Time Series Analysis
• Motivation

• Decomposition Models
  ◦ Additive models, multiplicative models

• Global Approaches
  ◦ Regression
  ◦ With and without seasonal component

• Local Approaches
  ◦ Moving Averages Smoothing
  ◦ With and without seasonal component

• Summary
Example: Temperatures data set (fictive)

- The plot shows the average temperature per day for 50 years.
- Is there any trend visible?
- How to extract seasonal effects?
 Decomposition Models

• The time series is given as a sequence of values

\[ y_1, \ldots, y_t, \ldots, y_n \]

• We assume that every \( y_t \) is a composition of (some of) the following components:
  \[ g_t \] trend component
  \[ s_t \] seasonal component
  \[ c_t \] cyclical variation
  \[ \epsilon_t \] irregular component (random factors, noise)

• Assume a functional dependency:

\[ y_t = f(g_t, s_t, c_t, \epsilon_t) \]
Components of Time Series

**Trend Component**
- Reflects long-term developments.
- Often assumed to be a monotone function of time.
- Represents the actual component we are interested in.

**Cyclic Component**
- Reflects mid-term developments.
- Models economical cycles such as booms and recessions.
- Variable cycle length.
- We do not consider this component here.

Remark: Often, both components are combined.
Components of Time Series

**Seasonal Component**
- Reflects short-term developments.
- Constant cycle length (i.e., 12 months)
- Represents changes that (re)occur rather regularly.

**Irregular Component**
- Represents everything else that cannot be related to the other components.
- Combines irregular changes, random noise and local fluctuations.
- We assume that the values are small and have an average of zero.
Decompositions

Additive Decomposition

\[ y_t = g_t + s_t + \epsilon_t \]

- Pure trend model:  \( y_t = g_t + \epsilon_t \) (stock market, no season)
- Possible extension:  \( y_t = g_t + s_t + x_t \beta + \epsilon_t \) (calendar effects)

Multiplicative Decomposition

\[ y_t = g_t \cdot s_t \cdot \epsilon_t \]

- Seasonal changes may increase with trend.
- Transform into additive model:

\[ \tilde{y}_t = \log y_t + \log s_t + \log \epsilon_t \]
**Goal:** Estimate the components from a given time series, i.e.

\[ \hat{y}_t + \hat{s}_t + \epsilon_t \approx y_t \]

**Application:** With the estimates, we can compute the

- trend-adjusted series: \( y_t - \hat{g}_t \)
- season-adjusted series: \( y_t - \hat{s}_t \)
- We only consider additive models here.

⇒ Additional assumptions necessary in order to find ways to infer the desired components.
• **Global approach:** There is a fix functional dependence throughout the entire time range. (⇒ regression models)

• **Local approach:** We do not postulate a global model and rather use local estimations to describe the respective components.

• **Seasonal effects:** We have to decide beforehand whether to assume a seasonal component or not.

<table>
<thead>
<tr>
<th></th>
<th>Global</th>
<th>Local</th>
</tr>
</thead>
<tbody>
<tr>
<td>without Season</td>
<td>Regression</td>
<td>Smoothing Averages</td>
</tr>
<tr>
<td>with Season</td>
<td>Dummy Variables</td>
<td>Smoothing Averages</td>
</tr>
</tbody>
</table>
Global Approach (without Season)

Model: \[ y_t = g_t + \epsilon_t \]

Assumptions:

- No seasonal component: \( s_t = 0 \)
- Depending on \( g_t \), use regression analysis to estimate the parameter(s) to define the trend component.
  - linear trend: \( g_t = \beta_0 + \beta_1 t \)
  - quadratic trend: \( g_t = \beta_0 + \beta_1 t + \beta_2 t^2 \)
  - polynomial trend: \( g_t = \beta_0 + \beta_1 t + \cdots + \beta_q t^q \)
  - exponential trend: \( g_t = \beta_0 \exp(\beta_1 t) \)
  - logistic trend: \( g_t = \frac{\beta_0}{\beta_1 + \exp(-\beta_2 t)} \)
Global Approach (with Season)

Model: \( y_t = s_t + \epsilon_t \)  (no trend)

Assumptions:

• No trend component: \( g_t = 0 \)

• Seasonal component does not change from period to period.

• Introduce *dummy variables* for every time span (here: months) that serve as indicator functions to determine to which month a specific \( t \) belongs:

\[
s_m(t) = \begin{cases} 
1, & \text{if } t \text{ belongs to month } m \\
0, & \text{otherwise}
\end{cases}
\]

• The seasonal component is then set up as

\[
s_t = \sum_{m=1}^{12} \beta_m s_m(t).
\]

• Determine the *monthly effects* \( \beta_m \) with normal least squares method.
Global Approach (with Season)

Model:  \[ y_t = g_t + s_t + \epsilon_t \]

Assumptions:

- Estimate \( \hat{g}_t \) while temporarily ignoring \( s_t \).
- Estimate \( s_t \) from the trend-adjusted \( \tilde{y}_t = y_t - \hat{g}_t \).

Model:  \[ y_t = \alpha_1 t + \cdots + \alpha_q t^q + \cdots + \beta_1 s_1(t) + \cdots + \beta_{12} s_{12}(t) + \epsilon_t \]

Assumptions:

- Seasonal component does not change from period to period.
- Model the seasonal effects with trigonometric functions:

\[ s_t = \beta_0 + \sum_{m=1}^{6} \beta_m \cos\left(2\pi \frac{m}{12} t\right) + \sum_{m=1}^{5} \gamma_m \sin\left(2\pi \frac{m}{12} t\right) \]

- Determine \( \alpha_1, \ldots, \alpha_q, \beta_0, \ldots, \beta_6 \) and \( \gamma_1, \ldots, \gamma_5 \) with normal least squares method.
Local Approach (without Season)

**General Idea:** Smooth the time series.

- Estimate the trend component $g_t$ at time $t$ as the average of the values around time $t$.

For a given time series $y_1, \ldots, y_n$, the **Smoothing Average** $y_t^*$ of order $r$ is defined as follows:

$$
y_t^* = \begin{cases} 
\frac{1}{2k + 1} \cdot \sum_{j=-k}^{k} y_{t+j}, & \text{if } r = 2k + 1 \\
\frac{1}{2k} \cdot \left( \frac{1}{2} y_{t-k} + \sum_{j=-k+1}^{k-1} y_{t+j} + \frac{1}{2} y_{t+k} \right), & \text{if } r = 2k
\end{cases}
$$
Local Approach (without Season)

Model: \[ y_t = g_t + \epsilon_t \]

Assumptions:

- In every time frame of width \( 2k + 1 \) the time series can be assumed to be linear.
- \( \epsilon_t \) averages to zero.
- Then we use the smoothing average to estimate the trend component:

\[ \hat{g}_t = y_t^* \]
Local Approach (with Season)

Model: \[ y_t = g_t + s_t + \epsilon_t \]

Assumptions:

- Seasonal component has period length \( p \) (repeats after \( p \) points):
  \[ s_t = s_{t+p}, \quad t = 1, \ldots, n - p \]

- Sum of seasonal values is zero:
  \[ \sum_{j=1}^{p} s_j = 0 \]

- Trend component is linear in time frames of width \( p \) (if \( p \) is odd)
  or \( p + 1 \) (if \( p \) is even).

- Irregular component averages to zero.
Local Approach (with Season)

Let $k = \frac{p - 1}{2}$ (for odd $p$) or $k = \frac{p}{2}$ (for even $p$).

Then:

- Estimate the trend component with smoothing average:

  $$\hat{g}_t = y_t^*, \quad k + 1 \leq t \leq n - k$$

- Estimate the seasonal components $s_1, \ldots, s_p$ as follows:

  $$\hat{s}_i = \tilde{s}_i - \frac{1}{p} \sum_{j=1}^{p} \tilde{s}_j \quad \text{with} \quad \tilde{s}_t \frac{1}{m_i - l_i + 1} \sum_{j=l}^{m_i} (y_{i+jp} - y_{i+jp}^*), \quad 1 \leq i \leq p$$

  where

  $$m_i = \max \{m \in \mathbb{N}_0 \mid i + mp \leq n - k\}$$

  and

  $$l_i = \min \{l \in \mathbb{N}_0 \mid i + lp \geq k + 1\}$$
We can extract an increase and decrease of 1 degree during 50 years even though the amount of noise is more than twice as large than the actual trend.
Example

5 years period, trend ±8 degrees, noise amount ±2 degrees
Example

100 years period, trend ±1 degree, noise amount ±3 degrees
Summary

- **Definition of the problem domain**
  - Consider a time series to be composed of subcomponents.
  - Additive and multiplicative models.

- **Global and local approaches**
  - With and without seasonal components.

- **Robust to noise**
  - Noise can be higher than the trend component itself.