Chapter 465

Exponential Smoothing – Horizontal

Introduction

Simple exponential smoothing forecasts horizontal series: those without trends or seasonal patterns. It is appropriate for short-term forecasts of series using a weighted average of the most recent observations.

The forecasting algorithm makes use of the following formulas:

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

Here α is the smoothing constant which is between zero and one.

The forecast at time T for the value at time $T+k$ is F_T.

Another form of the above equation which shows how this procedure received its name is

$$F_t = \alpha X_t + \alpha(1-\alpha)X_{t-1} + \alpha(1-\alpha)^2 X_{t-2} + \alpha(1-\alpha)^3 X_{t-3} + \cdots$$

From this equation we see that the method constructs a weighted average of the observations. The weight of each observation decreases exponentially as we move back in time. Hence, since the weights decrease exponentially and averaging is a form of smoothing, the technique was name exponential smoothing.

Smoothing Constants

Notice that the *smoothing constant*, α, determines how fast the weights of the series decays. The value may be chosen either subjectively or objectively. Values near one put almost all weight on the most recent observations. Values of the smoothing constant near zero allow the distant past observations to have a large influence.

When selecting the smoothing constant *subjectively*, you use your own experience with this, and similar, series. Also, specifying the smoothing constant yourself lets you tune the forecast to your own beliefs about the future of the series. If you believe that the mechanism generating the series has recently gone through some fundamental changes, use a smoothing constant value of 0.9 which will cause distant observations to be ignored. If, however, you think the series is fairly stable and only going through random fluctuations, use a value of 0.1.

To select the value of the smoothing constant *objectively*, you search for a value that is best in some sense. Our program searches for that value that minimizes the size of the combined forecast errors of the currently available series. Three methods of summarizing the amount of error in the forecasts are available: the mean square error (MSE), the mean absolute error (MAE), and the mean absolute percent error (MAPE). The forecast error is the difference between the forecast of the current period made at the last period and the value of the series at the current period. This is written as

$$e_t = X_t - F_{t-1}$$

Using this formulation, we can define the three error-size criterion as follows:

$$MSE = \frac{1}{n}\sum e_t^2$$

$$MAE = \frac{1}{n}\sum \left|e_t\right|$$

$$MAPE = \frac{100}{n}\sum \left|\frac{e_t}{X_t}\right|$$

To find the value of the smoothing constants objectively, we select one of these criterion and search for that value of α that minimize this function. The program conducts a search for the appropriate values using an efficient grid-searching algorithm.

Initial Values

Exponential smoothing requires special initialization since the forecast for period one requires the forecast at period zero, which we do not, by definition, have. Several methods have been proposed for generating starting values. We have adopted the backcasting method which is currently considered to be one of the best. Backcasting is simply reversing the series so that we forecast into the past instead of into the future. This produces the required starting value. Once we have done this, we can then switch the series back and apply the algorithm in the regular manor.

Relationship to ARIMA Method

It can be shown that both exponential smoothing is equivalent to the ARIMA(0,1,1) model (see Kendall and Ord (1990) page 130). This is why backcasting is recommended for initial values.

Assumptions and Limitations

This algorithm is useful for short-term forecasting of nonseasonal time series with no apparent upward or downward. The series is assumed to have a changing (or evolving) mean that is not fixed over all time. We assume that future values of this average are unpredictable, so that the current level (current average or mean) of the series is the best forecast of future values.

Data Structure

The data are entered in a single variable.

Missing Values

When missing values are found in the series, they are either replaced or omitted. The replacement value is the average of the nearest observation in the future and in the past or the nearest non-missing value in the past.

If you do not feel that this is a valid estimate of the missing value, you should manually enter a more reasonable estimate before using the algorithm. These missing value replacement methods are particularly poor for seasonal data. We recommend that you replace missing values manually before using the algorithm.

Example 1 – Horizontal Exponential Smoothing

This section presents an example of how to generate a forecast of a horizontal series. The data in the Intel dataset gives price and volume data for Intel stock during August, 1995. We will forecast values for daily volumes. These values are contained in the variable Intel_Volume.

Setup

To run this example, complete the following steps:

1 Open the Intel example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Intel** and click **OK**.

2 Specify the Exponential Smoothing – Horizontal procedure options

- Find and open the **Exponential Smoothing – Horizontal** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings template. To load this template, click **Open Example Template** in the Help Center or File menu.

Option	Value
Variables Tab	
Time Series Variable(s)	**Intel_Volume**
Reports Tab	
Forecast Report	**Data and Forecasts**

3 Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Forecast Summary Section

Forecast Summary Section

Variable	Intel_Volume
Number of Rows	20
Missing Values	None
Mean	10974.54
Pseudo R-Squared	0.000000
Mean Square Error	1.632774E+07
Mean |Error|	2876.168
Mean |Percent Error|	25.98573
Alpha Search	Mean Square Error
Alpha	0.3769887
Forecast	13100.84

This report summarizes the forecast equation.

Variable

The name of the variable for which the forecasts are generated.

Number of Rows

The number of rows that were in the series. This is provided to allow you to double-check that the correct series was used.

Missing Values

If missing values were found, this option lists the method used to estimate them.

Mean

The mean of the variable across all time periods.

Pseudo R-Squared

This value generates a statistic that acts like the R-Squared value in multiple regression. A value near zero indicates a poorly fitting model, while a value near one indicates a well-fitting model. The statistic is calculated as follows:

$$R^2 = 100\left(1 - \frac{SSE}{SST}\right)$$

where *SSE* is the sum of square residuals and *SST* is the total sum of squares after correcting for the mean.

Mean Square Error

The average squared residual (MSE) is a measure of how closely the forecasts track the actual data. The statistic is popular because it shows up in analysis of variance tables. However, because of the squaring, it tends to exaggerate the influence of outliers (points that do not follow the regular pattern).

Mean |Error|

The average absolute residual (MAE) is a measure of how closely the forecasts track the actual data without the squaring.

Mean |Percent Error|

The average percent absolute residual (MAPE) is a measure of how closely the forecasts track the actual data put on a percentage basis.

Alpha Search

If a search was made to find the best value of the smoothing constant, this row gives the criterion used during the search.

Alpha

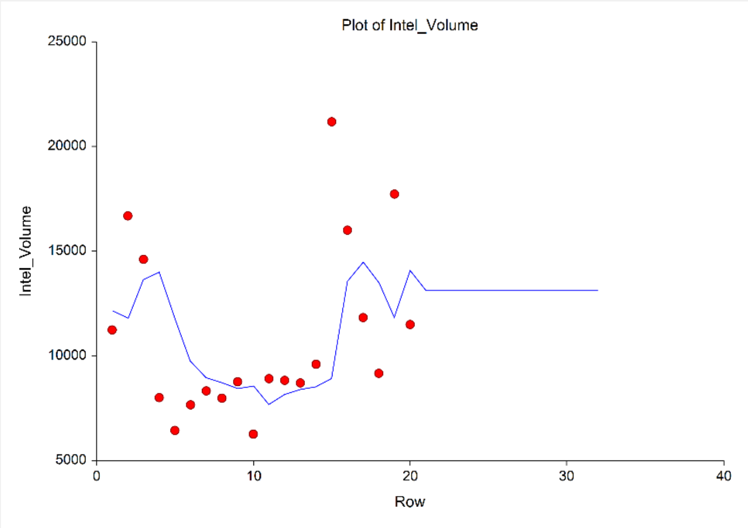
The value of the smoothing constant that was used to generate the forecasts.

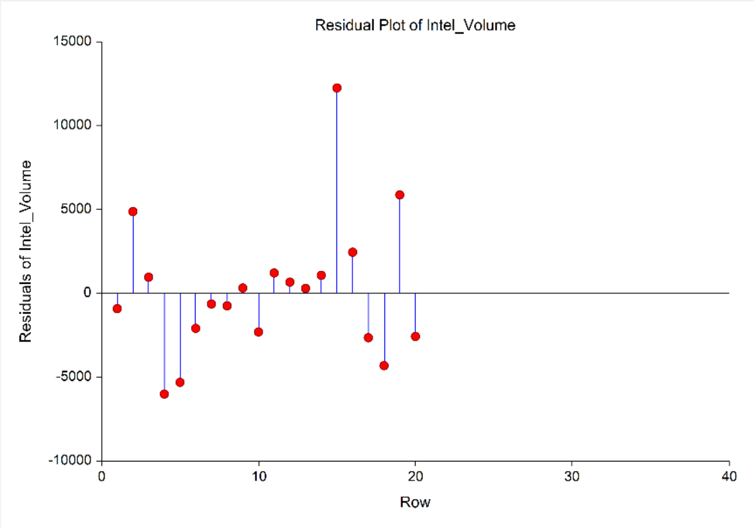
Forecast

The value of the forecast. This is the value that used to forecast future values from this point on. Remember that this method does not adjust for trend or seasonality, so only the current average is used for forecasting.

Forecast and Residuals Plots

Forecast and Residuals Plots





Forecast Plot

The forecast plot lets you analyze how closely the forecasts track the data. The plot also shows the forecasts at the end of the data series.

Residual Plot

This plot lets you analyze the residuals themselves. You are looking for patterns, outliers, or any other information that may help you improve the forecasting model. The first thing to compare is the scale of the Residual Plot versus the scale of the Forecast Plot. If your forecasting is working well, the vertical scale of the Residual Plot will be much less than the scale of the Forecast Plot.

Forecasts Section

Forecasts Section

Row No.	Forecast Intel_Volume	Actual Intel_Volume	Residuals
1	12153.88	11242.2	-911.6835
2	11810.19	16689.9	4879.711
3	13649.79	14613.3	963.5151
4	14013.02	8009	-6004.019
5	11749.57	6441.8	-5307.772
6	9748.602	7664.5	-2084.102
7	8962.919	8330.3	-632.619
8	8724.429	7983	-741.4288
9	8444.919	8767.1	322.1815
10	8566.377	6266.4	-2299.977
11	7699.312	8915.3	1215.988
12	8157.726	8833	675.2744
13	8412.297	8709.7	297.4036
14	8524.414	9603	1078.586
15	8931.029	21185.2	12254.17
16	13550.71	16006.5	2455.787
17	14476.52	11832.4	-2644.117
18	13479.71	9168.1	-4311.615
19	11854.29	17729.3	5875.015
20	14069.1	11500.7	-2568.399
21	13100.84		
22	13100.84		
23	13100.84		
24	13100.84		
25	13100.84		
26	13100.84		
27	13100.84		
28	13100.84		
29	13100.84		
30	13100.84		
31	13100.84		
32	13100.84		

This section shows the values of the forecasts, the actual values, and the residuals.