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

Vilnius, December 22, 2017

Outline

Motivation and background

New test for white noise

Finite sample performance

Conclusions and future work

When do want to test for white noise?

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

Two approaches

- Time-domain tests based on autocovariances or autocorrelations.
- Frequency-domain tests based on spectral densities.

Autocovariances and autocorrelations

- $\{X_t\}_{t\in\mathbb{Z}}$ is a stationary sequence of random variables.
- $\{\gamma_h\}_{h\in\mathbb{Z}}$ are autocovariances defined by $\gamma_h = \mathrm{Cov}(X_h, X_0)$, for each $h \in \mathbb{Z}$.
- $\{\rho_h\}_{h\in\mathbb{Z}}$ are autocorrelations defined by $\rho_h = \gamma_h/\gamma_0$ for each $h \in \mathbb{Z}$.

Time-domain approach

The idea is to investigate the autocovariances or autocorrelations and to check if $\gamma_h = 0$ or $\rho_h = 0$ for each $h \neq 0$.

Estimating autocovariances

Definition

The sample autocovariance is defined by

$$\hat{\gamma}_h = n^{-1} \sum_{j=1}^{n-|h|} (X_{j+|h|} - \overline{X})(X_j - \overline{X})$$

for $|h| < n$, where $\overline{X} = n^{-1} \sum_{j=1}^{n} X_j$.

γ_h is estimated using $n - |h|$ observations with $|h| < n$.

Estimating autocorrelations

Definition

The sample autocorrelation is defined by

$$\hat{\rho}_h = \frac{\hat{\gamma}_h}{\hat{\gamma}_0}$$

for $|h| < n$.

The portmanteau test

Box and Pierce (1970) proposed to use the following test statistic

$$Q_{BP} = n \sum_{j=1}^{h} \hat{\rho}_j^2,$$

where the parameter h is called the lag truncation number.

If X_t's are iid random variables, then $Q_{BP} \xrightarrow{d} \chi_h^2$ as $n \to \infty$.

If X_t's are residuals of ARMA(p, q) model with iid errors, then $Q_{BP} \xrightarrow{d} \chi^2_{h-(p+q)}$ as $n \to \infty$.

Univariate and multivariate cases

Univariate case

- Box and Pierce (1970), Pierce (1972), Davies, Triggs and Newbold (1977), Ljung and Box (1978), McLeod and Li (1983), Ljung (1986), Peña and Rodríguez (2002).

Multivariate case

- Chitturi (1974, 1976), Hosking (1980, 1981), Li and MacLeod (1981), Mahdi and MacLeod (2010).

Noncorrelation vs independence

- Box and Pierce (1970) proposed test works under the assumption of iid random variables.
- If the errors are uncorrelated but not independent, the test is not reliable (see Romano and Thombs (1996) and Francq, Roy and Zakoïan (2005)).

Frequency-domain approach

The idea is to compare the spectral density corresponding to the sequence of the random variables and the spectral density of white noise.

Spectral density

Definition

The spectral density is a discrete-time Fourier transform of $\{\gamma_h\}_{j\in\mathbb{Z}}$ given by

$$f(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j e^{-i\omega j}$$

for each $\omega \in [-\pi, \pi]$ provided that $\sum_{j=-\infty}^{\infty} |\gamma_j|$, where $i = \sqrt{-1}$.

Inverse of discrete-time Fourier transform

Proposition

Suppose that the sequence $\{\gamma_h\}_{h\in\mathbb{Z}}$ is absolutely summable. Then

$$\gamma_h = \int_{-\pi}^{\pi} f(\omega) e^{i\omega h} d\omega$$

for each $h \in \mathbb{Z}$, where $i = \sqrt{-1}$.

The spectral density f and $\{\gamma_h\}_{h\in\mathbb{Z}}$ form a Fourier pair.

Spectral density of white noise

If X_t's are uncorrelated, i.e. $\gamma_h = 0$ for each $h \neq 0$, then

$$f(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j e^{-i\omega j} = \frac{\gamma_0}{2\pi}$$

for each $\omega \in [-\pi, \pi]$.

Frequency-domain test for white noise

Hong (1996) proposes a test for white noise based on the divergence measure

$$Q^2(f, f_0) = 2\pi \int_{-\pi}^{\pi} \left| \frac{f(\omega)}{\gamma_0} - \frac{1}{2\pi} \right|^2 d\omega.$$

f is estimated using a kernel estimator and the appropriately standardised test statistic, under certain assumptions, is asymptotically standard normal if X_t's are iid.

Frequency-domain approach

Developed by Durlauf (1991), Hong (1996), Deo (2000), Chen and Deo (2004), Dette, Kinsvater and Vetter (2010), Shao (2011).

Tests for functional time series

Time-domain tests for independence

- Gabrys and Kokoszka (2007), Gabrys, Horváth, and Kokoszka (2010);
- Horváth, Hušková, and Rice (2013).

Frequency-domain test for white noise

- Zhang (2016).

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 truncation number 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 autocovariance 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 nor 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

- A frequency-domain test for white noise in functional time series.
- The asymptotic distribution of our test statistic is simple and our approach does not need bootstrap to obtain the critical values.
- We do not need to choose the number of functional principal components nor do we need to choose 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 stationary $L^2[0,1]$-valued random elements.

Definition

The autocovariance kernels $\{\gamma_h\}_{h\in\mathbb{Z}}$ of $\{X_t\}_{t\in\mathbb{Z}}$ are defined by

$$\gamma_h(\tau,\sigma) = \mathrm{Cov}[X_h(\tau), X_0(\sigma)]$$

for each $\tau,\sigma \in [0,1]$ and $h \in \mathbb{Z}$.

White noise and hypothesis

Definition

$\{X_t\}_{t\in\mathbb{Z}}$ is white noise if X_t's are uncorrelated, i.e. if $\gamma_h = 0$ for each $h \neq 0$.

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

Spectral density kernel

Definition

The spectral density kernel is a discrete-time Fourier transform of $\{\gamma_h\}_{h\in\mathbb{Z}}$ defined by

$$f_\omega = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma_j e^{-i\omega j}$$

for $\omega \in [-\pi, \pi]$ provided that $\sum_{j=-\infty}^{\infty}\|\gamma_j\|_2 < \infty$.

If X_t's are uncorrelated, then $f_\omega = \gamma_0/(2\pi)$.

The spectral density kernel is investigated by Panaretos and Tavakoli (2013).

Distance to white noise

We measure the distance between f_ω, $\omega \in [-\pi, \pi]$, and $\gamma_0/(2\pi)$ using the following distance function

$$m^2 = \int_{-\pi}^{\pi} \|f_\omega - \gamma_0/(2\pi)\|_2^2 d\omega.$$

Also, we have that

$$m^2 = \int_{-\pi}^{\pi} \|f_\omega\|_2^2 d\omega - \frac{1}{2\pi}\|\gamma_0\|_2^2 = \frac{1}{2\pi}\sum_{j\neq 0}\|\gamma_j\|_2^2.$$

The last equality clearly shows that the distance is equal to 0 if and only if X_t's are uncorrelated.

Hypothesis

The hypothesis that we test is as follows

$$H_0 \ : \ m^2 = 0 \quad \text{versus} \quad H_1 \ : \ m^2 > 0.$$

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

fDFT and periodogram kernel

Definition

The functional discrete Fourier transform (fDFT) is defined as

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

for $\omega \in [-\pi, \pi]$ and $T \geq 1$.

Definition

The periodogram kernel is defined as

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

for each $\tau, \sigma \in [0, 1]$, where $\bar{x}$ is the complex conjugate of $x \in \mathbb{C}$.

Estimator of minimum distance

To estimate the distance m^2, we avoid direct estimation of the spectral density kernel and propose to use sums of inner-products and norms of the periodogram kernels.

The estimator of m^2 is defined as

$$\hat{m}_T = 2\pi \left[\frac{2}{T} \sum_{k=2}^{\lfloor T/2 \rfloor} \langle p_{\omega_k}^{(T)}, p_{\omega_{k-1}}^{(T)} \rangle - \left\| \frac{1}{T} \sum_{k=1}^{\lfloor T/2 \rfloor} [p_{\omega_k}^{(T)} + \bar{p}_{\omega_k}^{(T)}] \right\|_2^2 \right],$$

where ω_k are the Fourier frequencies defined by $\omega_k = 2\pi k/T$ for $1 \leq k \leq \lfloor T/2 \rfloor$ and $T \geq 1$.

Intuition behind Riemann sums

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

$$\mathsf{E}\left[\frac{2}{T}\sum_{k=2}^{\lfloor T/2\rfloor}\langle p_{\omega_k}^{(T)}, p_{\omega_{k-1}}^{(T)}\rangle\right] \to \frac{1}{2\pi}\int_{-\pi}^{\pi}\|f_\omega\|_2^2 \mathrm{d}\omega$$

and

$$\mathsf{E}\left\|\frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor}[p_{\omega_k}^{(T)} + \bar{p}_{\omega_k}^{(T)}]\right\|_2^2 \to \frac{1}{(2\pi)^2}\|\gamma_0\|_2^2$$

as $T \to \infty$.

Intuition behind estimator

Recall that

$$m^2 = \int_{-\pi}^{\pi} \|f_\omega\|_2^2 \mathrm{d}\omega - \frac{1}{2\pi}\|\gamma_0\|_2^2.$$

If

$$2\pi \,\mathrm{E}\Big[\frac{2}{T}\sum_{k=2}^{\lfloor T/2 \rfloor}\langle p_{\omega_k}^{(T)}, p_{\omega_{k-1}}^{(T)}\rangle\Big] \approx \int_{-\pi}^{\pi} \|f_\omega\|_2^2 \mathrm{d}\omega$$

and

$$2\pi \,\mathrm{E}\Big\|\frac{1}{T}\sum_{k=1}^{\lfloor T/2 \rfloor}[p_{\omega_k}^{(T)} + \bar{p}_{\omega_k}^{(T)}]\Big\|_2^2 \approx \frac{1}{2\pi}\|\gamma_0\|_2^2,$$

we might expect that

$$\hat{m}_T \approx m^2.$$

Asymptotic distribution of the estimator

Theorem

Suppose that

(i) $\{X_t\}_{t\in\mathbb{Z}}$ *is strictly stationary sequence of* $L^2[0,1]$*-valued random elements such that* $\operatorname{E}\|X_0\|_2^k < \infty$ *for each* $k \geq 1$*;*

(ii) $\int_0^1 \int_0^1 \sum_{t_1,t_2,t_3\in\mathbb{Z}} |\operatorname{E}[X_{t_1}(\tau)X_{t_2}(\sigma)X_{t_3}(\tau)X_0(\sigma)]| d\tau d\sigma < \infty$*;*

(iii) $\sum_{t_1,\ldots,t_{k-1}\in\mathbb{Z}}(1+|t_j|)\|\operatorname{cum}(X_{t_1},\ldots,X_{t_{k-1}},X_0)\|_2 < \infty$ *for* $j = 1,2,\ldots,k-1$ *and all* $k \geq 2$*.*

Then

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

where v^2 *is the asymptotic variance. Under the null hypothesis,* v^2 *is given by* $v^2_{H_0} = 8\pi^2\|f_0\|_2^4$*.*

Rejection rule

A consistent estimator of the asymptotic standard deviation under the null hypothesis is given by

$$\widehat{v_{H_0}} = \frac{4\pi}{T} \sum_{k=2}^{\lfloor T/2 \rfloor} \langle p_{\omega_k}^{(T)}, p_{\omega_{k-1}}^{(T)} \rangle$$

for $T \geq 1$.

The null hypothesis is rejected if

$$\hat{m}_T > \frac{\widehat{v_{H_0}}}{\sqrt{T}} z_{1-\alpha},$$

where $z_{1-\alpha}$ is the $(1-\alpha)$-quantile of the standard normal distribution.

Approximation of the power function

We have that

$$P\Big(\hat{m}_T > \frac{\widehat{v_{H_0}}}{\sqrt{T}} z_{1-\alpha}\Big) \approx \Phi\Big(\sqrt{T}\frac{m^2}{v} - \frac{v_{H_0}}{v} z_{1-\alpha}\Big)$$

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 periodogram kernels 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 $\rho : L^2[0,1] \to L^2[0,1]$ is an integral operator defined by

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

for $f \in L^2[0,1]$ and $\tau \in [0,1]$ with some kernel $\mathcal{K} \in L^2[0,1]^2$ 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 (noncorrelation) in functional time series with a simple asymptotic distribution.
- Our test neither requires a choice of the lag truncation number nor the choice of 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.