Introduction to
Time Series Diagnostic Tests

CHUNG-MING KUAN
Department of Finance & CRETA
National Taiwan University

March 20, 2011
Outline

1. Introduction

2. **Q-Type Tests**
   - Conventional *Q* Tests
   - Modified *Q* Test

3. **Spectral Tests**
   - Durlauf’s Spectral Test
   - Modified Spectral Test

4. **Other Tests of Serial Independence**
   - Variance Ratio Test
   - BDS Test

5. **Test for Time Reversibility**
Time series properties of \( y_t \)

- **Serial uncorrelatedness**: \( y_t \) uncorrelated with \( y_{t-i}, \ i = 1, 2, \ldots \).
- **Martingale difference**: \( y_t \) uncorrelated with any function on \( y_{t-i}, \ i = 1, 2, \ldots \).
- **Serial independence**: No relation between any function of \( y_t \) and any function of \( y_{t-i}, \ i = 1, 2, \ldots \).
- **Time reversibility**: Distributions are invariant wrt the reversal of time indices.

**Diagnostic testing**

- Testing results provide information on how these raw data may be modeled.
- When a model is estimated, diagnostic tests on model residuals yield information about model adequacy.
A weakly stationary time series \( \{y_t\} \) is such that its autocovariances (autocorrelations) depend on \( i \) but not on \( t \).

- **Autocovariances:** \( \gamma(i) = \text{cov}(y_t, y_{t-i}), i = 0, 1, 2, \ldots \)
- **Autocorrelations:** \( \rho(i) = \gamma(i)/\gamma(0), i = 0, 1, 2, \ldots \)

**Estimates:** Sample autocovariances are

\[
\hat{\gamma}(i) = \frac{1}{T} \sum_{t=1}^{T-i} (y_t - \bar{y})(y_{t+i} - \bar{y});
\]

sample autocorrelations are \( \hat{\rho}(i) = \hat{\gamma}(i)/\hat{\gamma}(0) \).

The null hypothesis of serial uncorrelatedness is:

\[
H_0: \rho(1) = \rho(2) = \cdots = 0.
\]

We test this hypothesis by checking sample autocorrelations.
Asymptotic Properties

Let $\rho_m = (\rho(1), \ldots, \rho(m))'$ and $\hat{\rho}_m = (\hat{\rho}(1), \ldots, \hat{\rho}(m))'$. We have

$$\sqrt{T}(\hat{\rho}_m - \rho_m) \xrightarrow{D} \mathcal{N}(0, V),$$

where the $(i,j)$-th elements of $V$ are

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

with

$$c_{i+1,j+1} = \sum_{k=-\infty}^{\infty} \mathbb{E} \left[ (y_t - \mu)(y_{t+i} - \mu)(y_{t+k} - \mu)(y_{t+k+j} - \mu) \right] - \mathbb{E} \left[ (y_t - \mu)(y_{t+i} - \mu) \right] \mathbb{E} \left[ (y_{t+k} - \mu)(y_{t+k+j} - \mu) \right].$$
• Under the null, $\rho_m = 0$, so that $\sqrt{T} V^{-1/2} \hat{\rho}_m \overset{D}{\rightarrow} \mathcal{N}(0, I_m)$ and

$$T \hat{\rho}'_m V^{-1} \hat{\rho}_m \overset{D}{\rightarrow} \chi^2(m).$$

This result holds when $V$ is replaced by a consistent estimator $\hat{V}$.

• When $\rho(i) = 0$ for all $i$, $v_{ij} = c_{i+1,j+1}/\gamma(0)^2$ with

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

This can be further simplified when more conditions are imposed.
Conventional $Q$ Tests

When $y_t$ are serially independent,

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

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

In this case, $\mathbf{V}$ simplifies to $\mathbf{I}_m$.

Box and Pierce (1970)

Under serial independence,

$$Q_T = T\hat{\rho}_m' \hat{\rho}_m \xrightarrow{D} \chi^2(m).$$
Fuller (1976, p. 242):

$$\text{cov}(\sqrt{T} \hat{\rho}(i), \sqrt{T} \hat{\rho}(j)) = \left\{ \begin{array}{ll} \frac{T-i}{T} + O(T^{-1}), & i = j \neq 0, \\ O(T^{-1}), & i \neq j. \end{array} \right.$$ 

That is, \((T - i)/T\) is a better estimate of the diagonal elements \(v_{ii}\) in finite samples. This suggests that the finite-sample power may be improved if \(\hat{\rho}(i)^2\) are normalized by \((T - i)/T\).

Ljung and Box (1978)

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

This test is also computed as: 
$$T(T + 2) \sum_{i=1}^m \frac{\hat{\rho}(i)^2}{T - i}.$$
Without the serial independence condition, we assume

\[ \mathbb{E} \left[ (y_t - \mu)(y_{t+i} - \mu)(y_{t+k} - \mu)(y_{t+k+j} - \mu) \right] = 0, \]

for each \( k \) when \( i \neq j \) and for \( k \neq 0 \) when \( i = j \).

We have: \( c_{i+1,j+1} = 0 \) when \( i \neq j \), and

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

when \( i = j \). Hence, \( \mathbf{V} \) is diagonal with \( v_{ii} = c_{i+1,i+1}/\gamma(0)^2 \).

Estimate of \( v_{ii} \):

\[ \hat{v}_{ii} = \frac{1}{T} \sum_{t=1}^{T-i} (y_t - \bar{y})(y_{t+i} - \bar{y})^2 \left[ \frac{1}{T} \sum_{t=1}^{T} (y_t - \bar{y})^2 \right]^{-1}. \]
Lobato et al. (2001)

\[ Q^*_T = T \sum_{i=1}^{m} \frac{\hat{\rho}(i)^2}{\hat{\nu}_{ii}} \xrightarrow{D} \chi^2(m). \]

- Under conditional homoskedasticity,
  \[ c_{i+1,i+1} = \mathbb{E}[(y_t - \mu)^2(y_{t+i} - \mu)^2] = \gamma(0)^2. \] Estimating \( c_{i+1,i+1} \) thus makes the \( Q^* \) test more robust to conditional heteroskedasticity, such as ARCH and GARCH processes.

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

- Instead of testing only \( m \) autocorrelations, it would be nice if one can test all correlation coefficients. To this end, note that the spectral density function is the Fourier transform of the autocorrelations:

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

Under the null, \( f(\omega) = (2\pi)^{-1}. \)

- The spectral test compares the sample counterpart of \( f \) (also known as the periodogram) with \((2\pi)^{-1}\), i.e.,

\[
\frac{1}{2\pi} \left( \sum_{j=-(T-1)}^{T-1} \hat{\rho}(j) e^{-ij\omega} - 1 \right).
\]
Recall that \( \exp(-ij\omega) = \cos(j\omega) - i \sin(j\omega) \), where \( \sin \) is an odd function and \( \cos \) an even function. 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 wrt \( \omega \) on \([0, a]\), \(0 \leq a \leq \pi\), we obtain

\[
\frac{1}{\pi} \sum_{j=1}^{T-1} \hat{\rho}(j) \frac{\sin(ja)}{j},
\]

which are the cumulated differences and also a process in \( a \).
• Durlauf's test is based on normalized, cumulated differences:

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

(1)

where \( \pi t = a \) and \( m(T) \) grows with \( T \) at a slower rate.

• The spectral representation of the standard Brownian motion \( B \) is:

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

where \( \epsilon_t \) are i.i.d. \( \mathcal{N}(0, 1) \) r.v. Then, \( W_T(1) = \epsilon \) and

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

where \( B^0 \) denotes the Brownian bridge.
Using $\sqrt{T} \hat{\rho}(i) \approx \mathcal{N}(0,1)$, we have $D_T(t) \Rightarrow B^0(t), \ t \in [0, 1]$. A test can be constructed by applying a functional to measure the fluctuation of $D_T$.

**Durlauf (1991)**

1. **Anderson-Darling test:**

   $$AD_T = \int_0^1 \frac{[D_T(t)]^2}{t(1-t)} \, dt \Rightarrow \int_0^1 \frac{[B^0(t)]^2}{t(1-t)} \, dt.$$

2. **Cramér-von Mises test:**

   $$\text{CvM}_T = \int_0^1 [D_T(t)]^2 \, dt \Rightarrow \int_0^1 [B^0(t)]^2 \, dt.$$

3. **Kolmogorov-Smirnov test:**

   $$\text{KS}_T = \sup |D_T(t)| \Rightarrow \sup |B^0(t)|.$$
as in Lobato et al. (2001), Deo (2000) also finds the asymptotic variance of \( \sqrt{T} \hat{\rho}(j) \) is \( \mathbb{I} \mathbb{E}(y_t^2 y_{t-j}^2)/\gamma(0)^2 \) under conditional heteroskedasticity. Normalized using \( \hat{\nu}_{jj} \) we have

\[
D^c_T(t) = \frac{\sqrt{2T}}{\pi} \sum_{j=1}^{m(T)} \frac{\hat{\rho}(j)}{\sqrt{\hat{\nu}_{jj}}} \frac{\sin(j\pi t)}{j}.
\]

The test based on \( D^c_T \) ought to be more robust to conditional heteroskedasticity than Durlauf’s test.

**Deo (2000)**

\[
CvM^c_T = \int_0^1 [D^c_T(t)]^2 \, dt \Rightarrow \int_0^1 [B^0(t)]^2 \, dt.
\]
Variance Ratio Test

- Null hypothesis: i.i.d.
- \( y_t \) are i.i.d. with mean zero and variance \( \sigma^2 \). For any \( k \),
  \[
  \text{var}(y_t + \cdots + y_{t-k+1}) = k\sigma^2.
  \]
- Let \( \tilde{\sigma}_k^2 \) denote an estimator of \( \text{var}(y_t + \cdots + y_{t-k+1}) \) and \( \hat{\sigma}^2 \) the sample variance of \( y_t \). Then, \( \tilde{\sigma}_k^2 / k \) and \( \hat{\sigma}^2 \) should be close to each other under the null.
- The variance ratio test of Cochrane (1988) is simply a normalized version of \( \tilde{\sigma}_k^2 / (k\hat{\sigma}^2) \).
- Notation: \( \eta_t \) is the partial sum of \( y_i \) such that \( y_t = \eta_t - \eta_{t-1} \).
  Suppose there are \( kT + 1 \) observations \( \eta_0, \eta_1, \ldots, \eta_{kT} \).
The standard estimator of $\sigma^2$: 

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

which is both consistent and asymptotically efficient under the null.

An estimator of $\sigma_k^2 = \text{var}(\eta_t - \eta_{t-k})$ is

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

where $\bar{y}_t = \sum_{kt-k+1}^{kt} y_i / k$. Clearly, $\tilde{\sigma}_k^2 / k$ is consistent for $\sigma^2$ but not asymptotically efficient under the null.

Under the null, $\sqrt{kT}(\hat{\sigma}^2 - \sigma^2) \overset{D}{\to} \mathcal{N}(0, 2\sigma^4)$ and 

$$\sqrt{T}(\tilde{\sigma}_k^2 - k\sigma^2) \overset{D}{\to} \mathcal{N}(0, 2k^2\sigma^4).$$
Hausman (1978) test: Let $\hat{\theta}_e$ be a consistent and asymptotically efficient estimator of $\theta$ and $\hat{\theta}_c$ a consistent but not asymptotically efficient estimator. Then, $\hat{\theta}_e$ 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 $\hat{\theta}_e$.

A decomposition:

$$\frac{1}{\sqrt{k}} \sqrt{T} (\tilde{\sigma}_k^2 - k\sigma^2) = \sqrt{kT} \left( \frac{\tilde{\sigma}_k^2}{k} - \sigma^2 \right)$$

$$= \sqrt{kT} \left( \frac{\tilde{\sigma}_k^2}{k} - \hat{\sigma}^2 \right) + \sqrt{kT} (\hat{\sigma}^2 - \sigma^2).$$

The LHS is $\mathcal{N}(0, 2k\sigma^4)$, and the second term on the RHS is $\mathcal{N}(0, 2\sigma^4)$. 
The first term on the RHS is thus:

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

The normalized variance ratio is

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

**Cochrane (1988)**

Letting $VR = \frac{\tilde{\sigma}^2_k}{(k\hat{\sigma}^2)}$, we have under the null of i.i.d.,

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

- Null hypothesis: i.i.d.
A strictly stationary process \( \{y_t\} \) is said to be **time reversible (TR)** if
\[
F_{t_1, t_2, \ldots, t_n} (c_1, c_2, \ldots, c_n) = F_{t_n, t_{n-1}, \ldots, t_1} (c_1, c_2, \ldots, c_n).
\]

Examples: Independent sequences, Gaussian ARMA processes.

When the condition fails, \( \{y_t\} \) is said to be **time irreversible**.

- A linear, non-Gaussian process is time irreversible in general.
- Tong (1990): “time irreversibility is the rule rather than the exception when it comes to nonlinearity” (p. 197).

A test of time reversibility can be viewed as a joint test of **linearity** and **Gaussianity** or a test of **independence**, e.g., Ramsey and Rothman (1996) and Chen, Chou, and Kuan (2000).
Cox (1981): When \( \{y_t\} \) is TR, the marginal distribution of \( y_t - y_{t-k} \) must be symmetric about the origin for any \( k \).

Existing tests of the symmetry of \( y_t - y_{t-k} \):
- Testing the third central moment being zero.
- Testing the bi-covariances being zero, because

\[
\mathbb{E}(y_t - y_{t-k})^3 = -3 \mathbb{E}(y_t^2 y_{t-k}) + 3 \mathbb{E}(y_t y_{t-k}^2).
\]

Note: These are all necessary conditions of distribution symmetry.

Drawbacks: Such tests require the data to possess high-order moments.
A distribution is symmetric iff the imaginary part of its characteristic function is zero. Hence, time reversibility implies

\[ h_k(\omega) := \mathbb{E}[\sin(\omega(y_t - y_{t-k}))] = 0, \quad \forall \omega \in \mathbb{R}^+. \]

We may integrate out \( \omega \) with a positive and integrable weighting function \( g \):

\[
\int_{\mathbb{R}^+} h_k(\omega) g(\omega) \, d\omega = \int_{\mathbb{R}} \left( \int_{\mathbb{R}^+} \sin(\omega(y_t - y_{t-k})) g(\omega) \, d\omega \right) \, dF = 0,
\]

where \( F \) is the cdf of \( y_t \).

We can test if \( \mathbb{E}[\psi_g(y_t - y_{t-k})] = 0 \), with

\[
\psi_g(y_t - y_{t-k}) = \int_{\mathbb{R}^+} \sin(\omega(y_t - y_{t-k})) g(\omega) \, d\omega.
\]
CCK Test

Chen, Chou, and Kuan (2000)

\[ C_{g,k} = \sqrt{T_k} \bar{\psi}_{g,k} / \bar{\sigma}_{g,k} \xrightarrow{D} \mathcal{N}(0, 1). \]

where \( T_k = T - k \), \( \bar{\psi}_{g,k} = \sum_{t=k+1}^{T} \psi_g(y_t - y_{t-k}) / T_k \), and \( \bar{\sigma}_{g,k}^2 \) is a consistent estimator of the asymptotic variance of \( \sqrt{T_k \bar{\psi}_{g,k}} \).

- \( \psi_g \) is bounded so that no moment condition is needed for the CLT; this test is thus robust to moment failure.
- Setting \( g(\omega) = \exp(-\omega/\beta)/\beta \) with \( \beta > 0 \) (exponential dist),
  \[ \psi_{\exp}(y_t - y_{t-k}) = \frac{\beta(y_t - y_{t-k})}{1 + \beta^2(y_t - y_{t-k})^2}; \]
  a rule of thumb is to set \( \beta = 1/\sigma_y \).
Figure: $h(\omega)$ of centered exponential distributions with $\beta = 0.5$ (line 0), $\beta + 1$ (line 1) and $\beta = 2$ (line 2).
Some Remarks

- When this test is applied to model residuals, it is difficult to estimate the asymptotic variance of $\sqrt{T_k \bar{\psi}_{g,k}}$. An easy way is to bootstrap the standard error.

- Chen and Kuan (2002): This test is powerful against asymmetric dependence in data, such as volatility asymmetry, but the existing Q-type and BDS tests are not. Thus, this test may be used to distinguish between EGARCH and GARCH models.

- We may test the $L_2$ norm of $h_k(\cdot)$: $\int_{\mathbb{R}^+} h_k(\omega)^2 \, d\omega = 0$. The resulting test does not have an analytic form and usually has a data-dependent distribution.