3. Seasonality

We continue working with the log-transformed Singapore Industrial Production Index, 1983M1 to 2017M12. The figure below shows both the log and first difference of that series.

3.1 Seasonality refers to fixed-period patterns that occur in a time series, usually for “mechanical” reasons. The seasonality in the log(SG_IP) series is noticeable in the original (not shown here) and log-transformed series. It is quite obvious in the first difference series, though perhaps it is not immediately clear that the spikes occur at fixed periods.

There are other ways to visually explore seasonality in time series data. We show in the next two diagrams the “Seasonal” and the “Seasonal Subseries” plots for both the first difference and original series. The seasonal plot shows one line per year, arranged by month. The seasonal subseries plot shows a sequence of lines, one per month, each containing observations for all years for that month.
Seasonalities usually have periods of one year for quarterly, monthly, and possibly higher
certainty data. Daily data often show seasonality with periods of one week. There are some
seasonalities that do not have the usual periods. For instance, the Hari Raya holidays follow
the (roughly) 11-month lunar cycle, and any seasonality associated with this holiday will have
an 11-month period. It is possible that there are both 11-month and 12-month period
seasonalities in a time series. There could be seasonalities that mostly follow a certain period.
The Chinese Lunar New Year holidays usually occur in February, but sometimes in January.

3.2 One way of modelling seasonality is to use seasonal dummies. These are indicator
variables marking the month of the data (for monthly data). For monthly data, the January
dummy takes value ‘1’ for January observations, ‘0’ otherwise, and so on. The following 12
columns represents the first 18 observations for each of the twelve monthly seasonal dummies:

<table>
<thead>
<tr>
<th>Jan 83</th>
<th>Feb 83</th>
<th>Mar 83</th>
<th>Apr 83</th>
<th>May 83</th>
<th>Jun 83</th>
<th>Jul 83</th>
<th>Aug 83</th>
<th>Sep 83</th>
<th>Oct 83</th>
<th>Nov 83</th>
<th>Dec 83</th>
<th>Jan 84</th>
<th>Feb 84</th>
<th>Mar 84</th>
<th>Apr 84</th>
<th>May 84</th>
<th>Jun 84</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0</td>
<td>0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0</td>
<td>0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0</td>
<td>0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0</td>
<td>0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0</td>
<td>0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0</td>
<td>0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
There are many ways to use these seasonal dummies. The following are three equivalent specifications, where $d_{Jan,t}$ represents the January dummy variable, they give the exact same fit, though the interpretation of the parameters are different:

$$Y_t = \beta_1 d_{Jan,t} + \beta_2 d_{Feb,t} + \beta_3 d_{Mar,t} + \cdots + \beta_{12} d_{Dec,t} + \varepsilon_t$$  \hspace{1cm} (3.2.1a)

$$Y_t = \alpha_0 + \alpha_2 d_{Feb,t} + \alpha_3 d_{Mar,t} + \cdots + \alpha_{12} d_{Dec,t} + \varepsilon_t$$  \hspace{1cm} (3.2.1b)

$$Y_t = \delta_0 + \delta_2(d_{Feb,t} - \frac{1}{12}) + \delta_3(d_{Mar,t} - \frac{1}{12}) + \cdots + \delta_{12}(d_{Dec,t} - \frac{1}{12}) + \varepsilon_t$$  \hspace{1cm} (3.2.1c)

We leave the interpretation of the parameters as an exercise. Note that if you include the constant term, you must leave out at least one dummy variable. Including all monthly dummy variables and a constant would lead to perfectly collinear regressors (sometimes referred to as the “dummy variable trap”). In our example we left out the January dummy variable. You can leave out any one particular dummy variable, only the interpretations change.

We can combine seasonal dummies with other variables. The following is the fit of the “quadratic deterministic trend + seasonal dummy model” model

$$Y_t = \alpha_0 + \alpha_2 d_{Feb,t} + \alpha_3 d_{Mar,t} + \cdots + \alpha_{12} d_{Dec,t} + \beta_1 t + \beta_2 t^2 + \varepsilon_t$$  \hspace{1cm} (3.2.2)

What remains (the residuals) is the de-trended, seasonally-adjusted SG_IP series, showing only cycles. We will discuss cycles in a later session. The $R^2$ turns out to 0.9752, though it should be pointed out that a very high $R^2$ is common when modelling trending series.
3.4 Statistical agencies often provide seasonally-adjusted (s.a.) series time series data alongside the non-seasonally-adjusted (n.s.a.) versions. We often work with the seasonally-adjusted version, since the seasonality itself is not typically of much interest. Seasonally-adjusted series are produced by methods somewhat more sophisticated than the seasonal dummy approach which we do not discuss here. Later we will consider non-deterministic seasonalities.

**Exercises**

1. How do cycles differ from seasonality? The seasonal dummies approach makes several assumptions about the nature of the seasonality. Can you think of some of these assumptions?

2. Explore visually the series SG_ELEC_GEN in the file TS_1.csv, which contains monthly electricity generated in Singapore, and report your observations. What do you think is driving the seasonality? Fit an appropriate deterministic trend + seasonal dummy model to the series, and report your estimation results.

3. Find interpretations for the parameters in each of the three seasonal dummy models 3.2.1a, 3.2.1b, and 3.2.1c. In 3.2.1b, how will your interpretations change if the December dummy is left out instead of the January dummy? How would you test for the presence of seasonality in each of the three models? (State the exact hypothesis you would test)