



EconS 450 Advanced Farm Management

Forecasting – Lecture 2



Simple Exponential Smoothing

Exponential smoothing is, like the moving average forecast, a simple and often used forecasting technique

Exponential Smoothing models continually adjust according to the *signs of the forecast errors*.



Simple Exponential Smoothing

A simple exponential smoothing model adjusts forecasts according to the sign of the forecast error:

- if $(X_t - F_t)$ is positive, the forecast is increased
- if $(X_t - F_t)$ is negative, the forecast is decreased



Exponential Smoothing Formula

An exponential smoothing forecast is calculated as:

$$F_{t+1} = F_t + \alpha e_t$$

where α is the smoothing parameter, and $0 < \alpha < 1$



Exponential Smoothing Formula

In words, with exponential smoothing, the new forecast is equal to the old forecast plus a fraction of the forecast error.

The magnitude of the adjustment is dependent on the value of α .



Exponential Smoothing Formula

The question is, how do you choose α ?

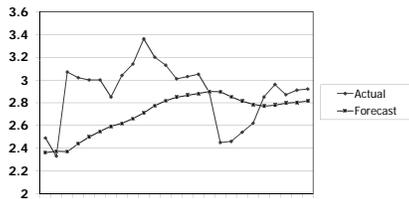
- in practice, you will *never* have an α less than 0.1
- If $\alpha=1$, you have the *naive model*
- a smaller α means that the forecast will not change as much

Choosing a Value for α

If the mean of a time series is not changing, and most of the fluctuations are random, a low α value is appropriate.

Choosing a Value for α

This graph illustrates a low α value, $\alpha = 0.1$:



Exponential Smoothing Example

Assuming an $\alpha = 0.1$ and the following data:

Price	Forecast	Error
28	30	-2
27		
33		
25		

Exponential Smoothing Example

The forecast for period 2 is:

$$F_2 = 30 + 0.1 \times (-2) = 29.8$$

The forecast for period 3 would be:

$$F_3 = 29.8 + 0.1 \times (-2.8) = 29.52$$

Where Did F_1 Come From?

Clearly, exponential smoothing forecasts are dependent upon the “first” forecast.

If F_1 is not representative of the data, all subsequent forecasts will be distorted.

Selecting F_1

a good rule of thumb is to set F_1 equal to the mean of the warm-up sample





Choosing α

In practice, choosing α can be done by trial and error. Follow these steps:

1. calculate forecasts for 3 different α values and calculate the MSE over the warm up sample for each



Choosing α

2. Pick an α between those with the lowest 2 MSE's from step 1). Calculate the forecasts and MSE for this α .
3. The MSE from step 2 will allow you to narrow your search for the best α .



Simple Exponential Smoothing

The simple exponential smoothing model is quite robust, it gives *good performance* on many different kinds of time-series.

It is, however *slow to recognize shifts in the mean* of the time series.



Simple Exponential Smoothing

If the *mean* of your data *increases* suddenly, the forecasts will underestimate the actual price.

The same thing will happen if there is a long-term trend in the data.



Simple Exponential Smoothing

If you are trying to use simple exponential smoothing to forecast prices that have been rising over time (*due to inflation*) you will have to deflate the price series.

The other option is to *modify the model* to work with trending data.



Smoothing Linear Trends

There are slightly more sophisticated smoothing models that can be used for series with a trend.

The first one we will look at will smooth both the *level* and the *trend* (or you can think of it as smoothing both the intercept and the slope).

Smoothing Linear Trends
 Smoothing a linear trend is done by following a sequence of equations:

1. $S_t = F_{t-1} + \alpha_1 e_{t-1}$

where: S_t = the smoothed level at the end of time t
 F_{t-1} = the forecast for t-1
 α_1 = the smoothing parameter for the level
 e_{t-1} = the forecast error in t-1

Smoothing Linear Trends

2. $T_t = T_{t-1} + \alpha_2 e_{t-1}$

where: T_t = smoothed trend at t
 T_{t-1} = smoothed trend at t-1
 α_2 = smoothing parameter for the trend
 e_{t-1} = error in time t

Smoothing Linear Trends

3. $F_t = S_t + T_t$

Where F_t is our forecast for time t.

Like the simple model, we need some starting values, in particular, we need values for S_0 and T_0 .



Determining the α values

The α values can be determined by trial and error as before.

- set $\alpha_1 = 0.5$ and adjust α_2
- once you have found an α_2 that minimizes MSE, begin adjusting α_1
- try a few new values for α_2
- repeat the process



Determining the α values

Note:

α_1 and α_2 *do not have to be equal*, and in most cases should not be.

As with the simple model,

$0 < \alpha_i < 1$



How About Starting Values?

Like the simple model, we need some starting values, in particular, we need values for S_0 and T_0

Where do we get these?

One convenient way is to plug in the values from a regression model of $P = f(\text{time})$.



Smoothing Nonlinear Trends

Forecast accuracy can often be improved by using a *nonlinear trend model*.

Nonlinear trend forecasts can be generated by a simple modification to the linear trend smoothing model.



Smoothing Nonlinear Trends

We will consider two types of nonlinear trends:

- exponential
- damped



Smoothing Nonlinear Trends

The exponential trend

- the amount of growth in each period will increase

The damped trend

- the amount of growth in each period will decline



Smoothing Nonlinear Trends

The exponential smoothing model for nonlinear trends consists of three parts (like our linear trend model).

The first equation is:

$$1. S_t = F_{t-1} + \alpha_1 e_{t-1}$$

This is identical to the formula we used for smoothing linear trends.



Smoothing Nonlinear Trends

The second equation is:

$$2. T_t = \phi T_{t-1} + \alpha_2 e_{t-1}$$

This is also the same as before, except the trend is multiplied by the parameter ϕ . The effect of this parameter is to *accelerate* or *decelerate* the trend.



Smoothing Nonlinear Trends

Given the preceding formula:

- when $\phi > 1$, we have an exponential trend
- when $0 < \phi < 1$, we have a damped trend
- when $\phi = 1$, we have a linear trend



Smoothing Nonlinear Trends

The forecast is calculated as before, from the sum of the two equations:

$$3.F_t = S_t + \phi T_t$$

The *only difference* between these equations and the ones we used for smoothing a linear trend is the presence of ϕ .



Smoothing Nonlinear Trends

If we are only concerned about one-step-ahead forecasting, i.e. predicting one-period into the future then-

selecting a trend is straightforward.

Simply pick the ϕ , α_1 and α_2 that minimize MSE, and adjust them at each step.



One Step Ahead Forecasts

Predicting several steps into the future with nonlinear trend models can cause problems.

For example, the damped trend will eventually predict *growth stops*.



Simple Linear Regression

The term “regression” was introduced by Francis Galton.

He found the height of children born to either tall or short parents tended to “regress” toward the average height of the population.



Simple Linear Regression

The modern interpretation of regression is quite different:

Regression analysis is concerned with the study of the dependence of one variable on one or more other variables.



Simple Linear Regression

Regression is carried out in order to predict the mean value of the *dependent variable* in terms of values of the known (or fixed in repeated sampling) *explanatory variables*.



Simple Linear Regression

Regression is concerned with explaining changes in a *stochastic* variable from changes in *deterministic* (or non-stochastic) variables.

This is a *statistical*, not functional, relationship.



Simple Linear Regression

Statistical relationships deal with random or stochastic variables, that is, variables that have *probability distributions*.

Deterministic relationships are not random. An example would be Newton's law of gravity.



Simple Linear Regression

Regression analysis does not necessarily imply *causation*.

One example of this is that we will use time to explain prices.

Correlation between prices and time may exist, but time *does not cause* prices



Simple Linear Regression

Because the dependent variable is assumed stochastic, we must append a stochastic term to the right hand side of our estimation equation.

This is known as the error term.

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + e$$



The Error Term

The error term (e) serves as a surrogate for all those variables that are omitted from our model, but collectively affect Y.



The Error Term

Why don't we just include all the important variables in the model?

- our theory regarding the behavior of Y may be incomplete
- we may not have data on these variables



The Error Term

Why don't we just include all the important variables in the model?

- combinations of included variables may be important
- even if we include all variables, there will be some intrinsic randomness in Y



The Error Term

Why don't we just include all the important variables in the model?

- The data we have may not be measured accurately
- Occam's Razor



Terminology

Once estimates have been obtained from the regression, the parameters, typically referred to as betas, are given "hats." For simplicity we will refer to the estimator of beta as "b"



Some Definitions

The “standard” regression model is written as:

$$y = b_0 + b_1X + e$$

- y is the dependent variable that we wish to forecast
- X is a set of independent variables that explain y
- e is random error
