LECTURE ON

TIME SERIES DIAGNOSTIC TESTS

CHUNG-MING KUAN

Institute of Economics
Academia Sinica

This version: March 25, 2003

Address for correspondence: Institute of Economics, Academia Sinica, Taipei 115, Taiwan.

E-mail: ckuan@econ.sinica.edu.tw; URL: www.sinica.edu.tw/as/ssrc/ckuan

Contents

1 Introduction

Diagnostic tests are important tools for the time series modeling. Most of existing diagnostic tests are designed to examine the dependence structure of a time series. If a time series is serially uncorrelated, no linear function of the variables in the information set can account for the behavior of the current variable. If a time series is a martingale difference sequence (to be defined later), no function, linear or nonlinear, of the variables in the information set can characterize the current variable. Yet a function of the current variable may still depend on the past information. A serially independent time series implies that there is absolutely no relationship between the current variable and past information. Diagnostic testing on raw data series thus provides information regarding how these data might be modeled. When a model is estimated, we can also apply diagnostic tests to model residuals and check whether the estimated model can be further improved upon.

In practice, there are three classes of diagnostic tests, each focusing on a specific dependence structure of a time series. The tests of serial uncorrelatedness include the well known Q tests of Box and Pierce (1970) and Ljung and Box (1978), the robust Q^* test of Lobato, Nankervis, and Savin (2001), the spectral tests of Durlauf (1991), and the robust spectral test of Deo (2000). There are also tests of the martingale difference hypothesis, including Bierens (1982, 1984), Bierens and Ploberger (1997), Hong (1999), Dominguez and Lobato (2000), Whang (2000, 2001), Kuan and Lee (2004), and Park and Whang (2005). For the hypothesis of serial independence, two leading tests are the variance ratio test of Cochrane (1988) and the so-called BDS test of Brock, Dechert, and Scheinkman (1987); see also Campbell, Lo, and MacKinlay (1997) and Brock, Dechert, Scheinkman, and LeBaron (1996). Skaug and Tjostheim (1993), Pinkse (1998), and Hong (1999) also proposed non-parametric tests of serial independence. In addition to these tests, there are tests of time reversibility, which may also be interpreted as tests of independence; see e.g., Ramsey and Rothman (1996) and Chen, Chou, and Kuan (2000). It has been shown that a test of time reversibility is particularly powerful against asymmetric dependence.

In this note we introduce various diagnostic tests for time series. We will not discuss non-parametric tests because they are not asymptotically pivotal, in the sense that their asymptotic distributions are data dependent. This note proceeds as follows. Section 2.1 focuses on the tests of serial uncorrelatedness. In Section 3, we discuss the tests of the martingale difference hypothesis. Section 4 presents the variance ratio test and the BDS test of serial independence. The tests of time reversibility are discussed in Section 5.

2 Tests of Serial Uncorrelatedness

Given a weakly stationary time series $\{y_t\}$, let μ denote its mean and $\gamma(\cdot)$ denote its autocovariance function, where $\gamma(i) = \text{cov}(y_t y_{t-i})$ for $i = 0, 1, 2, \ldots$. The autocorrelation function $\rho(\cdot)$ is such that $\rho(i) = \gamma(i)/\gamma(0)$. The series $\{y_t\}$ is serially uncorrelated if, and only if, its autocorrelation function is identically zero.

2.1 Q Tests

As testing all autocorrelations is infeasible in practice, existing tests of serial uncorrelatedness focus on a given number of autocorrelations and ignore $\rho(i)$ for large i. The null hypothesis is

$$H_0\colon \rho(1) = \cdots = \rho(m) = 0,$$

where m is a pre-specified number. Let $\bar{y}_T$ denote the sample mean of y_t, $\hat{\gamma}_T(i)$ the i th sample autocovariance:

$$\hat{\gamma}_T(i) = \frac{1}{T} \sum_{t=1}^{T-i} (y_t - \bar{y})(y_{t+i} - \bar{y}),$$

and $\hat{\rho}_T(i) = \hat{\gamma}(i)/\hat{\gamma}(0)$ the i th sample autocorrelation. For notation convenience, we shall suppress the subscript T and simply write $\bar{y}$, $\hat{\gamma}(i)$ and $\hat{\rho}(i)$. Writing $\boldsymbol{\rho}_m = \big(\rho(1), \ldots, \rho(m)\big)'$, the null hypothesis is $\boldsymbol{\rho}_m = \boldsymbol{o}$, and the estimator of $\boldsymbol{\rho}_m$ is $\hat{\boldsymbol{\rho}}_m = \big(\hat{\rho}(1), \ldots, \hat{\rho}(m)\big)'$. Under quite general conditions, it can be shown that as T tends to infinity,

$$\sqrt{T}\big(\hat{\boldsymbol{\rho}}_m - \boldsymbol{\rho}_m\big) \xrightarrow{D} \mathcal{N}(\boldsymbol{o}, \boldsymbol{V}),$$

where $\xrightarrow{D}$ stands for convergence in distribution, and the (i,j) th element of $\boldsymbol{V}$ is

$$v_{ij} = \frac{1}{\gamma(0)^2} \big[c_{i+1,j+1} - \rho(i) c_{1,j+1} - \rho(j) c_{1,i+1} + \rho(i)\rho(j) c_{1,1}\big],$$

with

$$c_{i+1,j+1} = \sum_{k=-\infty}^{\infty} \mathbb{E}\big[(y_t - \mu)(y_{t+i} - \mu)(y_{t+k} - \mu)(y_{t+k+j} - \mu)\big] - \mathbb{E}\big[(y_t - \mu)(y_{t+i} - \mu)\big]\, \mathbb{E}\big[(y_{t+k} - \mu)(y_{t+k+j} - \mu)\big];$$

see e.g., Lobato et al. (2001). Fuller (1976, p. 256) presents for a different expression of $\boldsymbol{V}$.[1] It follows from the continuous mapping theorem that

$$T(\hat{\boldsymbol{\rho}}_m - \boldsymbol{\rho}_m)'\boldsymbol{V}^{-1}(\hat{\boldsymbol{\rho}}_m - \boldsymbol{\rho}_m) \xrightarrow{D} \chi^2(m). \tag{2.1}$$

This distribution result is fundamental for the tests presented in this section.

Under the null hypothesis, $\boldsymbol{V}$ can be simplified such that $v_{ij} = c_{i+1,j+1}/\gamma(0)^2$ with

$$c_{i+1,j+1} = \sum_{k=-\infty}^{\infty} \mathbb{E}\big[(y_t - \mu)(y_{t+i} - \mu)(y_{t+k} - \mu)(y_{t+k+j} - \mu)\big].$$

In particular, when y_t are serially independent,

$$c_{i+1,j+1} = \begin{cases} 0, & i \neq j, \\ \gamma(0)^2, & i = j. \end{cases}$$

In this case, $\boldsymbol{V}$ is an identity matrix, and the normalized sample autocorrelations $\sqrt{T}\hat{\rho}(i)$, $i = 1, \ldots, m$, are independent $\mathcal{N}(0,1)$ random variables asymptotically. Many computer programs now raw a confidence interval for sample autocorrelations so as to provide a quick, visual check of the significance of $\hat{\rho}(i)$. For example, the 90% (95%) confidence interval of $\hat{\rho}(i)$ is $\pm 1.645/\sqrt{T}$ ($\pm 1.96/\sqrt{T}$). A joint test is

$$\mathcal{Q}_T = T\hat{\boldsymbol{\rho}}_m'\hat{\boldsymbol{\rho}}_m = T\sum_{i=1}^{m} \hat{\rho}(i)^2 \xrightarrow{D} \chi^2(m), \tag{2.2}$$

under the null hypothesis; $\mathcal{Q}_T$ is the well known Q test of Box and Pierce (1970).

When y_t are independent random variables with mean zero, variance σ^2, and finite 6^{th} moment, we have from a result of Fuller (1976, p. 242) that

$$\operatorname{cov}\big(\sqrt{T}\hat{\rho}(i),\, \sqrt{T}\hat{\rho}(j)\big) = \begin{cases} \frac{T-i}{T} + O(T^{-1}), & i = j \neq 0, \\ O(T^{-1}), & i \neq j. \end{cases}$$

This result provides an approximation up to $O(T^{-1})$. Then for sufficiently large T, the diagonal elements of $\boldsymbol{V}$ are approximately $(T-i)/T$, whereas the off-diagonal elements essentially vanish. This leads to the modified Q test of Ljung and Box (1978):

$$\widetilde{\mathcal{Q}}_T = T^2 \sum_{i=1}^{m} \frac{\hat{\rho}(i)^2}{T-i} \xrightarrow{D} \chi^2(m), \tag{2.3}$$

[1] Fuller (1976, p. 256) shows that the $(i,j)^{\text{th}}$ element of $\boldsymbol{V}$ is

$$v_{ij} = \sum_{k=-\infty}^{\infty} \rho(k)\rho(k-i+j) + \rho(k+j)\rho(k-i) - 2\rho(k)\rho(j)\rho(k-i) - 2\rho(k)\rho(i)\rho(k-j) + 2\rho(i)\rho(j)\rho(k)^2.$$

cf. (2.2). Clearly, the Box-Pierce Q test and the Ljung-Box $\widetilde{Q}$ test are asymptotically equivalent, yet the latter ought to have better finite-sample performance because it employs a finite-sample correction of $\boldsymbol{V}$. In practice, the Ljung-Box Q statistic is usually computed as

$$T(T+2)\sum_{i=1}^{m}\frac{\hat{\rho}(i)^2}{(T-i)}, \tag{2.4}$$

which is of course asymptotically equivalent to (2.3).

Another modification of the Q test can be obtained by assuming that

$$\mathbb{E}\big[(y_t-\mu)(y_{t+i}-\mu)(y_{t+k}-\mu)(y_{t+k+j}-\mu)\big]=0, \tag{2.5}$$

for each k when $i \neq j$ and for $k \neq 0$ when $i = j$. Given this assumption, $c_{i+1,j+1} = 0$ when $i \neq j$, but

$$c_{i+1,j+1} = \mathbb{E}\big[(y_t-\mu)^2(y_{t+i}-\mu)^2\big].$$

when $i = j$. Under this assumption, $\boldsymbol{V}$ is diagonal with the diagonal element $v_{ii} = c_{i+1,i+1}/\gamma(0)^2$, which can be consistently estimated by

$$\hat{v}_{ii} = \frac{\frac{1}{T}\sum_{t=1}^{T-i}(y_t-\bar{y})^2(y_{t+i}-\bar{y})^2}{[\frac{1}{T}\sum_{t=1}^{T}(y_t-\bar{y})^2]^2}.$$

Then under the null hypothesis,

$$\mathcal{Q}_T^* = T\sum_{i=1}^{m}\hat{\rho}(i)^2/\hat{v}_{ii} \xrightarrow{D} \chi^2(m); \tag{2.6}$$

Lobato et al. (2001) refer to this test as the Q^* test. Note that the Q^* test does not require y_t to be serially independent, in contrast with the Box-Pierce and Ljung-Box Q tests. The Q^* test is therefore more suitable for testing processes that are serially uncorrelated but serially dependent, such as GARCH processes (see Example in Section 3.1.

Remark:

1. The asymptotic distribution of the Box-Pierce and Ljung-Box Q tests is derived under the assumption that $\{y_t\}$ is serially independent. This distribution result is also valid when $\{y_t\}$ is a martingale difference sequence (a precise definition will be given in Section 3.1) with additional moment conditions. These Q tests can also be interpreted as independence (or martingale-difference) tests with a focus on autocorrelations. Although the Q^* test does not require the independence assumption, the condition (2.5) is difficult to verify in practice.

2. When the Q-type tests are applied to the residuals of an ARMA(p,q) model, the asymptotic null distribution becomes $\chi^2(m-p-q)$.

3. The asymptotic null distribution of the Q-type tests is valid provided that data possess at least finite $(4+\delta)$ th moment for some $\delta > 0$. Many financial time series, unfortunately, may not satisfy this moment requirement; see e.g., de Lima (1997).

2.2 The Spectral Tests

Instead of testing a fixed number of autocorrelations $\rho(j)$, it is also possible to test if $\rho(j)$ are all zero:

$$H_0 \colon \rho(1) = \rho(2) = \cdots = 0.$$

Recall that the spectral density function is the Fourier transform of the autocorrelations:

$$f(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \rho(j) e^{-ij\omega}, \qquad \omega \in [-\pi, \pi],$$

where $i = (-1)^{1/2}$. When the autocorrelations are all zero, the spectral density reduces to the constant $(2\pi)^{-1}$ for all ω. It is then natural to base a test of all autocorrelations by comparing the sample counterpart of $f(\omega)$ and $(2\pi)^{-1}$.

Let $I_T(\omega)$ denote the *periodogram*, the sample spectral density, of the time series $\{y_t\}$. The difference between $I_T(\omega)$ and $(2\pi)^{-1}$ is

$$\frac{1}{2\pi} \left(\sum_{j=-(T-1)}^{T-1} \hat{\rho}(j) e^{-ij\omega} - 1 \right).$$

Recall that $\exp(-ij\omega) = \cos(j\omega) - i \sin(j\omega)$, where sin is an odd function such that $\sin(j\omega) = -\sin(-j\omega)$, and cos is an even function such that $\cos(j\omega) = \cos(-j\omega)$. Thus,

$$\frac{1}{2\pi} \left(\sum_{j=-(T-1)}^{T-1} \hat{\rho}(j) e^{-ij\omega} - 1 \right) = \frac{1}{\pi} \sum_{j=1}^{T-1} \hat{\rho}(j) \cos(j\omega).$$

Integrating this function with respect to ω on $[0, a]$, $0 \le a \le \pi$, we obtain the cumulated differences:

$$\frac{1}{\pi} \sum_{j=1}^{T-1} \hat{\rho}(j) \frac{\sin(ja)}{j}.$$

Consider now the normalized, cumulated differences:

$$D_T(t) = \frac{\sqrt{2T}}{\pi} \sum_{j=1}^{m(T)} \hat{\rho}(j) \frac{\sin(j\pi t)}{j},$$

where $\pi t = a$ and $m(T)$ is less than T but grows with T at a slower rate.

Recall that a sequence of i.i.d. $\mathcal{N}(0,1)$ random variables $\{\epsilon_t\}$ can approximate the Brownian motion B via

$$W_T(t) = \epsilon_0 t + \frac{\sqrt{2}}{\pi} \sum_{j=1}^{T} \epsilon_j \frac{\sin(j\pi t)}{j} \Rightarrow B(t), \quad t \in [0,1]$$

where $\Rightarrow$ stands for weak convergence. Then,

$$W_T(t) - tW_T(1) = \frac{\sqrt{2}}{\pi} \sum_{j=1}^{T} \epsilon_j \frac{\sin(j\pi t)}{j} \Rightarrow B^0(t), \quad t \in [0,1], \tag{2.7}$$

where B^0 denotes the Brownian bridge. It can also be shown that when y_t are conditionally homoeskedastic, $\sqrt{T}\hat{\rho}(j)$, $j = 1, \ldots, m$, converge in distribution to independent $\mathcal{N}(0,1)$ random variables under the null hypothesis. In view of the approximation (2.7) and the asymptotic normality of $T^{1/2}\hat{\rho}(j)$, we obtain

$$D_T(t) \Rightarrow B^0(t), \qquad t \in [0,1]. \tag{2.8}$$

The spectral tests proposed by Durlauf (1991) are constructed from taking various functionals on D_T.

Durlauf (1991) considered the following test statistics:
(1) the Anderson-Darling test:

$$\mathrm{AD}_T = \int_0^1 \frac{[D_T(t)]^2}{t(1-t)} \, \mathrm{d}t \Rightarrow \int_0^1 \frac{[B^0(t)]^2}{t(1-t)} \, \mathrm{d}t;$$

(2) the Cramér-von Mises test:

$$\mathrm{CVM}_T = \int_0^1 [D_T(t)]^2 \, \mathrm{d}t \Rightarrow \int_0^1 [B^0(t)]^2 \, \mathrm{d}t;$$

(3) the Kolmogorov-Smirnov test:

$$\mathrm{KS}_T = \sup |D_T(t)| \Rightarrow \sup |B^0(t)|;$$

(4) the Kuiper test:

$$\mathrm{Ku}_T = \sup_{s,t} |D_T(t) - D_T(s)| \Rightarrow \sup |B^0(t) - B^0(s)|.$$

The limits of these test statistics are direct consequences of (2.8) and the continuous mapping theorem. These limits are also the limits of the well-known goodness-of-fit tests in the statistics literature, and their critical values have been tabulated in, e.g., Shorack and Wellner (1986).

A condition ensuring the result (2.8) is that $\sqrt{T}\hat{\rho}(j)$ are independent $\mathcal{N}(0,1)$ asymptotically. Similar to the finding of Lobato et al. (2001), Deo (2000) noted that when y_t are conditionally heteroskedastic, the asymptotic variance of $\sqrt{T}\hat{\rho}(j)$ would be $\mathbb{E}(y_t^2 y_{t-j}^2)/\gamma(0)^2$. In this case, (2.8) fails to hold because D_T is not properly normalized; instead, D_T converges weakly to a different process. As a result, the limiting distributions of the tests considered by Durlauf (1991) would have thicker right-tails under conditional heteroskedasticity and render these tests over-sized; that is, these tests reject too often under the null hypothesis than they should. Deo (2000) proposed the following modification of D_T:

$$D_T^c(t) = \frac{\sqrt{2T}}{\pi} \sum_{j=1}^{m(T)} \frac{\hat{\rho}(j)}{\hat{v}_{jj}} \frac{\sin(j\pi t)}{j},$$

where

$$\hat{v}_{jj} = \frac{1}{\hat{\gamma}(0)} \left(\frac{1}{T-j} \sum_{t=1}^{T-j} (y_t - \bar{y})^2 (y_{t+j} - \bar{y})^2 \right)^{1/2},$$

and $\hat{\gamma}(0)$ is the sample variance of y_t. With additional regularity conditions, the Cramér-von Mises test based on D_T^c has the same limit as the original CVM_T test, i.e.,

$$\mathrm{CVM}_T^c = \int_0^1 [D_T^c(t)]^2 \,\mathrm{d}t \Rightarrow \int_0^1 [B^0(t)]^2 \,\mathrm{d}t.$$

This modification is analogous to the Q^* test of Lobato et al. (2001). The simulation results of Deo (2000) demonstrate that the modified spectral test is indeed robust to some series that are conditionally heteroskedastic.

3 Tests of the Martingale Difference Hypothesis

3.1 The Martingale Difference Hypothesis

Let $\{\mathcal{F}^t\}$ denote a sequence of information sets such that $\mathcal{F}^s \subseteq \mathcal{F}^t$ for all $s < t$. Such a sequence is known as a *filtration*. A sequence of integrable random variables $\{y_t\}$ is said to be a *martingale difference sequence* with respect to the filtration $\{\mathcal{F}^t\}$ if, and only if, $\mathbb{E}(y_t \mid \mathcal{F}^{t-1}) = 0$ for all t. When $\{y_t\}$ is a martingale difference sequence, its cumulated sums, $\eta_t = \sum_{s=1}^t y_s$, are such that $\mathbb{E}(\eta_t \mid \mathcal{F}^{t-1}) = \eta_{t-1}$ and form the process known as a *martingale*. By the law of iterated expectations, a martingale difference sequence must have mean zero. This implication is not restrictive, as we can always evaluate the "centered" series $\{y_t - \mathbb{E}(y_t)\}$ when y_t have non-zero means.

The concept of martingale difference can be related to time series *non-predictability.* We say that $\{y_t\}$ is not predictable (in the mean-squared-error sense) if, and only if, there is a filtration $\{\mathcal{F}^t\}$ such that the conditional expectations $\mathbb{E}(y_t \mid \mathcal{F}^{t-1})$ are the same as the unconditional expectations $\mathbb{E}(y_t)$. That is, the conditioning variables in $\mathcal{F}^{t-1}$ do not help improving the forecast of y_t so that the best L_2 forecast is not different from the naive forecast. Clearly, this definition is equivalent to requiring $\{y_t - \mathbb{E}(y_t)\}$ being a martingale difference sequence with respect to $\{\mathcal{F}^t\}$. In the time series analysis, the information sets $\mathcal{F}^t$ are understood as the σ-algebras generated by $\boldsymbol{Y}^t = \{y_t, y_{t-1}, \dots, y_1\}$. Many practitioners believe that they can "beat the market" if they are able to predict future returns. Examining the martingale difference property thus allows us to evaluate whether a time series is indeed predictable.

It is easily verified that $\{y_t\}$ is a martingale difference sequence if, and only if, y_t are uncorrelated with $h(\boldsymbol{Y}^{t-1})$ for any measurable function h; i.e.,

$$\mathbb{E}[y_t\, h(\boldsymbol{Y}^{t-1})] = 0. \tag{3.1}$$

Taking h in (3.1) as the linear function, we immediately see that y_t must be serially uncorrelated with $y_{t-1}, \dots, y_1$. Thus, a martingale difference sequence must be serially uncorrelated; the converse need not be true. For example, consider the following nonlinear moving average process:

$$y_t = \varepsilon_{t-1}\varepsilon_{t-2}(\varepsilon_{t-2} + \varepsilon_t + 1),$$

where ε_t are i.i.d. $N(0,1)$ random variables. It is clear that $\mathrm{corr}(y_t, y_{t-j}) = 0$ for all j, yet $\{y_t\}$ is not a martingale difference process. Recall that a white noise is a sequence of uncorrelated random variables that have zero mean and a constant variance. Thus, a martingale difference sequence need not be a white noise because the former does not have any restriction on variance or other high-order moments.

If $\{y_t\}$ is a sequence of serially independent random variables with zero mean, we have

$$\mathbb{E}[y_t\, h(\boldsymbol{Y}^{t-1})] = \mathbb{E}(y_t)\, \mathbb{E}[h(\boldsymbol{Y}^{t-1})] = 0,$$

so that $\{y_t\}$ is necessarily a martingale difference sequence. The converse need not be true. To see this, consider a simple ARCH (autoregressive conditional heteroskedasticity) process $\{y_t\}$ such that $y_t = v_t^{1/2}\, \varepsilon_t$, where ε_t are i.i.d. random variables with mean zero and variance 1, and

$$v_t = a + b y_{t-1}^2,$$

with $a, b > 0$. It can be seen that $\{y_t\}$ is a martingale difference sequence, yet it is serially dependent due to the correlaions among y_t^2. We therefore conclude that serial independence implies the martingale difference property which in turn implies serial uncorrelatedness.

4 Tests of Serial Independence

We have seen that serial independence is a much more stringent requirement than the martingale difference property. The ARCH example in Section 3.1, for instance, is a martingale difference sequence but serially dependent. Given that the ARCH and GARCH (generalized ARCH) models are popular in financial applications, it is also of interest to consider testing a special form of serial dependence, viz., the correlations among squared returns. McLeod and Li (1983) suggested testing whether the first m autocorrelations of y_t^2 are zero using a Q test. That is, one computes (2.3) or (2.4) with $\hat{\rho}(i)$ the sample autocorrelations of y_t^2: where

$$\hat{\rho}(i) = \frac{\frac{1}{T}\sum_{t=1}^{T-i}(y_t^2 - m_2)(y_{t+i}^2 - m_2)}{\frac{1}{T}\sum_{t=1}^{T}(y_t^2 - m_2)^2},$$

with m_2 the sample mean of y_t^2. The asymptotic null distributions of the resulting Q test remains $\chi^2(m)$. While the McLeod-Li test also checks a necessary condition of serial independence, its validity requires an even stronger moment condition (finite 8 th moment). The tests discussed below, on the other hand, focus on the i.i.d. condition which is sufficient for serial independence. Failing to reject the null hypothesis is consistent with serial independence, yet rejecting the null hypothesis does not imply serial dependence.

4.1 The Variance Ratio Test

The variance-ratio test of Cochrane (1988) is a convenient diagnostic test of the i.i.d. assumption. Suppose that y_t are i.i.d. random variables with mean zero and variance σ^2. Then for any k, $\mathrm{var}(y_t + \cdots + y_{t-k+1})$ is simply $k\sigma^2$. Let $\tilde{\sigma}_k^2$ denote an estimator of $\mathrm{var}(y_t + \cdots + y_{t-k+1})$ and $\hat{\sigma}^2$ the sample variance of σ^2. Under the null hypothesis, $\tilde{\sigma}_k^2/k$ and $\hat{\sigma}^2$ should be close to each other. The variance ratio test is simply a normalized version of $\tilde{\sigma}_k^2/(k\hat{\sigma}^2)$.

Let η_t be the partial sum of y_i such that $y_t = \eta_t - \eta_{t-1}$. Suppose there are $kT + 1$ observations $\eta_0, \eta_1, \ldots, \eta_{kT}$. Define the sample average of y_t as

$$\bar{y} = \frac{1}{kT}\sum_{t=1}^{kT}(\eta_t - \eta_{t-1}) = \frac{1}{kT}(\eta_{kT} - \eta_0).$$

The variance estimator of $\sigma^2 = \mathrm{var}(\eta_t - \eta_{t-1})$ is

$$\hat{\sigma}^2 = \frac{1}{kT}\sum_{t=1}^{kT}\big(\eta_t - \eta_{t-1} - \bar{y}\big)^2.$$

Under the i.i.d. null hypothesis, $\sqrt{kT}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{D} \mathcal{N}(0,\, 2\sigma^4)$. Consider now the following estimator of $\sigma_k^2 = \mathrm{var}(\eta_t - \eta_{t-k})$:

$$\tilde{\sigma}_k^2 = \frac{1}{T}\sum_{t=1}^{T}\big(\eta_{kt} - \eta_{kt-k} - k\bar{y}\big)^2 = \frac{1}{T}\sum_{t=1}^{T}\big[k(\bar{y}_t - \bar{y})\big]^2,$$

where $\bar{y}_t = \sum_{kt-k+1}^{kt} y_i/k$. Under the i.i.d. hypothesis, $\sigma_k^2 = k\sigma^2$, so that

$$\sqrt{T}(\tilde{\sigma}_k^2 - k\sigma^2) \xrightarrow{D} \mathcal{N}(0,\, 2k^2\sigma^4),$$

or equivalently,

$$\begin{aligned}
\frac{1}{\sqrt{k}}\sqrt{T}(\tilde{\sigma}_k^2 - k\sigma^2) &= \sqrt{kT}\left(\frac{\tilde{\sigma}_k^2}{k} - \sigma^2\right) \\
&= \sqrt{kT}\left(\frac{\tilde{\sigma}_k^2}{k} - \hat{\sigma}^2\right) + \sqrt{kT}\big(\hat{\sigma}^2 - \sigma^2\big) \\
&\xrightarrow{D} \mathcal{N}(0,\, 2k\sigma^4).
\end{aligned}$$

While $\hat{\sigma}^2$ is both consistent and asymptotically efficient for σ^2 under the null hypothesis, $\tilde{\sigma}_k^2/k$ is consistent but not asymptotically efficient. Following Hausman (1978),[2] we conclude that $\hat{\sigma}^2$ must be asymptotically uncorrelated with $\tilde{\sigma}_k^2/k - \hat{\sigma}^2$ and hence

$$\sqrt{kT}\left(\frac{\tilde{\sigma}_k^2}{k} - \hat{\sigma}^2\right) \xrightarrow{D} \mathcal{N}\big(0,\, 2(k-1)\sigma^4\big).$$

It follows that

$$\sqrt{kT}\left(\frac{\tilde{\sigma}_k^2}{k\hat{\sigma}^2} - 1\right) \xrightarrow{D} \mathcal{N}\big(0,\, 2(k-1)\big).$$

Denote the ratio $\tilde{\sigma}_k^2/(k\hat{\sigma}^2)$ as VR. The statistic

$$\sqrt{kT}(\mathrm{VR} - 1)/\sqrt{2(k-1)} \xrightarrow{D} \mathcal{N}(0,\, 1),$$

under the null hypothesis.

[2]Given the parameter of interest θ, let $\hat{\theta}_e$ be a consistent and asymptotically efficient estimator of θ and $\hat{\theta}_c$ is a consistent estimator but not asymptotically efficient. Hausman (1978) showed that $\hat{\theta}_e$ is asymptotically uncorrelated with $\hat{\theta}_c - \hat{\theta}_e$. For if not, there would exist a linear combination of $\hat{\theta}_e$ and $\hat{\theta}_c - \hat{\theta}_e$ that is asymptotically more efficient than $\hat{\theta}_e$.

In practice, one may employ other estimators for the variance ratio test. For example, σ_k^2 may be estimated by

$$\tilde{\sigma}_k^2 = \frac{1}{kT} \sum_{t=k}^{kT} (\eta_t - \eta_{t-k} - k\bar{y})^2.$$

One may also correct the bias of variance estimators and compute

$$\hat{\sigma}^2 = \frac{1}{kT-1} \sum_{t=1}^{kT} (\eta_t - \eta_{t-1} - \bar{y})^2,$$

$$\tilde{\sigma}_k^2 = \frac{1}{M} \sum_{t=k}^{kT} (\eta_t - \eta_{t-k} - k\bar{y})^2,$$

where $M = (kT - k + 1)(1 - 1/T)$. For more detailed discussion, we refer to Campbell, Lo, and MacKinlay (1997).

4.2 The BDS Test

The BDS test of serial independence also checks whether a sequence of random variables are i.i.d. Let $\boldsymbol{Y}_t^n = (y_t, y_{t+1}, \ldots, y_{t+n-1})$. Define the *correlation integral* with the *dimension* n and *distance* ϵ as:

$$C(n, \epsilon) = \lim_{T \to \infty} \binom{T-n}{2}^{-1} \sum\nolimits_{\forall s<t} \mathbf{I}_\epsilon(\boldsymbol{Y}_t^n, \boldsymbol{Y}_s^n),$$

where $\mathbf{I}_\epsilon(Y_t^n, Y_s^n) = 1$ if the maximal norm $\|\boldsymbol{Y}_t^n - \boldsymbol{Y}_s^n\| < \epsilon$ and 0 otherwise. The correlation integral is a measure of the proportion that any pairs of n-vectors ($\boldsymbol{Y}_t^n$ and $\boldsymbol{Y}_s^n$) are within a certain distance ϵ. If y_t are indeed i.i.d., $\boldsymbol{Y}_t^n$ should exhibit no pattern in the n-dimensional space, so that $C(n, \epsilon) = C(1, \epsilon)^n$. The BDS test is then designed to check whether the sample counterparts of $C(n, \epsilon)$ and $C(1, \epsilon)^n$ are sufficiently close. Specifically, the BDS statistic reads

$$B_T(n, \epsilon) = \sqrt{T - n + 1}(C_T(n, \epsilon) - C_T(1, \epsilon)^n)/\hat{\sigma}(n, \epsilon),$$

where

$$C_T(n, \epsilon) = \binom{T-n}{2}^{-1} \sum\nolimits_{\forall s<t} \mathbf{I}_\epsilon(\boldsymbol{Y}_t^n, \boldsymbol{Y}_s^n),$$

and $\hat{\sigma}^2(n, \epsilon)$ is a consistent estimator of the asymptotic variance of $\sqrt{T - n + 1} C_T(n, \epsilon)$; see Brock et al. (1996) for details. The asymptotic null distribution of the BDS test is $\mathcal{N}(0, 1)$.

The performance of the BDS test depends on the choice of n and ϵ. There is, however, no criterion to determine these two parameters. In practice, one may consider several values of n and set ϵ as a proportion to the sample standard deviation s_T of the data, i.e., $\epsilon = \delta s_T$ for some δ. Common choices of δ are 0.75, 1, and 1.5. An advantage of the BDS test is that it is robust to random variables that do not possess high-order moments. The BDS test usually needs a large sample to ensure proper performance. Moreover, it has been found that the BDS test has low power against various forms of nonlinearity; see, e.g., Hsieh (1989, 1991, 1993), Rothman (1992), Brooks and Heravi (1999), and Brooks and Henry (2000). In particular, the BDS test is not sensitive to certain class of self-exciting threshold AR processes (Rothman, 1992) and neglected asymmetry in volatility (Hsieh, 1991; Brooks and Henry, 2000; Chen and Kuan, 2002).

5 Tests of Time Reversibility

A different type of diagnostic test focuses on the property of *time reversibility.* A strictly stationary process $\{y_t\}$ is said to be time reversible if its finite-dimensional distributions are all invariant to the reversal of time indices. That is,

$$F_{t_1,t_2,\ldots,t_n}(c_1, c_2, \ldots, c_n) = F_{t_n,t_{n-1},\ldots,t_1}(c_n, c_{n-1}, \ldots, c_n).$$

When this condition does not hold, $\{y_t\}$ is said to be *time irreversible.* Clearly, independent sequences and stationary Gaussian ARMA processes are time reversible. Rejecting the null hypothesis of time reversibility thus implies that the data can not be serially independent. As such, the test of time reversibility can also be interpreted as a test of serial independence.

Time irreversibility indicates some time series characteristics that can not be described by the autocorrelation function. When $\{y_t\}$ is time reversible, it can be shown that for any k, the marginal distribution of $y_t - y_{t-k}$ must be symmetric about the the origin; see e.g., Cox (1981) and Chen, Chou, and Kuan (2000). That is, $y_t - y_{t-k}$ and $y_{t-k} - y_t$ should have the same distributions for each k. If this symmetry condition fails for some k, there is some asymmetric dependence between y_t and y_{t-k}, in the sense that the effect of y_{t-k} on y_t is different from that of y_t on y_{t-k}. In view of this property, we infer that nonlinear time series are time irreversible in general. Indeed, Tong (1990) states that "time irreversibility is the rule rather than the exception when it comes to nonlinearity" (p. 197). Moreover, linear and stationary processes with non-Gaussian innovations are typically time irreversible. Thus, testing time reversibility can complement existing tests of independence.

Existing tests of time reversibility aim at symmetry of $y_t - y_{t-k}$. A necessary condition of distribution symmetry is its third central moment being zero. One may then test time reversibility by evaluating whether the sample third moment is sufficiently close to zero. Observe that by stationarity,

$$\begin{aligned}\mathbb{E}(y_t - y_{t-k})^3 &= \mathbb{E}(y_t^3) - 3\,\mathbb{E}(y_t^2 y_{t-k}) + 3\,\mathbb{E}(y_t y_{t-k}^2) - \mathbb{E}(y_{t-k}^3)\\ &= -3\,\mathbb{E}(y_t^2 y_{t-k}) + 3\,\mathbb{E}(y_t y_{t-k}^2),\end{aligned}$$

where the two terms on the right-hand side are referred to as the *bi-covariances*. Ramsey and Rothman (1996) base their test of time reversibility on the sample bi-covariances. Note that both the third-moment test and bi-covariance test require the data to possess at least finite 6^{th} moment. Unfortunately, most financial time series do not satisfy this moment condition. On the other hand, Chen, Chou and Kuan (2000) consider a different testing approach that is robust to the failure of moment conditions.

It is well known that a distribution is symmetric if, and only if, the imaginary part of its characteristic function is zero. Hence, time reversibility of $\{y_t\}$ implies that

$$h_k(\omega) := \mathbb{E}\big[\sin\big(\omega(y_t - y_{t-k})\big)\big] = 0, \qquad \text{for all } \omega \in \mathbb{R}^+. \tag{5.1}$$

Note that (5.1) include infinitely moment conditions indexed by ω. Let g be a positive function such that $\int g(\omega)\,\mathrm{d}\omega < \infty$. By changing the orders of integration, (5.1) implies that

$$\int_{\mathbb{R}^+} h_k(\omega) g(\omega)\,\mathrm{d}\omega = \int_{\mathbb{R}} \left(\int_{\mathbb{R}^+} \sin\big(\omega(y_t - y_{t-k})\big) g(\omega)\,\mathrm{d}\omega \right) \mathrm{d}F = 0,$$

where F is the distribution function of y_t. This condition is equivalent to

$$\mathbb{E}[\psi_g(y_t - y_{t-k})] = 0, \tag{5.2}$$

where

$$\psi_g(y_t - y_{t-k}) = \int_{\mathbb{R}^+} \sin\big(\omega(y_t - y_{t-k})\big) g(\omega)\,\mathrm{d}\omega.$$

To test (5.2), Chen, Chou, and Kuan (2000) suggest using the sample average of $\psi_g(y_t - y_{t-k})$:

$$\mathcal{C}_{g,k} = \sqrt{T_k}\,\bar{\psi}_{g,k}/\bar{\sigma}_{g,k}, \tag{5.3}$$

where $T_k = T - k$ with T the sample size, $\bar{\psi}_{g,k} = \sum_{t=k+1}^{T} \psi_g(y_t - y_{t-k})/T_k$, and $\bar{\sigma}^2_{g,k}$ is a consistent estimator of the asymptotic variance of $\sqrt{T_k}\,\bar{\psi}_{g,k}$. A suitable central limit theorem then ensures that $\mathcal{C}_{g,k}$ is asymptotically distributed as $\mathcal{N}(0,1)$ under the null hypothesis (5.2).

A novel feature of this test is that, because $\psi_g(\cdot)$ is bounded between 1 and -1, *no* moment condition is needed when the central limit theorem is invoked. Yet a major drawback of $\mathcal{C}_{g,k}$ is that the null hypothesis (5.2) is only a necessary condition of (5.1). Thus, h_k may be integrated to zero by some g function even when h_k is not identically zero. For such a g function, the resulting $\mathcal{C}_{g,k}$ test does not have power against asymmetry of $y_t - y_{t-k}$. Choosing a proper g function is therefore crucial for implementing this test. Chen, Chou, and Kuan (2000) observed that for absolutely continuous distributions, $h_k(\omega)$ is a damped sine wave and eventually decays to zero as $\omega \to \infty$; see, e.g., Figure **??** for h_k of various "centered" exponential distributions. This suggests choosing g as the density function of a random variable on $\mathbb{R}^+$. In particular, when g is the density of the exponential distribution ($g = \exp$) with the parameter $\beta > 0$, i.e., $g(\omega) = \exp(-\omega/\beta)/\beta$, it can be shown that ψ_g has an analytic expression:

$$\psi_{\exp}(y_t - y_{t-k}) = \frac{\beta(y_t - y_{t-k})}{1 + \beta^2(y_t - y_{t-k})^2}. \tag{5.4}$$

The closed form (5.4) renders the computation of $\mathcal{C}_{\exp,k}$ test quite easy. One simply plugs the data into (5.4) and calculates their sample average and sample standard deviation. The test statistic is now readily computed as (5.3). Chen, Chou, and Kuan (2000) demonstrate that the $\mathcal{C}_{\exp,k}$ test performs strikingly well in finite samples and is very robust to the data without higher-order moments. The third-moment-based test and the bi-covariance test, on the other hand, have little power when the data do not have proper moments.

Apart from computation simplicity, the $\mathcal{C}_{\exp,k}$ test is very flexible. By varying the value of β, $\mathcal{C}_{\exp,k}$ is able to check departures from (5.2) in different ways. When β is small, this test concentrates on $h_k(\omega)$ for smaller ω. By contrast, more $h_k(\omega)$ can be taken into account as β increases. How to choose an optimal β that maximizes the test power remains an unsolved problem, however.

Remarks:

1. One may consider testing a condition equivalent to (5.1). For example, a Cramér-von Mises type condition is based on

   $$\int_{\mathbb{R}^+} h_k(\omega)^2 g(\omega)\, \mathrm{d}\omega,$$

   which is zero if, and only if, (5.1) holds. This condition, however, does not permit changing the orders of integration. The resulting test is more difficult to implement and usually has a data-dependent distribution.

2. To apply the test (5.3) to model residuals, Chen and Kuan (2002) noted that one must compute the sample standard deviation with care. Chen and Kuan (2002) also demonstrated that the $\mathcal{C}_{\exp,k}$ test is particularly powerful against asymmetric dependence in data, but other diagnostic tests are not. Specifically, they simulated data from an EGARCH process which exhibits volatility asymmetry but estimated a GARCH model which is a model of symmetric volatility pattern. It is shown that the $\mathcal{C}_{\exp,k}$ test on the standardized GARCH residuals can reject the null hypothesis of time reversibility with high probability, yet the Q-type tests and the BDS test have low power and fail to reject the null hypothesis of serial uncorrelatedness (independence). As far as model diagnostic is concerned, the latter tests are unable to distinguish between GARCH and EGARCH models.

References

Bierens, Herman J. (1982). Consistent model specification tests, *Journal of Econometrics*, **20**, 105–134.

Bierens, Herman J. (1984). Model specification testing of time series regressions, *Journal of Econometrics*, **26**, 323–353.

Bierens, Herman J. and Werner Ploberger (1997). Asymptotic theory of integrated conditional moment tests, *Econometrica*, **65**, 1129–1151.

Box, G. E. P. and David A. Pierce (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models, *Journal of the American Statistical Association*, **65**, 1509–1526.

Brock, W. A., W. D. Dechert and J. A. Scheinkman (1987). A test for independence based on the correlation dimension, Working paper, Department of Economics, University of Wisconsin at Madison.

Brock, W. A., W. D. Dechert, J. A. Scheinkman, and B. LeBaron (1996). A test for independence based on the correlation dimension, *Econometric Reviews*, **15**, 197–235.

Brooks, Chris and Saeed M. Heravi (1999). The effect of (mis-specified) GARCH filters on the finite sample distribution of the BDS test, *Computational Economics*, **13**, 147–162.

Brooks, Chris and Óland T. Henry (2000). Can portmanteau nonlinearity tests serve as general mis-specification tests? Evidence from symmetric and asymmetric GARCH models, *Economics Letters*, **67**, 245–251.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay (1997). *The Econometrics of Financial Markets*, Princeton, NJ: Princeton University Press.

Chen, Yi-Ting, Ray Chou, and Chung-Ming Kuan (2000). Testing time reversibility without moment restrictions, *Journal of Econometrics*, **95**, 199–218.

Chen, Yi-Ting and Chung-Ming Kuan (2002). Time irreversibility and EGARCH effects in U.S. stock index returns, *Journal of Applied Econometrics*, **17**, 565–578.

Cochrane, John H. (1988). How big is the random walk in GNP? *Journal of Political Economy*, **96**, 893–920.

Cox, D. R. (1981). Statistical analysis of time series: Some recent developments, *Scandinavian Journal of Statistics*, **8**, 93–115.

de Lima, Pedro J. F. (1997). On the robustness of nonlinearity tests to moment condition failure, *Journal of Econometrics*, **76**, 251–280.

Deo, Rohit S. (2000). Spectral tests of the martingale hypothesis under conditional heteroscedasticity, *Journal of Econometrics*, **99**, 291–315.

Dominguez, Manuel A. and Ignacio N. Lobato (2000). A consistent test for the martingale difference hypothesis, Working Paper, Instituto Tecnologico Autonomo de Mexico.

Durlauf, Steven N. (1991). Spectral based testing of the martingale hypothesis, *Journal of Econometrics*, **50**, 355–376.

Fuller, Wayne A. (1976). *Introduction to Statistical Time Series*, New York, NY: Wiley.

Hausman, Jerry A. (1978). Specification tests in econometrics, *Econometrica*, **46**, 1251–1271.

Hong, Yongmiao (1999). Hypothesis testing in time series via the empirical characteristic function: A generalized spectral density approach, *Journal of the American Statistical Association*, **94**, 1201–1220.

Hsieh, David A. (1989). Testing for nonlinear dependence in daily foreign exchange rates, *Journal of Business*, **62**, 339–368.

Hsieh, David A. (1991). Chaos and nonlinear dynamics: Application to financial markets, *Journal of Finance*, **46**, 1839–1877.

Hsieh, David A. (1993). Implications of nonlinear dynamics for financial risk management, *Journal of Financial and Quantitative Analysis*, **28**, 41–64.

Ljung, G. M. and G. E. P. Box (1978). On a measure of lack of fit in time series models, *Biometrika*, **65**, 297–303.

Lobato, Ignacio, John C. Nankervis, and N. E. Savin (2001). Testing for autocorrelation using a modified Box-Pierce Q test, *International Economic Review*, **42**, 187–205.

McLeod, A. I. and W. K. Li (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations, *Journal of Time Series Analysis*, **4**, 269–273.

Pinkse, Joris (1998). A consistent nonparametric test for serial independence, *Journal of Econometrics*, **84**, 205–231.

Ramsey, James B. and Philip Rothman (1996). Time irreversibility and business cycle asymmetry, *Journal of Money, Credit, and Banking*, **28**, 3–20.

Rothman, Philip (1992). The comparative power of the tr test against simple threshold models, *Journal of Applied Econometrics*, **7**, S187–S195.

Shorack, Galen R. and John A. Wellner (1986). *Empirical Processes with Applications to Statistics*, New York, NY: Wiley.

Skaug, Hans J. and Dag Tjostheim (1993). A nonparametric test of serial independence based on the empirical distribution function, *Biometrika*, **80**, 591–602.

Tong, Howell (1990). *Non-linear Time Series, A Dynamic System Approach*, New York, NY: Oxford University Press.

Whang, Yoon-Jae (2000). Consistent bootstrap tests of parametric regression functions, *Journal of Econometrics*, **98**, 27–46.

Whang, Yoon-Jae (2001). Consistent specification testing for conditional moment restrictions, *Economics Letters*, **71**, 299–306.