Testing for white noise in functional time series

Vaidotas Characiejus[1]

Joint work with Pramita Bagchi[2] and Holger Dette[2]

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

[2]Fakultät für Mathematik, Ruhr-Universität Bochum, Germany

The joint UCL-ULB seminar
Louvain-la-Neuve, March 10, 2017

Outline

Two common problems

- The goodness of fit of a statistical model (testing whether the errors of the model are independent or uncorrelated).
- The validity of a statistical method (testing whether the data is a simple random sample).

Two approaches

- Time domain correlation-based tests;
- Frequency domain periodogram-based tests.

The Box-Ljung-Pierce approach

- Introduced by Box and Pierce (1970) and Ljung and Box (1978).
- The idea is to verify that all autocovariances and/or autocorrelations up to lag h are suitably close to 0.
- Such tests are typically referred to as "portmanteau" tests.

Generalisations to the multivariate case

Univariate case

- Box and Pierce (1970), Pierce (1972), Ljung and Box (1978).

Multivariate case

- Chitturi (1974, 1976), Hosking (1980), Li and MacLeod (1981).

Uncorrelatedness vs independence

- Box and Pierce (1970) and Ljung and Box (1978) proposed tests work under the assumption of iid errors.
- If the errors are uncorrelated but not independent, the tests are not reliable (see Romano and Thombs (1996) and Francq, Roy and Zakoïan (2005)).

Frequency-domain approach

- Developed by Durlauf (1991), Hong (1996), Deo (2000), Dette, Kinsvater and Vetter (2010), Shao (2011).
- The idea is to compare the spectral density corresponding to the sequence of the random variables and the spectral density of white noise.

Time-domain tests for functional observations

Gabrys and Kokoszka (2007)

- A portmanteau test of independence and identical distribution of functional observations.
- Based on the Karhunen–Loéve expansion.
- Need to choose the number of principal components p and the lag parameter h.
- Extended by Gabrys, Horváth, Kokoszka (2010) to test for independence in the errors of a functional linear model.

Time-domain tests for functional observations (2)

Horváth, Hušková, and Rice (2013)

- A test that is based on the sum of the L^2 norms of the empirical correlation functions.
- There is no need to choose the number of principal components p and the lag parameter h goes to infinity as the sample size increases.

Frequency-domain tests for functional observations

Zhang (2016)

- A Cramér-von Mises type test based on the L^2 norm of the functional periodogram function.
- Does not involve the choices of the functional principal components or the lag truncation number.
- The approach is robust to dependence within white noise.
- The limiting distribution of the test statistic is non-pivotal and a block bootstrap procedure is needed to obtain the critical values.

Our test

- We propose a frequency-domain based test for white noise in functional time series with a simple asymptotic distribution.
- Our approach does not need bootstrap to obtain the critical values.
- We do not need to choose the number of functional principal components or the lag truncation number.
- Our test is a generalisation of the test proposed by Dette, Kinsvater and Vetter (2010).

Functional time series

$\{X_t\}_{t\in\mathbb{Z}}$ are *strictly* stationary $L^2([0,1],\mathbb{R})$-valued random elements.

We denote the *mean curve* by

$$\mu(\tau) = \operatorname{E} X_0(\tau)$$

and the *autocovariance kernel* at lag $t \in \mathbb{Z}$ by

$$r_t(\tau,\sigma) = \operatorname{Cov}[X_t(\tau), X_0(\sigma)]$$

for $\tau, \sigma \in [0,1]$ provided that $\operatorname{E} \|X_0\|_2^2 < \infty$.

The Hilbert-Schmidt operators

Theorem

A bouned linear operator $A : L^2([0,1]^k, \mathbb{C}) \to L^2([0,1]^k, \mathbb{C})$ *is a Hilbert-Schmidt operator if and only if there exists a kernel* $k_A \in L^2([0,1]^{2k}, \mathbb{C})$ *such that*

$$Af(x) = \int_{[0,1]^k} k_A(x,y) f(y) \mathrm{d}y$$

a.e. in $[0,1]^k$ *for each* $f \in L^2([0,1]^k, \mathbb{C})$.

Weidmann (1980)

The Hilbert-Schmidt norm

The Hilbert-Schmidt norm of a Hilbert-Schmidt operator $A : L^2([0,1]^k, \mathbb{C}) \to L^2([0,1]^k, \mathbb{C})$ with the kernel k_A is given by

$$|||A|||_2^2 = \|k_A\|_2^2 = \int_{[0,1]^k} \int_{[0,1]^k} |k_A(x,y)|^2 \,\mathrm{d}x\mathrm{d}y.$$

Spectral density kernel

The *spectral density kernel* is defined as

$$f_\omega = \frac{1}{2\pi} \sum_{t \in \mathbb{Z}} \exp(-i\omega t) r_t$$

for $\omega \in \mathbb{R}$ provided that $\sum_{t \in \mathbb{Z}} \|r_t\|_2 < \infty$.

If $\{X_t\}_{t \in \mathbb{Z}}$ are uncorrelated, $f_\omega = r_0/(2\pi)$.

This definition was proposed by Panaretos and Tavakoli (2013).

Spectral density operator

The *spectral density operator* $\mathcal{F}_\omega : L^2([0,1],\mathbb{R}) \to L^2([0,1],\mathbb{C})$ is a Hilbert-Schmidt operator defined as

$$\mathcal{F}_\omega f(\tau) = \int_0^1 f_\omega(\tau,\sigma) f(\sigma) \mathrm{d}\sigma$$

for each $\omega \in \mathbb{R}$, $\tau \in [0,1]$ and $f \in L^2([0,1],\mathbb{R})$.

The hypothesis

Let us consider the null hypothesis

$$H_0 \ : \ \mathcal{F}_\omega = \mathcal{F} \quad a.e.$$

against the alternative

$$H_A \ : \ \mathcal{F}_\omega \neq \mathcal{F} \quad \text{on a set of positive Lebesgue measure}$$

for some Hilbert-Schmidt operator $\mathcal{F} : L^2([0,1],\mathbb{R}) \to L^2([0,1],\mathbb{C})$.

The distance function

We consider the problem of approximating $\mathcal{F}_\omega$ by a constant self-adjoint Hilbert-Schmidt operator $\mathcal{F}$ (corresponding to a white noise functional process) by the distance function

$$M^2(\mathcal{F}) = \int_{-\pi}^{\pi} |||\mathcal{F}_\omega - \mathcal{F}|||_2^2 \mathrm{d}\omega.$$

The kernel $\tilde{f}$ and the operator $\widetilde{\mathcal{F}}$

We define the kernel $\tilde{f}$ by setting

$$\tilde{f}(\tau,\sigma) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_\omega(\tau,\sigma)\mathrm{d}\omega$$

for each $\tau,\sigma \in [0,1]$. The kernel $\tilde{f}$ is symmetric, positive definite and $\|\tilde{f}\|_2 < \infty$. The operator $\widetilde{\mathcal{F}}$ is induced by $\tilde{f}$ and is given by

$$\widetilde{\mathcal{F}}f(\tau) = \int_0^1 \tilde{f}(\tau,\sigma)f(\sigma)\mathrm{d}\sigma$$

for each $\tau \in [0,1]$ and $f \in L^2([0,1],\mathbb{R})$.

The minimum distance

Lemma

Suppose that $\mathcal{F} : L^2([0,1],\mathbb{R}) \to L^2([0,1],\mathbb{C})$ is a Hilbert-Schmidt operator. Then

$$\begin{aligned} M^2(\mathcal{F}) &= \int_{-\pi}^{\pi} |||\mathcal{F}_\omega - \mathcal{F}|||_2^2 \mathrm{d}\omega \\ &= \int_{-\pi}^{\pi} |||\mathcal{F}_\omega - \widetilde{\mathcal{F}}|||_2^2 \mathrm{d}\omega + \int_{-\pi}^{\pi} |||\widetilde{\mathcal{F}} - \mathcal{F}|||_2^2 \mathrm{d}\omega. \end{aligned}$$

In particular, $M^2(\mathcal{F})$ is minimized at $\widetilde{\mathcal{F}}$.

The minimum distance in terms of f_ω

Let us denote $m^2 = M^2(\widetilde{\mathcal{F}})$.

Lemma

The minimum distance is given by

$$m^2 = \int_{-\pi}^{\pi} |||\mathcal{F}_\omega - \widetilde{\mathcal{F}}|||_2^2 \mathrm{d}\omega$$
$$= \int_0^1 \int_0^1 \left[\int_{-\pi}^{\pi} |f_\omega(\tau,\sigma)|^2 \mathrm{d}\omega - \frac{1}{2\pi} \left| \int_{-\pi}^{\pi} f_\omega(\tau,\sigma) \mathrm{d}\omega \right|^2 \right] \mathrm{d}\tau \mathrm{d}\sigma,$$

where f_ω is the spectral density kernel.

Testing for white noise

Since $m^2 = 0$ if $\{X_t\}_{t\in\mathbb{Z}}$ is white noise, we want to test the null hypothesis

$$H_0 \ : \ m^2 = 0$$

against the alternative

$$H_A \ : \ m^2 > 0.$$

To perform this test, we need an estimator for the minimum distance m^2.

fDFT and periodogram

The *functional discrete Fourier transform* (fDFT) of the observations $\{X_t\}_{t=0}^{T-1}$ is defined as

$$\widetilde{X}_{\omega}^{(T)} = \frac{1}{\sqrt{2\pi T}} \sum_{t=0}^{T-1} X_t \exp(-i\omega t)$$

for $\omega \in \mathbb{R}$ and $T \geq 1$.

The *periodogram kernel* is defined as

$$p_{\omega}^{(T)}(\tau, \sigma) = [\widetilde{X}_{\omega}^{(T)}(\tau)][\overline{\widetilde{X}_{\omega}^{(T)}(\sigma)}]$$

for each $\omega \in \mathbb{R}$, $\tau, \sigma \in [0, 1]$ and $T \geq 1$.

Riemann sums of periodograms

The periodogram kernel is not a consistent estimator of the spectral density kernel.

We propose to use the Riemann sums of periodograms. Let us define

$$S_{T,1}(\tau,\sigma) = \frac{1}{T}\sum_{k=1}^{\lfloor T/2 \rfloor}(p_{\omega_k}^{(T)}(\tau,\sigma) + \bar{p}_{\omega_k}^{(T)}(\tau,\sigma))$$

and

$$S_{T,2}(\tau,\sigma) = \frac{2}{T}\sum_{k=1}^{\lfloor T/2 \rfloor} p_{\omega_k}^{(T)}(\tau,\sigma)\bar{p}_{\omega_{k-1}}^{(T)}(\tau,\sigma),$$

where $\omega_k = 2\pi k/T$ with $k = 1, \ldots, \lfloor T/2 \rfloor$ and $T \geq 1$.

Intuition behind the Riemann sums

Using the results of Panaretos and Tavakoli (2013), we obtain

$$\mathsf{E}\, S_{T,1}(\tau,\sigma) \to \frac{1}{2\pi}\int_{-\pi}^{\pi} f_\omega(\tau,\sigma)\mathrm{d}\omega$$

and

$$\mathsf{E}\, S_{T,2}(\tau,\sigma) \to \frac{1}{2\pi}\int_{-\pi}^{\pi} |f_\omega(\tau,\sigma)|^2\mathrm{d}\omega$$

a.e. as $T \to \infty$.

Estimator of minimum distance

We have that

$$m^2 = \int_0^1 \int_0^1 \Big[\underbrace{\int_{-\pi}^{\pi} |f_\omega(\tau,\sigma)|^2 \mathrm{d}\omega}_{\approx 2\pi \, \mathsf{E}\, S_{T,2}(\tau,\sigma)} - \frac{1}{2\pi} \Big| \underbrace{\int_{-\pi}^{\pi} f_\omega(\tau,\sigma) \mathrm{d}\omega}_{\approx 2\pi \, \mathsf{E}\, S_{T,1}(\tau,\sigma)} \Big|^2 \Big] \mathrm{d}\tau \mathrm{d}\sigma.$$

Hence, we define the estimator of m^2 as

$$\hat{m}_T^2 = 2\pi \int_0^1 \int_0^1 [S_{T,2}(\tau,\sigma) - |S_{T,1}(\tau,\sigma)|^2] \mathrm{d}\tau \mathrm{d}\sigma$$

for $T \geq 1$.

Asymptotic distribution of the estimator

Theorem

Suppose that $\{X_k\}_{k\in\mathbb{Z}}$ *is a strictly stationary time series with values in* $L^2([0,1],\mathbb{R}$, $\mathrm{E}\,\|X_0\|_2^k < \infty$ *for each* $k \geq 1$*,*

(i) $\sum_{t_1,t_2,t_3=-\infty}^{\infty} \mathrm{E}[\|X_{t_1}X_{t_2}\|_1 \|X_{t_3}X_{t_0}\|_1] < \infty$,

(ii) $\sum_{t_1,t_2,\ldots,t_{k-1}=-\infty}^{\infty}(1+|t_j|)\|\,\mathrm{cum}(X_{t_1},\ldots,X_{t_{k-1}},X_0)\|_2 < \infty$ *for* $j = 1,2,\ldots,k-1$ *and all* $k \geq 1$.

Then

$$\sqrt{T}(\hat{m}_T^2 - m^2) \xrightarrow{d} N(0,v^2) \quad \textit{as} \quad T \to \infty,$$

where v^2 *under the null hypothesis is given by*

$$v_H^2 = 8\pi^2 \left[\int_0^1 \int_0^1 |f_0(\tau,\sigma)|^2 \mathrm{d}\tau \mathrm{d}\sigma \right]^2.$$

Properties of the estimator

The definition

$$S_{T,2}(\tau,\sigma) = \frac{2}{T} \sum_{k=1}^{\lfloor T/2 \rfloor} p_{\omega_k}^{(T)}(\tau,\sigma)\bar{p}_{\omega_{k-1}}^{(T)}(\tau,\sigma)$$

with ω_{k-1} makes the estimator unbiased, but the value of the test statistic might be complex.

Properties of the estimator (2)

The test statistic is real with a probability converging to 1 as $T \to \infty$. Hence, our main result might be viewed as

$$\Re(\sqrt{T}(\hat{m}_T^2 - m^2)) \xrightarrow{d} N(0, v^2) \quad \text{as} \quad T \to \infty$$

and

$$\Im(\sqrt{T}(\hat{m}_T^2 - m^2)) \xrightarrow{p} 0 \quad \text{as} \quad T \to \infty.$$

Rejection rule

We reject the null hypothesis if

$$\Re(\hat{m}_T^2) > \frac{\hat{v}_H}{\sqrt{T}} z_{1-\alpha},$$

where $z_{1-\alpha}$ denotes the $(1-\alpha)$-quantile of the standard normal distribution and $\hat{v}_H$ is an appropriate estimator of v_H.

Estimation of the asymptotic variance

Since

$$\mathrm{E}\, S_{T,2}(\tau,\sigma) \to \frac{1}{2\pi}\int_{-\pi}^{\pi} |f_\omega(\tau,\sigma)|^2 \mathrm{d}\omega$$

a.e. as $T \to \infty$, a consistent estimator of the asymptotic variance under the null hypothesis is given by

$$\hat{v}_H^2 = 2\pi \int_0^1 \int_0^1 S_{T,2}(\tau,\sigma)\mathrm{d}\tau\mathrm{d}\sigma.$$

Approximation of the power function

We have that

$$P\Big(\Re(\hat{m}_T^2) > \frac{\hat{v}_{H,T}}{\sqrt{T}} z_{1-\alpha}\Big) \approx \Phi\Big(\sqrt{T}\frac{m^2}{v} - \frac{v_H}{v} z_{1-\alpha}\Big)$$

and this shows that our test is consistent.

Simulation study

- Simulation setup is similar to that of Zhang (2016).
- The sample size T is chosen to be equal to 128, 256, 512 or 1024.
- The number of the Monte Carlo replications is 1000.
- The data is generated on a grid on 1000 equispaced points in $[0,1]$ for each functional observation.
- The kernels $S_{T,1}$ and $S_{T,2}$ are calculated at 1000×1000 equispaced points in $[0,1]^2$.

Functional time series under the null hypothesis

We simulate iid

- standard Brownian motions;
- Brownian bridges.

We also simulate the FARCH(1) process defined as

$$X_t(\tau) = \varepsilon_t(\tau)\sqrt{\tau + \int_0^1 c_\psi \exp\left(\frac{\tau^2 + \sigma^2}{2}\right) X_{t-1}^2(\sigma)\mathrm{d}\sigma}$$

for $t \geq 1$ and $\tau \in [0,1]$, where $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ are iid standard Brownian motions and $c_\psi = 0.3418$.

Empirical rejection probabilities under the null

	Brownian motion			Brownian bridge			FARCH(1)		
T	10%	5%	1%	10%	5%	1%	10%	5%	1%
128	9.5 (11.0)	4.8 (4.2)	1.1 (0.8)	10.8 (11.0)	5.3 (5.4)	0.8 (1.1)	11.1 (10.7)	5.7 (5.9)	0.8 (0.9)
256	9.6 (10.0)	5.1 (4.2)	1.3 (0.9)	10.3 (9.5)	5.4 (4.8)	0.9 (0.7)	10.9 (11.1)	5.5 (5.2)	0.7 (0.9)
512	10.1 (9.9)	5.1 (4.7)	0.8 (0.6)	9.7 (10.3)	5.1 (5.9)	1.0 (1.3)	10.9 (11.1)	5.3 (4.9)	0.8 (0.7)
1024	9.8 (10.0)	4.9 (4.9)	0.9 (0.8)	9.9 (9.9)	5.2 (5.1)	0.8 (1.1)	10.5 (9.8)	5.2 (4.8)	0.7 (1.2)

The numbers in brackets give the corresponding results of the test of Zhang (2016)

Functional time series under the alternative hypothesis

We simulate observations from the FAR(1) model

$$X_t - \mu = \rho(X_{t-1} - \mu) + \varepsilon_t$$

for $t \geq 1$, where $\rho : L^2([0,1], \mathbb{R}) \to L^2([0,1], \mathbb{R})$ is an integral operator defined by

$$\rho f(\tau) = \int_0^1 \mathcal{K}(\tau, \sigma) f(\sigma) \mathrm{d}\sigma$$

for $f \in L^2([0,1], \mathbb{R})$ and $\tau \in [0,1]$ with some kernel $\mathcal{K} \in L^2([0,1]^2, \mathbb{R})$ and iid errors $\{\varepsilon_t\}_{t \in \mathbb{Z}}$.

Functional time series under the alternative hypothesis

We consider four different FAR(1) models where the errors are either Brownian motions or Brownian bridges and the kernel of the integral operator is either the Gaussian kernel

$$\mathcal{K}_G(\tau, \sigma) = c_G \exp\left(\frac{\tau^2 + \sigma^2}{2}\right)$$

or the Wiener kernel

$$\mathcal{K}_W(\tau, \sigma) = c_W \min(\tau, \sigma),$$

where the constants c_G and c_W are chosen such that the correspoding Hilbert-Schmidt norm is equal to 0.3.

Empirical rejection probabilities under the alternative

ε_t	Brownian motion					
$\mathcal{K}$	Gaussian			Wiener		
T	10%	5%	1%	10%	5%	1%
128	82.6 (86.1)	80.7 (83.7)	65.9 (58.5)	87.6 (89.9)	82.4 (83.1)	66.9 (59.7)
256	99.0 (99.6)	98.2 (99.2)	98.2 (99.0)	99.4 (99.9)	98.3 (99.5)	94.2 (98.6)
512	99.8 (99.7)	99.6 (99.5)	99.6 (99.0)	99.9 (99.9)	99.9 (99.8)	99.6 (99.1)
1024	100.0 (100.0)	99.9 (100.0)	99.7 (99.8)	100.0 (100.0)	100.0 (99.8)	99.8 (99.5)

The numbers in brackets give the corresponding results of the test of Zhang (2016)

Empirical rejection probabilities under the alternative

ε_t	Brownian bridge					
$\mathcal{K}$	Gaussian			Wiener		
T	10%	5%	1%	10%	5%	1%
128	80.1 (79.2)	77.4 (68.3)	60.1 (54.4)	87.6 (80.2)	79.9 (65.8)	61.2 (58.1)
256	100.0 (100.0)	97.0 (98.2)	95.5 (97.2)	99.9 (100.0)	98.3 (99.1)	98.1 (98.8)
512	100.0 (100.0)	99.3 (98.7)	99.3 (98.1)	100.0 (100.0)	100.0 (100.0)	98.8 (99.1)
1024	100.0 (100.0)	100.0 (100.0)	100.0 (99.4)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)

The numbers in brackets give the corresponding results of the test of Zhang (2016)

Concluding remarks

- We propose a frequency-domain based test for white noise (uncorrelatedness) in functional time series with a simple asymptotic distribution.
- Our test does not require a choice of the lag truncation number or the number of functional principal components.
- Critical values of the test statistic can be easily obtained, there is no need for bootstrap.
- The finite sample performance in testing for white noise is very similar to that of Zhang (2016).

Future work

- Adapt the test for the situation when we do not observe the random elements directly but we only have residuals, i.e. adapt the test for model diagnostic checking.
- Establish the asymptotic distribution of the test statistic under simpler and weaker assumptions.

Thank you!