Abrupt change in mean avoiding variance estimation and block bootstrap

Barbora Peštová
The Czech Academy of Sciences
Institute of Computer Science

Motivation

- To know **whether a change has happened in some unknown time** is a task that is not only interesting, but also desirable for many scientific fields, e.g., in econometrics, biology, or climatology
- **Statistical hypothesis testing** is used for this detection purpose
- Sequences of dependent observations are **naturally ordered in time**
- Our approach to detect the unknown change lies in usage of so-called **ratio type test statistics**

Main objectives and aims

- Changes in the mean structures are studied, while random deviations from the mean structure are assumed to posses **common unknown variance**
- Using **ratio type test statistics** of the form

$$\max \frac{\max \mathit{Num}}{\max \mathit{Denom}}$$

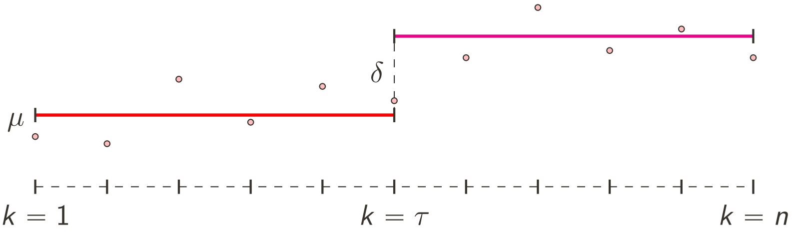
  Num and *Denom* are **functionals of residuals' partial sums**
- An advantage of the ratio type test statistics is **no need to estimate variability** of the underlying stochastic model
  - dependent random errors
  - even iid case under alternative
- A reasonable alternative to classical (non-ratio) statistics, when it is **difficult to find a suitable variance estimate**
- Proposed by Horváth et al. (2008)

Abrupt change in mean

- The **location model** for observations $Y_1, \ldots, Y_n$ with at most one abrupt change in mean

$$Y_k = \mu + \delta \, \mathcal{I}\{k > \tau\} + \varepsilon_k, \qquad k = 1, \ldots, n,$$

  where μ, $\delta \equiv \delta_n$ and $\tau \equiv \tau_n$ are unknown parameters
- τ is called the **change point**
- $\varepsilon_1, \ldots, \varepsilon_n$, we denote the **random errors** (possibly dependent)


μ
δ
$k = 1$
$k = \tau$
$k = n$

The null and the alternative

- Test the **null hypothesis** that **no change** occurred

  $$H_0 : \tau = n$$

- The **alternative** that **change occurred at some unknown time-point** τ

  $$H_1 : \tau < n, \; \delta \neq 0$$

- Ideas described in Horváth et al. (2008), Hušková (2007), and Hušková and Marušiaková (2012)

Ratio type test statistic based on M-residuals

$$\mathcal{R}_n(\psi,\gamma,\varphi) = \max_{n\gamma \leqslant k \leqslant n-n\gamma} \left(\frac{n-k}{k}\right)^{\varphi} \frac{\max_{1 \leqslant i \leqslant k} \left|\sum_{j=1}^{i} \psi(Y_j - \widehat{\mu}_{1k}(\psi))\right|}{\max_{k \leqslant i \leqslant n-1} \left|\sum_{j=i+1}^{n} \psi(Y_j - \widehat{\mu}_{2k}(\psi))\right|}$$

- $0 < \gamma < 1/2$ and $\varphi \in \mathbb{R}$ are **given constants**
- Considering different **score functions** ψ, we may construct similar statistics, but more **robust against outliers** and more suitable for **heavy-tailed distributions**
- $\widehat{\mu}_{1k}(\psi)$ is an ***M*-estimate** of parameter μ based on $Y_1, \ldots, Y_k$ and $\widehat{\mu}_{2k}(\psi)$ is an M-estimate of μ based on $Y_{k+1}, \ldots, Y_n$
- $\psi_{L_2}(x) = x$ and $\varphi = 0$ studied in Horváth et al. (2008)

Assumptions on the score function and errors

- Assumption 1: $\{\varepsilon_i\}_{i\in\mathbb{N}}$ form a **strictly stationary α-mixing sequence** with a distribution function F, that is **symmetric around zero** and for some $\chi > 0$, $\chi' > 0$ there exists a constant $C_1(\chi, \chi') > 0$ such that

$$\sum_{h=0}^{\infty} (h+1)^{\chi/2} \alpha(h)^{\chi'/(2+\chi+\chi')} \leqslant C_1(\chi, \chi')$$

  where $\alpha(k)$, $k = 0, 1, \ldots$ are the α-mixing coefficients
- Assumption 2: The **score function** ψ is a **non-decreasing** and **antisymmetric** function

Assumptions on the score function and errors

- Assumption 3:
$$\int |\psi(x)|^{2+\chi+\chi'} dF(x) < \infty$$
and for $|t_j| \leqslant C_3(\chi, \chi'),\ j = 1, 2$:
$$\int |\psi(x+t_2) - \psi(x+t_1)|^{2+\chi+\chi'} dF(x) \leqslant C_2(\chi, \chi')|t_2 - t_1|^{\eta}$$
where $1 \leqslant \eta \leqslant 2 + \chi + \chi',\ \chi > 0,\ \chi' > 0$
- Assumption 4: $\lambda(t) = -\int \psi(e-t)dF(e),\ t \in \mathbb{R}$ satisfies $\lambda(0) = 0$, $\lambda'(0) > 0$ and that $\lambda'(\cdot)$ is Lipschitz in the neighborhood of 0
- Assumption 5: **Long-run variance**
$$0 < \sigma^2(\psi) = \mathsf{E}\psi^2(\varepsilon_1) + 2\sum_{i=1}^{\infty} \mathsf{E}\psi(\varepsilon_1)\psi(\varepsilon_{i+1}) < \infty$$

Remarks on assumptions

- Assumption 1 is satisfied for example for **ARMA processes with continuously distributed stationary innovations** and bounded variance (Doukhan, 1994)
- The conditions regarding ψ reduce to moment restrictions for $\psi_{L_2}(x) = x$ (**L_2-method**)
- For $\psi_{L_1}(x) = \mathrm{sgn}(x)$ (**L_1-method**), the conditions reduce to F being a symmetric distribution and having continuous density f in a neighborhood of 0 with $f(0) > 0$
- We may consider the derivative of the **Huber loss function**

$$\psi_H(x) = x\,\mathcal{I}\{|x| \leqslant K\} + K\,\mathrm{sgn}(x)\,\mathcal{I}\{|x| > K\}$$

for some $K > 0$

Under the null

Theorem

Assumptions 1–5 and hypothesis H_0 hold. Then,

$$\mathcal{R}_n(\psi,\gamma,\varphi) \xrightarrow[n\to\infty]{\mathscr{D}} \left(\frac{1-\gamma}{\gamma}\right)^{|\varphi-1/2|} \frac{\sup\limits_{0\leqslant t\leqslant 1} |\mathcal{B}(t)|}{\sup\limits_{0\leqslant t\leqslant 1} |\mathcal{B}'(t)|},$$

where $\{\mathcal{B}(t), 0\leqslant t\leqslant 1\}$ and $\{\mathcal{B}'(t), 0\leqslant t\leqslant 1\}$ are independent Brownian bridges.

- The null hypothesis is **rejected for large values** of $\mathcal{R}_n(\psi,\gamma,\varphi)$
- Explicit form of the limit distribution is not known
- To obtain **critical values**: either a **simulation** from the limit distribution or **resampling** methods may be used

Under the alternative

Theorem

Assumptions 1–5 and alternative H_1 hold. If $\tau \equiv \tau_n = [n\zeta]$ for some $\zeta:\ \gamma < \zeta < 1-\gamma$ and $\sqrt{n}|\delta| \equiv \sqrt{n}|\delta_n| \to \infty$, then

$$\mathcal{R}_n(\psi,\gamma,\varphi) \xrightarrow[n\to\infty]{\mathsf{P}} \infty.$$

- Test statistic **explodes over all bounds** under the alternative
- **Consistency** $\Rightarrow$ the asymptotic distribution from Theorem "Under the null" can be used to construct the test

Resampling

- To avoid a simulation of the asymptotic distribution of the test statistic
- **Circular moving block bootstrap with replacement**
- **Overlapping blocks** of consequent observations are formed from the original observations
- The first few consequent observations from the original sequence are appended after the last observation
- For a sequence of length n, we always have n possible blocks of subsequent observations to choose from, cf. Kirch (2006)
- L – number of blocks, K – block length
- **Bootstrap version** of $\mathcal{R}_n(\psi, \gamma, \varphi)$ is defined as $\mathcal{R}^*_{L,K}(\psi, \gamma, \varphi)$

Validity of the bootstrap procedure

Theorem

Under some assumptions, as $L \to \infty$,

$$\mathrm{P}\left[\mathcal{R}^*_{L,K}(\psi,\gamma,\varphi) \leqslant y \middle| Y_1,\ldots,Y_n\right] \xrightarrow{\mathrm{P}} \mathrm{P}\left[\left(\frac{1-\gamma}{\gamma}\right)^{|\varphi-1/2|} \frac{\sup_{0\leqslant t\leqslant 1}|\mathcal{B}(t)|}{\sup_{0\leqslant t\leqslant 1}|\mathcal{B}'(t)|} \leqslant y\right].$$

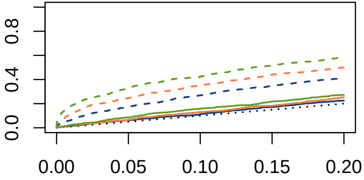
- $\mathcal{R}^*_{L,K}(\psi,\gamma,\varphi)$ provides **asymptotically correct critical values** for the test based on $\mathcal{R}_n(\psi,\gamma,\varphi)$, when observations follow **either the null or alternative**

Simulation scenarios

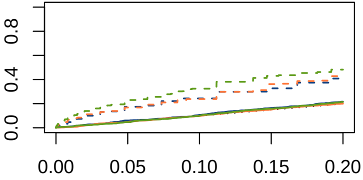
- Performance of the test based on test statistic $\mathcal{R}_n(\psi, \gamma, \varphi)$ with $\psi_{L_2}(x) = x$ and $\psi_{L_1}(x) = \mathrm{sgn}(x)$
- Comparison of the **circular moving block bootstrap** and the **simulation from the limit distribution**
- **Size-power** plots (Kirch, 2006)
- The ideal situation **under the null hypothesis** is depicted by the straight dotted line
- **Under the alternative**, the desired situation would be a steep function with values close to 1

Simulation scenarios (cont.)

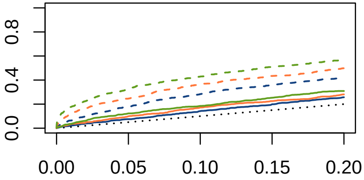
- 10000 independent samples generated to compute asymptotic critical values
- When bootstrapping, for each sample 1000 bootstrap samples used to compute bootstrap critical values
- 1000 repetitions in simulations of rejection rates
- $n = 200$, $\tau = n/2$, $\gamma = 0.1$, $\varphi = 0$, and $\delta = 1$
- **Errors** are **AR(1)** with coefficients 0.3 (red) and 0.5 (green), or **iid** (blue)
- **Innovations** are $\mathrm{N}(0, 1)$ and t_5
- **Rejection rates** based on **asymptotic** critical values ... dashed line, based on block **bootstrap** with block length $K = 5$ ... solid line



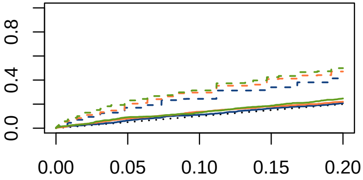
(a) L_2, N$(0, 1)$



(b) L_1, N$(0, 1)$



(c) L_2, t_5

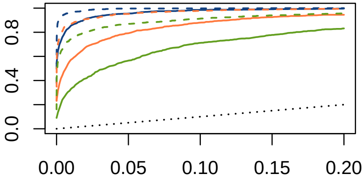


(d) L_1, t_5

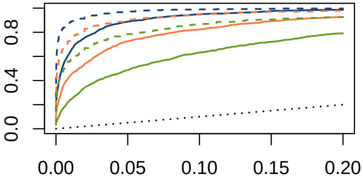
Figure: Null hypothesis, $n = 200$.

Simulation results under H_0

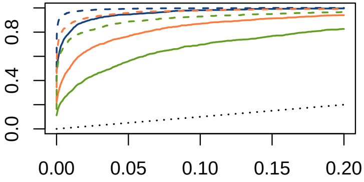
- Comparing to the critical values obtained by simulations from the asymptotic distribution, the critical values obtained by **bootstrapping are more accurate**, especially for the AR(1) sequences
- When comparing the accuracy of α-errors for different choices of the score function ψ, the L_1 **method seems to perform better** than the L_2 method
- With the choice of ψ_{L_2}, the simulated rejection rates under H_0 are higher than the corresponding theoretical α-levels for **larger values of the autoregression coefficient**, while for the L_1 method they remain much more stable



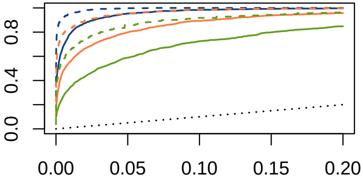
(a) L_2, $\mathrm{N}(0,1)$



(b) L_1, $\mathrm{N}(0,1)$



(c) L_2, t_5

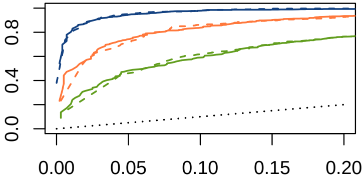


(d) L_1, t_5

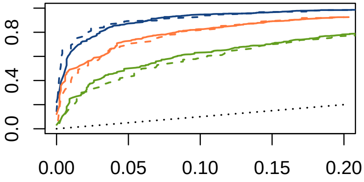
Figure: Alternative, $n = 200$.

Simulation findings under H_1

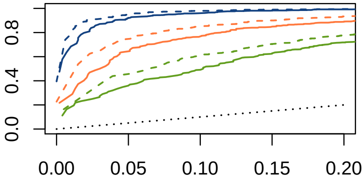
- L_1-method's **power** of the test **slightly decreases**
- Comparing the case of $\mathrm{N}(0,1)$ **innovations** with the case of t_5 innovations, the rejection rates for the L_1 version of the test statistic tend to be slightly higher for the t_5 distribution, while they remain more or less the same for the L_2 version
- Not demonstrated here:
  - As expected, the accuracy of the critical values tends to be better for **larger** n
  - $\gamma = 0.2$ seems to provide more accurate critical values than $\gamma = 0.1$, but the test power is larger in the latter case
  - With **larger abrupt change**, the power of the test increases



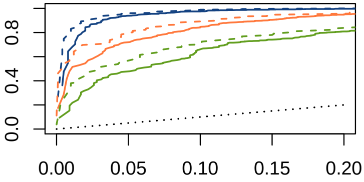
(a) L_2, $\mathrm{N}(0,1)$



(b) L_1, $\mathrm{N}(0,1)$



(c) L_2, t_5



(d) L_1, t_5

Figure: Empirical (adjusted) size-power plots, $n = 200$.

Adjusted α-errors

- The **empirical size-power plots** display the empirical size of the test (i.e., 1−sensitivity) on the *x*-axis versus the empirical power of the test (i.e., specificity) on the *y*-axis
- The ideal shape of the curve is **as steep as possible**
- The empirical size-power plots demonstrate that the bootstrap ratio type test statistic $\mathcal{R}^{*}_{L,K}(\psi,\gamma,\varphi)$ gives **approximately the same** (only slightly smaller) **empirical powers for the adjusted empirical sizes** comparing to the original test statistic $\mathcal{R}_{n}(\psi,\gamma,\varphi)$
- This is due to **two opposing facts**: $\mathcal{R}^{*}_{L,K}(\psi,\gamma,\varphi)$ keeps the significance level of the test better, but $\mathcal{R}_{n}(\psi,\gamma,\varphi)$ gives higher power of the test

Summary

- Abrupt change in mean model for sequences of time ordered observations, where the **mean can change at unknown time point**
- Testing procedures rely on **maximum ratio type statistics**
- The **main advantage** is that they provide an alternative to non-ratio type statistics in situations, in which **variance estimation is problematic** or cumbersome
- Asymptotic behavior of the test statistic is derived **under the null hypothesis** as well as **under alternatives**
- To **calculate critical values**, one can use **simulations and resampling methods**
- **Validity** of the block bootstrap procedure is shown

Thank you !

pestova@cs.cas.cz