2.2 Elimination of Trend and Seasonality

Here we assume that the TS model is additive and there exist both trend and seasonal components, that is

$$X_t = m_t + s_t + Y_t, \tag{2.6}$$

where the noise fluctuates about zero, i.e.,

$$\mathrm{E}(Y_t) = 0,$$

the seasonality component s_t is such that

$$s_t = s_{t-d},$$

where d denotes the length of the period and

$$\sum_{k=1}^{d} s_k = 0.$$

With a seasonal effect of a constant period length d it is convenient to index the data by the number of the season and the number within the season, for example monthly data ($d = 12$) for b years would be denoted by

$$x_{jk}, \quad j = 1, \dots, b, \quad k = 1, \dots, 12.$$

As before we want to extract the residuals $\widehat{Y}_t$ in order to examine their statistical properties. The following methods allow for estimation of the trend and the seasonal components.

2.2.1 Small Trend Method

This method is useful when the time series has a small trend and we may assume that the trend within each period is constant. Then, due to the assumptions of model 2.6, the period average is an unbiased estimator of the trend, that is

$$\widehat{m}_j = \frac{1}{d}\sum_{k=1}^{d} X_{jk}.$$

The seasonal component estimator, which satisfies the model assumptions is

$$\widehat{s}_k = \frac{1}{b}\sum_{j=1}^{b} (X_{jk} - \hat{m}_j).$$

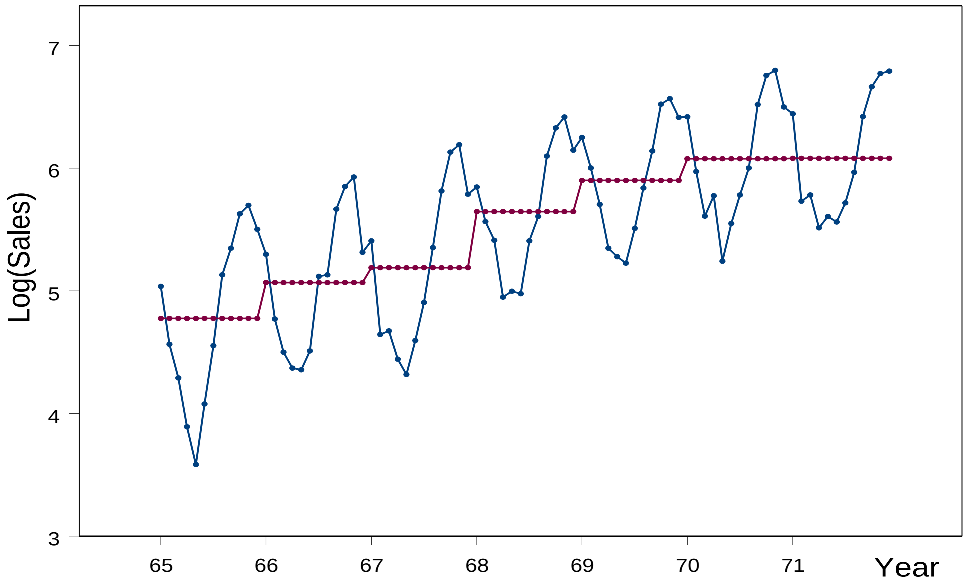

Figure 2.10: Transformed data and the Trend: Sales of an industrial heater.

Removing the estimates of trend and seasonality from the TS, we obtain the residuals

$$\widehat{Y}_{jk} = X_{jk} - \widehat{m}_j - \widehat{s}_k, \quad \text{for } j = 1, \ldots, b, \; k = 1, \ldots, d.$$

In Figures 2.10 - 2.14 there are plotted the transformed data representing sales of an industrial heater in successive months in years 1965 - 1971, constant trend within a year, detrended observations $x_{jk} - \widehat{m}_j$, the estimated seasonal components $\widehat{s}_k$, deseasonlized data $x_{jk} - \widehat{s}_k$ and the residuals $\widehat{y}_{jk}$ respectively.

2.2.2 Classical Decomposition

The method, presented by Brockwell and Davis (2002), consists of the following steps:

Step 1 Estimate trend using a moving average filter of the period length d, that is estimate trend by

$$\widehat{m}_t = \begin{cases} \frac{1}{d}\left(\frac{1}{2}X_{t-q} + X_{t-q+1} + \ldots + \frac{1}{2}X_{t+q}\right) & \text{for } d = 2q, \; q < t < n - q, \\ \frac{1}{d}\left(X_{t-q} + X_{t-q+1} + \ldots + X_{t+q}\right) & \text{for } d = 2q + 1, \; q + 1 < t < n - q. \end{cases}$$

Step 2 Estimate seasonal effects s_k for $k = 1, \ldots, d$:
compute the averages of the detrended values $(X_l - \hat{m}_l)$, $q < l = k + jd \leq$

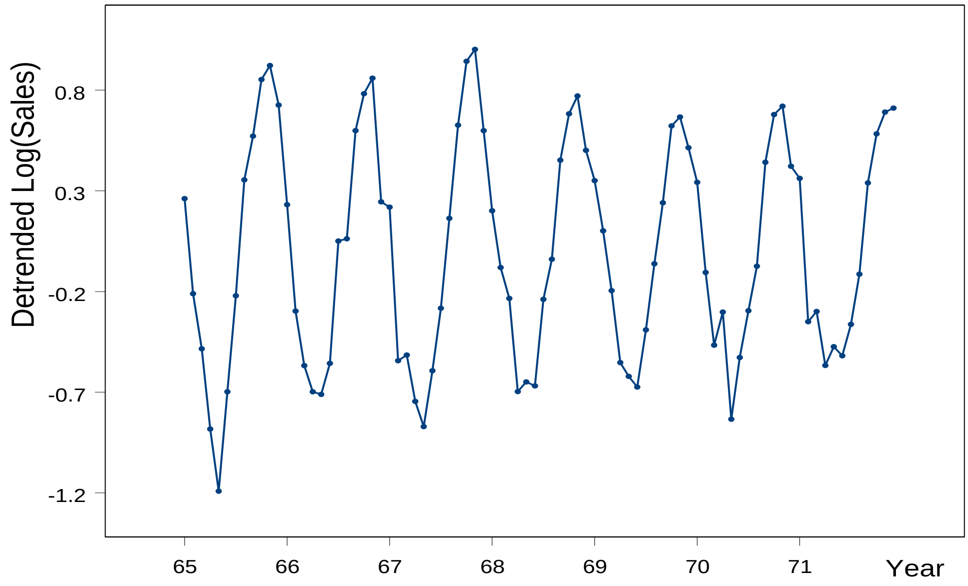

Figure 2.11: Detrended data: Sales of an industrial heater.

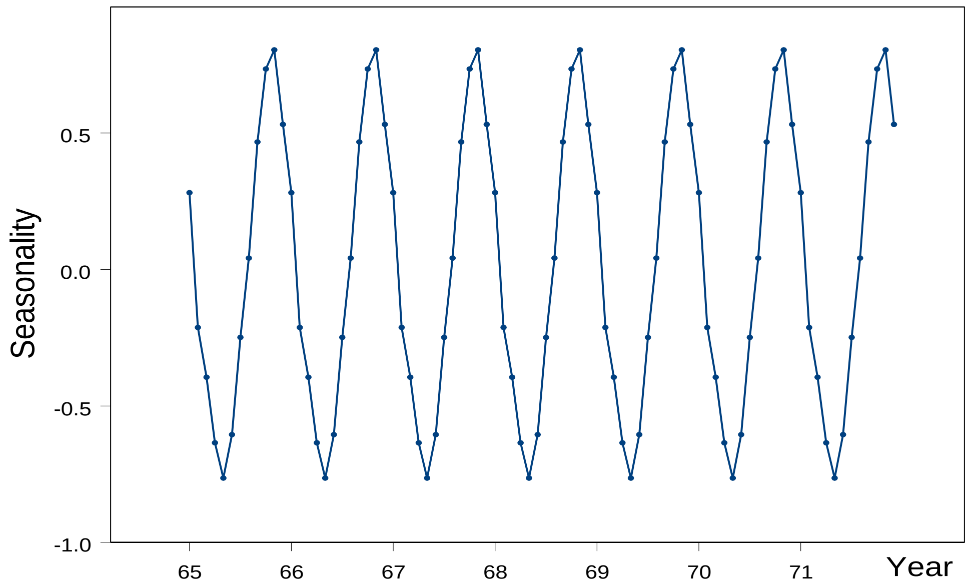

Figure 2.12: Seasonal component: Sales of an industrial heater.

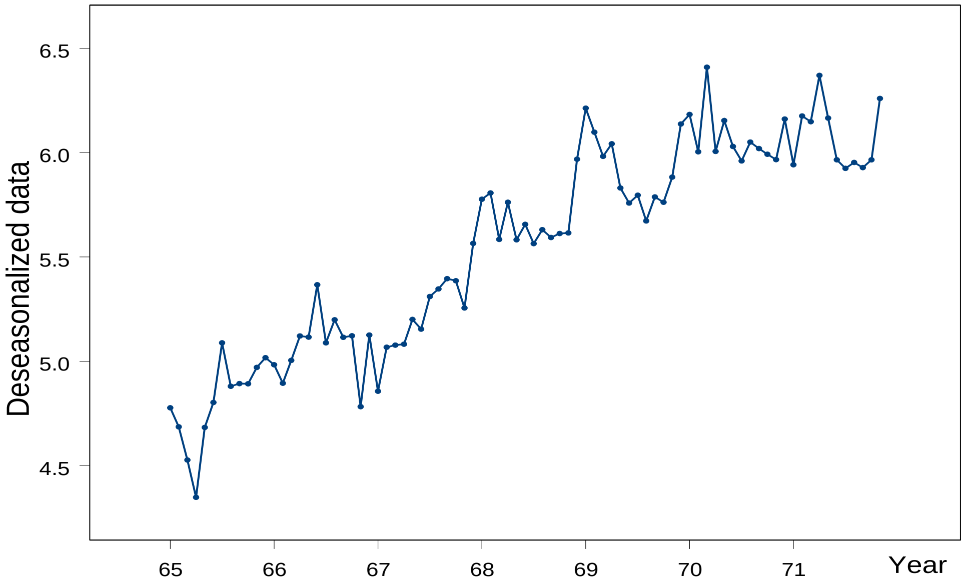

Figure 2.13: Deseasonalized data: Sales of an industrial heater.

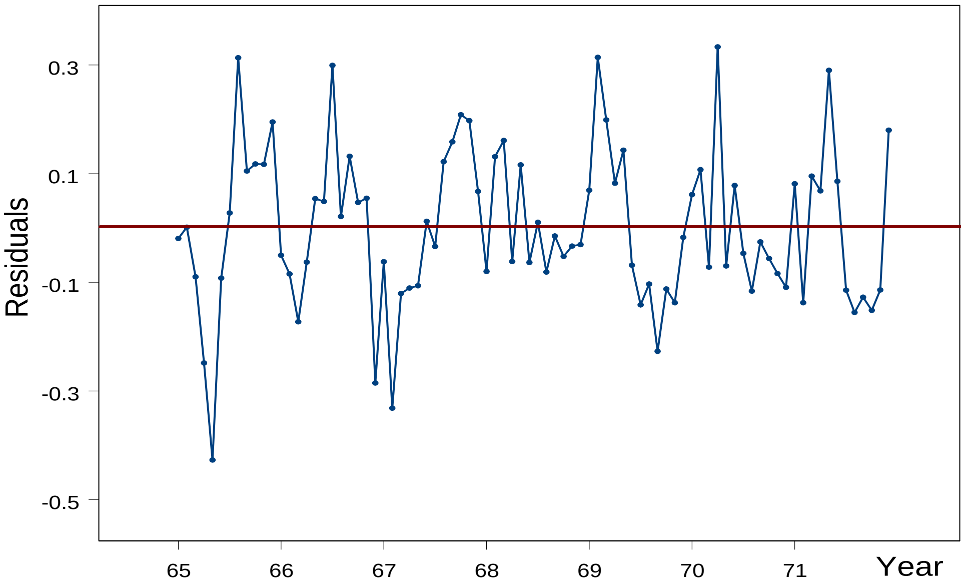

Figure 2.14: Residuals after removing trend and seasonal effects: Sales of an industrial heater.

$n-q$, $j = 1, \ldots, b$ and adjust them so as the seasonal effects meet the model assumptions, that is estimate the seasonal component s_k as

$$\widehat{s}_k = \begin{cases} \overline{(X_l - \widehat{m}_l)}_k - \frac{1}{d}\sum_{i=1}^{d}\overline{(X_l - \widehat{m}_l)}_i, & k = 1, \ldots, d, \\ \widehat{s}_{k-d}, & k > d. \end{cases}$$

Step 3 Remove the seasonality to obtain

$$D_t = X_t - \widehat{s}_t, \quad t = 1, \ldots, n.$$

Step 4 Re-estimate the trend $\widehat{m}_t$ from the deseasonalized variables $\{D_t\}$.

Step 5 Calculate the residuals

$$\widehat{Y}_t = X_t - \widehat{m}_t - \widehat{s}_t.$$

The decomposition used in MINITAB follows the same ideas but it is done in a different order. Also, it uses median, not mean, for estimating seasonal effects. It involves the following steps:

1. Fit a trend line to the data, using least squares regression.

2. Detrend the data by subtracting the trend component from the data (additive model).

3. Smooth the data using a centered moving average with a length equal to the length of the seasonal cycle.

4. Subtract the m.a. from the detrended data to obtain what are often referred to as raw seasonals.

5. Within each seasonal period, the median value of the raw seasonals is found. The medians are adjusted so that their sum is zero.
   These adjusted medians constitute the so called **seasonal indices**.

6. The seasonal indices are used in turn to seasonally adjust the data.

2.2.3 Differencing at lag d

We define the **lag-d differencing operator** by

$$\nabla_d X_t = X_t - X_{t-d} = (1 - B^d)X_t. \tag{2.7}$$

Applying the lag-d operator to the model (2.6) we obtain

$$\begin{aligned}\nabla_d X_t &= (m_t + s_t + Y_t) - (m_{t-d} + s_{t-d} + Y_{t-d}) \\ &= m_t - m_{t-d} + Y_t - Y_{t-d}.\end{aligned}$$

This removes the seasonal effect. Then to remove trend we may apply one of the methods described in Section 2.1, for example the method using the lag-1 difference operator ∇.

Later during the course we will be modelling the residuals. Knowing their properties will allow us to forecast future noise values in terms of their past values (if they are dependent) and so to forecast future values of the variable of interest X. First however, we need to recall some properties of random variables and their distributions.