

Outliers and Calendar Effects

ESTP Training
17-19 October 2023

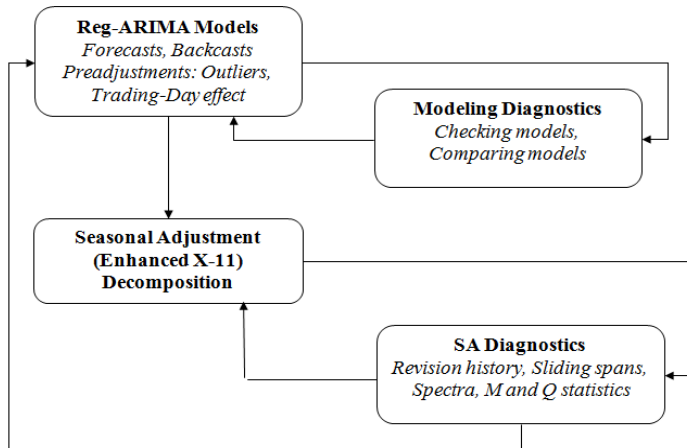
THE CONTRACTOR IS ACTING UNDER A FRAMEWORK CONTRACT CONCLUDED WITH THE COMMISSION

On the Menu today

- The RegARIMA model.
- The various kinds of outliers and ruptures.
- Trading-day and Calendar Effects
- Playing with JD+: Is there any calendar effects in your data?

X-13ARIMA-SEATS

Two building blocks: RegARIMA and X11.



The RegARIMA model

A usual additive model for SA is:

$$X_t = TC_t + S_t + O_t + TD_t + MH_t + I_t,$$

Where all the components are orthogonal and are non-seasonal (except S_t of course).

A SA in 2 steps:

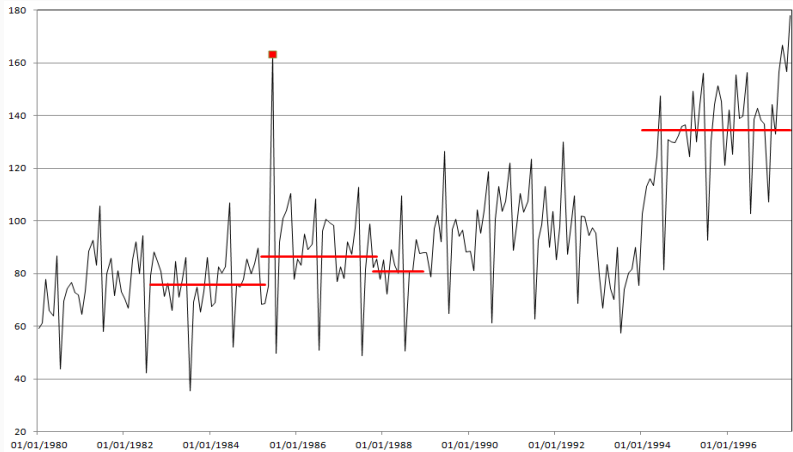
Step 1: $X_t = O_t + TD_t + MH_t + Z_t$; where Z_t follows an ARIMA model.

Step 2: Decomposition of $Z_t = TC_t + S_t + I_t$

Outliers

A First Example

Various ruptures in the series.



Additive outliers (AO) represent a one-off peak or trough in the time series at a single observation are defined as:

$$AO_t = \begin{cases} 1 & \text{if } t = t_0 \\ 0 & \text{if } t \neq t_0 \end{cases} \quad (1)$$

Temporary Changes

Temporary changes (TC) where the change in the level of the time series is not of permanent but temporary nature, i.e. the effect decreases exponentially ($0 < \alpha < 1$) with the course of time, have the following representation:

$$TC_t = \begin{cases} 0 & \text{if } t < t_0 \\ \alpha^{t-t_0} & \text{if } t \geq t_0 \end{cases} \quad (2)$$

Level Shifts

Level shifts (LS) permanently increase or decrease the (transformed) data by some constant factor prior to a certain observation can be written as:

$$LS_t = \begin{cases} -1 & \text{if } t < t_0 \\ 0 & \text{if } t \geq t_0 \end{cases} \quad (3)$$

Ramps

Ramps (RP) allow for a smooth, linear transition between two time points (t_0 is the start date and t_1 the end date) unlike the abrupt change associated with level shifts are given by:

$$RP_t = \begin{cases} -1 & \text{if } t \leq t_0 \\ (t - t_0)/(t_1 - t_0) - 1 & \text{if } t_0 < t < t_1 \\ 0 & \text{if } t \geq t_1 \end{cases} \quad (4)$$

Temporary Level Shifts

Temporary level shifts (TL) where the level shift is of temporary rather than permanent nature it follows that:

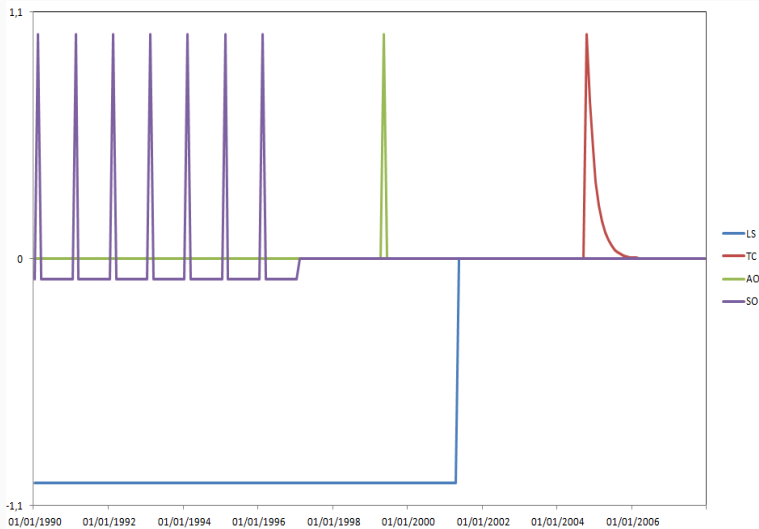
$$TL_t = \begin{cases} 0 & \text{if } t \leq t_0 \\ 1 & \text{if } t_0 < t \leq t_1 \\ 0 & \text{if } t > t_1 \end{cases} \quad (5)$$

Seasonal Outliers

Seasonal outliers (SO) do not affect the level of the time series but the seasonal pattern by allowing for an abrupt increase or decrease of the seasonal component for a specific month ($s = 12$) or quarter ($s = 4$), hence a permanent seasonal break, of which the formula is:

$$SO_t = \begin{cases} 0 & \text{if } t \geq t_0 \\ -1 & \text{if } t < t_0 \text{ and } t \text{ same month/quarter as } t_0 \\ 1/(s-1) & \text{otherwise} \end{cases} \quad (6)$$

Regressors



Characteristics of different types of outliers

| Type of outlier | Component | Durability of impact | Visible in SA data? | Frequency |
|-----------------------|-------------|----------------------|---------------------|-----------|
| Additive outlier | Irregular | Temporary | Yes | Common |
| Temporary change | Irregular | Temporary | Yes | Rare |
| Level shift | Trend-cycle | Permanent | Yes | Common |
| Ramp | Trend-cycle | Permanent | Yes | Rare |
| Temporary level shift | Trend-cycle | Temporary | Yes | Rare |
| Seasonal outlier | Seasonal | Permanent | No | Sometimes |

Calendars

Calendars

Calendars are normally based on astronomical events; and the 2 most important astronomical objects are the Sun and the Moon.

- Solar calendars (Gregorian, Julian) are based on the motion of the Earth around the Sun. The calendar year approximates the “tropical year”; 12 months strongly linked to the seasons (Seasonal effect).
- Lunar calendars (Islamic) are based on the motion of the Moon around the Earth. A month approximates the “synodic month”, the time between 2 “new moons”.
- Solar/lunar calendars (Chinese, Hebrew, Hindu) try to achieve the 2 approximations.

The Various Calendar Effects

Months and Quarters

Most of economic series are recorded “according to the Gregorian calendar” on a monthly or quarterly basis

- But months (quarters) are not really comparable:
 - ☐ Not the same number of days
 - ☐ Not the same number of Mondays, Tuesdays
 - ☐ National public holidays
- And this may have an impact on the observed variable.
 - ☐ Retail sales, Transportation etc.
- Important temporal and spatial differences

Temporal differences

Big differences if we take into account the National specificities.

Example: Number of working days (Monday to Friday) in France by quarter.

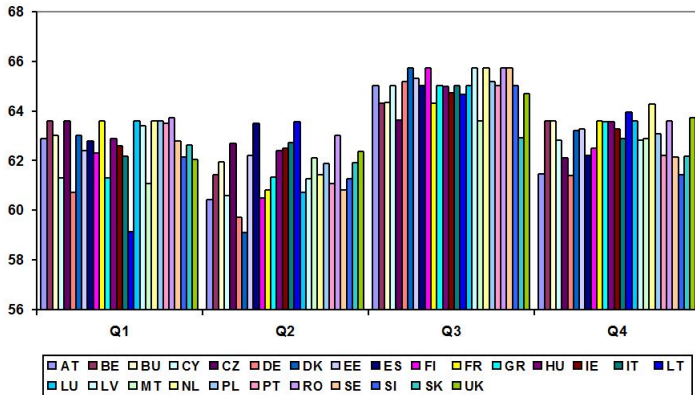
- Compare 2017Q2 to 2016Q2 (annual growth rate)
- Compare 2017Q2 to 2017Q1 (quarterly growth rate)

| Gregorian calendar | | | | | | National calendar | | | | |
|--------------------|----|----|----|----|-------|-------------------|----|----|----|-------|
| Year | Q1 | Q2 | Q3 | Q4 | Total | Q1 | Q2 | Q3 | Q4 | Total |
| 2016 | 65 | 65 | 66 | 65 | 261 | 63 | 63 | 64 | 63 | 253 |
| 2017 | 65 | 65 | 65 | 65 | 260 | 65 | 60 | 63 | 63 | 251 |
| 2018 | 65 | 65 | 65 | 66 | 261 | 64 | 60 | 64 | 64 | 252 |
| 2019 | 64 | 65 | 66 | 66 | 261 | 63 | 60 | 65 | 63 | 251 |
| 2020 | 65 | 65 | 66 | 66 | 262 | 64 | 60 | 65 | 64 | 253 |

Geographical differences

Number of working-days (Monday to Friday) by quarter for EU countries.

A very busy graph \Leftrightarrow large spatial differences.



The various calendar effects

- The seasonality (due to the Gregorian solar calendar)
- Length-of-Month (LOM) effect
- “Day of the week” (Trading-day) effect
 - ☐ Ex.: Retail trade turnover is likely to be more important on Saturdays than on other days.
- Public Holidays
 - ☐ Most of statutory holidays are linked to a date, not to a day of the week (Christmas).
- Moving Holidays
 - ☐ Some holidays “move” across the year (Easter, Ramadan, Chinese New Year, Diwali etc.).
 - ☐ Usually occurs when you mixed 2 calendars.

Detecting “pure” Trading-Day Effects

Detecting “pure” Trading-Day Effects

Hopefully, the Gregorian calendar is quite regular:

- The structure of the week is periodic with period 7;
- The length of the year is of 365 or 366 days. Leap Year?
 - ☐ $(\text{MOD}(\text{year},4)=0 \text{ AND } \text{MOD}(\text{year},100) \neq 0) \text{ OR } \text{MOD}(\text{year},400)=0.$
 - ☐ 97 leap years in a 400-year cycle.
 - ☐ The length of the year is periodic with period 400 years.
- Integer number of weeks in a complete cycle (20871);
- Thus the Gregorian calendar is periodical with period 400 years.
- Detection of a trading-day effect:
 - ☐ Using the spectrum.
 - ☐ Using a Reg-ARIMA model.

Exploring some distributions

- A perpetual calendar can easily be generated
- Easy to compute some distributions, basic statistics and spectra, on 1 or more complete 400-year cycles:
 - ☐ # of Mondays (T, W, T, F, S, S) per month or quarter
 - ☐ # of weekdays per month or quarter
- # of days distributions (and spectra) are very similar (“symmetry” on a cycle) but neither equal nor independent
- Warning: Distributions may be seasonal.

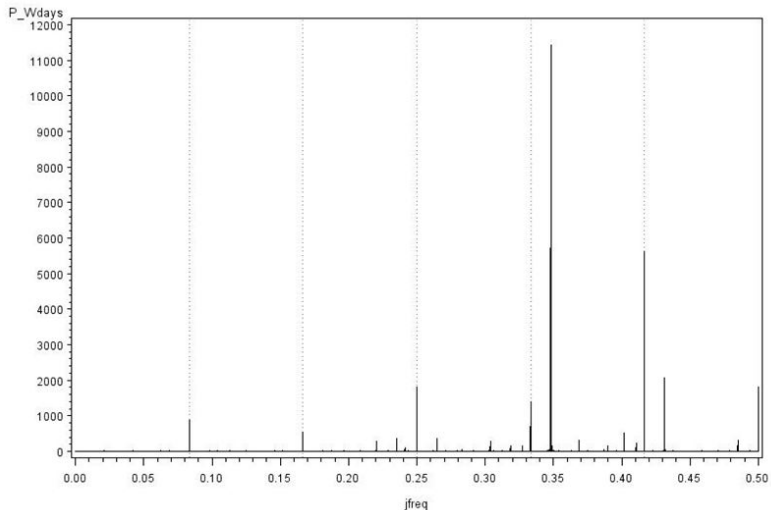
Basic Characteristics: Monthly Data

| | Sunday | Monday | Tuesday | Wednesday | Thursday | Friday | Saturday | Week Days |
|----------------|---------|---------|---------|-----------|----------|---------|----------|-----------|
| Mean | 4.3481 | 4.3481 | 4.3481 | 4.3481 | 4.3481 | 4.3481 | 4.3481 | 21.7406 |
| STD | 0.4764 | 0.4764 | 0.4764 | 0.4764 | 0.4764 | 0.4764 | 0.4764 | 0.9614 |
| N | 33600 | 33600 | 33600 | 33600 | 33600 | 33600 | 33600 | 33600 |
| Corr Sunday | 1 | 0.4115 | -0.1659 | -0.5340 | -0.5340 | -0.1668 | 0.4088 | -0.4902 |
| Corr Monday | 0.4115 | 1 | 0.4088 | -0.1687 | -0.5340 | -0.5340 | -0.1677 | 0.0853 |
| Corr Tuesday | -0.1659 | 0.4088 | 1 | 0.4106 | -0.1650 | -0.5340 | -0.5340 | 0.5552 |
| Corr Wednesday | -0.5340 | -0.1687 | 0.4106 | 1 | 0.4106 | -0.1677 | -0.5340 | 0.7358 |
| Corr Thursday | -0.5340 | -0.5340 | -0.1650 | 0.4106 | 1 | 0.4097 | -0.1659 | 0.5556 |
| Corr Friday | -0.1668 | -0.5340 | -0.5340 | -0.1677 | 0.4097 | 1 | 0.4115 | 0.0862 |
| Corr Saturday | 0.4088 | -0.1677 | -0.5340 | -0.5340 | -0.1659 | 0.4115 | 1 | -0.4907 |
| Corr Week Days | -0.4902 | 0.0853 | 0.5552 | 0.7358 | 0.5556 | 0.0862 | -0.4907 | 1 |

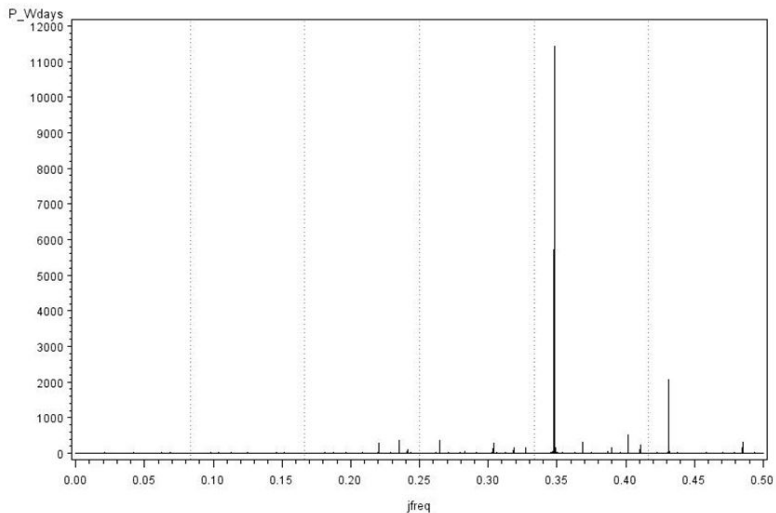
Basic Characteristics: Quarterly Data

| | | Sunday | Monday | Tuesday | Wednesday | Thursday | Friday | Saturday | Week Days |
|------|-----------|---------|---------|---------|-----------|----------|---------|----------|-----------|
| Mean | | 13.0444 | 13.0444 | 13.0444 | 13.0444 | 13.0444 | 13.0444 | 13.0444 | 65.2219 |
| STD | | 0.3101 | 0.3121 | 0.3101 | 0.3121 | 0.3101 | 0.3101 | 0.3101 | 0.6663 |
| N | | 11200 | 11200 | 11200 | 11200 | 11200 | 11200 | 11200 | 11200 |
| Corr | Sunday | 1 | -0.0203 | -0.0205 | -0.0203 | -0.0205 | -0.0205 | -0.0205 | -0.0477 |
| Corr | Monday | -0.0203 | 1 | -0.0203 | -0.0202 | -0.0203 | -0.0203 | -0.0203 | 0.4306 |
| Corr | Tuesday | -0.0205 | -0.0203 | 1 | -0.0203 | -0.0205 | -0.0205 | -0.0205 | 0.4273 |
| Corr | Wednesday | -0.0203 | -0.0202 | -0.0203 | 1 | -0.0203 | -0.0203 | -0.0203 | 0.4306 |
| Corr | Thursday | -0.0205 | -0.0203 | -0.0205 | -0.0203 | 1 | -0.0205 | -0.0205 | 0.4273 |
| Corr | Friday | -0.0205 | -0.0203 | -0.0205 | -0.0203 | -0.0205 | 1 | -0.0205 | 0.4273 |
| Corr | Saturday | -0.0205 | -0.0203 | -0.0205 | -0.0203 | -0.0205 | -0.0205 | 1 | -0.0477 |
| Corr | Week Days | -0.0477 | 0.4306 | 0.4273 | 0.4306 | 0.4273 | 0.4273 | -0.0477 | 1 |

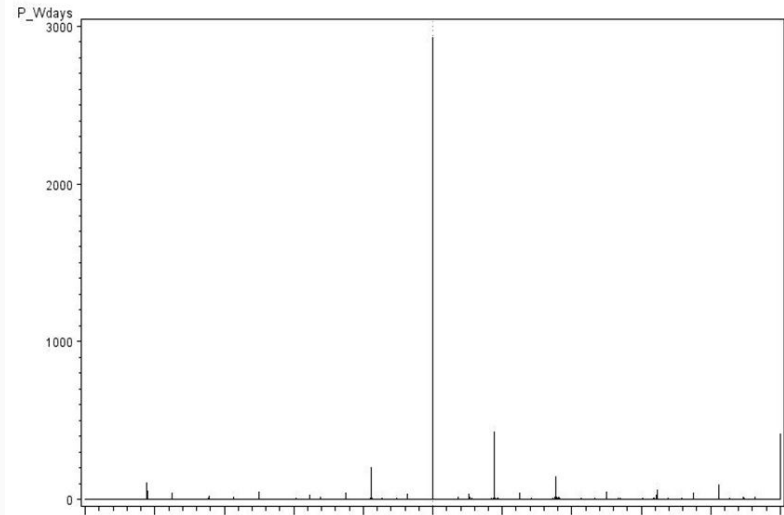
Spectra of the Monthly Weekday Series



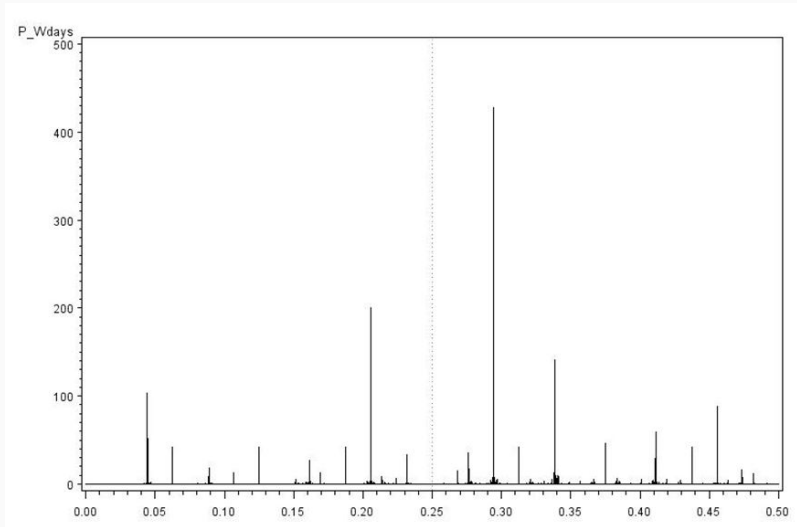
Spectra of the Centered Monthly Weekday Series



Spectra of the Quarterly Weekday Series



Spectra of the Centered Quarterly Weekday Series



Main trading-day frequencies (Monthly)

| | Sunday | Monday | Tuesday | Wednesday | Thursday | Friday | Saturday | Week Days |
|----|----------|----------|----------|-----------|----------|----------|----------|-----------|
| 1 | 0.348125 | 0.348125 | 0.348125 | 0.348125 | 0.348125 | 0.348125 | 0.348125 | 0.348125 |
| 2 | 0.431458 | 0.431458 | 0.431458 | 0.431458 | 0.431458 | 0.431458 | 0.431458 | 0.431458 |
| 3 | 0.220417 | 0.303750 | 0.303750 | 0.303750 | 0.303750 | 0.303750 | 0.303750 | 0.401875 |
| 4 | 0.303750 | 0.220417 | 0.220417 | 0.220417 | 0.220417 | 0.220417 | 0.220417 | 0.264792 |
| 5 | 0.401875 | 0.401875 | 0.401875 | 0.401875 | 0.401875 | 0.401875 | 0.401875 | 0.235208 |
| 6 | 0.264792 | 0.264792 | 0.235208 | 0.264792 | 0.235208 | 0.264792 | 0.235208 | 0.485208 |
| 7 | 0.235208 | 0.235208 | 0.264792 | 0.235208 | 0.264792 | 0.235208 | 0.264792 | 0.368958 |
| 8 | 0.485208 | 0.368958 | 0.485208 | 0.368958 | 0.485208 | 0.485208 | 0.485208 | 0.303750 |
| 9 | 0.368958 | 0.485208 | 0.368958 | 0.485208 | 0.368958 | 0.368958 | 0.368958 | 0.220417 |
| 10 | 0.347917 | 0.347917 | 0.347917 | 0.347917 | 0.410625 | 0.347917 | 0.347917 | 0.347917 |

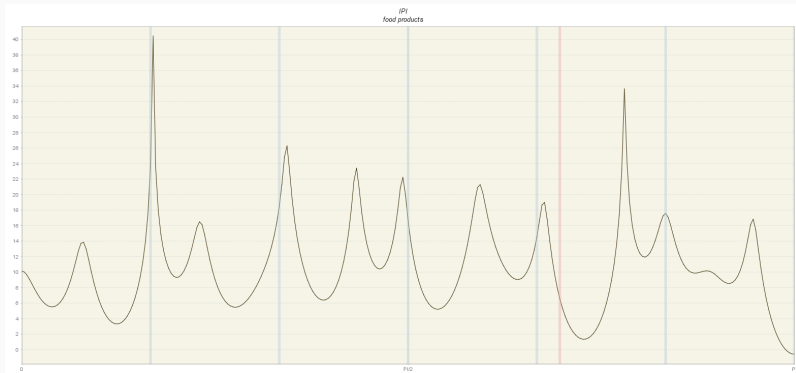
Main trading-day frequencies (Quarterly)

| | Sunday | Monday | Tuesday | Wednesday | Thursday | Friday | Saturday | Week Days |
|----|----------|----------|----------|-----------|----------|----------|----------|-----------|
| 1 | 0.294375 | 0.294375 | 0.294375 | 0.294375 | 0.294375 | 0.294375 | 0.294375 | 0.294375 |
| 2 | 0.338750 | 0.338750 | 0.338750 | 0.338750 | 0.338750 | 0.338750 | 0.338750 | 0.205625 |
| 3 | 0.205625 | 0.205625 | 0.205625 | 0.205625 | 0.205625 | 0.205625 | 0.205625 | 0.338750 |
| 4 | 0.383125 | 0.411250 | 0.383125 | 0.411250 | 0.411250 | 0.383125 | 0.411250 | 0.044375 |
| 5 | 0.411250 | 0.383125 | 0.411250 | 0.383125 | 0.383125 | 0.411250 | 0.383125 | 0.455625 |
| 6 | 0.044375 | 0.044375 | 0.044375 | 0.044375 | 0.044375 | 0.044375 | 0.044375 | 0.411250 |
| 7 | 0.366875 | 0.455625 | 0.366875 | 0.455625 | 0.366875 | 0.366875 | 0.366875 | 0.375000 |
| 8 | 0.455625 | 0.366875 | 0.455625 | 0.366875 | 0.455625 | 0.455625 | 0.455625 | 0.125000 |
| 9 | 0.320625 | 0.276250 | 0.276250 | 0.276250 | 0.320625 | 0.320625 | 0.320625 | 0.062500 |
| 10 | 0.276250 | 0.320625 | 0.161250 | 0.320625 | 0.276250 | 0.276250 | 0.276250 | 0.187500 |

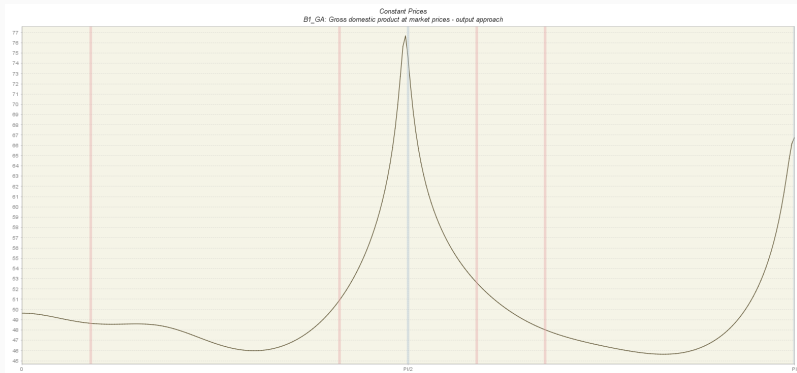
Checking for a trading-day effect

- X-13Arima-Seats displays the spectrum of the (transformed) raw data, the SA series and the irregular components.
 - ☐ Monthly series: X12 checks for the 2 main frequencies
 - ☐ Quarterly series: no more checking.
- Visual “6-star” checking
- JDemetra+: see examples hereafter;
- Testing significance of a spectral peak?
 - ☐ McElroy and Holan (2005), A Nonparametric Test for Assessing Spectral Peaks, USCB, Research Report.
 - ☐ Canova-Hansen test

Example of Autoregressive Spectrum in JD+ (Monthly)



Example of Autoregressive Spectrum in JD+ (Quarterly)



Did you know that

2 pages from a Swedish almanac from 1712.

- Top left corner: "Februarius", At the top of the right page: XXX
- And at the bottom weather forecasts !!! (Snöö = snow)



Moving Holidays

Moving Holidays

- Basically due to the use of 2 different calendars at the same time
- Christian Easter
 - Easter Sunday is the first Sunday after the first full moon after vernal equinox.
 - Linked to the astronomic calendar
- Ramadan, Aid-el-Fitr, Aid-el-Adha etc.
 - Linked to the Islamic calendar (lunar)
- Chinese New Year, Diwali, Pesach etc.
 - Lunar/solar calendars

An Example: Easter

- Catholic Easter
 - Sunday Easter can fall between March 22 (last time: 1818, next time: 2285) and April 25 (last time: 1943, next time: 2038) i.e. during the first or second quarter. WARNING: Much more in April!!!
- Orthodox Easter (Greece, Romania)
 - Between April 4 (last time: 2010, next time after 7000) and May 9 (last time before 1600, next time: 2173).
- In Christian countries, Easter affects the production of many economic sectors:
 - Statutory holidays (transportation, Tourism sector: hotel)
 - Consumption habits (chocolate, flowers, lamb meat etc.)
- Very similar in principle to Diwali (October-November), Chinese New Year (January-February) etc.

Calculating Easter Dates

Algorithms exist. An example that gives correct dates from 1532.

$$G = \text{year} \bmod 19$$

For the Julian calendar:

$$I = (19G + 15) \bmod 30$$

$$J = (\text{year} + \frac{\text{year}}{4} + I) \bmod 7$$

For the Gregorian calendar:

$$C = \frac{\text{year}}{100}$$

$$H = (C - \frac{C}{4} - \frac{8C+13}{25} + 19G + 15) \bmod 30$$

$$I = H - \frac{H}{28} (1 - \frac{29}{H+1} \frac{21-G}{11})$$

$$J = (\text{year} + \frac{\text{year}}{4} + I + 2 - C + \frac{C}{4}) \bmod 7$$

Thereafter, for both calendars:

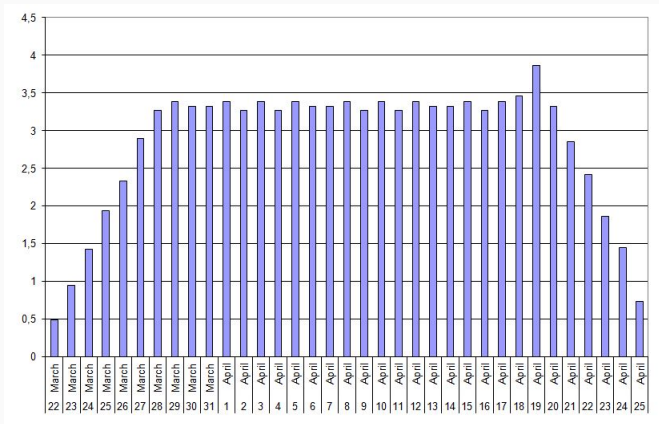
$$L = I - J$$

$$\text{EasterMonth} = 3 + \frac{L+40}{44}$$

$$\text{EasterDay} = L + 28 - 31 \frac{\text{EasterMonth}}{4}$$

Catholic Easter Dates

Much more in April than in March \Rightarrow seasonal effect.



Example: Calculating Islamic Festival Dates

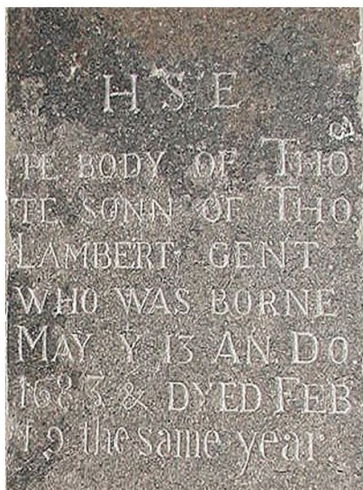
- One year has 354 or 355 days; 12 months with 29 or 30 days;
- In “Leap Years”, the last month (Jumada al Oula) counts 30 days;
- The years follow a 30-year cycle;
- Several “rules” for the “leap years”:
 - 2, 5, 7, 10, 13, 15, 18, 21, 24, 26, et 29 (Kuwait algorithm)
 - 2, 5, 7, 10, 13, 16, 18, 21, 24, 26, et 29 (common version)
 - 2, 5, 8, 10, 13, 16, 19, 21, 24, 27, et 29 (Indian tables)
 - 2, 5, 8, 11, 13, 16, 19, 21, 24, 27, et 30
- On this 30-year cycle the year average length is very close to the length of the lunar year (<35s).
- Ras-El-Am = MDY(1,25,1583) and you follow the cycle
- But you might have to adapt this calendar to the real observation of the Moon.

Building a National Calendar

National Calendars

- Building a National calendar is not easy, it could even become a nightmare very quickly.
- A calendar is alive:
 - ☐ Public or bank holidays appear and disappear;
 - ☐ Substitute days (UK, Canada);
 - ☐ Special days (UK: Queen's Jubilee)
- Exception is the Rule!
 - ☐ Provincial specificities (Germany, Canada, India) etc.
- How to construct a European calendar?
 - ☐ Weighting problem
- Is it really worth it? If it is complex, keep it simple!

Did you know that



- The tombstone tells about a boy who was born on 13 May 1683 and died on 19 February of the same year;
- In 1683 in England, new year's day was 25 March;
- Indeed, would not be Spring a much better date to begin a new year?

Modeling the Trading-Day Effect

A Basic Regression Model

A usual additive model for SA is:

$$X_t = TC_t + S_t + O_t + TD_t + MH_t + I_t,$$

Where all the components are orthogonal and are non-seasonal (except S_t of course).

A basic model for trading-day could be:

$$X_t = \sum_{i=1}^{i=7} \alpha_i N_{it} + \epsilon_t,$$

Where:

- N_{it} = # of Mondays (i=1), ..., Sundays (i=7) in month t.
- α_i is the effect of day i. (constant on the period)
- ϵ_t is a white noise (X11, X12) or an ARIMA model (X12, TS)

Improving the Model

The (Reg-ARIMA) model presents 3 problems:

- The regressor covariance matrix is ill-conditioned and its inverse is unstable \implies estimates are also unstable.
- TD and seasonality are not independent
- We also would like that the TD effect is equal to 0 over a week

To solve most of these problems, the most popular model (X12, TS) uses contrasts and do the estimation on the raw data:

$$X_t = \beta_0 LY_t + \sum_{j=1}^{j=6} \beta_j (N_{jt} - N_{7t}) + \epsilon_t \quad \text{and} \quad \hat{\beta}_7 = - \sum_{j=1}^{j=6} \hat{\beta}_j$$

Stability of the Estimates

The inverse covariance matrix of the contrasts is more stable.

Ratio[$\min(\lambda)/\max(\lambda)$], computed over the 400-year cycle; all series of length 5 to 25 years.

| Length (in years) | Gregorian Calendar | | | French National Calendar | |
|----------------------|--------------------|-------------|-------------|--------------------------|-------------|
| | 7 regressors | 6 contrasts | 2 contrasts | 6 contrasts | 2 contrasts |
| 5 | 0.0005 | 0.0321 | 0.0243 | 0.1773 | 0.1444 |
| 10 | 0.0005 | 0.0329 | 0.0243 | 0.1935 | 0.1442 |
| 15 | 0.0005 | 0.0331 | 0.0243 | 0.1975 | 0.1440 |
| 20 | 0.0005 | 0.0331 | 0.0243 | 0.1997 | 0.1437 |
| All | 0.0005 | 0.0328 | 0.0243 | 0.1919 | 0.1441 |

The ratio is multiplied by about 6 which shows a much better stability.

Seasonality

Contrasts remove a large part of TD seasonality:

- The seasonal part is the average number of a specific day over the 400-year cycle;
- The averages (see Table) are quite close and taking the difference reduces the seasonality.

| Month | Monday | Tuesday | Wednesday | Thursday | Friday | Saturday | Sunday |
|-------|--------|---------|-----------|----------|--------|----------|--------|
| 1 | 4.4300 | 4.4250 | 4.4300 | 4.4275 | 4.4300 | 4.4300 | 4.4275 |
| 2 | 4.0325 | 4.0375 | 4.0325 | 4.0375 | 4.0325 | 4.0350 | 4.0350 |
| 3 | 4.4300 | 4.4275 | 4.4300 | 4.4250 | 4.4300 | 4.4275 | 4.4300 |
| 4 | 4.2850 | 4.2875 | 4.2850 | 4.2875 | 4.2850 | 4.2850 | 4.2850 |
| 5 | 4.4300 | 4.4250 | 4.4300 | 4.4275 | 4.4300 | 4.4300 | 4.4275 |
| 6 | 4.2850 | 4.2875 | 4.2850 | 4.2850 | 4.2850 | 4.2850 | 4.2875 |
| 7 | 4.4300 | 4.4275 | 4.4300 | 4.4300 | 4.4275 | 4.4300 | 4.4250 |
| 8 | 4.4275 | 4.4300 | 4.4250 | 4.4300 | 4.4275 | 4.4300 | 4.4300 |
| 9 | 4.2875 | 4.2850 | 4.2875 | 4.2850 | 4.2850 | 4.2850 | 4.2850 |
| 10 | 4.4250 | 4.4300 | 4.4275 | 4.4300 | 4.4300 | 4.4275 | 4.4300 |
| 11 | 4.2875 | 4.2850 | 4.2850 | 4.2850 | 4.2850 | 4.2875 | 4.2850 |
| 12 | 4.4275 | 4.4300 | 4.4300 | 4.4275 | 4.4300 | 4.4250 | 4.4300 |

- The 7-regressor model:

$$X_t = \beta_0 LY_t + \sum_{j=1}^{j=6} \beta_j (N_{jt} - N_{7t}) + \epsilon_t \quad \text{and} \quad \hat{\beta}_7 = - \sum_{j=1}^{j=6} \hat{\beta}_j$$

- The 2-regressor model (weekdays vs. weekend)

$$\beta_6 = \beta_7 \quad \text{and} \quad \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5$$
$$\sum_{j=1}^{j=7} \beta_j = 0 \Rightarrow \beta_6 = \beta_7 = -\frac{5}{2}\beta_1$$

$$X_t = \beta_0 LY_t + \beta_1 \left[\sum_{j=1}^{j=5} N_{jt} - \frac{5}{2}(N_{6t} + N_{7t}) \right] + \epsilon_t$$

- An important remark on the Leap Year effect.

Limits of these (too) simple models

- Do not take into account National specificities;
- 2 sets of regressors only : 1 or 6 contrasts. Reality might be more complex (Saturday for Retail Trade)
- Effects are constant over time. But things and laws change (Retail Trade and Sundays)
- More flexibility is needed to construct and adapt regressors.
 - ☐ Groupings;
 - ☐ Take then care of potential seasonality.

Taking into account the National calendar

- Complete Model with 15 variables: distinction between Monday off and Monday in.

$$TD_t = \beta_0 LY_t + \sum_{j=1}^{j=14} \beta_j N_{jt}$$

- Where N_{jt} are the # of Mondays in, Tuesdays in, ..., Sundays in, Mondays off, Tuesdays off, ..., Sundays off.
- But not very parsimonious: hypothesis on the various days.

Example

- All days off are equivalent to a Sunday.

$$\beta_7 = \beta_8 = \beta_9 = \cdots = \beta_{13} = \beta_{14}$$

- And we obtain the following model:

$$TD_t = \beta_0 LY_t + \sum_{j=1}^{j=6} \beta_j \left[N_{jt} - \frac{1}{8} \left(\sum_{j=7}^{j=14} N_{jt} \right) \right]$$

- But contrasts do not remove the seasonality anymore: Need to seasonally adjust the regressors using the 400-year periodicity of the calendar or the periodicity of the moving holiday (Easter, Ramadan, Chinese New Year, Diwali, Pesach etc.).

Modeling moving holidays

Modeling moving holidays

Moving Holidays might have 2 different effects:

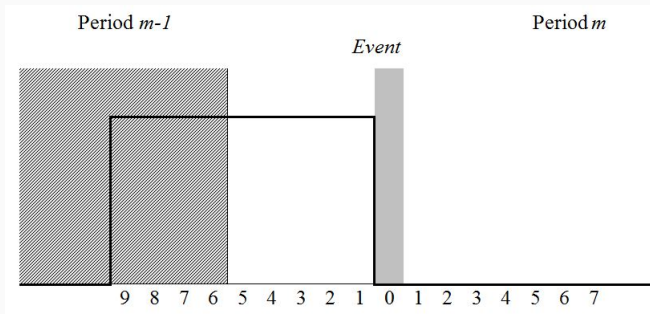
- Days off are often associated to these holidays to celebrate the event. This effect is accounted for in the Trading-Day effect;
- They might have an impact on the economy. For examples, members of a family might travel to celebrate the event together, the prices of some commodities might increase before the event etc.

This is the second effect we want to capture here using an “impact model”.

An Example: Easter

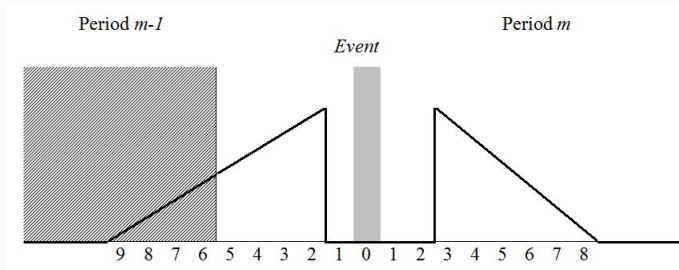
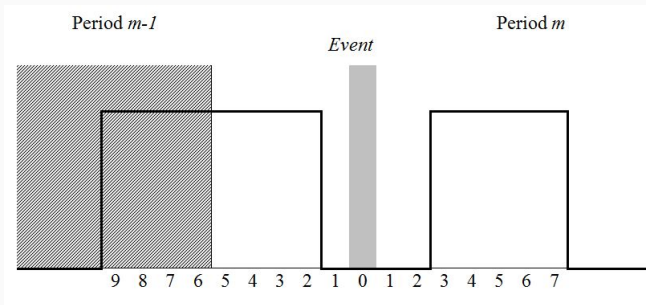
- We will take the example of Easter; very similar to Pesach, Diwali in principle.
- Easter[w], the most popular model already implemented in JDemetra+:
 - ☐ The activity is affected the w days before Easter Sunday;
 - ☐ The activity level remains constant on the w -day period.
 - ☐ For example, Catholic Easter can move between March 22 and April 25. Therefore, the 2 months (as well as Q1 and Q2) can be impacted according to the year.

The Easter_[w] Model



- If i denotes the year, j the month. Let us note $1 \leq w \leq 25$ and n_{ij} the number of the w days falling in month j of year i .
- The regressor associated to this model is: $X_{ij} = n_{ij}/w$.

More General Models



Reminders on Regression Models

If you want to estimate something, you need to observe it!!!!

- The Leap Year effect
 - ☐ Of course one more day impacts production, turnover etc.
 - ☐ What if your series starts in 2015?
- Orthodox Easter (between April 4 and May 9) and quarterly time series
 - ☐ Once again, a very important festival that might impact the economy.
 - ☐ Good Friday, Easter Monday, Whit Monday etc. Trading-Day or Seasonal component?
 - ☐ And what can we say on the impact model?

| Julian Easter | Differences | 6 days before (TS) | 8 days before (X13) |
|---------------|-------------|--------------------|---------------------|
| 04/04/2010 | 4 | 29/03/2010 | 27/03/2010 |
| 08/04/2007 | 8 | 02/04/2007 | 31/03/2007 |
| 08/04/2018 | 8 | 02/04/2018 | 31/03/2018 |
| 08/04/2029 | 8 | 02/04/2029 | 31/03/2029 |
| 11/04/1999 | 11 | 05/04/1999 | 03/04/1999 |
| 11/04/2004 | 11 | 05/04/2004 | 03/04/2004 |
| 12/04/2015 | 12 | 06/04/2015 | 04/04/2015 |
| 12/04/2026 | 12 | 06/04/2026 | 04/04/2026 |