This introduction will cover basic forecasting methods, how to set the parameters of those methods, and how to measure forecast accuracy.

We will use the following terminology:

\( F_t \) Forecast of demand in period \( t \).

\( A_t \) Actual demand in period \( t \).

1 Forecast Error

The way we will measure how well a forecast performs is by using measures of forecast error that you’ve probably seen before. Define \( E_t \) as the difference between the forecast and the actual:

\[
E_t = F_t - A_t.
\]

As you can see, it is very easy to use. All you need to decide ahead of time is what value to use for \( \alpha \), and you just need to know what the most recent demand was and what the most recent forecast was, and that’s all you need to make a new forecast.

2 What is Exponential Smoothing?

The basic idea in exponential smoothing is that we take an average of our old estimate of some quantity, and some new information about that quantity. In exponential smoothing, we are assuming that there is no growth, no trend to the data. So every period, we are just making new estimates of the intercept.

\[
F_t = \alpha \cdot D_{t-1} + (1 - \alpha) \cdot F_{t-1}.
\]

The new demand gives us another data point about what the intercept might be, and the old forecast is our old estimate of the intercept.

We can (and we are going to) use this idea to update estimates of other things, like a trend. Although the formula will look different, the idea is the same:
New Estimate = α \cdot \text{New information} + (1 - α) \cdot \text{Old Estimate}.

3 Forecasting with a Trend

When our demand has a trend, there are two main methods that we can use.

3.1 Linear Regression

I assume that you are all familiar with linear regression from your statistics classes. Basically, we assume that there is a linear relationship between one output (dependent) variable, $Y$, and the input (independent) variable, $X$. In our case, we will be looking at the independent variable as being time, $t$, and we think that demand is generally growing over time. We do a linear regression to get a formula like this:

$$Y(t) = a + bt.$$  

For any time value, $t$, we put it into the equation, and get a straight-line forecast of the demand for that period.

How well the data are approximated by the line is represented in the term $R^2$. $R^2$ can be literally interpreted as the "percentage of changes in $Y$ that can be explained by changes in $X$.

To do a linear regression in Excel, there are four ways you could do, presented in the order of the things I like the least to the way I think is the best.

First, you could dust off your statistics book and type in the formulas from it. That sounds like a lot of work, and there are lots of opportunities to make a mistake when typing in those big formulas. But if you get it all in correctly, the spreadsheet will update automatically when you enter new data points.

Secondly, could also use the “Data Analysis ToolPak.” You should find that under “Tools — Data Analysis.” If it is not there, go to “Tools — Add-Ins.” In that dialog box, check the box by Analysis ToolPak. If that does not appear in the dialog box, you need to get out your CDs and install that part of Excel. After you go to “Tools — Data Analysis,” a dialog box comes up, where you tell it which cells are the X’s, and which are the Y’s. Tell it where to put the output, but be careful. If you put it on the sheet you are working on, the following 18 rows will get written over with the output.

In that output, you will want to look at where the “Intercept” row and the “Coefficient” column intersect. That is the intercept. The slope is the row below that, in the “X Variable” column, where it meets the “Coefficient” column. R-squared is in the second row of numbers, under “R Square.”

One problem with this method is that when you add new data point to the spreadsheet, you have to go back up to “Tools — Data Analysis” every time to re-run the LR. The spreadsheet can’t update automatically.

Thirdly, another way to get the intercept and slope is to create a graph of the data. Right click on the data line, and select “Add Trendline.” In that dialog box, add a linear trendline, and under “Options,” you can have the equation of the trendline displayed on the graph, and also R-squared. The trouble with doing the LR in the graph, is that you can’t make use of the numbers that appear in the graph in any calculations.
Finally, the best way to do the LR is to use the SLOPE and INTERCEPT functions. SLOPE(range of x value, range of y values) gives you the slope, and INTERCEPT(x values, y values) gives the intercept. To find out R-Squared, use RSQ(x values, y values).

3.2 Exponential Smoothing with a Trend

The only problem with Linear Regression is that it gives all the demand points equal weight when trying to fit a line. Really, we would like it to try hardest to fit the line to the most recent data points, and not worry quite so much about fitting the line to the oldest data points. Linear regression cannot do that.

However, we do remember that exponential smoothing had that type of behavior: give the most weight to the most recent. There is a way we can adapt exponential smoothing to work with a trend. (If you really want to be fancy, this is also known as Double Exponential Smoothing, or Holt’s Method.) We will define two terms:

- \( S_t \) our estimate of the level at time \( t \) (kind of like an intercept, but not exactly).
- \( T_t \) our estimate of the trend in period \( t \), or the slope.

\[ TAF_{t+1} = S_t + T_t. \]

our estimate of what the demand will actually be in period \( t \).

In each period, we will revise our estimate of the level, and our estimate of the trend. To do that, we will use two different smoothing constants, \( \alpha \), and \( \beta \). Like \( \alpha \), \( \beta \) is a number between 0.1 and 0.3, usually.

1. At the end of a period, compute a new intercept:

\[ S_t = TAF_t + \alpha (A_t - TAF_t). \]

We can also write this as:

\[ S_t = (1 - \alpha) * TAF_t + \alpha A_t \]

2. Compute a new, smoothed estimate of the trend

\[ T_t = T_{t-1} + \beta (TAF_t - TAF_{t-1} - T_{t-1}). \]

3. Use these two new terms to predict the demand for period \( t + 1 \):

\[ TAF_{t+1} = S_t + T_t. \]

If you want to make a forecast for the period that is \( n \) periods into the future, just add on \( n \) periods worth of growth:

\[ TAF_{t+n} = S_t + T_t * n. \]

3.3 Is this Right?

When I first saw the equation for the trend, I thought “why doesn’t it look more like the equation for the intercept?” The equation for the level, (the first way I wrote it), looks like the previous estimate \( TAF_t \) plus \( \alpha \) times the amount of the error.
But look at our formula for revising the trend, $T_t$.

\[ T_t = T_{t-1} + \beta (TAF_t - TAF_{t-1} - T_{t-1}). \]

We have the previous estimate of the trend, $T_{t-1}$, plus $\beta$, (instead of $\alpha$) multiplied by something that could be thought of as being an error term on our estimate of the trend. How? Well, $TAF_t - TAF_{t-1}$ tells us how much the forecast has gone up from the last period to the current one, which is very much a real-time snapshot of how quickly our forecasts are increasing. Last period, we thought things were growing by $T_t$, but now it turns out they are growing by $TAF_t - TAF_{t-1}$, so we can think of $(TAF_t - TAF_{t-1} - T_{t-1})$ as being the difference between how much things are actually growing and how much we used to think things were growing.

So even though the equation for the trend doesn’t look that much like the equation for the trend, the two are very much using the same spirit.