SOME PROCEDURES FOR DETECTION OF STRUCTURAL BREAKS

Marie Hušková

Charles University, Prague

October 18, 2011
Karlsruhe

Outline

1. Introduction

Outline

Outline

Outline

Structural breaks

Marie Hušková

Outline

Introduction

I. Retrospective procedures

I.1. Location models

I.2. Regression model

I.3. Asymptotic Intermezzo

I.4. Remarks and applications

I.5. Applications and simulations

II. Sequential procedures

II.1.

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

INTRODUCTION

Procedures for detection of structural breaks—
procedures on stability of statistical models, disorders, segmented regression, switching regression, change point problem, etc.

Retrospective procedures:
all observations available at the beginning of data analysis

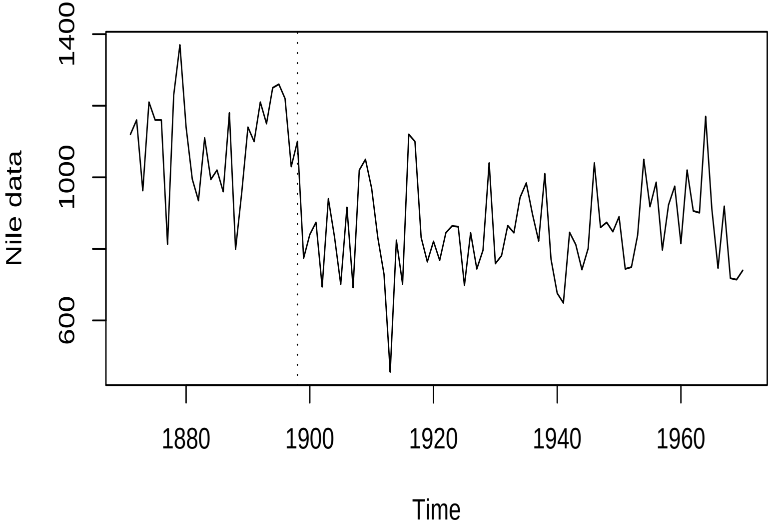
Sequential procedures:
observations are arriving sequentially, decisions made after each new observation.

Applications – meteorology, climatology, hydrology or environmental studies, econometric time series, statistical quality control, application in medical care, etc.

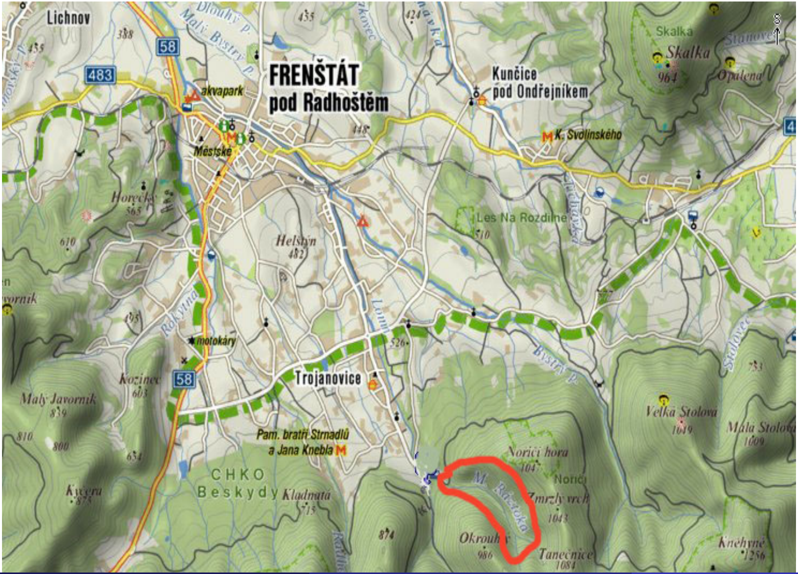
Testing and estimation, theoretical and computational problems.

Nile data

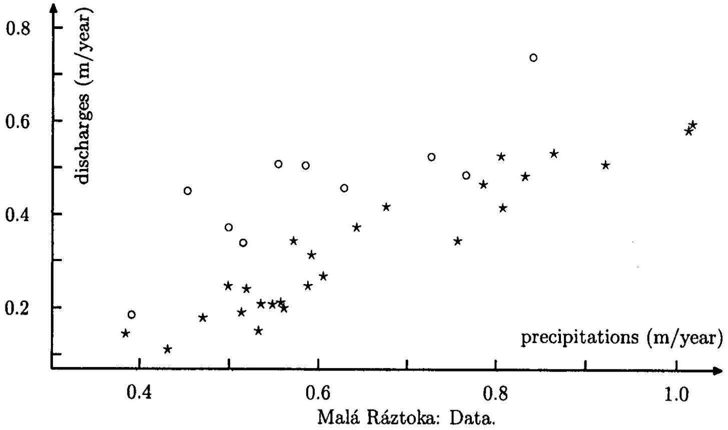
Annual water discharges at Aswan (Nile river)



Ráztoka river data



annual water discharges versus annual total amount of precipitation



Outline

I. RETROSPECTIVE PROCEDURES:

Observations $Y_1, \ldots, Y_n$ obtained at the ordered time points $t_1 < \cdots < t_n$ such that
$Y_1, \ldots, Y_{k^*}$ —- model I
$Y_{k^*}, \ldots, Y_n$ —- model II
k^* — **change point** — unknown

The problem: to detect (to test H_0: no change & H_1: there is a change), to identify k^* (to estimate k^*) and to estimate the model before and after the change.

• Many variants – multiple changes, abrupt changes, gradual changes, changes in various parameters, changes in distributions, independent and dependent observations.

• Construction of tests and estimators — various approaches as in most of the statistical problems.

Outline

Detection changes in location models

Location model

$$X_k = \mu_k + \epsilon_k \qquad 1 \le k \le n,$$

• $\mu_1, \ldots, \mu_n$ — means of the respective observations

• $\epsilon_1, \ldots, \epsilon_n$ random error terms with zero means with additional properties.

Testing problem

$$H_0 : \quad \mu_k = \mu \quad 1 \le k \le n$$

versus

H_A : there is $1 \le k^* < n$ such that $\mu_1 = \mu_2 = \ldots = \mu_k^* \neq \mu_{k^*+1} = \ldots = \mu_n$.

Estimator of the change point k^*

Typically test procedures based on functionals of properly standardized cumulative sums (CUSUM) used, e.g., H_0 is rejected for large

$$T_n = \max_{1\le k\le n} \Big|\sum_{i=1}^{k}(X_i - \overline{X}_n)\Big|/(n\sigma_n^2)^{1/2}$$

$$T_{n1} = \max_{1<k<n}\Big\{\sqrt{\frac{n}{k(n-k)}}\frac{1}{\sigma_n}\Big|\sum_{i=1}^{k}(X_i - \overline{X}_n)\Big|\Big\}$$

$$T_{n2}(\varepsilon) = \max_{\varepsilon n<k<(1-\varepsilon)n}\Big\{\frac{1}{\sqrt{n}}\frac{1}{\sigma_n}\Big|\sum_{i=1}^{k}(X_i - \overline{X}_n)\Big|\Big\}$$

$\overline{X}_n = (1/n)\sum_{1\le i\le n} X_i$

σ_n – an estimator of the scale σ

The large values indicate that the null hypothesis is not true.

- Approximation for critical values:

(i) based on limit distribution,

(ii) bootstrap based— with replacement or without replacement, modified block bootstrap (adjusted to a possible change)

- Under H_0 and certain assumptions, as $n \to \infty$:

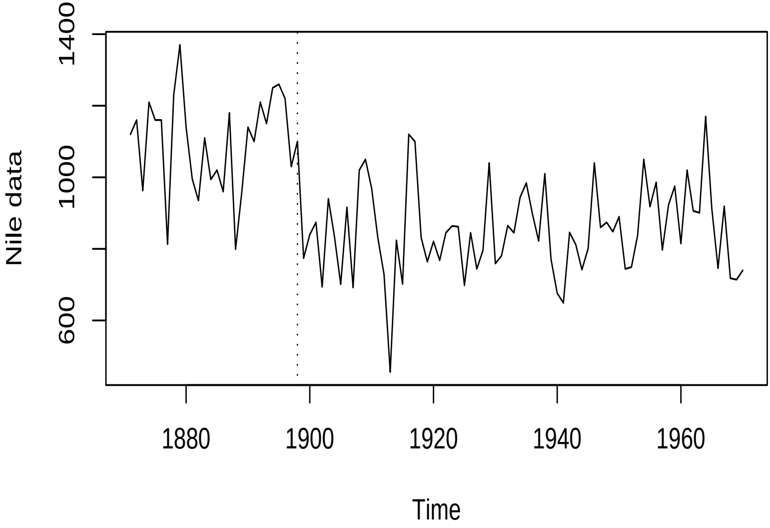
$$T_n \to^d \sup_{1<t<1} \{|B(t)|\}$$

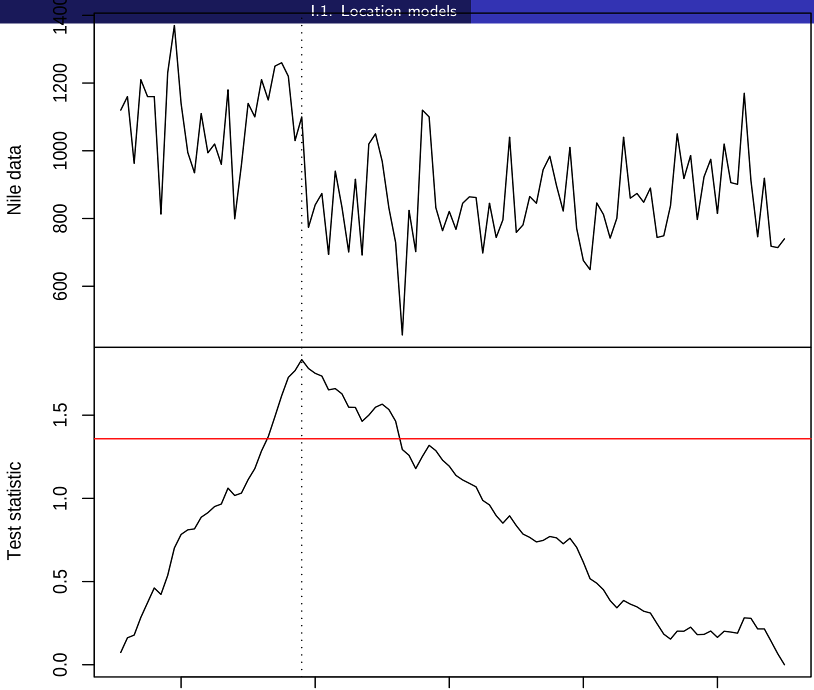
$\{B(t), 0 < t < 1\}$ – Brownian bridge

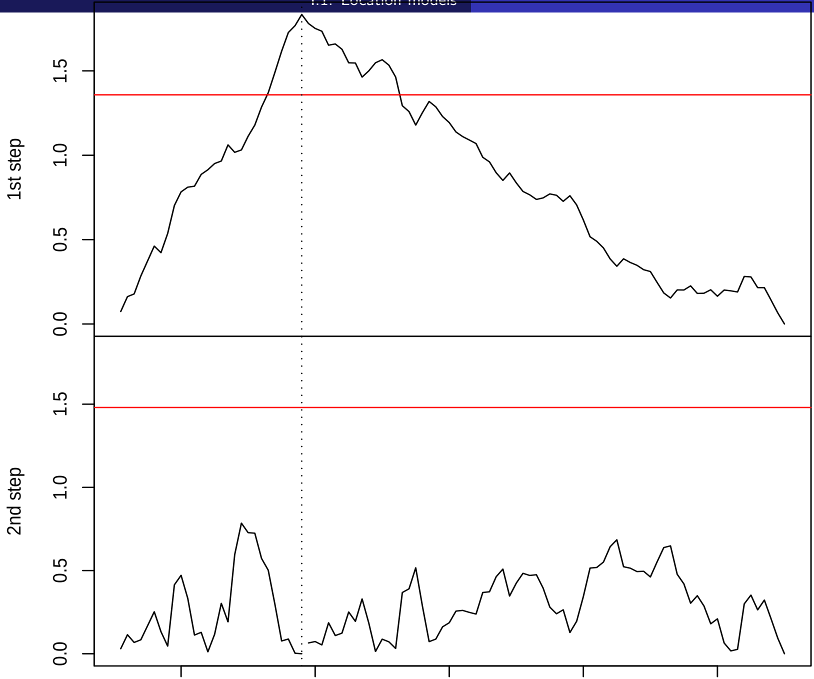
- Problem to find σ_n^2 – a Bartlett type estimator adjusted to a possible change –not always reasonable results.

Nile data

Annual water discharges at Aswan (Nile river)




Nile data
Test statistic
1400
1200
1000
800
600
1.5
1.0
0.5
0.0


1st step
2nd step
0.0
0.5
1.0
1.5
0.0
0.5
1.0
1.5

Outline

I.2. REGRESSION MODEL

$Y_1, \dots, Y_n$ are observed at time points $t_1 < \dots < t_n$:

$$\begin{aligned} Y_i \quad &= \mathbf{x}_i^T \boldsymbol{\beta} + e_i, & i = 1 \dots, k^* \\ &= \mathbf{x}_i^T \boldsymbol{\beta} + \mathbf{x}_i^T \boldsymbol{\delta} + e_i, & i = k^* + 1 \dots, n, \end{aligned}$$

$e_1, \dots, e_n$ — innovations, usually zero mean, nonzero variance σ^2 and finite $E|e_i|^{2+\Delta}$ with some $\Delta > 0$ plus restrictions on dependency

$\boldsymbol{\beta}, \boldsymbol{\delta} \neq \mathbf{0}$ — parameters

k^* .... change point

$\mathbf{x}_1, \ldots, \mathbf{x}_n$ – p-dim. design points (random or nonrandom):

nontrending regression: $\frac{1}{n}\sum_{i=1}^{k} \mathbf{x}_i \mathbf{x}_i^T \approx \frac{k}{n}\mathbf{C},\ k \leq n$

trending regression: $\mathbf{x}_i = \mathbf{h}(i/n), i = 1, \ldots, n$, $\mathbf{h}$ smooth nonconstant vector function

Main problems:

(i) $H_0:\ k^* = n \quad \& \quad H_1:\ k^* < n$

(ii) estimators of change points and model

Test statistics

$$T_n = \max_{p \le k < n-p} \left\{ \left(\widehat{\boldsymbol{\beta}}_k - \widehat{\boldsymbol{\beta}}_k^0 \right)^T \widehat{\boldsymbol{\Sigma}}_k^{-1} \left(\widehat{\boldsymbol{\beta}}_k - \widehat{\boldsymbol{\beta}}_k^0 \right) \right\}$$

- $\widehat{\boldsymbol{\beta}}_k$ — LSE of $\boldsymbol{\beta}$ based on $Y_1, \ldots, Y_k$
- $\widehat{\boldsymbol{\beta}}_k^0$ — LSE of $\boldsymbol{\beta}$ based on $Y_{k+1}, \ldots, Y_n$
- $\widehat{\boldsymbol{\Sigma}}_k^{-1}$ is an estimator of the variance matrix of $\widehat{\boldsymbol{\beta}}_k - \widehat{\boldsymbol{\beta}}_k^0$

Equivalently

$$T_n = \max_{p \le k < n-p} \left\{ \mathbf{S}_k^T \mathbf{C}_k^{-1} \mathbf{C}_n (\mathbf{C}_k^0)^{-1} \mathbf{S}_k \frac{1}{\widehat{\sigma}_n^2} \right\},$$

$\mathbf{S}_k = \sum_{i=1}^k \mathbf{x}_i \widehat{e}_i, \quad k = 1, \ldots, n,$
$\widehat{e}_i = Y_i - \mathbf{x}_i^T \widehat{\boldsymbol{\beta}}_n, \; i = 1, \ldots, n$ – - residuals

$\mathbf{C}_k = \sum_{i=1}^k \mathbf{x}_i \mathbf{x}_i^T, \quad \mathbf{C}_k^0 = \mathbf{C}_n - \mathbf{C}_k$

Test statistics

$$T_n = \max_{p \le k < n-p} \left\{ \left(\widehat{\beta}_k - \widehat{\beta}_k^0 \right)^T \widehat{\mathbf{\Sigma}}_k^{-1} \left(\widehat{\beta}_k - \widehat{\beta}_k^0 \right) \right\}$$

- $\widehat{\beta}_k$ — LSE of β based on $Y_1, \dots, Y_k$
- $\widehat{\beta}_k^0$ — LSE of β based on $Y_{k+1}, \dots, Y_n$
- $\widehat{\mathbf{\Sigma}}_k^{-1}$ is an estimator of the variance matrix of $\widehat{\beta}_k - \widehat{\beta}_k^0$

Equivalently

$$T_n = \max_{p \le k < n-p} \left\{ \mathbf{S}_k^T \mathbf{C}_k^{-1} \mathbf{C}_n (\mathbf{C}_k^0)^{-1} \mathbf{S}_k \frac{1}{\widehat{\sigma}_n^2} \right\},$$

$\mathbf{S}_k = \sum_{i=1}^k \mathbf{x}_i \widehat{e}_i, \quad k = 1, ..., n,$
$\widehat{e}_i = Y_i - \mathbf{x}_i^T \widehat{\beta}_n,\ i = 1, \dots, n$ – - residuals

$\mathbf{C}_k = \sum_{i=1}^k \mathbf{x}_i \mathbf{x}_i^T, \quad \mathbf{C}_k^0 = \mathbf{C}_n - \mathbf{C}_k$
$\widehat{\sigma}_n^2$ **-suitable standardization**

Critical regions

$$T_n > c_n(\alpha)$$

$c_n(\alpha)$ — critical value

α — level

Approximation of the critical values:

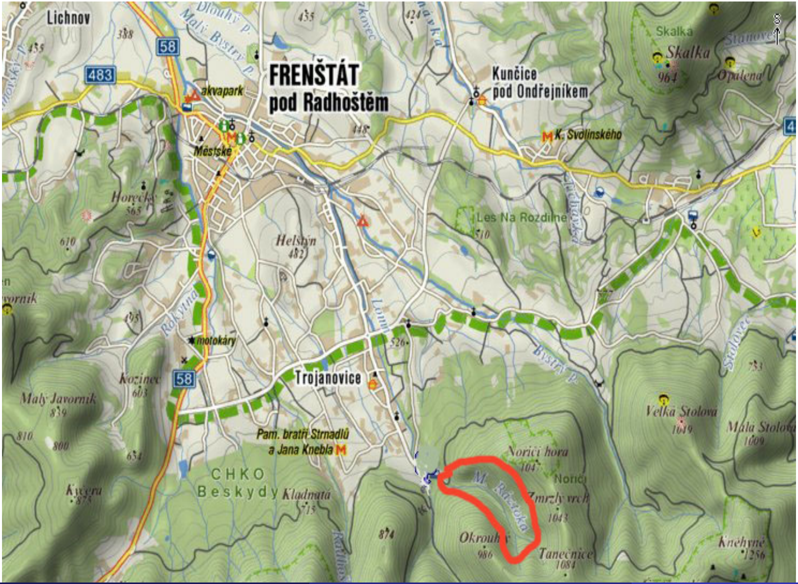
(i) limit distribution of T_n under H_0;

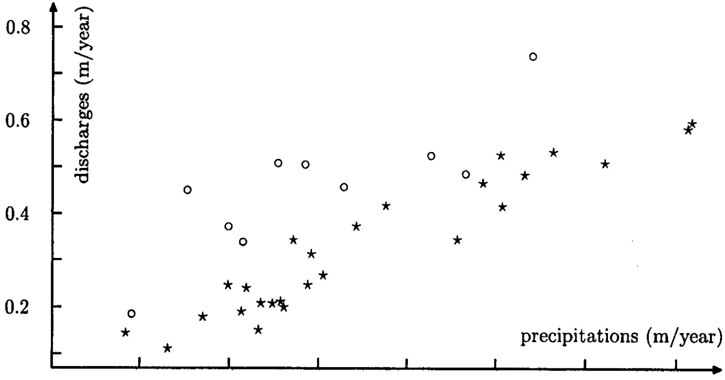
(ii) resampling methods (bootstrap)

Estimator $\widehat{k}^*$ of the change point k^* defined as such $\widehat{k}^*$ it maximizes w.r.t. k

$$\left\{(\widehat{\boldsymbol{\beta}}_k - \widehat{\boldsymbol{\beta}}_k^0)^T \widehat{\boldsymbol{\Sigma}}_k^{-1} (\widehat{\boldsymbol{\beta}}_k - \widehat{\boldsymbol{\beta}}_k^0)\right\}$$

Ráztoka river data





Malá Ráztoka: Data.

rainfall-runoffs relations – Ráztoka

Data small river Malá Ráztoka in 1954 — 1989 (36 years)
$(x_i, Y_i),\ i = 1, \ldots, 36$

x_i — annual total amount of precipitation

Y_i — annual water discharges

Question: was there a change in dependence of annual water discharges on annual total amount of precipitation

$$\begin{aligned} E\,Y_i &= \beta_1 + \beta_2 x_i, & i &= 1, \ldots, m^* \\ E\,Y_i &= \beta_1 + \beta_2 x_i + \delta_1 + \delta_2 x_i, & i &= m^* + 1, \ldots, n \end{aligned}$$

$(\beta_1, \beta_2), (\delta_1, \delta_2), k^*$ – parameters

nontrending regression

Conclusion: there was a change

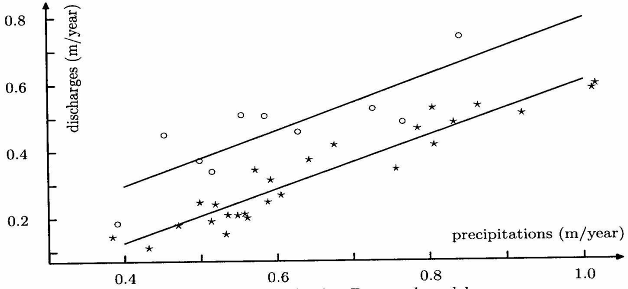
estimator of the change point – 26 years

estimator before the change : -193.6 0.8

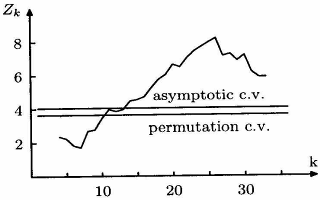
estimator after the change: -33.1 0.82

Figure; $T_n = \max_{2<k<n-1} Z_k, \quad T_n^0 = \max_{2<k<n-1} Z_k^0$

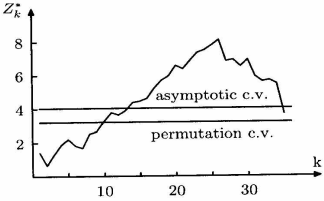
Z_k^0— partial sum of standardized residuals



Malá Ráztoka: Data and model.



Statistics Z_k.



Statistics Z_k^*.

Outline

I.3. ASYMPTOTIC INTERMEZZO

Test statistics:

$$T_n = \max_{p \le k < n-p} \left\{ \left(\widehat{\beta}_k - \widehat{\beta}_k^0 \right)^T \widehat{\mathbf{\Sigma}}_k^{-1} \left(\widehat{\beta}_k - \widehat{\beta}_k^0 \right) \right\}$$

$$T_n(\varepsilon) = \max_{\varepsilon n \le k \le (1-\varepsilon)n} \left\{ \dots \right\}, \quad 0 < \varepsilon < 1/2$$

nontrending and trending regression:

$$\lim_{n \to \infty} P(a(\log n)(T_n)^{1/2} \le t + b_p(\log n))$$

$$= \exp\{-2\exp\{-t\}\}, \; t \in R^1,$$

$a(y) = (2\log y)^{1/2}$,

$b_p(y) = 2\log y + \frac{p}{2}\log\log y - \log(2\Gamma(p/2)), y > 1,$

nontrending regression

$$T_n(\varepsilon) \to^d \sup_{\varepsilon<t<1-\varepsilon} \left\{ \frac{\sum_{i=1}^p B_i^2(t)}{t(1-t)} \right\}$$

$\{B_j(t); t \in (0,1)\}$, $j = 1, ..., p$, — independent Brownian bridges

trending regression $\mathbf{x}_i = \mathbf{h}(i/n)$

$$T_n(\varepsilon) \to^d \sup_{\varepsilon<t<1-\varepsilon} \mathbf{S}^T(t)\mathbf{C}(t)\mathbf{C}^{-1}(1)\mathbf{C}^0(t)\mathbf{S}(t)$$

$$\mathbf{S}(t) = \int_0^t \mathbf{h}(x)dB(x) - \mathbf{C}(t)\mathbf{C}^{-1}(1)\int_0^1 \mathbf{h}(x)dB(x), \quad t \in [0,1]$$

with $\{B(x),\ x \in [0,1]\}$ being a Brownian bridge,
$\mathbf{C}(t) = \lim_{n\to\infty} \frac{1}{n}\mathbf{C}_{\lfloor nt \rfloor}$.

Outline

I.4. REMARKS

• Other type procedures - sum type procedures–Bayesian ones, L_1- procedures (L_1-estimators and L_1- residuals), $M-$ procedures

• Suitable bootstrap provides good approximations for critical values.

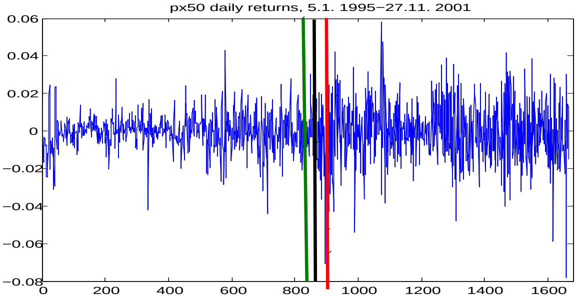
• Other models — time series models, nonlinear models, nonparametric regression models,...

• Problem of standardization — problem of choice of σ^2 — Bartlett type estimator adjusted to a possible change

Outline

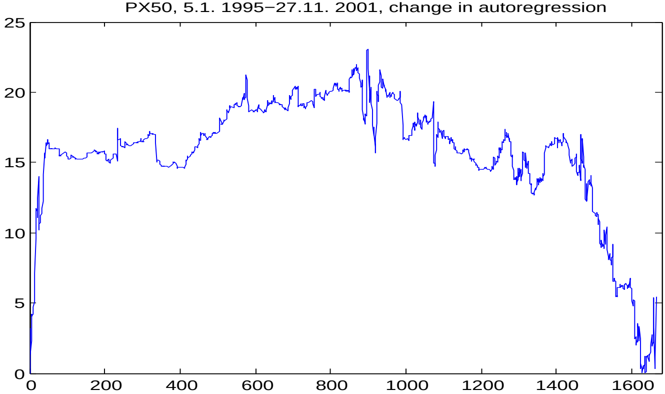
I.5. Application 1

Daily return of PX50 in 5.1.1995 - 27.11. 2001.



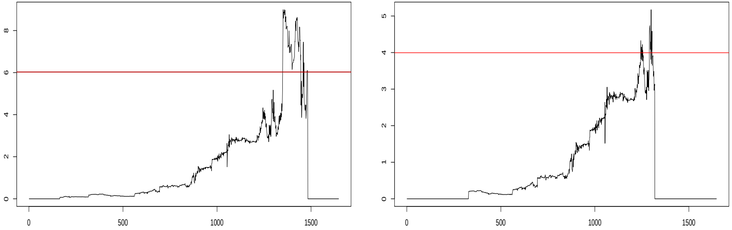
Test for a change in autoregression
Read line — estimated change point (1.9. 1998)

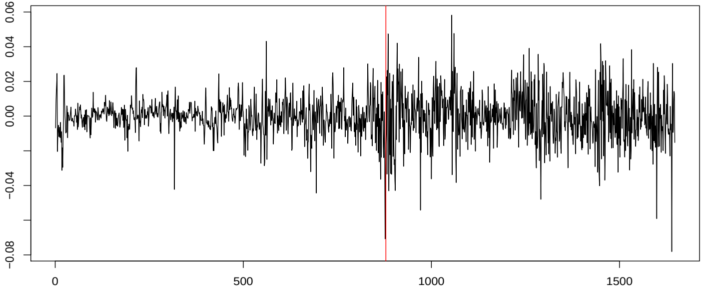
Green line — estimated change point in volatility based on 8.11. 1994 - 24.5. 1999
Black line — estimated change point in volatility based on 5.1.1995 - 27.11. 2001 (18.6. 1998)



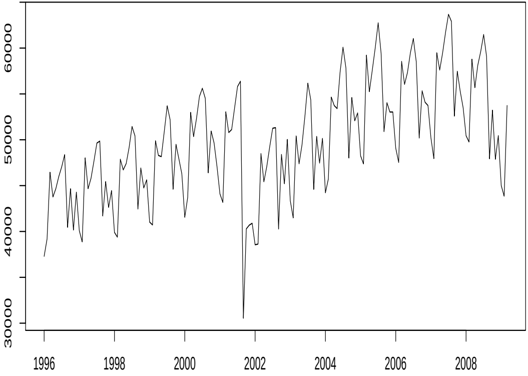
Tests statistics for a change in autoregression:
Above: classical test statistic with estimated σ^2

Next page: ratio type tests statistics, left $\gamma = 0.1$, right $\gamma = 0.2$


0
2
4
6
8
0
500
1000
1500
0
1
2
3
4
5
0
500
1000
1500


PX50 daily returns, 1.2.1995 – 27.9.2001
0.06
0.04
0.02
0.00
−0.04
−0.08
0
500
1000
1500

Simulation and Application 2
Monthly air traffic data



Simulation

$$Y_j = 4 + \frac{6j}{n} + \frac{3}{2}\cos(\frac{2\pi j}{12n}) + \frac{3}{2}\sin(\frac{2\pi}{12n}) + \frac{9}{10}\cos(\frac{2\pi}{4n} + \frac{1}{2}\sin(\frac{2\pi}{4n}) + e_j,$$

$j = 1, \dots, 200$

e_j –either AR(1) or MA(1), normal distr.

Change point $k^* = 100$ either in the intercept or in one harmonic regressor

1000 repetitions

critical value obtained through circular block bootstrap

$H_A^{(1)}$ change in intercept, AR (1) or MA(1)
$H_A^{(1)}$ change in one regressor, AR (1) or MA(1)

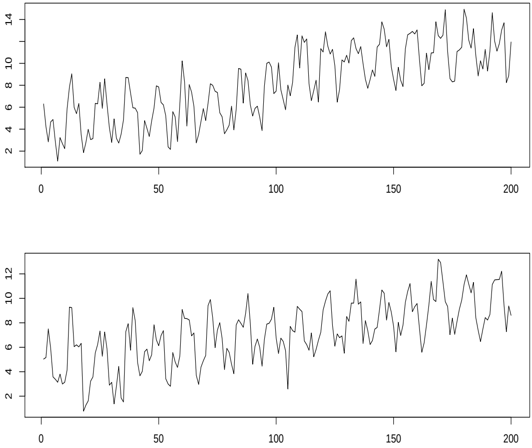


Figure 4: Time series plots of the processes under $H_A^{(1)}$ with $\delta = 2$ (upper panel) and $H_A^{(2)}$ (lower

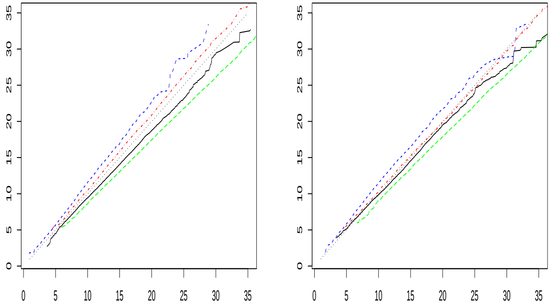


Figure 2: Circular bootstrap distribution versus finite sample distribution for AR(1) innovations (left) and MA(1) innovations (right) with scalings $\varphi^2 = 2$ (- - -), 1.5 (-·-·-), 1.0 (—), 0.5 (- - -).

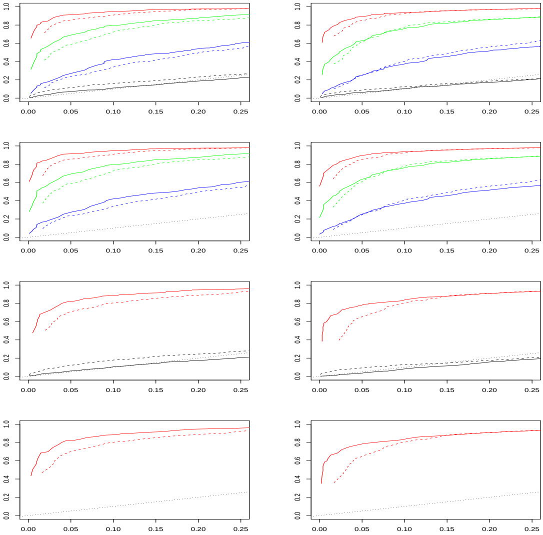
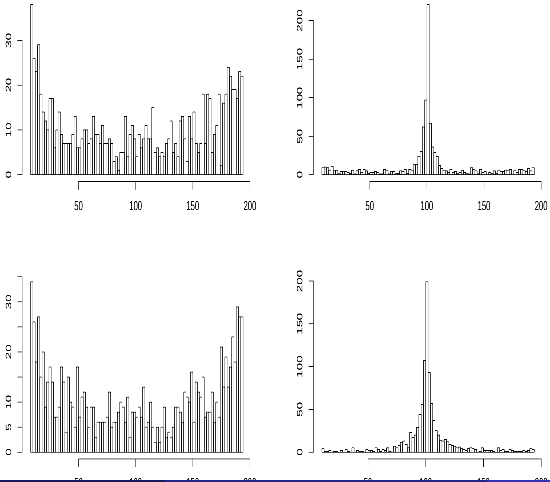
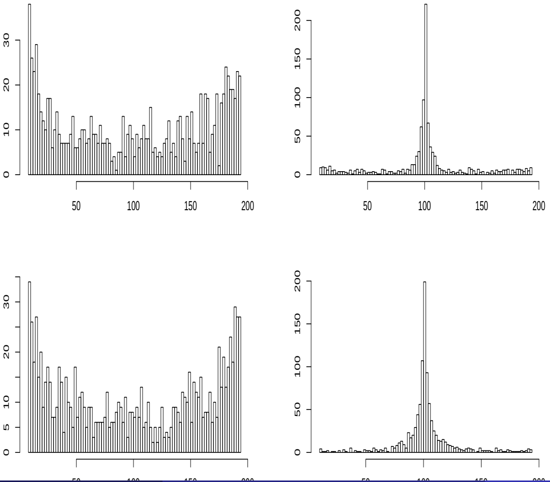


Figure 3: Size-power curves for $H_A^{(1)}$ (upper half) and $H_A^{(2)}$ (lower half) for the asymptotic modification (first and third line) and the the circular bootstrap (second and fourth line) with AR(1) innovations (left) and MA(1) innovations (right).


0
10
20
30
50
100
150
200
0
50
100
150
200
50
100
150
200
0
5
10
20
30
0
50
100
150
200


0
10
20
30
50
100
150
200
0
50
100
150
200
50
100
150
200
0
5
10
20
30
0
50
100
150
200

Data

Monthly air traffic data

model through the root of data, $n = 159$

$$Y_j = \beta_0 + \beta_1 j/n + \sum_{\ell=1}^{q} \Big(\beta_\ell^c \cos(2\pi\omega_{ell} j/n) + \beta_\ell^{(s)} \sin(2\pi\omega_{ell} j/n) \Big) + e_j$$

$q = 4$, $\omega_1 = 2/160$, $\omega_3 = 13/160$, $\omega_3 = 40/160$, $\omega_1 = 80/160$

ω_2 – annual cycle

ω_3 – quarterly cycle

ω_4 —two months cycle

$\widehat{k}^* = 69$

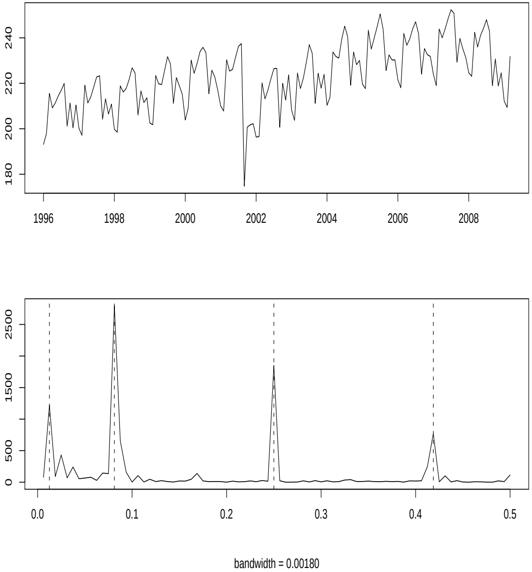
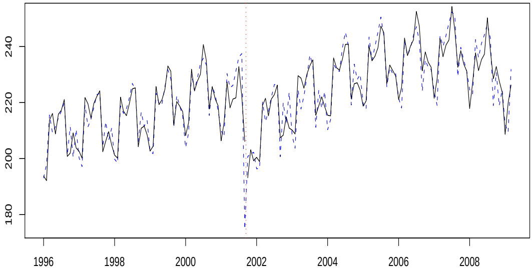


Figure 6: Square root transformation of the monthly air carrier traffic data (upper panel) and its



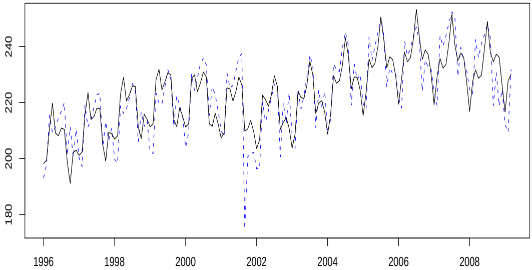


Figure 7: The fitted model based on the proposed data segmentation procedure (upper panel) and

Outline

II. SEQUENTIAL PROCEDURES

• **Sequential setup** – data arrives sequentially and after each new observation a decision is made (either ”data indicate a break” or ”data do not indicate a break”). We want to reveal a change as soon as possible, however to avoid a false alarm.

• **Historical (training) data** of size m without any instability (no change) are assumed.

• **Monitoring schemes** to detect an instability in parameters in regression model and autoregressive sequence.

• Chu et al (1996), Leisch et al. (2000, 2005).

• Such problems occurs:

statistical quality control

medicine (e.g., monitoring of patients in intensive medical care, monitoring elderly people at home)

detection of instability in financial and econometric time series (e.g. an instability in the CAPM models).

Outline

II.1. REGRESSION MODELS

$$Y_i = \mathbf{X}_i^T \boldsymbol{\beta}_i + e_i, \quad 1 \le i < \infty \tag{1}$$

$e_1, e_2, \dots$ – (i.i.d.) random errors
$\mathbf{X}_1, \mathbf{X}_2, \dots$ – p-dimensional design points
$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots$ – parameters

"noncontamination" assumption: $Y_1, \dots, Y_m$ – historical data with $\boldsymbol{\beta}_1 = \dots = \boldsymbol{\beta}_m = \boldsymbol{\beta}_0$ ($\boldsymbol{\beta}_0$ –unknown)

Testing problem

$H_0 : \boldsymbol{\beta}_i = \boldsymbol{\beta}_0$, for all $i > m$

$H_A : \ \textit{there is } k^* \ge 1 \quad \boldsymbol{\beta}_i = \boldsymbol{\beta}_0, \quad i = m+1, \dots, m+k^*$

$\boldsymbol{\beta}_i = \boldsymbol{\beta}^0, \quad i \ge k^* + 1, \ \boldsymbol{\beta}_0 \ne \boldsymbol{\beta}^0$

$\boldsymbol{\beta}^0 \ne \boldsymbol{\beta}_0, k^*$ – unknown

Outline

Stopping rule:

$$\tau(m,N) = \begin{cases} \inf\{1 \le k \le N : \, |Q(m,k)| \ge cq(k/m)\} \\ \infty \text{ if } |Q(m,k)| < c\sqrt{m}q(k/m) \text{ for all } k = 1, 2, \dots, N \end{cases}$$

$Q(m,k)$ – statistics (detectors) based on $(\mathbf{X}_i, Y_i),\ i = 1, \dots, m+k$,

$q(t),\ t \in (0,1)$ – (stopping) boundary function

c – tuning constant

$N = N(m) \to \infty$ as $m \to \infty$

---

the null hypothesis rejected and observation is stopped if

$$|Q(m,k)| \ge cq(k/m),$$

otherwise observations continue

Requested:

(i) under H_0:

$$\lim_{m\to\infty} P_{H_0}[\tau(m,N) < \infty] = \alpha \qquad (2)$$

(ii) under H_A:

$$\lim_{m\to\infty} P_{H_A}[\tau(m,N) < \infty] = 1, \qquad (3)$$

$\alpha \in (0,1)$ – level of the test (α close to 0)

$\alpha = 1/2$, $N < \infty$— then (2) can be interpreted as request N is median of $\tau(m,N)$

Classical sequential analysis works in terms of expectation of $\tau(m,N)$.

The major problem – the choice of detectors and the boundary functions

Choice of $c = c_m(\alpha)$:

$$P_{H_0}(|Q(m,k)| \geq cq(k/m)) \approx \alpha$$

Classes of detectors:

(I) detectors based on partial sums of residuals (CUSUM)

(II) detectors based on quadratic forms of partial sums of weighted residuals (differences of estimators)

(III) detectors based on partial sums of predictors

(IV) detectors based on partial sums of recursive residuals

Class of boundary functions q:

(B.1) $q(t) = q_\gamma(t) = (1+t)\,(t/(t+1))^\gamma,\ t \in (0,\infty)$ where $\gamma \in [0, 1/2)$.

γ — a tuning parameter,

$\gamma = 0$ — for expected late changes

γ close to $1/2$ – for expected early changes

Description of test procedures

Procedure (I) CUSUM (cumulative sums) test procedure based on partial sums L_2-residuals

$$\widehat{e}_i = Y_i - \mathbf{X}_i^T \widehat{\beta}_m, \tag{4}$$

$\widehat{\beta}_n$ is the LSE of β based on the first n observations
Detectors

$$Q_I(m,k) = \frac{1}{\widehat{\sigma}_m} \sum_{i=m+1}^{m+k} \widehat{e}_i, \tag{5}$$

$\widehat{\sigma}_m^2$ – suitable standardization based on $Y_1, \ldots, Y_m$, in i.i.d. case

$$\widehat{\sigma}_m^2 = \frac{1}{m-p} \sum_{i=1}^{m} \widehat{e}_i^2. \tag{6}$$

By theoretical results below $c = c_I(\alpha, \gamma)$ is a solution of:

$$P\left[\sup_{0\le t\le 1} \frac{|W(t)|}{t^\gamma} \ge c\right] = \alpha, \tag{7}$$

$\{W(t),\, 0 \le t \le 1\}$ – Wiener process.

The related stopping rule $\tau_I(m, \gamma)$ has the level α

while the consistency holds under additional assumptions,

this can behave a quite poorly

Procedure (II) detectors based on weighted partial sums:

$$Q_{II}(m,k) = \frac{1}{\widehat{\sigma}_m^2}\Big(\sum_{i=m+1}^{m+k} \mathbf{X}_i \widehat{e}_i\Big)^T \mathbf{C}_m^{-1}\Big(\sum_{i=m+1}^{m+k} \mathbf{X}_i \widehat{e}_i\Big) \qquad (8)$$

where $\widehat{e}_i$ and $\widehat{\sigma}_m^2$ defined above, respectively, and

$$\mathbf{C}_k = \sum_{i=1}^{k} \mathbf{X}_i \mathbf{X}_i^T, \quad k = 1, 2, \dots.$$

equivalent expression:

$$Q_{II}(m,k) = \frac{1}{\widehat{\sigma}_m^2}\Big(\widehat{\boldsymbol{\beta}}_{m,m+k} - \widehat{\boldsymbol{\beta}}_m\Big)^T \Big(\mathbf{C}_{m+k} - \mathbf{C}_m\Big)\mathbf{C}_m^{-1}$$

$$\Big(\mathbf{C}_{m+k} - \mathbf{C}_m\Big)\Big(\widehat{\boldsymbol{\beta}}_{m,m+k} - \widehat{\boldsymbol{\beta}}_m\Big),$$

$\widehat{\boldsymbol{\beta}}_{m,m+k}$ is the LSE of $\boldsymbol{\beta}$ based on $Y_{m+1}, \dots, Y_{m+k}$

By Theorem 1 below the constant $c = c_{II}(\alpha, \gamma)$ is a solution of the equation:

$$P\left[\sup_{0\le t\le 1} \frac{\sum_{j=1}^{p} W_j^2(t)|}{t^{2\gamma}} \ge c\right] = \alpha, \tag{9}$$

$\{W_j(t),\ 0 \le t \le 1\}$, $j = 1, \dots, p$, – independent Wiener processes.

The related stopping rule $\tau_{II}(m, \gamma)$ has both desired properties (2) and (3).

Procedure (III) detectors based on combination of prediction approach of Clark and McFadden (2005) and ideas applied in *Procedure I*.

The detectors:

$$Q_{III}(m,k) = \frac{1}{\sqrt{m}\widehat{\eta}_m}\Big(\sum_{i=1}^{k}(Y_{m+i} - \widehat{Y}_{m+i})^2 - \frac{k}{m}\sum_{i=p+1}^{m}(Y_i - \widehat{Y}_i)^2\Big), \quad k = 1, \dots \tag{10}$$

where $\widehat{Y}_i$ is a prediction of the ith observation based on $i-1$ previous observations, i.e.,

$$\widehat{Y}_i = \mathbf{X}_i^T \widehat{\beta}_{i-1}, \quad i = p+1, \dots \tag{11}$$

and $\widehat{\eta}_m^2$ is the estimator of η^2 defined by

$$\widehat{\eta}_m^2 = \frac{1}{m}\sum_{i=p+1}^{m}(Y_i - \widehat{Y}_i)^4 - \Big(\frac{1}{m}\sum_{i=p+1}^{m}(Y_i - \widehat{Y}_i)^2\Big)^2. \tag{12}$$

By Theorem 1 below the constant $c = c_{III}(\alpha, \gamma)$ is a solution of the equation (7) w.r.t. c.

Under the considered assumptions the related stopping rule $\tau(m) = \tau_{III}(m, \gamma)$ has both desired properties (2) and (3).

The explicit expressions for the probabilities on the l.h.s. of (7) and (9) are known only for $\gamma = 0$, otherwise their approximations can be obtained through simulations.

Limit properties

Assumptions (regression):

Assumptions on $\{e_i, 1 \le i < \infty\}$ and $\{\mathbf{X}_i^T, 1 \le i < \infty\}$:

(A.1a) $\{e_i\}_{i=1}^{\infty}$ i.i.d. with $\mathrm{E}\, e_1 = 0$, $0 < \mathrm{Var}\, e_1 = \sigma^2 < \infty$ and $\mathrm{E}\, |e_1|^{\nu} < \infty$ for some $\nu > 2$,

(A.1b) $\{e_i\}_{i=1}^{\infty}$ i.i.d. with $\mathrm{E}\, e_1 = 0$, $0 < \eta = \mathrm{var}\,(e^2) < \infty$ and $\mathrm{E}\, |e_1|^{\lambda} < \infty$ for some $\lambda > 4$.

(A.2) $\{\mathbf{X}_i^T\}_{i=1}^{\infty}$ – strictly stationary sequence of p-dimensional vectors $\mathbf{X}_i^T = (1, X_{2i}, \dots, X_{pi})$, which is independent of $\{e_i, 1 \le i < \infty\}$,

(A.3) there exist a positive definite matrix $\mathbf{C}$ and a constant $\tau > 0$ such that

$$\left| \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i \mathbf{X}_i^T - \mathbf{C} \right| = O(n^{-\tau}), \quad a.s.$$

Ass. (A.2) and (A.3) satisfied, e.g., for autoregressive sequences $\{X_{ji}\}_i$, $j = 2, \dots, p$, with finite moments of order higher than 2.

Theorem 1. *Let $Y_1, Y_2, \dots,$ follow the model (1) and the assumptions (A.1), (A.3) and (B.1) be satisfied. Then under H_0*

$$\lim_{m\to\infty} P\left(\sup_{1\le k<\infty} \frac{|Q_I(m,k)|}{q_\gamma(k/m)} \le x\right) = P\left(\sup_{0\le t\le 1} \frac{|W_1(t)|}{t^\gamma} \le x\right) \quad (13)$$

$$\lim_{m\to\infty} P\left(\sup_{1\le k<\infty} \frac{Q_{II}(m,k)}{q^2_\gamma(k/m)} \le x\right) = P\left(\sup_{0\le t\le 1} \frac{\sum_{i=1}^p W_i^2(t)}{t^{2\gamma}} \le x\right) \quad (14)$$

and

$$\lim_{m\to\infty} P\left(\sup_{1\le k<\infty} \frac{|Q_{III}(m,k)|}{q_\gamma(k/m)} \le x\right) = P\left(\sup_{0\le t\le 1} \frac{|W_1(t)|}{t^\gamma} \le x\right) \quad (15)$$

for all x, where $\{W_i(t);\, 0 \le t \le 1\}$, $i = 1, \dots, p$ are independent Wiener processes.

Theorem 2. *$Y_1, Y_2, \dots,$ follow the model (1). Let the assumptions (A.2), (A.3) and (B.1) be satisfied. Under H_A with*

$$\lim_{m\to\infty} m\boldsymbol{\delta}_m^T\boldsymbol{\delta}_m = \infty \tag{16}$$

then, as $m \to \infty$

$$\sup_{1\le k<\infty} \frac{Q_{II}(m,k)}{q_\gamma^2(k/m)} \xrightarrow{P} \infty \tag{17}$$

if, moreover, (A.1b) is satisfied then

$$\sup_{1\le k<\infty} \frac{Q_{III}(m,k)}{q_\gamma(k/m)} \xrightarrow{P} \infty. \tag{18}$$

Outline

II.3. Remarks and modifications

• Theorem 1 provides approximations for the critical values (or the tuning parameters) c for the described procedures.

• The explicit form for the distributions of $\sup_{0\le t\le 1}\{|W_1(t)|t^{-\gamma}\}$ and $\sup_{0\le t\le 1}\{\sum_{i=1}^{p} W_i^2(t)t^{-2\gamma}\}$ known only for $\gamma = 0$, otherwise simulations of Wiener processes. Tables of some simulated critical values can be found, e.g. in Horváth et al. (2004) and Koubková (2004) for $p = 1$ and in Hušková and Koubková (2005) for $p = 2$.

• Theorems 1 and 2 imply that under respective assumptions the procedures described in Section 2 have the desired properties (2) and (3).

• The assumption (A.1a) in the assertion (13) in Theorem 1 can be relaxed. The assertion remains true if the assumption (A.1a) is replaced by the assumptions: there exist constants $\xi < 1/2$ and $\sigma_0 > 0$ and for each m there exist independent Wiener processes $\{W_{m,j}(t),\ t \in [0,\infty)\}$, $j = 1,2$ such that

$$\max_{1\le k<\infty} k^{-\xi} \Big| \sum_{i=m+1}^{m+k} e_i - \sigma_0 W_{m,1}(k) \Big| = O_P(1), \quad (m \to \infty) \tag{19}$$

,

$$m^{-\xi} \Big| \sum_{i=1}^{m} e_i - \sigma_0 W_{m,2}(m) \Big| = O_P(1), \quad (m \to \infty). \tag{20}$$

and

$$\sup_{1\le k<\infty} k^{-\xi} \Big| \sum_{i=m+1}^{m+k} \mathbf{X}_i e_i \Big| = O_P(1). \tag{21}$$

Similarly, the assertion on *Procedure III* in Theorem 1 remains true the assumption (A.1.b) under weaker assumptions.

• Moving sums procedures based on $\sum_{i=m+k+1}^{m+k+G}(.)$, $k = 1, \ldots,$ with a suitably prechosen G (MOSUM type procedures) instead partial sums of the form $\sum_{i=m+1}^{m+k}(.)$, $k = 1, \ldots$ (Leisch et al. (2000), Horváth, Kühn, Steinebach (2007)).

• The presented procedures can be viewed as L_2 procedures. The results can be extended to

• L_1 procedures (Koubková (2006),
• procedures generated by some loss function ρ,
• Rao score type tests.

• Bootstrap approximation for critical values proposed by Kirch (2008) and Hušková and Kirch (2009). Simulations have good performance.

- **Fluctuation tests** can be developed along the line of Section 2. The tests for the alternatives:

$$\boldsymbol{\beta}_i = \boldsymbol{\beta}_0 + m^{-\mu}\mathbf{g}(i/m), \quad i = 1, \dots,$$

$\mu \in (0, 1/2)$ and $\mathbf{g}$ is a function with finite variation, such that $\mathbf{g}(t) \neq \mathbf{0}, \quad t \in [0,1]$ and such that for each $\mathbf{d} \neq \mathbf{0}$

$$0 < \int_1^\infty (\mathbf{d}^T \mathbf{g}(t))^2 dt < \infty.$$

For details see , e.g., Leisch [2000].

• The main steps of the proofs rely on the asymptotic representations.

Procedure I:

$$\sum_{i=m+1}^{m+k} \widehat{e}_i = \sum_{i=m+1}^{m+k} (e_i - \overline{e}_n) + R_I(m,k).$$

Procedure II:

$$\sum_{i=m+1}^{m+k} \mathbf{X}_i \widehat{e}_i = \sum_{i=m+1}^{m+k} \mathbf{X}_i e_i - (\mathbf{C}_{m+k} - \mathbf{C}_m)\mathbf{C}_m^{-1} \sum_{j=1}^{m} \mathbf{X}_j e_j + R_{II}(m,k).$$

Procedure III:

$$\sum_{i=m+1}^{m+k} (Y_i - \widehat{Y}_i)^2 = \sum_{i=m+1}^{m+k} e_i^2 - \frac{k}{m} \sum_{j=1}^{m} e_j^2 + R_{III}(m,k).$$

Here $R_I(m,k)$, $R_{II}(m,k)$ and $R_{III}(m,k)$ are reminder terms that

• **Limit properties of the stopping rules under alternatives** If $k^* \approx m^\beta \;\; 1 > \beta \geq 0$ -small

$$\tau(m) - k^* = O_P\left(m^{(1-2\gamma)/(2-2\gamma)}\right), \quad (m \to \infty)$$

and for k^* small, δ_m – amount of change in location model

$$\frac{\tau_m - a_m}{b_m} \to^d N(0,1)$$

$$a_m = \left(\frac{cm^{1/2-\gamma}}{|\delta_m|}\right)^{1/(1-\gamma)}$$

$$b_m = \frac{\sigma}{(|\delta_m|(1-\gamma))} a_m^{1/2}$$

number of papers: Aue, Horváth, Steinebach, Reimherr, Koubková, Prášková

Outline

II.4. Simulation study

$$Y_i = \beta_{0,i} + X_i\beta_{1,i} + e_i$$

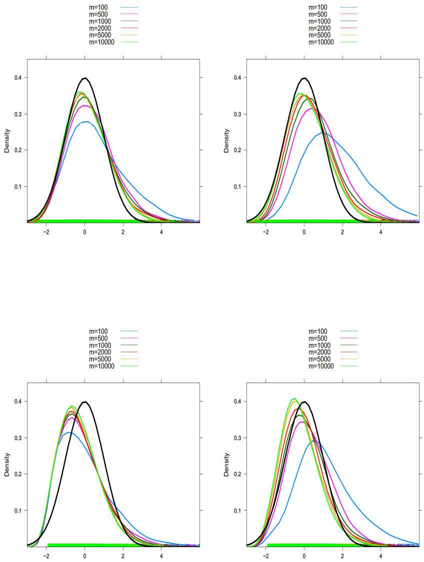
$e_i - N(0,1), \quad X_i - N(0,1)$

$m = 100, 500, 1000, 2000, \qquad \gamma = 0,\ 0.25,\ 0.49$

change in intercept, $k^* = 1,\ 5$

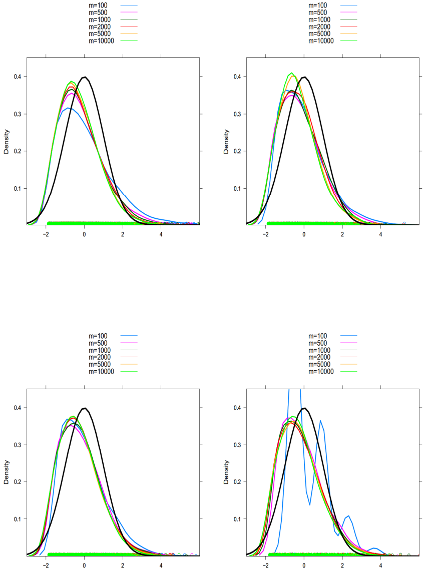
δ_m	100	500	1000	2000	5000
$3m^{-1/4}$ (S)	0.9487	0.6344	0.5335	0.4486	0.356
$12m^{-1/2.5}$ (M1)	1.9019	0.9991	0.7571	0.5738	0.397
$7m^{-1/3}$ (M2)	1.5081	0.8819	0.7000	0.5556	0.409
$6m^{-1/4.3}$(L)	2.0560	1.4141	1.2036	1.0244	0.827
$10m^{-1/4}$ (XL)	3.1623	2.1147	1.7783	1.4953	1.189

Table: Size of change depending on the training period

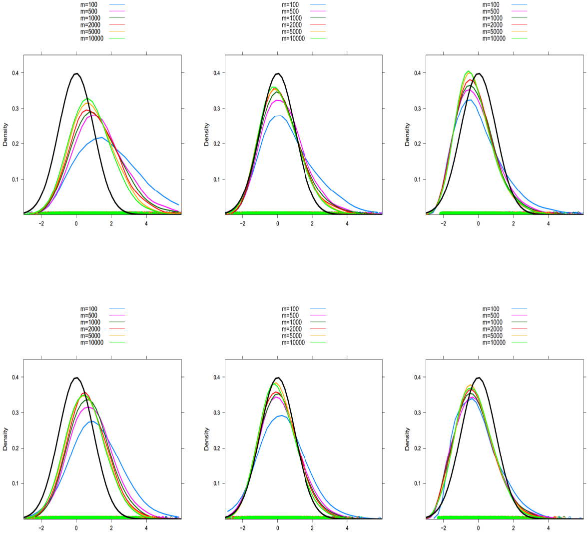


Small change

Left: $k^* = 1$, right: $k^* = 5$; upper: $\gamma = 0.25$, lower: $\gamma = 0.49$



$k^* = 1$, $\gamma = 0.49$. small (S), mild (M1), large (L), very large (XL)



$k^* = 1.$

$\gamma = 0, 0.25, 0.49$; upper: small change (S), lower: large change

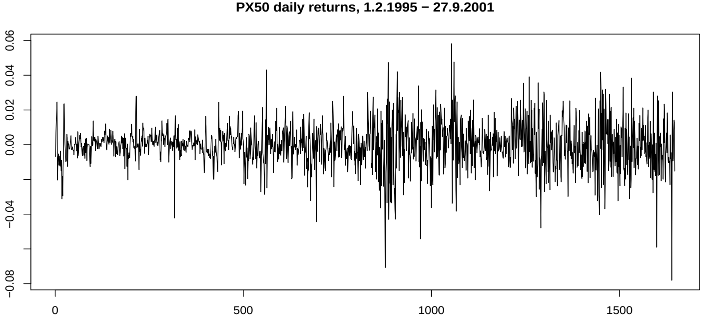
γ		100	500	1000	2000	5000
0	Min	13.00	62.0	105.0	172.0	369.
	Med	38.00	113.0	184.0	302.0	584.
	Mean	39.36	115.9	187.4	306.1	591.
	Max	121.00	247.0	350.0	514.0	913.
0.25	Min	1.00	13.00	16.00	53.0	104.
	Med	22.00	61.00	94.00	147.0	268.
	Mean	23.84	63.42	97.88	151.8	274.
	Max	87.00	174.00	249.00	347.0	596.
0.49	Min	1.00	1.00	1.00	1.00	1.00
	Med	13.00	26.00	37.00	51.00	81.0
	Mean	14.85	30.03	41.25	57.73	89.9
	Max	88.00	153.00	197.00	261.00	352.0

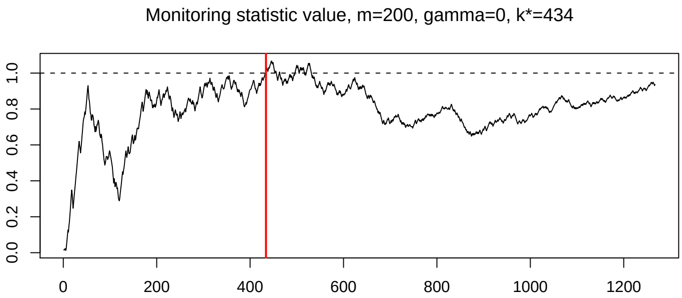
Table: Values of stopping times, small change in the intercept, $X_i \sim N(0,1)$, $e_i \sim N(0,1)$, $k^* = 1$

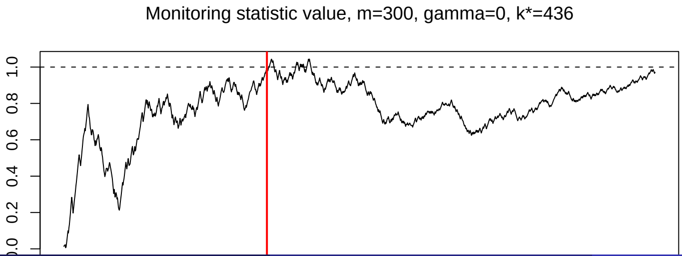
γ		100	500	1000	2000	5000
0	Min	16.00	65.0	105.0	168.0	368.0
	Med	43.00	118.0	189.0	307.0	589.5
	Mean	44.86	120.6	192.2	310.6	595.9
	Max	116.00	247.0	355.0	522.0	949.0
0.25	Min.	1.00	15.00	25.0	58.0	108.0
	Med	29.00	67.00	101.0	153.0	273.0
	Mean	30.28	69.44	103.7	157.5	280.2
	Max	96.00	180.00	250.0	349.0	596.0
0.49	Min	1.00	1.00	1.00	1.00	1.00
	Med	21.00	36.00	46.00	62.00	91.0
	Mean	23.45	38.56	50.03	66.11	98.3
	Max	96.00	151.00	190.00	304.00	342.0

Table: Values of stopping times, small change in the intercept, $X_i \sim N(0,1)$, $e_i \sim N(0,1)$, $k^* = 5$

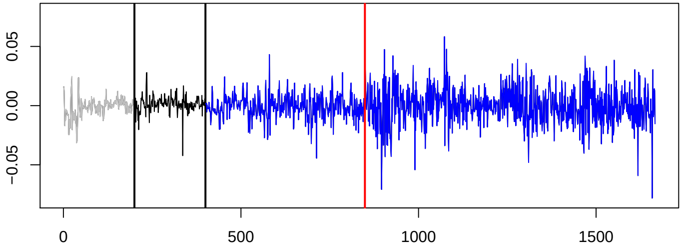
Application
Daily returns PX 5.1.1995 - 27.9.2001



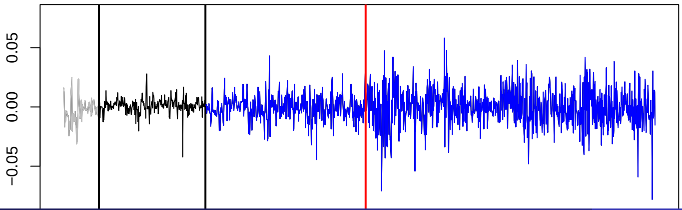

Monitoring statistic value, m=200, gamma=0, k*=434
0.0
0.2
0.4
0.6
0.8
1.0
0
200
400
600
800
1000
1200


Monitoring statistic value, m=300, gamma=0, k*=436
0.0
0.2
0.4
0.6
0.8
1.0

px50 daily returns, 5.1.1995 − 27.9.2001, out=200, m=200, gamma=0, k*=434



px50 daily returns, 5.1.1995 − 27.9.2001, out=100, m=300, gamma=0, k*=436



Outline

References

Csörgő, M. and, L. Horváth (1997). *Limit Theorems in Change–Point Analysis.* Wiley, Chichester.

Kim, J.-Y. (2000). Detection of change in persistence of a linear time series. *Journal of Econometrics* **95**, 97–116.

Kim, J.-Y. , Belaire-Franch, J. and Amador, R.B. (2002). Corrigendum to “Detection of change in persistence of a linear time series”. **109**, 389–392.

Perron, P. (2006). Dealing with structural breaks. In: Palgrave Handbook of Econometrics, Volume 1: Econometric Theory.

Berkes, I., Gombay, E., Horváth, L. and Kokoszka P. Sequential Change Point Detection in GARCH (p, q) Models. Econometric Theory, 2004.

Aue, A., Horváth,L, Hušková: On distinguishing between walk and change in the mean, 2009.

Horváth, L., Horváth, Z. and Hušková, M. Ratio tests for change point detection. Festschrift for P.K. Sen. (2008)

Aue, A., and Horváth, L. (2004) Delay time in sequential detection of change, *Statistics and Probability Letters* **67**, 221-231.

Aue, A., Horváth, L., Hušková, M., and Kokoszka, P. (2006) Change-point monitoring in linear models, *Econometrics Journal* **9**, 373-403.

Aue, A., Horváth, L., Kokoszka, P., and Steinebach, J. (2008) Monitoring shifts in mean: Asymptotic normality of stopping times, *Test* **17**, 515-530.

Aue, A., Horváth, L., and Reimherr, M.L. (2009) Delay times of sequential procedures for multiple time series regression models. *Journal of Econometrics* **149**, 174-190.

Berkes, I., Gombay, E., Horváth, L., and Kokoszka, P. (2004) Sequential change-point detection in GARCH(p, q) models, *Econometric Theory* **20**, 1140-1167.

Chu, C.-S., Stinchcombe, M., and White, H. (1996) Monitoring structural change, *Econometrica* **64**, 1045-1065.

Horváth, L., Hušková, M., Kokoszka, P., and Steinebach, J. (2004) Monitoring changes in linear models, *Journal of Statistical Planning and Inference* **126**, 225-251.

Horváth, L., Kühn, M., and Steinebach, J. (2008) On the performance of the fluctuation test for structural change, *Sequential Analysis* **27**, 126-140.

Hušková, M., Prášková, Z., and Steinebach, J. (2007) On the detection of changes in autoregressive time series, I. Asymptotics, *Journal of Statistical Planning and Inference* **137**, 1243-1259.

Hušková, M., Kirch, C., Prášková, Z., and Steinebach, J. (2008) On the detection of changes in autoregressive time series, II. Resampling procedures, *Journal of Statistical Planning and Inference* **138**, 1697-1721.

Hušková, M., and Koubková, A. (2005) Monitoring jump changes in linear models, *Journal of Statistical Research* **39**, 59-78.

Hušková, M., and Koubková, A. (2006) Sequential procedures for detection of changes in autoregressive sequences, Proceedings of *Prague Stochastics 2006*, MATFYZPRESS, Charles University Prague, 437-447.

Kirch, C. (2006) *Resampling Methods for the Change Analysis of Dependent Data*, PhD thesis, University of Cologne.

Kokoszka, P., and Leipus, R. (1998) Change-point in the mean of dependent observations, *Statistics and Probability Letters* **40**, 385-393.

Koubková, A. (2006a) Sequential change-point detection in linear models, Proceedings of *Prague Stochastics 2006*, MATFYZPRESS, Charles University Prague, 484-494.

Koubková, A. (2006b) *Sequential Change-Point Analysis*, PhD

Kvesic, M. (2006) *Strukturanalyse linearer Modelle*, Diploma thesis, University of Cologne.

Leisch, F., Hornik, K., and Kuan, C.-M. (2000) Monitoring structural changes with the generalized fluctuation test, *Econometric Theory* **16**, 835-854.

Schmitz, A. (2007) *Monitoring Changes in Dependent Data*, Diploma thesis, University of Cologne.

Zeileis, A., Leisch, F., Kleiber, C., and Hornik, K. (2005) Monitoring structural change in dynamic econometric models, *Journal of Applied Econometrics* **20**, 99-121.

Černíková, A., Hušková, M., Prášková, Z., Steinebach,J. Delay time in monitoring jump changes in linear models. Submitted (2009).

THANK YOU !!!!!