future, so now if we want to know the forecast not only for this period but the period after that and the period after that, you would simply multiply that trend term by however many in the period's in the future you want to go.

So here are our equations, and now we have two equations for double smoothing: F_t here is the- looks similar to what it was in single smoothing but we've done some algebra on that and included this trend term, so now it looks like this. We're taking α times the amount of the old demand that just happened plus (1- α) times everything else. Everything else is the forecast that we had based average forecast the last period plus the trend that we thought would happen, so

(F_{t-1} + T_{t-1}) equals FIT_{t-1}, so what is this? It's α times what just happened plus (1- α) times what we thought would just happen, okay? For the- after you do that, you update the trend term, and notice the trend term uses what you just computed, so you take the new F_t minus the old F_t; that difference is a trend, up or down, you multiple that by β. The rest, (1-β), is multiplied by the last estimate you had for growth.

Anyway, those are the equations, and let's look at an example here.

[On Screen]

Example

The initial forecast average (F_1) is assumed to be 1000 units, and the initial trend (T_1) is assumed to be 50 units. Demand in period 1 was 900 units and in period 2 was 1200 units. Using double exponential smoothing with α = 0.3 and β = 0.2 units, what is the forecast for period 3?

$F_1 = 1000, T_1 = 50, A_1 = 900, A_2 = 1200$

$F_2 = 0.3(900) + (1\text{-}0.3)(1000 + 50)= 1005$

$T_2 = 0.2(1005 – 1000) + (1\text{-} 0.2)(50) = 41$

$F_3 = (0.3(1200) + (1 – 0.3)(1005 + 41)\ 1092$

$T_3 = 0.2(1092 – 1005) + (1 – 0.2)(41) = 50$

$FIT_3 = 1092 + 50 = 1142$

$FIT_4 = 1142 + 50 = 1192$

$FIT_5 + 1142 + 2(50) = 1242$

I've typed it all out for us. Similar numbers are the one we just had for single smoothing. We're going to assume an initial 1000 unit forecast. One thing we're adding here is that we assume there's growth going on and our best estimate at that time was that it was growing 50 units per period. Again 900 units in period 1 and 1200 in period 2. In addition to our alpha we've got a β term for the trend, and we are using .2 for that. Question is, what is the forecast for period 3? So this first line is what we know: the initial forecast, the initial forecast of trend, forecast average, forecast trend, and then the two actuals that we know. So going back to the previous page, F_2 then is α times the demand in the previous period, the 900, plus (1- α) times the F of the previous period, which is 1000, plus the T of the previous period, which is 50, so the forecast we actually had made for period 1 was 1,050. We take 70% of that plus 30% of what actually happened, which was 900. That equals 1,005. Now for the trend estimate for period 2 you take β times the forecast you just made, so that's 1,005, minus the forecast you had made the previous period, which was 1,000, plus (1-β) times the last trend forecast that you had made, which was 50. So the new T, the $T_{2,}$ is 41. So what is our actual forecast for period 2? That's FIT_2, which is 1,005 + 41 or 1,046, okay? So you do the same kind of thing for period 3; you would take the demand, actual demand of 1200 for the previous period, that comes from way up there, plus (1- α) times the F_2, which was our 1,005, and the 41, which was that trend that we just calculated. That's

1,092. For T_3 we start with that, subtract the 1,005, which was $F_{2,}$ and then (1-β) times this T_2. So if you can see where all those arrows are going, that's where everything's coming from. So F_3 is 1,092, T_3 is 50, so the actual forecast then for period 3 is 1,092 + 50 or 1,142, so that's the answer to the question. I could look even forward from there if I want to estimate demand for period 4. The best guess is the last thing that I made, which was 1,142 for the previous period plus the growth that we expected to have of 50, so that's 1,192. The period after that would be 50 more, etc., etc. And that's a quick look at exponential smoothing.

There is something called triple exponential smoothing, which includes a third Greek letter, gamma, and with that, that includes seasonality factor like winter, spring, summer, fall. I'll leave that to you to study on your own, but hopefully this helps clear things up. Thank you.