Testing for structural changes in LMSV time series

Annika Betken (joint work with Rafał Kulik)

Ruhr-University Bochum

Oaxaca, June 18th 2018
Self-Similarity, Long-Range Dependence and Extremes

Research supported by the *German Academic Scholarship Foundation* and *Collaborative Research Center SFB 823*.

Outline

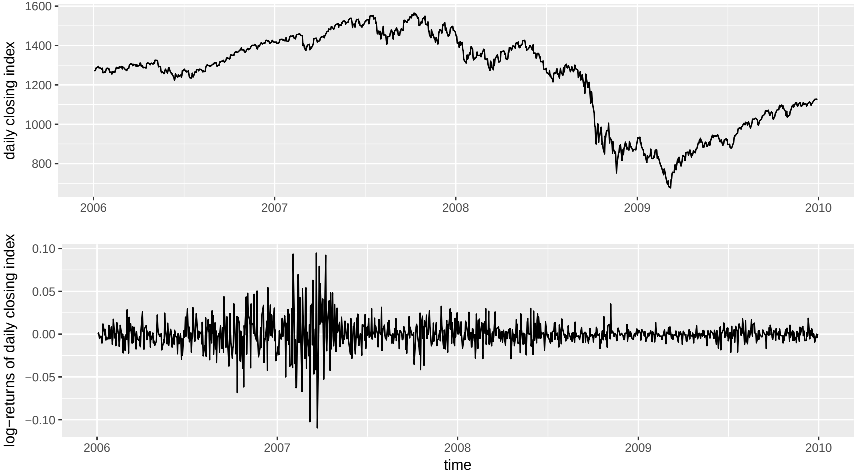
1. **L**ong **M**emory **S**tochastic **V**olatility model
   - Data example
   - Model assumptions
2. Change-point problem
   - CUSUM-based change-point tests

     Partial sum process

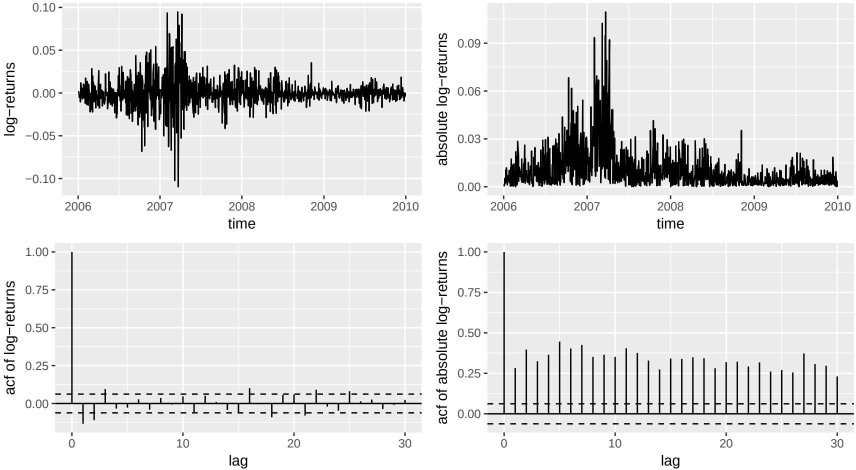
     Self-normalized CUSUM test
   - Wilcoxon-based change-point tests

     Two-parameter empirical process

     Self-normalized Wilcoxon test



Daily closing index of Standard & Poor's 500 and its log-returns from January 2006 to December 2009.



Log-returns and absolute log-returns of Standard & Poor's 500 daily closing index from January 2006 to December 2009 and its sample autocorrelations.

Long Memory Stochastic Volatility model

Stochastic **V**olatility model: time series X_j, $j \geq 1$,

$$X_j = \sigma(Y_j)\varepsilon_j, \quad j \geq 1 ,$$

where

- ε_j, $j \geq 1$, is an i.i.d. sequence with $\operatorname{E}\varepsilon_1 = 0$;
- σ is a non-negative measurable function;
- Y_j, $j \geq 1$, is a stationary Gaussian long-memory process;

Long **M**emory:

$$Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k} , \quad \sum_{k=1}^{\infty} c_k^2 = 1 ,$$

for an i.i.d. Gaussian sequence η_j, $j \in \mathbb{Z}$, with $\operatorname{E}\eta_1 = 0$, $\operatorname{Var}\eta_1 = 1$, and

$$\gamma_Y(k) = \operatorname{Cov}(Y_j, Y_{j+k}) = k^{-D} L_\gamma(k),$$

where $D \in (0, 1)$ and L_γ is slowly varying at ∞.

Given observations

$$X_j = \sigma(Y_j)\varepsilon_j, \quad Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k}, \ j \geq 1 \ .$$

LMSV model: Assume that $\{\varepsilon_j, j \geq 1\}$ and $\{\eta_j, j \in \mathbb{Z}\}$ are mutually independent.
(Introduced in: *Breidt, Crato, de Lima (1998)*, *Harvey (2002)*.)

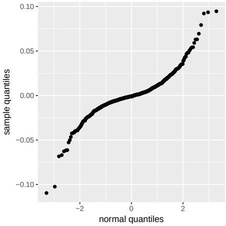
Observe that

$$\begin{aligned} &\mathrm{Cov}(X_1, X_{k+1}) = 0, \ k \geq 1, \\ &\mathrm{Cov}(|X_1|, |X_{k+1}|) = (\mathrm{E}\,|\varepsilon_1|)^2 \,\mathrm{Cov}(|\sigma(Y_1)|, |\sigma(Y_{k+1})|), \end{aligned}$$

i.e. the X_j, $j \geq 1$, are uncorrelated, while the absolute values of the variables inherit the dependence structure from Y_j, $j \geq 1$.

LMSV model with leverage: Assume that $\{(\varepsilon_j, \eta_j), j \geq 1\}$ is a sequence of i.i.d. vectors.
Then, Y_i and ε_i are independent for fixed i, but Y_i may not be independent of ε_j, $j < i$.

Heavy tails:



Q-Q plot for the log-returns of Standard & Poor's 500 daily closing index from January 2006 to December 2009.

Given observations

$$X_j = \sigma(Y_j)\varepsilon_j, \quad Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k}, \ j \geq 1 \ .$$

Assumption: $P(\varepsilon_1 > x) = x^{-\alpha} L(x)$ for some $\alpha > 0$ and a slowly varying function L.

Breiman's Lemma: If $\mathrm{E}\,\sigma^{\alpha+\delta}(Y_1) < \infty$ for some $\delta > 0$, then

$$P(X_1 > x) \sim \mathrm{E}\,\sigma^{\alpha}(Y_1) P(\varepsilon_1 > x), \text{ as } x \to \infty.$$

Change-point problem

Given: observations $X_1, \ldots, X_n$ and a measurable function ψ, consider $Z_i = \psi(X_i)$, $i = 1, \ldots n$.

Testing problem:

$$\mathbf{H_0} : \operatorname{E} Z_1 = \ldots = \operatorname{E} Z_n$$

against

$$\mathbf{H_1} : \operatorname{E} Z_1 = \ldots = \operatorname{E} Z_k \neq \operatorname{E} Z_{k+1} = \ldots = \operatorname{E} Z_n$$
$$\text{for some } k \in \{1, \ldots, n-1\}.$$

Change-point problem

Given: observations $X_1, \ldots, X_n$ and a measurable function ψ, consider $Z_i = \psi(X_i)$, $i = 1, \ldots n$.

Testing problem:

$$\mathbf{H_0} : \operatorname{E} Z_1 = \ldots = \operatorname{E} Z_n$$

against

$$\mathbf{H_1} : \operatorname{E} Z_1 = \ldots = \operatorname{E} Z_k \neq \operatorname{E} Z_{k+1} = \ldots = \operatorname{E} Z_n$$
$$\text{for some } k \in \{1, \ldots, n-1\}.$$

Examples:

- $\psi(x) = x$ in order to detect changes in the mean;
- $\psi(x) = x^2$ in order to detect changes in the variance.

CUSUM-based change-point tests

CUSUM change-point test: rejects for large values of

$$C_n = \max_{1 \leq k < n} C_n(k), \text{ where } C_n(k) = \left| \sum_{i=1}^{k} Z_i - \frac{k}{n} \sum_{i=1}^{n} Z_i \right|.$$

Observe that

$$C_n(k) = \left| \sum_{i=1}^{k} (Z_i - \operatorname{E} Z_1) - \frac{k}{n} \sum_{i=1}^{n} (Z_i - \operatorname{E} Z_1) \right|.$$

CUSUM-based change-point tests

CUSUM change-point test: rejects for large values of

$$C_n = \max_{1 \leq k < n} C_n(k), \text{ where } C_n(k) = \left| \sum_{i=1}^{k} Z_i - \frac{k}{n} \sum_{i=1}^{n} Z_i \right|.$$

Observe that

$$C_n(k) = \left| \sum_{i=1}^{k} (Z_i - \operatorname{E} Z_1) - \frac{k}{n} \sum_{i=1}^{n} (Z_i - \operatorname{E} Z_1) \right|.$$

$\Rightarrow$ Consider the partial sum process

$$S_n(t) := \sum_{j=1}^{\lfloor nt \rfloor} (Z_j - \operatorname{E} Z_1), \ t \in [0, 1].$$

Partial sum process:

$$S_n(t) := \sum_{j=1}^{\lfloor nt \rfloor} (Z_j - \mathrm{E}\, Z_1)\,,\ t \in [0,1],$$

$$Z_j = \psi(X_j),\ X_j = \sigma(Y_j)\varepsilon_j,\quad Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k},\ j \geq 1\ .$$

Theorem (Beran, Feng, Ghosh, Kulik (2013))

Let $\mathcal{F}_j = \sigma(\varepsilon_j, \varepsilon_{j-1}, \varepsilon_{j-2}, \ldots, \eta_j, \eta_{j-1}, \eta_{j-2}, \ldots)$. Suppose X_n, $n \in \mathbb{N}$, is an LMSV time series and let ψ be a measurable function. Assume that $\nu^2 = \mathrm{E}(Z_1^2) < \infty$. Given the previous assumptions (+ technical assumptions), under $\mathbf{H_0}$,

- *if $\mathrm{E}(Z_1 \mid \mathcal{F}_0) = 0$, then*

$$n^{-\frac{1}{2}} S_n(t) \xrightarrow{\mathcal{D}} \nu B_{\frac{1}{2}}(t)$$

- *if $\mathrm{E}(Z_1 \mid \mathcal{F}_0) \neq 0$, then*

$$n^{\frac{D}{2}-1} L_\gamma^{-\frac{1}{2}}(n) S_n(t) \xrightarrow{\mathcal{D}} C_{\psi,\sigma,D} B_H(t)$$

in $D[0,1]$, where B_H denotes a fractional Brownian motion, $H = 1 - \frac{D}{2}$, and $C_{\psi,\sigma,D}$ is an unknown constant.

CUSUM change-point test: rejects for large values of

$$C_n = \max_{1 \le k < n} C_n(k), \text{ where } C_n(k) = \left| \sum_{i=1}^{k} Z_i - \frac{k}{n} \sum_{i=1}^{n} Z_i \right|.$$

Asymptotic distribution of the CUSUM test statistic:

Corollary (B., Kulik (2017))

Suppose X_n, $n \in \mathbb{N}$, is an LMSV time series and let ψ be a measurable function. Assume that $\nu^2 = \mathrm{E}(Z_1^2) < \infty$. Given the previous assumptions (+ technical assumptions), under $\mathbf{H_0}$*,*

- *if* $\mathrm{E}(Z_1 \mid \mathcal{F}_0) = 0$*, then*

$$n^{-\frac{1}{2}} \max_{1 \le k < n} C_n(k) \xrightarrow{\mathcal{D}} \nu \sup_{0 \le t \le 1} \left| B_{\frac{1}{2}}(t) - tB_{\frac{1}{2}}(1) \right| ;$$

- *if* $\mathrm{E}(Z_1 \mid \mathcal{F}_0) \neq 0$*, then*

$$n^{\frac{D}{2}-1} L_\gamma^{-\frac{1}{2}}(n) \max_{1 \le k < n} C_n(k) \xrightarrow{\mathcal{D}} |C_{\psi,\sigma,D}| \sup_{0 \le t \le 1} |B_H(t) - tB_H(1)| .$$

Recall that

$$X_j = \sigma(Y_j)\varepsilon_j, \quad j \geq 1 \,,$$

where ε_j, $j \geq 1$, is an i.i.d. sequence with $\operatorname{E}\varepsilon_1 = 0$, $\operatorname{Var}\varepsilon_1 = 1$ and $Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k}$ for an i.i.d. sequence η_j, $j \in \mathbb{Z}$ and that $\mathcal{F}_j = \sigma\left(\varepsilon_j, \varepsilon_{j-1}, \varepsilon_{j-2}, \ldots, \eta_j, \eta_{j-1}, \eta_{j-2}, \ldots\right)$.

$\Rightarrow$ ε_j is independent of $\mathcal{F}_{j-1}$ and Y_j is $\mathcal{F}_{j-1}$-measurable.

- **Change in mean:** $\psi(x) = x \Rightarrow \operatorname{E}(\psi(X_1) \mid \mathcal{F}_0) = \sigma(Y_1)\operatorname{E}(\varepsilon_1) = 0$.
  Then
  $$n^{-\frac{1}{2}} \max_{1 \leq k < n} C_n(k) \xrightarrow{\mathcal{D}} \nu \sup_{0 \leq t \leq 1} \left| B_{\frac{1}{2}}(t) - t B_{\frac{1}{2}}(1) \right| \,.$$

- **Change in variance:** $\psi(x) = x^2 \Rightarrow \operatorname{E}(\psi(X_1) \mid \mathcal{F}_0) = \sigma^2(Y_1)\operatorname{E}(\varepsilon_1^2) \neq 0$.
  Then
  $$n^{\frac{D}{2}-1} L_\gamma^{-\frac{1}{2}}(n) \max_{1 \leq k < n} C_n(k) \xrightarrow{\mathcal{D}} |C_{\psi,\sigma,D}| \sup_{0 \leq t \leq 1} |B_H(t) - t B_H(1)| \,.$$

CUSUM change-point test: rejects for large values of

$$C_n = \max_{1 \le k < n} C_n(k), \text{ where } C_n(k) = \left| \sum_{i=1}^{k} Z_i - \frac{k}{n} \sum_{i=1}^{n} Z_i \right| .$$

Asymptotic distribution of the CUSUM test statistic:

Corollary (B., Kulik (2017))

Given the asumptions of the previous theorem, under $\mathbf{H_0}$*,*

- *if* $\mathrm{E}(Z_1 \mid \mathcal{F}_0) = 0$*, then*

$$n^{-\frac{1}{2}} \max_{1 \le k < n} C_n(k) \xrightarrow{\mathcal{D}} \nu \sup_{0 \le t \le 1} \left| B_{\frac{1}{2}}(t) - tB_{\frac{1}{2}}(1) \right| ;$$

- *if* $\mathrm{E}(Z_1 \mid \mathcal{F}_0) \neq 0$*, then*

$$n^{\frac{D}{2}-1} L_{\gamma}^{-\frac{1}{2}}(n) \max_{1 \le k < n} C_n(k) \xrightarrow{\mathcal{D}} |C_{\psi,\sigma,D}| \sup_{0 \le t \le 1} |B_H(t) - tB_H(1)| .$$

CUSUM change-point test: rejects for large values of

$$C_n = \max_{1 \le k < n} C_n(k), \text{ where } C_n(k) = \left| \sum_{i=1}^{k} Z_i - \frac{k}{n} \sum_{i=1}^{n} Z_i \right|.$$

Asymptotic distribution of the CUSUM test statistic:

Corollary (B., Kulik (2017))

Given the asumptions of the previous theorem, under $\mathbf{H_0}$*,*

– *if* $\mathrm{E}(Z_1 \mid \mathcal{F}_0) = 0$*, then*

$$n^{-\frac{1}{2}} \max_{1 \le k < n} C_n(k) \xrightarrow{\mathcal{D}} \nu \sup_{0 \le t \le 1} \left| B_{\frac{1}{2}}(t) - t B_{\frac{1}{2}}(1) \right| ;$$

– *if* $\mathrm{E}(Z_1 \mid \mathcal{F}_0) \neq 0$*, then*

$$n^{\frac{D}{2}-1} L_\gamma^{-\frac{1}{2}}(n) \max_{1 \le k < n} C_n(k) \xrightarrow{\mathcal{D}} |C_{\psi,\sigma,D}| \sup_{0 \le t \le 1} |B_H(t) - t B_H(1)| .$$

- Hurst parameter H/LRD parameter D
- slowly varying function L_γ
- coefficients ν, $C_{\psi,\sigma,D}$

CUSUM change-point test: rejects for large values of

$$C_n = \max_{1 \le k < n} C_n(k), \text{ where } C_n(k) = \left| \sum_{i=1}^{k} Z_i - \frac{k}{n} \sum_{i=1}^{n} Z_i \right| .$$

Asymptotic distribution of the CUSUM test statistic:

Corollary (B., Kulik (2017))

Given the asumptions of the previous theorem, under $\mathbf{H_0}$*,*

- *if* $\mathrm{E}(Z_1 \mid \mathcal{F}_0) = 0$*, then*

$$n^{-\frac{1}{2}} \max_{1 \le k < n} C_n(k) \xrightarrow{\mathcal{D}} \nu \sup_{0 \le t \le 1} \left| B_{\frac{1}{2}}(t) - tB_{\frac{1}{2}}(1) \right| ;$$

- *if* $\mathrm{E}(Z_1 \mid \mathcal{F}_0) \neq 0$*, then*

$$n^{\frac{D}{2}-1} L_{\gamma}^{-\frac{1}{2}}(n) \max_{1 \le k < n} C_n(k) \xrightarrow{\mathcal{D}} |C_{\psi,\sigma,D}| \sup_{0 \le t \le 1} |B_H(t) - tB_H(1)| .$$

- Hurst parameter H/LRD parameter D unknown!
- slowly varying function L_γ unknown!
- coefficients ν, $C_{\psi,\sigma,D}$ unknown!

Self-normalized CUSUM test statistic (*Shao (2011)*):

$$\mathrm{SC}_n = \max_{1 \le k < n} C_n^\star(k),$$

$$C_n^\star(k) = V_n^{-1}(k) \left| \sum_{i=1}^{k} Z_i - \frac{k}{n} \sum_{i=1}^{n} Z_i \right|$$

with

$$V_n(k) = \left\{ \frac{1}{n} \sum_{t=1}^{k} S_t^2(1,k) + \frac{1}{n} \sum_{t=k+1}^{n} S_t^2(k+1,n) \right\}^{\frac{1}{2}}$$

where $S_t(j,k) = \sum_{h=j}^{t} \left(Z_h - \bar{Z}_{j,k} \right)$, $\quad \bar{Z}_{j,k} = \frac{1}{k-j+1} \sum_{t=j}^{k} Z_t$.

Remark: SC_n only depends on the realizations $Z_1, \ldots, Z_n$, i.e. SC_n does not depend on any unknown parameters.

Asymptotic distribution of the self-normalized CUSUM test statistic:

Corollary (B., Kulik (2017))

Given the asumptions of the previous theorem, under $\mathbf{H_0}$, *it follows that* $SC_n \xrightarrow{\mathcal{D}} SC_H$, *where*

$$SC_H = \sup_{r \in [0,1]} \frac{|B_H(r) - rB_H(1)|}{\left\{\int_0^r (V_H(r';0,r))^2 \, dr' + \int_r^1 (V_H(r';r,1))^2 \, dr'\right\}^{\frac{1}{2}}}$$

with $V_H(r;r_1,r_2) = B_H(r) - B_H(r_1) - \frac{r-r_1}{r_2-r_1}\{B_H(r_2) - B_H(r_1)\}$ *for* $r \in [r_1, r_2]$, $0 < r_1 < r_2 < 1$, *and with*

- $H = \frac{1}{2}$ *if* $\mathrm{E}(Z_1 \mid \mathcal{F}_0) = 0$;
- $H = 1 - \frac{D}{2}$ *if* $\mathrm{E}(Z_1 \mid \mathcal{F}_0) \neq 0$.

Remark: The limit SC_H only depends on the Hurst parameter H, i.e. it does not depend on the unknown constants $C_{\Psi,\sigma,D}$ and ν.

Simulations

Change in volatility:

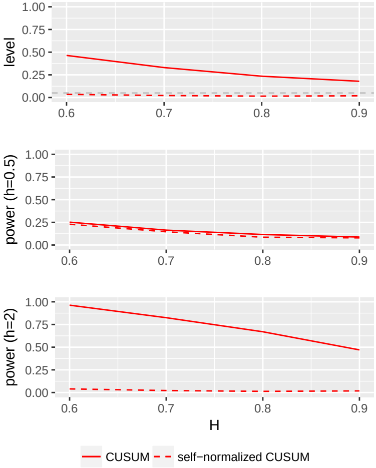
$$Z_j = \psi(X_j) \text{ with } \psi(x) = x^2$$
$$X_j = \sigma(Y_j)\varepsilon_j,$$

where

- ε_j, $j \geq 1$, i.i.d. centered Pareto$(\alpha, 1)$ distributed with $\alpha = 4.5$
- Y_j, $j \geq 1$, is a fractional Gaussian noise sequence with Hurst parameter H
- $\sigma(z) = \exp(z)$

Under H_1:

- shift in $\tau = 0.25$ of heights $h = 0.5$ and $h = 2$



Wilcoxon-based change-point tests

Wilcoxon change-point test: rejects for large values of

$$W_n = \max_{1 \leq k < n} W_n(k), \text{ where } W_n(k) = \left| \sum_{i=1}^{k} \sum_{j=k+1}^{n} \left(1_{\{Z_i \leq Z_j\}} - \frac{1}{2} \right) \right|.$$

Observe that

$$\begin{aligned} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=\lfloor nt \rfloor+1}^{n} \left(1_{\{Z_i \leq Z_j\}} - \frac{1}{2} \right) =& \lfloor nt \rfloor \sum_{j=\lfloor nt \rfloor+1}^{n} F_{1,\lfloor nt \rfloor}(Z_j) - \int_{\mathbb{R}} F_{Z_1}(x) dF_{Z_1}(x) \\ =& (n - \lfloor nt \rfloor) \int_{\mathbb{R}} \lfloor nt \rfloor \left(F_{1,\lfloor nt \rfloor}(x) - F_{Z_1}(x) \right) dF_{\lfloor nt \rfloor+1,n}(x) \\ &+ \int_{\mathbb{R}} F_{Z_1}(x) d \left(F_{\lfloor nt \rfloor+1,n}(x) - F_{Z_1}(x) \right) \end{aligned}$$

where $F_{k,l}(x) = \sum_{j=k}^{l} 1_{\{Z_j \leq x\}}$.

Wilcoxon-based change-point tests

Wilcoxon change-point test: rejects for large values of

$$W_n = \max_{1 \le k < n} W_n(k), \text{ where } W_n(k) = \left| \sum_{i=1}^{k} \sum_{j=k+1}^{n} \left(1_{\{Z_i \le Z_j\}} - \frac{1}{2} \right) \right|.$$

Observe that

$$\sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=\lfloor nt \rfloor+1}^{n} \left(1_{\{Z_i \le Z_j\}} - \frac{1}{2} \right) = \lfloor nt \rfloor \sum_{j=\lfloor nt \rfloor+1}^{n} F_{1,\lfloor nt \rfloor}(Z_j) - \int_{\mathbb{R}} F_{Z_1}(x) dF_{Z_1}(x)$$

$$= (n - \lfloor nt \rfloor) \int_{\mathbb{R}} \lfloor nt \rfloor \left(F_{1,\lfloor nt \rfloor}(x) - F_{Z_1}(x) \right) dF_{\lfloor nt \rfloor+1,n}(x)$$

$$+ \int_{\mathbb{R}} F_{Z_1}(x) d\left(F_{\lfloor nt \rfloor+1,n}(x) - F_{Z_1}(x) \right)$$

where $F_{k,l}(x) = \sum_{j=k}^{l} 1_{\{Z_j \le x\}}$.

$\Rightarrow$ Consider the two-parameter empirical process

$$\lfloor nt \rfloor \left(F_{1,\lfloor nt \rfloor}(x) - F_{Z_1}(x) \right) = \sum_{j=1}^{\lfloor nt \rfloor} \left(1_{\{Z_j \le x\}} - F_{Z_1}(x) \right), \; t \in [0,1], \; x \in [-\infty, \infty].$$

Two-parameter empirical process limit theorems

$$\sum_{j=1}^{\lfloor nt \rfloor} \left(1_{\{Z_j \leq x\}} - F_{Z_1}(x) \right), \; t \in [0,1], \; x \in [-\infty, \infty].$$

- Independent observations
  - *Müller (1970), Kiefer (1972).*
- Short-range dependent observations
  - *Berkes and Philipp (1977)*: for strong mixing processes.
  - *Berkes, Hörmann and Schauer (2009)*: for S-mixing processes.
- Long-range dependent observations
  - *Dehling and Taqqu (1989)*: for subordinated Gaussian processes.
  - *Giraitis and Surgailis (2002)*: for linear processes.

A two-parameter empirical process limit theorem for subordinated LMSV time series

Theorem (B., Kulik (2017))

Suppose X_n, $n \in \mathbb{N}$, is an LMSV time series and let ψ be a measurable function. Define $\Psi_x(y) := P(\psi(y\varepsilon_1) \leq x)$ and assume that $\Psi_x(y)$ is differentiable. Given the previous assumptions (+ technical assumptions), under $\mathbf{H_0}$,

$$n^{\frac{D}{2}-1} L_\gamma^{-1/2}(n) \sum_{j=1}^{\lfloor nt \rfloor} \left(1_{\{\psi(X_j) \leq x\}} - F_{\psi(X_1)}(x) \right) \xrightarrow{\mathcal{D}} J\left(\Psi_x'(y) \circ \sigma\right) B_H(t),$$

in $D([-\infty, \infty] \times [0,1])$ with B_H denoting a fractional Brownian motion, $H = 1 - \frac{D}{2}$, and $J(G) = \mathrm{E}\left(G(Y_1) Y_1\right)$.

Proof

Recall that

$$X_j = \sigma(Y_j)\varepsilon_j, \;\; Y_j = \sum_{k=1}^{\infty} c_k \eta_{j-k}, \;\; j \geq 1 \; ,$$

where ε_j, $j \geq 1$, and η_j, $j \in \mathbb{Z}$ are i.i.d. sequences and that

$$\mathcal{F}_j := \sigma\left(\varepsilon_j, \varepsilon_{j-1}, \ldots, \eta_j, \eta_{j-1}, \ldots\right),$$

such that ε_j is independent of $\mathcal{F}_{j-1}$ and Y_j is $\mathcal{F}_{j-1}$-measurable.

Consider the decomposition

$$\sum_{j=1}^{\lfloor nt \rfloor} \left(1_{\{\psi(X_j) \leq x\}} - F_{\psi(X_1)}(x)\right) = M_n(x,t) + R_n(x,t), \;\; x \in [-\infty, \infty]\, , \; t \in [0,1],$$

where

$$M_n(x,t) := \sum_{j=1}^{\lfloor nt \rfloor} \left(1_{\{\psi(X_j) \leq x\}} - \mathrm{E}\left(1_{\{\psi(X_j) \leq x\}} \mid \mathcal{F}_{j-1}\right)\right) \qquad \textbf{Martingale part},$$

$$R_n(x,t) := \sum_{j=1}^{\lfloor nt \rfloor} \left(\mathrm{E}\left(1_{\{\psi(X_j) \leq x\}} \mid \mathcal{F}_{j-1}\right) - F_{\psi(X_1)}(x)\right) \qquad \textbf{LRD part}.$$

Proof

Martingale part: Aldous' tightness condition for multiparameter martingales (*Ivanoff (1983)*) yields

$$\frac{1}{\sqrt{n}} M_n(x, t) = \mathcal{O}_P(1)$$

in $D([-\infty, \infty] \times [0, 1])$.

Proof

Martingale part: Aldous' tightness condition for multiparameter martingales (*Ivanoff (1983)*) yields

$$\frac{1}{\sqrt{n}} M_n(x, t) = \mathcal{O}_P(1)$$

in $D([-\infty, \infty] \times [0, 1])$.

Long-range dependent part: Define $\Psi_x(y) := P\left(\psi(y\varepsilon_1) \leq x\right)$.

$$\begin{aligned} n^{\frac{D}{2}-1} L_\gamma^{-\frac{1}{2}}(n) R_n(x, t) &= n^{\frac{D}{2}-1} L_\gamma^{-\frac{1}{2}}(n) \sum_{j=1}^{\lfloor nt \rfloor} \Big(\underbrace{\mathrm{E}\left(1_{\{\psi(\sigma(Y_j)\varepsilon_j) \leq x\}} \mid \mathcal{F}_{j-1}\right)}_{=\Psi_x(\sigma(Y_j))} - \underbrace{F_{\psi(X_1)}(x)}_{=\mathrm{E}\, \Psi_x(\sigma(Y_j))} \Big) \\ &= n^{\frac{D}{2}-1} L_\gamma^{-\frac{1}{2}}(n) \lfloor nt \rfloor \int_{\mathbb{R}} \Psi_x'(y) \left(F_{\lfloor nt \rfloor}(y) - \mathrm{E}\, F_{\lfloor nt \rfloor}(y)\right) dy, \end{aligned}$$

where $F_l(u) = \frac{1}{l} \sum_{j=1}^{l} 1_{\{\sigma(Y_j) \leq u\}}$.

Long-range dependent part:

$$n^{\frac{D}{2}-1}L_\gamma^{-\frac{1}{2}}(n)R_n(x,t) = -n^{\frac{D}{2}-1}L_\gamma^{-\frac{1}{2}}(n)\lfloor nt\rfloor \int_{\mathbb{R}} \Psi_x'(y)\left(F_{\lfloor nt\rfloor}(y) - \operatorname{E} F_{\lfloor nt\rfloor}(y)\right) dy$$

Theorem (Dehling, Taqqu (1989))

$$\sup_{t\in[0,1],x\in[-\infty,\infty]} \left| n^{\frac{D}{2}-1}L_\gamma^{-\frac{1}{2}}(n)\left\{\lfloor nt\rfloor\left(F_{\lfloor nt\rfloor}(y) - \operatorname{E} F_{\lfloor nt\rfloor}(y)\right) - J(\sigma;y)\sum_{j=1}^{\lfloor nt\rfloor} Y_j\right\}\right| \xrightarrow{P} 0,$$

where $J(\sigma;y) = \operatorname{E}\left(1_{\{\sigma(Y_1)\le y\}} Y_1\right)$.

It suffices to consider

$$-\int_{\mathbb{R}} \Psi_x'(y)J(\sigma;y)dy\; n^{\frac{D}{2}-1}L_\gamma^{-\frac{1}{2}}(n)\sum_{j=1}^{\lfloor nt\rfloor} Y_j.$$

Theorem (Taqqu (1975))

$$n^{\frac{D}{2}-1}L_\gamma^{-\frac{1}{2}}(n)\sum_{j=1}^{\lfloor nt\rfloor} Y_j \xrightarrow{\mathcal{D}} B_H(t) \text{ in } D[0,1],$$

where B_H denotes a fractional Brownian motion and $H = 1 - \frac{D}{2}$.

As a result,

$$-\int_{\mathbb{R}} \Psi_x'(y)J(\sigma;y)dy\; n^{\frac{D}{2}-1}L_\gamma^{-\frac{1}{2}}(n)\sum_{j=1}^{\lfloor nt\rfloor} Y_j \xrightarrow{\mathcal{D}} \underbrace{-\int_{\mathbb{R}} J(\sigma;y)\Psi_x'(y)dy}_{=J\left(\Psi_x'(y)\circ\sigma\right)}\; B_H(t).$$

Theorem (Taqqu (1975))

$$n^{\frac{D}{2}-1}L_\gamma^{-\frac{1}{2}}(n)\sum_{j=1}^{\lfloor nt\rfloor} Y_j \xrightarrow{\mathcal{D}} B_H(t) \text{ in } D\,[0,1],$$

where B_H denotes a fractional Brownian motion and $H = 1 - \frac{D}{2}$.

As a result,

$$-\int_{\mathbb{R}} \Psi'_x(y)J(\sigma;y)dy\; n^{\frac{D}{2}-1}L_\gamma^{-\frac{1}{2}}(n)\sum_{j=1}^{\lfloor nt\rfloor} Y_j \xrightarrow{\mathcal{D}} \underbrace{-\int_{\mathbb{R}} J(\sigma;y)\Psi'_x(y)dy}_{=J\left(\Psi'_x(y)\circ\sigma\right)}\; B_H(t).$$

Martingale part + Long-range dependent part:

$$n^{\frac{D}{2}-1}L_\gamma^{-\frac{1}{2}}(n)\sum_{j=1}^{\lfloor nt\rfloor}\left(1_{\{\psi(X_j)\leq x\}} - F_{\psi(X_1)}(x)\right)$$

$$= \underbrace{n^{\frac{D}{2}-\frac{1}{2}}L_\gamma^{-\frac{1}{2}}(n)}_{=o(1)}\underbrace{\frac{1}{\sqrt{n}}M_n(x,t)}_{=\mathcal{O}_P(1)} + \underbrace{n^{\frac{D}{2}-1}L_\gamma^{-\frac{1}{2}}(n)R_n(x,t)}_{\xrightarrow{\mathcal{D}} J\left(\Psi'_x(y)\circ\sigma\right)B_H(t)} \xrightarrow{\mathcal{D}} J\left(\Psi'_x(y)\circ\sigma\right)B_H(t).$$

Theorem (B., Kulik (2017))

Suppose X_n, $n \in \mathbb{N}$, is an LMSV time series and let ψ be a measurable function. Given the previous assumptions (+ technical assumptions), under $\mathbf{H_0}$*,*

$$n^{\frac{D}{2}-1} L_\gamma^{-1/2}(n) \sum_{j=1}^{\lfloor nt \rfloor} \left(1_{\{\psi(X_j) \leq x\}} - F_{\psi(X_1)}(x) \right) \xrightarrow{\mathcal{D}} J\left(\Psi'_x(y) \circ \sigma \right) B_H(t),$$

in $D([-\infty, \infty] \times [0, 1])$ with B_H denoting a fractional Brownian motion, $H = 1 - \frac{D}{2}$.

Theorem (B., Kulik (2017))

Suppose X_n, $n \in \mathbb{N}$, is an LMSV time series and let ψ be a measurable function. Given the previous assumptions (+ technical assumptions), under $\mathbf{H_0}$,

$$n^{\frac{D}{2}-1} L_\gamma^{-1/2}(n) \sum_{j=1}^{\lfloor nt \rfloor} \left(1_{\{\psi(X_j) \leq x\}} - F_{\psi(X_1)}(x) \right) \xrightarrow{\mathcal{D}} J\left(\Psi'_x(y) \circ \sigma\right) B_H(t),$$

in $D\left([-\infty, \infty] \times [0, 1]\right)$ with B_H denoting a fractional Brownian motion, $H = 1 - \frac{D}{2}$.

Asymptotic distribution of the Wilcoxon test statistic:

Corollary (B., Kulik (2017))

Given the asumptions of the previous theorem, under $\mathbf{H_0}$,

$$n^{\frac{D}{2}-2} L_\gamma^{-1/2}(n) \max_{1 \leq k < n} W_n(k) \xrightarrow{\mathcal{D}} |C_{\Psi,\sigma,D}| \sup_{0 \leq t \leq 1} |B_H(t) - tB_H(1)|.$$

- Hurst parameter H/LRD parameter D
- slowly varying function L_γ
- coefficient $C_{\psi,\sigma,D}$

Theorem (B., Kulik (2017))

Suppose X_n, $n \in \mathbb{N}$, is an LMSV time series and let ψ be a measurable function. Given the previous assumptions (+ technical assumptions), under $\mathbf{H_0}$,

$$n^{\frac{D}{2}-1} L_\gamma^{-1/2}(n) \sum_{j=1}^{\lfloor nt \rfloor} \left(1_{\{\psi(X_j) \leq x\}} - F_{\psi(X_1)}(x) \right) \xrightarrow{\mathcal{D}} J\left(\Psi'_x(y) \circ \sigma\right) B_H(t),$$

in $D([-\infty, \infty] \times [0, 1])$ with B_H denoting a fractional Brownian motion, $H = 1 - \frac{D}{2}$.

Asymptotic distribution of the Wilcoxon test statistic:

Corollary (B., Kulik (2017))

Given the asumptions of the previous theorem, under $\mathbf{H_0}$,

$$n^{\frac{D}{2}-2} L_\gamma^{-1/2}(n) \max_{1 \leq k < n} W_n(k) \xrightarrow{\mathcal{D}} |C_{\Psi,\sigma,D}| \sup_{0 \leq t \leq 1} |B_H(t) - tB_H(1)|.$$

- Hurst parameter H/LRD parameter D unknown!
- slowly varying function L_γ unknown!
- coefficient $C_{\psi,\sigma,D}$ unknown!

Define $R_i = \text{rank}(Z_i) = \sum_{j=1}^{n} 1_{\{Z_j \leq Z_i\}}$, $R_{j,k} = \frac{1}{k-j+1} \sum_{i=j}^{k} \bar{R}_i$.
Note that

$$\left| \sum_{i=1}^{k} \sum_{j=k+1}^{n} \left(1_{\{Z_i \leq Z_j\}} - \frac{1}{2} \right) \right| = \left| \sum_{i=1}^{k} R_i - \frac{k}{n} \sum_{i=1}^{n} R_i \right| .$$

Self-normalized Wilcoxon test statistic (*Betken (2016)*):

$$\text{SW}_n = \sup_{1 \leq k < n} W_n^\star(k),$$

$$W_n^\star(k) = V_n^{-1}(k) \left| \sum_{i=1}^{k} R_i - \frac{k}{n} \sum_{i=1}^{n} R_i \right| ,$$

with

$$V_n(k) = \left\{ \frac{1}{n} \sum_{t=1}^{k} S_t^2(1,k) + \frac{1}{n} \sum_{t=k+1}^{n} S_t^2(k+1,n) \right\}^{\frac{1}{2}}$$

where $S_t(j,k) = \sum_{h=j}^{t} \left(R_h - \bar{R}_{j,k} \right)$, $\quad \bar{R}_{j,k} = \frac{1}{k-j+1} \sum_{t=j}^{k} R_t$.

Remark: SW_n only depends on the realizations $X_1, \ldots, X_n$, i.e. SW_n does not depend on L_γ or H.

Asymptotic distribution of the self-normalized Wilcoxon test statistic

Corollary (B., Kulik (2017))

Given the asumptions of the previous theorem, under $\mathbf{H_0}$*, it follows that* $SW_n \xrightarrow{\mathcal{D}} SW_H$*, where*

$$SW_H = \sup_{r\in[0,1]} \frac{|B_H(r) - rB_H(1)|}{\left\{\int_0^r (V_H(r';0,r))^2 \, dr' + \int_r^1 (V_H(r';r,1))^2 \, dr'\right\}^{\frac{1}{2}}}$$

with $V_H(r;r_1,r_2) = B_H(r) - B_H(r_1) - \frac{r-r_1}{r_2-r_1}\{B_H(r_2) - B_H(r_1)\}$ *for* $r \in [r_1, r_2]$, $0 < r_1 < r_2 < 1$.

Remark: The limit SW_H only depends on the Hurst parameter H, i.e. it does not depend on the unknown constant $C_{\Psi,\sigma,D}$.

Simulations

Change in volatility:

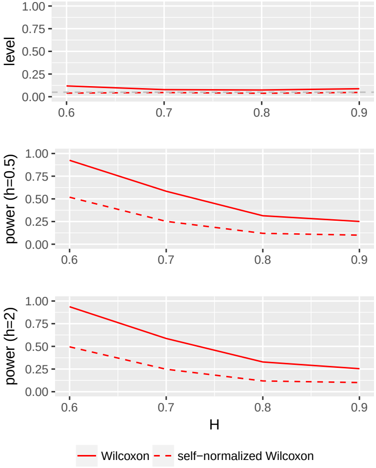
$$Z_j = \psi(X_j) \text{ with } \psi(x) = x^2$$
$$X_j = \sigma(Y_j)\varepsilon_j,$$

where

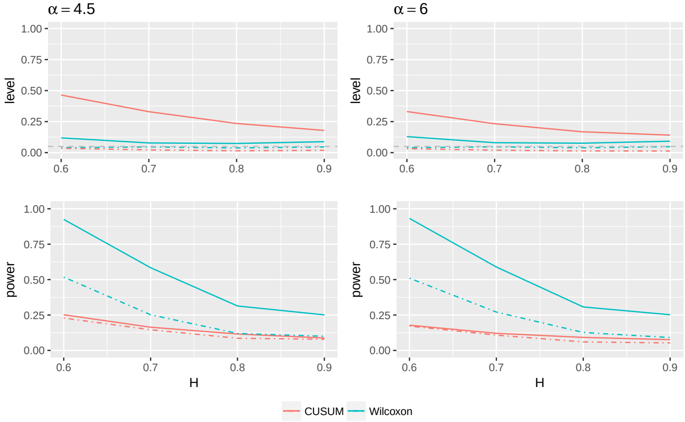
- ε_j, $j \geq 1$, i.i.d. centered Pareto$(\alpha, 1)$ distributed with $\alpha = 4.5$
- Y_j, $j \geq 1$, is a fractional Gaussian noise sequence with Hurst parameter H
- $\sigma(z) = \exp(z)$

Under H_1:

- shift in $\tau = 0.25$ of heights $h = 0.5$ and $h = 2$



Simulations



Rejection rates of the CUSUM and Wilcoxon tests for LMSV time series of length $n = 500$ with Hurst parameter H, tail index α and a shift in the variance in $\tau = 0.25$ with height $h = 0.5$. The calculations are based on 5000 simulation runs.

References

- A. BETKEN, R. KULIK (2016). Testing for change in stochastic volatility with long range dependence. *arXiv:1706.06351*.
- A. BETKEN (2016). Testing for change-points in long-range dependent times series by means of a self-normalized Wilcoxon test. *Journal of Time Series Analysis* 37:785 – 809.
- J. BERAN, Y. FENG, S. GHOSH, R. KULIK (2013). Long-Memory Processes. *Springer*.
- X. SHAO. (2011) A simple test of changes in mean in the possible presence of long-range dependence. *Journal of Time Series Analysis,* 32(6):598 – 606.
- R. KULIK, P. SOULIER (2011). The tail empirical process for long memory stochastic volatility sequences. *Stochastic Processes and their Applications*, 121:109 - 134.
- A. C. HARVEY (2002). Long memory in stochastic volatility. *In: Forecasting Volatility in the Financial Markets*, 307-320.
- F. BREIDT, N. CRATO, AND P. DE LIMA (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, 83:325 - 348.
- GAIL IVANOFF (1983). Stopping times and tightness in two dimensions. *Technical report*, University of Ottawa.