CHAPTER 4

EXPONENTIAL SMOOTHING METHODS

If you have to forecast, forecast often.

EDGAR R. FIEDLER, *American economist*

4.1 INTRODUCTION

We can often think of a data set as consisting of two distinct components: **signal** and **noise**. Signal represents any pattern caused by the intrinsic dynamics of the process from which the data are collected. These patterns can take various forms from a simple constant process to a more complicated structure that cannot be extracted visually or with any basic statistical tools. The constant process, for example, is represented as

$$y_t = \mu + \varepsilon_t, \tag{4.1}$$

where μ represents the underlying constant level of system response and ε_t is the noise at time t. The ε_t is often assumed to be uncorrelated with mean 0 and constant variance σ_ϵ^2.

We have already discussed some basic data smoothers in Section 2.2.2. **Smoothing** can be seen as a technique to separate the signal and the noise

Introduction to Time Series Analysis and Forecasting, Second Edition.
Douglas C. Montgomery, Cheryl L. Jennings and Murat Kulahci.

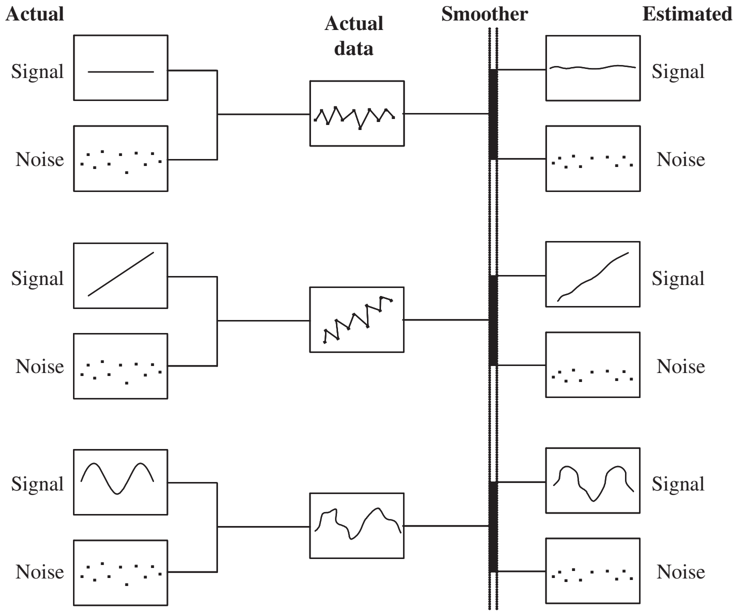


FIGURE 4.1 The process of smoothing a data set.

as much as possible and in that a smoother acts as a filter to obtain an "estimate" for the signal. In Figure 4.1, we give various types of signals that with the help of a smoother can be "reconstructed" and the underlying pattern of the signal is to some extent recovered. The smoothers that we will discuss in this chapter achieve this by simply relating the current observation to the previous ones. For a given data set, one can devise forward and/or backward looking smoothers but in this chapter we will only consider backward looking smoothers. That is, at any given T, the observation y_T will be replaced by a combination of observations at and before T. It does then intuitively make sense to use some sort of an "average" of the current and the previous observations to smooth the data. An obvious choice is to replace the current observation with the average of the observations at $T, T-1, \ldots, 1$. In fact this is the "best" choice in the least squares sense for a constant process given in Eq. (4.1).

A constant process can be smoothed by replacing the current observation with the best estimate for μ. Using the least squares criterion, we define the error sum of squares, SS_E, for the constant process as

$$SS_E = \sum_{t=1}^{T} (y_t - \mu)^2.$$

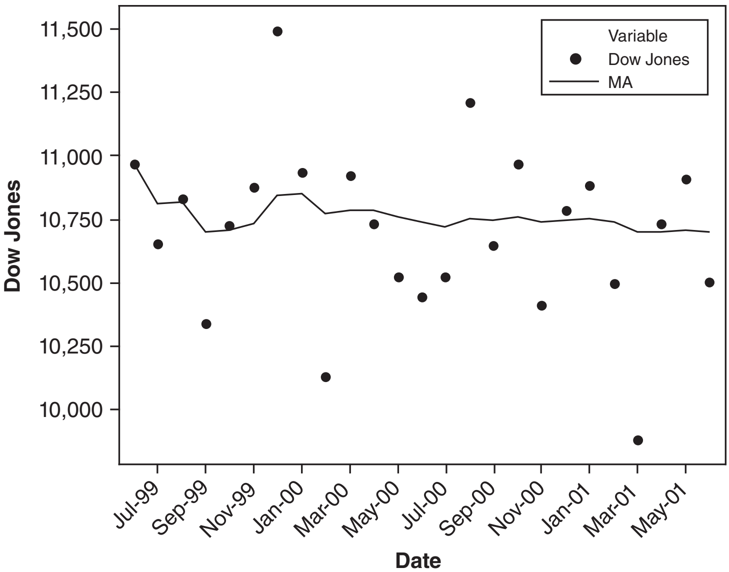


FIGURE 4.2 The Dow Jones Index from June 1999 to June 2001.

The least squares estimate of μ can be found by setting the derivative of SS with respect to μ to 0. This gives

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} y_t, \tag{4.2}$$

where $\hat{\mu}$ is the least squares estimate of μ. Equation (4.2) shows that the least squares estimate of μ is indeed the average of observations up to time T.

Figure 4.2 shows the monthly data for the Dow Jones Index from June 1999 to June 2001. Visual inspection suggests that a constant model can be used to describe the general pattern of the data.[1] To further confirm this claim, we use the smoother described in Eq. (4.2) for each data point by taking the average of the available data up to that point in time. The smoothed observations are shown by the line segments in Figure 4.2. It can be seen that the smoother in Eq. (4.2) indeed extracts the main pattern

[1] Please note that for this data the independent errors assumption in the constant process in Eq. (4.1) may have been violated. Remedies to check and handle such violations will be provided in the following chapters.

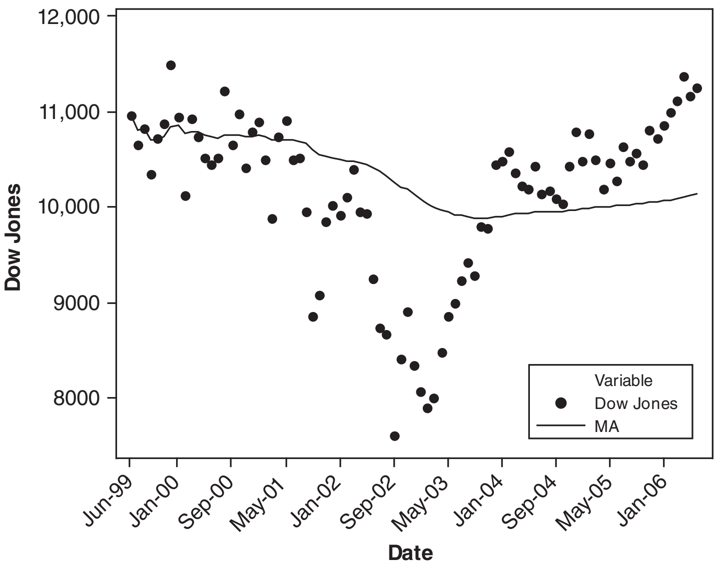


FIGURE 4.3 The Dow Jones Index from June 1999 to June 2006.

in the data and leads to the conclusion that during the 2-year period from June 1999 to June 2001, the Dow Jones Index was quite stable.

As we can see, for the constant process the smoother in Eq. (4.2) is quite effective in providing a clear picture of the underlying pattern. What happens if the process is not constant but exhibits a more complicated pattern? Consider again, for example, the Dow Jones Index from June 1999 to June 2006 given in Figure 4.3 (the complete data set is in Table 4.1). It is clear that the data do not follow the behavior typical of a constant behavior during this period. In Figure 4.3, we can also see the pattern that the smoother in Eq. (4.2) extracts for the same period. As the process changes, this smoother is having trouble keeping up with the process. What could be the reason for the poor performance after June 2001? The answer is quite simple: the constant process assumption is no longer valid. However, as time goes on, the smoother in Eq. (4.2) accumulates more and more data points and gains some sort of “inertia”. So when there is a change in the process, it becomes increasingly more difficult for this smoother to react to it.

How often is the constant process assumption violated? The answer to this question is provided by the Second Law of Thermodynamics, which in the most simplistic way states that if left on its own (free of external influences) any system will deteriorate. Thus the constant process is not

TABLE 4.1 Dow Jones Index at the End of the Month from June 1999 to June 2006

Date	Dow Jones	Date	Dow Jones	Date	Dow Jones	Date	Dow Jones
Jun-99	10,970.8	Apr-01	10,735	Feb-03	7891.08	Dec-04	10,783
Jul-99	10,655.2	May-01	10,911.9	Mar-03	7992.13	Jan-05	10,489.9
Aug-99	10,829.3	Jun-01	10,502.4	Apr-03	8480.09	Feb-05	10,766.2
Sep-99	10,337	Jul-01	10,522.8	May-03	8850.26	Mar-05	10,503.8
Oct-99	10,729.9	Aug-01	9949.75	Jun-03	8985.44	Apr-05	10,192.5
Nov-99	10,877.8	Sep-01	8847.56	Jul-03	9233.8	May-05	10,467.5
Dec-99	11,497.1	Oct-01	9075.14	Aug-03	9415.82	Jun-05	10,275
Jan-00	10,940.5	Nov-01	9851.56	Sep-03	9275.06	Jul-05	10,640.9
Feb-00	10,128.3	Dec-01	10,021.6	Oct-03	9801.12	Aug-05	10,481.6
Mar-00	10,921.9	Jan-02	9920	Nov-03	9782.46	Sep-05	10,568.7
Apr-00	10,733.9	Feb-02	10,106.1	Dec-03	10,453.9	Oct-05	10,440.1
May-00	10,522.3	Mar-02	10,403.9	Jan-04	10488.1	Nov-05	10,805.9
Jun-00	10,447.9	Apr-02	9946.22	Feb-04	10,583.9	Dec-05	10,717.5
Jul-00	10,522	May-02	9925.25	Mar-04	10,357.7	Jan-06	10,864.9
Aug-00	11,215.1	Jun-02	9243.26	Apr-04	10,225.6	Feb-06	10,993.4
Sep-00	10,650.9	Jul-02	8736.59	May-04	10,188.5	Mar-06	11,109.3
Oct-00	10,971.1	Aug-02	8663.5	Jun-04	10,435.5	Apr-06	11,367.1
Nov-00	10,414.5	Sep-02	7591.93	Jul-04	10,139.7	May-06	11,168.3
Dec-00	10,788	Oct-02	8397.03	Aug-04	10,173.9	Jun-06	11,247.9
Jan-01	10,887.4	Nov-02	8896.09	Sep-04	10,080.3		
Feb-01	10,495.3	Dec-02	8341.63	Oct-04	10,027.5		
Mar-01	9878.78	Jan-03	8053.81	Nov-04	10,428		

the norm but at best an exception. So what can we do to deal with this issue? Recall that the problem with the smoother in Eq. (4.2) was that it reacted too slowly to process changes because of its inertia. In fact, when there is a change in the process, earlier data no longer carry the information about the change in the process, yet they contribute to this inertia at an equal proportion compared to the more recent (and probably more useful) data. The most obvious choice is to somehow discount the older data. Also recall that in a simple average, as in Eq. (4.2), all the observations are weighted equally and hence have the same amount of influence on the average. Thus, if the weights of each observation are changed so that earlier observations are weighted less, a faster reacting smoother should be obtained. As mentioned in Section 2.2.2, a common solution is to use the **simple moving average** given in Eq. (2.3):

$$M_T = \frac{y_T + y_{T-1} + \cdots + y_{T-N+1}}{N} = \frac{1}{N} \sum_{t=T-N+1}^{N} y_t.$$

The most crucial issue in simple moving averages is the choice of the **span**, N. A simple moving average will react faster to the changes if N is small. However, we know from Section 2.2.2 that the variance of the simple moving average with uncorrelated observations with variance σ^2 is given as

$$\mathrm{Var}(M_T) = \frac{\sigma^2}{N}.$$

This means that as N gets small, the variance of the moving average gets bigger. This creates a dilemma in the choice of N. If the process is expected to be constant, a large N can be used whereas a small N is preferred if the process is changing. In Figure 4.4, we show the effect of going from a span of 10 observations to 5 observations. While the latter exhibits a more jittery behavior, it nevertheless follows the actual data more closely. A more thorough analysis on the choice of N can be performed based on the prediction error. We will explore this for exponential smoothers in Section 4.6.1, where we will discuss forecasting using exponential smoothing.

A final note on the moving average is that even if the individual observations are independent, the moving averages will be autocorrelated as two successive moving averages contain the same $N-1$ observations. In fact,

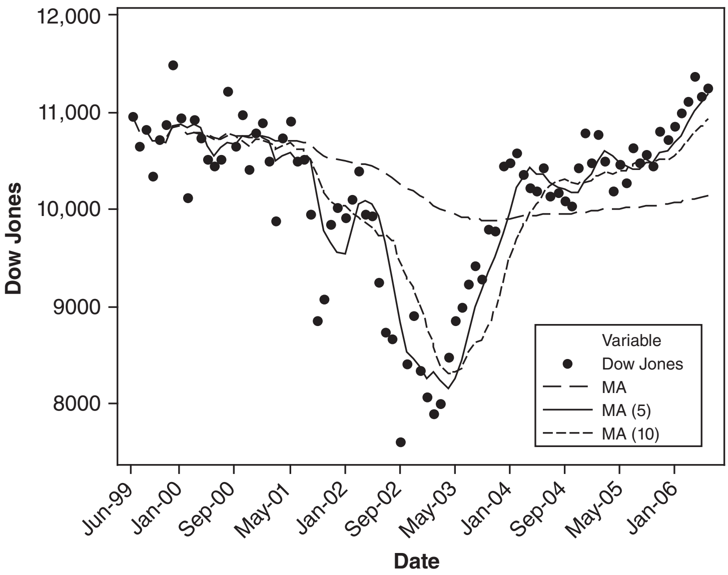


FIGURE 4.4 The Dow Jones Index from June 1999 to June 2006 with moving averages of span 5 and 10.

the autocorrelation function (ACF) of the moving averages that are k-lags apart is given as

$$\rho k = \begin{cases} 1 - \dfrac{|k|}{N}, & k < N \\ 0, & k \geq N \end{cases}.$$

4.2 FIRST-ORDER EXPONENTIAL SMOOTHING

Another approach to obtain a smoother that will react to process changes faster is to give geometrically decreasing weights to the past observations. Hence an exponentially weighted smoother is obtained by introducing a discount factor θ as

$$\sum_{t=0}^{T-1} \theta^t y_{T-t} = y_T + \theta y_{T-1} + \theta^2 y_{T-2} + \cdots + \theta^{T-1} y_1. \tag{4.3}$$

Please note that if the past observations are to be discounted in a geometrically decreasing manner, then we should have $|\theta| < 1$. However, the smoother in Eq. (4.3) is not an *average* as the sum of the weights is

$$\sum_{t=0}^{T-1} \theta^t = \frac{1-\theta^T}{1-\theta} \tag{4.4}$$

and hence does not necessarily add up to 1. For that we can adjust the smoother in Eq. (4.3) by multiplying it by $(1-\theta)/(1-\theta^T)$. However, for large T values, θ^T goes to zero and so the exponentially weighted average will have the following form:

$$\begin{aligned} \tilde{y}_T &= (1-\theta)\sum_{t=0}^{T-1} \theta^t y_{T-t} \\ &= (1-\theta)(y_T + \theta y_{T-1} + \theta^2 y_{T-2} + \cdots + \theta^{T-1} y_1) \end{aligned} \tag{4.5}$$

This is called a **simple** or **first-order exponential smoother**. There is an extensive literature on exponential smoothing. For example, see the books by Brown (1963), Abraham and Ledolter (1983), and Montgomery et al. (1990), and the papers by Brown and Meyer (1961), Chatfield and Yar (1988), Cox (1961), Gardner (1985), Gardner and Dannenbring (1980), and Ledolter and Abraham (1984).

An alternate expression in a recursive form for simple exponential smoothing is given by

$$\begin{aligned} \tilde{y}_T &= (1-\theta)y_T + (1-\theta)(\theta y_{T-1} + \theta^2 y_{T-2} + \cdots + \theta^{T-1} y_1) \\ &= (1-\theta)y_T + \theta \underbrace{(1-\theta)(y_{T-1} + \theta^1 y_{T-2} + \cdots + \theta^{T-2} y_1)}_{\tilde{y}_{T-1}} \\ &= (1-\theta)y_T + \theta \tilde{y}_{T-1}. \end{aligned} \tag{4.6}$$

The recursive form in Eq. (4.6) shows that first-order exponential smoothing can also be seen as the linear combination of the current observation and the smoothed observation at the previous time unit. As the latter contains the data from all previous observations, the smoothed observation at time T is in fact the linear combination of the current observation and the discounted sum of all previous observations. The simple exponential smoother is often represented in a different form by setting $\lambda = 1-\theta$,

$$\tilde{y}_T = \lambda y_T + (1-\lambda)\tilde{y}_{T-1} \tag{4.7}$$

In this representation the **discount factor**, λ, represents the weight put on the last observation and $(1-\lambda)$ represents the weight put on the smoothed value of the previous observations.

Analogous to the size of the span in moving average smoothers, an important issue for the exponential smoothers is the choice of the discount factor, λ. Moreover, from Eq. (4.7), we can see that the calculation of $\tilde{y}_1$ would require us to know $\tilde{y}_0$. We will discuss these issues in the next two sections.

4.2.1 The Initial Value, $\tilde{y}_0$

Since $\tilde{y}_0$ is needed in the recursive calculations that start with $\tilde{y}_1 = \lambda y_1 + (1-\lambda)\tilde{y}_0$, its value needs to be estimated. But from Eq. (4.7) we have

$$
\begin{aligned}
\tilde{y}_1 &= \lambda y_1 + (1-\lambda)\tilde{y}_0 \\
\tilde{y}_2 &= \lambda y_2 + (1-\lambda)\tilde{y}_1 = \lambda y_2 + (1-\lambda)(\lambda y_1 + (1-\lambda)\tilde{y}_0) \\
&= \lambda(y_2 + (1-\lambda)y_1) + (1-\lambda)^2\tilde{y}_0 \\
\tilde{y}_3 &= \lambda(y_3 + (1-\lambda)y_2 + (1-\lambda)^2 y_1) + (1-\lambda)^3\tilde{y}_0 \\
&\vdots \\
\tilde{y}_T &= \lambda(y_T + (1-\lambda)y_{T-1} + \cdots + (1-\lambda)^{T-1}y_1) + (1-\lambda)^T\tilde{y}_0,
\end{aligned}
$$

which means that as T gets large and hence $(1-\lambda)^T$ gets small, the contribution of $\tilde{y}_0$ to $\tilde{y}_T$ becomes negligible. Thus for large data sets, the estimation of $\tilde{y}_0$ has little relevance. Nevertheless, two commonly used estimates for $\tilde{y}_0$ are the following.

1. Set $\tilde{y}_0 = y_1$. If the changes in the process are expected to occur early and fast, this choice for the starting value for $\tilde{y}_T$ is reasonable.
2. Take the average of the available data or a subset of the available data, $\bar{y}$, and set $\tilde{y}_0 = \bar{y}$. If the process is at least at the beginning locally constant, this starting value may be preferred.

4.2.2 The Value of λ

In Figures 4.5 and 4.6, respectively, we have two simple exponential smoothers for the Dow Jones Index data with $\lambda = 0.2$ and $\lambda = 0.4$. It can be seen that in the latter the smoothed values follow the original observations more closely. In general, as λ gets closer to 1, and more emphasis is put on the last observation, the smoothed values will approach the original observations. Two extreme cases will be when $\lambda = 0$ and $\lambda = 1$. In the former, the smoothed values will all be equal to a constant, namely, y_0.

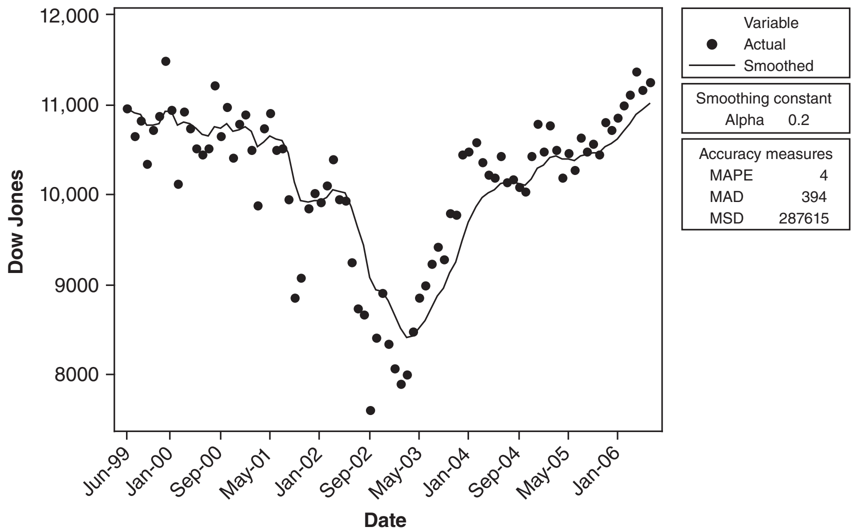


FIGURE 4.5 The Dow Jones Index from June 1999 to June 2006 with first-order exponential smoothing with $\lambda = 0.2$.

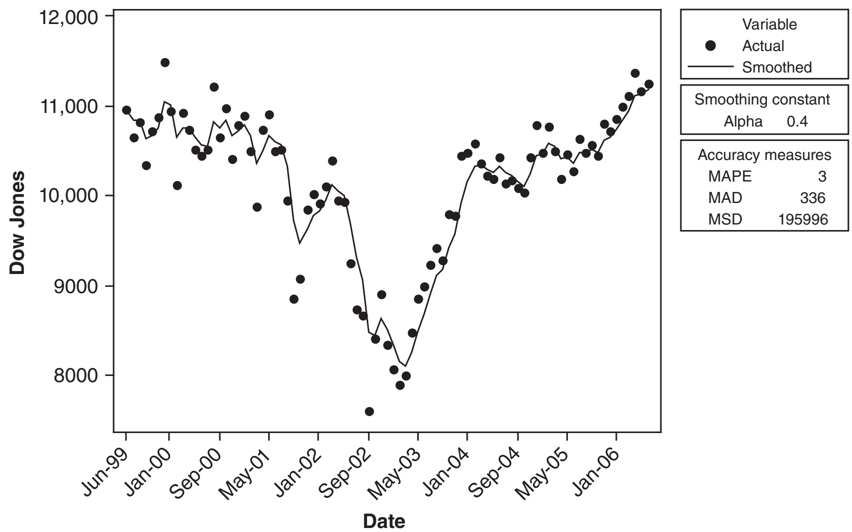


FIGURE 4.6 The Dow Jones Index from June 1999 to June 2006 with first-order exponential smoothing with $\lambda = 0.4$.

We can think of the constant line as the "smoothest" version of whatever pattern the actual time series follows. For $\lambda = 1$, we have $\tilde{y}_T = y_T$ and this will represent the "least" smoothed (or unsmoothed) version of the original time series. We can accordingly expect the variance of the simple exponential smoother to vary between 0 and the variance of the original time series based on the choice of λ. Note that under the independence and constant variance assumptions we have

$$\begin{aligned}
\mathrm{Var}(\tilde{y}_T) &= \mathrm{Var}\left(\lambda \sum_{t=0}^{\infty} (1-\lambda)^t y_{T-t}\right) \\
&= \lambda^2 \sum_{t=0}^{\infty} (1-\lambda)^{2t} \mathrm{Var}\,(y_{T-t}) \\
&= \lambda^2 \sum_{t=0}^{\infty} (1-\lambda)^{2t} \mathrm{Var}(y_T) \\
&= \mathrm{Var}(y_T)\lambda^2 \sum_{t=0}^{\infty} (1-\lambda)^{2t} \\
&= \frac{\lambda}{(2-\lambda)} \mathrm{Var}(y_T).
\end{aligned} \tag{4.8}$$

Thus the question will be how much smoothing is needed. In the literature, λ values between 0.1 and 0.4 are often recommended and do indeed perform well in practice. A more rigorous method of finding the right λ value will be discussed in Section 4.6.1.

Example 4.1 Consider the Dow Jones Index from June 1999 to June 2006 given in Figure 4.3. For first-order exponential smoothing we would need to address two issues as stated in the previous sections: how to pick the initial value y_0 and the smoothing constant λ. Following the recommendation in Section 4.2.2, we will consider the smoothing constants 0.2 and 0.4. As for the initial value, we will consider the first recommendation in Section 4.2.1 and set $\tilde{y}_0 = y_1$. Figures 4.5 and 4.6 show the smoothed and actual data obtained from Minitab with smoothing constants 0.2 and 0.4, respectively.

Note that Minitab reports several measures of accuracy; MAPE, MAD, and MSD. Mean absolute percentage error (MAPE) is the average absolute percentage change between the predicted value that is $\tilde{y}_{t-1}$ for a one-step-ahead forecast and the true value, given as

$$\mathrm{MAPE} = \frac{\sum_{t=1}^{T} |(y_t - \tilde{y}_{t-1})/y_t|}{T} \times 100 \quad (y_t \neq 0).$$

Mean absolute deviation (MAD) is the average absolute difference between the predicted and the true values, given as

$$\mathrm{MAD} = \frac{\sum_{t=1}^{T} |(y_t - \tilde{y}_{t-1})|}{T}.$$

Mean squared deviation (MSD) is the average squared difference between the predicted and the true values, given as

$$\mathrm{MSD} = \frac{\sum_{t=1}^{T} (y_t - \tilde{y}_{t-1})^2}{T}.$$

It should also be noted that the smoothed data with $\lambda = 0.4$ follows the actual data closer. However, in both cases, when there is an apparent linear trend in the data (e.g., from February 2003 to February 2004) the smoothed values consistently underestimate the actual data. We will discuss this issue in greater detail in Section 4.3.

As an alternative estimate for the initial value, we can also use the average of the data between June 1999 and June 2001, since during this period the time series data appear to be stable. Figures 4.7 and 4.8 show

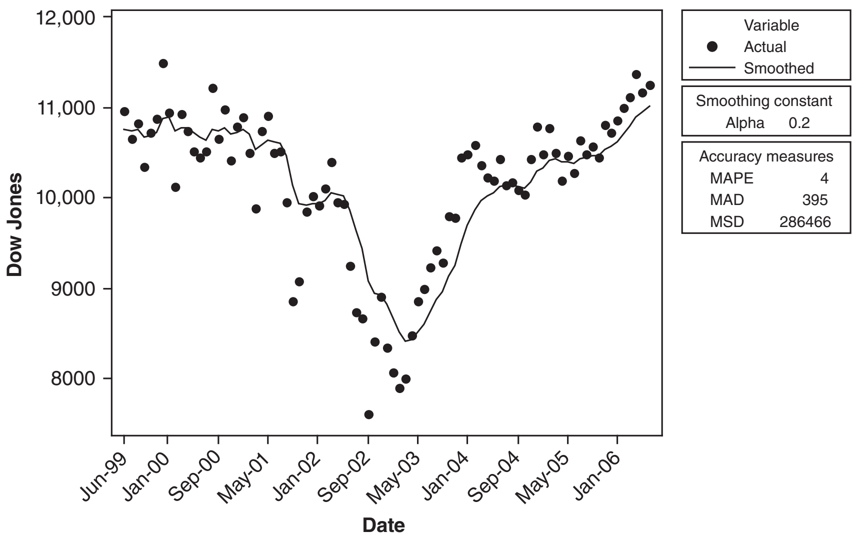


FIGURE 4.7 The Dow Jones Index from June 1999 to June 2006 with first-order exponential smoothing with $\lambda = 0.2$ and $\tilde{y}_0 = (\sum_{t=1}^{25} y_t/25)$ (i.e., initial value equal to the average of the first 25 observations).

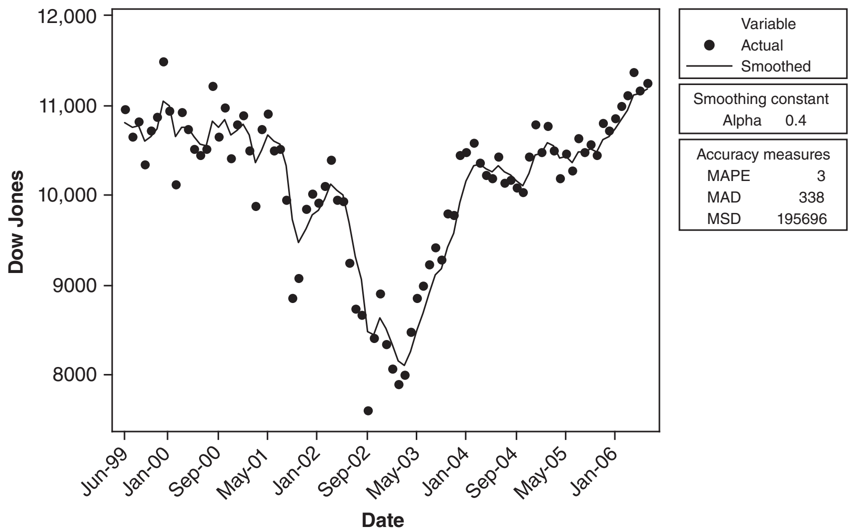


FIGURE 4.8 The Dow Jones Index from June 1999 to June 2006 with first-order exponential smoothing with $\lambda = 0.4$ and $\tilde{y}_0 = (\sum_{t=1}^{25} y_t/25)$ (i.e., initial value equal to the average of the first 25 observations).

the single exponential smoothing with the initial value equal to the average of the first 25 observations corresponding to the period between June 1999 and June 2001. Note that the choice of the initial value has very little effect on the smoothed values as time goes on.

4.3 MODELING TIME SERIES DATA

In Section 4.1, we considered the constant process where the time series data are expected to vary around a constant level with random fluctuations, which are usually characterized by uncorrelated errors with mean 0 and constant variance σ_ε^2. In fact the constant process represents a very special case in a more general set of models often used in modeling time series data as a function of time. The general class of models can be represented as

$$y_t = f(t; \beta) + \varepsilon_t, \tag{4.9}$$

where β is the vector of unknown parameters and ε_t represents the uncorrelated errors. Thus as a member of this general class of models, the constant process can be represented as

$$y_t = \beta_0 + \varepsilon_t, \tag{4.10}$$

where β_0 is equal to μ in Eq. (4.1). We have seen in Chapter 3 how to estimate and make inferences about the regression coefficients. The same principles apply to the class of models in Eq. (4.9). However, we have seen in Section 4.1 that the least squares estimates for β_0 at any given time T will be very slow to react to changes in the level of the process. For that, we suggested to use either the moving average or simple exponential smoothing.

As mentioned earlier, smoothing techniques are effective in illustrating the underlying pattern in the time series data. We have so far focused particularly on exponential smoothing techniques. For the class of models given in Eq. (4.9), we can find another use for the exponential smoothers: model estimation. Indeed for the constant process, we can see the simple exponential smoother as the estimate of the process level, or in regards to Eq. (4.10) an estimate of β_0. To show this in greater detail we need to introduce the sum of weighted squared errors for the constant process. Remember that the sum of squared errors for the constant process is given by

$$SS_E = \sum_{t=1}^{T} (y_t - \mu)^2.$$

If we argue that not all observations should have equal influence on the sum and decide to introduce a string of weights that are geometrically decreasing in time, the sum of squared errors becomes

$$SS_E^* = \sum_{t=0}^{T-1} \theta^t (y_{T-1} - \beta_0)^2, \tag{4.11}$$

where $|\theta|$ 1 < 1. To find the least squares estimate for β_0, we take the derivative of Eq. (4.11) with respect to β_0 and set it to zero:

$$\left.\frac{dSS_E^*}{d\beta_0}\right|_{\beta_0} = -2\sum_{t=0}^{T-1} \theta^t (y_{T-t} - \hat{\beta}_0) = 0. \tag{4.12}$$

The solution to Eq. (4.12), $\hat{\beta}_0$, which is the least squares estimate of β_0, is

$$\hat{\beta}_0 \sum_{t=0}^{T-1} \theta^t = \sum_{t=0}^{T-1} \theta^t y_{T-t}. \tag{4.13}$$

From Eq. (4.4), we have

$$\hat{\beta}_0 = \frac{1-\theta}{1-\theta^T} \sum_{t=0}^{T-1} \theta^t y_{T-t}. \tag{4.14}$$

Once again for large T, θ^T goes to zero. We then have

$$\hat{\beta}_0 = (1-\theta) \sum_{t=0}^{T-1} \theta^t y_{T-t}. \tag{4.15}$$

We can see from Eqs. (4.5) and (4.15) that $\beta_0 = \tilde{y}_T$. Thus the simple exponential smoothing procedure does in fact provide a weighted least squares estimate of β_0 in the constant process with weights that are exponentially decreasing in time.

Now we return to our general class of models given in Eq. (4.9) and note that $f(t; \beta)$ can in fact be any function of t. For practical purposes it is usually more convenient to consider the polynomial family for nonseasonal time series. For seasonal time series, we will consider other forms of $f(t; \beta)$ that fit the data and exhibit a certain periodicity better. In the polynomial family, the constant process is indeed the simplest model we can consider. We will now consider the next obvious choice: the linear trend model.

4.4 SECOND-ORDER EXPONENTIAL SMOOTHING

We will now return to our Dow Jones Index data but consider only the subset of the data from February 2003 to February 2004 as given in Figure 4.9. Evidently for that particular time period it was a bullish market and correspondingly the Dow Jones Index exhibits an upward linear trend as indicated with the dashed line.

For this time period, an appropriate model in time from the polynomial family should be the linear trend model given as

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t, \tag{4.16}$$

where the ε_t is once again assumed to be uncorrelated with mean 0 and constant variance $\sigma^2{}_\varepsilon$. Based on what we have learned so far, we may attempt to smooth/model this linear trend using the simple exponential smoothing procedure. The actual and fitted values for the simple exponential smoothing procedure are given in Figure 4.10. For the exponential

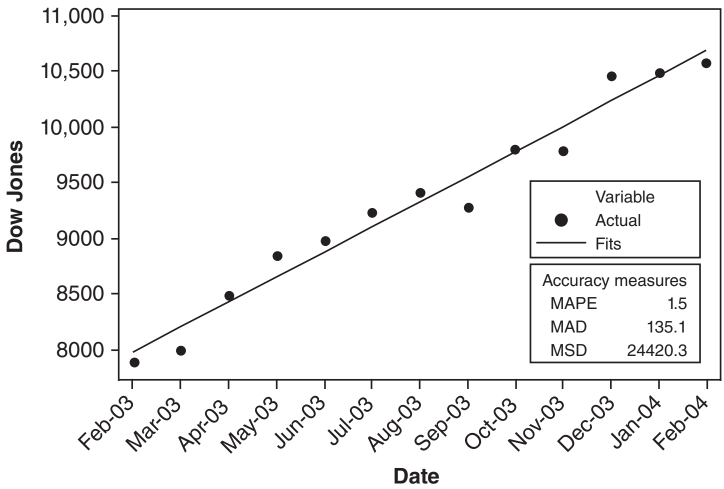


FIGURE 4.9 The Dow Jones Index from February 2003 to February 2004.

smoother, without any loss of generality, we used $\tilde{y}_0 = y_1$ and $\lambda = 0.3$. From Figure 4.10, we can see that while the simple exponential smoother was to some extent able to capture the slope of the linear trend, it also exhibits some bias. That is, the fitted values based on the exponential smoother are consistently underestimating the actual data. More interestingly, the amount of underestimation is more or less constant for all observations.

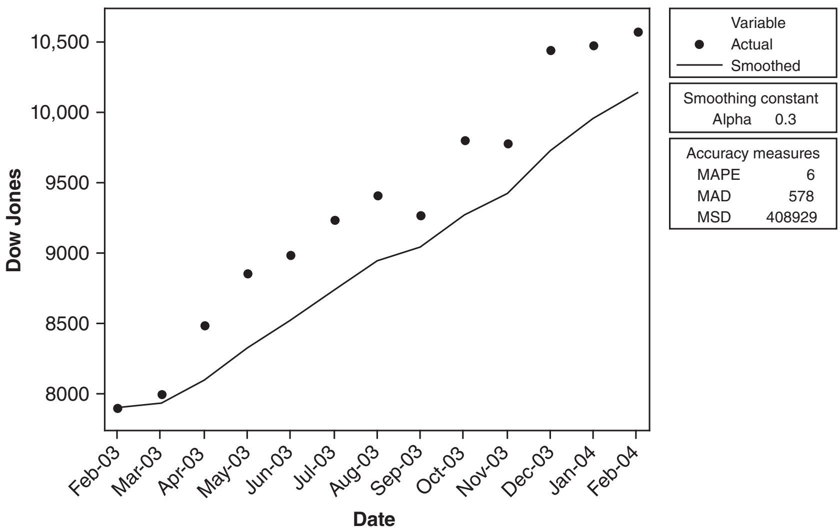


FIGURE 4.10 The Dow Jones Index from February 2003 to February 2004 with simple exponential smoothing with $\lambda = 0.3$.

In fact similar behavior for the simple exponential smoother can be observed in Figure 4.5 for the entire data from June 1999 to June 2006. Whenever the data exhibit a linear trend, the simple exponential smoother seems to over- or underestimates the actual data consistently. To further explore this, we will consider the expected value of $\tilde{y}_T$,

$$
\begin{aligned}
E(\tilde{y}_T) &= E\left(\lambda \sum_{t=0}^{\infty} (1-\lambda)^t y_{T-t}\right) \\
&= \lambda \sum_{t=0}^{\infty} (1-\lambda)^t E(y_{T-t}).
\end{aligned}
$$

For the linear trend model in Eq. (4.16), $E\,(y_t) = \beta_0 + \beta_1 t$. So we have

$$
\begin{aligned}
E(\tilde{y}_T) &= \lambda \sum_{t=0}^{\infty} (1-\lambda)^t (\beta_0 + \beta_1 (T-t)) \\
&= \lambda \sum_{t=0}^{\infty} (1-\lambda)^t (\beta_0 + \beta_1 T) - \lambda \sum_{t=0}^{\infty} (1-\lambda)^t (\beta_1 t) \\
&= (\beta_0 + \beta_1 T)\lambda \sum_{t=0}^{\infty} (1-\lambda)^t - \lambda\beta_1 \sum_{t=0}^{\infty} (1-\lambda)^t t.
\end{aligned}
$$

But for the infinite sums we have

$$
\sum_{t=0}^{\infty} (1-\lambda)^t = \frac{1}{1-(1-\lambda)} = \frac{1}{\lambda} \text{ and } \sum_{t=0}^{\infty} (1-\lambda)^t t = \frac{1-\lambda}{\lambda^2}.
$$

Hence the expected value of the simple exponential smoother for the linear trend model is

$$
\begin{aligned}
E(\tilde{y}_T) &= (\beta_0 + \beta_1 T) - \frac{1-\lambda}{\lambda}\beta_1 \\
&= E(y_T) - \frac{1-\lambda}{\lambda}\beta_1.
\end{aligned}
\tag{4.17}
$$

This means that the simple exponential smoother is a biased estimator for the linear trend model and the amount of bias is $-[(1-\lambda)/\lambda]\beta_1$. This indeed explains the underestimation in Figure 4.10. One solution will be to use a large λ value since $(1-\lambda)/\lambda \to 0$ as $\lambda \to 1$. In Figure 4.11, we show two simple exponential smoothers with $\lambda = 0.3$ and $\lambda = 0.99$. It can be

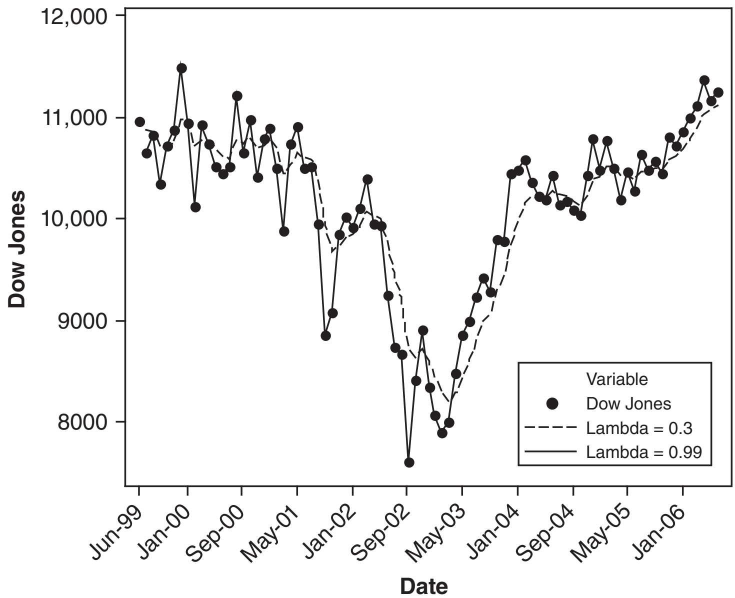


FIGURE 4.11 The Dow Jones Index from June 1999 to June 2006 using exponential smoothing with $\lambda = 0.3$ and 0.99.

seen that the latter does a better job in capturing the linear trend. However, it should also be noted that as the smoother with $\lambda = 0.99$ follows the actual observations very closely, it fails to smooth out the constant pattern during the first 2 years of the data. A method based on adaptive updating of the discount factor, λ, following the changes in the process is given in Section 4.6.4. In this section to model a linear trend model we will instead introduce the second-order exponential smoothing by applying simple exponential smoothing on $\tilde{y}_T$ as

$$\tilde{y}_T^{(2)} = \lambda \tilde{y}_T^{(1)} + (1 - \lambda)\tilde{y}_{T-1}^{(2)}, \tag{4.18}$$

where $\tilde{y}_T^{(1)}$ and $\tilde{y}_T^{(2)}$ denote the first- and second-order smoothed exponentials, respectively. Of course, in Eq. (4.18) we can use a different λ than in Eq. (4.7). However, for the derivations that follow, we will assume that the same λ is used in the calculations of both $\tilde{y}_T^{(1)}$ and $\tilde{y}_T^{(2)}$.

From Eq. (4.17), we can see that the first-order exponential smoother introduces bias in estimating a linear trend. It can also be seen in Figure 4.7 that the first-order exponential smoother for the linear trend model exhibits a linear trend as well. Hence the second-order smoother—that is, a first-order exponential smoother of the original first-order exponential smoother—should also have a bias. We can represent this as

$$E\left(\tilde{y}_T^{(2)}\right) = E\left(\tilde{y}_T^{(1)}\right) - \frac{1-\lambda}{\lambda}\beta_1. \tag{4.19}$$

From Eq. (4.19), an estimate for β_1 at time T is

$$\hat{\beta}_{1,T} = \frac{\lambda}{1-\lambda}\left(\tilde{y}_T^1 - \tilde{y}_T^2\right) \tag{4.20}$$

and for an estimate of β_0 at time T, we have from Eq. (4.17)

$$\begin{aligned} \tilde{y}_T^{(1)} &= \left(\hat{\beta}_{0,T} + \hat{\beta}_{1,T}T\right) - \frac{1-\lambda}{\lambda}\hat{\beta}_{1,T} \\ \Rightarrow \hat{\beta}_{0,T} &= \tilde{y}_T^{(1)} - T\hat{\beta}_{1,T} + \frac{1-\lambda}{\lambda}\hat{\beta}_{1,T}. \end{aligned} \tag{4.21}$$

In terms of the first- and second-order exponential smoothers, we have

$$\begin{aligned} \hat{\beta}_{0,T} &= \tilde{y}_T^{(1)} - T\frac{\lambda}{1-\lambda}\left(\tilde{y}_T^{(1)} - \tilde{y}_T^{(2)}\right) + \frac{1-\lambda}{\lambda}\left(\frac{\lambda}{1-\lambda}\left(\tilde{y}_T^{(1)} - \tilde{y}_T^{(2)}\right)\right) \\ &= \tilde{y}_T^{(1)} - T\frac{\lambda}{1-\lambda}\left(\tilde{y}_T^{(1)} - \tilde{y}_T^{(2)}\right) + \left(\tilde{y}_T^{(1)} - \tilde{y}_T^{(2)}\right) \\ &= \left(2 - T\frac{\lambda}{1-\lambda}\right)\tilde{y}_T^{(1)} - \left(1 - T\frac{\lambda}{1-\lambda}\right)\tilde{y}_T^{(2)}. \end{aligned} \tag{4.22}$$

Finally, combining Eq. (4.20) and (4.22), we have a predictor for y_T as

$$\begin{aligned} \tilde{y}_T &= \hat{\beta}_{0,T} + \hat{\beta}_{1,T}T \\ &= 2\tilde{y}_T^{(1)} - \tilde{y}_T^{(2)}. \end{aligned} \tag{4.23}$$

It can easily be shown that $\hat{y}_T$ is an unbiased predictor of y_T. In Figure 4.12, we use Eq. (4.23) to estimate the Dow Jones Index from February 2003 to February 2004. From Figures 4.10 and 4.12, we can clearly see that the second-order exponential smoother is doing a much better job in modeling the linear trend compared to the simple exponential smoother.

As in the simple exponential smoothing, we have the same two issues to deal with: initial values for the smoothers and the discount factors. The

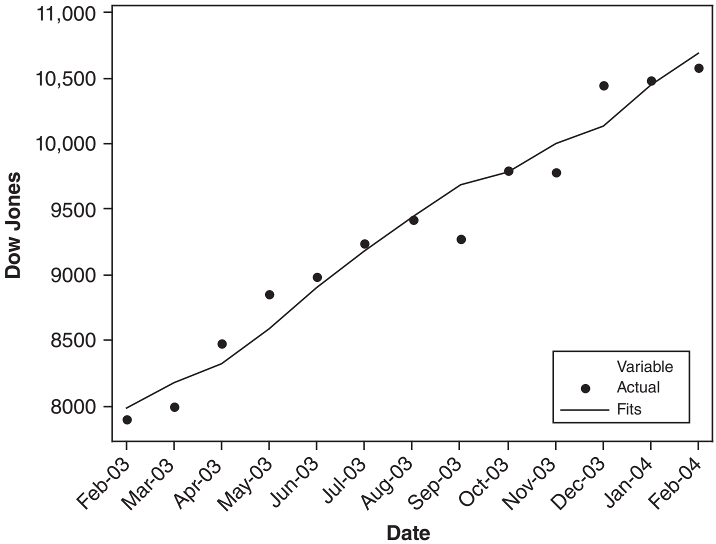


FIGURE 4.12 The Dow Jones Index from February 2003 to February 2004 with second-order exponential smoother with discount factor of 0.3.

latter will be discussed in Section 4.6.1. For the former we will combine Eqs. (4.17) and (4.19) as the following:

$$
\begin{aligned}
\tilde{y}_0^{(1)} &= \hat{\beta}_{0,0} - \frac{1-\lambda}{\lambda}\hat{\beta}_{1,0} \\
\tilde{y}_0^{(2)} &= \hat{\beta}_{0,0} - 2\left(\frac{1-\lambda}{\lambda}\right)\hat{\beta}_{1,0}.
\end{aligned}
\tag{4.24}
$$

The initial estimates of the model parameters are usually obtained by fitting the linear trend model to the entire or a subset of the available data. The least squares estimates of the parameter estimates are then used for $\hat{\beta}_{0,0}$ and $\hat{\beta}_{1,0}$.

Example 4.2 Consider the US Consumer Price Index (CPI) from January 1995 to December 2004 in Table 4.2. Figure 4.13 clearly shows that the data exhibits a linear trend. To smooth the data, following the recommendation in Section 4.2, we can use single exponential smoothing with $\lambda = 0.3$ as given in Figure 4.14.

As we expected, the exponential smoother does a very good job in capturing the general trend in the data and provides a less jittery (smooth) version of it. However, we also notice that the smoothed values are

TABLE 4.2 Consumer Price Index from January 1995 to December 2004

Month-Year	CPI	Month-Year	CPI	Month-Year	CPI	Month-Year	CPI	Month-Year	CPI
Jan-1995	150.3	Jan-1997	159.1	Jan-1999	164.3	Jan-2001	175.1	Jan-2003	181.7
Feb-1995	150.9	Feb-1997	159.6	Feb-1999	164.5	Feb-2001	175.8	Feb-2003	183.1
Mar-1995	151.4	Mar-1997	160	Mar-1999	165	Mar-2001	176.2	Mar-2003	184.2
Apr-1995	151.9	Apr-1997	160.2	Apr-1999	166.2	Apr-2001	176.9	Apr-2003	183.8
May-1995	152.2	May-1997	160.1	May-1999	166.2	May-2001	177.7	May-2003	183.5
Jun-1995	152.5	Jun-1997	160.3	Jun-1999	166.2	Jun-2001	178	Jun-2003	183.7
Jul-1995	152.5	Jul-1997	160.5	Jul-1999	166.7	Jul-2001	177.5	Jul-2003	183.9
Aug-1995	152.9	Aug-1997	160.8	Aug-1999	167.1	Aug-2001	177.5	Aug-2003	184.6
Sep-1995	153.2	Sep-1997	161.2	Sep-1999	167.9	Sep-2001	178.3	Sep-2003	185.2
Oct-1995	153.7	Oct-1997	161.6	Oct-1999	168.2	Oct-2001	177.7	Oct-2003	185
Nov-1995	153.6	Nov-1997	161.5	Nov-1999	168.3	Nov-2001	177.4	Nov-2003	184.5
Dec-1995	153.5	Dec-1997	161.3	Dec-1999	168.3	Dec-2001	176.7	Dec-2003	184.3
Jan-1996	154.4	Jan-1998	161.6	Jan-2000	168.8	Jan-2002	177.1	Jan-2004	185.2
Feb-1996	154.9	Feb-1998	161.9	Feb-2000	169.8	Feb-2002	177.8	Feb-2004	186.2
Mar-1996	155.7	Mar-1998	162.2	Mar-2000	171.2	Mar-2002	178.8	Mar-2004	187.4
Apr-1996	156.3	Apr-1998	162.5	Apr-2000	171.3	Apr-2002	179.8	Apr-2004	188
May-1996	156.6	May-1998	162.8	May-2000	171.5	May-2002	179.8	May-2004	189.1
Jun-1996	156.7	Jun-1998	163	Jun-2000	172.4	Jun-2002	179.9	Jun-2004	189.7
Jul-1996	157	Jul-1998	163.2	Jul-2000	172.8	Jul-2002	180.1	Jul-2004	189.4
Aug-1996	157.3	Aug-1998	163.4	Aug-2000	172.8	Aug-2002	180.7	Aug-2004	189.5
Sep-1996	157.8	Sep-1998	163.6	Sep-2000	173.7	Sep-2002	181	Sep-2004	189.9
Oct-1996	158.3	Oct-1998	164	Oct-2000	174	Oct-2002	181.3	Oct-2004	190.9
Nov-1996	158.6	Nov-1998	164	Nov-2000	174.1	Nov-2002	181.3	Nov-2004	191
Dec-1996	158.6	Dec-1998	163.9	Dec-2000	174	Dec-2002	180.9	Dec-2004	190.3

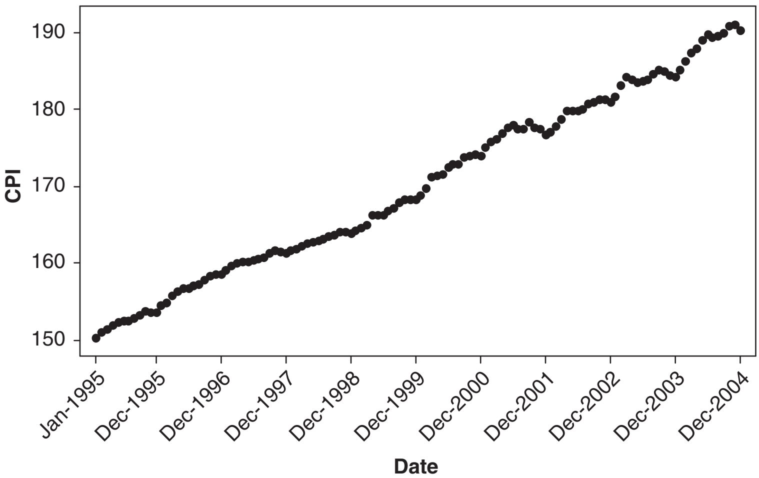


FIGURE 4.13 US Consumer Price Index from January 1995 to December 2004.

consistently below the actual values. Hence there is an apparent bias in our smoothing. To fix this problem we have two choices: use a bigger λ or **second-order** exponential smoothing. The former will lead to less smooth estimates and hence defeat the purpose. For the latter, however, we can use $\lambda = 0.3$ to calculate and $\tilde{y}_T^{(1)}$ and $\tilde{y}_T^{(2)}$ as given in Table 4.3.

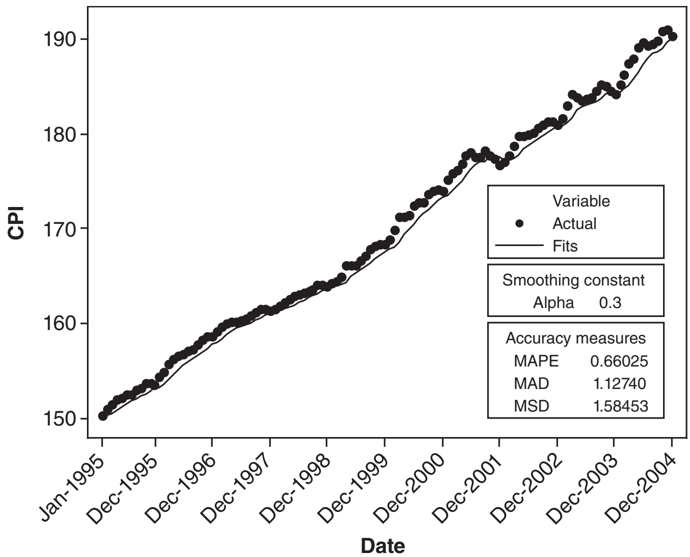


FIGURE 4.14 Single exponential smoothing of the US Consumer Price Index (with $\tilde{y}_0 = y_1$).

TABLE 4.3 Second-Order Exponential Smoothing of the US Consumer Price Index (with $\lambda = 0.3, \tilde{y}_0^{(1)} = y_1$, and $\tilde{y}_0^{(2)} = \tilde{y}_0^{(1)}$

Date	y_t	$\tilde{y}_T^{(1)}$	$\tilde{y}_T^{(2)}$	$\hat{y}_T = 2\tilde{y}_T^{(1)} - \tilde{y}_T^{(2)}$
Jan-1995	150.3	150.300	150.300	150.300
Feb-1995	150.9	150.480	150.354	150.606
Mar-1995	151.4	150.756	150.475	151.037
Apr-1995	151.9	151.099	150.662	151.536
May-1995	152.2	151.429	150.892	151.967
Nov-2004	191.0	190.041	188.976	191.106
Dec-2004	190.3	190.119	189.319	190.919

Note that we used $\tilde{y}_0^{(1)} = y_1$, and $\tilde{y}_0^{(2)} = \tilde{y}_0^{(1)}$ as the initial values of $\tilde{y}_T^{(1)}$ and $\tilde{y}_T^{(2)}$. A more rigorous approach would involve fitting a linear regression model in time to the available data that give

$$\begin{aligned}\hat{y}_t &= \hat{\beta}_{0,T} + \hat{\beta}_{1,T} t \\ &= 149.89 + 0.33t,\end{aligned}$$

where t goes from 1 to 120. Then from Eq. (4.24) we have

$$\begin{aligned}\tilde{y}_0^{(1)} &= \hat{\beta}_{0,0} - \frac{1-\lambda}{\lambda}\hat{\beta}_{1,0} \\ &= 149.89 - \frac{1-0.3}{0.3}0.33 = 146.22 \\ \tilde{y}_0^{(2)} &= \hat{\beta}_{0,0} - 2\left(\frac{1-\lambda}{\lambda}\right)\hat{\beta}_{1,0} \\ &= 149.89 - 2\left(\frac{1-0.3}{0.3}\right)0.33 = 142.56.\end{aligned}$$

Figure 4.15 shows the second-order exponential smoothing of the CPI. As we can see, the second-order exponential smoothing not only captures the trend in the data but also does not exhibit any bias.

The calculations for the second-order smoothing for the CPI data are performed using Minitab. We first obtained the first-order exponential smoother for the CPI, $\tilde{y}_T^{(1)}$, using $\lambda = 0.3$ and $\tilde{y}_0^{(1)} = y_1$. Then we obtained $\tilde{y}_T^{(2)}$ by taking the first-order exponential smoother $\tilde{y}_T^{(1)}$ using $\lambda = 0.3$ and $\tilde{y}_0^{(2)} = \tilde{y}_1^{(1)}$. Then using Eq. (4.23) we have $\hat{y}_T = 2\tilde{y}_T^{(1)} - \tilde{y}_T^{(2)}$.

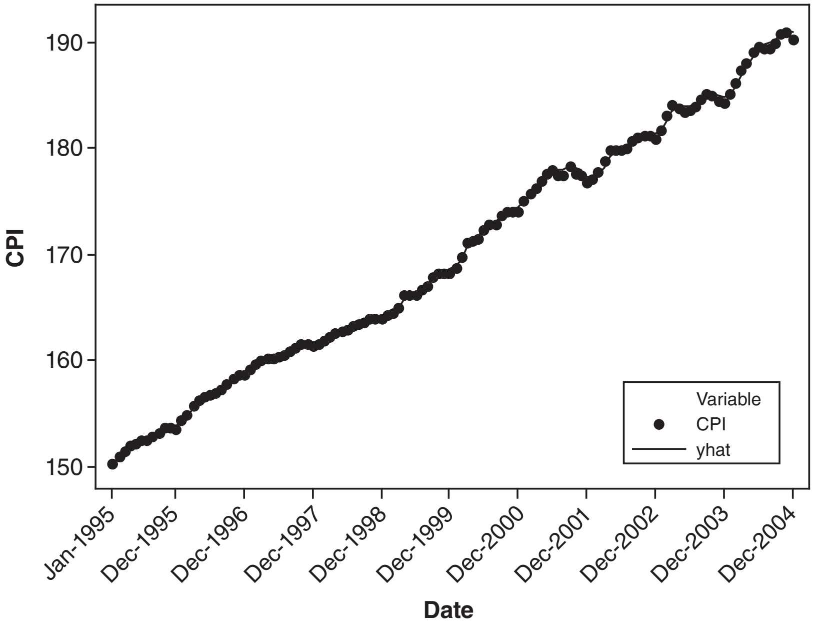


FIGURE 4.15 Second-order exponential smoothing of the US Consumer Price Index (with $\lambda = 0.3$, $\tilde{y}_0^{(1)} = y_1$, and $\tilde{y}_0^{(2)} = \tilde{y}_1^{(1)}$).

The "Double Exponential Smoothing" option available in Minitab is a slightly different approach based on Holt's method (Holt, 1957). This method divides the time series data into two components: the level, L_t, and the trend, T_t. These two components can be calculated from

$$L_t = \alpha y_t + (1 - \alpha)(L_{t-1} + T_{t-1})$$
$$T_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1}$$

Hence for a given set of α and γ, these two components are calculated and L_t is used to obtain the double exponential smoothing of the data at time t. Furthermore, the sum of the level and trend components at time t can be used as the one-step-ahead ($t + 1$) forecast. Figure 4.16 shows the actual and smoothed data using the double exponential smoothing option in Minitab with $\alpha = 0.3$ and $\gamma = 0.3$.

In general, the initial values for the level and the trend terms can be obtained by fitting a linear regression model to the CPI data with time as

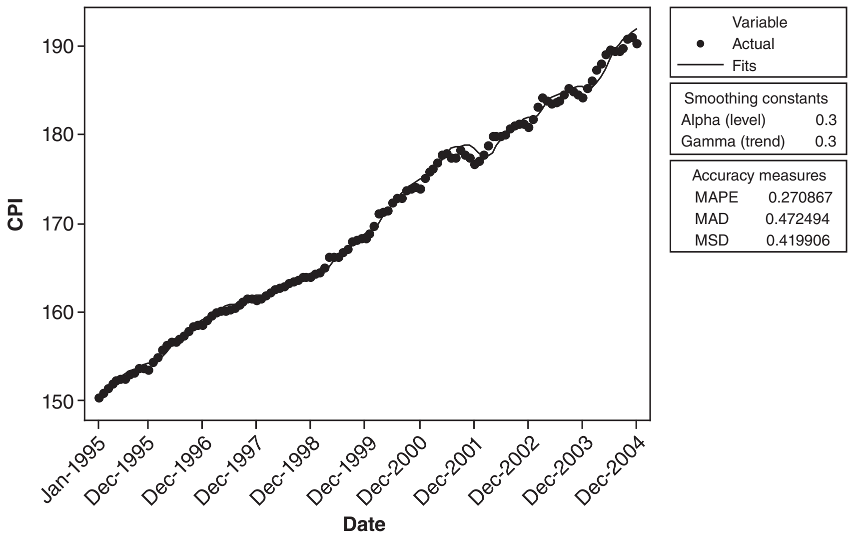


FIGURE 4.16 The double exponential smoothing of the US Consumer Price Index (with $\alpha = 0.3$ and $\gamma = 0.3$).

the regressor. Then the intercept and the slope can be used as the initial values of L_t and T_t respectively.

Example 4.3 For the Dow Jones Index data, we observed that first-order exponential smoothing with low values of λ showed some bias when there were linear trends in the data. We may therefore decide to use the second-order exponential smoothing approach for this data as shown in Figure 4.17. Note that the bias present with first-order exponential smoothing has been eliminated. The calculations for second-order exponential smoothing for the Dow Jones Index are given in Table 4.4.

4.5 HIGHER-ORDER EXPONENTIAL SMOOTHING

So far we have discussed the use of exponential smoothers in estimating the constant and linear trend models. For the former we employed the **simple** or **first-order** exponential smoother and for the latter the **second-order** exponential smoother. It can further be shown that for the general nth-degree polynomial model of the form

$$y_t = \beta_0 + \beta_1 t + \frac{\beta_2}{2!}t^2 + \cdots + \frac{\beta_n}{n!}t^n + \varepsilon_t, \qquad (4.25)$$

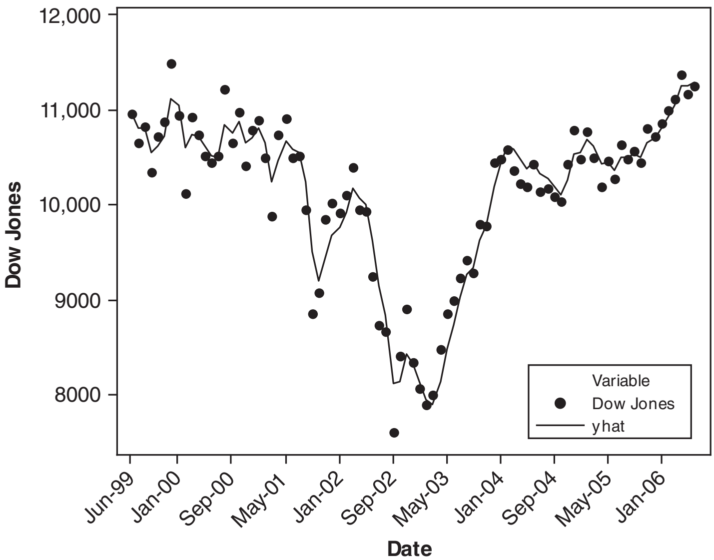


FIGURE 4.17 The second-order exponential smoothing of the Dow Jones Index (with $\lambda = 0.3$, $\tilde{y}_0^{(1)} = y_1$, and $\tilde{y}_0^{(2)} = \tilde{y}_1^{(1)}$).

where the ε_t is assumed to be independent with mean 0 and constant variance $\sigma_\varepsilon{}^2$, we employ $(n + 1)$-order exponential smoothers

$$
\begin{aligned}
\tilde{y}_T^{(2)} &= \lambda y_T + (1 - \lambda)\tilde{y}_{T-1}^{(1)} \\
\tilde{y}_T^{(2)} &= \lambda \tilde{y}_T^{(1)} + (1 - \lambda)\tilde{y}_{T-1}^{(2)} \\
&\vdots \\
\tilde{y}_T^{(n)} &= \lambda \tilde{y}_T^{(n-1)} + (1 - \lambda)\tilde{y}_{T-1}^{(n)}
\end{aligned}
$$

TABLE 4.4 Second-Order Exponential Smoothing of the Dow Jones Index (with $\lambda = 0.3$, $\tilde{y}_0^{(1)} = y_1$, and $\tilde{y}_0^{(2)} = \tilde{y}_1^{(1)}$)

Date	$\tilde{y}_t$	$\tilde{y}_T^1$	$\tilde{y}_T^2$	$\hat{y}_T = 2\tilde{y}_T^{(1)} - \tilde{y}_T^{(2)}$
Jun-1999	10,970.8	10,970.8	10,970.8	10,970.8
Jul-1999	10,655.2	10,876.1	10,942.4	10,809.8
Aug-1999	10,829.3	10,862.1	10,918.3	10,805.8
Sep-1999	10,337.0	10,704.6	10,854.2	10,554.9
Oct-1999	10,729.9	10,712.2	10,811.6	10,612.7
May-2006	11,168.3	11,069.4	10,886.5	11,252.3
Jun-2006	11,247.9	11,123.0	10,957.4	11,288.5

to estimate the model parameters. For even the quadratic model (second-degree polynomial), the calculations get quite complicated. Refer to Montgomery et al. (1990), Brown (1963), and Abraham and Ledolter (1983) for the solutions to higher-order exponential smoothing problems. If a high-order polynomial does seem to be required for the time series, the autoregressive integrated moving average (ARIMA) models and techniques discussed in Chapter 5 can instead be considered.

4.6 FORECASTING

We have so far considered exponential smoothing techniques as either visual aids to point out the underlying patterns in the time series data or to estimate the model parameters for the class of models given in Eq. (4.9). The latter brings up yet another use of exponential smoothing—forecasting future observations. At time T, we may wish to forecast the observation in the next time unit, $T + 1$, or further into the future. For that, we will denote the τ-step-ahead forecast made at time T as $\hat{y}_{T+\tau}(T)$. In the next two sections and without any loss of generality, we will once again consider first- and second-order exponential smoothers as examples for forecasting time series data from the constant and linear trend processes.

4.6.1 Constant Process

In Section 4.2 we discussed first-order exponential smoothing for the constant process in Eq. (4.1) as

$$\tilde{y}_T = \lambda y_T + (1 - \lambda)\tilde{y}_{T-1}.$$

In Section 4.3 we further showed that the constant level in Eq. (4.1), β_0, can be estimated by $\tilde{y}_T$. Since the constant model consists of two parts—β_0 that can be estimated by the first-order exponential smoother and the random error that cannot be predicted—our forecast for the future observation is simply equal to the current value of the exponential smoother

$$\hat{y}_{T+\tau}(T) = \tilde{y}_T = \tilde{y}_T. \tag{4.26}$$

Please note that, for the constant process, the forecast in Eq. (4.26) is the same for all future values. Since there may be changes in the level of the constant process, forecasting all future observations with the same value will most likely be misleading. However, as we start accumulating more observations, we can update our forecast. For example, if the data at $T+1$ become available, our forecast for the future observations becomes

$$\tilde{y}_{T+1} = \lambda y_{T+1} + (1-\lambda)\tilde{y}_T$$

or

$$\hat{y}_{T+1+\tau}(T+1) = \lambda y_{T+1} + (1-\lambda)\hat{y}_{T+\tau}(T) \tag{4.27}$$

We can rewrite Eq. (4.27) for $\tau = 1$ as

$$\begin{aligned} \hat{y}_{T+2}(T+1) &= \hat{y}_{T+1}(T) + \lambda(y_{T+1} - \hat{y}_{T+1}(T)) \\ &= \hat{y}_{T+1}(T) + \lambda e_{T+1}(1), \end{aligned} \tag{4.28}$$

where $e_{T+1}(1) = y_{T+1} - \hat{y}_{T+1}(T)$ is called the one-step-ahead forecast or prediction error. The interpretation of Eq. (4.28) makes it easier to understand the forecasting process using exponential smoothing: our forecast for the next observation is simply our previous forecast for the current observation plus a fraction of the forecast error we made in forecasting the current observation. The fraction in this summation is determined by λ. Hence how fast our forecast will react to the forecast error depends on the discount factor. A large discount factor will lead to fast reaction to the forecast error but it may also make our forecast react fast to random fluctuations. This once again brings up the issue of the choice of the discount factor.

Choice of λ We will define the sum of the squared one-step-ahead forecast errors as

$$SS_E(\lambda) = \sum_{t=1}^{T} e_t^2(1). \tag{4.29}$$

For a given historic data, we can in general calculate SS_E values for various values of λ and pick the value of λ that gives the smallest sum of the squared forecast errors.

Prediction Intervals Another issue in forecasting is the uncertainty associated with it. That is, we may be interested not only in the "point estimates" but also in the quantification of the prediction uncertainty. This is usually achieved by providing the prediction intervals that are expected at a specific confidence level to contain the future observations. Calculations of the prediction intervals will require the estimation of the variance of the forecast errors. We will discuss two different techniques in estimating prediction error variance in Section 4.6.3. For the constant process, the 100 $(1 - \alpha/2)$ percent prediction intervals for any lead time τ are given as

$$\tilde{y}_T \pm Z_{\alpha/2}\hat{\sigma}_e,$$

where $\tilde{y}_T$ is the first-order exponential smoother, $Z_{\alpha/2}$ is the $100(1 - \alpha/2)$ percentile of the standard normal distribution, and $\hat{\sigma}_e$ is the estimate of the standard deviation of the forecast errors.

It should be noted that the prediction interval is constant for all lead times. This of course can be (and probably is in most cases) quite unrealistic. As it will be more likely that the process goes through some changes as time goes on, we would correspondingly expect to be less and less "sure" about our predictions for large lead times (or large τ values). Hence we would anticipate prediction intervals that are getting wider and wider for increasing lead times. We propose a remedy for this in Section 4.6.3. We will discuss this issue further in Chapter 6.

Example 4.4 We are interested in the average speed on a specific stretch of a highway during nonrush hours. For the past year and a half (78 weeks), we have available weekly averages of the average speed in miles/hour between 10 AM and 3 PM. The data are given in Table 4.5. Figure 4.18 shows that the time series data follow a constant process. To smooth out the excessive variation, however, first-order exponential smoothing can be used. The "best" smoothing constant can be determined by finding the smoothing constant value that minimizes the sum of the squared one-step-ahead prediction errors.

The sum of the squared one-step-ahead prediction errors for various λ values is given in Table 4.6. Furthermore, Figure 4.19 shows that the minimum SS_E is obtained for $\lambda = 0.4$.

Let us assume that we are also asked to make forecasts for the next 12 weeks at week 78. Figure 4.20 shows the smoothed values for the first 78 weeks together with the forecasts for weeks 79–90 with prediction intervals. It also shows the actual weekly speed during that period. Note that since the constant process is assumed, the forecasts for the next 12 weeks are the same. Similarly, the prediction intervals are constant for that period.

TABLE 4.5 The Weekly Average Speed During Nonrush Hours

Week	Speed	Week	Speed	Week	Speed	Week	Speed
1	47.12	26	46.74	51	45.71	76	45.69
2	45.01	27	46.62	52	43.84	77	44.59
3	44.69	28	45.31	53	45.09	78	43.45
4	45.41	29	44.69	54	44.16	79	44.75
5	45.45	30	46.39	55	46.21	80	45.46
6	44.77	31	43.79	56	45.11	81	43.73
7	45.24	32	44.28	57	46.16	82	44.15
8	45.27	33	46.04	58	46.50	83	44.05
9	46.93	34	46.45	59	44.88	84	44.83
10	47.97	35	46.31	60	45.68	85	43.93
11	45.27	36	45.65	61	44.40	86	44.40
12	45.10	37	46.28	62	44.17	87	45.25
13	43.31	38	44.11	63	45.18	88	44.80
14	44.97	39	46.00	64	43.73	89	44.75
15	45.31	40	46.70	65	45.14	90	44.50
16	45.23	41	47.84	66	47.98	91	45.12
17	42.92	42	48.24	67	46.52	92	45.28
18	44.99	43	45.59	68	46.89	93	45.15
19	45.12	44	46.56	69	46.01	94	46.24
20	46.67	45	45.02	70	44.98	95	46.15
21	44.62	46	43.67	71	45.76	96	46.57
22	45.11	47	44.53	72	45.38	97	45.51
23	45.18	48	44.37	73	45.33	98	46.98
24	45.91	49	44.62	74	44.07	99	46.64
25	48.39	50	46.71	75	44.02	100	44.31

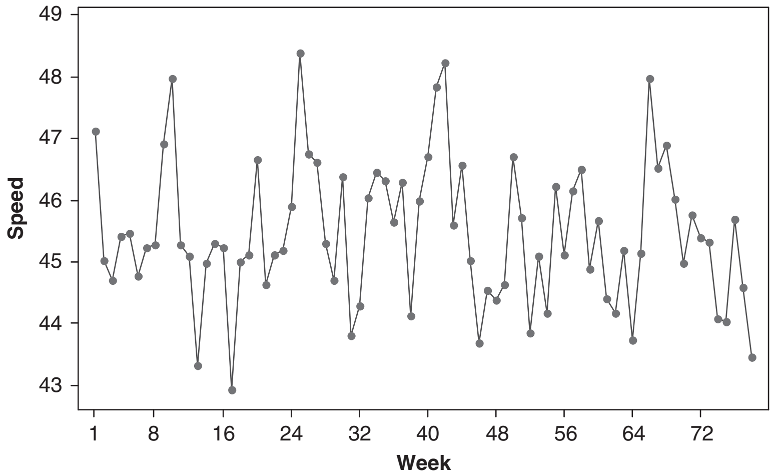


FIGURE 4.18 The weekly average speed during nonrush hours.

TABLE 4.6 SS_E for Different λ Values for the Average Speed Data

λ		0.1		0.2		0.3		0.4		0.5		0.9	
Week	Speed	Forecast	$e(t)$	Forecast	$e(t)$	Forecast	$e(t)$	Forecast	$e(t)$	Forecast	$e(t)$	Forecast	$e(t)$
1	47.12	47.12	0.00	47.12	0.00	47.12	0.00	47.12	0.00	47.12	0.00	47.12	0.00
2	45.01	47.12	−2.11	47.12	−2.11	47.12	−2.11	47.12	−2.11	47.12	−2.U	47.12	−2.11
3	44.69	46.91	−2.22	46.70	−2.01	46.49	−1.80	46.28	−1.59	46.07	−1.38	45.22	−0.53
4	45.41	46.69	−1.28	46.30	−0.89	45.95	−0.54	45.64	−0.23	45.38	0.03	44.74	0.67
5	45.45	46.56	−1.11	46.12	−0.67	45.79	−0.34	45.55	−0.10	45.39	0.06	45.34	0.11
6	44.77	46.45	−1.68	45.99	−1.22	45.69	−0.92	45.51	−0.74	45.42	−0.65	45.44	−0.67
7	45.24	46.28	−1.04	45.74	−0.50	45.41	−0.17	45.21	0.03	45.10	0.14	44.84	0.40
8	45.27	46.18	−0.91	45.64	−0.37	45.36	−0.09	45.22	0.05	45.17	0.10	45.20	0.07
9	46.93	46.09	0.84	45.57	1.36	45.33	1.60	45.24	1.69	45.22	1.71	45.26	1.67
10	47.97	46.17	1.80	45.84	2.13	45.81	2.16	45.92	2.05	46.07	1.90	46.76	1.21
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
75	44.02	45.42	−1.40	45.30	−1.28	45.12	−1.10	44.93	−0.91	44.75	−0.73	44.20	−0.18
76	45.69	45.28	0.41	45.05	0.64	44.79	0.90	44.56	1.13	44.39	1.30	44.04	1.65
77	44.59	45.32	−0.73	45.18	−0.59	45.06	−0.47	45.01	−0.42	45.04	−0.45	45.52	−0.93
78	43.45	45.25	−1.80	45.06	−1.61	44.92	−1.47	44.84	−1.39	44.81	−1.36	44.68	−1.23
SS_E			124.14		118.88		117.27		116.69		116.95		128.98

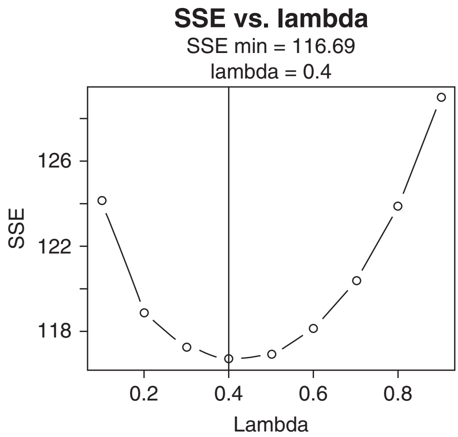


FIGURE 4.19 Plot of SS_E for various λ values for average speed data.

4.6.2 Linear Trend Process

The t-step-ahead forecast for the linear trend model is given by

$$\begin{aligned}\hat{y}_{T+\tau}(T) &= \hat{\beta}_{0,T} + \hat{\beta}_{1,T}(T+\tau) \\ &= \hat{\beta}_{0,T} + \hat{\beta}_{1,T}T + \hat{\beta}_{1,T}\tau \\ &= \hat{y}_T + \hat{\beta}_{1,T}\tau.\end{aligned} \tag{4.30}$$

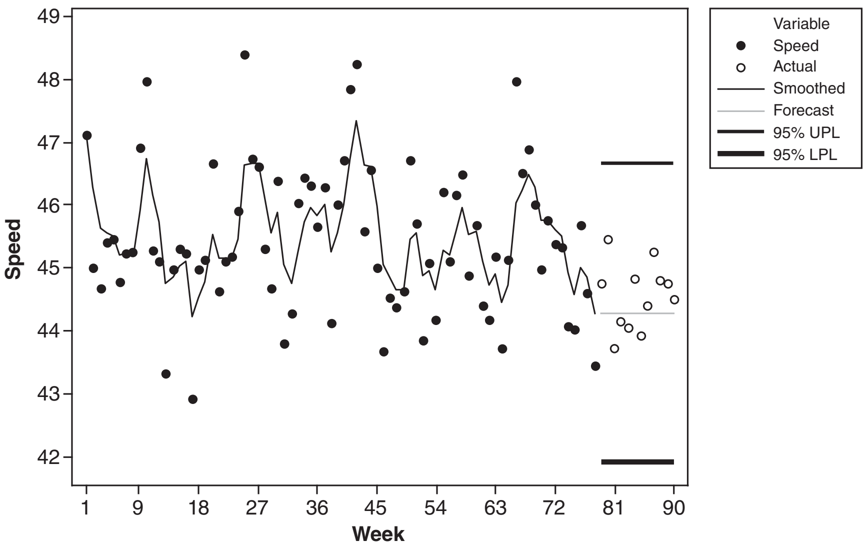


FIGURE 4.20 Forecasts for the weekly average speed data for weeks 79–90.

In terms of the exponential smoothers, we can rewrite Eq. (4.30) as

$$\begin{aligned}\hat{y}_{T+\tau}(\tau) &= \left(2\tilde{y}_T^{(1)} - \tilde{y}_T^{(2)}\right) + \tau\frac{\lambda}{1-\lambda}\left(\tilde{y}_T^{(1)} - \tilde{y}_T^{(2)}\right) \\ &= \left(2 + \frac{\lambda}{1-\lambda}\tau\right)\tilde{y}_T^{(1)} - \left(1 + \frac{\lambda}{1-\lambda}\tau\right)\tilde{y}_T^{(2)}.\end{aligned} \tag{4.31}$$

It should be noted that the predictions for the trend model depend on the lead time and, as opposed to the constant model, will be different for different lead times. As we collect more data, we can improve our forecasts by updating our parameter estimates using

$$\begin{aligned}\hat{\beta}_{0,T+1} &= \lambda(1+\lambda)y_{T+1} + (1-\lambda)^2(\hat{\beta}_{0,T} + \hat{\beta}_{1,T}) \\ \hat{\beta}_{1,T+1} &= \frac{\lambda}{(2-\lambda)}\left(\hat{\beta}_{0,T+1} - \hat{\beta}_{0,T}\right) + \frac{2\,(1-\lambda)}{(2-\lambda)}\hat{\beta}_{1,T}\end{aligned} \tag{4.32}$$

Subsequently, we can update our τ-step-ahead forecasts based on Eq. (4.32). As in the constant process, the discount factor, λ, can be estimated by minimizing the sum of the squared one-step-ahead forecast errors given in Eq. (4.29).

In this case, the $100(1 - \alpha/2)$ percent prediction interval for any lead time τ is

$$\left(2 + \frac{\lambda}{1-\lambda}\tau\right)\hat{y}_T^{(1)} - \left(1 + \frac{\lambda}{1-\lambda}\tau\right)\hat{y}_T^{(2)} \pm Z_{\alpha/2}\frac{c_\tau}{c_1}\hat{\sigma}_e,$$

where

$$c_i^2 = 1 + \frac{\lambda}{(2-\lambda)^3}[(10 - 14\lambda + 5\lambda^2) + 2i\lambda(4 - 3\lambda) + 2i^2\lambda^2].$$

Example 4.5 Consider the CPI data in Example 4.2. Assume that we are currently in December 2003 and would like to make predictions of the CPI for the following year. Although the data from January 1995 to December 2003 clearly exhibit a linear trend, we may still like to consider first-order exponential smoothing first. We will then calculate the "best" λ value that minimizes the sum of the squared one-step-ahead prediction errors. The predictions and prediction errors for various λ values are given in Table 4.7.

Figure 4.21 shows the sum of the squared one-step-ahead prediction errors (SS_E) for various values of λ.

TABLE 4.7 The Predictions and Prediction Errors for Various λ Values for CPI Data

		$\lambda = 0.1$		$\lambda = 0.2$		$\lambda = 0.3$			$\lambda = 0.9$		$\lambda = 0.99$	
Month-Year	CPI	Prediction	Error	Prediction	Error	Prediction	Error		Prediction	Error	Prediction	Error
Jan-1995	150.3	150.30	0.00	150.30	0.00	150.30	0.00	…	150.30	0.00	150.30	0.00
Feb-1995	150.9	150.30	0.60	150.30	0.60	150.30	0.60	…	150.30	0.60	150.30	0.60
Mar-1995	151.4	150.36	1.04	150.42	0.98	150.48	0.92	…	150.84	0.56	150.89	0.51
Apr-1995	151.9	150.46	1.44	150.62	1.28	150.76	1.14	…	151.34	0.56	151.39	0.51
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮
Nov-2003	184.5	182.29	2.21	183.92	0.58	184.45	0.05	…	185.01	−0.51	185.00	−0.50
Dec-2003	184.3	182.51	1.79	184.03	0.27	184.46	−0.16	…	184.55	−0.25	184.51	−0.21
SS_E			1061.50		309.14		153.71			31.90		28.62

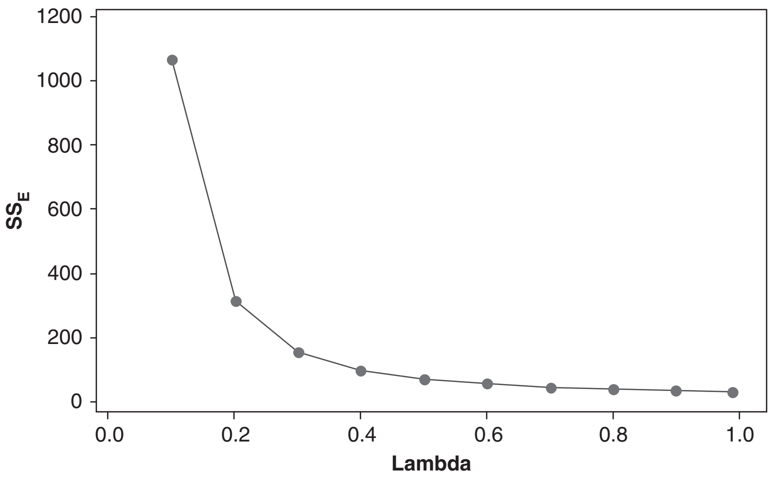


FIGURE 4.21 Scatter plot of the sum of the squared one-step-ahead prediction errors versus λ.

We notice that the SS_E keeps on getting smaller as λ gets bigger. This suggests that the data are highly autocorrelated. This can be clearly seen in the ACF plot in Figure 4.22. In fact if the "best" λ value (i.e., λ value that minimizes SS_E) turns out to be high, it may indeed be better to switch to a higher-order smoothing or use an ARIMA model as discussed in Chapter 5.

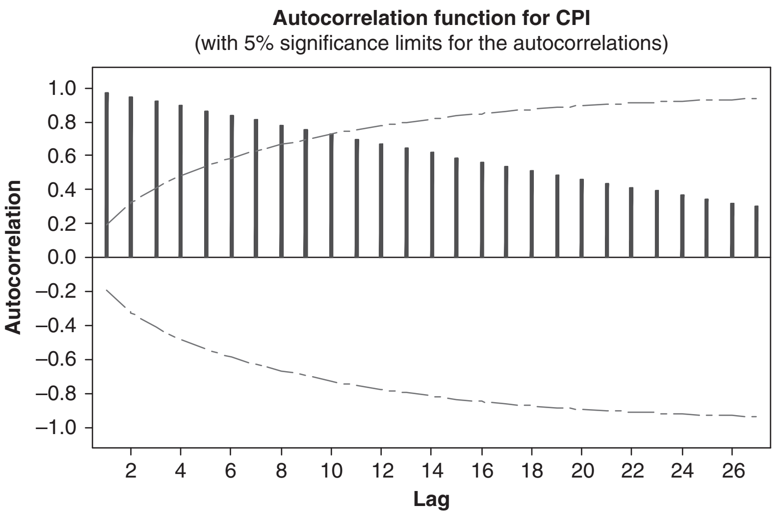


FIGURE 4.22 ACF plot for the CPI data (with 5% significance limits for the autocorrelations).

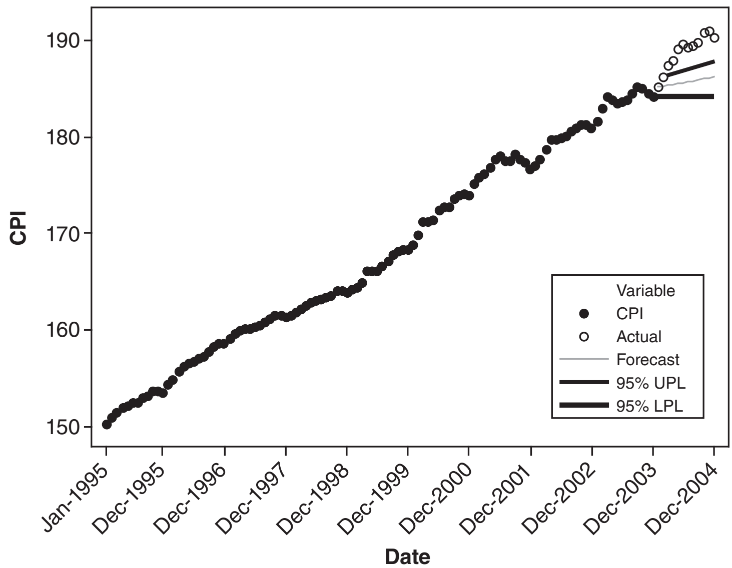


FIGURE 4.23 The 1- to 12-step-ahead forecasts of the CPI data for 2004.

Since the first-order exponential smoothing is deemed inadequate, we will now try the second-order exponential smoothing to forecast next year's monthly CPI values. Usually we have two options:

1. On December 2003, make forecasts for the entire 2004 year; that is, 1-step-ahead, 2-step-ahead, … , 12-step-ahead forecasts. For that we can use Eq. (4.30) or equivalently Eq. (4.31). Using the double exponential smoothing option in Minitab with $\lambda = 0.3$, we obtain the forecasts given in Figure 4.23.

Note that the forecasts further in the future (for the later part of 2004) are quite a bit off. To remedy this we may instead use the following strategy.

2. In December 2003, make the one-step-ahead forecast for January 2004. When the data for January 2004 becomes available, then make the one-step-ahead forecast for February 2004, and so on. We can see from Figure 4.24 that forecasts when only one-step-ahead forecasts are used and adjusted as actual data becomes available perform better than in the previous case where, for December 2003, forecasts are made for the entire following year.

The JMP software package also has an excellent forecasting capability. Table 4.8 shows output from JMP for the CPI data for double

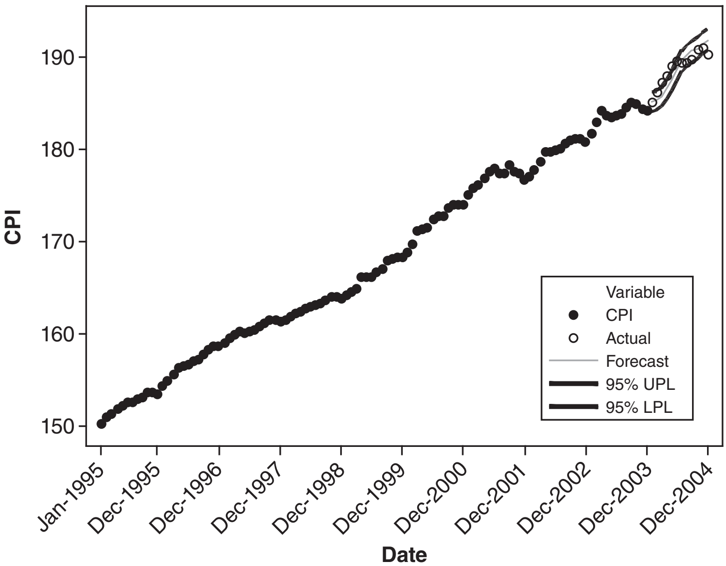
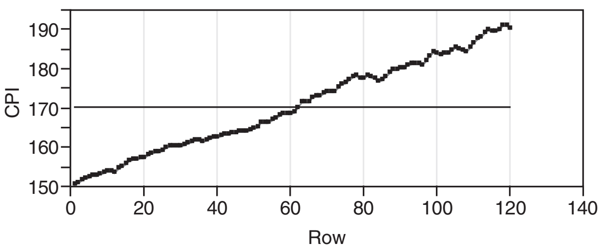


FIGURE 4.24 The one-step-ahead forecasts of the CPI data for 2004.

exponential smoothing. JMP uses the double smoothing procedure that employs a single smoothing constant. The JMP output shows the time series plot and summary statistics including the sample ACF. It also provides a sample partial ACF, which we will discuss in Chapter 5. Then an optimal smoothing constant is chosen by finding the value of λ that

TABLE 4.8 JMP Output for the CPI Data

Time series CPI



Mean	170.13167
Std	11.629323
N	120
Zero Mean ADF	8.4844029
Single Mean ADF	-0.075966
Trend ADF	-2.443095

(*continued*)

TABLE 4.8 (*Continued*)

Time series basic diagnostics

Lag	AutoCorr	Plot autocorr	Ljung-box Q	p-Value
0	1.0000		.	.
1	0.9743		116.774	<.0001
2	0.9472		228.081	<.0001
3	0.9203		334.053	<.0001
4	0.8947		435.091	<.0001
5	0.8694		531.310	<.0001
6	0.8436		622.708	<.0001
7	0.8166		709.101	<.0001
8	0.7899		790.659	<.0001
9	0.7644		867.721	<.0001
10	0.7399		940.580	<.0001
11	0.7161		1009.46	<.0001

Lag	AutoCorr	Plot autocorr	Ljung-box Q	p-Value
12	0.6924		1074.46	<.0001
13	0.6699		1135.85	<.0001
14	0.6469		1193.64	<.0001
15	0.6235		1247.84	<.0001
16	0.6001		1298.54	<.0001
17	0.5774		1345.93	<.0001
18	0.5550		1390.14	<.0001
19	0.5324		1431.24	<.0001
20	0.5098		1469.29	<.0001
21	0.4870		1504.36	<.0001
22	0.4637		1536.48	<.0001
23	0.4416		1565.91	<.0001
24	0.4205		1592.87	<.0001
25	0.4000		1617.54	0.0000

Lag	Partial	plot partial
0	1.0000	
1	0.9743	
2	–0.0396	
3	–0.0095	
4	0.0128	
5	–0.0117	
6	–0.0212	
7	–0.0379	
8	–0.0070	
9	0.0074	
10	0.0033	
11	–0.0001	
12	–0.0116	
13	0.0090	
14	–0.0224	
15	–0.0220	
16	–0.0139	
17	–0.0022	
18	–0.0089	
19	–0.0174	
20	–0.0137	
21	–0.0186	
22	–0.0234	
23	0.0074	
24	0.0030	
25	–0.0036	

TABLE 4.8 (*Continued*)

Model Comparison

Model	DF	Variance	AIC
Double (Brown) Exponential Smoothing	117	0.247119	171.05558

SBC	RSquare	-2LogLH	AIC	Rank	SBC Rank	MAPE	MAE
173.82626	0.998	169.05558	0	0	0.216853	0.376884	

Model: Double (Brown) Exponential Smoothing

Model Summary

DF	117
Sum of Squared Errors	28.9129264
Variance Estimate	0.24711903
Standard Deviation	0.49711068
Akaike's 'A' Information Criterion	171.055579
Schwarz's Bayesian Criterion	173.826263
RSquare	0.99812888
RSquare Adj	0.99812888
MAPE	0.21685285
MAE	0.37688362
-2LogLikelihood	169.055579

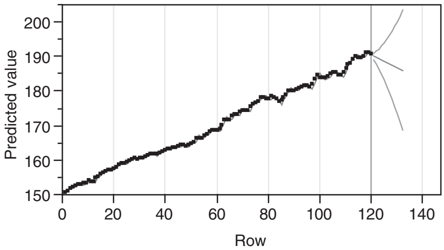
Stable Yes

Invertible Yes

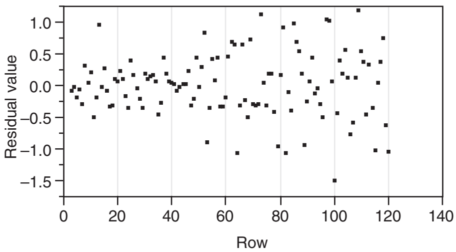
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Level Smoothing Weight	0.81402446	0.0919040	8.86	<.0001

Forecast



Residuals



(*continued*)

TABLE 4.8 (*Continued*)

Lag	AutoCorr plot autocorr	Ljung-box Q	p-Value
0	1.0000	.	.
1	0.0791	0.7574	0.3841
2	–0.3880	19.1302	<.0001
3	–0.2913	29.5770	<.0001
4	–0.0338	29.7189	<.0001
5	0.1064	31.1383	<.0001
6	0.1125	32.7373	<.0001
7	0.1867	37.1819	<.0001
8	–0.1157	38.9063	<.0001
9	–0.3263	52.7344	<.0001
10	–0.1033	54.1324	<.0001
11	0.2149	60.2441	<.0001
12	0.2647	69.6022	<.0001
13	–0.0773	70.4086	<.0001
14	0.0345	70.5705	<.0001
15	–0.1243	72.6937	<.0001
16	–0.1429	75.5304	<.0001
17	0.0602	76.0384	<.0001
18	0.1068	77.6533	<.0001
19	0.0370	77.8497	<.0001
20	–0.0917	79.0656	<.0001
21	–0.0363	79.2579	<.0001
22	–0.0995	80.7177	<.0001
23	–0.0306	80.8570	<.0001
24	0.2602	91.0544	<.0001
25	0.1728	95.6007	<.0001
Lag	**Partial plot partial**		
0	1.0000		
1	0.0791		
2	–0.3967		
3	–0.2592		
4	–0.1970		
5	–0.1435		
6	–0.0775		
7	0.1575		
8	–0.1144		
9	–0.2228		
10	–0.1482		
Lag	**AutoCorr plot autocorr**	**Ljung-box Q**	**p-Value**
11	–0.0459		
12	0.0368		
13	–0.1335		
14	0.2308		
15	–0.0786		
16	0.0050		
17	0.0390		
18	–0.0903		
19	–0.0918		
20	0.0012		
21	–0.0077		
22	–0.1935		
23	–0.0665		
24	0.1783		
25	0.0785		

minimizes the error sum of squares. The value selected is $\lambda = 0.814$. This relatively large value is not unexpected, because there is a very strong linear trend in the data and considerable autocorrelation. Values of the forecast for the next 12 periods at origin December 2004 and the associated prediction interval are also shown. Finally, the residuals from the model fit are shown along with the sample ACF and sample partial ACF plots of the residuals. The sample ACF indicates that there may be a small amount of structure in the residuals, but it is not enough to cause concern.

4.6.3 Estimation of σ_e^2

In the estimation of the variance of the forecast errors, σ_e^2, it is often assumed that the model (e.g., constant, linear trend) is correct and constant in time. With these assumptions, we have two different ways of estimating σ_e^2:

1. We already defined the one-step-ahead forecast error as $e_T(1) = y_T - \hat{y}_T(T-1)$. The idea is to apply the model to the historic data and obtain the forecast errors to calculate:

$$\begin{aligned} \hat{\sigma}_e^2 &= \frac{1}{T}\sum_{t=1}^{T} e_t^2(1) \\ &= \frac{1}{T}\sum_{t=1}^{T} (y_t - \hat{y}_t(t-1))^2 \end{aligned} \tag{4.33}$$

It should be noted that in the variance calculations the mean adjustment was not needed, since for the correct model the forecasts are unbiased; that is, the expected value of the forecast errors is 0.

As more data are collected, the variance of the forecast errors can be updated as

$$\hat{\sigma}_{eT+1}^2 = \frac{1}{T+1}\left(T\hat{\sigma}_{e,T}^2 + e_{T+1}^2(1)\right). \tag{4.34}$$

As discussed in Section 4.6.1, it may be counterintuitive to have a constant forecast error variance for all lead times. We can instead define $\sigma_e^2(\tau)$ as the τ-step-ahead forecast error variance and estimate it by

$$\hat{\sigma}_e^2(\tau) = \frac{1}{T-\tau+1}\sum_{t=\tau}^{T} e_1^2(\tau). \tag{4.35}$$

Hence the estimate in Eq. (4.35) can instead be used in the calculations of the prediction interval for the τ-step-ahead forecast.

2. For the second method of estimating σ_e^2 we will first define the *mean absolute deviation* Δ as

$$\Delta = E(|e - E(e)|) \tag{4.36}$$

and, assuming that the model is correct, calculate its estimate by

$$\hat{\Delta}_T = \delta|e_T(1)| + (1-\delta)\hat{\Delta}_{T-1}. \tag{4.37}$$

Then the estimate of the σ_e^2 is given by

$$\hat{\sigma}_{e,T} = 1.25\hat{\Delta}_T. \tag{4.38}$$

For further details, see Montgomery et al. (1990).

4.6.4 Adaptive Updating of the Discount Factor

In the previous sections we discussed estimation of the "best" discount factor, $\hat{\lambda}$, by minimizing the sum of the squared one-step-ahead forecasts errors. However, as we have seen with the Dow Jones Index data, changes in the underlying time series model will make it difficult for the exponential smoother with fixed discount factor to follow these changes. Hence a need for monitoring and, if necessary, modifying the discount factor arises. By doing so, the discount factor will adapt to the changes in the time series model. For that we will employ the procedure originally described by Trigg and Leach (1967) for single discount factor. As an example we will consider the first-order exponential smoother and modify it as

$$\hat{y}_T = \lambda_T y_T + (1-\lambda_T)\tilde{y}_{T-1}. \tag{4.39}$$

Please note that in Eq. (4.39), the discount factor λ_T is given as a function of time and hence it is allowed to adapt to changes in the time series model. We also define the *smoothed error* as

$$Q_T = \delta e_T(1) + (1-\delta)Q_{T-1}, \tag{4.40}$$

where δ is a smoothing parameter.

Finally, we define the tracking signal as

$$\frac{Q_T}{\hat{\Delta}_T}, \tag{4.41}$$

where $\hat{\Delta}_T$ is given in Eq. (4.37). This ratio is expected to be close to 0 when the forecasting system performs well and to approach ± 1 as it starts to fail. In fact, Trigg and Leach (1967) suggest setting the discount factor to

$$\lambda_T = \left| \frac{Q_T}{\hat{\Delta}_T} \right| \tag{4.42}$$

Equation (4.42) will allow for automatic updating of the discount factor.

Example 4.6 Consider the Dow Jones Index from June 1999 to June 2006 given in Table 4.1. Figure 4.2 shows that the data do not exhibit a single regime of constant or linear trend behavior. Hence a single exponential smoother with adaptive discount factor as given in Eq. (4.42) can be used. Figure 4.25 shows two simple exponential smoothers for the Dow Jones Index: one with fixed $\lambda = 0.3$ and another one with adaptive updating based on the Trigg–Leach method given in Eq. (4.42).

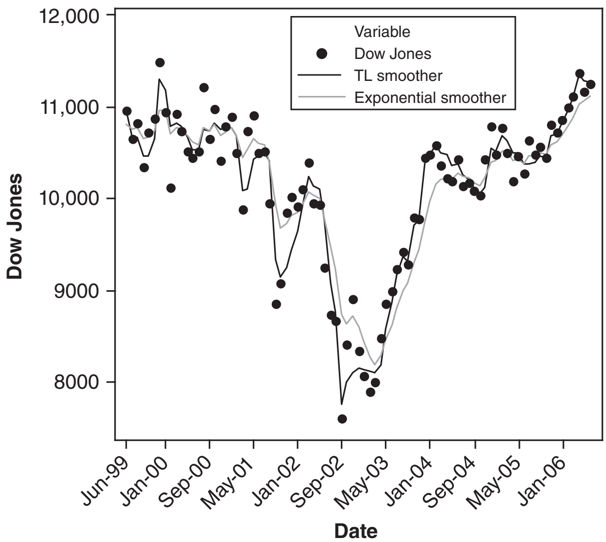


FIGURE 4.25 Time series plot of the Dow Jones Index from June 1999 to June 2006, the simple exponential smoother with $\lambda = 0.3$, and the Trigg–Leach (TL) smoother with $\delta = 0.3$.

TABLE 4.9 The Trigg–Leach Smoother for the Dow Jones Index

Date	Dow Jones	Smoothed	λ	Error	Q_t	D_t
Jun-99	10,970.8	10,970.8	1		0	0
Jul-99	10,655.2	10,655.2	1	−315.6	−94.68	94.68
Aug-99	10,829.3	10,675.835	0.11853	174.1	−14.046	118.506
Sep-99	10,337	10,471.213	0.6039	−338.835	−111.483	184.605
Oct-99	10,729.9	10,471.753	0.00209	258.687	−0.43178	206.83
⋮	⋮	⋮	⋮	⋮	⋮	⋮
May-06	11,168.3	11,283.962	0.36695	−182.705	68.0123	185.346
Jun-06	11,247.9	11,274.523	0.26174	−36.0619	36.79	140.561

This plot shows that a better smoother can be obtained by making automatic updates to the discount factor. The calculations for the Trigg–Leach smoother are given in Table 4.9.

The adaptive smoothing procedure suggested by Trigg and Leach is a useful technique. For other approaches to adaptive adjustment of exponential smoothing parameters, see Chow (1965), Roberts and Reed (1969), and Montgomery (1970).

4.6.5 Model Assessment

If the forecast model performs as expected, the forecast errors should not exhibit any pattern or structure; that is, they should be uncorrelated. Therefore it is always a good idea to verify this. As noted in Chapter 2, we can do so by calculating the sample ACF of the forecast errors from

$$rk = \frac{\sum_{t=k}^{T-1} [e_t(1) - \bar{e}] \left[e_{t-k}(1) - \bar{e}\right]}{\sum_{T=0}^{T-1} [e_t(1) - \bar{e}]^2}, \tag{4.43}$$

where

$$\bar{e} = \frac{1}{n} \sum_{t=1}^{T} e_t(1).$$

If the one-step-ahead forecast errors are indeed uncorrelated, the sample autocorrelations for any lag k should be around 0 with a standard error $1/\sqrt{T}$. Hence a sample autocorrelation for any lag k that lies outside the $\pm 2/\sqrt{T}$ limits will require further investigation of the model.

4.7 EXPONENTIAL SMOOTHING FOR SEASONAL DATA

Some time series data exhibit cyclical or seasonal patterns that cannot be effectively modeled using the polynomial model in Eq. (4.25). Several approaches are available for the analysis of such data. In this chapter we will discuss exponential smoothing techniques that can be used in modeling seasonal time series. The methodology we will focus on was originally introduced by Holt (1957) and Winters (1960) and is generally known as Winters' method, where a seasonal adjustment is made to the linear trend model. Two types of adjustments are suggested—additive and multiplicative.

4.7.1 Additive Seasonal Model

Consider the US clothing sales data given in Figure 4.26. Clearly, for certain months of every year we have high (or low) sales. Hence we can conclude that the data exhibit seasonality. The data also exhibit a linear trend as the sales tend to get higher for the same month as time goes on. As the final observation, we note that the amplitude of the seasonal pattern, that is, the range of the periodic behavior within a year, remains more or

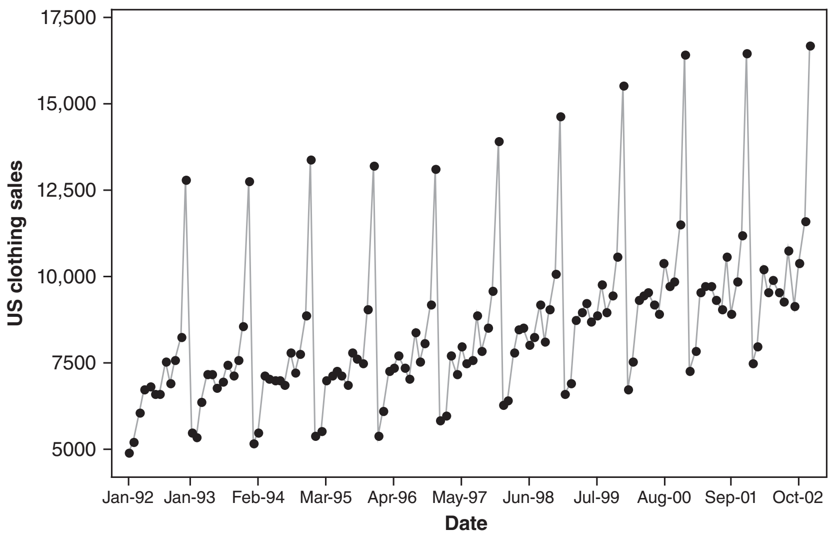


FIGURE 4.26 Time series plot of US clothing sales from January 1992 to December 2003.

less constant in time and remains independent of the average level within a year.

We will for this case assume that the seasonal time series can be represented by the following model:

$$y_t = L_t + S_t + \varepsilon_t, \tag{4.44}$$

where L_t represents the level or linear trend component and can in turn be represented by $\beta_0 + \beta_1 t$; S_t represents the seasonal adjustment with $S_t = S_{t+s} = S_{t+2s} = \dots$ for $t = 1, \dots, s-1$, where s is the length of the season (period) of the cycles; and the ε_t are assumed to be uncorrelated with mean 0 and constant variance σ_ε^2. Sometimes the level is called the *permanent component*. One usual restriction on this model is that the seasonal adjustments add to zero during one season,

$$\sum_{t=1}^{s} S_t = 0. \tag{4.45}$$

In the model given in Eq. (4.44), for forecasting the future observations, we will employ first-order exponential smoothers with different discount factors. The procedure for updating the parameter estimates once the current observation y_T is obtained is as follows.

Step 1. Update the estimate of L_T using

$$\hat{L}_T = \lambda_1(y_T - \hat{S}_{T-s}) + (1 - \lambda_1)\left(\hat{L}_{T-1} + \hat{\beta}_{1,T-1}\right), \tag{4.46}$$

where $0 < \lambda_1 < 1$. It should be noted that in Eq. (4.46), the first part can be seen as the "current" value for L_T and the second part as the forecast of L_T based on the estimates at $T - 1$.

Step 2. Update the estimate of β_1 using

$$\hat{\beta}_{1,T} = \lambda_2(\hat{L}_T - \hat{L}_{T-1}) + (1 - \lambda_2)\hat{\beta}_{1,T-1}, \tag{4.47}$$

where $0 < \lambda_2 < 1$. As in Step 1, the estimate of β_1 in Eq. (4.47) can be seen as the linear combination of the "current" value of β_1 and its "forecast" at $T - 1$.

Step 3. Update the estimate of S_t using

$$\hat{S}_T = \lambda_3(y_T - \hat{L}_T) + (1 - \lambda_3)\hat{S}_{T-s}, \tag{4.48}$$

where $0 < \lambda_3 < 1$.

Step 4. Finally, the τ-step-ahead forecast, $\hat{y}_{T+\tau}(T)$, is

$$\hat{y}_{T+\tau}(T) = \hat{L}_T + \hat{\beta}_{1,T}\tau + \hat{S}_T(\tau - s). \tag{4.49}$$

As before, estimating the initial values of the exponential smoothers is important. For a given set of historic data with n seasons (hence ns observations), we can use the least squares estimates of the following model:

$$y_t = \beta_0 + \beta_1 t + \sum_{i=1}^{s-1} \gamma_i (I_{t,i} - I_{t,s}) + \varepsilon_t, \tag{4.50}$$

where

$$I_{t,i} = \begin{cases} 1, & t = i, i+s, i+2s, \dots \\ 0, & \text{otherwise} \end{cases}. \tag{4.51}$$

The least squares estimates of the parameters in Eq. (4.50) are used to obtain the initial values as

$$\begin{aligned}
&\hat{\beta}_{0,0} = \hat{L}_0 = \hat{\beta}_0 \\
&\hat{\beta}_{1,0} = \hat{\beta}_1 \\
&\hat{S}_{j-s} = \hat{Y}_{\mathrm{j}} \quad \text{for } 1 \le j \le \mathrm{s} - 1 \\
&\hat{S}_0 = -\sum_{j=1}^{s-1} \hat{y}_j
\end{aligned}$$

These are initial values of the model parameters at the *original* origin of time, $t = 0$. To make forecasts from the correct origin of time the permanent component must be shifted to time T by computing $\hat{L}_T = \hat{L}_0 + ns\hat{\beta}_1$. Alternatively, one could smooth the parameters using equations (4.46)–(4.48) for time periods $t = 1, 2, \dots, \mathrm{T}$.

Prediction Intervals As in the nonseasonal smoothing case, the calculations of the prediction intervals would require an estimate for the prediction error variance. The most common approach is to use the relationship between the exponential smoothing techniques and the ARIMA models of Chapter 5 as discussed in Section 4.8, and estimate the prediction error variance accordingly. It can be shown that the seasonal exponential smoothing using the three parameter Holt–Winters method is optimal for an ARIMA $(0, 1, s+1) \times (0, 1, 0)_s$, process, where s represents the length of the period of the seasonal cycles. For further details, see Yar and Chatfield (1990) and McKenzie (1986).

An alternate approach is to recognize that the additive seasonal model is just a linear regression model and to use the ordinary least squares (OLS) regression procedure for constructing prediction intervals as discussed in Chapter 3. If the errors are correlated, the regression methods for autocorrelated errors could be used instead of OLS.

Example 4.7 Consider the clothing sales data given in Table 4.10. To obtain the smoothed version of this data, we can use the Winters' method option in Minitab. Since the amplitude of the seasonal pattern is constant over time, we decide to use the additive model. Two issues we have encountered in previous exponential smoothers have to be addressed in this case as well—initial values and the choice of smoothing constants. Similar recommendations as in the previous exponential smoothing options can also be made in this case. Of course, the choice of the smoothing constant, in particular, is a bit more concerning since it involves the estimation of three smoothing constants. In this example, we follow our usual recommendation and choose smoothing constants that are all equal to 0.2. For more complicated cases, we recommend seasonal ARIMA models, which we will discuss in Chapter 5.

Figure 4.27 shows the smoothed version of the seasonal clothing sales data. To use this model for forecasting, let us assume that we are currently in December 2002 and we are asked to make forecasts for the following year. Figure 4.28 shows the forecasted sales for 2003 together with the actual data and the 95% prediction limits. Note that the forecast for December 2003 is the 12-step-ahead forecast made in December 2002. Even though the forecast is made further in the future, it still performs well since in the "seasonal" sense it is in fact a one-step-ahead forecast.

4.7.2 Multiplicative Seasonal Model

If the amplitude of the seasonal pattern is proportional to the average level of the seasonal time series, as in the liquor store sales data given in Figure 4.29, the following multiplicative seasonal model will be more appropriate:

$$y_t = L_t S_t + \varepsilon_t, \tag{4.52}$$

where L_t once again represents the permanent component (i.e., $\beta_0 + \beta_1 t$); S_t represents the seasonal adjustment with $S_t = S_{t+s} = S_{t+2s} = \cdots$ for

TABLE 4.10 US Clothing Sales from January 1992 to December 2003

Date	Sales	Date	Sales	Date	Sales	Date	Sales	Date	Sales
Jan-92	4889	Aug-94	7824	Mar-97	7695	Oct-99	9481	May-02	9906
Feb-92	5197	Sep-94	7229	Apr-97	7161	Nov-99	10577	Jun-02	9530
Mar-92	6061	Oct-94	7772	May-97	7978	Dec-99	15552	Jul-02	9298
Apr-92	6720	Nov-94	8873	Jun-97	7506	Jan-00	6726	Aug-02	10,755
May-92	6811	Dec-94	13397	Jul-97	7602	Feb-00	7514	Sep-02	9128
Jun-92	6579	Jan-95	5377	Aug-97	8877	Mar-00	9330	Oct-02	10,408
Jul-92	6598	Feb-95	5516	Sep-97	7859	Apr-00	9472	Nov-02	11,618
Aug-92	7536	Mar-95	6995	Oct-97	8500	May-00	9551	Dec-02	16,721
Sep-92	6923	Apr-95	7131	Nov-97	9594	Jun-00	9203	Jan-03	7891
Oct-92	7566	May-95	7246	Dec-97	13952	Jul-00	8910	Feb-03	7892
Nov-92	8257	Jun-95	7140	Jan-98	6282	Aug-00	10378	Mar-03	9874
Dec-92	12,804	Jul-95	6863	Feb-98	6419	Sep-00	9731	Apr-03	9920
Jan-93	5480	Aug-95	7790	Mar-98	7795	Oct-00	9868	May-03	10,431
Feb-93	5322	Sep-95	7618	Apr-98	8478	Nov-00	11512	Jun-03	9758
Mar-93	6390	Oct-95	7484	May-98	8501	Dec-00	16422	Jul-03	10,003
Apr-93	7155	Nov-95	9055	Jun-98	8044	Jan-01	7263	Aug-03	11,055

(*continued*)

TABLE 4.10 (*Continued*)

Date	Sales	Date	Sales	Date	Sales	Date	Sales	Date	Sales
May-93	7175	Dec-95	13,201	Jul-98	8272	Feb-01	7866	Sep-03	9941
Jun-93	6770	Jan-96	5375	Aug-98	9189	Mar-01	9535	Oct-03	10,763
Jul-93	6954	Feb-96	6105	Sep-98	8099	Apr-01	9710	Nov-03	12058
Aug-93	7438	Mar-96	7246	Oct-98	9054	May-01	9711	Dec-03	17535
Sep-93	7144	Apr-96	7335	Nov-98	10,093	Jun-01	9324		
Oct-93	7585	May-96	7712	Dec-98	14668	Jul-01	9063		
Nov-93	8558	Jun-96	7337	Jan-99	6617	Aug-01	10,584		
Dec-93	12,753	Jul-96	7059	Feb-99	6928	Sep-01	8928		
Jan-94	5166	Aug-96	8374	Mar-99	8734	Oct-01	9843		
Feb-94	5464	Sep-96	7554	Apr-99	8973	Nov-01	11,211		
Mar-94	7145	Oct-96	8087	May-99	9237	Dec-01	16,470		
Apr-94	7062	Nov-96	9180	Jun-99	8689	Jan-02	7508		
May-94	6993	Dec-96	13109	Jul-99	8869	Feb-02	8002		
Jun-94	6995	Jan-97	5833	Aug-99	9764	Mar-02	10,203		
Jul-94	6886	Feb-97	5949	Sep-99	8970	Apr-02	9548		

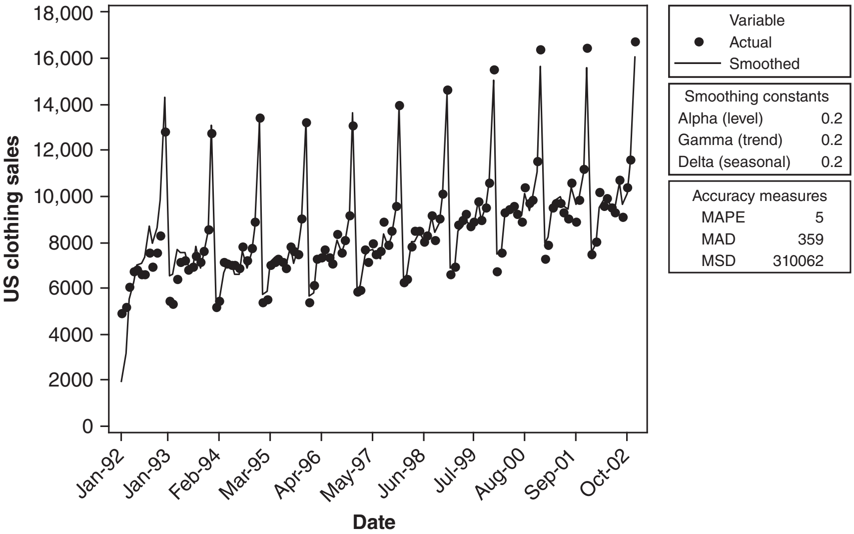


FIGURE 4.27 Smoothed data for the US clothing sales from January 1992 to December 2003 using the additive model.

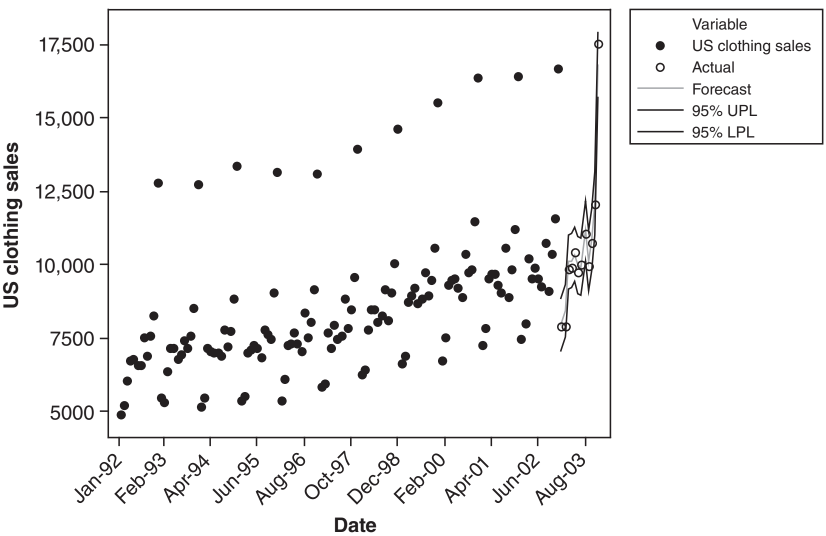


FIGURE 4.28 Forecasts for 2003 for the US clothing sales.

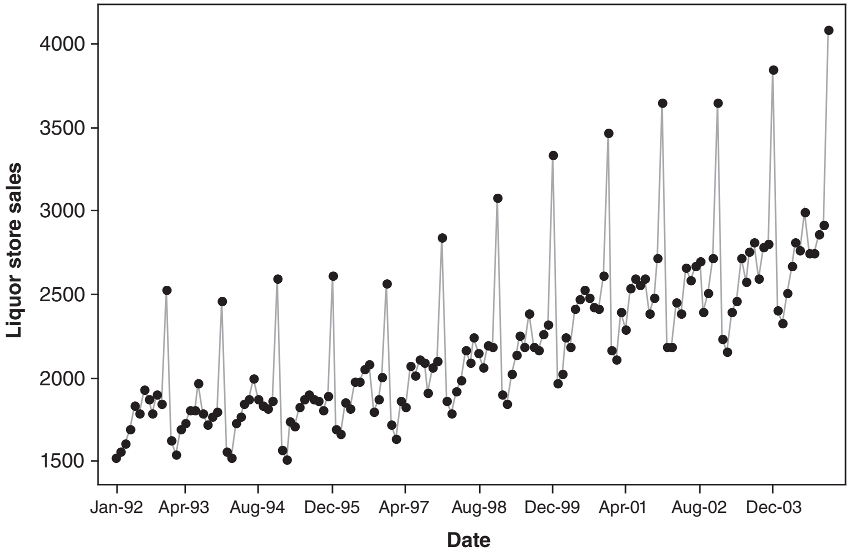


FIGURE 4.29 Time series plot of liquor store sales data from January 1992 to December 2004.

$t = i, \ldots, s - 1$, where s is the length of the period of the cycles; and the ε_t are assumed to be uncorrelated with mean 0 and constant variance σ_ε^2. The restriction for the seasonal adjustments in this case becomes

$$\sum_{t}^{s} S_t = s. \tag{4.53}$$

As in the additive model, we will employ three exponential smoothers to estimate the parameters in Eq. (4.52).

Step 1. Update the estimate of L_T using

$$\hat{L}_T = \lambda_1 \frac{y_T}{\hat{S}_{T-s}} + (1 - \lambda_1)(\hat{L}_{T-1} + \hat{\beta}_{1,T-1}), \tag{4.54}$$

where $0 < \lambda_1 < 1$. Similar interpretation as in the additive model can be made for the exponential smoother in Eq. (4.54).

Step 2. Update the estimate of β_1 using

$$\hat{\beta}_{1,T} = \lambda_2(\hat{L}_T - \hat{L}_{T-1}) + (1 - \lambda_2)\hat{\beta}_{1,T-1}, \tag{4.55}$$

where $0 < \lambda_2 < 1$.

Step 3. Update the estimate of S_t using

$$\hat{S}_T = \lambda_3 \frac{y_T}{\hat{L}_T} + (1 - \lambda_3)\hat{S}_{T-s}, \tag{4.56}$$

where $0 < \lambda_3 < 1$.

Step 4. The τ-step-ahead forecast, $\hat{y}_{T+\tau}(T)$, is

$$\hat{y}_{T+\tau}(T) = (\hat{L}_T + \hat{\beta}_{1,T}\tau)\hat{S}_T(\tau - s). \tag{4.57}$$

It will almost be necessary to obtain starting values of the model parameters. Suppose that a record consisting of n seasons of data is available. From this set of historical data, the initial values, $\hat{\beta}_{0,0}$, $\hat{\beta}_{1,0}$, and $\hat{S}_0$, can be calculated as

$$\hat{\beta}_{0,0} = \hat{L}_0 = \frac{\bar{y}_n - \bar{y}_1}{(n-1)s},$$

where

$$\bar{y}_i = \frac{1}{s} \sum_{t=(i-1)s+1}^{is} y_t$$

and

$$\hat{\beta}_{1,0} = \bar{y}_1 - \frac{s}{2}\hat{\beta}_{0,0}$$

$$\hat{S}_{j-s} = s \frac{\hat{S}_j^*}{\sum_{i=1}^{s} \hat{S}_i^*} \text{for } 1 \leq j \leq s,$$

where

$$\hat{S}_j^* = \frac{1}{n} \sum_{t=1}^{\text{n}} \frac{y_{(t-1)s+j}}{\bar{y}_t - ((s+1)/2 - j)\,\hat{\beta}_0}.$$

For further details, please see Montgomery et al. (1990) and Abraham and Ledolter (1983).

Prediction Intervals Constructing prediction intervals for the multiplicative model is much harder than the additive model as the former is nonlinear. Several authors have considered this problem, including Chatfield and Yar (1991), Sweet (1985), and Gardner (1988). Chatfield and Yar (1991) propose an empirical method in which the length of the prediction interval depends on the point of origin of the forecast and may decrease in length near the low points of the seasonal cycle. They also discuss the case where the error is assumed to be proportional to the seasonal effect rather than constant, which is the standard assumption in Winters' method. Another approach would be to obtain a "linearized" version of Winters' model by expanding it in a first-order Taylor series and use this to find an approximate variance of the predicted value (statisticians call this the delta method). Then this prediction variance could be used to construct prediction intervals much as is done in the linear regression model case.

Example 4.8 Consider the liquor store data given in Table 4.11. In Figure 4.29, we can see that the amplitude of the periodic behavior gets larger as the average level of the seasonal data gets larger due to a linear trend. Hence the multiplicative model will be more appropriate. Figures 4.30 and 4.31 show the smoothed data with additive and multiplicative models, respectively. Based on the performance of the smoothers, it should therefore be clear that the multiplicative model should indeed be preferred.

As for forecasting using the multiplicative model, we can assume as usual that we are currently in December 2003 and are asked to forecast the sales in 2004. Figure 4.32 shows the forecasts together with the actual values and the prediction intervals.

4.8 EXPONENTIAL SMOOTHING OF BIOSURVEILLANCE DATA

Bioterrorism is the use of biological agents in a campaign of aggression. The use of biological agents in warfare is not new; many centuries ago plague and other contagious diseases were employed as weapons. Their use today is potentially catastrophic, so medical and public health officials are designing and implementing biosurveillance systems to monitor populations for potential disease outbreaks. For example, public health officials collect syndrome data from sources such as hospital emergency rooms, outpatient clinics, and over-the-counter medication sales to detect disease outbreaks, such as the onset of the flu season. For an excellent and highly readable introduction to statistical techniques for biosurveillance and syndromic surveillance, see Fricker (2013). Monitoring of syndromic data is also a type of epidemiologic surveillance in a biosurveillance process,

TABLE 4.11 Liquor Store Sales from January 1992 to December 2004

Date	Sales	Date	Sales	Date	Sales	Date	Sales	Date	Sales
Jan-92	1519	Aug-94	1870	Mar-97	1862	Oct-99	2264	May-02	2661
Feb-92	1551	Sep-94	1834	Apr-97	1826	Nov-99	2321	Jun-02	2579
Mar-92	1606	Oct-94	1817	May-97	2071	Dec-99	3336	Jul-02	2667
Apr-92	1686	Nov-94	1857	Jun-97	2012	Jan-00	1963	Aug-02	2698
May-92	1834	Dec-94	2593	Jul-97	2109	Feb-00	2022	Sep-02	2392
Jun-92	1786	Jan-95	1565	Aug-97	2092	Mar-00	2242	Oct-02	2504
Jul-92	1924	Feb-95	1510	Sep-97	1904	Apr-00	2184	Nov-02	2719
Aug-92	1874	Mar-95	1736	Oct-97	2063	May-00	2415	Dec-02	3647
Sep-92	1781	Apr-95	1709	Nov-97	2096	Jun-00	2473	Jan-03	2228
Oct-92	1894	May-95	1818	Dec-97	2842	Jul-00	2524	Feb-03	2153
Nov-92	1843	Jun-95	1873	Jan-98	1863	Aug-00	2483	Mar-03	2395
Dec-92	2527	Jul-95	1898	Feb-98	1786	Sep-00	2419	Apr-03	2460
Jan-93	1623	Aug-95	1872	Mar-98	1913	Oct-00	2413	May-03	2718
Feb-93	1539	Sep-95	1856	Apr-98	1985	Nov-00	2615	Jun-03	2570
Mar-93	1688	Oct-95	1800	May-98	2164	Dec-00	3464	Jul-03	2758
Apr-93	1725	Nov-95	1892	Jun-98	2084	Jan-01	2165	Aug-03	2809

(*continued*)

TABLE 4.11 (*Continued*)

Date	Sales	Date	Sales	Date	Sales	Date	Sales	Date	Sales
May-93	1807	Dec-95	2616	Jul-98	2237	Feb-01	2107	Sep-03	2597
Jun-93	1804	Jan-96	1690	Aug-98	2146	Mar-01	2390	Oct-03	2785
Jul-93	1962	Feb-96	1662	Sep-98	2058	Apr-01	2292	Nov-03	2803
Aug-93	1788	Mar-96	1849	Oct-98	2193	May-01	2538	Dec-03	3849
Sep-93	1717	Apr-96	1810	Nov-98	2186	Jun-01	2596	Jan-04	2406
Oct-93	1769	May-96	1970	Dec-98	3082	Jul-01	2553	Feb-04	2324
Nov-93	1794	Jun-96	1971	Jan-99	1897	Aug-01	2590	Mar-04	2509
Dec-93	2459	Jul-96	2047	Feb-99	1838	Sep-01	2384	Apr-04	2670
Jan-94	1557	Aug-96	2075	Mar-99	2021	Oct-01	2481	May-04	2809
Feb-94	1514	Sep-96	1791	Apr-99	2136	Nov-01	2717	Jun-04	2764
Mar-94	1724	Oct-96	1870	May-99	2250	Dec-01	3648	Jul-04	2995
Apr-94	1769	Nov-96	2003	Jun-99	2186	Jan-02	2182	Aug-04	2745
May-94	1842	Dec-96	2562	Jul-99	2383	Feb-02	2180	Sep-04	2742
Jun-94	1869	Jan-97	1716	Aug-99	2182	Mar-02	2447	Oct-04	2863
Jul-94	1994	Feb-97	1629	Sep-99	2169	Apr-02	2380	Nov-04	2912
								Dec-04	4085

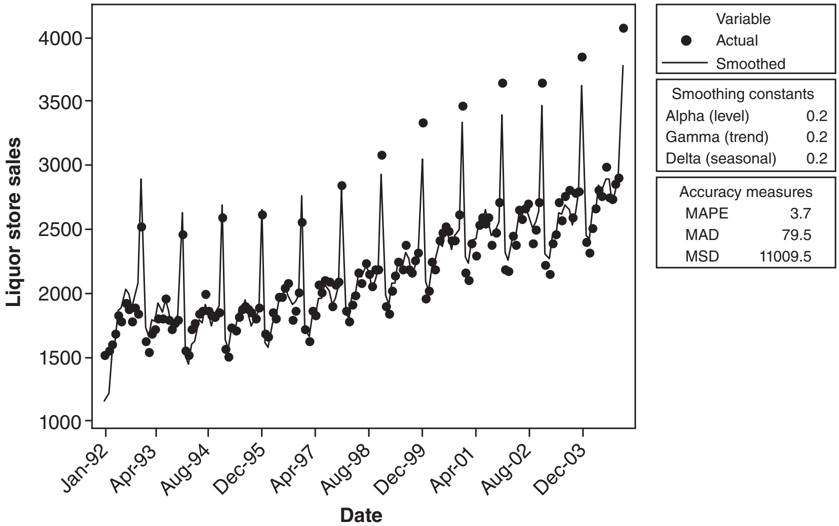


FIGURE 4.30 Smoothed data for the liquor store sales from January 1992 to December 2004 using the additive model.

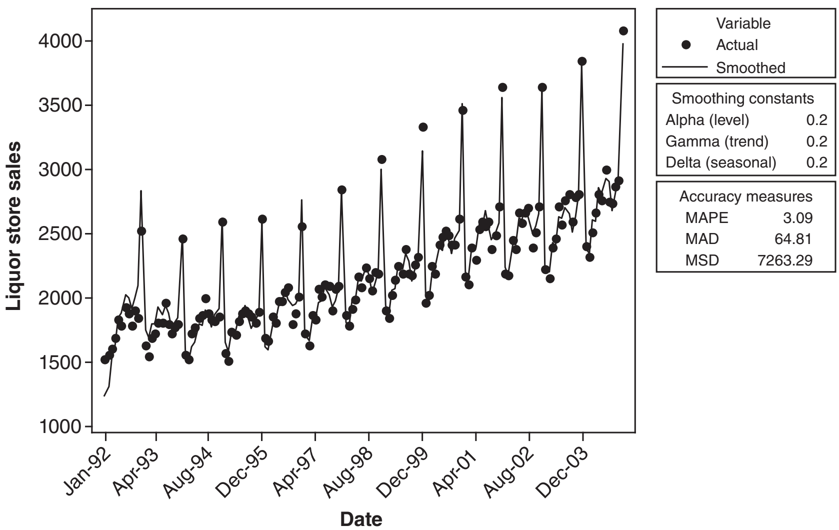


FIGURE 4.31 Smoothed data for the liquor store sales from January 1992 to December 2004 using the multiplicative model.

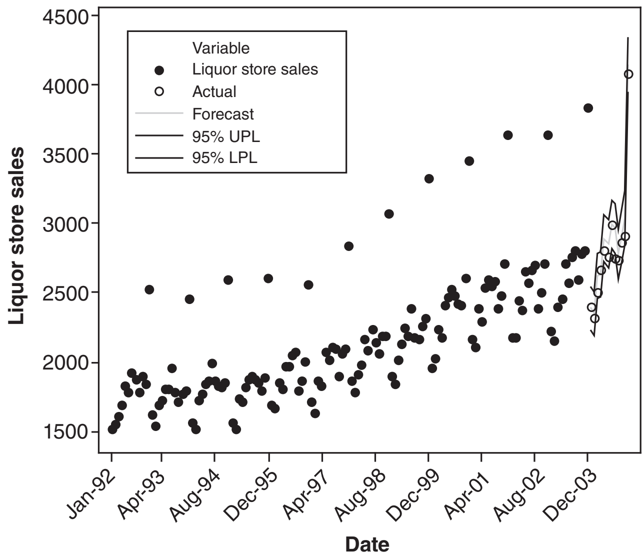


FIGURE 4.32 Forecasts for the liquor store sales for 2004 using the multiplicative model.

where significantly higher than anticipated counts of influenza-like illness might signal a potential bioterrorism attack.

As an example of such syndromic data, Fricker (2013) describes daily counts of respiratory and gastrointestinal complaints for more than 2 $^1/_2$ years at several hospitals in a large metropolitan area. Table 4.12 presents the respiratory count data from one of these hospitals. There are 980 observations. Fifty observations were missing from the original data set. The missing values were replaced with the last value that was observed on the same day of the week. This type of data imputation is a variation of "Hot Deck Imputation" discussed in Section 1.4.3 and in Fricker (2013). It is also sometimes called last observation (or Value) carried forward (LOCF). For additional discussion see the web site: http://missingdata.lshtm.ac.uk/.

Figure 4.33 is a time series plot of the respiratory syndrome count data in Table 4.12. This plot was constructed using the Graph Builder feature in JMP. This software package overlays a smoothed curve on the data. The curve is fitted using **locally weighted regression**, often called **loess**. This is a variation of kernel regression that uses a weighted average of the data in a local neighborhood around a specific location to determine the value to plot at that location. Loess usually uses either first-order linear regression or a quadratic regression model for the weighted least squares fit. For more information on kernel regression and loess see Montgomery, et al. (2012).

TABLE 4.12 Counts of Respiratory Complaints at a Metropolitan Hospital

Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count
1	17	101	30	201	31	301	12	401	28	501	35	601	26	701	19	801	41	901	29
2	29	102	21	202	23	302	16	402	26	502	27	602	31	702	12	802	50	902	26
3	31	103	32	203	13	303	24	403	28	503	33	603	23	703	17	803	42	903	36
4	34	104	32	204	18	304	21	404	29	504	30	604	24	704	22	804	56	904	31
5	18	105	43	205	36	305	14	405	33	505	30	605	27	705	20	805	36	905	25
6	43	106	25	206	23	306	15	406	36	506	29	606	24	706	22	306	51	906	31
7	34	107	32	207	22	307	23	407	62	507	30	607	31	707	21	807	40	907	32
8	23	108	31	208	23	308	10	403	31	508	22	608	29	708	24	808	29	908	30
9	23	109	33	209	26	309	16	409	30	509	40	609	36	709	16	809	61	909	31
10	39	110	40	210	22	310	11	410	31	510	40	610	31	710	14	810	42	910	29
11	25	111	37	211	21	311	16	411	27	511	41	611	30	711	14	811	56	911	30
12	15	112	34	212	25	312	16	412	35	512	34	612	27	712	30	812	60	912	35
13	29	113	29	213	20	313	12	413	45	513	30	613	27	713	24	813	38	913	24
14	20	114	50	214	18	314	23	414	37	514	33	614	25	714	25	814	52	914	27
15	21	115	27	215	26	315	10	415	23	515	17	615	34	715	17	815	32	915	22
16	22	116	28	216	32	315	15	416	31	516	32	616	33	716	27	816	43	916	33
17	24	117	23	217	41	317	11	417	33	517	40	617	36	717	25	817	54	917	29
18	19	118	27	218	30	318	17	418	27	518	30	618	26	718	14	818	36	913	37
19	28	119	27	219	34	319	13	419	28	519	27	619	20	719	25	819	51	919	29
20	29	120	41	220	38	320	14	420	46	520	30	620	27	720	25	820	57	920	32
21	26	121	29	221	22	321	20	421	39	521	38	621	25	721	26	821	48	921	27
22	22	122	26	222	35	322	10	422	53	522	22	622	36	722	20	822	70	922	22
23	21	123	28	223	36	323	15	423	33	523	27	623	30	723	21	823	48	923	33

(*continued*)

TABLE 4.12 (*Continued*)

Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count
24	29	124	30	224	37	324	14	424	32	524	19	624	39	724	29	824	54	924	29
25	25	125	49	225	27	325	6	425	45	525	19	625	26	725	16	825	36	925	37
26	20	126	43	226	23	326	17	426	21	526	33	626	20	726	24	826	51	926	29
27	20	127	27	227	31	327	17	427	47	527	45	627	27	727	42	827	52	927	20
28	29	128	32	228	39	328	17	423	23	528	34	628	36	728	44	828	48	928	13
29	29	129	13	229	39	329	23	429	39	529	27	629	43	729	34	829	70	929	27
30	32	130	26	230	31	330	9	430	32	530	31	630	46	730	33	830	48	930	23
31	16	131	34	231	43	331	21	431	27	531	19	631	33	731	26	831	57	931	17
32	25	132	27	232	35	332	13	432	29	532	22	632	26	732	29	832	38	932	26
33	20	133	33	233	41	333	13	433	37	533	23	633	33	733	33	833	44	933	23
34	22	134	42	234	24	334	14	434	32	534	13	634	24	734	34	834	34	934	27
35	27	135	29	235	39	335	25	435	28	535	29	635	23	735	42	835	50	935	28
36	32	136	29	236	44	336	15	436	42	536	13	636	51	736	43	836	39	936	21
37	23	137	29	237	35	337	18	437	33	537	20	637	35	737	33	837	65	937	20
38	31	138	28	238	30	338	21	438	36	538	20	638	26	738	31	838	55	938	25
39	22	139	35	239	29	339	18	439	25	539	23	639	32	739	30	839	46	939	30
40	21	140	33	240	13	340	12	440	19	540	17	640	29	740	35	840	57	940	13
41	27	141	3S	241	23	341	10	441	34	541	31	641	24	741	34	841	43	941	19
42	37	142	23	242	19	342	10	442	34	542	21	642	18	742	43	842	50	942	20
43	28	143	28	243	24	343	17	443	33	543	29	643	36	743	21	843	39	943	27
44	41	144	23	244	19	344	12	444	26	544	20	644	15	744	42	844	55	944	14
45	45	145	31	245	27	345	24	445	43	545	21	645	33	745	30	845	38	945	21
46	40	146	29	246	20	345	22	446	31	546	25	646	21	746	29	846	29	946	32
47	32	147	24	247	19	347	14	447	30	547	35	647	25	747	29	347	32	947	18
48	45	148	22	248	28	348	14	448	41	548	24	648	25	748	41	848	27	948	25

49	48	149	30	249	19	349	9	449	15	549	25	649	19	749	35	849	22	949	13
50	51	150	21	250	29	350	19	450	23	550	23	650	23	750	29	850	23	950	25
51	51	151	24	251	24	351	15	451	25	551	27	651	18	751	37	851	25	951	19
52	21	152	21	252	33	352	9	452	27	552	35	652	26	752	31	852	19	952	27
53	43	153	30	253	20	353	18	453	40	553	36	653	27	753	24	853	29	953	27
54	42	154	25	254	29	354	17	454	40	554	33	654	11	754	47	854	34	954	18
55	56	155	17	255	17	355	15	455	34	555	27	655	20	755	3	855	27	955	25
55	51	155	22	255	19	355	21	455	42	555	33	656	13	755	34	355	30	956	26
57	51	157	18	257	23	357	22	457	12	557	25	657	20	757	35	857	30	957	39
58	60	158	19	258	26	358	17	458	24	558	32	658	23	758	39	858	24	958	59
59	35	159	20	259	25	359	21	459	20	559	23	659	19	759	29	859	33	959	34
60	43	160	22	260	32	360	26	460	26	560	42	660	21	760	41	860	29	960	34
61	42	161	39	261	21	361	23	461	46	561	25	661	21	761	36	861	36	961	24
62	55	162	35	262	15	362	20	462	35	562	33	662	29	762	50	862	29	962	25
63	46	163	29	263	20	363	28	463	46	563	19	663	18	763	33	863	27	963	40
64	49	164	24	264	19	364	34	464	33	564	40	664	25	764	38	864	32	964	19
65	40	165	22	265	13	365	23	465	27	565	35	665	24	765	40	865	30	965	35
66	33	166	26	266	25	366	20	466	35	566	36	666	19	766	41	866	23	966	34
67	45	167	27	267	19	367	37	467	33	567	33	667	15	767	34	867	25	967	33
68	37	168	28	268	9	368	22	468	29	568	25	668	23	768	42	868	23	968	29
69	44	169	36	269	20	369	32	469	45	569	33	669	14	769	40	869	29	969	23
70	50	170	31	270	20	370	41	470	18	570	33	670	16	770	50	870	26	970	29
71	37	171	31	271	21	371	35	471	21	571	27	671	16	771	30	871	29	971	25
72	36	172	34	272	21	372	41	472	35	572	33	672	22	772	34	872	22	972	19
73	43	173	19	273	20	373	43	473	39	573	33	673	13	773	28	873	16	973	34
74	49	174	37	274	13	374	33	474	40	574	39	674	19	774	21	874	25	974	37
75	4C	175	39	275	25	375	32	475	33	575	30	675	23	775	24	875	26	975	34

(continued)

TABLE 4.12 (*Continued*)

Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count	Day	Count
76	65	176	32	276	29	376	28	476	35	576	33	676	14	776	37	876	25	976	29
77	49	177	36	277	16	377	42	477	20	577	26	677	16	777	44	877	24	977	27
78	49	178	42	278	18	378	27	478	36	578	26	678	10	778	39	878	29	978	18
79	34	179	31	279	32	379	25	479	34	579	23	679	14	779	37	879	34	979	26
80	33	180	28	280	32	380	32	480	35	580	24	680	13	780	35	880	35	980	28
81	29	181	35	281	19	381	27	481	36	581	32	681	15	781	32	881	29		
82	32	182	36	282	24	382	35	482	29	582	24	682	15	782	41	882	39		
83	57	183	35	283	18	383	26	483	19	583	32	683	11	783	41	883	31		
84	43	184	32	284	20	384	32	484	36	584	41	634	11	784	51	884	26		
85	40	185	26	285	20	385	42	485	35	585	26	685	18	785	43	885	24		
86	46	186	29	286	20	386	38	486	31	586	28	686	16	786	35	886	31		
87	33	187	25	287	24	387	36	487	23	587	25	687	18	787	33	887	24		
88	30	188	23	288	15	388	26	488	31	588	29	688	15	788	33	888	29		
89	41	189	29	289	22	389	26	489	29	589	40	689	16	789	31	889	26		
90	38	190	29	290	16	390	24	490	44	590	34	690	11	790	43	890	45		
91	29	191	26	291	14	391	30	491	42	591	41	691	11	791	45	891	36		
92	41	192	18	292	17	392	32	492	31	592	37	692	23	792	43	892	29		
93	28	193	19	293	15	393	14	493	31	593	36	693	20	793	42	893	22		
94	47	194	17	294	8	394	27	494	24	594	26	694	18	794	36	894	31		
95	42	195	22	295	23	395	26	495	30	595	42	695	24	795	34	895	38		
96	34	196	25	296	17	396	25	496	26	596	40	696	14	796	30	896	36		
97	40	197	33	297	13	397	23	497	26	597	34	697	22	797	46	897	33		
98	35	198	10	298	15	398	27	498	39	598	41	698	16	798	54	898	34		
99	40	199	25	299	15	399	36	499	35	599	37	699	26	799	52	899	34		
100	24	200	25	300	13	400	40	500	34	600	36	700	17	800	39	900	25		

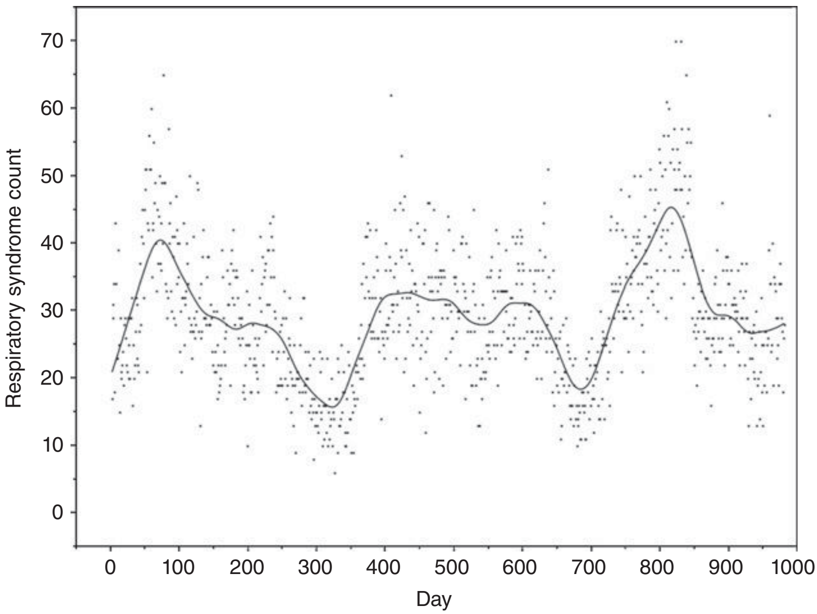


FIGURE 4.33 Time series plot of daily respiratory syndrome count, with kernel-smoothed fitted line. ($\alpha = 0.1$).

Over the 2 $^1/_2$ year period, the daily counts of the respiratory syndrome appear to follow a weak seasonal pattern, with the highest peak in November–December (late fall), a secondary peak in March–April, and then decreasing to the lowest counts in June–August (summer). The amplitude, or range within a year, seems to vary, but counts do not appear to be increasing or decreasing over time.

Not immediately evident from the time series plots is a potential day effect. The box plots of the residuals from the loess smoothed line in Figure 4.33 are plotted in Figure 4.34 versus day of the week. These plots exhibit variation that indicates slightly higher-than-expected counts on Monday and slightly lower-than-expected counts on Thursday, Friday, and Saturday.

The exponential smoothing procedure in JMP was applied to the respiratory syndrome data. The results of first-order or simple exponential smoothing are summarized in Table 4.13 and Figure 4.35, which plots only the last 100 observations along with the smoothed values. JMP reported the value of the smoothing constant that produced the minimum value of the error sum of squares as $\lambda = 0.21$. This value also minimizes the AIC and BIC criteria, and results in the smallest values of the mean absolute

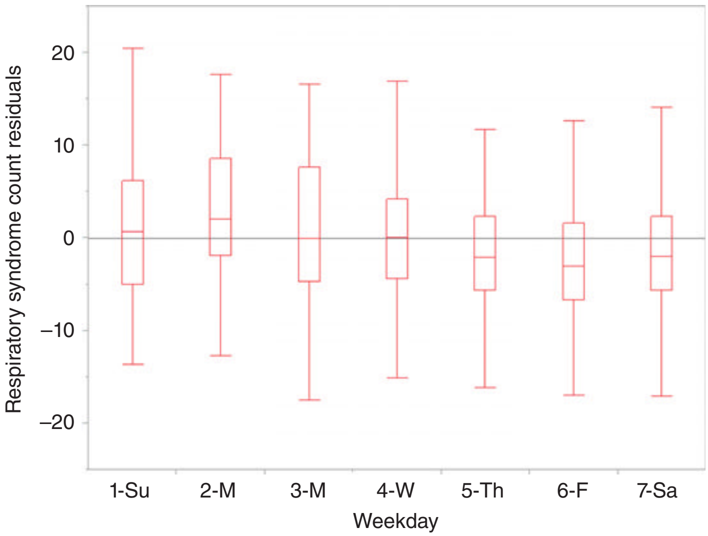


FIGURE 4.34 Box plots of residuals from the kernel-smoothed line fit to daily respiratory syndrome count.

prediction error and the mean absolute, although there is very little difference between the optimal value of $\lambda = 0.21$ and the values $\lambda = 0.1$ and $\lambda = 0.4$.

The results of using second-order exponential smoothing are summarized in Table 4.14 and illustrated graphically for the last 100 observations in Figure 4.36. There is not a lot of difference between the two procedures, although the optimal first-order smoother does perform slightly better and the larger smoothing parameters in the double smoother perform more poorly.

Single and double exponential smoothing do not account for the apparent mild seasonality observed in the original time series plot of the data.

TABLE 4.13 First-Order Simple Exponential Smoothing Applied to the Respiratory Data

Model	Variance	AIC	BIC	MAPE	MAE
First-Order Exponential (min SSE, $\lambda = 0.21$)	52.66	6660.81	6665.70	21.43	5.67
First-Order Exponential ($\lambda = 0.1$)	55.65	6714.67	6714.67	22.23	5.85
First-Order Exponential ($\lambda = 0.4$)	55.21	6705.63	6705.63	21.87	5.82

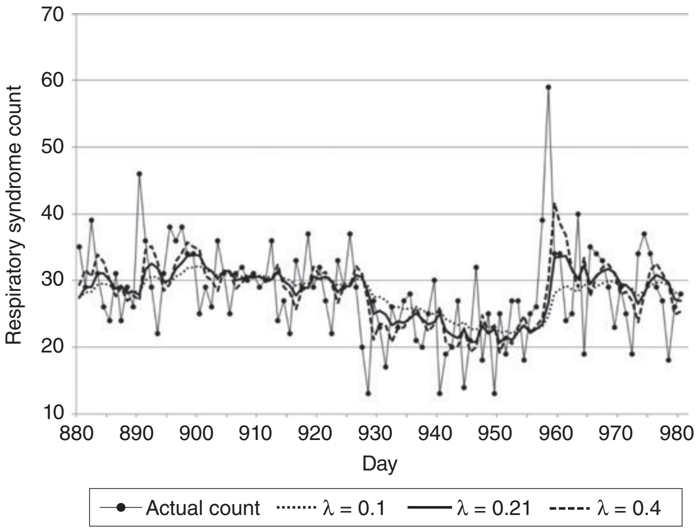


FIGURE 4.35 Respiratory syndrome counts using first-order exponential smoothing with $\lambda = 0.1$, $\lambda = 0.21$ (min SSE), and $\lambda = 0.4$.

We used JMP to fit Winters' additive seasonal model to the respiratory syndrome count data. Because the seasonal patterns are not strong, we investigated seasons of length 3, 7, and 12 periods. The results are summarized in Table 4.15 and illustrated graphically for the last 100 observations in Figure 4.37. The 7-period season works best, probably reflecting the daily seasonal pattern that we observed in Figure 4.34. This is also the best smoother of all the techniques that were investigated. The values of $\lambda = 0$ for the trend and seasonal components in this model are an indication that there is not a significant linear trend in the data and that the seasonal pattern is relatively stable over the period of available data.

TABLE 4.14 Second-Order Simple Exponential Smoothing Applied to the Respiratory Data

Model	Variance	AIC	BIC	MAPE	MAE
Second-Order Exponential (min SSE, $\lambda = 0.10$)	54.37	6690.98	6695.86	21.71	5.78
Second-Order Exponential ($\lambda = 0.2$)	58.22	6754.37	6754.37	22.44	5.98
Second-Order Exponential ($\lambda = 0.4$)	74.46	6992.64	6992.64	25.10	6.74

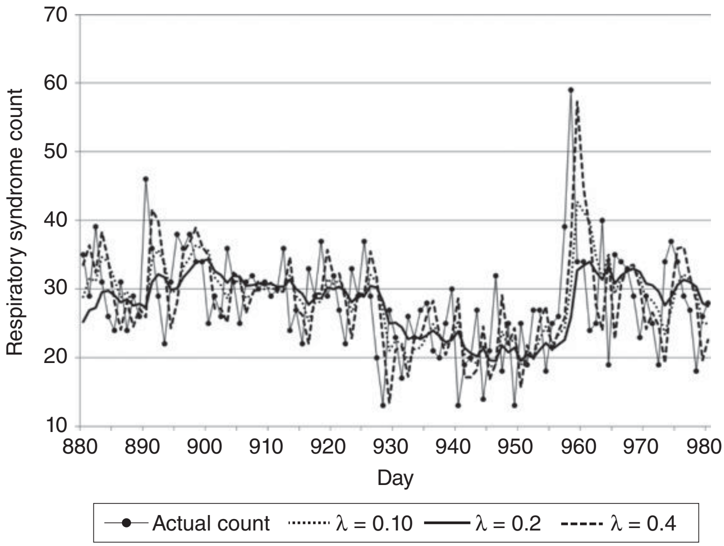


FIGURE 4.36 Respiratory syndrome counts using second-order exponential smoothing with $\lambda = 0.10$ (min SSE), $\lambda = 0.2$, and $\lambda = 0.4$.

TABLE 4.15 Winters' Additive Seasonal Exponential Smoothing Applied to the Respiratory Data

Model	Variance	AIC	BIC	MAPE	MAE
$S = 3$					
Winters Additive (min SSE, $\lambda 1 = 0.21$, $\lambda 2 = 0$, $\lambda 3 = 0$)	52.75	6662.75	6677.40	21.70	5.72
Winters Additive ($\lambda 1 = 0.2$, $\lambda 2 = 0.1$, $\lambda 3 = 0.1$)	57.56	6731.59	6731.59	22.38	5.94
$S = 7$					
Winters Additive (min SSE, $\lambda 1 = 0.22$, $\lambda 2 = 0$, $\lambda 3 = 0$)	49.77	6593.83	6608.47	21.10	5.56
Winters Additive ($\lambda 1 = 0.2$, $\lambda 2 = 0.1$, $\lambda 3 = 0.1$)	54.27	6652.57	6652.57	21.47	5.70
$S = 12$					
Winters Additive (min SSE, $\lambda 1 = 0.21$, $\lambda 2 = 0$, $\lambda 3 = 0$)	52.74	6635.58	6650.21	22.13	5.84
Winters Additive ($\lambda 1 = 0.2$, $\lambda 2 = 0.1$, $\lambda 3 = 0.1$)	58.76	6703.79	6703.79	22.77	6.08

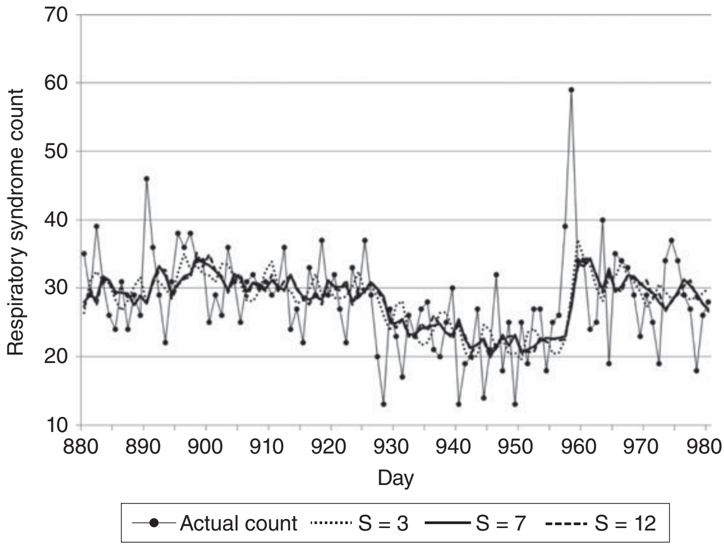


FIGURE 4.37 Respiratory syndrome counts using winters' additive seasonal exponential smoothing with $S = 3$, $S = 7$, and $S = 12$, and smoothing parameters that minimize SSE.

4.9 EXPONENTIAL SMOOTHERS AND ARIMA MODELS

The first-order exponential smoother presented in Section 4.2 is a very effective model in forecasting. The discount factor, λ, makes this smoother fairly flexible in handling time series data with various characteristics. The first-order exponential smoother is particularly good in forecasting time series data with certain specific characteristics.

Recall that the first-order exponential smoother is given as

$$\tilde{y}_T = \lambda y_T + (1 - \lambda)\tilde{y}_{T-1} \tag{4.58}$$

and the forecast error is defined as

$$e_T = y_T - \hat{y}_{T-1}. \tag{4.59}$$

Similarly, we have

$$e_{T-1} = y_{T-1} - \hat{y}_{T-2}. \tag{4.60}$$

By multiplying Eq. (4.60) by $(1 - \lambda)$ and subtracting it from Eq. (4.59), we obtain

$$\begin{aligned} e_T - (1-\lambda)e_{T-1} &= (y_T - \hat{y}_{T-1}) - (1-\lambda)(y_{T-1} - \hat{y}_{T-2}) \\ &= y_T - y_{T-1} - \hat{y}_{T-1} + \underbrace{\lambda y_{T-1} + (1-\lambda)\hat{y}_{T-2}}_{=\hat{y}_{T-1}} \\ &= y_T - y_{T-1} - \hat{y}_{T-1} + \hat{y}_{T-1} \\ &= y_T - y_{T-1}. \end{aligned} \tag{4.61}$$

We can rewrite Eq. (4.61) as

$$y_T - y_{T-1} = e_T - \theta e_{T-1}, \tag{4.62}$$

where $\theta = 1 - \lambda$. Recall from Chapter 2 the **backshift** *operator, B*, defined as $B(y_t) = y_{t-1}$. Thus Eq. (4.62) becomes

$$(1 - B)y_T = (1 - \theta B)e_T. \tag{4.63}$$

We will see in Chapter 5 that the model in Eq. (4.63) is called the **integrated moving average** model denoted as IMA(1,1), for the backshift operator is used only once on y_T and only once on the error. It can be shown that if the process exhibits the dynamics defined in Eq. (4.63), that is an IMA(1,1) process, the first-order exponential smoother provides minimum mean squared error (MMSE) forecasts (see Muth (1960), Box and Luceno (1997), and Box, Jenkins, and Reinsel (1994)). For more discussion of the equivalence between exponential smoothing techniques and the ARIMA models, see Abraham and Ledolter (1983), Cogger (1974), Goodman (1974), Pandit and Wu (1974), and McKenzie (1984).

4.10 R COMMANDS FOR CHAPTER 4

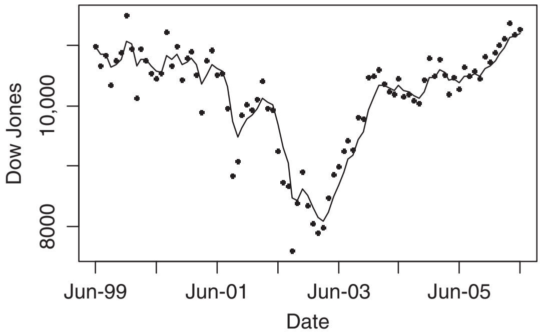
Example 4.1 The Dow Jones index data are in the second column of the array called dji.data in which the first column is the month of the year. We can use the following simple function to obtain the first-order exponential smoothing

```
firstsmooth<-function(y,lambda,start=y[1]){
      ytilde<-y
      ytilde[1]<-lambda*y[1]+(1-lambda)*start
      for (i in 2:length(y)){
            ytilde[i]<-lambda*y[i]+(1-lambda)*ytilde[i-1]
      }
ytilde
}
```

Note that this function uses the first observation as the starting value by default. One can change this by providing a specific start value when calling the function.

We can then obtain the smoothed version of the data for a specified lambda value and plot the fitted value as the following:

```
dji.smooth1<-firstsmooth(y=dji.data[,2],lambda=0.4)
plot(dji.data[,2],type="p", pch=16,cex=.5,xlab='Date',ylab='Dow
    Jones',xaxt='n')
axis(1, seq(1,85,12), dji.data[seq(1,85,12),1])
lines(dji.smooth1)
```



For the first-order exponential smoothing, measures of accuracy such as MAPE, MAD, and MSD can be obtained from the following function:

```
measacc.fs<- function(y,lambda){
            out<- firstsmooth(y,lambda)
            T<-length(y)
            #Smoothed version of the original is the one step
              ahead prediction
            #Hence the predictions (forecasts) are given as
            pred<-c(y[1],out[1:(T-1)])
            prederr<- y-pred
            SSE<-sum(prederr^2)
            MAPE<-100*sum(abs(prederr/y))/T
            MAD<-sum(abs(prederr))/T
            MSD<-sum(prederr^2)/T
            ret1<-c(SSE,MAPE,MAD,MSD)
            names(ret1)<-c("SSE","MAPE","MAD","MSD")
            return(ret1)
}

measacc.fs(dji.data[,2],0.4)
         SSE         MAPE           MAD           MSD
1.665968e+07 3.461342e+00 3.356325e+02 1.959962e+05
```

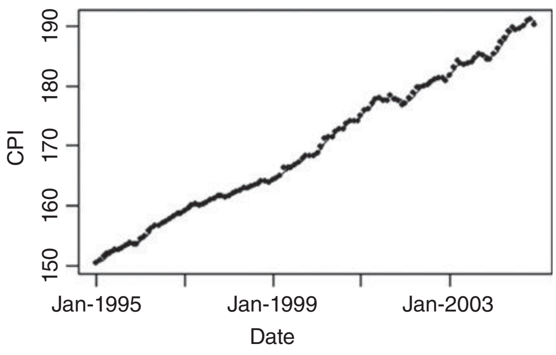
Note that alternatively we could use the Holt–Winters function from the stats package. The function requires three parameters (alpha, beta, and gamma) to be defined. Providing a specific value for alpha and setting beta and gamma to "FALSE" give the first-order exponential as the following

```
dji1.fit<-HoltWinters(dji.data[,2],alpha=.4, beta=FALSE, gamma=FALSE)
```

Beta corresponds to the second-order smoothing (or the trend term) and gamma is for the seasonal effect.

Example 4.2 The US CPI data are in the second column of the array called cpi.data in which the first column is the month of the year. For this case we use the firstsmooth function twice to obtain the double exponential smoothing as

```
cpi.smooth1<-firstsmooth(y=cpi.data[,2],lambda=0.3)
cpi.smooth2<-firstsmooth(y=cpi.smooth1,lambda=0.3)
cpi.hat<-2*cpi.smooth1-cpi.smooth2 #Equation 4.23
plot(cpi.data[,2],type="p", pch=16,cex=.5,xlab='Date',ylab='CPI',
  xaxt='n')
axis(1, seq(1,120,24), cpi.data[seq(1,120,24),1])
lines(cpi.hat)
```

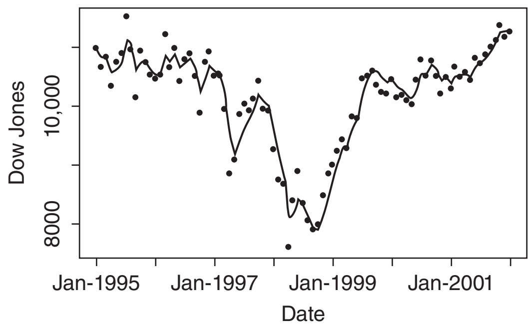


Note that the fitted values are obtained using Eq. (4.23). Also the corresponding command using Holt–Winters function is

```
HoltWinters(cpi.data[,2],alpha=0.3, beta=0.3, gamma=FALSE)
```

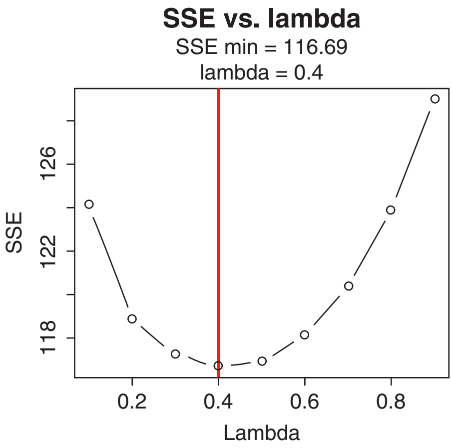
Example 4.3 In this example we use the firstsmooth function twice for the Dow Jones Index data to obtain the double exponential smoothing as in the previous example.

```
dji.smooth1<-firstsmooth(y=dji.data[,2],lambda=0.3)
dji.smooth2<-firstsmooth(y=dji.smooth1,lambda=0.3)
dji.hat<-2*dji.smooth1-dji.smooth2 #Equation 4.23
plot(dji.data[,2],type="p", pch=16,cex=.5,xlab='Date',ylab='Dow
 Jones',xaxt='n')
axis(1, seq(1,85,12), cpi.data[seq(1,85,12),1])
lines(dji.hat)
```



Example 4.4 The average speed data are in the second column of the array called speed.data in which the first column is the index for the week. To find the “best” smoothing constant, we will use the firstsmooth function for various lambda values and obtain the sum of squared one-step-ahead prediction error (SS_E) for each. The lambda value that minimizes the sum of squared prediction errors is deemed the “best” lambda. The obvious option is to apply firstsmooth function in a for loop to obtain SS_E for various lambda values. Even though in this case this may not be an issue, in many cases for loops can slow down the computations in R and are to be avoided if possible. We will do that using sapply function.

```
lambda.vec<-seq(0.1, 0.9, 0.1)
sse.speed<-function(sc){measacc.fs(speed.data[1:78,2],sc)[1]}
sse.vec<-sapply(lambda.vec, sse.speed)
opt.lambda<-lambda.vec[sse.vec == min(sse.vec)]
plot(lambda.vec, sse.vec, type="b", main = "SSE vs. lambda\n",
  xlab='lambda\n',ylab='SSE')
abline(v=opt.lambda, col = 'red')
mtext(text = paste("SSE min = ", round(min(sse.vec),2), "\n lambda
  = ", opt.lambda))
```



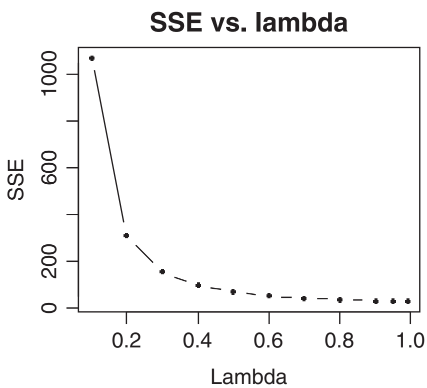
Note that we can also use Holt–Winters function to find the “best” value for the smoothing constant by not specifying the appropriate parameter as the following:

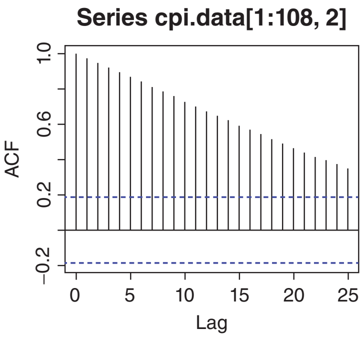
```
HoltWinters(speed.data[,2], beta=FALSE, gamma=FALSE)
```

Example 4.5 We will first try to find the best lambda for the CPI data using first-order exponential smoothing. We will also plot ACF of the data.

Note that we will use the data up to December 2003.

```
lambda.vec<-c(seq(0.1, 0.9, 0.1), .95, .99)
sse.cpi<-function(sc){measacc.fs(cpi.data[1:108,2],sc)[1]}
sse.vec<-sapply(lambda.vec, sse.cpi)
opt.lambda<-lambda.vec[sse.vec == min(sse.vec)]
plot(lambda.vec, sse.vec, type="b", main = "SSE vs. lambda\n",
  xlab='lambda\n',ylab='SSE', pch=16,cex=.5)
acf(cpi.data[1:108,2],lag.max=25)
```

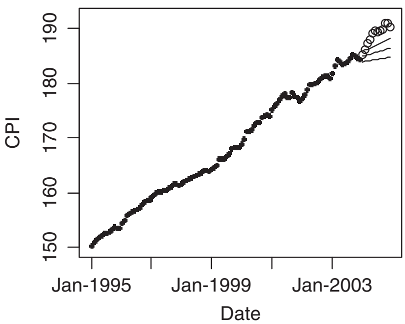




We now use the second-order exponential smoothing with lambda of 0.3. We calculate the forecasts using Eq. (4.31) for the two options suggested in the Example 4.5.

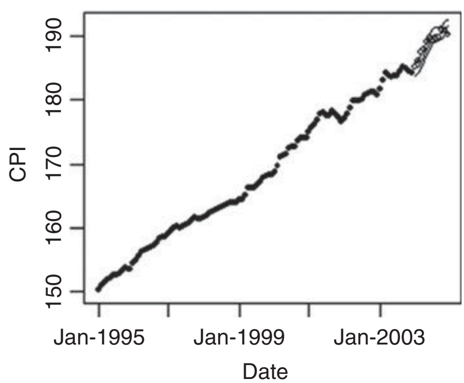
Option 1: On December 2003, make the forecasts for the entire 2004 (1- to 12-step-ahead forecasts).

```
lcpi<-0.3
cpi.smooth1<-firstsmooth(y=cpi.data[1:108,2],lambda=lcpi)
cpi.smooth2<-firstsmooth(y=cpi.smooth1,lambda=lcpi)
cpi.hat<-2*cpi.smooth1-cpi.smooth2
tau<-1:12
T<-length(cpi.smooth1)
cpi.forecast<-(2+tau*(lcpi/(1-lcpi)))*cpi.smooth1[T]-(1+tau*(lcpi/
  (1-lcpi)))*cpi.smooth2[T]
ctau<-sqrt(1+(lcpi/((2-lcpi)^3))*(10-14*lcpi+5*(lcpi^2)+2*tau*lcpi
  *(4-3*lcpi)+2*(tau^2)*(lcpi^2)))
alpha.lev<-.05
sig.est<- sqrt(var(cpi.data[2:108,2]- cpi.hat[1:107]))
cl<-qnorm(1-alpha.lev/2)*(ctau/ctau[1])*sig.est
plot(cpi.data[1:108,2],type="p", pch=16,cex=.5,xlab='Date',
  ylab='CPI',xaxt='n',xlim=c(1,120),ylim=c(150,192))
axis(1, seq(1,120,24), cpi.data[seq(1,120,24),1])
points(109:120,cpi.data[109:120,2])
lines(109:120,cpi.forecast)
lines(109:120,cpi.forecast+cl)
lines(109:120,cpi.forecast-cl)
```



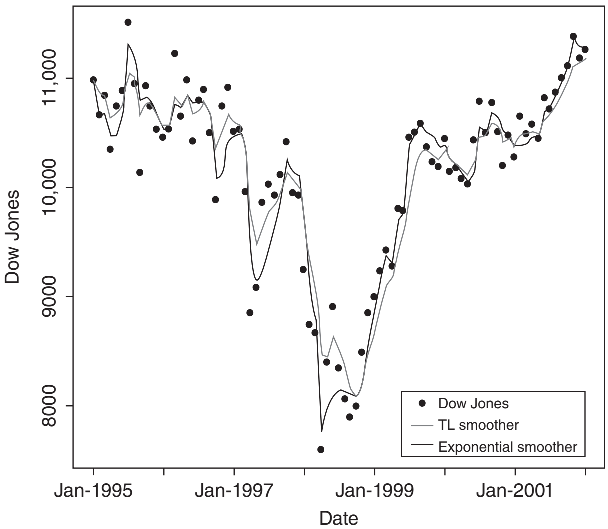
Option 2: On December 2003, make the forecast for January 2004. Then when January 2004 data are available, make the forecast for February 2004 (only one-step-ahead forecasts).

```
lcpi<-0.3
T<-108
tau<-12
alpha.lev<-.05
cpi.forecast<-rep(0,tau)
cl<-rep(0,tau)
cpi.smooth1<-rep(0,T+tau)
cpi.smooth2<-rep(0,T+tau)

for (i in 1:tau) {
      cpi.smooth1[1:(T+i-1)]<-firstsmooth(y=cpi.data[1:(T+i-1),2],
        lambda=lcpi)
      cpi.smooth2[1:(T+i-1)]<-firstsmooth(y=cpi.smooth1[1:(T+i-1)],
        lambda=lcpi)
      cpi.forecast[i]<-(2+(lcpi/(1-lcpi)))*cpi.smooth1[T+i-1]-
        (1+(lcpi/(1-lcpi)))*cpi.smooth2[T+i-1]
      cpi.hat<-2*cpi.smooth1[1:(T+i-1)]-cpi.smooth2[1:(T+i-1)]
      sig.est<- sqrt(var(cpi.data[2:(T+i-1),2]- cpi.hat[1:(T+i-2)]))
      cl[i]<-qnorm(1-alpha.lev/2)*sig.est
      }

plot(cpi.data[1:T,2],type="p", pch=16,cex=.5,xlab='Date',ylab='CPI',
  xaxt='n',xlim=c(1,T+tau),ylim=c(150,192))
axis(1, seq(1,T+tau,24), cpi.data[seq(1,T+tau,24),1])
points((T+1):(T+tau),cpi.data[(T+1):(T+tau),2],cex=.5)
lines((T+1):(T+tau),cpi.forecast)
lines((T+1):(T+tau),cpi.forecast+cl)
lines((T+1):(T+tau),cpi.forecast-cl)
```



Example 4.6 The function for the Trigg–Leach smoother is given as:

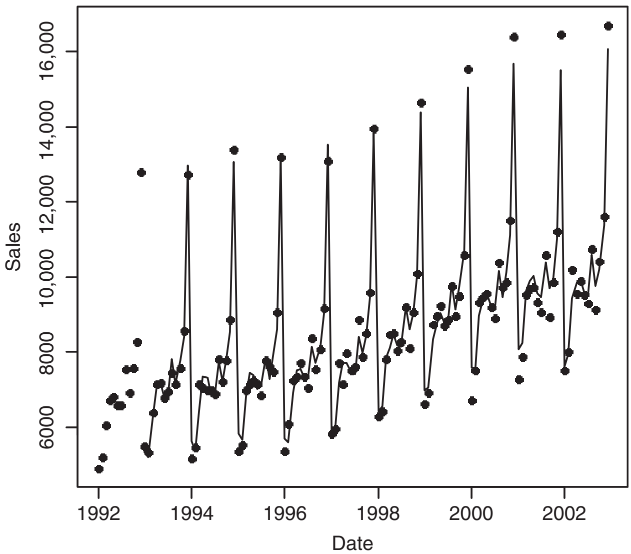
```
tlsmooth<-function(y,gamma,y.tilde.start=y[1],lambda.start=1){
      T<-length(y)

#Initialize the vectors
      Qt<-vector()
      Dt<-vector()
      y.tilde<-vector()
      lambda<-vector()
      err<-vector()
#Set the starting values for the vectors
      lambda[1]=lambda.start
      y.tilde[1]=y.tilde.start
      Qt[1]<-0
      Dt[1]<-0
      err[1]<-0

      for (i in 2:T){
            err[i]<-y[i]-y.tilde[i-1]
            Qt[i]<-gamma*err[i]+(1-gamma)*Qt[i-1]
            Dt[i]<-gamma*abs(err[i])+(1-gamma)*Dt[i-1]
            lambda[i]<-abs(Qt[i]/Dt[i])
            y.tilde[i]=lambda[i]*y[i] + (1-lambda[i])*y.tilde[i-1]
      }
return(cbind(y.tilde,lambda,err,Qt,Dt))
}

#Obtain the TL smoother for Dow Jones Index
out.tl.dji<-tlsmooth(dji.data[,2],0.3)
```



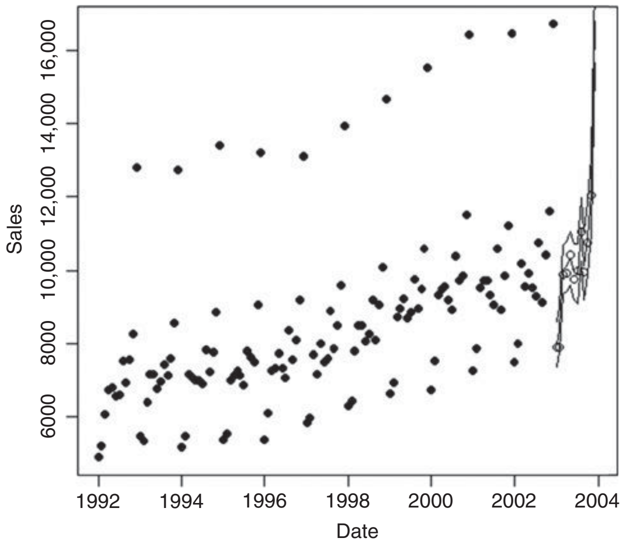
```
#Obtain the exponential smoother for Dow Jones Index
dji.smooth1<-firstsmooth(y=dji.data[,2],lambda=0.4)

#Plot the data together with TL and exponential smoother for
  comparison
plot(dji.data[,2],type="p", pch=16,cex=.5,xlab='Date',ylab='Dow
  Jones',xaxt='n')
axis(1, seq(1,85,12), cpi.data[seq(1,85,12),1])
lines(out.tl.dji[,1])
lines(dji.smooth1,col="grey40")
legend(60,8000,c("Dow Jones","TL Smoother","Exponential Smoother"),
  pch=c(16, NA, NA),lwd=c(NA,.5,.5),cex=.55,col=c("black",
  "black","grey40"))
```

Example 4.7 The clothing sales data are in the second column of the array called closales.data in which the first column is the month of the year. We will use the data up to December 2002 to fit the model and make forecasts for the coming year (2003). We will use Holt–Winters function given in stats package. The model is additive seasonal model with all parameters equal to 0.2.

```
dat.ts = ts(closales.data[,2], start = c(1992,1), freq = 12)
y1<-closales.data[1:132,]
# convert data to ts object
y1.ts<-ts(y1[,2], start = c(1992,1), freq = 12)
clo.hw1<-HoltWinters(y1.ts,alpha=0.2,beta=0.2,gamma=0.2,seasonal
  ="additive")
plot(y1.ts,type="p", pch=16,cex=.5,xlab='Date',ylab='Sales')
lines(clo.hw1$fitted[,1])
```

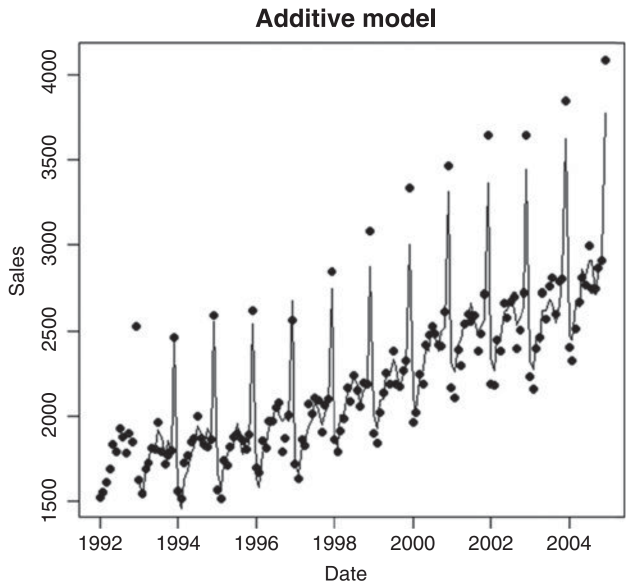


```
#Forecast the the sales for 2003
y2<-closales.data[133:144,]
y2.ts<-ts(y2[,2],start=c(2003,1),freq=12)

y2.forecast<-predict(clo.hw1, n.ahead=12, prediction.interval
  = TRUE)
plot(y1.ts,type="p", pch=16,cex=.5,xlab='Date',ylab='Sales',
  xlim=c(1992,2004))
points(y2.ts)
lines(y2.forecast[,1])
lines(y2.forecast[,2])
lines(y2.forecast[,3])
```

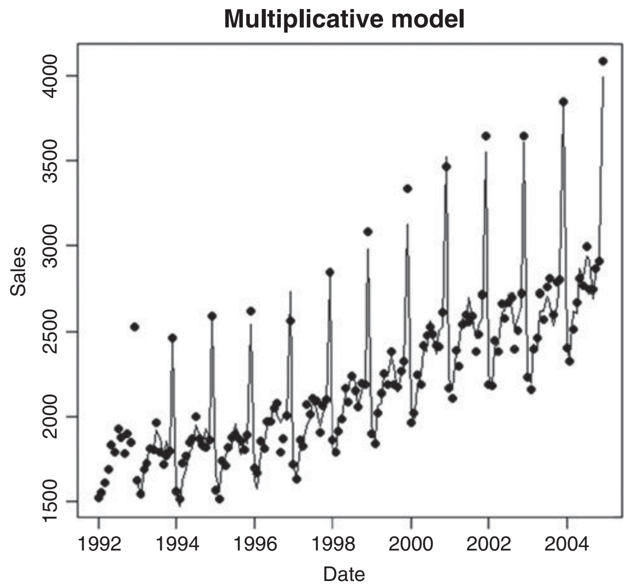


Example 4.8 The liquor store sales data are in the second column of the array called liqsales.data in which the first column is the month of the year. We will first fit additive and multiplicative seasonal models to the entire data to see the difference in the fits. Then we will use the data up to December 2003 to fit the multiplicative model and make forecasts for the coming year (2004). We will once again use Holt–Winters function given in stats package. In all cases we set all parameters to 0.2.

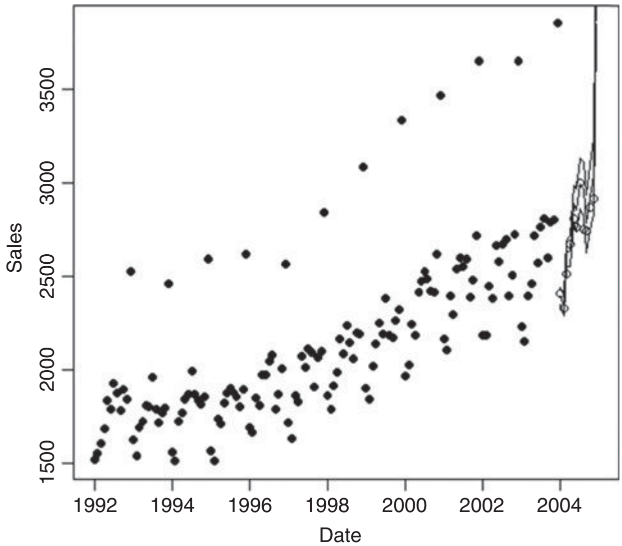
```
y.ts<- ts(liqsales.data[,2], start = c(1992,1), freq = 12)

liq.hw.add<-HoltWinters(y.ts,alpha=0.2,beta=0.2,gamma=0.2,
  seasonal="additive")
plot(y.ts,type="p", pch=16,cex=.5,xlab='Date',ylab='Sales',
  main="Additive Model")
lines(liq.hw.add$fitted[,1])
```




```
liq.hw.mult<-HoltWinters(y.ts,alpha=0.2,beta=0.2,gamma=0.2,
  seasonal="multiplicative")
plot(y.ts,type="p", pch=16,cex=.5,xlab='Date',ylab='Sales',
  main="Multiplicative Model")
lines(liq.hw.mult$fitted[,1])
```



```
y1<-liqsales.data[1:144,]
y1.ts<-ts(y1[,2], start = c(1992,1), freq = 12)
liq.hw1<-HoltWinters(y1.ts,alpha=0.2,beta=0.2,gamma=0.2,
  seasonal="multiplicative")
y2<-liqsales.data[145:156,]
y2.ts<-ts(y2[,2],start=c(2004,1),freq=12)

y2.forecast<-predict(liq.hw1, n.ahead=12, prediction.interval =
  TRUE)
plot(y1.ts,type="p", pch=16,cex=.5,xlab='Date',ylab='Sales',
  xlim=c(1992,2005))
points(y2.ts)
lines(y2.forecast[,1])
lines(y2.forecast[,2])
lines(y2.forecast[,3])
```



EXERCISES

4.1 Consider the time series data shown in Table E4.1.

a. Make a time series plot of the data.

b. Use simple exponential smoothing with $\lambda = 0.2$ to smooth the first 40 time periods of this data. How well does this smoothing procedure work?

c. Make one-step-ahead forecasts of the last 10 observations. Determine the forecast errors.

TABLE E4.1 Data for Exercise 4.1

Period	y_t	Period	y_t	Period	y_t	Period	y_t	Period	y_t
1	48.7	11	49.1	21	45.3	31	50.8	41	47.9
2	45.8	12	46.7	22	43.3	32	46.4	42	49.5
3	46.4	13	47.8	23	44.6	33	52.3	43	44.0
4	46.2	14	45.8	24	47.1	34	50.5	44	53.8
5	44.0	15	45.5	25	53.4	35	53.4	45	52.5
6	53.8	16	49.2	26	44.9	36	53.9	46	52.0
7	47.6	17	54.8	27	50.5	37	52.3	47	50.6
8	47.0	18	44.7	28	48.1	38	53.0	48	48.7
9	47.6	19	51.1	29	45.4	39	48.6	49	51.4
10	51.1	20	47.3	30	51.6	40	52.4	50	47.7

4.2 Reconsider the time series data shown in Table E4.1.

a. Use simple exponential smoothing with the optimum value of λ to smooth the first 40 time periods of this data (you can find the optimum value from Minitab). How well does this smoothing procedure work? Compare the results with those obtained in Exercise 4.1.

b. Make one-step-ahead forecasts of the last 10 observations. Determine the forecast errors. Compare these forecast errors with those from Exercise 4.1. How much has using the optimum value of the smoothing constant improved the forecasts?

4.3 Find the sample ACF for the time series in Table E4.1. Does this give you any insight about the optimum value of the smoothing constant that you found in Exercise 4.2?

4.4 Consider the time series data shown in Table E4.2.

a. Make a time series plot of the data.

b. Use simple exponential smoothing with $\lambda = 0.2$ to smooth the first 40 time periods of this data. How well does this smoothing procedure work?

c. Make one-step-ahead forecasts of the last 10 observations. Determine the forecast errors.

4.5 Reconsider the time series data shown in Table E4.2.

a. Use simple exponential smoothing with the optimum value of λ to smooth the first 40 time periods of this data (you can find the optimum value from Minitab). How well does this smoothing procedure work? Compare the results with those obtained in Exercise 4.4.

TABLE E4.2 Data for Exercise 4.4

Period	y_t	Period	y_t	Period	y_t	Period	y_t	Period	y_t
1	43.1	11	41.8	21	47.7	31	52.9	41	48.3
2	43.7	12	50.7	22	51.1	32	47.3	42	45.0
3	45.3	13	55.8	23	67.1	33	50.0	43	55.2
4	47.3	14	48.7	24	47.2	34	56.7	44	63.7
5	50.6	15	48.2	25	50.4	35	42.3	45	64.4
6	54.0	16	46.9	26	44.2	36	52.0	46	66.8
7	46.2	17	47.4	27	52.0	37	48.6	47	63.3
8	49.3	18	49.2	28	35.5	38	51.5	48	60.0
9	53.9	19	50.9	29	48.4	39	49.5	49	60.9
10	42.5	20	55.3	30	55.4	40	51.4	50	56.1

b. Make one-step-ahead forecasts of the last 10 observations. Determine the forecast errors. Compare these forecast errors with those from Exercise 4.4. How much has using the optimum value of the smoothing constant improved the forecasts?

4.6 Find the sample ACF for the time series in Table E4.2. Does this give you any insight about the optimum value of the smoothing constant that you found in Exercise 4.5?

4.7 Consider the time series data shown in Table E4.3.

a. Make a time series plot of the data.

b. Use simple exponential smoothing with $\lambda = 0.1$ to smooth the first 30 time periods of this data. How well does this smoothing procedure work?

TABLE E4.3 Data for Exercise 4.7

Period	y_t	Period	y_t	Period	y_t	Period	y_t	Period	y_t
1	275	11	297	21	231	31	255	41	293
2	245	12	235	22	238	32	255	42	284
3	222	13	237	23	251	33	229	43	276
4	169	14	203	24	253	34	286	44	290
5	236	15	238	25	283	35	236	45	250
6	259	16	232	26	283	36	194	46	235
7	268	17	206	27	245	37	228	47	275
8	225	18	295	28	234	38	244	48	350
9	246	19	247	29	273	39	241	49	290
10	263	20	227	30	293	40	284	50	269

c. Make one-step-ahead forecasts of the last 20 observations. Determine the forecast errors.

d. Plot the forecast errors on a control chart for individuals. Use a moving range chart to estimate the standard deviation of the forecast errors in constructing this chart. What conclusions can you draw about the forecasting procedure and the time series?

4.8 The data in Table E4.4 exhibit a linear trend.

a. Verify that there is a trend by plotting the data.

b. Using the first 12 observations, develop an appropriate procedure for forecasting.

c. Forecast the last 12 observations and calculate the forecast errors. Does the forecasting procedure seem to be working satisfactorily?

TABLE E4.4 Data for Exercise 4.8

Period	y_t	Period	y_t
1	315	13	460
2	195	14	395
3	310	15	390
4	316	16	450
5	325	17	458
6	335	18	570
7	318	19	520
8	355	20	400
9	420	21	420
10	410	22	580
11	485	23	475
12	420	24	560

4.9 Reconsider the linear trend data in Table E4.4. Take the first difference of this data and plot the time series of first differences. Has differencing removed the trend? Use exponential smoothing on the first 11 differences. Instead of forecasting the original data, forecast the first differences for the remaining data using exponential smoothing and use these forecasts of the first differences to obtain forecasts for the original data.

4.10 Table B.1 in Appendix B contains data on the market yield on US Treasury Securities at 10-year constant maturity.

a. Make a time series plot of the data.

b. Use simple exponential smoothing with $\lambda = 0.2$ to smooth the data, excluding the last 20 observations. How well does this smoothing procedure work?

c. Make one-step-ahead forecasts of the last 20 observations. Determine the forecast errors.

4.11 Reconsider the US Treasury Securities data shown in Table B.1.

a. Use simple exponential smoothing with the optimum value of λ to smooth the data, excluding the last 20 observations (you can find the optimum value from Minitab). How well does this smoothing procedure work? Compare the results with those obtained in Exercise 4.10.

b. Make one-step-ahead forecasts of the last 10 observations. Determine the forecast errors. Compare these forecast errors with those from Exercise 4.10. How much has using the optimum value of the smoothing constant improved the forecasts?

4.12 Table B.2 contains data on pharmaceutical product sales.

a. Make a time series plot of the data.

b. Use simple exponential smoothing with $\lambda = 0.1$ to smooth this data. How well does this smoothing procedure work?

c. Make one-step-ahead forecasts of the last 10 observations. Determine the forecast errors.

4.13 Reconsider the pharmaceutical sales data shown in Table B.2.

a. Use simple exponential smoothing with the optimum value of λ to smooth the data (you can find the optimum value from either Minitab or JMP). How well does this smoothing procedure work? Compare the results with those obtained in Exercise 4.12.

b. Make one-step-ahead forecasts of the last 10 observations. Determine the forecast errors. Compare these forecast errors with those from Exercise 4.12. How much has using the optimum value of the smoothing constant improved the forecasts?

c. Construct the sample ACF for these data. Does this give you any insight regarding the optimum value of the smoothing constant?

4.14 Table B.3 contains data on chemical process viscosity.

a. Make a time series plot of the data.

b. Use simple exponential smoothing with $\lambda = 0.1$ to smooth this data. How well does this smoothing procedure work?

c. Make one-step-ahead forecasts of the last 10 observations. Determine the forecast errors.

4.15 Reconsider the chemical process data shown in Table B.3.

a. Use simple exponential smoothing with the optimum value of λ to smooth the data (you can find the optimum value from either Minitab or JMP). How well does this smoothing procedure work? Compare the results with those obtained in Exercise 4.14.

b. Make one-step-ahead forecasts of the last 10 observations. Determine the forecast errors. Compare these forecast errors with those from Exercise 4.14.How much has using the optimum value of the smoothing constant improved the forecasts?

c. Construct the sample ACF for these data. Does this give you any insight regarding the optimum value of the smoothing constant?

4.16 Table B.4 contains data on the annual US production of blue and gorgonzola cheeses. This data have a strong trend.

a. Verify that there is a trend by plotting the data.

b. Develop an appropriate exponential smoothing procedure for forecasting.

c. Forecast the last 10 observations and calculate the forecast errors. Does the forecasting procedure seem to be working satisfactorily?

4.17 Reconsider the blue and gorgonzola cheese data in Table B.4 and Exercise 4.16. Take the first difference of this data and plot the time series of first differences. Has differencing removed the trend? Use exponential smoothing on the first differences. Instead of forecasting the original data, develop a procedure for forecasting the first differences and explain how you would use these forecasts of the first differences to obtain forecasts for the original data.

4.18 Table B.5 shows data for US beverage manufacturer product shipments. Develop an appropriate exponential smoothing procedure for forecasting these data.

4.19 Table B.6 contains data on the global mean surface air temperature anomaly.

a. Make a time series plot of the data.

b. Use simple exponential smoothing with $\lambda = 0.2$ to smooth the data. How well does this smoothing procedure work? Do you think this would be a reliable forecasting procedure?

4.20 Reconsider the global mean surface air temperature anomaly data shown in Table B.6 and used in Exercise 4.19.

a. Use simple exponential smoothing with the optimum value of λ to smooth the data (you can find the optimum value from either Minitab or JMP). How well does this smoothing procedure work? Compare the results with those obtained in Exercise 4.19.

b. Do you think using the optimum value of the smoothing constant would result in improved forecasts from exponential smoothing?

c. Take the first difference of this data and plot the time series of first differences. Use exponential smoothing on the first differences. Instead of forecasting the original data, develop a procedure for forecasting the first differences and explain how you would use these forecasts of the first differences to obtain forecasts for the original global mean surface air temperature anomaly.

4.21 Table B.7 contains daily closing stock prices for the Whole Foods Market.

a. Make a time series plot of the data.

b. Use simple exponential smoothing with $\lambda = 0.1$ to smooth the data. How well does this smoothing procedure work? Do you think this would be a reliable forecasting procedure?

4.22 Reconsider the Whole Foods Market data shown in Table B.7 and used in Exercise 4.21.

a. Use simple exponential smoothing with the optimum value of λ to smooth the data (you can find the optimum value from either Minitab or JMP). How well does this smoothing procedure work? Compare the results with those obtained in Exercise 4.21.

b. Do you think that using the optimum value of the smoothing constant would result in improved forecasts from exponential smoothing?

c. Use an exponential smoothing procedure for trends on this data. Is this an apparent improvement over the use of simple exponential smoothing with the optimum smoothing constant?

d. Take the first difference of this data and plot the time series of first differences. Use exponential smoothing on the first differences. Instead of forecasting the original data, develop a procedure for forecasting the first differences and explain how you would use these forecasts of the first differences to obtain forecasts for the stock price.

4.23 Unemployment rate data are given in Table B.8.

a. Make a time series plot of the data.

b. Use simple exponential smoothing with $\lambda = 0.2$ to smooth the data. How well does this smoothing procedure work? Do you think that simple exponential smoothing should be used to forecast this data?

4.24 Reconsider the unemployment rate data shown in Table B.8 and used in Exercise 4.23.

a. Use simple exponential smoothing with the optimum value of λ to smooth the data (you can find the optimum value from either Minitab or JMP). How well does this smoothing procedure work? Compare the results with those obtained in Exercise 4.23.

b. Do you think that using the optimum value of the smoothing constant would result in improved forecasts from exponential smoothing?

c. Use an exponential smoothing procedure for trends on this data. Is this an apparent improvement over the use of simple exponential smoothing with the optimum smoothing constant?

d. Take the first difference of this data and plot the time series of first differences. Use exponential smoothing on the first differences. Is this a reasonable procedure for forecasting the first differences?

4.25 Table B.9 contains yearly data on the international sunspot numbers.

a. Construct a time series plot of the data.

b. Use simple exponential smoothing with $\lambda = 0.1$ to smooth the data. How well does this smoothing procedure work? Do you think that simple exponential smoothing should be used to forecast this data?

4.26 Reconsider the sunspot data shown in Table B.9 and used in Exercise 4.25.

a. Use simple exponential smoothing with the optimum value of λ to smooth the data (you can find the optimum value from either Minitab or JMP). How well does this smoothing procedure work? Compare the results with those obtained in Exercise 4.25.

b. Do you think that using the optimum value of the smoothing constant would result in improved forecasts from exponential smoothing?

c. Use an exponential smoothing procedure for trends on this data. Is this an apparent improvement over the use of simple exponential smoothing with the optimum smoothing constant?

4.27 Table B.10 contains 7 years of monthly data on the number of airline miles flown in the United Kingdom. This is seasonal data.

a. Make a time series plot of the data and verify that it is seasonal.

b. Use Winters' multiplicative method for the first 6 years to develop a forecasting method for this data. How well does this smoothing procedure work?

c. Make one-step-ahead forecasts of the last 12 months. Determine the forecast errors. How well did your procedure work in forecasting the new data?

4.28 Reconsider the airline mileage data in Table B.10 and used in Exercise 4.27.

a. Use the additive seasonal effects model for the first 6 years to develop a forecasting method for this data. How well does this smoothing procedure work?

b. Make one-step-ahead forecasts of the last 12 months. Determine the forecast errors. How well did your procedure work in forecasting the new data?

c. Compare these forecasts with those found using Winters' multiplicative method in Exercise 4.27.

4.29 Table B.11 contains 8 years of monthly champagne sales data. This is seasonal data.

a. Make a time series plot of the data and verify that it is seasonal. Why do you think seasonality is present in these data?

b. Use Winters' multiplicative method for the first 7 years to develop a forecasting method for this data. How well does this smoothing procedure work?

c. Make one-step-ahead forecasts of the last 12 months. Determine the forecast errors. How well did your procedure work in forecasting the new data?

4.30 Reconsider the monthly champagne sales data in Table B.11 and used in Exercise 4.29.

a. Use the additive seasonal effects model for the first 7 years to develop a forecasting method for this data. How well does this smoothing procedure work?

b. Make one-step-ahead forecasts of the last 12 months. Determine the forecast errors. How well did your procedure work in forecasting the new data?

c. Compare these forecasts with those found using Winters' multiplicative method in Exercise 4.29.

4.31 Montgomery et al. (1990) give 4 years of data on monthly demand for a soft drink. These data are given in Table E4.5.

a. Make a time series plot of the data and verify that it is seasonal. Why do you think seasonality is present in these data?

b. Use Winters' multiplicative method for the first 3 years to develop a forecasting method for this data. How well does this smoothing procedure work?

c. Make one-step-ahead forecasts of the last 12 months. Determine the forecast errors. How well did your procedure work in forecasting the new data?

TABLE E4.5 Soft Drink Demand Data

Period	y_t	Period	y_t	Period	y_t	Period	y_t
1	143	13	189	25	359	37	332
2	191	14	326	26	264	38	244
3	195	15	289	27	315	39	320
4	225	16	293	28	362	40	437
5	175	17	279	29	414	41	544
6	389	18	552	30	647	42	830
7	454	19	674	31	836	43	1011
8	618	20	827	32	901	44	1081
9	770	21	1000	33	1104	45	1400
10	564	22	502	34	874	46	1123
11	327	23	512	35	683	47	713
12	235	24	300	36	352	48	487

4.32 Reconsider the soft drink demand data in Table E4.5 and used in Exercise 4.31.

a. Use the additive seasonal effects model for the first 3 years to develop a forecasting method for this data. How well does this smoothing procedure work?

b. Make one-step-ahead forecasts of the last 12 months. Determine the forecast errors. How well did your procedure work in forecasting the new data?

c. Compare these forecasts with those found using Winters' multiplicative method in Exercise 4.31.

4.33 Table B.12 presents data on the hourly yield from a chemical process and the operating temperature. Consider only the yield data in this exercise.

a. Construct a time series plot of the data.

b. Use simple exponential smoothing with $\lambda = 0.2$ to smooth the data. How well does this smoothing procedure work? Do you think that simple exponential smoothing should be used to forecast this data?

4.34 Reconsider the chemical process yield data shown in Table B.12.

a. Use simple exponential smoothing with the optimum value of λ to smooth the data (you can find the optimum value from either Minitab or JMP). How well does this smoothing procedure work? Compare the results with those obtained in Exercise 4.33.

b. How much has using the optimum value of the smoothing constant improved the forecasts?

4.35 Find the sample ACF for the chemical process yield data in Table B.12. Does this give you any insight about the optimum value of the smoothing constant that you found in Exercise 4.34?

4.36 Table B.13 presents data on ice cream and frozen yogurt sales. Develop an appropriate exponential smoothing forecasting procedure for this time series.

4.37 Table B.14 presents the CO_2 readings from Mauna Loa.

a. Use simple exponential smoothing with the optimum value of λ to smooth the data (you can find the optimum value from either Minitab or JMP).

b. Use simple exponential smoothing with $\lambda = 0.1$ to smooth the data. How well does this smoothing procedure work? Compare the results with those obtained using the optimum smoothing constant. How much has using the optimum value of the smoothing constant improved the exponential smoothing procedure?

4.38 Table B.15 presents data on the occurrence of violent crimes. Develop an appropriate exponential smoothing forecasting procedure for this time series.

4.39 Table B.16 presents data on the US. gross domestic product (GDP). Develop an appropriate exponential smoothing forecasting procedure for the GDP time series.

4.40 Total annual energy consumption is shown in Table B.17. Develop an appropriate exponential smoothing forecasting procedure for the energy consumption time series.

4.41 Table B.18 contains data on coal production. Develop an appropriate exponential smoothing forecasting procedure for the coal production time series.

4.42 Table B.19 contains data on the number of children 0−4 years old who drowned in Arizona.

a. Plot the data. What type of forecasting model seems appropriate?

b. Develop a forecasting model for this data?

4.43 Data on tax refunds and population are shown in Table B.20. Develop an appropriate exponential smoothing forecasting procedure for the tax refund time series.

4.44 Table B.21 contains data from the US Energy Information Administration on monthly average price of electricity for the residential sector in Arizona. This data have a strong seasonal component. Use the data from 2001–2010 to develop a multiplicative Winters-type exponential smoothing model for this data. Use this model to simulate one-month-ahead forecasts for the remaining years. Calculate the forecast errors. Discuss the reasonableness of the forecasts.

4.45 Use the electricity price data in Table B.21 from 2010–2010 and an additive Winters-type exponential smoothing procedure to develop a forecasting model.

a. Use this model to simulate one-month-ahead forecasts for the remaining years. Calculate the forecast errors. Discuss the reasonableness of the forecasts.

b. Compare the performance of this model with the multiplicative model you developed in Exercise 4.44.

4.46 Table B.22 contains data from the Danish Energy Agency on Danish crude oil production.

a. Plot the data and comment on any features that you observe from the graph. Calculate and plot the sample ACF and variogram. Interpret these graphs.

b. Use first-order exponential smoothing to develop a forecasting model for crude oil production. Plot the smoothed statistic on the same axes with the original data. How well does first-order exponential smoothing seem to work?

c. Use double exponential smoothing to develop a forecasting model for crude oil production. Plot the smoothed statistic on the same axes with the original data. How well does double exponential smoothing seem to work?

d. Compare the two smoothing models from parts b and c. Which approach seems preferable?

4.47 Apply a first difference to the Danish crude oil production data in Table B.22.

a. Plot the data and comment on any features that you observe from the graph. Calculate and plot the sample ACF and variogram. Interpret these graphs.

b. Use first-order exponential smoothing to develop a forecasting model for crude oil production. Plot the smoothed statistic on the same axes with the original data. How well does first-order exponential smoothing seem to work? How does this compare to the first-order exponential smoothing model you developed in Exercise 4.46 for the original (undifferenced) data?

4.48 Table B.23 shows weekly data on positive laboratory test results for influenza. Notice that these data have a number of missing values. In exercise you were asked to develop and implement a scheme to estimate the missing values. This data have a strong seasonal component. Use the data from 1997–2010 to develop a multiplicative Winters-type exponential smoothing model for this data. Use this model to simulate one-week-ahead forecasts for the remaining years. Calculate the forecast errors. Discuss the reasonableness of the forecasts.

4.49 Repeat Exercise 4.48 using an additive Winters-type model. Compare the performance of the additive and the multiplicative model from Exercise 4.48.

4.50 Data from the Western Regional Climate Center for the monthly mean daily solar radiation (in Langleys) at the Zion Canyon, Utah, station are shown in Table B.24. This data have a strong seasonal component. Use the data from 2003–2012 to develop a multiplicative Winters-type exponential smoothing model for this data. Use this model to simulate one-month-ahead forecasts for the remaining years. Calculate the forecast errors. Discuss the reasonableness of the forecasts.

4.51 Repeat Exercise 4.50 using an additive Winters-type model. Compare the performance of the additive and the multiplicative model from Exercise 4.50.

4.52 Table B.25 contains data from the National Highway Traffic Safety Administration on motor vehicle fatalities from 1966 to 2012. This data are used by a variety of governmental and industry groups, as well as research organizations.

a. Plot the fatalities data and comment on any features of the data that you see.

b. Develop a forecasting procedure using first-order exponential smoothing. Use the data from 1966–2006 to develop the model, and then simulate one-year-ahead forecasts for the remaining years. Compute the forecasts errors. How well does this method seem to work?

c. Develop a forecasting procedure using based on double exponential smoothing. Use the data from 1966–2006 to develop the model, and then simulate one-year-ahead forecasts for the remaining years. Compute the forecasts errors. How well does this method seem to work in comparison to the method based on first-order exponential smoothing?

4.53 Apply a first difference to the motor vehicle fatalities data in Table B.25.

a. Plot the differenced data and comment on any features of the data that you see.

b. Develop a forecasting procedure for the first differences based on first-order exponential smoothing. Use the data from 1966–2006 to develop the model, and then simulate one-year-ahead forecasts for the remaining years. Compute the forecasts errors. How well does this method seem to work?

c. Compare this approach with the two smoothing methods used in Exercise 4.52.

4.54 Appendix Table B.26 contains data on monthly single-family residential new home sales from 1963 through 2014.

a. Plot the home sales data. Comment on the graph.

b. Develop a forecasting procedure using first-order exponential smoothing. Use the data from1963–2000 to develop the model, and then simulate one-year-ahead forecasts for the remaining years. Compute the forecasts errors. How well does this method seem to work?

c. Can you explain the unusual changes in sales observed in the data near the end of the graph?

4.55 Appendix Table B.27 contains data on the airline's best on-time arrival and airport performance. The data are given by month from January 1995 through February 2013.

a. Plot the data and comment on any features of the data that you see.

b. Construct the sample ACF and variogram. Comment on these displays.

c. Develop an appropriate exponential smoothing model for these data.

4.56 Data from the US Census Bureau on monthly domestic automobile manufacturing shipments (in millions of dollars) are shown in Table B.28.

a. Plot the data and comment on any features of the data that you see.

b. Construct the sample ACF and variogram. Comment on these displays.

c. Develop an appropriate exponential smoothing model for these data. Note that there is some apparent seasonality in the data. Why does this seasonal behavior occur?

d. Plot the first difference of the data. Now compute the sample ACF and variogram for the differenced data. What impact has differencing had? Is there still some apparent seasonality in the differenced data?

4.57 Suppose that simple exponential smoothing is being used to forecast a process. At the start of period $t*$, the mean of the process shifts to a new level $\mu + \delta$. The mean remains at this new level for subsequent time periods. Show that the expected value of the exponentially smoothed statistic is

$$E(\hat{y}_t) = \begin{cases} \mu, & T \leq t^* \\ \mu + \delta - \delta(1 - \lambda)^{T - t^* + 1}, & T \geq t^* \end{cases}.$$

4.58 Using the results of Exercise 4.44, determine the number of periods following the step change for the expected value of the exponential smoothing statistic to be within 0.10 δ of the new time series level $\mu + \delta$. Plot the number of periods as a function of the smoothing constant. What conclusions can you draw?

4.59 Suppose that simple exponential smoothing is being used to forecast the process $y_t = \mu + \varepsilon_t$. At the start of period t^*, the mean of the process experiences a transient; that is, it shifts to a new level $\mu + \delta$, but reverts to its original level y at the start of the next period $t^* + 1$. The mean remains at this level for subsequent time periods. Show that the expected value of the exponentially smoothed statistic is

$$E(\hat{y}_t) = \begin{cases} \mu, & T \leq t^* \\ \mu + \delta\lambda(1-\lambda)^{T-t^*}, & T \geq t^* \end{cases}.$$

4.60 Using the results of Exercise 4.46, determine the number of periods that it will take following the impulse for the expected value of the exponential smoothing statistic to return to within 0.10 δ of the original time series level μ. Plot the number of periods as a function of the smoothing constant. What conclusions can you draw?