universität
innsbruck

Time Series Analysis

Time Series Regression Models

Time Series Regression Models

Contents

Contents

Time Series Regression Models

Motivation

Motivation

So far:

- Analysis of correlation structure of stationary time series.
- Nonstationarity captured by suitable differencing, i.e., focus on (seasonal and nonseasonal) stochastic trends.

However:

- Many time series exhibit other types of nonstationarities.
- Of particular interest are changes in the mean function that can be explained by regressors.
- Regressors might be known deterministic functions or other time series (i.e., stochastic).

Remark: Regressions with stochastic regressors may require special attention, especially when series are nonstationary. This may lead to spurious correlations etc.

Motivation

Idea: Combine regression model for the mean with ARIMA-type errors.

$$Y_t = X_t^\top \beta + Z_t$$

where $\{X_t\}$ is the (possibly multivariate) series of regressors with coefficients β and $\{Z_t\}$ is an ARIMA process.

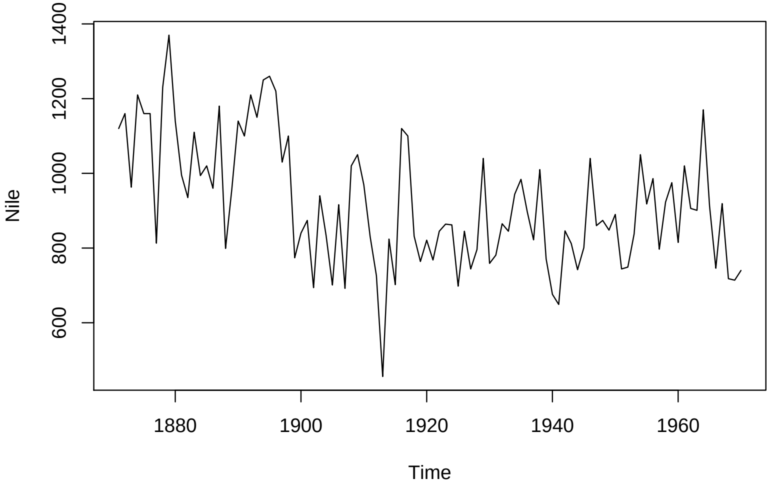
Jargon: ARIMAX model.

In R: `arima()` with `xreg` argument, e.g.,

```
R> arima(y, order = c(p, d, q), seasonal = c(P, D, Q), xreg = x)
```

Illustration: Annual flow of river Nile at Aswan from 1871 to 1970. Annual discharges in 10^8m^3.

Motivation

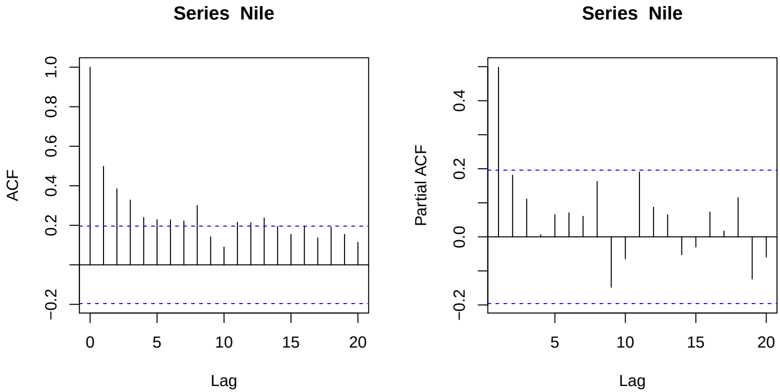


Motivation

Time series is not stationary but an integrated $I(1)$ model is also (weakly) rejected.

```
R> data("Nile", package = "datasets")
R> kpss.test(Nile)

        KPSS Test for Level Stationarity

data:  Nile
KPSS Level = 0.97, Truncation lag parameter = 4, p-value
= 0.01

R> adf.test(Nile)

        Augmented Dickey-Fuller Test

data:  Nile
Dickey-Fuller = -3.4, Lag order = 4, p-value = 0.06
alternative hypothesis: stationary
```

Motivation



Series Nile
ACF
1.0
0.8
0.6
0.4
0.2
−0.2
0
5
10
15
20
Lag
Series Nile
Partial ACF
0.4
0.2
0.0
−0.2
5
10
15
20
Lag

Motivation

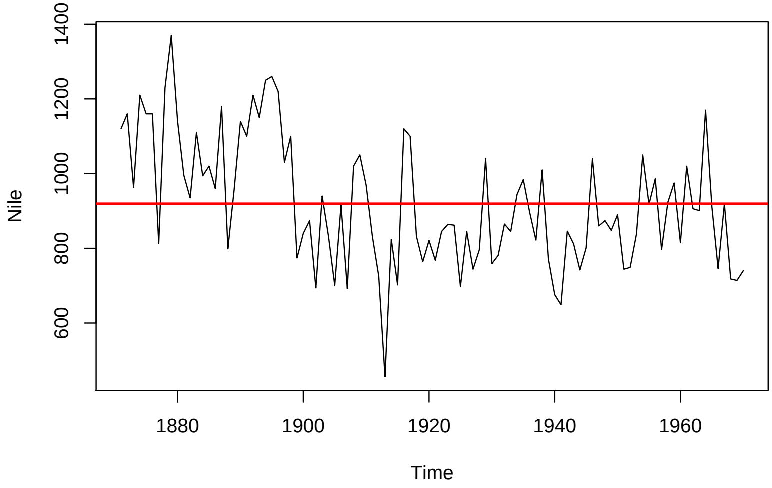
ACF/PACF might suggest an AR(1) model.

```
R> ar1 <- arima(Nile, order = c(1, 0, 0))
R> ar1

Call:
arima(x = Nile, order = c(1, 0, 0))

Coefficients:
        ar1  intercept
      0.506     919.57
s.e.  0.087      29.14

sigma^2 estimated as 21125:  log likelihood = -640,  aic = 1286
```

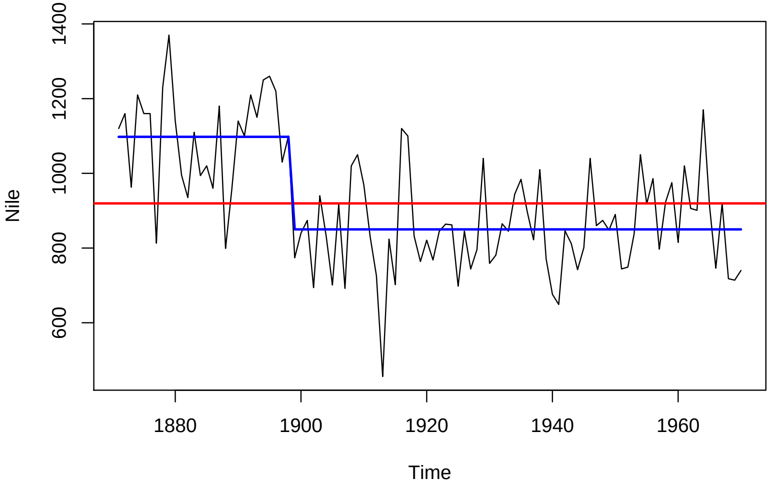
Also a BIC-based search would yield a similar ARIMA(1, 0, 1) model.

```
R> auto.arima(Nile, ic = "bic", approx = FALSE, stepwise = FALSE,
+    stationary = TRUE)
```

Motivation



Motivation



Motivation

Problem: The mean in the time series is neither constant nor following a (stochastic) trend.

Source: Aswan dam built in 1898.

Idea: Allow for a level shift by adding regressor $X_t = I(t > 1898)$, i.e., an indicator for the time period after introduction of the dam.

In R:

```
R> x <- as.numeric(time(Nile) > 1898)
R> ar1x <- arima(Nile, order = c(1, 0, 0), xreg = x)
R> coeftest(ar1x)
```

```
z test of coefficients:

           Estimate Std. Error z value Pr(>|z|)
ar1          0.1596     0.0986    1.62     0.11
intercept 1098.5169    27.8553   39.44  < 2e-16
x         -249.0750    32.8037   -7.59  3.1e-14
```

Motivation

```
R> ar1x
Call:
arima(x = Nile, order = c(1, 0, 0), xreg = x)

Coefficients:
        ar1  intercept       x
      0.160    1098.52  -249.1
s.e.  0.099      27.86    32.8

sigma^2 estimated as 15563:  log likelihood = -624.5,  aic = 1257
```

This suggests that the AR term is, in fact, not needed when the level change is accounted for.

```
R> ar0x <- arima(Nile, order = c(0, 0, 0), xreg = x)
```

Motivation

```
R> ar0x
Call:
arima(x = Nile, order = c(0, 0, 0), xreg = x)

Coefficients:
      intercept        x
        1097.75  -247.78
s.e.      23.89    28.15

sigma^2 estimated as 15975:  log likelihood = -625.8,  aic = 1258
```

This is also found in a BIC-based search which also selects an ARIMA(0, 0, 0) model.

```
R> auto.arima(Nile, xreg = x, ic = "bic", approx = FALSE,
+    stepwise = FALSE, stationary = TRUE)
```

Time Series Regression Models

Intervention Analysis

Intervention analysis

Idea: Mean function of a series may change after an intervention at time T.

Illustration: Drop in annual Nile flows after Aswan dam was built.

More generally: Changes in mean may not always be abrupt shifts, i.e., step-shaped. Formalized by Box & Tiao (1975) as *intervention analysis*.

Basic building blocks: Step function $S_t^{(T)}$ and pulse function $P_t^{(T)}$ for intervention at time T.

$$\begin{aligned} S_t^{(T)} &= I(t \geq T) \\ P_t^{(T)} &= S_t^{(T)} - S_{t-1}^{(T)} \end{aligned}$$

Based on these various conceivable mean functions m_t can be built.

Intervention analysis

Examples: Step response interventions.

- Permanent shift.

$$m_t = \omega S_t^{(T)}$$

- Permanent shift after a (known) delay of d time units.

$$m_t = \omega S_{t-d}^{(T)}$$

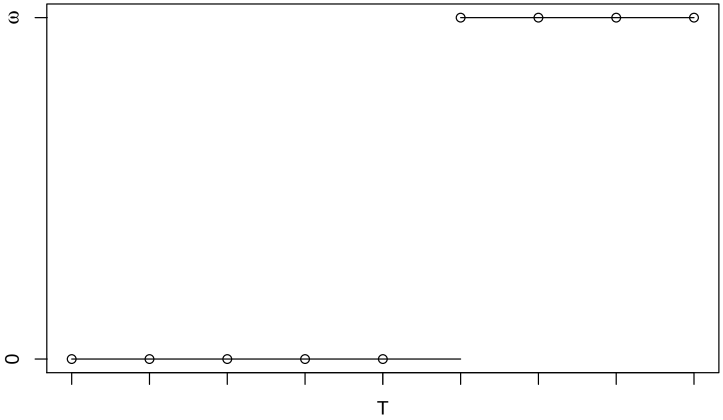
- Gradual shift with AR(1)-type structure and $0 < \delta < 1$.

$$\begin{aligned} m_t &= \delta m_{t-1} + \omega S_{t-1}^{(T)} \\ &= \omega \frac{1-\delta^{t-T}}{1-\delta} S_t^{(T)} \end{aligned}$$

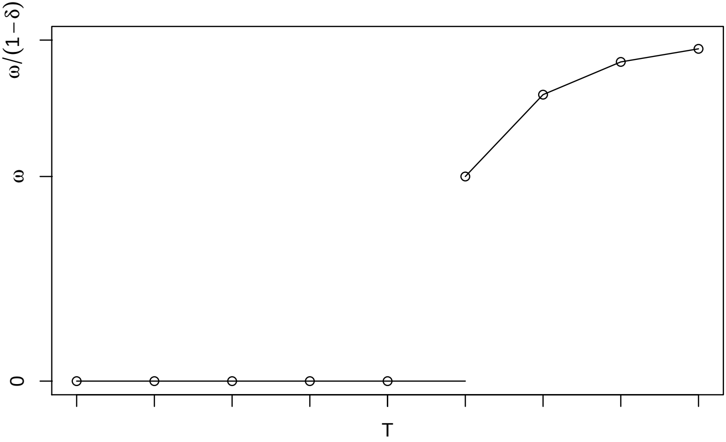
- Trend.

$$\begin{aligned} m_t &= m_{t-1} + \omega S_{t-1}^{(T)} \\ &= \omega(t-T) S_t^{(T)} \end{aligned}$$

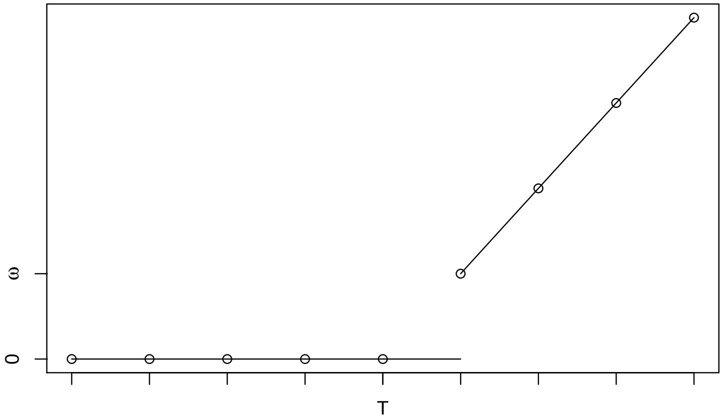
Intervention analysis



Intervention analysis



Intervention analysis



Intervention analysis

Examples: Pulse response interventions.

- Additive outlier.

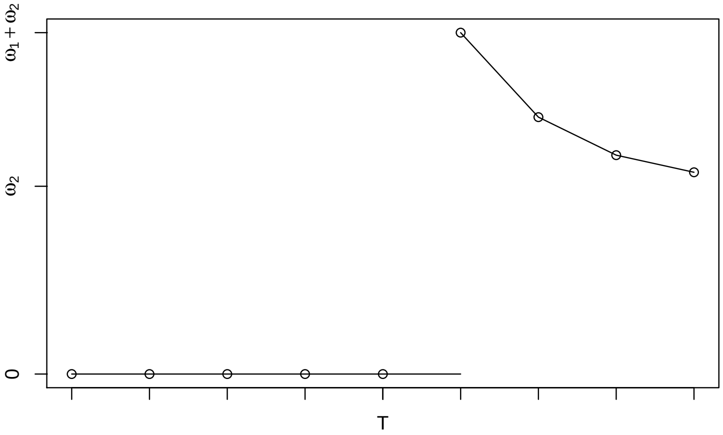
$$m_t \;=\; \omega P_t^{(T)}$$

- Gradual decay with AR(1)-type structure and $0 < \delta < 1$.

$$\begin{aligned} m_t &= \delta m_{t-1} \;+\; \omega P_{t-1}^{(T)} \\ &= \omega \delta^{T-t} S_t^{(T)} \end{aligned}$$

- Combinations of these, e.g., pulse response to $\omega_1 + \omega_2$ and then decay with rate δ to another level ω_2.
- Many further variations ...

Intervention analysis



Intervention analysis

Main problem:

- Timing of intervention T and potential delay d have to be known in advance.
- In many applications, it is hard to argue that these timings are really given endogenously.

Time Series Regression Models

Structural Changes

Structural changes

Goal: Framework for detection of changes in parameters of a model.

Typically: Most theory/applications for changes in coefficients β_t of linear regression models

$$Y_t = X_t^\top \beta_t + Z_t$$

Remarks:

- Note that in the notation the coefficients β_t are allowed to be time-varying.
- Most literature developed for OLS estimation of coefficients.
- Thus, $\{Z_t\}$ is either assumed to be white noise (so that standard inference can be used) or only weakly dependent (so that inference can be corrected "as usual").

Structural changes

Null hypothesis: *Structural stability* or *parameter stability*, i.e., the regression coefficients are constant/stable over time.

$$H_0 : \beta_t \ = \ \beta_0 \qquad (t = 1, \dots, n)$$

so that standard assumptions for OLS regression hold.

Alternative hypothesis: *Structural change* or *parameter instability*, i.e., coefficients somehow vary over time.

$$H_1 : \beta_t \ \neq \ \beta_0 \qquad \text{for at least one } t$$

In many applications, a single shift alternative is of interest, i.e., only a single *structural break* of unknown timing τ

$$H_1^* : \beta_t \ = \ \begin{cases} \beta^{(A)} & \text{for } t \leq \tau \\ \beta^{(B)} & \text{for } t > \tau \end{cases}$$

τ is also known as *breakpoint*, *changepoint*, *cutpoint*, ...

Structural changes

Remarks:

- Many more patterns of structural changes are conceivable.
  - Multiple abrupt shifts.
  - Random walks $\beta_t = \beta_{t-1} + \varepsilon_t$.
  - Gradual shifts and exponential decays.
  - ...
- Hence, the alternative is vast and no single test with “optimal” power against all patterns exists.
- Even for the rather narrow alternative H_1^*, optimal tests only exist in the rather weak sense that no other test uniformly dominates them.
- Therefore, it is important that the tests can be easily interpreted and can convey information about the pattern of change under the alternative.

Structural changes: OLS-based CUSUM

Idea:

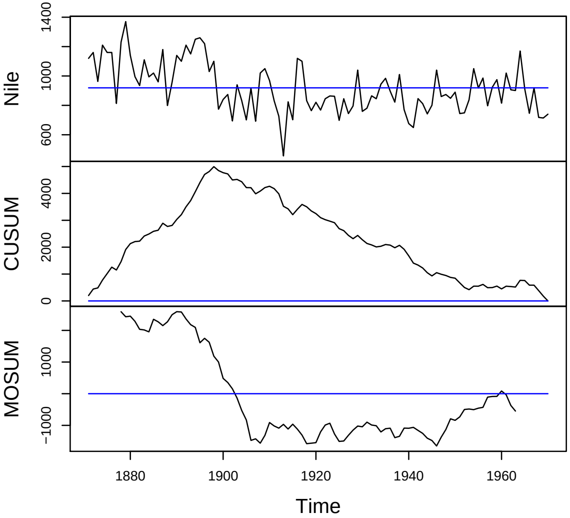
- Estimate coefficients $\hat{\beta}$, assuming that H_0 holds, i.e., parameters are stable throughout the sample period.
- Compute model deviations, e.g., OLS residuals $\widehat{Z}_t = Y_t - X_t^\top \hat{\beta}$.
- Assess whether there are systematic deviations over time by computing cumulative sums (CUSUMs) or moving sums (MOSUMs).

In R: `cumsum(residuals(...))` and `rollapply(residuals(...), k, sum)`.

Remark: MOSUM at time t is difference of CUSUM at t and $t - k$. Hence, focus on CUSUMs first.

Illustration: Annual flows of river Nile.

Structural changes: OLS-based CUSUM



Structural changes: OLS-based CUSUM

More formally: Consider *empirical fluctuation process* of suitably scaled cumulative sums of OLS residuals.

$$efp(z) = \frac{1}{\sqrt{n}\,\hat{\gamma}_0} \sum_{t=1}^{\lfloor zn \rfloor} \widehat{Z}_t \qquad (0 \le z \le 1)$$

Significance test: Aggregate fluctuation process to a scalar test statistic, e.g.,

$$\sup_{0 \le z \le 1} |efp(z)|$$

Question: What is the limiting distribution under the null hypothesis H_0?

Structural changes: OLS-based CUSUM

Excursion: Functional central limit theorems. Certain empirical processes (in discrete time) can be shown to converge to limiting stochastic processes (in continuous time) for $n \to \infty$.

Example: Random walk for white noise e_t.

$$\frac{1}{\sqrt{n}\;\sigma_e} \sum_{t=1}^{\lfloor zn \rfloor} e_t \;\stackrel{d}{\to}\; W(z) \qquad (0 \le z \le 1)$$

where $W(z)$ is a continuous process with Gaussian paths and

$$\begin{aligned} \mathsf{E}\{W(z)\} &= 0 \\ \mathsf{Var}\{W(z)\} &= z \\ \mathsf{Cov}\{W(v), W(z)\} &= \min(v, z) \end{aligned}$$

$W(z)$ is called *Brownian motion* or *Wiener process.*

Structural changes: OLS-based CUSUM

Example: Random walk for mean-adjusted white noise $e_t - \bar{e}$.

$$
\begin{aligned}
\frac{1}{\sqrt{n}\,\sigma_e} \sum_{t=1}^{\lfloor zn \rfloor} e_t - \bar{e} &= \frac{1}{\sqrt{n}\,\sigma_e} \sum_{t=1}^{\lfloor zn \rfloor} \left(e_t - \frac{1}{n} \sum_{t=1}^{n} e_t \right) \\
&= \frac{1}{\sqrt{n}\,\sigma_e} \left(\sum_{t=1}^{\lfloor zn \rfloor} e_t - \frac{\lfloor zn \rfloor}{n} \sum_{t=1}^{n} e_t \right) \\
&\stackrel{d}{\to} W(z) \;-\; zW(1)
\end{aligned}
$$

where $B(z) = W(z) - zW(1)$ is called *Brownian bridge* or *tied-down Brownian motion*. It also has Gaussian paths with

$$
\begin{aligned}
\mathrm{E}\{B(z)\} &= 0 \\
\mathrm{Var}\{B(z)\} &= z(1-z) \\
\mathrm{Cov}\{B(v), B(z)\} &= \min(v, z)(1 - \max(v, z))
\end{aligned}
$$

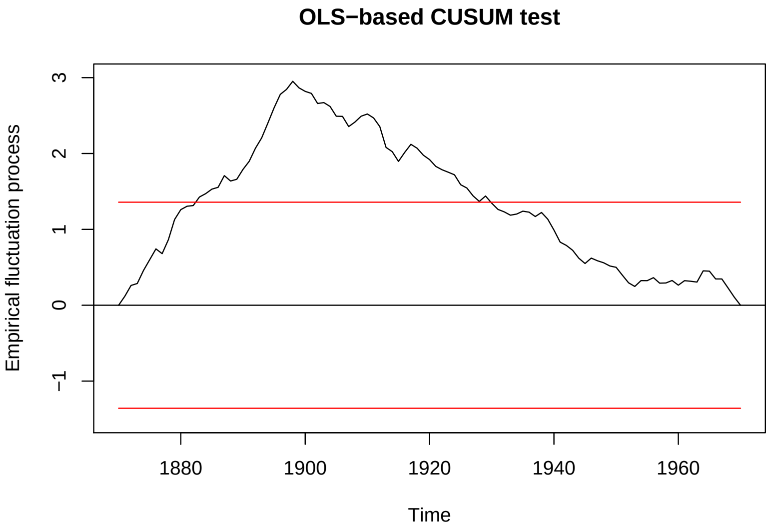
Structural changes: OLS-based CUSUM

Answer:

- Under H_0 all parameters (coefficients and variance) can be consistently estimated and OLS residuals essentially behave like mean-adjusted white noise.
- Hence, it can be shown that $efp(z) \stackrel{d}{\to} B(z)$.
- Moreover, for scalar functional $\lambda(\cdot)$: $\lambda(efp(z)) \stackrel{d}{\to} \lambda(B(z))$.
- Thus
$$\sup_{0\leq z\leq 1} |efp(z)| \stackrel{d}{\to} \sup_{0\leq z\leq 1} |B(z)|$$
- From this limiting distribution p values and critical values can be computed. It can be shown that

$$\mathrm{P}\left(\sup_{0\leq z\leq 1} |B(z)| > b\right) = 2\sum_{j=1}^{\infty}(-1)^{j+1}\exp(-2j^2b^2)$$

Structural changes: OLS-based CUSUM



Structural changes: OLS-based CUSUM

In R: `efp()` in *strucchange*.

OLS-based CUSUM process is computed first.

```
R> ocus <- efp(Nile ~ 1, type = "OLS-CUSUM")
```

The associated test can be performed graphically (using 5% critical values by default).

```
R> plot(ocus)
```

For a classical significance test with *p* value the function `sctest()` can be used.

```
R> sctest(ocus)

        OLS-based CUSUM test

data:  ocus
S0 = 3, p-value = 5e-08
```

Structural changes: Fluctuation tests

Fluctuation tests: Employ similar ideas for other types of empirical fluctuation processes.

- Assess stability by capturing fluctuation in partial sums of
  - residuals (OLS or recursive),
  - model scores (empirical estimating functions), or
  - parameter estimates (recursive or rolling).
- Idea: Under null hypothesis of parameter stability, resulting fluctuation processes exhibit limited fluctuation, under alternative of structural change, fluctuation is generally increased.
- Evidence for structural change if empirical fluctuation process crosses boundary that corresponding limiting process crosses only with probability α.
- Pattern of fluctuation typically conveys information about number and location of structural changes.

Structural changes: F tests

Tests based on F statistics:

- Designed to have good power for single-shift alternative H_1^* (with unknown timing).
- Basic idea is to compute an F statistic (or Chow statistic) for each conceivable breakpoint τ in given interval:

$$F(\tau) \; = \; \frac{\sum_{t=1}^{n} \widehat{Z}_t^2 - \sum_{t=1}^{n} \widetilde{Z}_t(\tau)^2}{\left(\sum_{t=1}^{n} \widetilde{Z}_t(\tau)^2\right) \; / \; (n - 2k)}$$

  where k is the number of regressors in X_t and $\widetilde{Z}_t(\tau)$ are the OLS residuals under H_1^* with breakpoint τ.
- Reject the null hypothesis of structural stability if
  - any of these statistics (sup$F$ test)
  - some other functional (mean$F$, exp$F$)

  exceeds critical value.

Structural changes: *F* tests

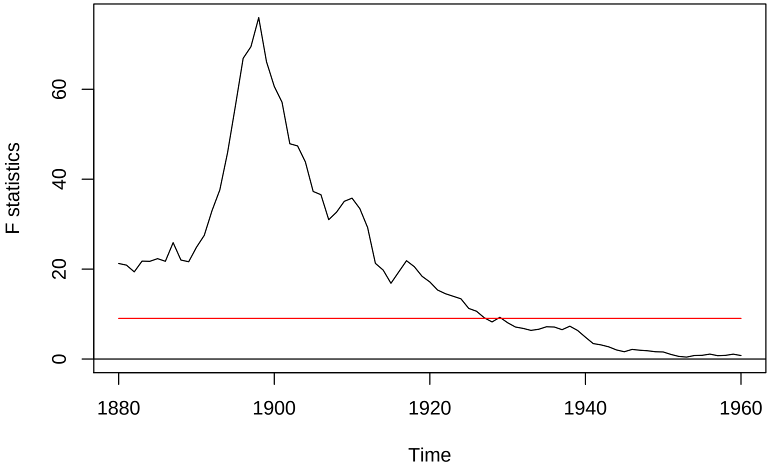
In R: Provided by *strucchange*.

- `efp()` computes empirical fluctuation processes of various `types`. Result is object of class "`efp`".
- `Fstats()` computes sequences of *F* statistics. Result is object of class "`Fstats`".
- Both classes have `plot()` methods for performing tests graphically.
- Both classes have `sctest()` methods (for structural change test) for performing traditional significance tests.
- Both `plot()` and `sctest()` take further parameters to perform various "flavors" of the tests.

Illustration: *F* statistics and sup*F* test with 10% trimming.

```
R> fs <- Fstats(Nile ~ 1, from = 0.1)
R> plot(fs)
```

Structural changes: *F* tests



Structural changes: Dating

Goal: Estimation of breakpoints between abrupt shifts. Use

$$Y_t = X_t^\top \beta^{(j)} + Z_t, \qquad t = \tau_{j-1}+1, \dots, \tau_j, \quad j = 1, \dots, m+1,$$

where

- $j = 1, \dots, m+1$ segment index,
- $\beta^{(j)}$ segment-specific set of regression coefficients,
- $\{\tau_1, \dots, \tau_m\}$ set of unknown breakpoints (convention: $\tau_0 = 0$ and $\tau_{m+1} = n$).

In R: Function `breakpoints()`.

- Uses dynamic programming algorithm based on Bellman principle.
- Finds those m breakpoints that minimize residual sum of squares of model with $m+1$ segments.
- Bandwidth parameter h determines minimal segment size of $h \cdot n$ observations.

Structural changes: Dating

If the timing of the intervention (of building the Aswan dam) becoming effective were unknown, it could easily be estimated.

```
R> bp <- breakpoints(Nile ~ 1)
R> plot(bp)
R> summary(bp)
         Optimal (m+1)-segment partition:

Call:
breakpoints.formula(formula = Nile ~ 1)

Breakpoints at observation number:

m = 1       28
m = 2       28          83
m = 3       28       68 83
m = 4       28 45 68 83
m = 5    15 30 45 68 83
```

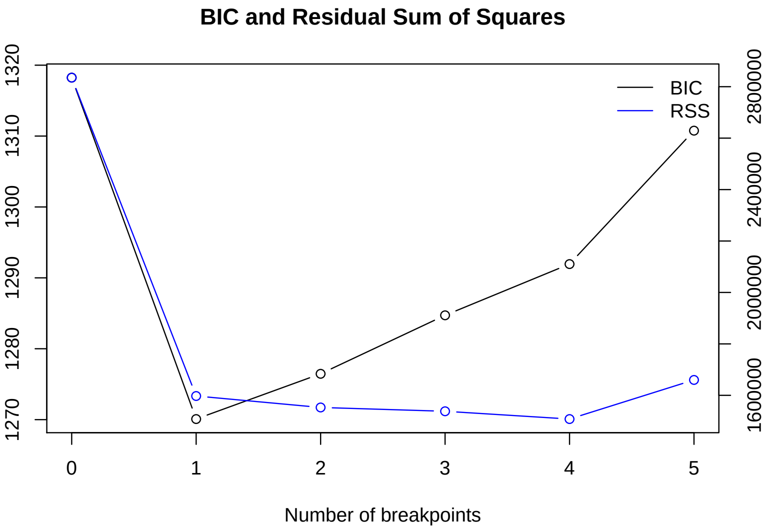
Structural changes: Dating

```
Corresponding to breakdates:

m = 1        1898
m = 2        1898                1953
m = 3        1898           1938 1953
m = 4        1898 1915 1938 1953
m = 5   1885 1900 1915 1938 1953

Fit:

m   0       1       2       3       4       5
RSS 2835157 1597457 1552924 1538097 1507888 1659994
BIC    1318    1270    1276    1285    1292    1311
```

Default in `coef()`, `residuals()`, `fitted()`, `confint()` is minimum BIC partition but `breaks = ...` can be added.

```
R> coef(bp)
            (Intercept)
1871 - 1898        1098
1899 - 1970         850
```

Structural changes: Dating



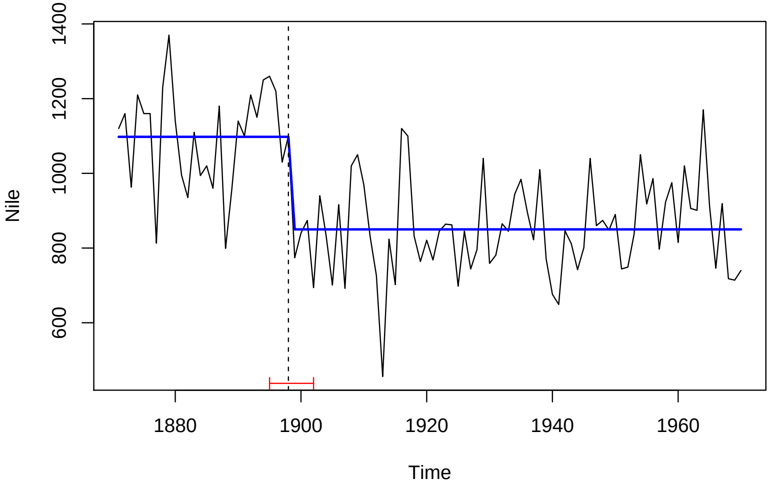
Structural changes: Dating

Furthermore: Using also non-standard asymptotic theory based on functional central limit theorems confidence intervals can be computed.

Visualization of data, fitted means, and confidence intervals for breakpoints.

```
R> plot(Nile)
R> lines(fitted(bp), col = 4, lwd = 2)
R> lines(confint(bp))
```

Structural changes: Dating



Structural changes: Seatbelt data

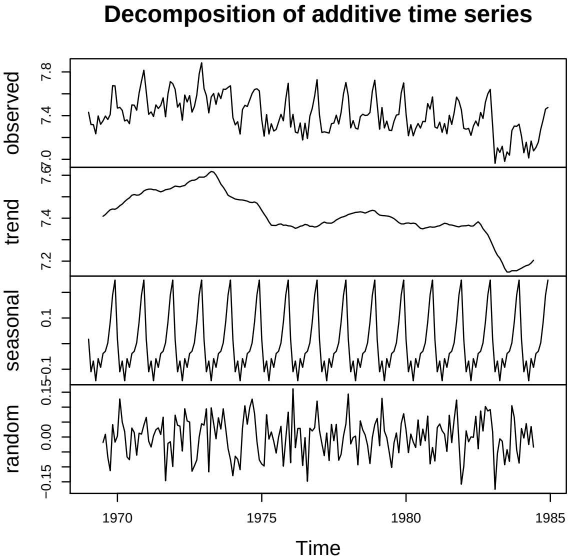
Illustration: `UKDriverDeaths`. Monthly totals of car drivers in Great Britain killed or seriously injured from 1969(1) to 1984(12). Taken from Harvey & Durbin (1986, *JRSS A*).

Remarks:

- Decrease in mean number of casualties after policy change (mandatory wearing of seatbelts) in January 1983.
- Parameters of time series model unlikely to be stable throughout sample period.
- Seasonality can be captured by SARIMA-type model.

```
R> dd <- log(UKDriverDeaths)
R> plot(decompose(dd))
```

Structural changes: Seatbelt data



Structural changes: Seatbelt data

Model: Similar to ARIMA$(1,0,0)(1,0,0)_{12}$ fitted by OLS

$$Y_t \;=\; \beta_1 \;+\; \beta_2 Y_{t-1} \;+\; \beta_3 Y_{t-12} \;+\; e_t \qquad t = 13, \dots, 192.$$

In R: Two approaches.

Approach 1: Use convenience interface `dynlm()`.

```
R> dynlm(dd ~ L(dd) + L(dd, 12))
```

```
Time series regression with "ts" data:
Start = 1970(1), End = 1984(12)

Call:
dynlm(formula = dd ~ L(dd) + L(dd, 12))

Coefficients:
(Intercept)        L(dd)    L(dd, 12)
      0.421        0.431        0.511
```

Structural changes: Seatbelt data

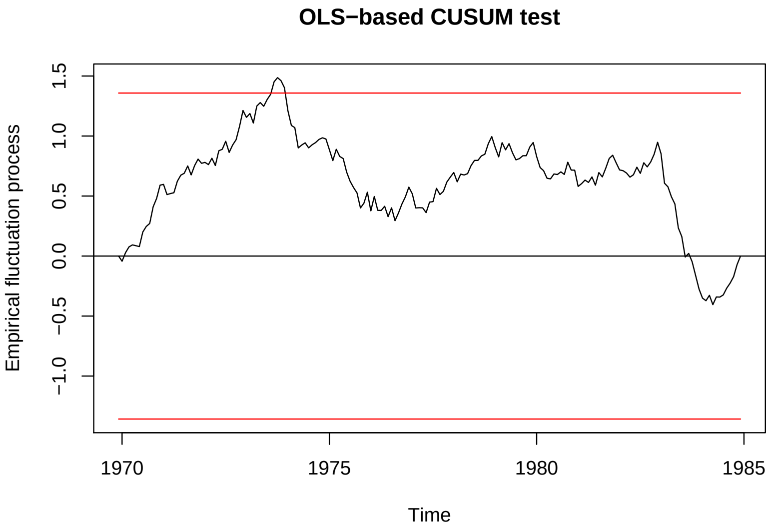
Approach 2: Set up regressors "by hand" and call `lm()`.

```
R> dd_dat <- ts.intersect(dd, dd1 = lag(dd, k = -1),
+    dd12 = lag(dd, k = -12))
R> lm(dd ~ dd1 + dd12, data = dd_dat)
Call:
lm(formula = dd ~ dd1 + dd12, data = dd_dat)

Coefficients:
(Intercept)          dd1         dd12
      0.421        0.431        0.511
```

The latter is required for employing the functions from *strucchange*.

```
R> dd_ocus <- efp(dd ~ dd1 + dd12, data = dd_dat,
+    type = "OLS-CUSUM")
R> sctest(dd_ocus)
        OLS-based CUSUM test

data:  dd_ocus
S0 = 1.5, p-value = 0.02
```

Structural changes: Seatbelt data



Structural changes: Seatbelt data

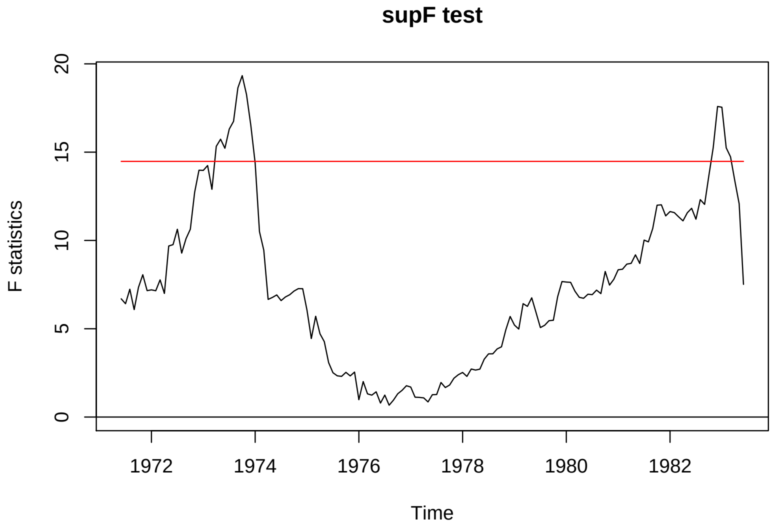
Additionally: Employ sup*F* test with 10% trimming.

```
R> dd_fs <- Fstats(dd ~ dd1 + dd12, data = dd_dat, from = 0.1)
R> sctest(dd_fs)

        supF test

data:  dd_fs
sup.F = 19, p-value = 0.007
```

Visualization:

```
R> plot(dd_fs, main = "supF test")
```

Structural changes: Seatbelt data



Structural changes: Seatbelt data

Dating: Identifies two changes associated with oil crisis and introduction of seatbelt regulation, respectively.

```
R> dd_bp <- breakpoints(dd ~ dd1 + dd12, data = dd_dat, h = 0.1)
R> coef(dd_bp, breaks = 2)
```

```
                    (Intercept)    dd1   dd12
1970(1) - 1973(10)        1.458 0.1173 0.6945
1973(11) - 1983(1)        1.534 0.2182 0.5723
1983(2) - 1984(12)        1.687 0.5486 0.2142
```

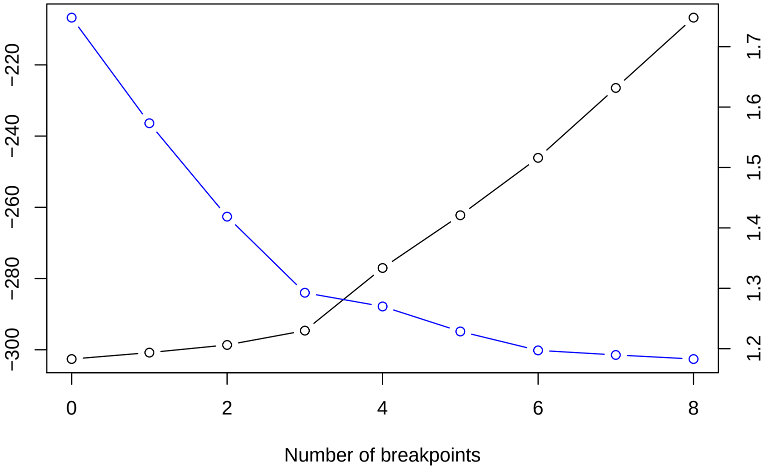
Visualization of dating results.

```
R> plot(dd_bp, legend = FALSE, main = "")
```

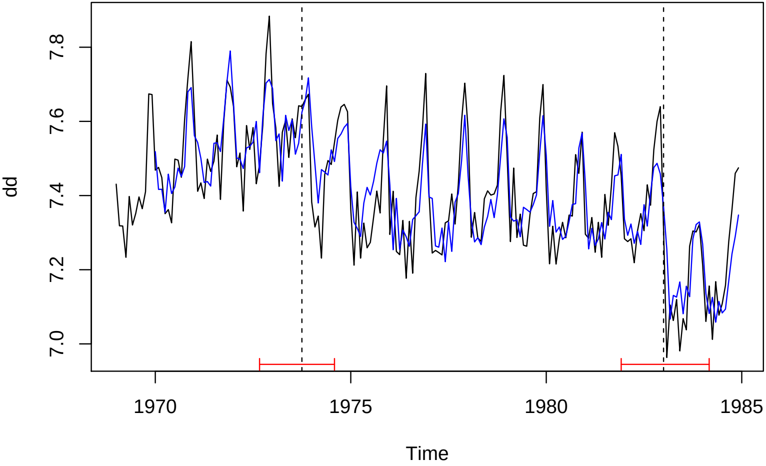
Visualization of 3-segment model.

```
R> plot(dd)
R> lines(fitted(dd_bp, breaks = 2), col = 4)
R> lines(confint(dd_bp, breaks = 2))
```

Structural changes: Seatbelt data



Structural changes: Seatbelt data



Time Series Regression Models

Outliers

Outliers

- Atypical observations with abrupt, short-term changes in the underlying process.
- Additive outliers:

$$Y_t' = Y_t + \omega_A P_t^{(T)}$$

- Innovative outliers:

$$e_t' = e_t + \omega_I P_t^{(T)}$$

which can be rewritten as

$$Y_t' = Y_t + \omega_I \psi_{t-T}$$

Time Series Regression Models

Spurious Correlation

Spurious correlation

- Goal of time series regression: Employ correlation not only within a series but across several series for modeling/prediction.
- For jointly stationary series $\{X_t\}$ and $\{Y_t\}$ employ cross correlation function $\varrho_k(X, Y) = \text{Cor}(X_t, Y_{t-k})$.
- However: If both series have autocorrelation of (or close to) 1, they may appear to be correlated although they are not.

Time Series Regression Models

Prewhitening and Stochastic Regression

Prewhitening and stochastic regression

- To avoid problems of spurious regression, remove autocorrelation in each series first.
- Employ residuals from an AR(p) model fit, with p sufficiently large to approximate the AR(∞) representation of ARIMA processes.