A Max-Correlation White Noise Test for Weakly Dependent Time Series

Jonathan B. Hill\textsuperscript{1} Kaiji Motegi\textsuperscript{2}

\textsuperscript{1}University of North Carolina at Chapel Hill
\textsuperscript{2}Kobe University

15th International Conference, WEAI
March 22, 2019
White noise = zero autocorrelations at all lags.

Testing the white noise hypothesis is of practical use:

1. Testing the weak form efficiency of stock markets. A rejection of the white noise hypothesis of stock returns would serve as a signal of arbitrage opportunity (Hill and Motegi, 2019).

2. Residual diagnostics of time series regressions. If a model fits data well, then the resulting residual should be white noise.

It is relatively easy to test the independence or MDS hypothesis, since the null hypothesis $H_0$ is strong enough to establish asymptotic theory (Box and Pierce, 1970).

It is challenging to test the white noise hypothesis, since $H_0$ is weak (only serial uncorrelatedness) and there are infinitely many zero restrictions.
Define

\[ \hat{\mu}_n = \frac{1}{n} \sum_{t=1}^{n} y_t, \quad \hat{\gamma}_n(0) = \frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{\mu}_n)^2, \]

\[ \hat{\gamma}_n(h) = \frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{\mu}_n)(y_{t-h} - \hat{\mu}_n), \quad \hat{\rho}_n(h) = \hat{\gamma}_n(h)/\hat{\gamma}_n(0). \]

Consider the classical Q-test:

\[ \hat{Q}_n = n \sum_{h=1}^{L} w_n(h) \hat{\rho}_n^2(h) \xrightarrow{d} \chi^2_L. \]

\[ w_n(h) = 1 \text{ in Box and Pierce (1970)}. \]
Recall:

\[ \hat{Q}_n = n \sum_{h=1}^{L} w_n(h) \hat{\rho}_n^2(h) \xrightarrow{d} \chi^2_L. \]

Recall that the white noise hypothesis is:

\[ H_0 : \rho(h) = 0 \quad \forall h \geq 1. \]

Two reasons why the Q-test is not a white noise test:

1. Asymptotic \( \chi^2 \) property requires the asymptotic independence of \( \{\hat{\rho}_n(1), \ldots, \hat{\rho}_n(L)\} \), which holds only when \( \{y_t\} \) is \textbf{serially independent}.

2. The Q-test cannot capture autocorrelations beyond lag \( L \).
Research question: How can we establish a formal white noise test?

1. How to formulate a test statistic?
2. How to choose a lag length?
3. How to compute a p-value?

Proposed solution: A bootstrapped max-correlation test with automatic lag selection.

1. The test statistic is $\hat{T}_n(\mathcal{L}_n^*) = \sqrt{n} \times \max_{1 \leq h \leq \mathcal{L}_n^*} |\hat{\rho}_n(h)|$.
3. An approximate p-value of Shao’s (2011) dependent wild bootstrap.
Table of contents

1 Introduction
2 Methodology
   - Max-correlation test statistic
   - P-value computation
   - Data-driven lag selection
3 Monte Carlo simulation
4 Empirical application
5 Conclusion
Methodology

- \( H_0 : \rho(h) = 0 \) for all \( h \geq 1 \).

- We propose a max-correlation test statistic:

\[
\hat{T}_n = \sqrt{n} \times \max \{ |\hat{\rho}_n(1)|, \ldots, |\hat{\rho}_n(L_n)| \},
\]

where \( L_n \to \infty \) as \( n \to \infty \) and \( L_n = o(n) \).

- An advantage of the maximum relative to the sum of squares is that the maximum leads to sharper performance when there exist autocorrelations at remote lags (e.g., seasonality).

- Shao (2011) proposed the dependent wild bootstrap (DWB), which is valid for white noise.

- We use DWB to compute an approximate p-value.
Methodology

1. Set a block size $b_n$ (typically $b_n = \sqrt{n}$).

2. Generate iid $\{\xi_1, \xi_2, \ldots, \xi_{n/b_n}\}$. Define an auxiliary variable:

$$\omega = \begin{bmatrix} \xi_1, \ldots, \xi_1 \quad \xi_2, \ldots, \xi_2 \quad \cdots \quad \xi_{n/b_n}, \ldots, \xi_{n/b_n} \end{bmatrix}'.$$

3. Compute bootstrapped autocorrelations:

$$\hat{\rho}_n^{(dw)}(h) = \frac{1}{\hat{\gamma}_n(0)} \times \frac{1}{n} \sum_{t=h+1}^{n} \omega_t [y_t y_{t-h} - \hat{\gamma}_n(h)], \quad h = 1, \ldots, L_n,$$

and $$\hat{T}_n^{(dw)} = \sqrt{n} \times \max_{1 \leq h \leq L_n} |\hat{\rho}_n^{(dw)}(h)|.$$

4. Repeat Steps 2-3 $M$ times and compute the bootstrapped p-value $$\hat{p}_{n,M}^{(dw)} = \left(1/M\right) \sum_{i=1}^{M} I(\hat{T}_{n,i}^{(dw)} \geq \hat{T}_n).$$

5. If $\hat{p}_{n,M}^{(dw)} < \alpha$, then reject $H_0$ at the $100\alpha\%$ level. Otherwise accept $H_0$. 
Choose the upper bound $\tilde{L}_n = o(n/\ln n)$ and a tuning parameter $q > 0$.

For each candidate $L \in \{1, \ldots, \tilde{L}_n\}$, compute the penalized max-correlation test statistic:

$$\hat{T}_n^P(L) = \hat{T}_n(L) - P_n(L),$$

where

$$P_n(L) = \begin{cases} \sqrt{(\ln n)L} & \text{if } \hat{T}_n(L) \leq \sqrt{q \ln n}, \\ \sqrt{2L} & \text{if } \hat{T}_n(L) > \sqrt{q \ln n}. \end{cases}$$

The optimal lag length $L^*_n$ is given by

$$L^*_n = \min \left\{ \arg \max_{L \in \{1, \ldots, \tilde{L}_n\}} \hat{T}_n^P(L) \right\}.$$
Monte Carlo simulation

- $\hat{T}^{dw}(L^*) = $ the proposed max-correlation test with DWB and automatic lag selection.

- $CvM^{dw} = $ Shao’s (2011) Cramér-von Mises [CvM] test with DWB, where the test statistic is

  $$C_n = n \int_0^\pi \left[ \sum_{h=1}^{n-1} \hat{\gamma}_n(h) \frac{\sin(h\lambda)}{h\pi} \right]^2 d\lambda.$$ 

- $CvM^{dw}$ turned out to be the strongest competitor to $\hat{T}^{dw}(L^*)$ among many other alternatives.
Monte Carlo simulation

<table>
<thead>
<tr>
<th>Model</th>
<th>IID $e_t = \nu_t$</th>
<th>AR(1) $e_t = 0.7e_{t-1} + \nu_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilin $y_t = 0.5e_{t-1}y_{t-2} + e_t$</td>
<td>$H_0$</td>
<td>$H_1$ (adjacent)</td>
</tr>
<tr>
<td>MA(6) $y_t = e_t + 0.25e_{t-6}$</td>
<td>$H_1$ (remote)</td>
<td>-</td>
</tr>
</tbody>
</table>

- $\nu_t \sim i.i.d. \, N(0, 1)$.
- Sample size is $n \in \{250, 1000\}$.
- For $\hat{T}^{dw}(\mathcal{L}^*)$, we set $\bar{L}_n = [1.5 \times n/(\ln n)^{4/3}]$ so that $\bar{L}_n \in \{38, 114\}$ for $n \in \{250, 1000\}$.
- Nominal size is $\alpha = 0.05$.
- The number of bootstrap samples is $M = 500$.
- The number of Monte Carlo iterations is $J = 1000$. 
<table>
<thead>
<tr>
<th>Rejection frequencies</th>
<th>( n = 250 )</th>
<th>( n = 1000 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{T}^{dw}(L^*) )</td>
<td>( C_{vM}^{dw} )</td>
<td>( \hat{T}^{dw}(L^*) )</td>
</tr>
<tr>
<td>( H_0 )</td>
<td>0.054</td>
<td>0.085</td>
</tr>
<tr>
<td>( H_1 ) (adjacent)</td>
<td>0.784</td>
<td>0.572</td>
</tr>
<tr>
<td>( H_1 ) (remote)</td>
<td>0.261</td>
<td>0.080</td>
</tr>
</tbody>
</table>

- Both tests have fairly accurate empirical size.
- \( \hat{T}^{dw}(L^*) \) is more powerful than \( C_{vM}^{dw} \), especially when there exists a remote autocorrelation.
Empirical application

- Testing the white noise hypothesis of log returns of major stock price indices:
  - S&P 500 (SPX)
  - FTSE 100 (FTSE)
  - Nikkei 225 (N225)
  - SSE Composite Index (SSEC)

- Daily close-to-close log returns from 11th Jan. 2018 through 10th Jan. 2019 (one year) are used.

- We perform $\hat{T}_{dw}(\mathcal{L}^*)$ with $q = 3.25$, $\bar{L}_n = [1.5 \times n/(\ln n)^{4/3}]$, and $M = 1000$.

- We also perform $CvM_{dw}$ for comparison.
Empirical application

SPX

FTSE

N225

SSEC
Empirical application

<table>
<thead>
<tr>
<th></th>
<th>$\mathcal{T}^{dw}(\mathcal{L}^*)$</th>
<th>$C_{VM}^{dw}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$\mathcal{L}_n^*$</td>
</tr>
<tr>
<td>SPX</td>
<td>250</td>
<td>1</td>
</tr>
<tr>
<td>FTSE</td>
<td>253</td>
<td>1</td>
</tr>
<tr>
<td>N225</td>
<td>246</td>
<td>1</td>
</tr>
<tr>
<td>SSEC</td>
<td>243</td>
<td>1</td>
</tr>
</tbody>
</table>

- For FTSE, the white noise hypothesis is **rejected** at the 10% level.
- It suggests that the recent UK stock market might be **inefficient** due to the large negative autocorrelation at lag 1.
- For other cases, weak form efficiency seems to hold.
Conclusion

- It is rather challenging to establish a formal white noise test due to the relatively weak null hypothesis and the infinitely many zero restrictions.

- Key issues are how to formulate a test statistic, how to choose a lag length, and how to compute a p-value.

- Our proposed solution is the bootstrapped max-correlation test with automatic lag selection.
  
  - The test statistic is based on the maximum of sample autocorrelations.
  - Escanciano and Lobato’s (2009) automatic lag selection is utilized.
  - Shao’s (2011) dependent wild bootstrap is used.
Our test is asymptotically valid under the null hypothesis of white noise and consistent under the alternative hypothesis.

Since our test allows for various filters such as AR and GARCH, it can be used for residual diagnostics.

The simulation results highlight that the proposed test has sharp size and high power. In particular, it is remarkably powerful against remote autocorrelations.

The empirical application on the weak form efficiency of stock markets suggests that the recent FTSE market might be inefficient due to the large negative autocorrelation at lag 1.
References


