Heteroscedasticity and Autocorrelation Robust Structural Change Detection

Zhou Zhou[1]

[1]Department of Statistics
University of Toronto

Argentina rainfall data

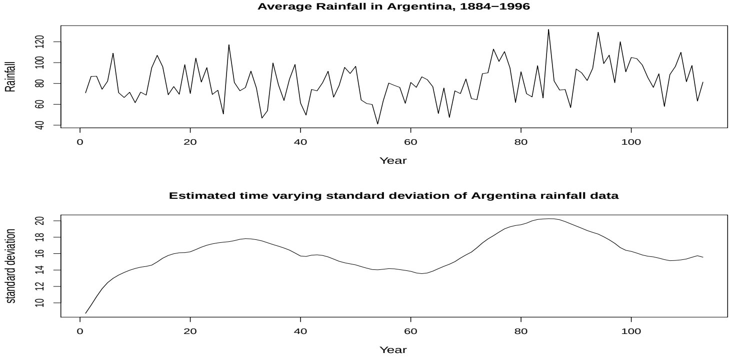

Figure : Upper panel: Time series plot of the Argentina rainfall data. Lower panel: Fitted marginal standard deviation of Argentina rainfall data.

Yearly rainfall in Tucumán Province, Argentina in millimeters from 1884 to 1996.

CUSUM test of change in mean.

Observe time series $\{X_i\}_{i=1}^n$ with $\mathbb{E}[X_i] = \mu_i$.

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_n = \mu, \quad \longleftrightarrow \quad H_a : \mu_i \neq \mu_j$ for some $1 \leq i < j \leq n$.

CUSUM test:

$$T_n = \max_{1 \leq i \leq n} |S_i - t_i S_n|, \text{ where } S_i = \sum_{j=1}^{i} X_j \text{ and } t_i = i/n. \qquad (2)$$

Piecewise locally stationary time series models

Definition

Write $X_i = \mu_i + e_i$. We say $\{e_i\}_{i=1}^n$ is piecewise locally stationary with r break points (PLS(r)) if there exist constants $0 = b_0 < b_1 < \cdots < b_r < b_{r+1} = 1$ and nonlinear filters $G_0, G_1, \cdots, G_r$, such that

$$e_i = G_j(t_i, \mathcal{F}_i), \text{ if } b_j < t_i \leq b_{j+1}, \tag{3}$$

where $t_i = i/n$, $\mathcal{F}_i = (\cdots, \varepsilon_0, \cdots, \varepsilon_{i-1}, \varepsilon_i)$ and ε_i's are i.i.d. random variables.

Piecewise locally stationary time series models

(i) (PLS linear processes)

$$G_k(t, \mathcal{F}_i) = \sum_{j=0}^{\infty} a_{k,j}(t)\varepsilon_{i-j} \quad b_k < t \le b_{k+1}, \tag{4}$$

where $a_{k,j}(\cdot)$'s are Lipschitz continuous functions.

(ii) (PLS nonlinear processes)

$$G_k(t, \mathcal{F}_i) = R_k(t, G_k(t, \mathcal{F}_{i-1}), \varepsilon_i), \quad b_k < t \le b_{k+1}, \tag{5}$$

where $R_k(t, \cdot, \cdot)$ is a smooth function of t.

Limiting behavior of CUSUM test for PLS time series

Define the long-run variance function

$$\sigma^2(t) = \sum_{k=-\infty}^{\infty} \mathrm{Cov}(G_i(t,\mathcal{F}_0), G_i(t,\mathcal{F}_k)) \text{ if } t \in (b_i, b_{i+1}].$$

Let $\sigma^2(0) = \lim_{t\downarrow 0} \sigma^2(t)$.

Theorem

Under mild conditions and the null hypothesis H_0, we have

$$T_n/\sqrt{n} \Rightarrow \sup_{0\le t\le 1} |U(t) - tU(1)|, \quad (6)$$

where $\Rightarrow$ stands for convergence in distribution and $U(t)$ is a zero-mean Gaussian process with covariance function $\gamma(t,s) = \int_0^{\min(t,s)} \sigma^2(r)\,dr$.

Behavior of traditional testing procedures under PLS

Lagged window estimate of long-run variance:

$$\hat{\sigma}^2_{LW} = \frac{m}{n-m+1} \sum_{j=1}^{n-m+1} \left(\frac{1}{m} \sum_{i=j}^{j+m-1} X_i - \frac{1}{n} S_n \right)^2. \quad (7)$$

Theorem

Under H_0, the assumption that $m \to \infty$ with $m/n \to 0$ and some mild conditions, we have

$$\hat{\sigma}^2_{LW} \to \int_0^1 \sigma^2(t)\, dt \text{ in probability.}$$

$$T_n/(\hat{\sigma}_{LW}\sqrt{n}) \Rightarrow \sup_{0 \le t \le 1} |U(t) - tU(1)| / \sqrt{\int_0^1 \sigma^2(t)\, dt}. \quad (8)$$

The robust bootstrap

for a fixed window size m, define the process

$$\Phi_{i,m} = \sum_{j=1}^{i} \frac{1}{\sqrt{m(n-m+1)}} (S_{j,m} - \frac{m}{n} S_n) R_j, \qquad i = 1, \cdots, n-m+1,$$

where $S_{j,m} = \sum_{r=j}^{j+m-1} X_r$ and $(R_i)_{i=1}^n$ are iid standard normal and are independent of $(X_i)_{i=1}^n$. On $[1/n, (n-m+1)/n]$, define the associated linear interpolation

$$\tilde{\Phi}_{m,n}(t) = \Phi_{t_* n,m} + n(t - t_*)(\Phi_{t^* n,m} - \Phi_{t_* n,m}),$$

where $t_* = \lfloor tn \rfloor / n$ and $t^* = t_* + 1/n$.

The robust bootstrap

Theorem

Assume that H_0 holds and $m \to \infty$ with $m/n \to 0$. Then conditional on $(X_i)_{i=1}^n$ and under some mild conditions, we have

$$\tilde{\Phi}_{m,n}(t) \Rightarrow U(t) \text{ on } \mathcal{C}(0,1) \text{ with the uniform topology.} \tag{9}$$

And, conditional on $(X_i)_{i=1}^n$,

$$\max_{m+1 \le i \le n-m+1} \left| \Phi_{i,m} - \frac{i}{n-m+1} \Phi_{n-m+1,m} \right| \Rightarrow \sup_{0 \le t \le 1} |U(t) - tU(1)|. \tag{10}$$

Implementation

- Select the window size m.

Implementation

- Select the window size m.
- Generate B (say 2000) conditionally i.i.d. copies $\{\Phi_{i,m}^{(r)}\}_{i=1}^{n-m+1}$, $r = 1, 2, \cdots, B$. Let $M_r = \max_{m+1 \le i \le n-m+1} |\Phi_{i,m}^{(r)} - \frac{i}{n-m+1} \Phi_{n-m+1,m}^{(r)}|$.

Implementation

- Select the window size m.
- Generate B (say 2000) conditionally i.i.d. copies $\{\Phi_{i,m}^{(r)}\}_{i=1}^{n-m+1}$, $r = 1, 2, \cdots, B$. Let $M_r = \max_{m+1 \le i \le n-m+1} |\Phi_{i,m}^{(r)} - \frac{i}{n-m+1}\Phi_{n-m+1,m}^{(r)}|$.
- Let $M_{(1)} \le M_{(2)} \le \cdots \le M_{(B)}$ be the ordered statistics of M_r $r = 1, 2, \cdots B$. Reject H_0 at level α if $T_n/\sqrt{n} > M_{(\lfloor B(1-\alpha) \rfloor)}$, where $\lfloor x \rfloor$ denotes the largest integer smaller or equal to x. Let $B^* = \max\{r : M_{(r)} \le T_n/\sqrt{n}\}$. The p-value of the test can be obtained by $1 - B^*/B$.

Accuracy of the robust bootstrap

Theorem

Under mild conditions, we have

$$\max_{i,j} \|Cov(\Phi_{i,m}, \Phi_{j,m} | \{X_i\}_{i=1}^n) - Cov(U(t_i), U(t_j))\| = O(\sqrt{\frac{m}{n}} + \frac{1}{m}). \quad (11)$$

Consequently the optimal window size $m = O(n^{1/3})$ *with optimal uniform MSE*$= O(n^{-1/3})$.

Power of the robust bootstrap

Theorem

Assume that $m \to \infty$ with $m/n \to 0$ and conditional on $(X_i)_{i=1}^n$. Then under the local alternative that $\mu(\cdot) = n^{-1/2} f(\cdot)$, where $f(\cdot)$ is a nonconstant piecewise Lipschitz continuous function, we have

$$\max_{m+1 \le i \le n-m+1} \left| \Phi_{i,m} - \frac{i}{n-m+1} \Phi_{n-m+1,m} \right| \Rightarrow \sup_{0 \le t \le 1} |U(t) - tU(1)|. \quad (12)$$

On the other hand, note that, under the above local alternative

$$T_n/\sqrt{n} \Rightarrow \sup_{0 \le t \le 1} |U(t) - tU(1) + \int_0^t f(s)\, ds - t \int_0^1 f(s)\, ds|.$$

Extensions to multidimensional statistics of PLS time series

Consider testing structure changes of

$\eta_i = \mathbb{E}\phi(\mathbf{X}_i, \mathbf{X}_{i-1}, \cdots, \mathbf{X}_{i-q})$, where $\{\mathbf{X}_i\}$ is a d-dimensional PLS process(

and $\phi : \mathbb{R}^{qd} \to \mathbb{R}^p$.
CUSUM test statistic

$$T_n(\phi) = \max_{1+q \le i \le n} |\mathbf{S}_i(\phi) - t_i \mathbf{S}_n(\phi)|, \text{ with } \mathbf{S}_i(\phi) = \sum_{j=q+1}^{i} \phi(\mathbf{X}_j, \mathbf{X}_{j-1}, \cdots, \mathbf{X}_{j-q})$$

Write $\mathbf{Y}_i = \phi(\mathbf{X}_i, \mathbf{X}_{i-1}, \cdots, \mathbf{X}_{i-q})$. Then

$$\mathbf{Y}_i = \eta_i + \mathbf{W}_j[t_i, (\cdots, \varepsilon_{i-1}, \varepsilon_i)], \text{ if } b_j < t_i \le b_{j+1}, \tag{14}$$

Write $\Sigma^2(t) = \sum_{k=-\infty}^{\infty} \mathrm{Cov}(\mathbf{W}_i(t, \mathcal{F}_0), \mathbf{W}_i(t, \mathcal{F}_k))$ if $t \in (b_i, b_{i+1}]$.

Extensions to multidimensional statistics of PLS time series

Theorem

i) : Under $H_0, T_n(\phi)/\sqrt{n} \Rightarrow \sup_{0\le t\le 1} |\mathbf{U}(t) - t\mathbf{U}(1)|$, *where* $\mathbf{U}(t)$ *is a p-dimensional zero-mean Gaussian process with covariance function* $\gamma(t,s) = \int_0^{\min(t,s)} \Sigma^2(r)\,dr$. *ii) : Let*

$$\Phi_{i,m}(\phi) = \sum_{j=q+1}^{i} \frac{1}{\sqrt{m(n-m-q+1)}}(\mathbf{S}_{j,m}(\phi) - \frac{m}{n}\mathbf{S}_n(\phi))R_j,$$

where $\mathbf{S}_{j,m}(\phi) = \sum_{r=j}^{j+m-1} \mathbf{Y}_r$ *and* $(R_i)_{i=1}^n$ *are i.i.d.* $N(0,1)$ *independent of* $(\mathbf{X}_i)_{i=1}^n$. *Under* H_0 *and conditional on* $(\mathbf{X}_i)_{i=1}^n$, *we have*

$$\max_{q+1\le i\le n-m+1} \left|\Phi_{i,m}(\phi) - \frac{i}{n-m+1}\Phi_{n-m+1,m}(\phi)\right| \Rightarrow \sup_{0\le t\le 1} |\mathbf{U}(t) - t\mathbf{U}(1)|.$$

Simulation: accuracy under stationarity

Scenarios: (a). AR(1) model with AR-coefficient *a*. (b). ARMA(1,1) model with AR-coefficient *a* and MA coefficient 0.5. (c). ARMA(1,1) model with AR-coefficient *a* and MA coefficient -0.6.

		$\alpha = 0.05$					$\alpha = 0.1$				
Case	*a*	LW1	LW2	SN	RB1	RB2	LW1	LW2	SN	RB1	RB2
(a)	0.2	6.7	4.1	5.6	7.0	3.9	13.5	9.7	10.1	13.9	9.3
(a)	0.5	5.8	4.6	5.9	5.1	3.4	13.1	11.0	11.2	11.6	9.6
(a)	0.8	3.5	3.0	5.6	8.3	2.3	10.2	9.3	14.4	8.5	7.5
(b)	0.2	4.6	3.4	5.4	4.7	3.2	10.5	9.5	10.7	10.6	8.9
(b)	0.5	4.5	4.2	6.4	3.4	3.1	11.3	11.2	12.1	9.5	9.3
(b)	0.8	2.3	2.9	9.4	1.9	2.3	8.8	9.6	14.9	7.2	7.9
(c)	0.2	0.4	1.0	2.3	0.5	1.2	1.4	3.9	5.4	1.5	4.1
(c)	0.5	1.3	3.1	4.4	1.3	2.5	3.3	9.0	8.3	3.4	7.7
(c)	0.8	19.3	6.2	8.2	19.4	5.8	29.4	14.1	14.0	29.7	13.0

Table 1. Simulated type I error rates (in %) for the five methods with nominal levels 5% and 10% under stationary models (a), (b) and (c). Series length $n = 200$.

Simulation: accuracy under stationarity

		$\alpha = 0.05$					$\alpha = 0.1$				
Case	a	LW1	LW2	SN	RB1	RB2	LW1	LW2	SN	RB1	RB2
(a)	0.2	6.6	5.8	5.1	6.5	5.6	11.8	11.0	9.4	12.2	11.2
(a)	0.5	6.9	6.5	6.0	6.1	5.9	12.9	12.4	11.0	12.3	11.8
(a)	0.8	5.2	6.0	6.2	4.5	5.2	11.4	12.6	10.9	10.0	11.3
(b)	0.2	4.7	4.7	4.8	4.4	4.3	9.8	9.7	9.4	9.4	9.3
(b)	0.5	5.2	5.0	5.7	4.6	4.6	10.9	10.8	10.8	10.0	9.8
(b)	0.8	4.3	6.3	6.8	3.3	5.3	10.5	12.6	11.6	9.0	11.6
(c)	0.2	0.6	1.5	4.0	0.5	1.4	1.8	4.5	7.6	1.8	4.6
(c)	0.5	1.6	3.8	5.5	1.7	3.4	4.2	9.0	9.7	4.3	8.6
(c)	0.8	18.9	6.8	6.0	19.0	6.0	29.6	13.8	11.2	29.7	12.9

Table 2. Simulated type I error rates (in %) for the five methods with nominal levels 5% and 10% under stationary models (a), (b) and (c). Series length $n = 500$.

Simulation: accuracy under non-stationarity

- Case (I):

  $X_i = V(t_i)X_i^*, \qquad i = 1, 2, \cdots, n,$ where $V(t) = 1 + 4I\{t > 0.75\}$,

  $\{X_i^*\}$ is a zero-mean AR(1) process with AR coefficient 0.5.

Simulation: accuracy under non-stationarity

- Case (I):

  $$X_i = V(t_i)X_i^*, \qquad i = 1, 2, \cdots, n, \text{ where } V(t) = 1 + 4I\{t > 0.75\},$$

  $\{X_i^*\}$ is a zero-mean AR(1) process with AR coefficient 0.5.
- Case (II): $X_i = G_0(t_i, \mathcal{F}_i)$ for $t_i \leq 1/3$ and $X_i = G_1(t_i, \mathcal{F}_i)$ for $t_i > 1/3$, where

  $$G_0(t, \mathcal{F}_i) = 0.5G_0(t, \mathcal{F}_{i-1}) + \varepsilon_i, \qquad G_1(t, \mathcal{F}_i) = -0.5G_1(t, \mathcal{F}_{i-1}) + \varepsilon_i,$$

  and ε_i's are i.i.d. $N(0, 1)$.

Simulation: accuracy under non-stationarity

- Case (I):

  $$X_i = V(t_i)X_i^*, \qquad i = 1, 2, \cdots, n, \text{ where } V(t) = 1 + 4I\{t > 0.75\},$$

  $\{X_i^*\}$ is a zero-mean AR(1) process with AR coefficient 0.5.
- Case (II): $X_i = G_0(t_i, \mathcal{F}_i)$ for $t_i \leq 1/3$ and $X_i = G_1(t_i, \mathcal{F}_i)$ for $t_i > 1/3$, where

  $$G_0(t, \mathcal{F}_i) = 0.5G_0(t, \mathcal{F}_{i-1}) + \varepsilon_i, \qquad G_1(t, \mathcal{F}_i) = -0.5G_1(t, \mathcal{F}_{i-1}) + \varepsilon_i,$$

  and ε_i's are i.i.d. $N(0, 1)$.
- Case (III): $X_i = G_0(t_i, \mathcal{F}_i)$, where

  $$G_0(t, \mathcal{F}_i) = 0.75\cos(2\pi t)G_0(t, \mathcal{F}_{i-1}) + \varepsilon_i,$$

  and ε_i's are i.i.d. $N(0, 1)$.

Simulation: accuracy under non-stationarity

- Case (I):

  $$X_i = V(t_i)X_i^*, \qquad i = 1, 2, \cdots, n, \text{ where } V(t) = 1 + 4I\{t > 0.75\},$$

  $\{X_i^*\}$ is a zero-mean AR(1) process with AR coefficient 0.5.
- Case (II): $X_i = G_0(t_i, \mathcal{F}_i)$ for $t_i \leq 1/3$ and $X_i = G_1(t_i, \mathcal{F}_i)$ for $t_i > 1/3$, where

  $$G_0(t, \mathcal{F}_i) = 0.5G_0(t, \mathcal{F}_{i-1}) + \varepsilon_i, \qquad G_1(t, \mathcal{F}_i) = -0.5G_1(t, \mathcal{F}_{i-1}) + \varepsilon_i,$$

  and ε_i's are i.i.d. $N(0, 1)$.
- Case (III): $X_i = G_0(t_i, \mathcal{F}_i)$, where

  $$G_0(t, \mathcal{F}_i) = 0.75\cos(2\pi t)G_0(t, \mathcal{F}_{i-1}) + \varepsilon_i,$$

  and ε_i's are i.i.d. $N(0, 1)$.
- Case (IV): $X_i = G_0(t_i, \mathcal{F}_i)$ for $t_i \leq 0.8$ and $X_i = G_1(t_i, \mathcal{F}_i)$ for $t_i > 0.8$, where

  $$G_0(t, \mathcal{F}_i) = 0.75\cos(2\pi t)G_0(t, \mathcal{F}_{i-1}) + \varepsilon_i, \; G_1(t, \mathcal{F}_i) = (0.5 - t)G_1(t, \mathcal{F}_{i-1}) + \varepsilon_i$$

  and ε_i's are i.i.d. $N(0, 1)$.

Simulation: accuracy under non-stationarity

	$\alpha = 0.05$					$\alpha = 0.1$				
Model	LW1	LW2	SN	RB1	RB2	LW1	LW2	SN	RB1	RB2
	$n = 200$									
I	18.4	16.9	22.9	3.9	3.5	29.4	29.0	30.8	10.5	10.3
II	17.6	9.2	12.0	10.7	3.7	25.1	19.1	19.3	19.8	10.4
III	38.5	12.9	13.1	33.1	3.7	51.0	26.7	21.4	46.1	12.1
IV	22.0	16.0	21.6	14.7	5.9	30.9	28.7	30.1	24.9	14.3
	$n = 500$									
I	20.8	21.5	23.0	6.4	6.5	29.4	29.9	30.7	12.1	12.0
II	14.6	9.3	13.0	9.7	4.2	23.4	18.4	20.1	17.8	10.5
III	31.4	13.2	11.9	21.9	5.9	43.5	24.0	19.2	37.9	12.4
IV	27.4	19.1	22.4	16.3	5.8	37.2	29.3	30.5	26.0	13.5

Table 3. Simulated type I error rates (in %) for the five methods with nominal levels 5% and 10% under non-stationary models I - IV. Series lengths $n = 200$ and 500.

Power comparison

(A) :

$$e_i = 0.5e_{i-1} + \varepsilon_i \text{ where } \varepsilon_i \sim \text{ i.i.d. } N(0,1).$$

(B) : $\{e_i\}$ follow the following PLS AR(1) model $e_i = G_0(t_i, \mathcal{F}_i)$ for $t_i \le 1/2$ and $e_i = G_1(t_i, \mathcal{F}_i)$ for $t_i > 1/2$, where

$$G_0(t, \mathcal{F}_i) = 0.75\cos(2\pi t)G_0(t, \mathcal{F}_{i-1}) + \varepsilon_i, \; G_1(t, \mathcal{F}_i) = (0.5 - t)G_1(t, \mathcal{F}_{i-1}) + \varepsilon_i.$$

Alternative:

$$\mu_i = \delta[t_i I\{0 \le t_i \le 0.5\} + (t_i - 1)I\{0.5 < t_i \le 1\}],$$

Power comparison

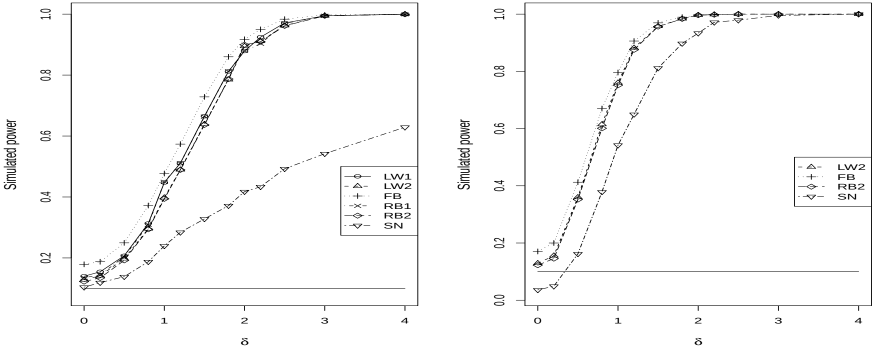

Figure : Left panel: simulated rejection rates for Case (A). Right panel: simulated rejection rates for Case (B). The horizontal lines represent the nominal level $y = 0.1$.

Argentina rainfall data

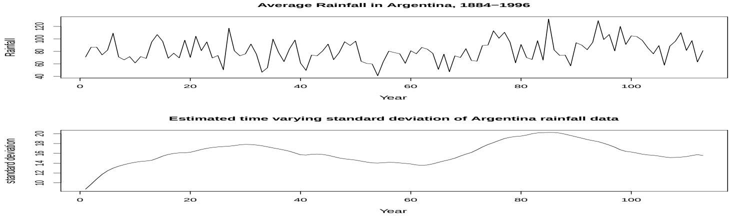

Figure : Upper panel: Time series plot of the Argentina rainfall data. Lower panel: Fitted marginal standard deviation of Argentina rainfall data.

Yearly rainfall in Tucumán Province, Argentina in millimeters from 1884 to 1996.
Classic long-run variance normalization: p-value $< 0.1\%$
Robust bootstrap: p-value $= 2\%$.

Interest rate data

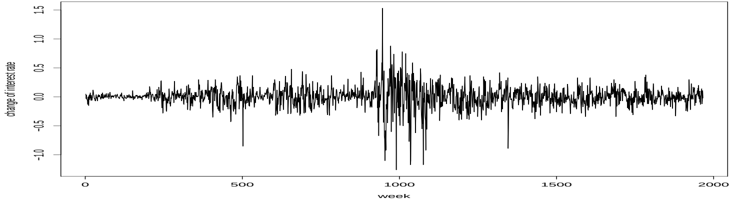

Figure : Change of weekly U.S. 1-year Treasury constant maturity rate from January 4, 1962 to September 10, 1999.

Classic long-run variance normalization for mean: p-value 12%
Robust bootstrap for mean: p-value 22%.
Classic long-run variance normalization for ACVF1: p-value 13%
Robust bootstrap for ACVF1: p-value 18%.

REFERENCES

Zhou, Z. (2013). Heteroscedasticity and autocorrelation robust structural change detection. *Journal of the American Statistical Association*, **108** 726-740.

Thank you!