Introduction to Seasonal adjustment

STATISTICS
DENMARK

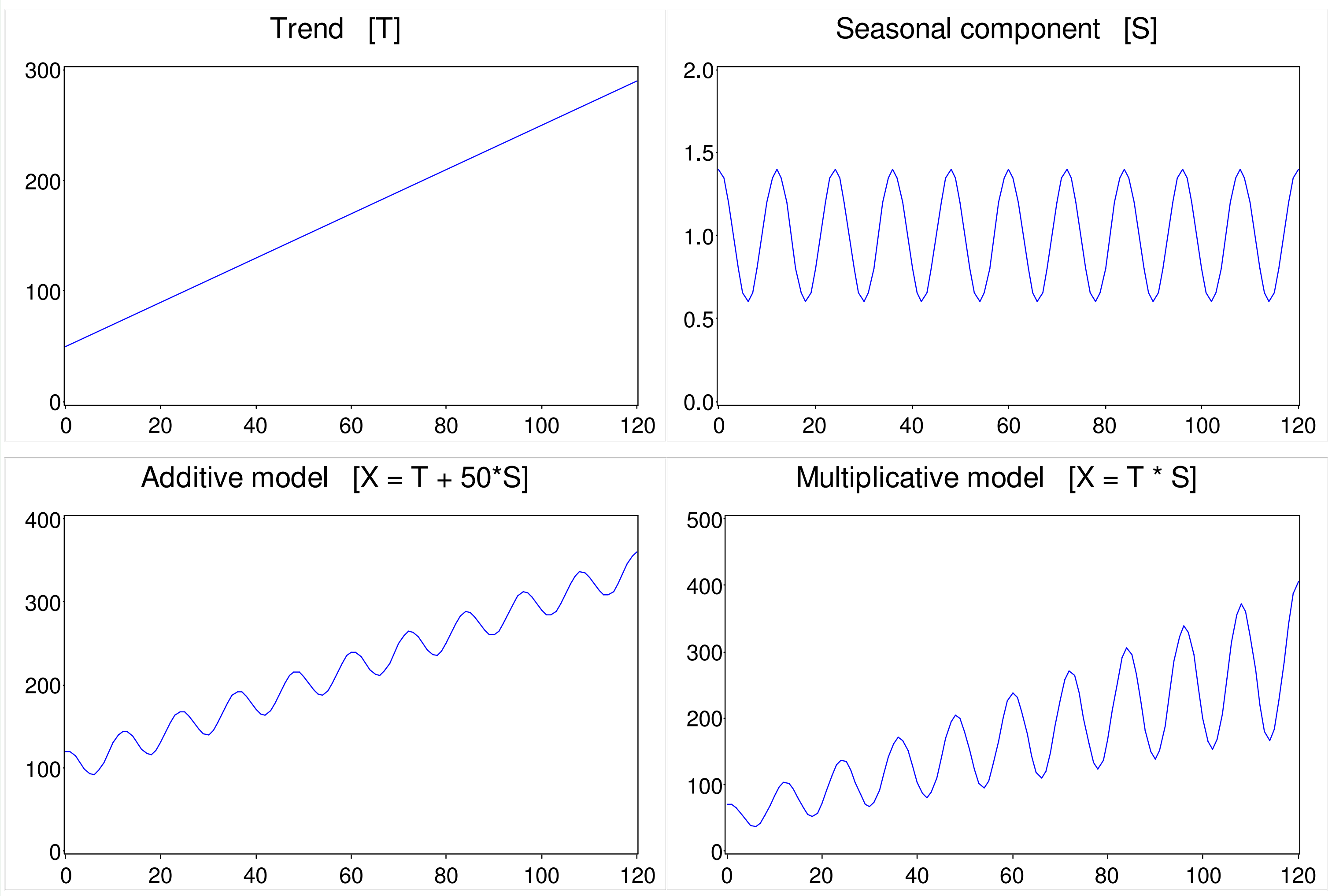
Decomposition of a time series

- What is seasonality?
- Unobserved components
- Additive or multiplicative model

What is seasonality

- No 100% definition!
- Fluctuations repeating themselves (more or less precisely) from year to year
- Intuitively, seasonality should sum to zero within a year (for an additive decomposition)

Simulated time series with seasonality



STATISTICS DENMARK

Model for unobservable components

- Additive model

  $$O_t = T_t + S_t + I_t = A_t + S_t$$

- Multiplicative model

  $$O_t = T_t \cdot S_t \cdot I_t = A_t \cdot S_t$$

Where *O* is the pre-adjusted series,
t is the time,
T is the trend-cycle,
S is the seasonal component,
I is the irregular component (white noise), and
A is the seasonally adjusted series.

No unique decomposition without conditioning!

Methods and software

- X-12-ARIMA
  - Relatively few assumptions
  - Non-parametric approach

- TRAMO/SEATS
  - Uses a mathematical model with many assumptions
  - Parametric approach

Demetra

- Developed by (not at) Eurostat
- Graphical interface for X-12-ARIMA and TRAMO/SEATS
- Currently under development – will feature functions for documenting seasonal adjustment

Pre-adjustment

Calendar effects and outliers

The flow in seasonal adjustment

1. Transformation (possibly)
   - To ensure stationarity (poss. log.transformation)
2. Pre-adjustment
   - Prepare for estimation of the season
3. ARIMA model
   - Find the model with the best description of the pre-adjusted time series
4. Forecast the series
   - Forecast with the ARIMA model
5. Decomposition
   - Split up in Trend, Season and Irregular component

STATISTICS DENMARK

Notation

Two possible models

- Additive: $O_t = T_t + S_t + I_t = A_t + S_t$
- Multiplikative (log.transformed):

$$O_t = T_t \cdot S_t \cdot I_t = A_t \cdot S_t$$

where O_t – the pre-adjusted time series

T_t – the trend

S_t – the seasonal component

I_t – the irregular component

A_t – the seasonal adjusted series

Regression

Normal regression $X_t = \beta Z_t + O_t$

- X_t the original series
- Z_t explanatory variable (one or more)
- β regression coefficient (one or more)
- O_t residuals (residual variation)

Pre-adjustment

$$X_t \quad = \beta_1 Z_{1t} + \ldots + \beta_k Z_{kt} \; + \; O_t$$

where X_t – the original series

Z_{it} – explicatory variables

β_i – regression coefficients

O_t – the preadjusted series (follows an ARIMA model)

Explicatory variables - Calendar

- Months/Quarters are not comparable
  - The length of the month differs
    - The more working days the higher production
    - Leap year
  - Type of the day differs
    - Sundays/Holy days versus normal trading days
    - Trade is bigger Saturday than Monday
  - Holydays can move (moving seasonality)
    - Easter
- Countries are not comparable
  - Easter in western and eastern Catholic church
  - Ramadan
  - Comparability in the EU (Eurostat)

Explicatory variables – External reasons

- Changes caused by external reasons:
  - Outliers (extreme/not typical observations)
    - Reduced sale of ice cream in a cold summer
    - Strike
    - Freezing days (construction business)
  - Level shift (permanent change)
    - Changes in duties and taxes e.g. on cars
    - Financial crisis
  - Transitory changes (temporary/provisional)
    - Felling of timber after a storm
    - Rise in price on e.g. coffee

Types of Pre-adjustment - 1

Two types of Pre-adjustments:

- **<u>Temporary</u>** corrections
  - Additive outliers (AO)
  - Level shift (LS)
  - Transitory change (TC)
  - Ramp
  - User Defined

  (Additive outliers, Level shift, Transitory change, Ramp: External reasons)

- **<u>Permanent</u>** corrections
  - Trading days
  - Easter
  - National Holy days
  - Leap Year
  - User Defined

  (Trading days, Easter, National Holy days, Leap Year: Calendar effects)

STATISTICS DENMARK

Types of Pre-adjustment - 2

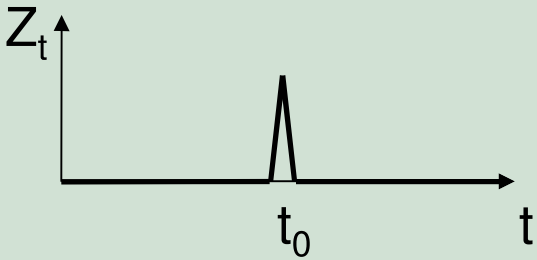
Difference between permanent and temporary corrections

After the decomposition

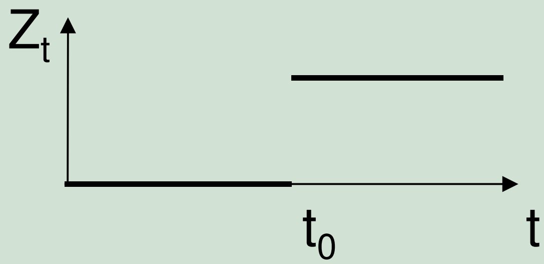
- temporary corrections will be taken back in the seasonal adjusted series (Level shifts e.g.)
  - These things have happened in the real world
- permanent corrections will not be taken back in the seasonal adjusted series
  - The purpose is to compare months/quarters (standard months, standard quarters)

Explicatory variables

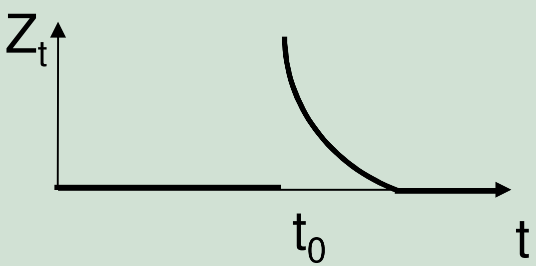
- <u>Temporary</u> corrections



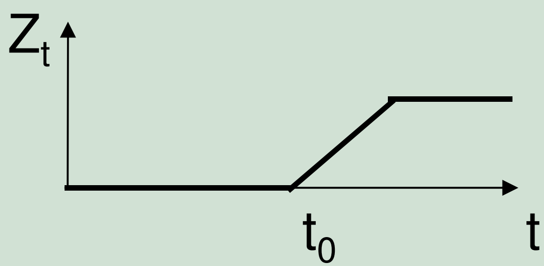
Additive outlier (AO)



Level shift (LS)



Transitory change (TC)



Ramp

Explicatory variables

- <u>Permanent</u> corrections

- The Production in February is on average 28.25 days.
- In normal years the Production will be increased from 28 days to 28.25 days
- In Leap years the Production will be reduced from 29 days to 28.25 days.

Explicatory variables

- Permanent corrections
  - Trading Days:
    - Z_t is the number of trading days in each month
      e.g. Z_t = (20,23,21,22,22,21,23,…)
    - Z_{1t} … Z_{7t} is the number of Mondays … Sundays
      e.g. Z_{1t} =(4,4,5,4,4,5,4,5,4,…)
    - All months are changed to include the same amount of trading days or days of each type (standard month).

Pre-adjustment in Demetra

- Automatic outlier/calendar correction:
  - LS/AO/TC, Easter, Leap year/ Trading days (7,6,2,1 regressors) via regression
  - Pre test for significance (t-test)
  - Effect included if significant

X-11

An algorithm for decomposition

$n{\times}m$-terms moving average

- An **n x m moving average** is an m-term simple average taken over n consecutive sequential spans.

$$\overline{X}_t^{3\times 3} = \frac{1}{3}\left(\frac{X_{t-2} + X_{t-1} + X_t}{3} + \frac{X_{t-1} + X_t + X_{t+1}}{3} + \frac{X_t + X_{t+1} + X_{t+2}}{3} \right)$$

$$= \frac{1}{9} X_{t-2} + \frac{2}{9} X_{t-1} + \frac{3}{9} X_t + \frac{2}{9} X_{t+1} + \frac{1}{9} X_{t+2}$$

Filters in X-11

- Estimation of trend-cycle from raw series: 2×12 (monthly figures) or 2×4 (quarterly figures)
- Estimation of seasonal component from detrended series (SI-ratio):

  $3\times3S$, $3\times5S$ or $3\times9S$ (only monthly series)

The X-11 algorithm

- Linear filters during three steps
  - Step 1: Initial estimates
  - Step 2: Final seasonal component
  - Step 3: Final trend and irregular component

- Let O_t be the pre-adjusted series and assume an additive decomposition model:

$$O_t = T_t + S_t + I_t = A_t + S_t$$

- The following slides show the decomposition of a monthly series…

X-11: Step 1 (1 of 3)

(i) Calculating initial trend estimate using 2×12 moving average:

$$T_t^{(1)} = \frac{1}{24}O_{t-6} + \frac{1}{12}O_{t-5} + \cdots + \frac{1}{12}O_t + \cdots + \frac{1}{12}O_{t+5} + \frac{1}{24}O_{t+6}$$

(ii) Calculating initial SI-ratio:

$$(S_t + I_t)^{(1)} = SI_t^{(1)} = O_t - T_t^{(1)}$$

X-11: Step 1 (2 of 3)

(iii) Calculating initial seasonal component using 3×3 seasonal average:

$$\tilde{S}_t^{(1)} = \frac{1}{9} SI_{t-24}^{(1)} + \frac{2}{9} SI_{t-12}^{(1)} + \frac{3}{9} SI_t^{(1)} + \frac{2}{9} SI_{t+12}^{(1)} + \frac{1}{9} SI_{t+24}^{(1)}$$

followed by normalization:

$$S_t^{(1)} = \tilde{S}_t^{(1)} - \left(\frac{1}{24} \tilde{S}_{t-6}^{(1)} + \frac{1}{12} \tilde{S}_{t-5}^{(1)} + \cdots + \frac{1}{12} \tilde{S}_{t-5}^{(1)} + \frac{1}{24} \tilde{S}_{t+6}^{(1)} \right)$$

X-11: Step 1 (3 of 3)

(iv) Calculating preliminary seasonally adjusted series:

$$A_t^{(1)} = O_t - S_t^{(1)}$$

- Choice of filters used during Step 1 does not depend on the series – they are fixed.

X-11: Step 1 (3 of 3)

(iv) Calculating preliminary seasonally adjusted series:

$$A_t^{(1)} = O_t - S_t^{(1)}$$

- Choice of filters used during Step 1 does not depend on the series – they are fixed.

X-11: Step 2 (1 of 3)

(i) Calculate intermediary trend using something called a Henderson filter. A different type of moving average, which empirically has proven useful.

(ii) Calculate SI-ratio:

$$(S_t + I_t)^{(2)} = SI_t^{(2)} = O_t - T_t^{(2)}$$

X-11: Step 2 (2 of 3)

(iii) Calculating seasonal factor using seasonal average (typically 3x5):

$$\tilde{S}_t^{(2)} = \frac{1}{15} SI_{t-36}^{(2)} + \frac{2}{15} SI_{t-24}^{(2)} + \frac{3}{15} SI_{t-12}^{(2)} + \frac{3}{15} SI_{t}^{(2)} + \frac{3}{15} SI_{t+12}^{(2)} + \frac{2}{15} SI_{t+24}^{(2)} + \frac{1}{15} SI_{t+36}^{(2)}$$

followed by normalization:

$$S_t^{(2)} = \tilde{S}_t^{(2)} - \left(\frac{1}{24} \tilde{S}_{t-6}^{(2)} + \frac{1}{12} \tilde{S}_{t-5}^{(2)} + \cdots + \frac{1}{12} \tilde{S}_{t-5}^{(2)} + \frac{1}{24} \tilde{S}_{t+6}^{(2)} \right)$$

X-11: Step 2 (3 of 3)

Seasonal filter is either $3{\times}3S$, $3{\times}5S$ or $3{\times}9S$ depending on changes in S_t relative to I_t. The shorter filters are preferred if the seasonal component changes a lot compared to the irregular component.

(iv) Calculating the final seasonally adjusted series (of the B iteration):

$$A_t^{(2)} = O_t - S_t^{(2)}$$

X-11: Step 3 (1 of 1)

(i) Calculating final trend using again a Henderson filter

$$T_t^{(3)} = \sum_{i=-H}^{H} h_i A_{t+i}^{(2)}$$

(ii) Calculating final irregular component:

$$I_t^{(3)} = A_t^{(2)} - T_t^{(3)}$$

such that the final decomposition is

$$\tilde{O}_t = T_t^{(3)} + S_t^{(2)} + I_t^{(3)}$$

Iterative use of the basic X-11 algorithm

- The X-11 algorithm is used 3 times to identify *extreme values* for temporary linearization of the series:
  - The B iteration finds preliminary extreme values
  - The C iteration finds final extreme values
  - The D iteration is the actual seasonal adjustment
- Extreme values of SI_t are replaced
- Therefore $\tilde{O}_t = T_t^{(3)} + S_t^{(2)} + I_t^{(3)} \neq O_t$
- $\tilde{O}_t$ from the B iteration is the starting point for the C iteration, and $\tilde{O}_t$ from the C iteration is the starting point for the D iteration.