Chapter 467

Exponential Smoothing – Trend & Seasonal

Introduction

This module forecasts seasonal series with upward or downward trends using the Holt-Winters exponential smoothing algorithm. Two seasonal adjustment techniques are available: additive and multiplicative.

Additive Seasonality

Given observations $X_1, X_2, \ldots, X_t$ of a time series, the Holt-Winters additive seasonality algorithm computes an evolving trend equation with a seasonal adjustment that is additive. *Additive* means that the amount of the adjustment is constant for all levels (average value) of the series.

The forecasting algorithm makes use of the following formulas:

$$a_t = \alpha(X_t - F_{t-s}) + (1-\alpha)(a_{t-1} + b_{t-1})$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}$$

$$F_t = \gamma(X_t - a_t) + (1-\gamma)F_{t-s}$$

Here α, β, and γ are smoothing constants which are between zero and one. Again, a_t gives the y-intercept (or level) at time *t*, while b_t is the slope at time *t*. The letter *s* represents the number of periods per year, so the quarterly data is represented by s = 4 and monthly data is represented by s = 12.

The forecast at time *T* for the value at time *T+k* is $a_T + b_T k + F_{[(T+k-1)/s]+1}$. Here *[(T+k-1)/s]* is means the remainder after dividing *T+k-1* by *s*. That is, this function gives the season (month or quarter) that the observation came from.

Multiplicative Seasonality

Given observations $X_1, X_2, \ldots, X_t$ of a time series, the Holt-Winters multiplicative seasonality algorithm computes an evolving trend equation with a seasonal adjustment that is multiplicative. *Multiplicative* means that the amount of the adjustment varies with the level (average value) of the series. Note that the nature of most economic time series makes the multiplicative model more popular than the additive model.

The forecasting algorithm makes use of the following formulas:

$$a_t = \alpha(X_t / F_{t-s}) + (1-\alpha)(a_{t-1} + b_{t-1})$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}$$

$$F_t = \gamma(X_t / a_t) + (1-\gamma)F_{t-s}$$

Here α, β, and γ are smoothing constants which are between zero and one. Again, a_t gives the y-intercept (or level) at time *t*, while b_t is the slope at time *t.* The letter *s* represents the number of periods per year, so the quarterly data is represented by s = 4 and monthly data is represented by s = 12.

The forecast at time *T* for the value at time *T+k* is $(a_T + b_T k)F_{[(T+k-1)/s]+1}$. Here *[(T+k-1)/s]* is means the remainder after dividing *T+k-1* by *s*. That is, this function gives the season (month or quarter) that the observation came from.

Smoothing Constants

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

Note that α is associated with the level of the series, β is associated with the trend, and γ is associated with the seasonality factors.

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 constants *objectively*, you search for values that are best in some sense. Our program searches for those values that minimize 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:

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

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

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

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

Initial Values

Winters method requires initialization since the forecast for period one requires the forecast at period zero, which we do not, by definition, have. It also requires the seasonal adjustment factors. Several methods have been proposed for generating starting values. **NCSS** uses the initialization method described in Bowerman and O'Connell (1993).

Relationship to ARIMA Method

The multiplicative seasonal adjustment model does not have an ARIMA counterpart, while the additive model does.

Assumptions and Limitations

These algorithms are useful for forecasting seasonal time series with (local or global) trend.

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 – Trend & Seasonal Exponential Smoothing

This section presents an example of how to generate forecasts of a series using Winters multiplicative seasonal model. The data in the Sales dataset will be used. We will forecast the values of the Sales variable for the next twelve months.

Setup

To run this example, complete the following steps:

1 Open the Sales example dataset

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

2 Specify the Exponential Smoothing – Trend / Seasonal procedure options

- Find and open the **Exponential Smoothing – Trend / Seasonal** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab

Time Series Variable(s) .... **Sales**
First Year **1970**

Report 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	Sales
Number of Rows	144
Missing Values	None
Mean	174.2847
Pseudo R-Squared	0.980145
Mean Square Error	16.10279
Mean |Error|	3.114085
Mean |Percent Error|	1.786407
Forecast Method	Winter's with multiplicative seasonal adjustment.
Search Iterations	532
Search Criterion	Mean Square Error
Alpha	0.3496813
Beta	5.572607E-05
Gamma	6.69852E-11
Intercept (A)	138.3665
Slope (B)	0.5871831
Season 1 Factor	0.9028255
Season 2 Factor	0.8556558
Season 3 Factor	0.9714928
Season 4 Factor	0.997539
Season 5 Factor	1.028944
Season 6 Factor	1.028335
Season 7 Factor	0.9977484
Season 8 Factor	1.005503
Season 9 Factor	0.9752801
Season 10 Factor	1.023642
Season 11 Factor	1.007328
Season 12 Factor	1.205706

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.

Forecast Method

This line shows which of the two possible seasonal adjustment algorithms was selected.

Search Iterations

This line shows how many iterations were needed to find the best values for the smoothing constants.

Search Criterion

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

Alpha

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

Beta

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

Gamma

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

Intercept (A)

The value of the y-intercept for time period one!

Slope (B)

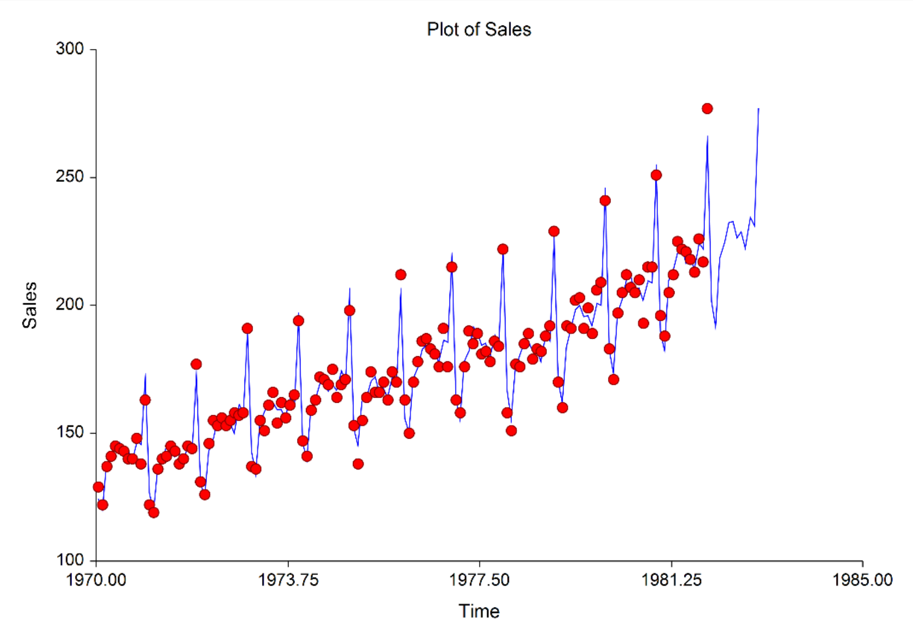
The value of the slope.

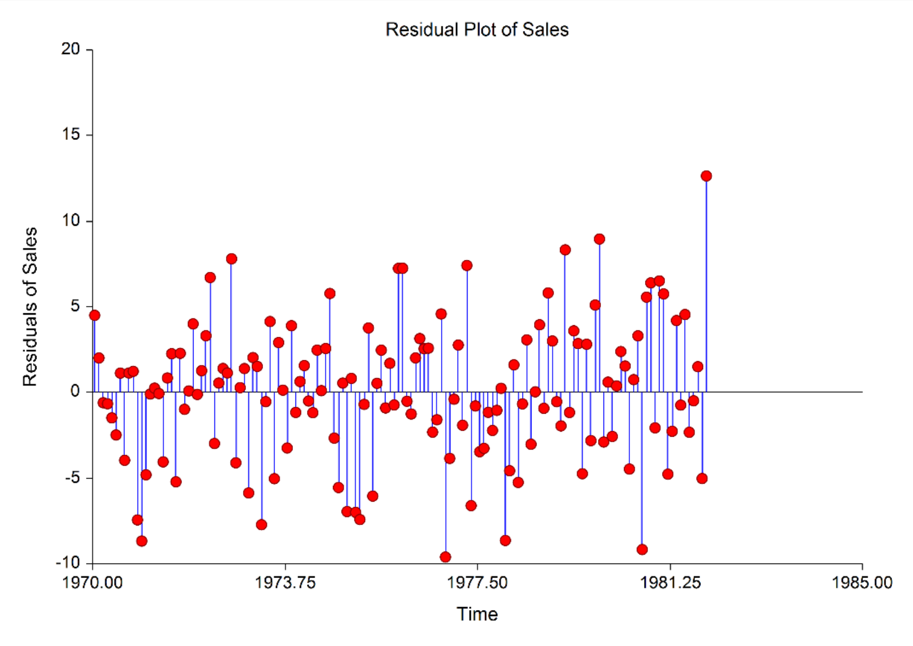
Season (1-12) Factor

The values of the multiplicative seasonal factors.

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 algorithm 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.	Date	Forecast Sales	Actual Sales	Residuals
1	19701	124.4976	129	4.502408
2	19702	119.9876	122	2.012394
3	19703	137.6008	137	-0.600761
4	19704	141.66	141	-0.6600129
5	19705	146.486	145	-1.486042
6	19706	146.4839	144	-2.483881
7	19707	141.8699	143	1.130069
8	19708	143.9612	140	-3.961171
9	19709	138.8632	140	1.136814
10	197010	146.7673	148	1.232683
11	197011	145.4439	138	-7.443867
12	197012	171.6791	163	-8.679084
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
140	19818	220.3259	218	-2.325934
141	19819	213.4872	213	-0.4872138
142	198110	224.4956	226	1.504373
143	198111	222.0268	217	-5.026803
144	198112	264.3555	277	12.64447
145	19821	201.7888		
146	19822	191.7484		
147	19823	218.2773		
148	19824	224.7152		
149	19825	232.394		
150	19826	232.8604		
151	19827	226.5199		
152	19828	228.8708		
153	19829	222.5643		
154	198210	234.2018		
155	198211	231.0607		
156	198212	277.2727		

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