3. Regression & Exponential Smoothing

3.1 Forecasting a Single Time Series

Two main approaches are traditionally used to model a single time series $z_1, z_2, \ldots, z_n$

1. Models the observation z_t as **a function of time** as

$$z_t = f(t, \boldsymbol{\beta}) + \varepsilon_t$$

   where $f(t, \boldsymbol{\beta})$ is a function of time t and unknown coefficients $\boldsymbol{\beta}$, and ε_t are uncorrelated errors.

* **Examples:**

  - The constant mean model: $z_t = \beta + \varepsilon_t$
  - The linear trend model: $z_t = \beta_0 + \beta_1 t + \varepsilon_t$
  - Trigonometric models for seasonal time series

$$z_t = \beta_0 + \beta_1 \sin \frac{2\pi}{12} t + \beta_2 \cos \frac{2\pi}{12} t + \varepsilon_t$$

2. **A time Series modeling approach**, the observation at time t is modeled as a linear combination of previous observations

* **Examples:**

  - The autoregressive model: $z_t = \sum_{j \geq 1} \pi_j z_{t-j} + \varepsilon_t$
  - The autoregressive and moving average model:
    $z_t = \sum_{j=1}^{p} \pi_j z_{t-j} + \sum_{j=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t$

Discounted least squares/general exponential smoothing

$$\sum_{t=1}^{n} w_t [z_t - f(t, \boldsymbol{\beta})]^2$$

- Ordinary least squares: $w_t = 1$.

- $w_t = w^{n-t}$, discount factor w determines how fast information from previous observations is discounted.

- Single, double and triple exponential smoothing procedures
  - The constant mean model
  - The Linear trend model
  - The quadratic model

3.2 Constant Mean Model

$$z_t = \beta + \varepsilon_t$$

- β: a constant mean level
- ε_t: a sequence of uncorrelated errors with constant variance σ^2.

If β, σ are ***known***, **the minimum mean square error forecast** of a future observation at time $n + l$, $z_{n+l} = \beta + \varepsilon_{n+l}$ is given by

$$z_n(l) = \beta$$

- $\mathrm{E}[z_{n+l} - z_n(l)] = 0$, $\mathrm{E}[z_{n+1} - z_n(l)]^2 = \mathrm{E}[\varepsilon_t^2] = \sigma^2$
- $100(1 - \lambda)$ **percent prediction intervals** for a future realization are given by

  $$[\beta - \mu_{\lambda/2}\sigma; \beta + \mu_{\lambda/2}\sigma]$$

  where $\mu_{\lambda/2}$ is the $100(1 - \lambda/2)$ percentage point of standard normal distribution.

If β, σ are ***unknown***, we use the least square estimate $\hat{\beta}$ to replace β

$$\hat{\beta} = \bar{z} = \frac{1}{n} \sum_{t=1}^{n} z_t$$

The **l-step-ahead forecast of z_{n+l}** from time origin n by

$$\hat{z}_n(l) = \bar{z}, \hat{\sigma}^2 = \frac{1}{n-1} \sum_{t=1}^{n} (z_t - \bar{z})^2$$

- $\mathrm{E}[z_{n+l} - \hat{z}_n(l)] = 0, \mathrm{E}\hat{\sigma}^2 = \sigma^2$, $\mathrm{E}[z_{n+1} - \hat{z}_n(l)]^2 = \sigma^2 \left(1 + \frac{1}{n}\right)$
- $100(1-\lambda)$ **percent prediction intervals** for a future realization are given by
$$\left[\beta - t_{\lambda/2}(n-1)\hat{\sigma}\left(1+\frac{1}{n}\right)^{\frac{1}{2}}; \beta + t_{\lambda/2}(n-1)\hat{\sigma}\left(1+\frac{1}{n}\right)^{\frac{1}{2}}\right]$$
where $t_{\lambda/2}(n-1)$ is the $100(1-\lambda/2)$ percentage point of t distribution with $n-1$ degree of freedom.

- Updating Forecasts

$$\begin{aligned}
\hat{z}_{n+1} &= \frac{1}{n+1}(z_1 + z_2 + \cdots + z_n + z_{n+1}) = \frac{1}{n+1}[z_{n+1} + n\hat{z}_n(1)] \\
&= \frac{n}{n+1}\hat{z}_n(1) + \frac{1}{n+1}z_{n+1} \\
&= \hat{z}_n(1) + \frac{1}{n+1}(z_{n+1} - \hat{z}_n(1))
\end{aligned}$$

- Checking the adequacy of the model
  Calculate the sample autocorrelation r_k of the residuals $z_t - \bar{z}$

$$r_k = \frac{\sum_{t=k+1}^{n} (z_t - \bar{z})(z_{t-k} - \bar{z})}{\sum_{t=1}^{n} (z_t - \bar{z})^2}, \quad k = 1, 2, \ldots$$

If $\sqrt{n}|r_k| > 2$: something might have gone wrong.

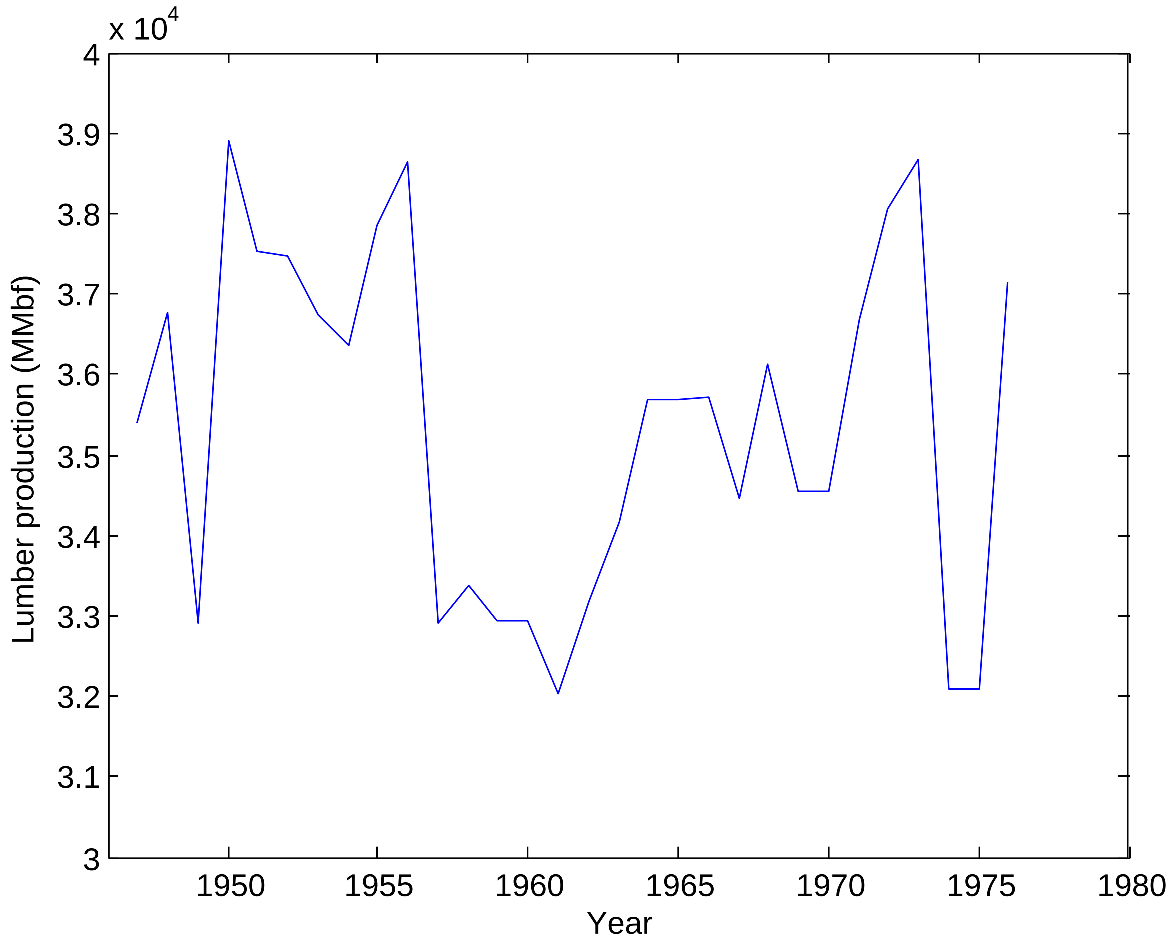
Example 3.1 Annual U.S. Lumber Production

Consider the annual U.S. lumber production from 1947 through 1976. The data were obtained from U.S. Department of Commerce Survey of Current Business. The 30 observations are listed in Table

Table 3.1: Annual Total U.S. Lumber Production (Millions of Broad Feet), 1947-1976 (Table reads from left to right)

35,404	36,762	32,901	38,902	37,515
37,462	36,742	36,356	37,858	38,629
32,901	33,385	32,926	32,926	32,019
33,178	34,171	35,697	35,697	35,710
34,449	36,124	34,548	34,548	36,693
38,044	38,658	32,087	32,087	37,153

Figure 3.1: Annual U.S. lumber production from 1947 to 1976(in millions of board feet)



- The plot of the data in Figure 3.1.

- The sample mean and the sample standard deviation are given by

$$\bar{z} = 35,625, \hat{\sigma} = \{\frac{1}{29}\sum(z_t - \bar{z})^2\}^{\frac{1}{2}} = 2037$$

- The sample auto correlations of the observations are list below

Lag k	1	2	3	4	5	6
Sample autocorrelation	.20	-.05	.13	.14	.04	-.17

Comparing the sample autocorrelations with their standard error $1/\sqrt{30} = .18$, we cannot find enough evidence to reject the assumption of uncorrelated error terms.

- The forecast from the constant mean model are the same for all forecast lead times and are given by

$$\hat{z}_{1976}(l) = \bar{z} = 35,652$$

The standard error of these forecast is given by

$$\hat{\sigma}\sqrt{1 + 1/n} = 2071$$

A 95 percent prediction interval is given by

$$[35,652 \pm (2.045)(2071)] \quad \text{or} \quad [31,417, 39,887]$$

- If new observation become available, the forecasts are easily updated. For example if Lumber production in 1977 was 37,520 million board feet. Then the revised forecasts are given by

$$\begin{aligned}
\hat{z}_{1977}(l) &= \hat{z}_{1976}(1) + \frac{1}{n+1}[z_{1977} - \hat{z}_{1976}(1)] \\
&= 35,652 + \frac{1}{31}[37,520 - 35,652] \\
&= 35,712
\end{aligned}$$

3.3 Locally Constant Mean Model and Simple Exponential Smoothing

- **Reason:** In many instances, the assumption of a time constant mean is restrictive. It is more reasonable to allow for a mean that moves slowly over time

- **Method:** Give more weight to the most recent observation and less to the observations in the distant past

$$\hat{z}_n(l) = c\sum_{t=0}^{n-1} w^t z_{n-t} = c[z_n + wz_{n-1} + \cdots + w^{n-1}z_1]$$

  $w(|w| < 1)$: *discount coefficient*, $c = (1-w)/(1-w^n)$ is needed to normalized sum of weights to 1.

- If $n \to \infty$ and $w < 1$, then $w^n \to 0$, then

$$\hat{z}_n(l) = (1-w)\sum_{j\geq 0} w^j z_{n-j}$$

- **Smoothing constant:** $\alpha = 1 - w$. **Smoothing statistics**

$$\begin{aligned} S_n &= S_n^{[1]} = (1-w)[z_n + w z_{n-1} + w^2 z_{n-2} + \cdots] \\ &= \alpha[z_n + (1-\alpha) z_{n-1} + (1-\alpha)^2 z_{n-2} + \cdots] \end{aligned}$$

- **Updating Forecasts:** (As easy as the constant mean model)

$$S_n = (1-w) z_n + w S_{n-1} = S_{n-1} + (1-w)[z_n - S_{n-1}]$$

$$\hat{z}_n(1) = (1-w) z_n + w \hat{z}_{n-1}(1) = \hat{z}_{n-1}(1) + (1-w)[z_n - \hat{z}_{n-1}(1)]$$

Actual Implementation of Simple Exp. Smoothing

- Initial value for S_0

$$S_n = (1-w)[z_n + wz_{n-1} + \cdots + w^{n-1}z_1] + w^n S_0$$

1. $S_0 = \bar{z}$, (mean change slowly, $\alpha \doteq 0$);
2. $S_0 = z_1$, (local mean changes quickly $\alpha \doteq 1$);
3. Backforecast

$$S_j^* = (1-w)z_j + wS_{j+1}^*, \quad S_{n+1}^* = z_n,$$

$$S_0 = z_0 = S_1^* = (1-w)z_1 + wS_2^*$$

- Choice of the Smoothing Constant: $\alpha = 1 - w$

$$e_{t-1}(1) = z_t - \hat{z}_{t-1}(1) = z_t - S_{t-1}, \quad (one - step - ahead\ forecast\ error).$$

Then minimize

$$\text{SSE}(\alpha) = \sum_{t=1}^{n} e_{t-1}^2(1).$$

- The smoothing constant that is obtained by simulation depends on the value of S_0

- Ideally, since the choice of α depend on S_0, one should choose α and S_0 jointly

- Examples

  - If $\alpha = 0$, one should choose $S_0 = \bar{z}$.
  - If $\alpha = 1$, one should choose $S_0 = z_1$
  - If $0 < \alpha < 1$, one could choose S_0 as the "backforecast" value:

$$S_0 = \alpha[z_1 + (1-\alpha)z_2 + \cdots + (1-\alpha)^{n-2}z_{n-1}] + (1-\alpha)^{n-1}z_n.$$

Example: Quarterly Iowa Nonfarm Income

As an example, we consider the quarterly Iowa nonfarm income for 1948-1979.

- The data exhibit exponential growth.
- Instead of analyzing and forecasting the original series, we first model the quarterly growth rates of nonfarm income.

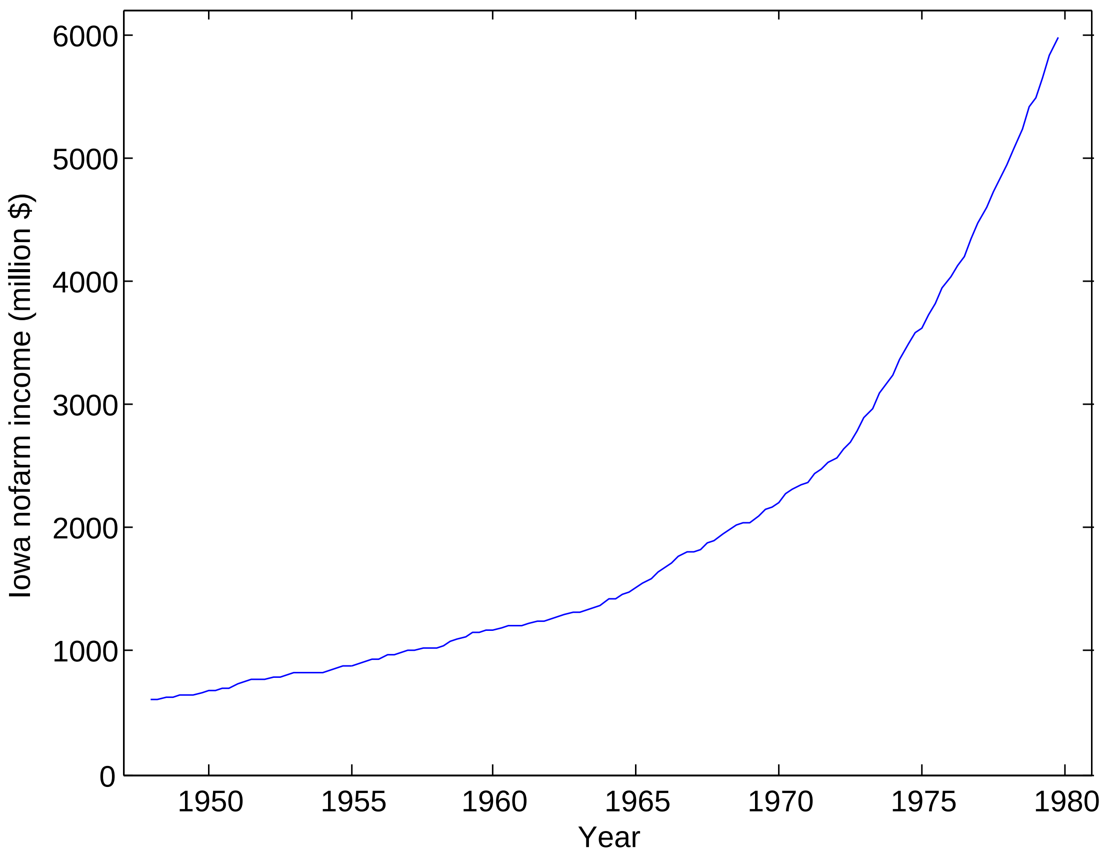
$$z_t = \frac{I_{t+1} - I_t}{I_t} 100 \approx 100 \log \frac{I_{t+1}}{I_t}$$

- The constant mean model would be clearly inappropriate. Compared with the standard error $1/\sqrt{127} = .089$, most autocorrelations are significantly different from zero

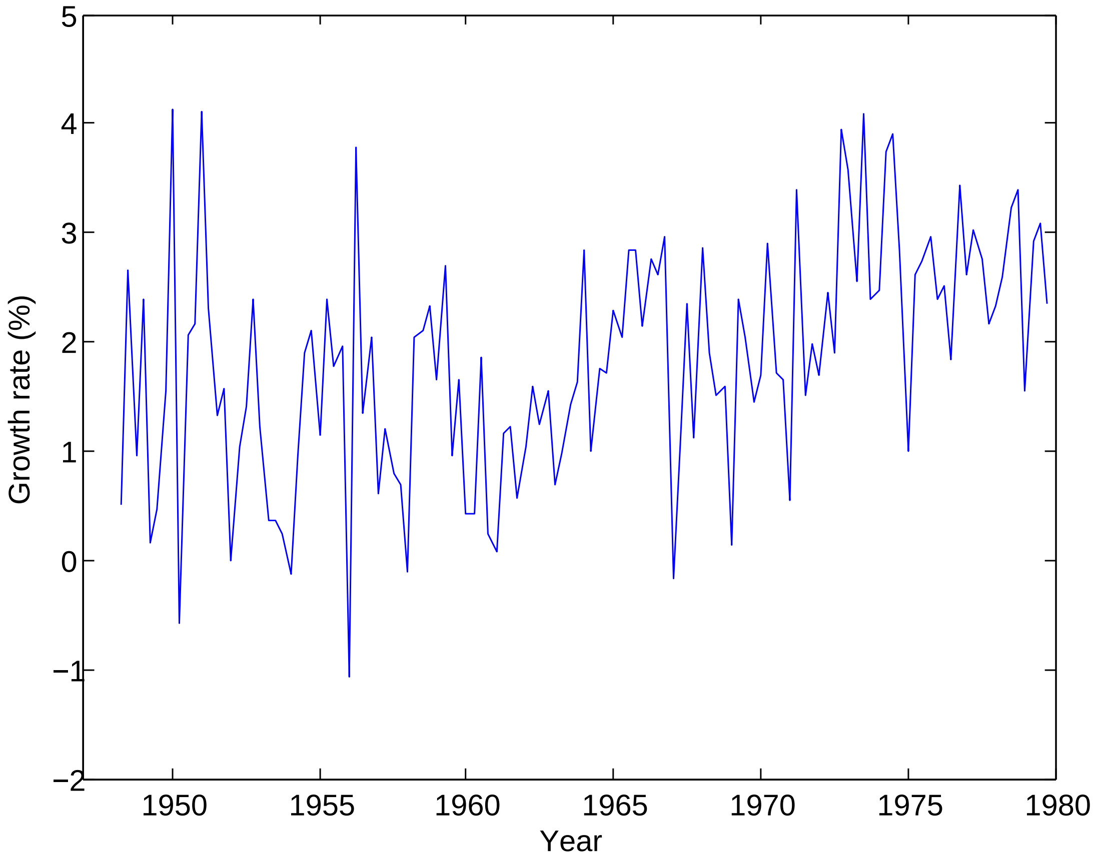
Table 3.2: Sample Autocorrelations r_k of Growth Rates of Iowa Nofarm Income (n=127)

Lag k	1	2	3	4	5	6
Sample autocorrelation r_k	.25	.32	.18	.35	.18	.22

Iowa nonfarm income, first quarter 1948 to fourth quarter 1979



Growth rates of Iowa nonfarm income, second quarter 1948 to fourth quarter 1979



- Since the mean is slowly changing, simple exponential smoothing appears to be an appropriate method.
- $\alpha = 0.11$ and $S_0 = \bar{z} = 1.829$.
- $\text{SSE}(.11) = \sum_{t=1}^{n} e_{t-1}^2(1) = (-1.329)^2 + (.967)^2 + \cdots + (.458)^2 + (-.342)^2 = 118.19$
- As a diagnostic check, we calculate the sample autocorrelations of the one-step-ahead forecast errors

$$r_k = \frac{\sum_{t=k}^{n-1} [e_t(1) - \bar{e}][e_{t-k}(1) - \bar{e}]}{\sum_{t=0}^{n-1} [e_t(1) - \bar{e}]^2}, \quad \bar{e} = \frac{1}{n} \sum_{t=0}^{n-1} e_t(1).$$

- To assess the significance of the mean of the forecast errors, we compare it with standard error $s/n^{1/2}(1/\sqrt{127} = .089)$, where

$$s^2 = \frac{1}{n} \sum_{t=0}^{n-1} [e_t(1) - \bar{e}]^2$$

Table 3.3. Simple Exponential Smoothing—Growth Rates of Iowa Nonfarm Income

Time t	Observation z_t	$\alpha = .11$ Smoothed Statistic S_t	$\alpha = .11$ One-Step-Ahead Forecast Error $e_{t-1}(1) = z_t - S_{t-1}$	$\alpha = .40$ Smoothed Statistic S_t	$\alpha = .40$ One-Step-Ahead Forecast Error $e_{t-1}(1) = z_t - S_{t-1}$
0		1.829		1.829	
1	0.50	1.683	−1.329	1.297	−1.329
2	2.65	1.789	0.967	1.838	1.353
3	0.97	1.699	−0.819	1.491	−0.868
4	2.40	1.776	0.701	1.855	0.909
5	0.16	1.598	−1.616	1.177	−1.695
6	0.47	1.474	−1.128	0.894	−0.707
	⋮	⋮	⋮	⋮	⋮
123	3.38	2.736	0.723	3.032	0.579
124	1.55	2.606	−1.186	2.439	−1.482
125	2.93	2.642	0.324	2.636	0.491
126	3.10	2.692	0.458	2.821	0.464
127	2.35	2.654	−0.342	2.633	−0.471
		SSE(.11) = 118.19		SSE(.40) = 132.56	

Table 3.4. Sums of Squared One-Step-Ahead Forecast Errors for Different Values of α; Simple Exponential Smoothing—Growth Rates of Iowa Nonfarm Income

α	SSE(α)	α	SSE(α)	α	SSE(α)
.01	140.78	.11	118.19	.21	120.95
.02	134.42	.12	118.24	.22	121.39
.03	128.84	.13	118.38	.23	121.86
.04	124.86	.14	118.57	.24	122.35
.05	122.22	.15	118.81	.25	122.86
.06	120.51	.16	119.09	.26	123.38
.07	119.44	.17	119.41	.27	123.93
.08	118.78	.18	119.75	.28	124.49
.09	118.41	.19	120.13	.29	125.07
.10	118.23	.20	120.53	.30	125.67

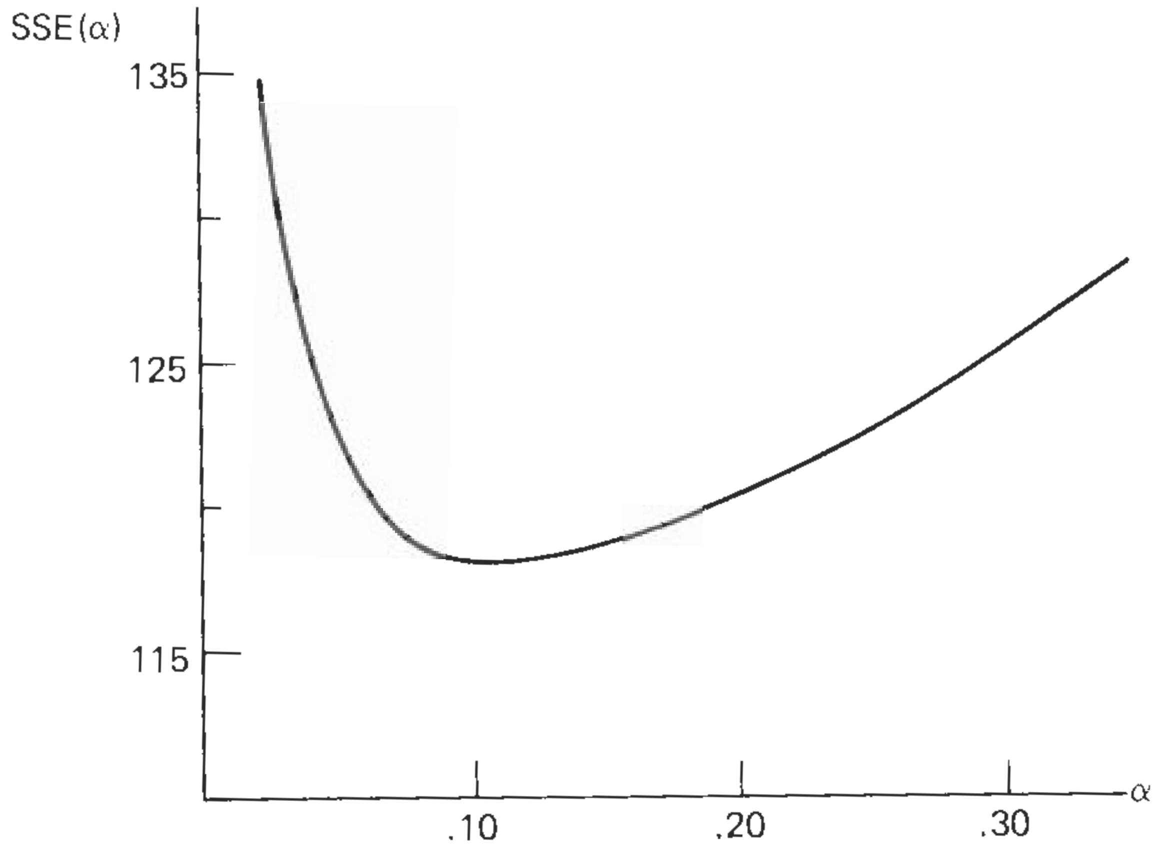


Figure 3.4. Plot of SSE(α), for simple exponential smoothing—growth rates of Iowa nonfar income.

Table 3.5. Means, Standard Errors, and Sample Autocorrelations of the One-Step-Ahead Forecast Errors from Exponential Smoothing (with $\alpha = .11$ and $\alpha = .40$)—Growth Rates of Iowa Nonfarm Income

Lag k	Sample Autocorrelations of One-Step-Ahead Forecast Errors	
	$\alpha = .11$	$\alpha = .40$
1	−.02	−.22
2	.08	−.02
3	−.09	−.19
4	.14	.13
5	−.07	−.11
6	.00	.00
Mean of historical forecast errors	.059	.016
Standard error of mean	.086	.091

3.4 Regression Models with Time as Independent Variable

$$z_{n+j} = \sum_{i=1}^{m} \beta_i f_i(j) + \varepsilon_{n+j} = \mathbf{f}'(j)\boldsymbol{\beta} + \varepsilon_{n+j}$$

- $\mathbf{f}(j+1) = \mathbf{L}\mathbf{f}(j), \mathbf{L} = (l_{ij})_{m \times m}$ full rank. (Difference equations).

- Equivalent model:

$$z_{n+j} = \sum_{i=1}^{m} \beta_i^* f_i(n+j) + \varepsilon_{n+j} = \mathbf{f}'(n+j)\boldsymbol{\beta}^* + \varepsilon_{n+j};$$

$$\mathbf{f}(n+j) = \mathbf{L}^n \mathbf{f}(j) \Rightarrow \boldsymbol{\beta} = \mathbf{L}^n \boldsymbol{\beta}^*.$$

- Examples: Constant Mean Model, Linear Trend Model, Quadratic Trend Model, k^{th} order Polynomial Trend Model, 12-point Sinusoidal Model

- **Estimation:** $\hat{\beta}_n$ minimizes $\sum_{j=1}^{n}[z_j - \mathbf{f}'(j-n)\boldsymbol{\beta}]^2$

$$\mathbf{y}' = (z_1, z_2, \ldots, z_n), \quad \mathbf{X}' = (\mathbf{f}(-n+1), \ldots, \mathbf{f}(0))$$

$$\mathbf{X}'\mathbf{X} = \sum_{j=0}^{n-1} \mathbf{f}(-j)\mathbf{f}'(-j) \hat{=} \mathbf{F}_n, \quad \mathbf{X}'\mathbf{y} = \sum_{j=0}^{n-1} \mathbf{f}(-j) z_{n-j} \hat{=} \mathbf{h}_n$$

$$\hat{\boldsymbol{\beta}}_n = \mathbf{F}_n^{-1}\mathbf{h}_n.$$

- **Prediction**

$$\hat{z}_n(l) = \mathbf{f}'(l)\hat{\boldsymbol{\beta}}_n, \mathrm{Var}(e_n(l)) = \sigma^2[1 + \mathbf{f}'(l)\mathbf{F}_n^{-1}\mathbf{f}(l)],$$

$$\hat{\sigma}^2 = \frac{1}{n-m}\sum_{j=0}^{n-1}(z_{n-j} - \mathbf{f}'(-j)\hat{\boldsymbol{\beta}}_n)^2.$$

$$100(1-\lambda)\% \text{ CI}: \quad \hat{z}_n(l) \pm t_{\lambda/2}(n-m)\hat{\sigma}[1 + \mathbf{f}'(l)\mathbf{F}^{-1}\mathbf{f}(l)]^{\frac{1}{2}}.$$

- **Updating Estimates and Forecasts:**

$$\hat{\boldsymbol{\beta}}_{n+1} = \mathbf{F}_{n+1}^{-1}\mathbf{h}_{n+1}.$$

$$\mathbf{F}_{n+1} = \mathbf{F}_n + \mathbf{f}(-n)\mathbf{f}'(-n);$$

$$\begin{aligned}
\mathbf{h}_{n+1} &= \sum_{j=0}^{n}\mathbf{f}(-j)z_{n+1-j} = \mathbf{f}(0)z_{n+1} + \sum_{j=0}^{n-1}\mathbf{f}(-j-1)z_{n-j} \\
&= \mathbf{f}(0)z_{n+1} + \sum_{j=0}^{n-1}\mathbf{L}^{-1}\mathbf{f}(-j)z_{n-j} = \mathbf{f}(0)z_{n+1} + \mathbf{L}^{-1}\mathbf{h}_n
\end{aligned}$$

$$\hat{z}_{n+1}(l) = \mathbf{f}'(l)\hat{\boldsymbol{\beta}}_{n+1}.$$

3.5 Discounted Least Square and General Exponential Smoothing

In *discounted least squares* or *general exponential smoothing*, the parameter estimates are determined by minimizing

$$\sum_{j=0}^{n-1} w^j [z_{n-j} - \mathbf{f}'(-j)\boldsymbol{\beta}]^2$$

The constant $w(|w| < 1)$ is a discount factor the discount past observation exponentially.
Define

$$\mathbf{W} = \text{diag}(w^{n-1} \quad w^{n-2} \quad \cdots \quad w \quad 1);$$

$$\mathbf{F}_n \hat{=} \mathbf{X}'\mathbf{W}\mathbf{X} = \sum_{j=0}^{n-1} w^j \mathbf{f}(-j)\mathbf{f}'(-j)$$

$$\mathbf{h}_n \hat{=} \mathbf{X}'\mathbf{W}\mathbf{y} = \sum_{j=0}^{n-1} w^j \mathbf{f}(-j) z_{n-j}$$

- **Estimation and Forecasts**

$$\hat{\boldsymbol{\beta}}_n = \mathbf{F}_n^{-1}\mathbf{h}_n, \quad \hat{z}_n(l) = \mathbf{f}'(l)\hat{\boldsymbol{\beta}}_n.$$

- **Updating Parameter Estimates and Forecasts**

$$\hat{\boldsymbol{\beta}}_{n+1} = \mathbf{F}_{n+1}^{-1}\mathbf{h}_{n+1}, \quad \mathbf{F}_{n+1} = \mathbf{F}_n + \mathbf{f}(-n)\mathbf{f}'(-n)w^n;$$

$$\mathbf{h}_{n+1} = \sum_{j=0}^{n} w^j \mathbf{f}(-j) z_{n+1-j} = \mathbf{f}(0)z_{n+1} + w\mathbf{L}^{-1}\mathbf{h}_n$$

If $n \longrightarrow \infty$, then $w^n\mathbf{f}(-n)\mathbf{f}'(-n) \longrightarrow 0$, and

$$\mathbf{F}_{n+1} = \mathbf{F}_n + \mathbf{f}(-n)\mathbf{f}'(-n)w^n \rightarrow \mathbf{F} \quad \text{as } n \rightarrow \infty.$$

Hence

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{n+1} &= \mathbf{F}^{-1}\mathbf{f}(0)z_{n+1} + [\mathbf{L}' - \mathbf{F}^{-1}\mathbf{f}(0)\mathbf{f}'(0)\mathbf{L}']\hat{\boldsymbol{\beta}}_n \\ &= \mathbf{L}'\hat{\boldsymbol{\beta}}_n + \mathbf{F}^{-1}\mathbf{f}(0)[z_{n+1} - \hat{z}_n(1)]. \\ \hat{z}_{n+1}(l) &= \mathbf{f}'(l)\hat{\boldsymbol{\beta}}_{n+1}. \end{aligned}$$

3.5 Locally Constant Linear Trend and Double Exp. Smoothing

- **Locally Constant Linear Trend Model**

$$z_{n+j} = \beta_0 + \beta_1 j + \varepsilon_{n+j}$$

- **Definition**

$$\mathbf{f}(j) = [1 \quad j]', \mathbf{L} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

$$F = \sum w^j \mathbf{f}(-j)\mathbf{f}'(-j) = \begin{bmatrix} \sum w^j & -\sum j w^j \\ -\sum j w^j & \sum j^2 w^j \end{bmatrix} = \begin{bmatrix} \frac{1}{1-w} & \frac{-w}{(1-w)^2} \\ \frac{-w}{(1-w)^2} & \frac{w(1+w)}{(1-w)^2} \end{bmatrix}$$

- **Discount least squares** that minimizing $\sum_{j=1}^{n-1} w^j [z_{n-j} - \mathbf{f}'(-j)\boldsymbol{\beta}]^2$, thus

$$\hat{\boldsymbol{\beta}}_n = \mathbf{F}^{-1}\mathbf{h}_n = \begin{bmatrix} 1-w^2 & (1-w)^2 \\ (1-w)^2 & \frac{(1-w)^3}{w} \end{bmatrix} \begin{bmatrix} \sum w^j z_{n-j} \\ -\sum j w^j z_{n-j} \end{bmatrix}$$

Thus

$$\hat{\beta}_{0,n} = (1-w^2)\sum w^j z_{n-j} - (1-w)^2 \sum j w^j z_{n-j}$$

$$\hat{\beta}_{1,n} = (1-w)^2 \sum w^j z_{n-j} - \frac{(1-w)^3}{w} \sum j w^j z_{n-j}$$

In terms of smoothing

$$S_n^{[1]} = (1-w)z_n + wS_{n-1}^{[1]} = (1-w)\sum w^j z_{n-j},$$

$$S_n^{[2]} = (1-w)S_n^{[1]} + wS_{n-1}^{[2]} = (1-w)^2 \sum (j+1) w^j z_{n-j},$$

$$S_n^{[k]} = (1-w)S_n^{[k-1]} + wS_{n-1}^{[k]}.$$

$$(S_n^{[0]} = \text{(no smoothing)} = z_n)$$

Then

$$\hat{\beta}_{0,n} = 2S_n^{[1]} - S_n^{[2]},$$

$$\hat{\beta}_{1,n} = \frac{1-w}{w}(S_n^{[1]} - S_n^{[2]}).$$

$$\hat{z}_n(l) = \hat{\beta}_{0,n} + \hat{\beta}_{1,n} \cdot l = (2 + \frac{1-w}{w} l) S_n^{[1]} - (1 + \frac{1-w}{w} l) S_n^{[2]}.$$

- **Updating:**

$$\begin{aligned}\hat{\beta}_{0,n+1} &= \hat{\beta}_{0,n} + \hat{\beta}_{1,n} + (1-w^2)[z_{n+1} - \hat{z}_n(1)],\\ \hat{\beta}_{1,n+1} &= \hat{\beta}_{1,n} + (1-w)^2[z_{n+1} - \hat{z}_n(1)];\end{aligned}$$

Or in another combination form:

$$\begin{aligned}\hat{\beta}_{0,n+1} &= (1-w^2)z_{n+1} + w^2(\hat{\beta}_{0,n} + \hat{\beta}_{1,n}),\\ \hat{\beta}_{1,n+1} &= \frac{1-w}{1+w}(\hat{\beta}_{0,n+1} - \hat{\beta}_{0,n}) + \frac{2w}{1+w}\hat{\beta}_{1,n}.\\ z_{n+1} - \hat{z}_n(1) &= \frac{1}{1-w^2}(\hat{\beta}_{0,n+1} - \hat{\beta}_{0,n} - \hat{\beta}_{1,n}).\end{aligned}$$

- **Implementation**
  - Initial Values for $S_0^{[1]}$ and $S_0^{[2]}$

$$
\begin{aligned}
S_0^{[1]} &= \hat{\beta}_{0,0} - \frac{w}{1-w}\hat{\beta}_{1,0}, \\
S_1^{[2]} &= \hat{\beta}_{0,0} - \frac{2w}{1-w}\hat{\beta}_{1,0};
\end{aligned}
$$

    where the $(\hat{\beta}_{0,0}, \hat{\beta}_{1,0})$ are usually obtained by considering a subset of the data fitted by the standard model

$$z_t = \beta_0 + \beta_1 t + \varepsilon_t.$$

  - Choice of the Smoothing Constant $\alpha = 1 - w$ The smoothing constant α is chosen to minimize the SSE:

$$
\begin{aligned}
\mathsf{SSE}(\alpha) &= \sum (z_t - \hat{z}_{t-1})^2 \\
&= \sum \left[z_t - \left(2 + \frac{\alpha}{1-\alpha}\right) S_{t-1}^{[1]} + \left(1 + \frac{\alpha}{1-\alpha}\right) S_{t-1}^{[2]} \right]^2.
\end{aligned}
$$

Example: Weekly Thermostat Sales

As an example for double exponential smoothing, we analyze a sequence of 52 weekly sales observations. The data are listed in Table and plotted in Figure. The data indicates an upward trend in the thermostat sales. This trend, however, does not appear to be constant but seems to change over time. A constant linear trend model would therefore not be appropriate.

Case Study II: University of Iowa Student Enrollments

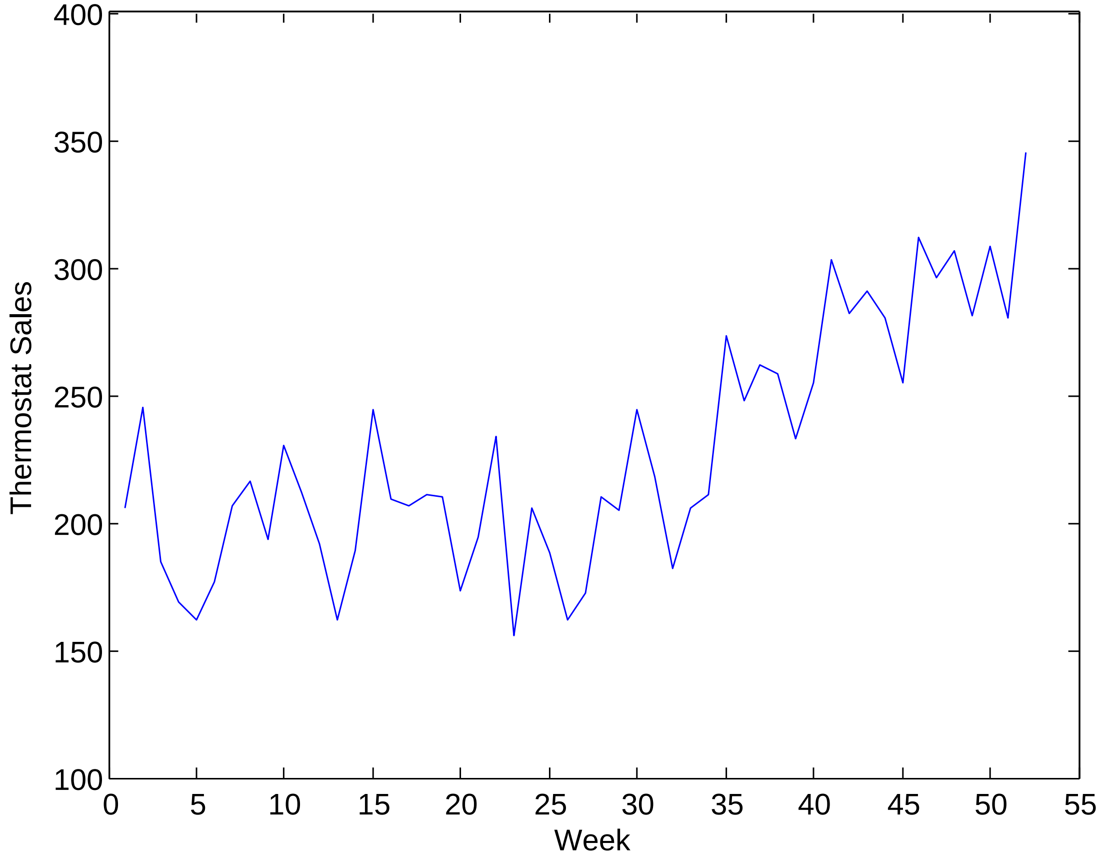
As another example, we consider the annual student enrollment (fall and spring semester combined) at the University of Iowa. Observations for last 29 years (1951/1952 through 1979/1980) are summarized in Table. A plot of the observations is given in Figure.

Table 3.6. Weekly Thermostat Sales, 52 Observations[a]

206	189	172	255
245	244	210	303
185	209	205	282
169	207	244	291
162	211	218	280
177	210	182	255
207	173	206	312
216	194	211	296
193	234	273	307
230	156	248	281
212	206	262	308
192	188	258	280
162	162	233	345

[a] Read downwards, left to right.

Source: Reprinted by permission of Prentice-Hall, Inc. from R. G. Brown (1962), *Smoothing, Forecasting and Prediction of Discrete Time Series*, p. 431.



Weekly thermostat sales

Table 3.7. Sample Autocorrelations of the Residuals from the Constant Linear Trend Model—Thermostat Sales

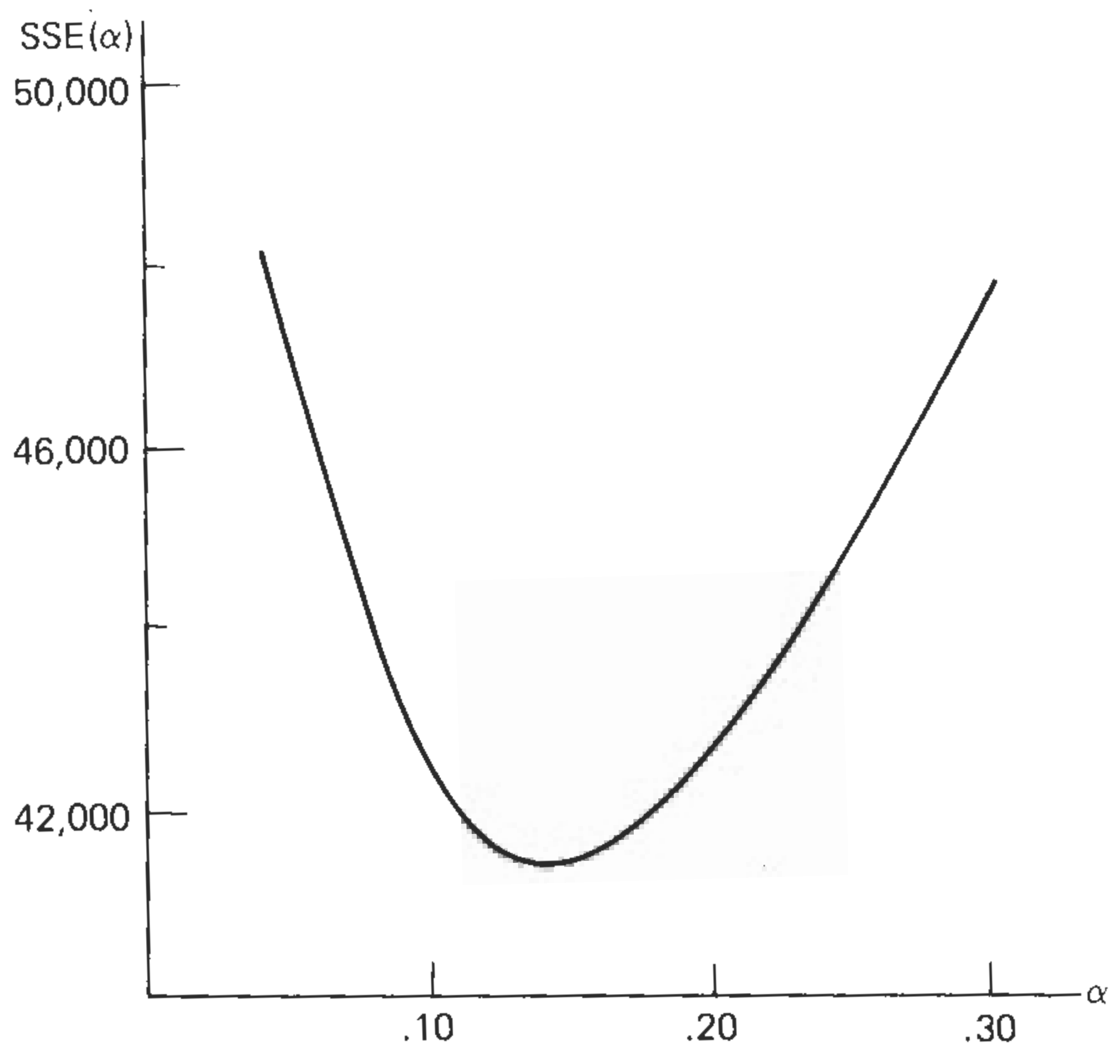
Lag k	1	2	3	4	5	6
Autocorrelation r_k	.41	.26	.18	.19	.27	.40

Table 3.8. Double Exponential Smoothing with Smoothing Constant $\alpha = .14$—Thermostat Sales

t	z_t	$S_t^{[1]}$	$S_t^{[2]}$	$\hat{z}_t(1)$	$\hat{z}_t(2)$	$e_{t-1}(1)$ [$= z_t - \hat{z}_{t-1}(1)$]
0		152.12	137.84	168.72	171.05	
1	206	159.66	140.89	181.48	184.54	37.28
2	245	171.61	145.19	202.32	206.62	63.52
3	185	173.48	149.15	201.77	205.73	−17.32
4	169	172.86	152.47	196.55	199.88	−32.77
5	162	171.34	155.11			−34.56
⋮	⋮	⋮	⋮	⋮	⋮	⋮
48	307	274.58	247.64	305.90	310.29	7.42
49	281	275.48	251.54	303.31	307.21	−24.90
50	308	280.03	255.53	308.52	312.51	4.69
51	280	280.03	258.96	304.53	307.96	−28.52
52	345	289.12	263.18	319.29	323.51	40.48
						SSE(.14) = 41,469

Table 3.9. **Sums of Squared One-Step-Ahead Forecast Errors for Different Values of α; Double Exponential Smoothing—Thermostat Sales**

α	SSE(α)	α	SSE(α)	α	SSE(α)
0.02	49,305	0.11	42,018	0.21	43,132
0.03	48,935	0.12	41,707	0.22	43,558
0.04	48,149	0.13	41,530	0.23	44,014
0.05	47,108	0.14	41,469	0.24	44,496
0.06	45,979	0.15	41,507	0.25	45,001
0.07	44,888	0.16	41,630	0.26	45,526
0.08	43,920	0.17	41,824	0.27	46,070
0.09	43,114	0.18	42,079	0.28	46,629
0.10	42,482	0.19	42,387	0.29	47,203
		0.20	42,740	0.30	47,790



ure 3.6. Plot of SSE(α), for double exponential smoothing—thermostat sales.

Table 3.10. Mean, Standard Error, and Sample Autocorrelations of the One-Step-Ahead Forecast Errors from Double Exponential Smoothing ($\alpha = .14$)—Thermostat Sales

Lag k	Sample Autocorrelations of One-Step-Ahead Forecast Errors r_k
1	.13
2	−.09
3	−.16
4	−.13
5	.05
Mean of historical forecast errors	1.86
Standard error of mean	3.95

3.6 Regression and Exponential Smoothing Methods to Forecast Seasonal Time Series

- **Seasonal Series:** Series that contain seasonal components are quite common, especially in economics, business, and the nature sciences.

- Much of seasonality can be explained on the basis of **physical reasons**. The earth's rotation around the sun, for example, introduces a yearly seasonal pattern into may of the meteorological variables.

- The seasonal pattern in certain variables, such as the one in meteorological variables, is usually quite **stable and deterministic** and repeats itself year after year. The seasonal pattern in business and economic series, however, is **frequently stochastic and changes with time**.

- Apart from a seasonal component, we observe in many series **an additional trend component**.

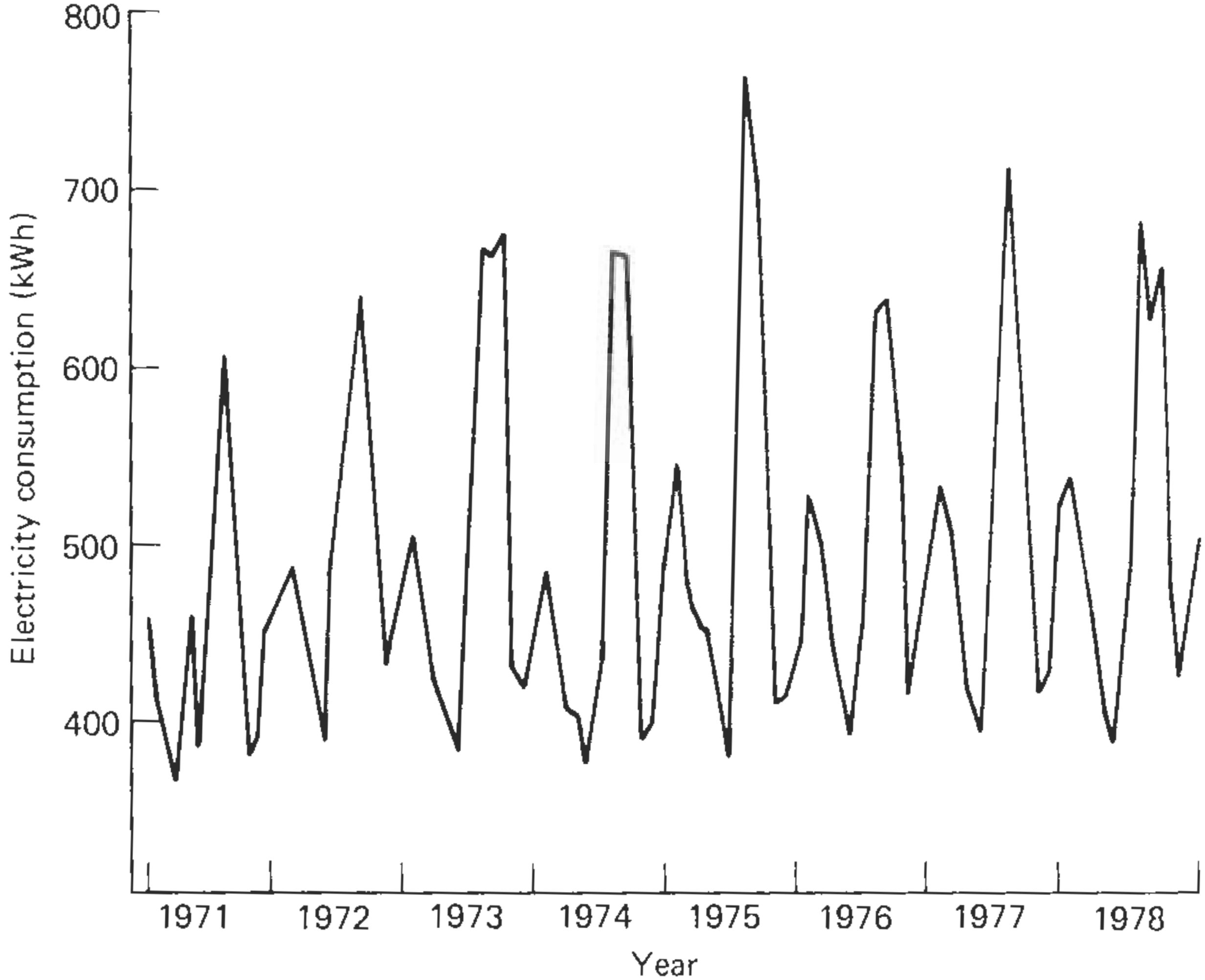


Figure 4.1. Monthly average residential electricity usage in Iowa City (in kilowatt-hour
January 1971 to December 1978.

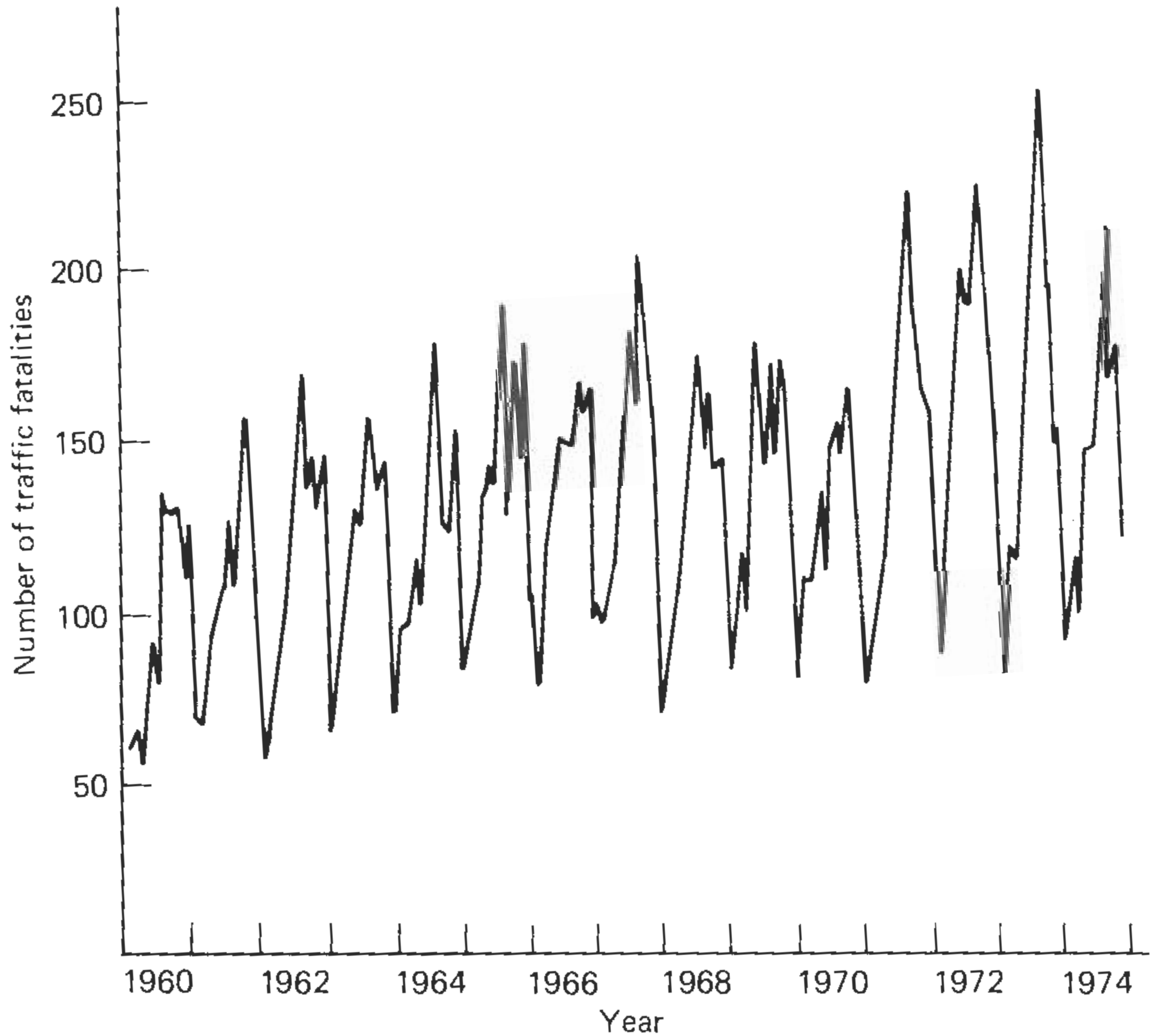


Figure 4.2. Monthly traffic fatalities in Ontario, January 1960 to December 1974.

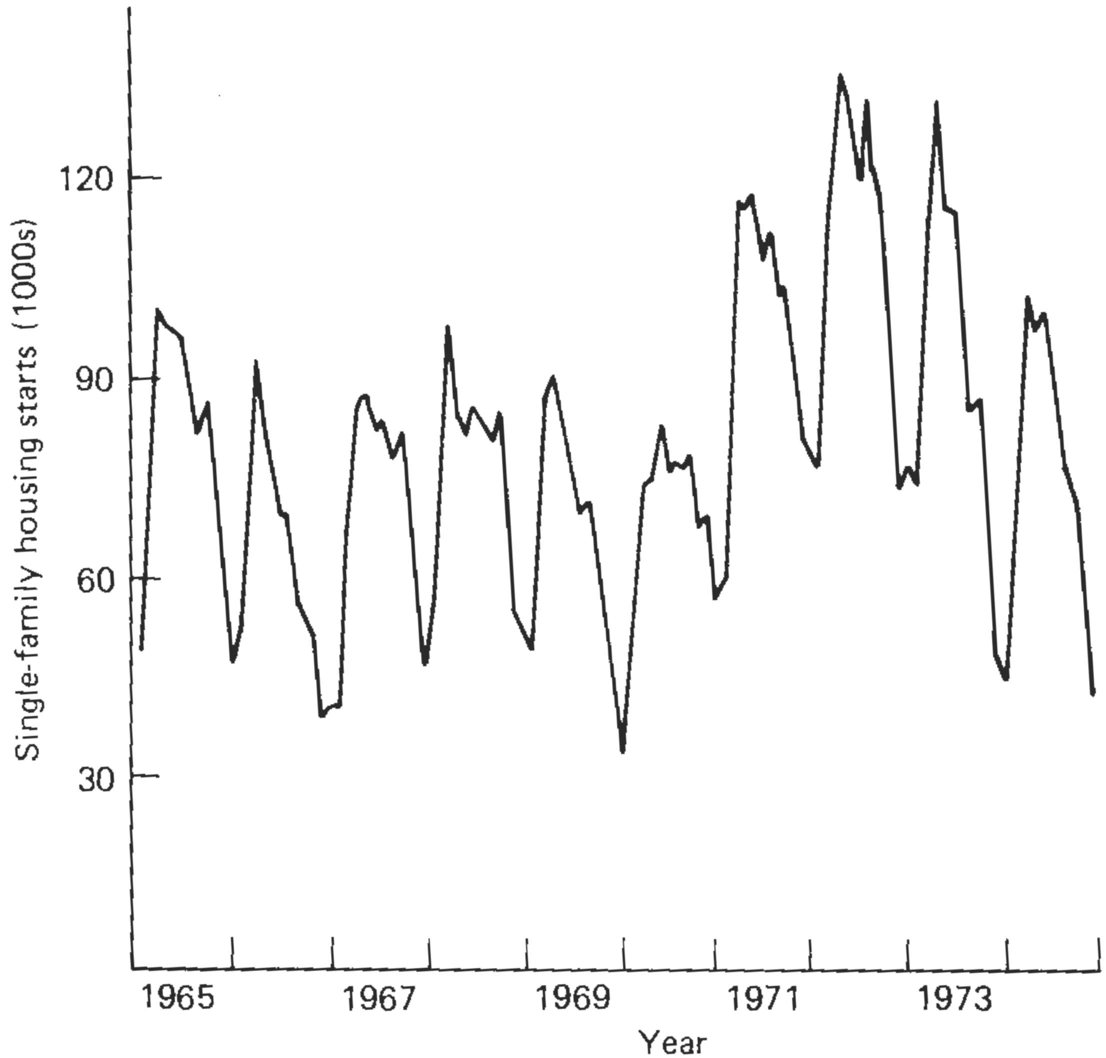


Figure 4.3. Monthly housing starts of privately owned single-family structures, January 1965

The traditional approach to modeling seasonal data is to decompose the series into three components: a trend T_t, a seasonal component S_t and an irregular (or error) component ε_t

- The additive decomposition approach

$$z_t = T_t + S_t + \varepsilon_t$$

- The multiplicative decomposition approach

$$z_t = T_t \times S_t \times \varepsilon_t$$

or

$$\log z_t = T_t^* + S_t^* + \varepsilon_t^*$$

- The other multiplicative model

$$z_t = T_t \times S_t + \varepsilon_t$$

3.6.1 Globally Constant Seasonal Models

Consider **the additive decomposition model** $z_t = T_t + S_t + \varepsilon_t$

- Traditionally the **trend component** T_t is modeled by low-order polynomials of time t:

$$T_t = \beta_0 + \sum_{i=1}^{k} \beta_i \frac{t^i}{i!}$$

- The seasonal component S_t can be described by **seasonal indicators**

$$S_t = \sum_{i=1}^{s} \delta_i \text{IND}_{ti}$$

where $\text{IND}_{ti} = 1$ if t corresponds to the seasonal period i, and 0 otherwise, or by **trigonometric functions**

$$S_t = \sum_{i=1}^{m} A_i \sin(\frac{2\pi i}{s} t + \phi_i).$$

where A_i and ϕ_i are amplitude and the phase shift of the sine function with frequency $f_i = 2\pi i / s$.

Modeling the Additive Seasonality with Seasonal Indicators

$$z_t = \beta_0 + \sum_{i=1}^{k} \beta_i \frac{t^i}{i!} + \sum_{i=1}^{s} \delta_i \text{IND}_{ti} + \varepsilon_t.$$

Since it uses $s+1$ parameters (β_0 and s seasonal indicators) to model s seasonal intercepts, restrictions have to be imposed before the parameters can be estimated. Several equivalent parameterizations are possible

- Omit the intercept: $\beta_0 = 0$.
- Restrict $\sum_{i=1}^{s} \delta_i = 0$.
- Set one of the δ's equal to zero; for example $\delta_s = 0$.

Mathematically these modified models are equivalent, but for convenience we usually choose (3).

Now we have the standard regression model:

$$z_t = \beta_0 + \sum_{i=1}^{k} \beta_i \frac{t^i}{i!} + \sum_{i=1}^{s-1} \delta_i \mathsf{IND}_{ti} + \varepsilon_t.$$

$$\boldsymbol{\beta}' = (\beta_0, \beta_1, \ldots, \beta_k, \delta_1, \ldots, \delta_{s-1});$$

Hence

$$\hat{\beta} = (\mathbf{X}\mathbf{X}')^{-1}\mathbf{X}'\mathbf{y}, \quad \mathbf{y}' = (z_1, z_2, \ldots, z_n);$$

$\mathbf{X}$ is an $n \times (k+s)$matrix with tth row given by

$$\mathbf{f}'(t) = \left(1, t, \frac{t^2}{2}, \cdots, \frac{t^k}{k!}, \mathrm{IND}_{t1}, \cdots, \mathrm{IND}_{t,s-1}\right)$$

The minimum mean square error forecast of z_{n+l} can be calculated from

$$\hat{z}_n(l) = \mathbf{f}'(n+l)\hat{\boldsymbol{\beta}}$$

$100(1-\alpha)\%$ prediction interval

$$\hat{z}_n(l) \pm t_{\lambda/2}(n-k-s)\hat{\sigma}[1 + \mathbf{f}'(n+l)(\mathbf{X}\mathbf{X}')^{-1}\mathbf{f}(n+l)]^{\frac{1}{2}},$$

where

$$\hat{\sigma}^2 = \frac{1}{n-k-s} \sum_{t=1}^{n} (z_t - \mathbf{f}'(t)\hat{\boldsymbol{\beta}})^2.$$

Change of Time Origin in the Seasonal Indicator Model

As we mentioned in previous sections, it might be easier to update the estimates/ prediction if we use the last observational time n as the time origin. In this case

$$z_{n+j} = \beta_0 + \sum_{i=1}^{k} \beta_i \frac{j^i}{i!} + \sum_{i=1}^{s-1} \delta_i \mathsf{IND}_{ji} + \varepsilon_{n+j}.$$

$$\mathbf{f}'(j) = \left(1, j, \frac{j^2}{2}, \cdots, \frac{j^k}{k!}, \mathrm{IND}_{j1}, \cdots, \mathrm{IND}_{j,s-1}\right)$$

$$\hat{\boldsymbol{\beta}}_n = \mathbf{F}_n^{-1} \mathbf{h}_n$$

$$\mathbf{F}_n = \sum_{j=0}^{n-1} \mathbf{f}(-j)\mathbf{f}'(-j), \quad , \mathbf{h}_n = \sum_{j=0}^{n-1} \mathbf{f}(-j) z_{n-j}.$$

Hence the prediction:

$$\hat{z}_n(l) = \mathbf{f}'(l)\hat{\boldsymbol{\beta}}$$

$100(1-\alpha)\%$ prediction interval

$$\hat{z}_n(l) \pm t_{\lambda/2}(n-k-s)\hat{\sigma}[1 + \mathbf{f}'(l)\mathbf{F}_n^{-1}\mathbf{f}(l)]^{\frac{1}{2}},$$

It can be shown that the forecast or fitting functions follow the difference equation $\mathbf{f}(j) = \mathbf{L}\mathbf{f}(j-1)$, where $\mathbf{L}$ is a $(k+s) \times (k+s)$ transition matrix

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_{11} & 0 \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{bmatrix}$$

As an illustration, let us consider a model with quadratic trend and seasonal period $s = 4$

$$z_{n+j} = \beta_0 + \beta_1 + \beta_2 \frac{j^2}{2} + \sum_{i=1}^{3} \delta_i \mathrm{IND}_{ji} + \varepsilon_{n+j}$$

Then the transition matrix $\mathbf{L}$ and the initial vector $\mathbf{f}(0)$ are given by

$$\mathbf{L}_{11} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1/2 & 1 & 1 \end{bmatrix} \quad \mathbf{L}_{21} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \mathbf{L}_{22} = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Successive application of the difference equation $\mathbf{f}(j) = \mathbf{L}\mathbf{f}(j-1)$ leads to $\mathbf{f}(j) = (1, j, j^2/2, \mathbf{IND}_{j1}, \mathbf{IND}_{j2}, \mathbf{IND}_{j3})'$.

Modeling the Seasonality with Trigonometric Functions

$$z_t = T_t + S_t + \varepsilon_t = \beta_0 + \sum_{i=1}^{k} \beta_i \frac{t^i}{i!} + \sum_{i=1}^{m} A_i \sin\left(\frac{2\pi i}{s} t + \phi_i\right) + \varepsilon_t,$$

where the number of harmonics m should not goes beyond $s/2$, i.e. half the seasonality. Monthly, quarterly data; $s/2$ harmonics are usually not necessary

This is illustrated in Figure, where we plot

$$\mathrm{E}(z_t) = \sum_{i=1}^{2} A_i \sin\left(\frac{2\pi i}{12} t + \phi_i\right)$$

for $A_1 = 1, \phi_1 = 0, A_2 = -0.70, \phi_2 = .6944\pi$.

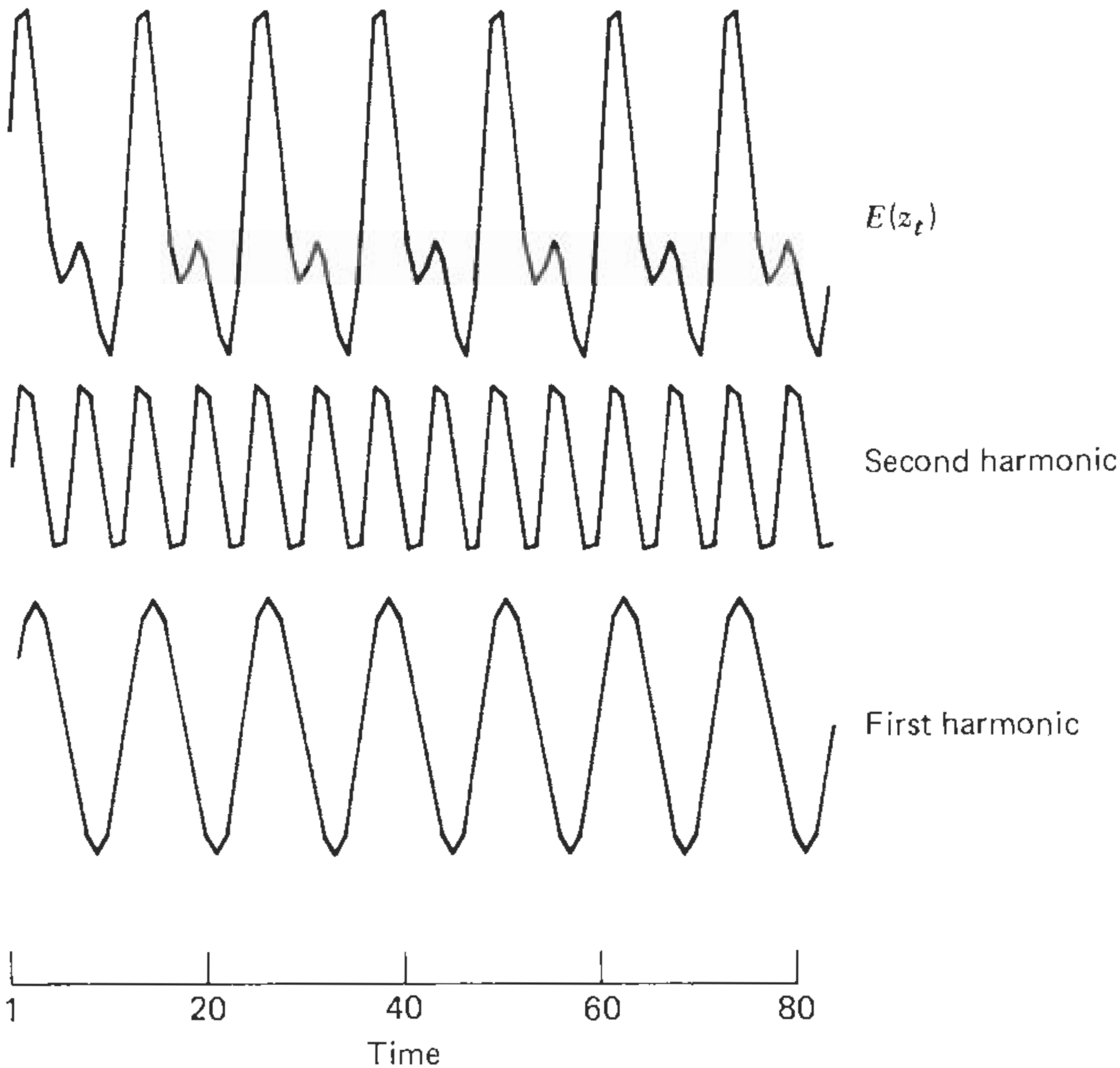


Figure 4.8. Plot of $E(z_t) = \sum_{i=1}^{2} A_i \sin\left(\frac{2\pi i t}{12} + \phi_i\right)$, for $A_1 = 1$, $\phi_1 = 0$, $A_2 = -0.70$, $\phi_2 = (0.6944)\pi$ (or 125°).

Change of Time Origin in the Seasonal Trigonometric Model

Examples:

- 12-point sinusoidal model $(k = 0, s = 12, m = 1)$

$$z_{n+j} = \beta_0 + \beta_{11} \sin \frac{2\pi j}{12} + \beta_{21} \cos \frac{2\pi j}{12} + \varepsilon_{n+j}$$

In this case:

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{3}/2 & 1/2 \\ 0 & -1/2 & \sqrt{3}/2 \end{bmatrix}, \quad \mathbf{f}(0) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

- Linear trend model with two superimposed harmonics $(k = 1, s = 12, m = 2)$:

$$z_{n+j} = \beta_0 + \beta_1 j + \beta_{11} \sin \frac{2\pi j}{12} + \beta_{21} \cos \frac{2\pi j}{12} + \beta_{12} \sin \frac{4\pi j}{12} + \beta_{22} \cos \frac{4\pi j}{12} + \varepsilon_{n+j}$$

3.6.2 Locally Constant Seasonal Models

$$z_{n+j} = \mathbf{f}'(j)\boldsymbol{\beta} + \varepsilon_{n+j}$$

- Target: Minimizing

$$S(\boldsymbol{\beta}, n) = \sum_{j=0}^{n-1} w^j [z_{n-j} - \mathbf{f}'(-j)\boldsymbol{\beta}]^2$$

- updating:

$$\hat{\boldsymbol{\beta}}_{n+1} = \mathbf{L}'\hat{\boldsymbol{\beta}}_n + \mathbf{F}^{-1}\mathbf{f}(\mathbf{0})[z_{n+1} - \hat{z}_n(1)], \quad (\mathbf{F} = \sum_{j \geq 0} w^j \mathbf{f}'(-j)\mathbf{f}(-j))$$

- A collection of infinite sums needed to cacluate $\mathbf{F}$ for seasonal models is given in the following Table.

Table 4.5. Infinite Sums Needed in General Exponential Smoothing

$$\Sigma\omega^j = \frac{1}{1-\omega} \qquad \Sigma\omega^j j = \frac{\omega}{(1-\omega)^2} \qquad \Sigma\omega^j j^2 = \frac{\omega(1+\omega)}{(1-\omega)^3}$$

$$\Sigma\omega^j j^3 = \frac{\omega(1+4\omega+\omega^2)}{(1-\omega)^4} \qquad \Sigma\omega^j j^4 = \frac{\omega(1+11\omega+11\omega^2+\omega^3)}{(1-\omega)^5}$$

$$\Sigma\omega^j \sin fj = \frac{\omega \sin f}{g_1} \qquad \Sigma\omega^j \cos fj = \frac{1-\omega\cos f}{g_1}$$

$$\Sigma\omega^j j \sin fj = \frac{\omega(1-\omega^2)\sin f}{g_1^2} \qquad \Sigma\omega^j j \cos fj = \frac{\omega(1+\omega^2)\cos f - 2\omega^2}{g_1^2}$$

$$\Sigma\omega^j \sin f_1 j \sin f_2 j = \frac{1}{2}\left[\frac{1-\omega\cos(f_1-f_2)}{g_2} - \frac{1-\omega\cos(f_1+f_2)}{g_3}\right]$$

$$\Sigma\omega^j \sin f_1 j \cos f_2 j = \frac{1}{2}\left[\frac{\omega\sin(f_1-f_2)}{g_2} + \frac{\omega\sin(f_1+f_2)}{g_3}\right]$$

$$\Sigma\omega^j \cos f_1 j \cos f_2 j = \frac{1}{2}\left[\frac{1-\omega\cos(f_1-f_2)}{g_2} + \frac{1-\omega\cos(f_1+f_2)}{g_3}\right]$$

where

$$g_1 = 1 - 2\omega\cos f + \omega^2$$

$$g_2 = 1 - 2\omega\cos(f_1 - f_2) + \omega^2$$

$$g_3 = 1 - 2\omega\cos(f_1 + f_2) + \omega^2$$

- It is usually suggested that the least squares estimate of $\boldsymbol{\beta}$ in the regression model $z_t = \mathbf{f}'(t)\boldsymbol{\beta} + \varepsilon_t$ be taken as initial vector $\hat{\boldsymbol{\beta}}_0$.

- To update the estimates, a smoothing constant must be determined.
  - As Brown (1962) suggest that the value of w should lie between $(.70)^{1/g}$ and $(.95)^{1/g}$
  - If sufficient historical data are available, one can estimate $w = 1 - \alpha$ by simulation and choose the smoothing constant that minimizes the sum of the squared one-step-ahead forecast errors

$$\text{SSE}(\alpha) = \sum_{t=1}^{n} [z_t - \hat{z}_{t-1}(1)]^2$$

- After estimating the smoothing constant α, one should always check the adequacy of the model. The sample autocorrelation function of the one-step-ahead forecast errors should be calculated . Significant autocorrelations indicate that the particular forecast model is not appropriate.

Locally Constant Seasonal Models Using Seasonal Indicators

$$z_{n+j} = \beta_0 + \beta_1 j + \sum_{i=1}^{3} \delta_i \mathrm{IND}_{ji} + \varepsilon_{n+j} = \mathbf{f}'(j)\beta$$

where $\mathbf{f}(j) = [1 \ j \ \mathrm{IND}_{j1} \ \mathrm{IND}_{j2} \ \mathrm{IND}_{j3}]', \mathbf{f}(0) = [1 \ 0 \ 0 \ 0 \ 0]'$. Then

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Hence the updating weights in

$$\hat{\boldsymbol{\beta}}_{n+1} = \mathbf{L}'\hat{\boldsymbol{\beta}}_n + \mathbf{F}^{-1}\mathbf{f}(0)[z_{n+1} - \hat{z}_n(1)]$$

can be calculated from $\mathbf{f}(0)$ and the symmetric matrix

$$\mathbf{F} = \begin{bmatrix} \frac{1}{1-w} & \frac{-w}{(1-w)^2} & \frac{w^3}{1-w^4} & \frac{w^2}{1-w^4} & \frac{w}{1-w^4} \\ & \frac{w(1+w)}{(1-w)^3} & \frac{-w^3(3+w^4)}{(1-w^4)^2} & \frac{-w^2(2+2w^4)}{(1-w^4)^2} & \frac{-w(1+3w^4)}{(1-w^4)^2} \\ & & \frac{w^3}{1-w^4} & 0 & 0 \\ & \text{symmetric} & & \frac{w^2}{1-w^4} & 0 \\ & & & & \frac{w}{1-w^4} \end{bmatrix}$$

Implications of $\alpha \to 1$

In this situation $\mathbf{F}$ $\to$singular as $w \to 0$. But we have that

$$\begin{aligned} \lim_{w\to 0} \mathbf{F}^{-1}\mathbf{f} &= \mathbf{f}^* = \left(1, \frac{1}{s}, -\frac{1}{s}, -\frac{2}{s}, \cdots, -\frac{s-1}{s}\right)'; \\ \hat{\boldsymbol{\beta}}_{n+1} &= \mathbf{L}'\hat{\boldsymbol{\beta}}_n + \mathbf{f}^*[z_{n+1} - \hat{z}_n(1)]; \\ \hat{z}_n(1) &= z_n + z_{n+1-s} - z_{n-s} = z_{n+1-s} + (z_n - z_{n-s}) \\ \hat{z}_n(l) &= \hat{z}_n(l-1) + \hat{z}_n(l-s) - \hat{z}_n(l-s-1) \end{aligned}$$

Example: Car Sales

Consider the monthly car sales in Quebec from January 1960 through December 1967 ($n = 96$ observations). The remaining 12 observations (1968) are used as a holdcut period to evaluate the forecast performance. An initial inspection of the series in Figure shows the data may be described by an additive model with a linear trend ($k = 1$) and a yearly seasonal pattern; the trend and the seasonal components appear fairly constant.

$$z_t = \beta_0 + \beta_1 t + \sum_{i=1}^{11} \delta_i \mathrm{IND}_{ti} + \varepsilon_t.$$

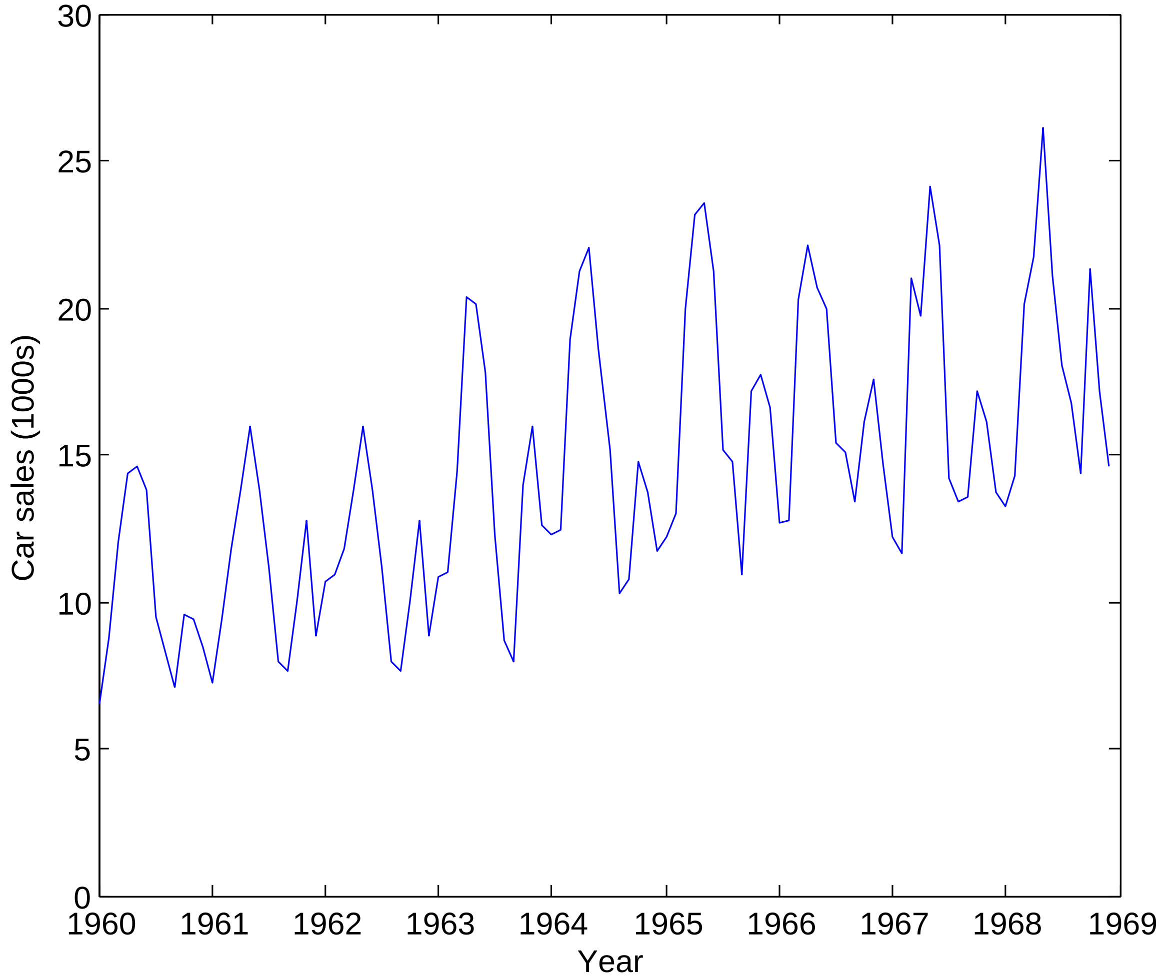


Figure: Monthly car sales in Quebec, Canada; January 1960 to December 1968

Table 4.1. Least Squares Estimation and Forecasting Results for Quebec Monthly Car Sales (1960–1967; $n = 96$)

$$z_t = \beta_0 + \beta_1 t + \sum_{i=1}^{11} \delta_i \, \mathrm{IND}_{ti} + \varepsilon_t$$

Coefficient	Estimate	Standard Error	t Ratio
β_0	7.40	.59	12.62
β_1	.088	.0054	16.45
δ_1	−.60	.72	−.84
δ_2	−.05	.72	−.07
δ_3	5.34	.72	7.38
δ_4	7.52	.72	10.40
δ_5	8.69	.72	12.03
δ_6	6.31	.72	8.73
δ_7	1.41	.72	1.95
δ_8	−.87	.72	−1.21
δ_9	−2.28	.72	−3.16
δ_{10}	1.79	.72	2.48
δ_{11}	2.36	.72	3.27

ANOVA Table

Source	SS	df	MS
Regression	1671.8	12	139.32
Error	172.9	83	2.08
Total (corrected for mean)	1844.7	95	

Autocorrelations of the Residuals

Lag k	r_k	Lag k	r_k
1	.29	7	−.02
2	.20	8	−.02
3	.12	9	.04
4	.11	10	.02
5	.08	11	.17
6	.14	12	.09

Forecasts

Lead Time l

l	1	2	3	4	5	6
z_{96+l}	13.21	14.25	20.14	21.73	26.10	21.08
$\hat{z}_{96}(l)$	15.34	15.99	21.46	23.73	24.99	22.70
$z_{96+l} - \hat{z}_{96}(l)$	−2.13	−1.74	−1.32	−2.00	1.11	−1.62

l	7	8	9	10	11	12
z_{96+l}	18.02	16.72	14.39	21.34	17.18	14.58
$\hat{z}_{96}(l)$	17.89	15.69	14.37	18.53	19.19	16.92
$z_{96+l} - \hat{z}_{96}(l)$	.13	1.03	.02	2.81	−2.01	−2.34

Table 4.3. Analysis of Variance Table—Car Sales (1960–1967; $n = 96$)

$$z_t = \beta_0 + \beta_1 t + \sum_{i=1}^{5} \left(\beta_{1i} \sin \frac{2\pi i}{12} t + \beta_{2i} \cos \frac{2\pi i}{12} t \right) + \beta_{26} \cos \pi t + \varepsilon_t$$

Source	SS		df
Regression	1671.8		12
trend		517.8	1
1st harmonic		659.0	2
2nd harmonic		451.4	2
3rd harmonic		17.1	2
4th harmonic		0.8	2
5th harmonic		25.7	2
6th harmonic		0.0	1
Error	172.9		83
Total (corrected for mean)	1844.7		95

Table 4.4. Least Squares Estimation and Forecasting Results—Car Sales (1960–1967; $n = 96$)

$$z_t = \beta_0 + \beta_1 t + \sum_{i=1}^{2}\left(\beta_{1i}\sin\frac{2\pi i}{12}t + \beta_{2i}\cos\frac{2\pi i}{12}t\right) + \varepsilon_t$$

Coefficient	Estimate	Standard Error	t Ratio
β_0	9.88	.32	30.80
β_1	0.088	.0057	15.30
β_{11}	2.57	.22	11.45
β_{21}	−2.66	.22	−11.90
β_{12}	−2.96	.22	−13.19
β_{22}	0.83	.22	3.71

Autocorrelations of the Residuals

Lag k	r_k	Lag k	r_k
1	.13	7	.09
2	.13	8	.02
3	.11	9	.02
4	.10	10	−.01
5	.16	11	.07
6	−.08	12	.25

Forecasts

Lead Time l

l	1	2	3	4	5	6
z_{96+l}	13.21	14.25	20.14	21.73	26.10	21.08
$\hat{z}_{96}(l)$	15.24	16.42	20.32	24.38	25.33	22.34
$z_{96+l} - \hat{z}_{96}(l)$	−2.03	−2.17	−.18	−2.65	.77	−1.26

l	7	8	9	10	11	12
z_{96+l}	18.02	16.72	14.39	21.34	17.18	14.58
$\hat{z}_{96}(l)$	17.81	15.15	15.70	17.78	18.66	17.54
$z_{96+l} - \hat{z}_{96}(l)$	.21	1.57	−1.31	3.56	−1.48	−2.96

Table 4.10. Sums of the Squared One-Step-Ahead Forecast Errors from the Model

$$z_{n+j} = \beta_0 + \beta_1 j + \sum_{i=1}^{2} (\beta_{1i} \sin f_i j + \beta_{2i} \cos f_i j) + \varepsilon_{n+j}$$

for Various Values of the Smoothing Constant—Car Sales (n = 96)

α	SSE(α)
.03	233.4
.04	233.8
.05	234.4
.06	235.9
.07	238.7
.08	242.7
.09	247.9
.10	254.3
.15	301.4
.20	374.7

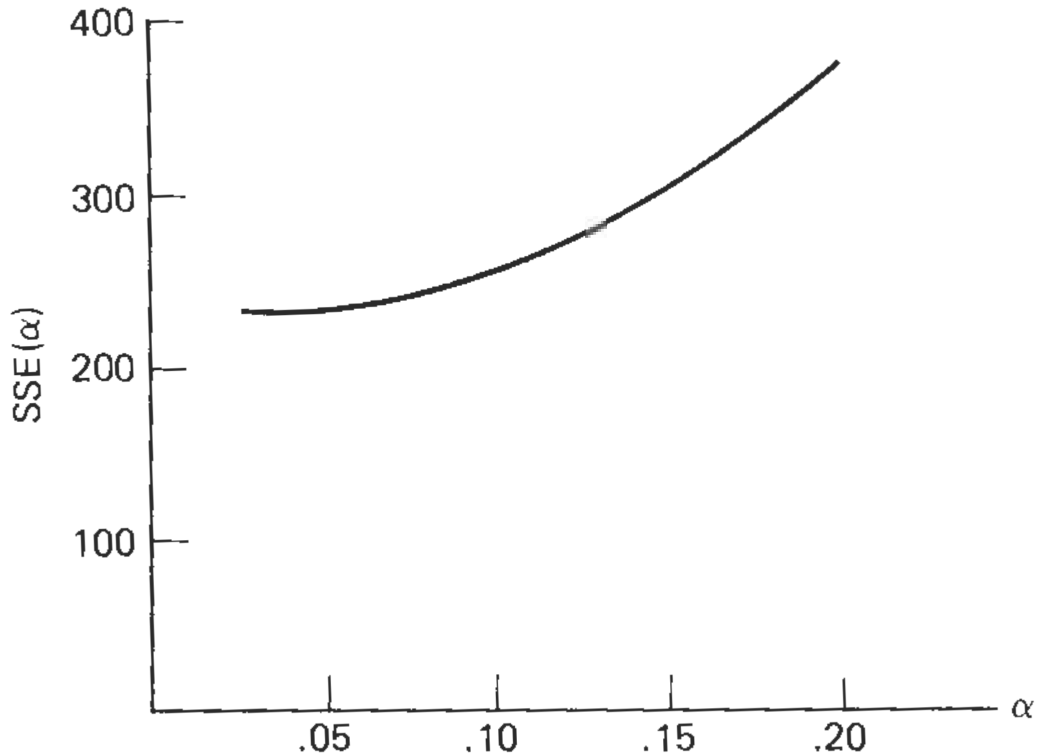


Figure 4.10. Plot of the sum of the squared one-step-ahead forecast errors from the model

$$z_{n+j} = \beta_0 + \beta_1 j + \sum_{i=1}^{2} (\beta_{1i} \sin f_i j + \beta_{2i} \cos f_i j) + \varepsilon_{n+j}$$

Car sales.

Table 4.11. General Exponential Smoothing ($\alpha = .03$) for the Car Sales Model

$$z_{n+j} = \beta_0 + \beta_1 j + \sum_{i=1}^{2} (\beta_{1i} \sin f_i j + \beta_{2i} \cos f_i j) + \varepsilon_{n+j}$$

Time	z_t	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_{11}$	$\hat{\beta}_{21}$	$\hat{\beta}_{12}$	$\hat{\beta}_{22}$	$\hat{z}_t(1)$	$e_{t-1}(1)$ $[= z_t - \hat{z}_{t-1}(1)]$
0		9.877	.088	2.575	−2.665	−2.956	.832	6.800	
1	6.550	9.950	.088	3.561	−1.034	−2.200	−2.158	7.939	−.250
2	8.728	10.082	.088	3.606	.929	.775	−2.940	11.978	.789
3	12.026	10.174	.088	2.658	2.610	2.934	−.797	15.994	.048
⋮									
95	16.119	17.939	.081	.764	−3.554	−.902	3.194	16.140	−1.753
96	13.713	17.884	.079	2.425	−2.831	−3.235	.681	14.263	−2.427
97	13.210	17.904	.078	3.510	−1.297	−2.216	−2.519	15.435	−1.053
⋮									
107	17.180	18.767	.080	.352	−3.551	−1.099	3.143	16.568	−1.417
108	14.577	18.735	.078	2.070	−3.010	−3.287	.510	14.650	−1.991

Example: New Plant and Equipment Expenditures

Consider quarterly new plant and equipment expenditures for the first quarter of 1964 through the fourth quarter of 1974 ($n = 44$). The time series plot in Figure indicate s that the size of the seasonal swings increases with the level of the series; hence a logarithmic transformation must considered. The next Figure shows that this transformation has stabilized the variance.

$$z_t = \ln y_t = \beta_0 + \beta_1 + \sum_{i=1}^{3} \delta_i \text{IND}_{ti} + \varepsilon_t$$

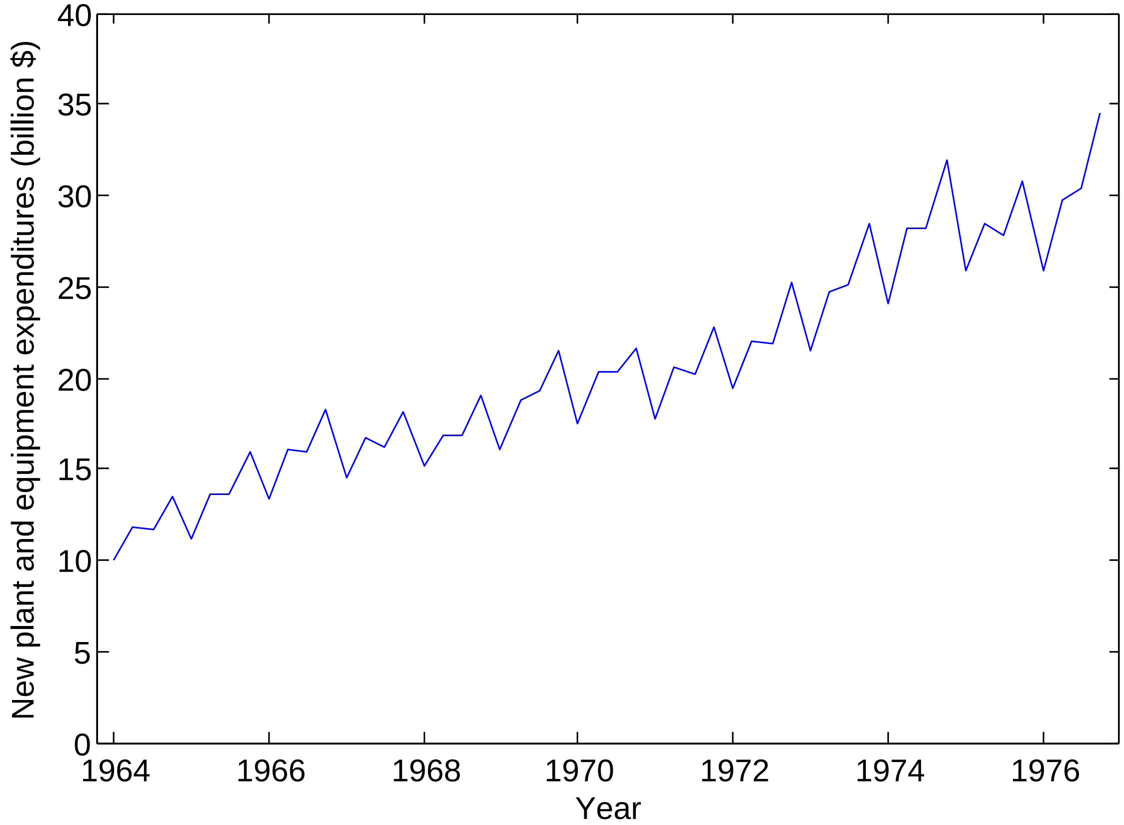
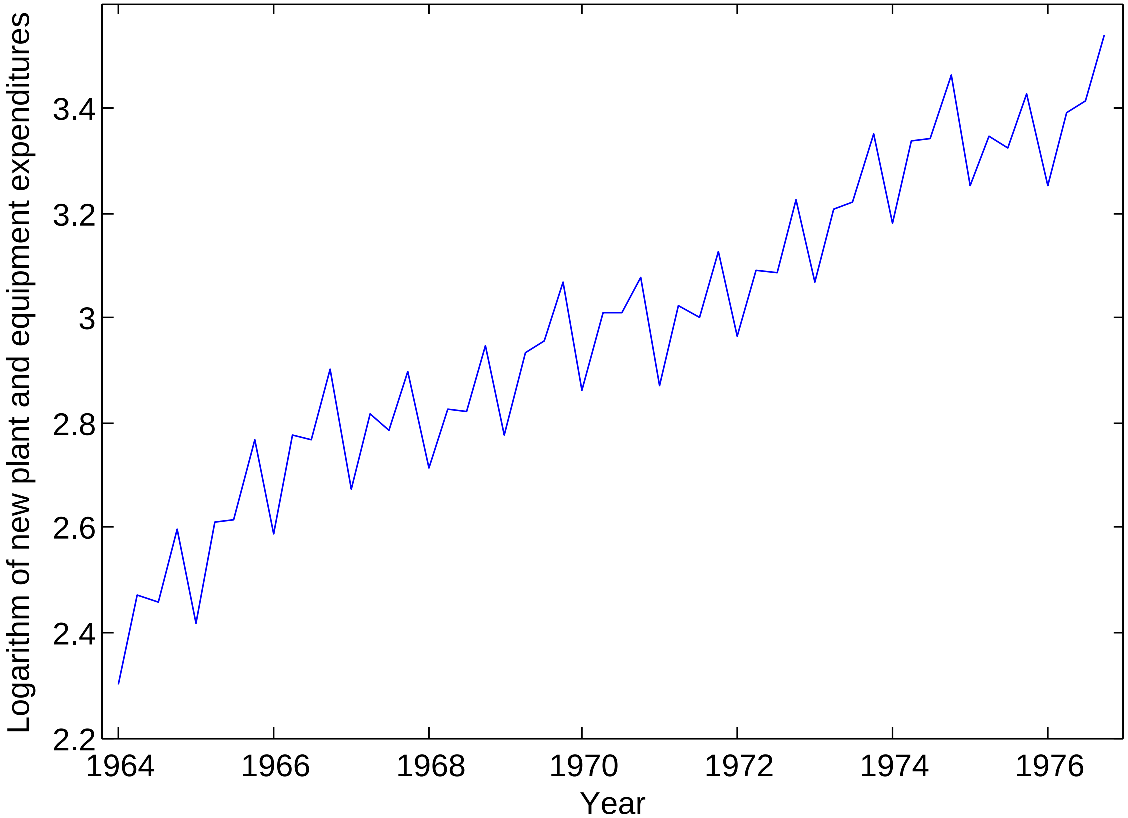


Figure: Quarterly new plant and equipment expenditure in U.S. industries (in billions of dollars), first quarter 1 to fourth quarter 1976



Logarithm of quarterly new plant and equipment expenditures,
frist quarter 1964 to fourth quarter 1976

Table 4.2. Least Squares Estimation Results—New Plant and Equipment Expenditures (1964–1974; $n = 44$)

$$\ln y_t = \beta_0 + \beta_1 t + \sum_{i=1}^{3} \delta_i \mathrm{IND}_{ti} + \varepsilon_t$$

Coefficient	Estimate	Standard Error	t Ratio
β_0	2.580	.020	128.82
β_1	.019	.00057	33.20
δ_1	−.216	.021	−10.45
δ_2	−.081	.021	−3.93
δ_3	−.104	.021	−5.07

ANOVA Table

Source	SS	df	MS
Regression	2.973	4	.74
Error	.091	39	.0023
Total (corrected for mean)	3.063	43	

Autocorrelations of the Residuals

Lag k	r_k	Lag k	r_k
1	.84	5	−.02
2	.68	6	−.18
3	.47	7	−.33
4	.24	8	−.42

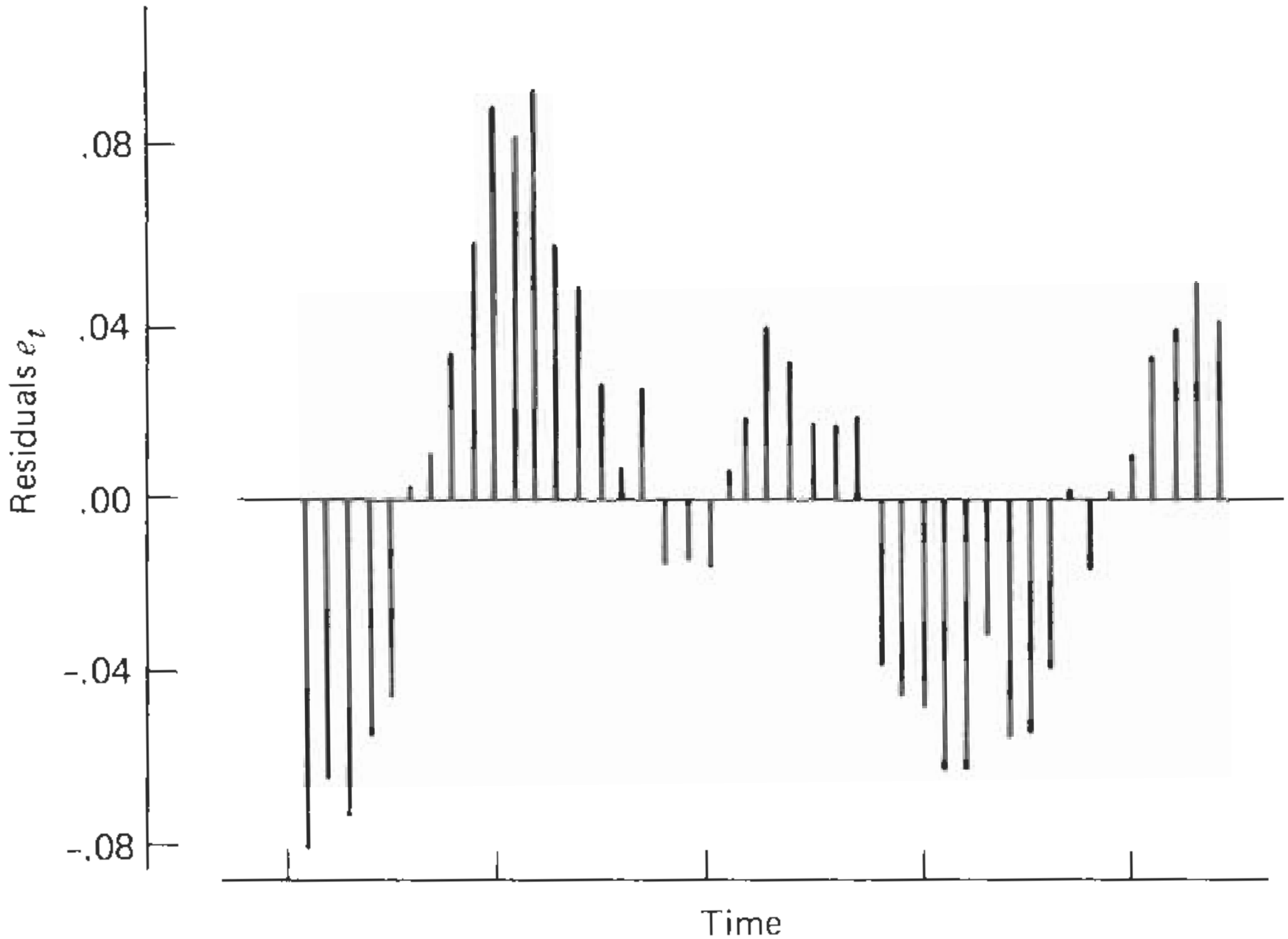


Figure 4.7. Time series plot of the residuals from the model

$$\ln y_t = \beta_0 + \beta_1 t + \sum_{i=1}^{3} \delta_i \text{IND}_{ti} + \varepsilon_t$$

New plant and equipment expenditures.

Table 4.6. Sums of the Squared One-Step-Ahead Forecast Errors from the Model

$$z_{n+j} = \ln y_{n+j} = \beta_0 + \beta_1 j + \sum_{i=1}^{3} \delta_i \, \mathrm{IND}_{ji} + \varepsilon_{n+j}$$

for Various Values of the Smoothing Constant α.— New Plant and Equipment Expenditures (n = 44)

α	SSE(α)
.10	.1114
.30	.0904
.50	.0691
.70	.0540
.90	.0455
1.10	.0421
1.20	.0418
1.30	.0422
1.50	.0445

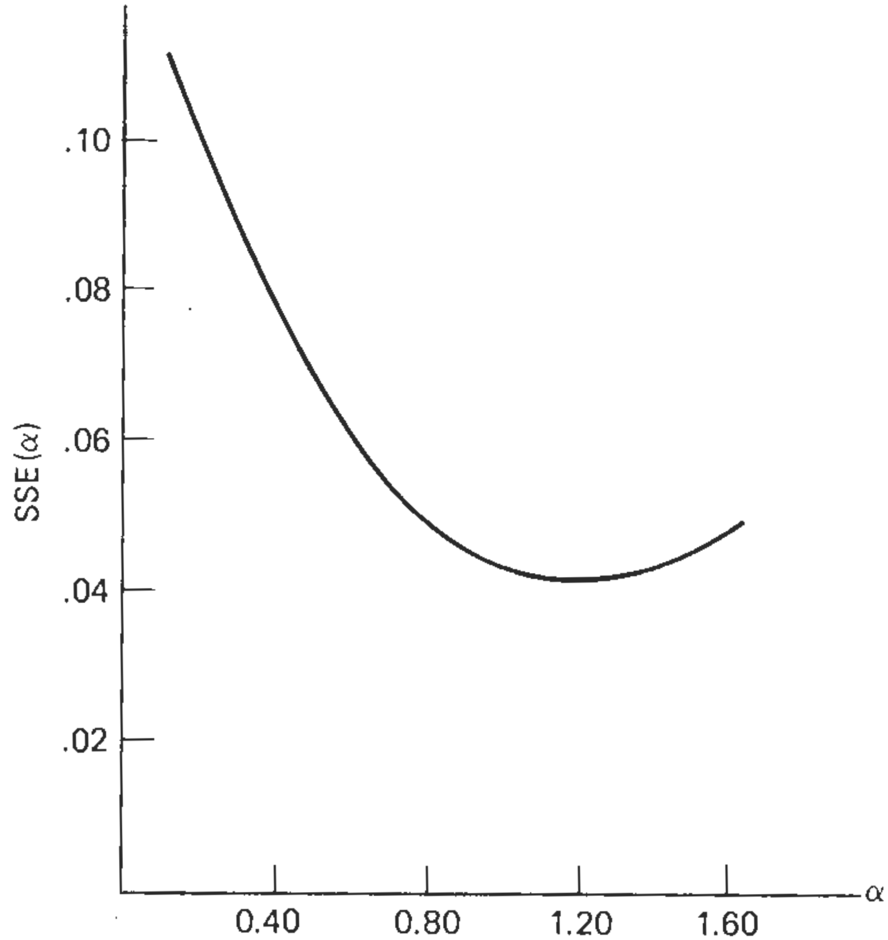


Figure 4.9. Plot of the sum of the squared one-step-ahead forecast errors from the model

$$z_{n+j} = \ln y_{n+j} = \beta_0 + \beta_1 j + \sum_{i=1}^{3} \delta_i \mathrm{IND}_{ji} + \varepsilon_{n+j}$$

New plant and equipment expenditures.

Table 4.7. General Exponential smoothing (α = 1.20) for the Model[a]

$$z_{n+j} = \ln y_{n+j} = \beta_0 + \beta_1 j + \sum_{i=1}^{3} \delta_i \text{IND}_{ji} + \varepsilon_{n+j}$$

New Plant and Equipment Expenditures

Time	z_t	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\delta}_1$	$\hat{\delta}_2$	$\hat{\delta}_3$	$\hat{z}_t(1)$	$e_{t-1}(1)$ [$= z_t - \hat{z}_{t-1}(1)$]
0		2.580	.019	−.216	−.081	−.104	2.384	
1	2.303	2.303	−.005	.143	.144	.272	2.440	−.081
2	2.472	2.472	.004	−.002	.117	−.165	2.474	.032
3	2.460	2.460	.000	.120	−.157	.013	2.580	−.015
⋮								
43	3.340	3.340	.029	.097	−.100	.027	3.466	−.013
44	3.463	3.463	.028	−.196	−.069	−.095	3.295	−.003
45	3.251	3.251	.015	.132	.119	.227	3.398	−.044
46	3.347	3.347	.000	−.008	.115	−.097	3.340	−.050
47	3.325	3.325	−.005	.125	−.083	.018	3.444	−.015
48	3.426	3.426	−.010	−.206	−.099	−.111	3.210	−.019
49	3.253	3.253	.003	.102	.077	.175	3.358	.043
50	3.391	3.391	.013	−.029	.060	−.125	3.375	.033
51	3.415	3.415	.025	.084	−.113	.001	3.524	.040
52	3.542	3.541	.030	−.199	−.091	−.097	3.373	.018

[a]Smoothing constant α and initial value $\hat{\beta}_0$ are determined from the first 44 observations.

Table 4.8. Sample Autocorrelations of the One-Step-Ahead Forecast Errors from the Model

$$z_{n+j} = \ln y_{n+j} = \beta_0 + \beta_1 j + \sum_{i=1}^{3} \delta_i \, \mathrm{IND}_{ji} + \varepsilon_{n+j}$$

with Smoothing Constant $\alpha = 1.20$.—New Plant and Equipment Expenditures ($n = 44$)

Lag k	r_k	Lag k	r_k
1	−.02	5	−.18
2	.07	6	−.08
3	.22	7	−.17
4	−.27	8	−.39

Table 4.9. Forecasts $\hat{y}_t(l) = \exp[\hat{z}_t(l)]$ of Original Data y_{t+l}.— New Plant and Equipment Expenditures

Time[a] t	$\hat{y}_t(1)$	$\hat{y}_t(2)$	$\hat{y}_t(3)$	$\hat{y}_t(4)$	y_t
44	27.00	31.53	31.63	35.77	31.92
45	29.90	29.96	33.89	27.41	25.82
46	28.22	31.91	25.82	28.42	28.43
47	31.31	25.33	27.91	27.28	27.79
48	24.78	27.28	26.66	29.49	30.74
49	28.73	28.07	31.06	26.15	25.87
50	29.22	32.33	27.19	31.22	29.70
51	33.92	28.53	32.75	33.55	30.41
52	29.17	33.48	34.26	38.90	34.52

[a] $t = 44$ corresponds to the fourth quarter of 1964.