

A TEACHER'S GUIDE TO THE MODELS USED IN TIME SERIES

MODULE OF iNZight

Introduction

The Time Series module of the FREE software package iNZight uses two different statistical models. The model used to obtain the series decomposition is called a Seasonal Trend Lowess and the model used to calculate predictions is a Holt-Winters model. This guide is to give teachers a brief summary of the models used but the new standard has no expectation that students need to know any theoretical background to the models.

Seasonal Trend Lowess

(Lowess – Locally weighted regression scatterplot smoothing)

Smoothing or filtering a Time Series is best thought of as similar to the idea of filtering music through an amplifier. We can amplify certain sounds or we can suppress certain sounds. Similarly, we can suppress (remove) certain features in a Time Series, such as seasonality, in order to model the trend and/or cycle. Once we have built a suitable model for the smoothed series, we can add back the appropriate seasonal component in order to produce predictions.

A common method for smoothing a Time Series is to use moving averages, which is what has traditionally be taught in schools for AS 3.1. One drawback of moving averages is that our moving average series becomes shorter than the original Time Series. If we have monthly data, our first moving average value is calculated on observations 1 to 12, and the second moving average value is calculated on observations 2 to 13. We then average these two values to get our first moving average value which then replaces observation 7 in our original series. Similarly, at the end of our series, there are six observations that we have no moving average values for.

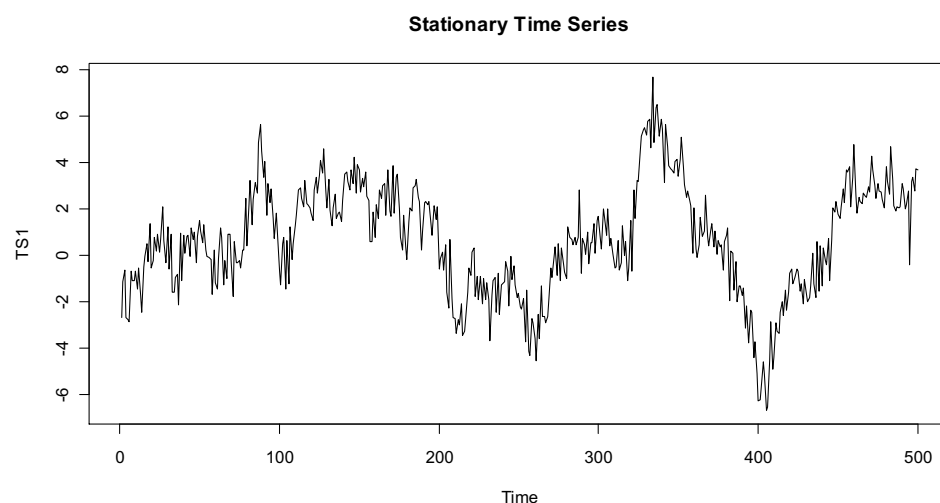
A more useful tool for isolating and then removing the seasonal component of a Time Series is **Seasonal Trend Lowess** a decomposition function in R (the programming language that iNZight is written in). The method used is to first smooth the trend and cycle using a lowess smoother (fitting a local regression to a window of points and using the point on the fitted regression line as the value of the smooth for the time value in the middle of the window). The regression that is used is “weighted”, in that observations near the edge of the window are given less weight than observations near the centre of the window when determining the local regression line. Then a separate lowess smoother is used on each seasonal sub-series (i.e. all the January observations, all the February observations, ...). The “trend and cycle” smoothed value and the appropriate “seasonal” smoothed value can be subtracted from the original observation to yield the remainder or random component for that

observation. iNZight produces a plot of the decomposition that shows the original series, the seasonal component, the trend and cycle and finally, the random component.

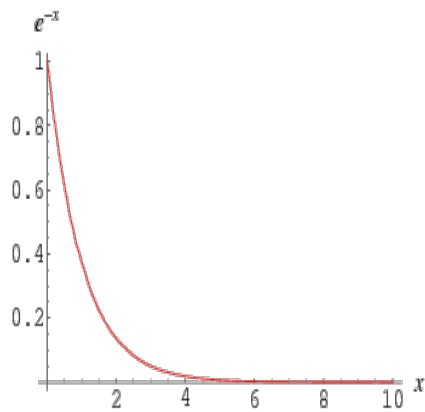
A third option for smoothing data is exponential smoothing and it is this technique that is used in the Holt-Winters model.

Holt-Winters Model

This model, often referred to as a procedure, was first proposed in the early 1960s. It uses a process known as exponential smoothing. All data values in a series contribute to the calculation of the prediction model.



Exponential smoothing in its simplest form should only be used for non-seasonal time series exhibiting a constant trend (or what is known as a stationary time series). It seems a reasonable assumption to give more weight to the more recent data values and less weight to the data values from further in the past. An intuitive set of weights is the set of weights that decrease each time by a constant ratio. Strictly speaking this implies an infinite number of past observations but in practice there will be a finite number. Such a procedure is known as **exponential smoothing** since the weights lie on an exponential curve.



If the smoothed series is denoted by S_t

α denotes the smoothing parameter, the exponential smoothing constant, $0 < \alpha < 1$

The smoothed series is given by: $S_t = \alpha y_t + (1 - \alpha)S_{t-1}$

where $S_1 = y_1$

The smaller the value of α , the smoother the resulting series.

It can be shown that: $S_t = \alpha y_t + \alpha(1 - \alpha)y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + \dots + (1 - \alpha)^{t-1} y_1$

Consider the following Time Series:

14 24 5 18 10 17 23 17 23 ...

Using the formulae above, with an exponential smoothing constant, $\alpha = 0.1$

$$S_1 = y_1 = 14$$

$$S_2 = \alpha y_2 + (1 - \alpha)S_1 = 0.1(24) + 0.9(14) = 15$$

$$S_3 = \alpha y_3 + (1 - \alpha)S_2 = 0.1(5) + 0.9(15) = 14$$

$$S_4 = \alpha y_4 + (1 - \alpha)S_3 = 0.1(18) + 0.9(14) = 14.4$$

etc

Thus the smoothed series depends on all previous values, with the most weight given to the most recent values.

Exponential smoothing requires a large number of observations.

Exponential smoothing is not appropriate for data that has a seasonal component, cycle or trend. However, modified methods of exponential smoothing are available to deal with data containing these components.

The Holt-Winters model uses a modified form of exponential smoothing. It applies three exponential smoothing formulae to the series. Firstly, the level (or mean) is smoothed to give a local average value for the series. Secondly, the trend is smoothed and lastly each seasonal sub-series (ie all the January values, all the February values..... for monthly data) is smoothed separately to give a seasonal estimate for each of the seasons. A combination of these three series is used to calculate the predictions output by iNZight.

The exponential smoothing formulae applied to a series with a trend and constant seasonal component using the Holt-Winters additive technique are:

$$a_t = (Y_t - s_{t-p}) + (1 - \alpha)(a_{t-1} + b_{t-1})$$

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

$$s_t = (Y_t - a_t) + (1 - \gamma)s_{t-p}$$

where:

α , β and γ are the smoothing parameters

a_t is the smoothed level at time t

b_t is the change in the trend at time t

s_t is the seasonal smooth at time t

p is the number of seasons per year

The Holt-Winters algorithm requires starting (or initialising) values. Most commonly:

$$a_p = \frac{1}{p}(Y_1 + Y_2 + \dots + Y_p)$$

$$b_p = \frac{1}{p} \left(\frac{Y_{p+1} - Y_1}{p} + \frac{Y_{p+2} - Y_2}{p} + \dots + \frac{Y_{p+p} - Y_p}{p} \right)$$

$$s_1 = Y_1 - a_p, \quad s_2 = Y_2 - a_p, \quad \dots, \quad s_p = Y_p - a_p$$

The Holt-Winters forecasts are then calculated using the latest estimates from the appropriate exponential smooths that have been applied to the series.

So we have our forecast for time period $T + 1$:

$$\hat{y}_{T+1} = a_T + b_T + s_T$$

where: a_T is the smoothed estimate of the level at time T

b_T is the smoothed estimate of the change in the trend value at time T

s_T is the smoothed estimate of the appropriate seasonal component at T

As mentioned earlier the Holt-Winters model assumes that the seasonal pattern is relatively constant over the time period. Students would be expected to notice changes in the seasonal pattern and identify this as a potential problem with the model, particularly if long-term predictions are made. In practice this is dealt with by transforming the original data and modelling the transformed series or using a multiplicative model. Students are not expected to know this, but are required to identify a variable seasonal pattern as a potential problem. The exponential smoothing formulae applied to a series using Holt-Winters Multiplicative models are:

$$a_t = \frac{Y_t}{s_{t-p}} + (1 - \alpha)(a_{t-1} + b_{t-1})$$

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

$$s_t = \frac{Y_t}{a_t} + (1 - \gamma)s_{t-p}$$

The initialising values are as for the additive model, except:

$$s_1 = \frac{Y_1}{a_p}, \quad s_2 = \frac{Y_2}{a_p}, \quad \dots, \quad s_p = \frac{Y_p}{a_p}$$

So we have our prediction for time period $T + 1$:

$$\hat{y}_{T+1} = (a_T + b_T)s_T$$

Calculation of Prediction Intervals for Holt Winters

Reference Yar,M. & Chatfield, C. (1990) Prediction intervals for the Holt-Winters forecasting procedure, *International Journal of Forecasting*, Vol. 6,pp 127-137, North Holland.

Rachel Passmore, Royal Society Endeavour Teacher Fellow,
Department of Statistics, University of Auckland

There are many situations where it is important to give interval predictions, rather than point predictions, as a means of assessing future uncertainty. An interval prediction associated with a prescribed probability is sometimes called a confidence interval, but it is recommended that the term prediction interval is used in the context of time series analysis. This is because prediction interval is more descriptive and because the term ‘confidence interval’ is usually applied to interval estimates of model parameters. Unfortunately it is relatively common to see predictions made without any reference to prediction intervals. This may be because there are a number of different ways that prediction intervals can be calculated. The paper above provides not only details of how the prediction intervals for Holt-Winters are produced but also compares the authors’ preferred method with several alternative methods. It also compares the prediction intervals calculated for the same data set by a variety of different models.

Derivation and details of prediction interval calculation can be found in Yar & Chatfield’s (1990) paper – see section 3 and 4 on page 129.

In one example given in the paper, a monthly index of employment in manufacturing in Canada, a prediction for three years after the end of the actual data is provided of 115.9. A prediction interval of [113.95,117.85] is also calculated. A suggested interpretation of this prediction interval (P.I.) is

‘There is a 95% chance that the true index value of employment in manufacturing in Canada in three years time will be between 113.95 and 117.85.’

The details given in this paper apply to an additive Holt-Winters model only.

Assessing non-stationary model forecasts

The test of any prediction model is how well does it predict when compared to actual data values. To do this either remove the last few given observations or find the next few actual observations. Different prediction models can then be compared using a statistic known as the Root Mean Squared Error of Prediction (RMSEP). The formula for calculating this statistic is given below

$$RMSEP = \sqrt{\frac{1}{\tau} \sum_{t=1}^{\tau} (y_t - \hat{y}_t)^2}$$

where τ is the number of predictions we are using in the calculation of *RMSEP*.

Students are not expected to calculate RMSEP.

Want to know more? There are several Youtube clips that explain exponential smoothing and Holt-Winters using Excel or R if you are interested.

Summary of what students need to know

- Holt –Winters Additive model assumes seasonal pattern is reasonably constant

Rachel Passmore, Royal Society Endeavour Teacher Fellow,
Department of Statistics, University of Auckland

- Holt-Winters Model uses a technique of exponential smoothing, which is a weighted sum of previous values in a series. More weight is given to more recent values and less weight is given to values from the distant past.
- Holt-Winters Additive model exponentially smooths three series in order to produce predictions – the level, the trend and the seasonal sub-series.
- Students should be able to identify cyclical components and inconsistent seasonal patterns. They should note that such features are incompatible with assumptions underlying Holt-Winters Additive model and suggest a multiplicative model be considered instead. Such a comment would be expected at Excellence level only. Students are NOT expected to calculate a multiplicative model.

Rachel Passmore

Endeavour Teacher Fellow

Email : passm@vodafone.co.nz

August 2012

With thanks to Mike Forster, Department of Statistics,
University of Auckland