# Estimation / elimination of both trend and seasonality

Suppose we have a trend-seasonal model

$$Y_t = m_t + S_t + e_t \;; \quad E(e_t) = 0$$
$$V(e_t) = \sigma^2 < \infty$$

Note that if $d$ is the period of seasonality, then we assume that

$$\cdots = S_{t-2d} = S_{t-d} = S_t = S_{t+d} = S_{t+2d} = \cdots$$

and also $\sum_{j=1}^{d} S_j = 0$

Suppose, for illustration, that we have monthly data with 12-month period.

Let us write the time index $t$ as

$$t = 12(j-1) + k$$

$j$ : year no. $j = 1(1)J$

$k$ : month no. $k = 1(1)12$

$Y_t$ is written as $Y_{j,k} = Y_{12(j-1)+k}$

## Method 1 : Slow trend method

In case the trend is slow, it is assumed that the trend remains constant during a particular year i.e. $m_j$ for a particular year $j$ is constant

**Step I** : estimate trend as

$$\hat{m}_j = \frac{1}{12} \sum_{k=1}^{12} Y_{j,k} \quad \left(\text{since } \sum_{k=1}^{12} S_k = 0\right)$$

$j = 1(1)J$

Step II : Estimate seasonal factors as

$$\hat{s}_K = \frac{1}{J}\sum_{j=1}^{J}\left(y_{j,K} - \hat{m}_j\right)$$

logic

data

Years

$$\begin{array}{c|cccc}
 & 1 & 2 & \cdots & J \\
\hline
1 & y_{1,1} & y_{2,1} & & y_{J,1} \\
2 & y_{1,2} & y_{2,2} & & y_{J,2} \\
\vdots & \vdots & \vdots & & \vdots \\
12 & y_{1,12} & y_{2,12} & & y_{J,12} \\
 & \downarrow & \downarrow & & \downarrow \\
 & \hat{m}_1 & \hat{m}_2 & & \hat{m}_J
\end{array}$$

Months

de-trended data

$$\longrightarrow \quad \begin{array}{ccc}
\text{year \#1} & & \text{year \# J} \\
\begin{pmatrix} y_{1,1}-\hat{m}_1 \\ y_{1,2}-\hat{m}_1 \\ \vdots \\ y_{1,12}-\hat{m}_1 \end{pmatrix} & \cdots & \begin{pmatrix} y_{J,1}-\hat{m}_J \\ \vdots \\ y_{J,12}-\hat{m}_J \end{pmatrix}
\end{array}$$

Seasonal factor data.

Average values, over years, for the months are (for $K^{th}$ month), $\hat{s}_K$'s

$$\hat{s}_K = \frac{1}{J\,\cancel{12}}\sum_{j=1}^{J}\left(y_{j,K} - \hat{m}_j\right)$$

$$\hat{e}_{j,K} = y_{j,K} - \hat{m}_j - \hat{s}_K \;; \quad j = 1(1)J, \; K = 1(1)12$$

Note that

$$\sum_{K=1}^{12}\hat{S}_K = \frac{1}{J\,\cancel{12}}\sum_{K=1}^{12}\sum_{j=1}^{J}\left(y_{j,K} - \hat{m}_j\right)$$

$= 0$ (Similar to the true ones $s_K$'s).

Method 2 : Fast trend method

In case there is a significant trend which can not be assumed to be constant for a year, we proceed in the following way :

Step I : Obtain rough estimate of trend

Use an MA filter, filter coefficients are such that seasonal component is eliminated and noise is dampened (i.e. the output process has lower variance than the original time series)

For a monthly data with period of seasonality 12, use a 12 point moving average to achieve the above.

$d = 12 \ (= 2q)$

$\underline{q=6}$

$$\hat{m}_t = \frac{1}{12}\left(\frac{1}{2} y_{t-6} + y_{t-5} + \cdots + y_t + \cdots + y_{t+5} + \frac{1}{2} y_{t+6}\right)$$

In general,

$$\hat{m}_t = \frac{1}{2q}\left(\frac{1}{2} y_{t-q} + y_{t-q+1} + \cdots + y_t + \cdots + y_{t+q-1} + \frac{1}{2} y_{t+q}\right)$$

or $$\hat{m}_t = \frac{1}{2q+1}\left(y_{t-q} + \cdots + y_t + \cdots + y_{t+q}\right)$$

depending on the period of seasonality (even or odd)

Step II : Estimation of seasonal components

For each month K (K = 1, .., 12) ; compute the average (say $w_K$) of deviations

$$\left\{ y_{12(j-1)+K} - \hat{m}_{12(j-1)+K} : j = 1, \ldots, J \right\}$$

over the J years.

Estimate $s_k$ as

$$\hat{s}_k = \omega_k - \frac{1}{12}\sum_{k=1}^{12}\omega_k$$

$$\& \quad \hat{s}_k = \hat{s}_{k-d} \quad \forall\ k > d$$

In general $$\hat{s}_k = \omega_k - \frac{1}{d}\sum_{k=1}^{d}\cancel{s}\,\omega_k$$

Remark: Note that $\omega_k$s that we obtained here are similar to the estimates of $s_k$s obtained in Method 1. However, the $\{\omega_k\}$ sequence that we have obtained here are not used as estimates of $s_k$ under the current setup as $\sum \omega_k \neq 0$. With the centering, we have $\sum_{k=1}^{d}\hat{s}_k = 0$.

Step iii : Deseasonalize the data

$$d_t = y_t - \hat{s}_t \quad ; \quad t = 1, \ldots, n$$

and get $(d_1, \ldots, d_n)$

Step iv :

Re-estimate trend using $(d_1, \ldots, d_n)$ using any of the trend estimation methods

Remark: Iterate, if required.

## Method 3: Elimination of trend & seasonality

using differencing, we can eliminate both trend and seasonality from the data (if they are present)

Define a lag d difference operator

$$\nabla_d y_t = y_t - y_{t-d} = (1-B^d)\, y_t$$

Apply $\nabla_d$ to $y_t = m_t + s_t + e_t$ ; $E(e_t) = 0$

where d is the period of seasonality, $V(e_t) = \sigma^2 < \infty$ hence

$$\ldots = s_{t-d} = s_t = s_{t+d} = \ldots$$

$$Z_t = \nabla^d y_t = \nabla^d m_t + \nabla^d s_t + \nabla^d e_t$$

$$Z_t = \underbrace{(m_t - m_{t-d})}_{\substack{\text{deterministic time}\\ \text{trend component}}} + \underbrace{\cancel{(s_t - s_{t-d})}}_{0} + \underbrace{(e_t - e_{t-d})}_{\substack{\text{irregular}\\ \text{component}}}$$

Trend $(m_t - m_{t-d})$ term in $Z_t$ can be eliminated through differencing of appropriate power of $\nabla$ operator.

## Mathematical formulation of a time series

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $T$ be an index set

<u>Def$^n$</u>: A real valued time series is a real valued function $X(t,\omega)$ defined on $T \times \Omega$, $\ni$ for a fixed $t$, $X(t,\omega)$ $\left(= X_t(\omega) = X_t \text{, say}\right)$ is a random variable defined on $(\Omega, \mathcal{F}, P)$.

A time series is thus a collection $\{X_t : t \in T\}$ of random variables.