A general white noise test based on kernel lag-window estimates of the spectral density operator

Vaidotas Characiejus[a]

Joint work with Gregory Rice[b]

[a]Département de Mathématique, Université libre de Bruxelles, Belgium
vaidotas.characiejus@ulb.ac.be

[b]Department of Statistics and Actuarial Science, University of Waterloo, Canada

ISNPS2018
Salerno, June 15, 2018

Outline

Problem

Method and test statistic

Simulation study

Summary

$\mathbb{H}$-valued time series

$\{X_t\}_{t\in\mathbb{Z}}$ is a stationary sequence of random elements with values in a real separable Hilbert space $\mathbb{H}$ such that $\operatorname{E} X_0 = 0$.

Definition

The autocovariance operators $\{\mathcal{C}(j)\}_{j\in\mathbb{Z}}$ of $\{X_t\}_{t\in\mathbb{Z}}$ are defined by

$$\mathcal{C}(j) = \operatorname{E}[X_j \otimes X_0] = \operatorname{E}[\langle \cdot, X_0\rangle X_j]$$

for $j \in \mathbb{Z}$.

White noise and hypothesis

Definition

$\{X_t\}_{t\in\mathbb{Z}}$ is white noise if X_t's are uncorrelated, i.e. if $\mathcal{C}(j) = 0$ for each $j \neq 0$.

We are interested in testing the hypothesis that $\{X_t\}_{t\in\mathbb{Z}}$ is white noise.

Available tests

Time-domain tests for independence

- Gabrys and Kokoszka [2007], Gabrys, Horváth, and Kokoszka [2010];
- Horváth, Hušková, and Rice [2013].

Frequency-domain tests for white noise

- Zhang [2016];
- Bagchi, Ch., and Dette [2018].

Test that we propose

The idea is to measure the distance between the stationary sequence $\{X_t\}_{t\in\mathbb{Z}}$ and white noise.

Such a test was proposed by Hong (1996) in the univariate setting.

Spectral density function

Definition

The spectral density function is a discrete-time Fourier transform of $\{\mathcal{C}(j)\}_{j\in\mathbb{Z}}$ defined by

$$\mathcal{F}(\omega) = (2\pi)^{-1} \sum_{j\in\mathbb{Z}} \mathcal{C}(j) e^{-ij\omega}$$

for $\omega \in [-\pi, \pi]$ provided that $\sum_{j\in\mathbb{Z}} \|\mathcal{C}(j)\|_2 < \infty$, where $i = \sqrt{-1}$ and $\|\cdot\|_2$ is the Hilbert-Schmidt norm.

If $\{X_t\}_{t\in\mathbb{Z}}$ is white noise, then $\mathcal{F}(\omega) = (2\pi)^{-1}\mathcal{C}(0)$ for $\omega \in [-\pi, \pi]$.

Distance function

The distance between $\mathcal{F}$ and $(2\pi)^{-1}\mathcal{C}(0)$ is measured by

$$Q^2 = 2\pi \int_{-\pi}^{\pi} |||\mathcal{F}(\omega) - (2\pi)^{-1}\mathcal{C}(0)|||_2^2 d\omega,$$

where $||| \cdot |||_2$ is the Hilbert-Schmidt norm.

We have that $Q^2 = \sum_{h \neq 0} |||\mathcal{C}(h)|||_2^2$.

Hypothesis

The hypothesis that we want to test is as follows

$$H_0 \ : \ Q = 0 \quad \text{versus} \quad H_1 \ : \ Q > 0.$$

To perform the test, we need an estimator of Q.

Sample autocovariance operators

Definition

The sample autocovariance operators are defined by

$$\hat{\mathcal{C}}_n(j) = n^{-1} \sum_{t=j+1}^{n} X_t \otimes X_{t-j}$$

for $0 \leq j < n$ and by $\hat{\mathcal{C}}_n(j) = \hat{\mathcal{C}}_n^*(-j)$ for $-n < j < 0$.

Estimator of spectral density function

Definition

The kernel lag-window estimator of the spectral density function is defined by

$$\hat{\mathcal{F}}_n(\omega) = (2\pi)^{-1} \sum_{|j|<n} k(j/p_n)\hat{\mathcal{C}}_n(j)e^{-ij\omega}$$

for $\omega \in [-\pi, \pi]$, where $k : \mathbb{R} \to [-1, 1]$ is a kernel and $\{p_n\}_{n\geq 1}$ is a bandwidth.

Estimator of the distance to white noise

The estimator of Q is defined by

$$\hat{Q}_n^2 = 2\pi \int_{-\pi}^{\pi} |||\hat{\mathcal{F}}_n(\omega) - (2\pi)^{-1}\hat{\mathcal{C}}_n(0)|||_2^2 d\omega.$$

Alternatively, the estimator $\hat{Q}_n$ can be expressed as

$$\hat{Q}_n^2 = 2 \sum_{j=1}^{n-1} k^2(j/p_n) |||\hat{\mathcal{C}}_n(j)|||_2^2.$$

Test statistic

We propose to use the test statistic T_n defined by

$$T_n = T_n(k, p_n) = \frac{2^{-1} n \hat{Q}_n^2 - \hat{\sigma}_n^4 C_n(k)}{\|\hat{\mathcal{C}}_n(0)\|_2^2 \sqrt{2 D_n(k)}}$$

for $n \geq 1$, where $\hat{\sigma}^2 = n^{-1} \sum_{t=1}^{n} \|X_t\|^2$,

$$C_n(k) = \sum_{j=1}^{n-1} (1 - j/n) k^2(j/p_n),$$

$$D_n(k) = \sum_{j=1}^{n-2} (1 - j/n)(1 - (j+1)/n) k^4(j/p_n).$$

Asymptotic distribution of the statistic

Theorem

Suppose that

(i) $\{X_t\}_{t\in\mathbb{Z}}$ *are iid* $\mathbb{H}$*-valued random elements such that* $\operatorname{E} X_0 = 0$ *and* $\operatorname{E}\|X_0\|^4 < \infty$*;*

(ii) k *is an even function that is continuous at zero and at all but finite number of points, with* $k(0) = 1$ *and* $k(x) = O(x^{-\alpha})$ *for some* $\alpha > 1/2$ *as* $x \to \infty$*;*

(iii) $p_n \to \infty$ *and* $p_n/n \to 0$ *as* $n \to \infty$*.*

Then

$$T_n \xrightarrow{d} N(0, 1)$$

as $n \to \infty$*.*

Special cases

Hong (1996)

We have that

$$T_n = \frac{\hat{\sigma}_n^4}{\|\|\hat{\mathcal{C}}_n(0)\|\|_2^2} \cdot \frac{2^{-1} n \hat{\sigma}_n^{-4} \hat{Q}_n^2 - C_n(k)}{\sqrt{2D_n(k)}}$$

for $n \geq 1$. If $\mathbb{H} = \mathbb{R}$, then

$$\frac{\hat{\sigma}_n^4}{\|\|\hat{\mathcal{C}}_n(0)\|\|_2^2} \xrightarrow{p} 1$$

and we recover the test statistic proposed by Hong (1996).

Special cases (cont.)

Horváth, Hušková, and Rice (2013)

If $\mathbb{H} = L^2([0,1], \mathbb{R})$ and $k = \mathbf{1}_{\{|x| \leq 1\}}$, then T_n is asymptotically equivalent to

$$T_n^* = \frac{n \sum_{j=1}^{p_n} |||\hat{\mathcal{C}}_n(j)|||_2^2 - \hat{\sigma}_n^4 p_n}{|||\hat{\mathcal{C}}_n(0)|||_2^2 \sqrt{2p_n}}$$

which is the test statistic considered in Horváth, Hušková, and Rice (2013).

Consistency of the test

Theorem

Suppose that

(i) $\{X_t\}_{t\in\mathbb{Z}}$ *is a fourth order stationary sequence of zero mean* $\mathbb{H}$*-valued random elements such that* $\sum_{j=-\infty}^{\infty} \|\|\mathcal{C}(j)\|\|_1^2 < \infty$ *and* $\sup_{j\in\mathbb{Z}} \sum_{h=-\infty}^{\infty} \|\|\mathcal{K}_{h+j,h,j}\|\|_1 < \infty$*, where* $\|\|\cdot\|\|_1$ *is the nuclear norm and* $\{\mathcal{K}_{j_1,j_2,j_3}\}_{j_1,j_2,j_3\in\mathbb{Z}}$ *are the fourth order cumulant operators;*

(ii) $p_n \to \infty$ *and* $p_n/n \to 0$ *as* $n \to \infty$*.*

Then

$$(p_n^{1/2}/n)T_n \xrightarrow{p} \frac{2^{-1}Q^2}{\|\|\mathcal{C}(0)\|\|_2^2(2D(k))^{1/2}}$$

as $n \to \infty$*.*

Square root transformation

The transformed test statistic is given by

$$\begin{aligned} T_n^{SQ} &= T_n^{SQ}(k, p_n) \\ &= \Big[\frac{2\hat{\sigma}_n^4 C_n(k)}{D_n(k) \|\hat{\mathcal{C}}_n(0)\|_2^4}\Big]^{1/2} [(2^{-1} n \hat{Q}_n^2)^{1/2} - (\hat{\sigma}_n^4 C_n(k))^{1/2}]. \end{aligned}$$

Under the same assumptions, we have that

$$T_n^{SQ} \xrightarrow{d} N(0, 1)$$

as $n \to \infty$.

Simulation setup

We investigate the case when $\mathbb{H} = L^2([0,1], \mathbb{R})$.

The following data generating processes are considered

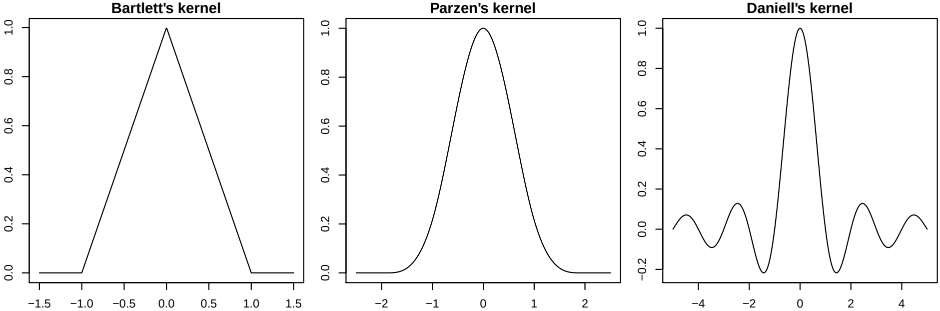
(i) IID-BM;

(ii) fGARCH$(1,1)$ (Aue, Horváth, and Pellatt (2016));

(iii) FAR$(1, S)$-BM with the kernel of the operator given by $\varphi_c(t,s) = c\exp\{(t^2+s^2)/2\}$ for $t, s \in [0,1]$ and the constant c is chosen so that $\|\varphi_c\| = S \in (0,1)$.

Each random function was generated on 100 equally spaced points. The burn-in sample for fGARCH$(1,1)$ and FAR$(1, S)$-BM was 100. The number of the Monte Carlo replication was 1000.

Kernels


Bartlett's kernel
Parzen's kernel
Daniell's kernel
0.0
0.2
0.4
0.6
0.8
1.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
−2
−1
0
1
2
−0.2
−4
4

Bandwidth selection

Similarly as in Bühlmann (1996), we consider bandwidths of the form

$$p_n = n^{1/(2q+1)}$$

and

$$p_n = \hat{M} n^{1/(2q+1)},$$

where q is the order of the kernel and $\hat{M}$ is a constant estimated from the data.

Monte Carlo simulation

DGP:	IID-BM				fGARCH(1,1)				FAR(1,0.3)-BM			
	$n=100$		$n=250$		$n=100$		$n=250$		$n=100$		$n=250$	
Stat/Nominal Size	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%	5%	1%
$T_n(k_B, n^{1/3})$	60	30	73	39	145	85	126	78	826	733	996	986
$T_n(k_B, \hat{M}n^{1/3})$	61	38	74	33	136	87	119	76	751	633	990	981
$T_n(k_P, n^{1/5})$	59	32	71	26	139	87	132	81	841	773	998	992
$T_n(k_P, \hat{M}n^{1/5})$	64	28	74	39	138	90	130	77	789	681	996	986
$T_n(k_D, n^{1/5})$	61	35	65	28	139	87	131	76	843	770	999	996
$T_n(k_D, \hat{M}n^{1/5})$	63	33	72	33	140	85	117	77	836	763	998	995
$T_n^{SQ}(k_B, n^{1/3})$	37	13	54	15	110	51	102	48	771	618	993	980
$T_n^{SQ}(k_B, \hat{M}n^{1/3})$	44	16	41	13	90	42	79	33	793	631	998	987
$T_n^{SQ}(k_P, n^{1/5})$	43	15	45	13	99	46	93	48	789	640	996	981
$T_n^{SQ}(k_P, \hat{M}n^{1/5})$	38	13	54	20	100	50	90	44	739	592	990	977
$T_n^{SQ}(k_D, n^{1/5})$	43	19	41	14	101	45	94	48	802	666	996	982
$T_n^{SQ}(k_D, \hat{M}n^{1/5})$	42	16	41	16	99	44	87	43	798	657	997	987
$Z_n(10)$	48	9	49	11	50	12	41	5	708	386	992	913
BCD_n	28	13	38	12	46	22	59	21	197	109	433	301

Summary

- A general test for white noise for $\mathbb{H}$-valued time series.
- The asymptotic distribution under independence and the consistency of the test.
- Better power against functional autoregressive alternatives compared to the existing tests.
- Not well sized for general weak white noise in function space such as for functional GARCH processes.

Preprint: `https://arxiv.org/abs/1803.09501`