Estimation and Inference

- Some asymptotic results based on m.d.s

  * Estimation - Inference on μ

  * Estimation - Inference on γ_k, ρ_k

  * Estimation - Inference on the parameters of an AR(1) and of a MA(1)

Some Asymptotic Results

B-N decomposition:

$$X_t = C(L)\varepsilon_t = [C(1) + (1-L)\tilde{C}(L)]\varepsilon_t$$

What is $\tilde{C}(L)$?

$$\sum_{j=0}^{\infty} c_j L^j = \left(\sum_{j=0}^{\infty} c_j - \sum_{j=1}^{\infty} c_j\right)$$

$$+\left(\sum_{j=1}^{\infty} c_j - \sum_{j=2}^{\infty} c_j\right)L$$

$$+\left(\sum_{j=2}^{\infty} c_j - \sum_{j=3}^{\infty} c_j\right)L^2$$

$$+ \dots$$

$$= \sum_{j=0}^{\infty} c_j - \left(\sum_{j=1}^{\infty} c_j\right)(1-L)$$

$$-\left(\sum_{j=2}^{\infty} c_j\right)L(1-L)$$

$$-\left(\sum_{j=3}^{\infty} c_j\right)L^2(1-L)$$

$$= \boxed{\sum_{j=0}^{\infty} c_j} + \boxed{\sum_{j=0}^{\infty} \tilde{c}_j L^j}(1-L)$$

with $\tilde{c}_j = -\sum_{s=j+1}^{\infty} c_s$.

Lemma: $\tilde{C}(L) = \sum_{j=0}^{\infty} \tilde{c}_j L^j$; $\tilde{c}_j = \sum_{j+1}^{\infty} c_s$.

(a) $\sum_0^{\infty} |j|^{\frac{1}{2}} |c_j| < \infty \Rightarrow \sum_0^{\infty} \tilde{c}_j^2 < \infty$ $\quad (\Leftarrow \sum_1^{\infty} j^2 c_j^2 < \infty.)$

$\Uparrow$

(b) $\sum_0^{\infty} |j|\,|c_j| < \infty \Rightarrow \sum_0^{\infty} |\tilde{c}_j| < \infty$

Law of Large Numbers

Thm (SLLN for m.d.s): Let ε_t be a m.d.s with

$$V(\varepsilon_t) = \sigma_t^2 < \infty$$
$$\sum_{t=1}^{\infty} \frac{\sigma_t^2}{t^2} < \infty$$

Then $\frac{1}{n} \sum_{t=1}^{n} \varepsilon_t \xrightarrow{a.s} 0$ as $n \to \infty$

Thm (SLLN for Linear Processes)

If $X_t = C(L)\varepsilon_t$ with $\sum_{j=1}^{\infty} j^2 c_j^2 < \infty$ and ε_t a m.d.s with

$$\sup_t E|\varepsilon_t|^2 < \infty,$$

then $\frac{1}{n} \sum_{t=1}^{n} X_t \xrightarrow{a.s} 0$ as $n \to \infty$

Proof:

$$X_t = C(1)\varepsilon_t + (1-L)\tilde{C}(L)\varepsilon_t$$

$$\frac{1}{n}\sum_{t=1}^{n} X_t = C(1)\frac{\sum_{t=1}^{n}\varepsilon_t}{n} + \frac{1}{n}\left(\tilde{\varepsilon}_0 - \tilde{\varepsilon}_n\right)$$

where $\tilde{\varepsilon}_j = \tilde{C}(L)\varepsilon_j$

By the SLLN for m.d.s

$$\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t \xrightarrow{a.s} 0$$

Now we only have to prove that

$$\frac{1}{n}\left(\tilde{\varepsilon}_0-\tilde{\varepsilon}_n\right)\xrightarrow{a.s} 0$$

Take for instance $\dfrac{\tilde{\varepsilon}_n}{n}$:

$$\sum_{n=1}^{\infty} P\left(\frac{|\tilde{\varepsilon}_n|}{n}>\delta\right) < \sum_{n=1}^{\infty}\frac{E(\tilde{\varepsilon}_n^{\,2})}{n^2\delta^2} = \frac{1}{\delta^2}\sum_{1}^{\infty}\frac{1}{n^2}\left(\sum_{0}^{\infty}\tilde{c}_j^{\,2}\right)K$$

$$< \infty$$

By Markov's Inequality

$$\left(P(|X|\geq\varepsilon)\leq\frac{E|X|^p}{\varepsilon^p}\text{ if } E|X|^p<\infty \text{ and } p>0\right)$$

where $K=\sup_t E(|\varepsilon_t|^2)<\infty$ and $\sum_0^\infty \tilde{c}_j^{\,2}<\infty$.

By Borel-Cantelli Lemma $\dfrac{\tilde{\varepsilon}_n}{n}\xrightarrow{a.s} 0$

$$\left(\sum_{n=1}^{\infty} P(|X_n(w)-X(w)|\geq\varepsilon)<\infty,\text{ then } X_n\xrightarrow{a.s} X\right)$$

Central Limit Theorems

Thm (CLT for m.d.s) Let ε_t be a strictly stationary and ergodic m.d.s with

$$V(\varepsilon_t) = \sigma^2 < \infty.$$

Then
$$\frac{\sum_{t=1}^{n} \varepsilon_t}{\sqrt{n}} \xrightarrow{d} N(0, \sigma^2)$$

Thm (CLT for linear processes)

If $X_t = C(L)\varepsilon_t$ with $\sum_{j=1}^{\infty} |j| |c_j| < \infty$

and ε_t a strictly stationary and ergodic m.d.s with $V(\varepsilon_t) = \sigma^2 < \infty$,

then
$$\sqrt{n} \frac{\sum_{t=1}^{n} X_t}{n} \xrightarrow{d} N(0, \underbrace{C^2(1)\sigma^2}_{\text{LONG RUN VARIANCE}})$$

Proof

$$\sqrt{n}\,\bar{X}_n = \frac{\sum_{t=1}^{n} X_t}{\sqrt{n}} = C(1) \underbrace{\frac{\sum_{t=1}^{n} \varepsilon_t}{\sqrt{n}}}_{\text{By CLT for m.d.s}} + \frac{1}{\sqrt{T}}\left(\tilde{\varepsilon}_0 - \tilde{\varepsilon}_n\right)$$

$$\xrightarrow{d} N(0, C^2(1)\,\sigma_\varepsilon^2)$$

We need to prove that $\frac{1}{\sqrt{T}}(\tilde{\varepsilon}_0 - \tilde{\varepsilon}_n) = o_p(1)$

Notice that we cannot use now the Borel-Cantelli argument used to prove the SLLN.

This will be done by proving $E|\tilde{\varepsilon}_n| < \infty$ and then applying Markov's inequality

$$E|\tilde{\varepsilon}_n| = E\left|\sum_{j=0}^{\infty} \tilde{c}_j \varepsilon_{n-j}\right| \leq \sum_{j=0}^{\infty} |\tilde{c}_j| E|\varepsilon_{n-j}| < \infty$$

Example: AR(1) ; $X_t = \phi X_{t-1} + \varepsilon_t$

and ε_t a m.d.s satisfying the conditions of the previous theorem

$$C(1) = \frac{1}{1-\phi L}, \text{ then } \sqrt{n}(\bar{X}_n) \xrightarrow{d} N\left(0, \frac{\sigma^2}{(1-\phi)^2}\right)$$

or

$$\sqrt{n}(\bar{X}_n) \xrightarrow{d} N\left(0, \gamma(X_t)\frac{1+\phi}{1-\phi}\right)$$

because $V(X_t) = \frac{\sigma^2}{1-\phi^2}$.

Notice the problem when $\phi = 1$.

Some Results for sample correlations

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (X_t - \bar{X}_n)(X_{t+h} - \bar{X}_n) \quad 0 \le h \le n-1$$

and

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

Thm: If $\{X_t\}$ is the stationary process

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j} \qquad \{\varepsilon_t\} \sim iid(0, \sigma^2)$$

where $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ and (*) $\boxed{E\,\varepsilon_t^4 < \infty}$, then

for each $h \in \{1, 2, \ldots\}$ we have

$$\sqrt{n}\,(\hat{\rho}(h) - \rho(h)) \xrightarrow{d} N(0, W)$$

where

$$\hat{\rho}(h) = \begin{bmatrix} \hat{\rho}(1) \\ \hat{\rho}(2) \\ \vdots \\ \hat{\rho}(h) \end{bmatrix} \quad \text{and} \quad \rho(h) = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(h) \end{bmatrix}$$

and W is the covariance matrix whose $(i-j)$ element is given by Bartlett's formula

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

(*) This condition can be relaxed at the expense of $\sum |j| \psi_j^2 < \infty$.

Simple algebra shows

$$W_{ij} = \sum_{k=1}^{\infty} \{ \rho(k+i) + \rho(k-i) - 2\rho(i)\rho(k) \}$$
$$\times \{ \rho(k+j) + \rho(k-j) - 2\rho(j)\rho(k) \}$$

and

$$W_{ii} = \sum_{k=1}^{\infty} \left(\rho_{i+k} + \rho_{i-k} - 2\rho_i \rho_k \right)^2 .$$

Confidence intervals will be formed by

$$\boxed{\pm C_\alpha \sqrt{\frac{W_{ii}}{n}}}$$ with $C_\alpha = CV_{\frac{\alpha}{2}}$ from $N(0,1)$

Examples

White noise : $W_{ii} = 1$ $\quad \boxed{\pm 1.96 \frac{1}{\sqrt{n}}}$ 95%

$MA(q)$: $W_{ii} = [1 + 2\rho^2(1) + 2\rho^2(2) + \dots + 2\rho^2(q)]$
$i > q$

$MA(1)$: $\pm 1.96 \sqrt{\frac{[1 + 2\hat{\rho}^2(1)]}{n}}$ $\quad i > 1$

$AR(1)$: $W_{ii} = \sum_{k=1}^{i} \phi^{2i} (\phi^{-k} - \phi^{k})^2 + \sum_{k=i+1}^{\infty} \phi^{2k} (\phi^{-i} - \phi^{i})^2$

$$= (1 - \phi^{2i})(1 + \phi^2)(1 - \phi^2)^{-1} - 2i\phi^{2i} \quad {}_{i=1,2}.$$

$$\boxed{\simeq (1+\phi^2)/(1-\phi^2)}$$ for i large

Estimation and Inference of an AR(1)

Suppose X_t is a stationary autoregressive process of order one satisfying

$$X_t = \phi X_{t-1} + \varepsilon_t$$

where $|\phi| < 1$ and $\varepsilon_t \sim iid$ $V(\varepsilon_t) = \sigma^2 < \infty$.

Then $X_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$ is strictly stationary and ergodic.

$$\hat{\phi} = \frac{\sum X_t X_{t-1}}{\sum X_{t-1}^2} = \phi + \frac{\sum X_{t-1} \varepsilon_t}{\sum X_{t-1}^2}$$

$$\sqrt{n}\left(\hat{\phi} - \phi\right) = \frac{\frac{1}{\sqrt{n}} \sum_{t=1}^{n} X_{t-1} \varepsilon_t}{\frac{1}{n} \sum X_{t-1}^2}$$

$$\frac{1}{n} \sum_{t}^{n} X_{t-1}^2 \xrightarrow{a.s} E(X_t^2) = \frac{\sigma^2}{1-\phi^2}$$ by the Ergodic Theorem

Observe that $Z_t = \varepsilon_t X_{t-1}$ is a m.d.s with respect $\mathcal{F}(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots)$

$$E\left[Z_t \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \dots\right] = X_{t-1} E(\varepsilon_t) = 0.$$

Z_t as a function of ε_t, y_{t-1} is clearly strictly stationary and ergodic.

$$V(\varepsilon_t X_{t-1}) = E(\varepsilon_t^2 X_{t-1}^2) = \sigma \, E(X_{t-1}^2) = \frac{\sigma^4}{1-\phi^2} < \infty$$

Applying the CLT for m.d.s

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} X_{t-1} \varepsilon_t \xrightarrow{d} N\left(0, \frac{\sigma^4}{1-\phi^2}\right).$$

Therefore

$$\sqrt{n}\,(\hat{\phi}_n - \phi) \xrightarrow{d} N(0, 1-\phi^2)$$

What if $\phi = 1$? Problem again

Estimation and Inference MA(1)

$$Y_t = \varepsilon_t + \theta \varepsilon_{t-1} \quad |\theta|<1, \ \varepsilon_t \sim iid(0,\sigma^2)$$

The first estimator can be obtained from

$$\rho_1 = \frac{\theta}{(1+\theta^2)} .$$

Estimating ρ_1 by

$$\hat{\rho}_1 = \frac{\sum_2^n (Y_t - \bar{Y}_n)(Y_{t-1} - \bar{Y}_n)}{\sum_1^n (Y_t - \bar{Y}_n)^2}$$

we obtain

$$\hat{\theta}_r = \begin{cases} \dfrac{1-[1-4\hat{\rho}_1]^{\frac{1}{2}}}{2\hat{\rho}_1} & 0 < |\hat{\rho}_1| < 0.5 \\ -1 & \hat{\rho}_1 < -.5 \\ 1 & \hat{\rho}_1 > .5 \\ 0 & \hat{\rho}_1 = 0 \end{cases}$$

More efficient estimator can be obtained by LS or by ML.

By LS :

$$\varepsilon_t = -\theta \varepsilon_{t-1} + Y_t$$

$$Y_t = -\sum_{j=1}^{t-1} (-\theta)^j Y_{t-j} - (-\theta)^t \varepsilon_0 + \varepsilon_t$$

$$= \boxed{f_t(Y; \theta, \varepsilon_0) + \varepsilon_t}$$

where $f_t(Y;\theta, \varepsilon_0) = \theta \varepsilon_0$

$$f_t(Y;\theta,\varepsilon_t) = -\sum_{j=1}^{t-1} (-\theta) Y_{t-j} - (-\theta)^t \varepsilon_0$$

LS in $y_t = f_t(y; \theta, \varepsilon_0) + \varepsilon_t$

i) inside the classical NLS set-up that you have learned in previous econometric courses. In fact because we can obtain an initial consistent estimator of θ^0 we will obtain our asymptotic results from a one-step minimization procedure (one-step Gauss-Newton estimator)

Let's assume we have an initial estimator $\tilde{\theta}$ satisfying $\tilde{\theta} - \theta = o_p(n^{-\frac{1}{4}})$ and $\tilde{\varepsilon}_0 = O_p(1)$

The one-step Gauss-Newton estimator of θ is obtained by regressing

$$\varepsilon_t(y; \tilde{\theta}) = y_t - f_t(y; \tilde{\theta}; \hat{\varepsilon}_0) = \sum_{j=0}^{t-1} (-\tilde{\theta})^j y_{t-j} + (-\tilde{\theta})^t \tilde{\varepsilon}_0$$

on the first derivative of $f_t(y; \theta, \varepsilon_0)$ evaluated at $\theta = \tilde{\theta}$, that derivative is

$$W_t(y; \tilde{\theta}) = \begin{cases} \tilde{\varepsilon}_0 & t = 1 \\ \sum_{j=1}^{t-1} j(-\tilde{\theta})^{j-1} y_{t-j} + t(-\tilde{\theta})^{t-1} \tilde{\varepsilon}_0 ; & t = 2, 3, \ldots n \end{cases}$$

Regressing $\varepsilon_t(y; \tilde{\theta})$ on $W_t(y; \tilde{\theta})$ we obtain an estimator of $\theta - \tilde{\theta}$. The improved estimator of θ is then

$$\hat{\theta} = \tilde{\theta} + \Delta\hat{\theta} \qquad (*)$$

where $\Delta\hat{\theta} = \left[\sum_{t=1}^{n} [W_t(y; \tilde{\theta})]^2\right]^{-1} \sum_{t=1} \varepsilon_t(y; \tilde{\theta}) W_t(y; \tilde{\theta})$

Thm (Asymptotic Normality for MA(1))

Let $Y_t = \varepsilon_t + \theta^0 \varepsilon_{t-1}$, where $|\theta^0| < 1$, $\varepsilon_t \text{ iid}(0, \sigma^2)$ with $E\{|\varepsilon_t|^{2+\nu}\} < L < \infty$ for some $\nu > 0$.

Let $\tilde{\varepsilon}_0$ and $\tilde{\theta}$ be initial estimators satisfying $\tilde{\varepsilon}_0 = O_p(1)$, $\tilde{\theta} - \theta = o_p(n^{-\frac{1}{4}})$, and $|\tilde{\theta}| < 1$.

Then

$$\sqrt{n}\,(\hat{\theta} - \theta^0) \xrightarrow{d} N(0, 1 - (\theta^0)^2)$$

where $\hat{\theta}$ is defined in (*). Also,

$\hat{\sigma}^2 \xrightarrow{p} (\sigma^0)^2$, where σ^0 is the true value of σ and

$$\hat{\sigma}^2 = \frac{\sum_{t=1} \varepsilon_t^2(Y; \hat{\theta})}{n}.$$

The Hannan-Rissanen algorithm

(1) $\hat{\varepsilon}_t = y_t - P[y_t \mid y_{t-1}, y_{t-2}, \dots]$

(2) $y_t = \theta\, \hat{\varepsilon}_{t-1} + u_t$

$$\hat{\theta} = \frac{\sum y_t \hat{\varepsilon}_{t-1}}{\sum \hat{\varepsilon}_{t-1}^2}$$

Check how good performs this alternative estimator.

Box-Pierce with the true errors

Basic Idea: Use the correlations

$$\hat{\rho}_i = \frac{\sum_{t=1}^{n} y_t y_{t-i}}{\sum_{t=1}^{n} y_{t-i}^2}$$ to test

$H_0: y_t = \varepsilon_t$ white noise

By the CLT on correlations we have $\sqrt{n}\,(\underset{(k\times 1)}{\hat{\rho}_k} - \underset{(k\times 1)}{\rho_k}) \overset{d}{\sim} N(0, \underset{(k\times k)}{W})$

Under H_0,

$$\sqrt{n}\,(\hat{\rho}_k) \overset{d}{\sim} N(0, \underset{k\times k}{I})$$

Therefore

$$\boxed{n \sum_{j=1}^{k} (\hat{\rho}_j)^2 \xrightarrow{d} \chi^2_k} \equiv Q$$ Box-pierce statistic

If we had fitted an ARMA(p,q) the # d.freedom $= k-(p+q)$.

This test is testing

$H_0: \mathrm{Cov}(y_t\, y_{t-j}) = 0 \quad j = 1 \ldots k$

$H_a: \mathrm{Cov}(y_t\, y_{t-j}) \neq 0$ for at least one j.

A modification $Q^* = n(n+2) \sum_{j=1}^{k} (n-j)^{-1} (\hat{\rho}_j)$ Ljung-Box

Box-Pierce with sample autocorrelations calculated from residuals

A more realistic case is to assume that we have a model

$$y_t = X_t'\beta + \varepsilon_t \qquad t = 1 \dots n$$

and we want to test if the errors from this model are white noise.

We do not observe the errors but the residuals $\hat{\varepsilon}_t$.

Before
$$\begin{cases} \rho_j \equiv \dfrac{\gamma_j}{\gamma_0} \quad (j = 1, 2 \dots) \\ \gamma_j \equiv \dfrac{1}{n} \displaystyle\sum_{t=j+1}^{n} \varepsilon_t \varepsilon_{t-j} \end{cases}$$

Now
$$\begin{cases} \hat{\rho}_j \equiv \dfrac{\hat{\gamma}_j}{\hat{\gamma}_0} \\ \hat{\gamma}_j \equiv \dfrac{1}{n} \displaystyle\sum_{t=j+1}^{n} \hat{\varepsilon}_t \hat{\varepsilon}_{t-j} \end{cases}$$

Is it right to use $\hat{\rho}_j$ (calculated from the residuals) instead of ρ_j and the residual-based Q statistics derived from $\{\hat{\rho}_j\}$ for testing for serial correlation?

YES,

but only if the regressors are strictly exogenous: $E(\varepsilon_i | X) = 0 \quad i = 1 \dots n$

This assumption is too strong for time series data. The strict exogeneity assumption implies that for any regressor k, $E(x_{jk}\, \varepsilon_i) = 0$ $\forall i, j$ not only $i = j$. For instance an AR(1) model

$$y_i = \beta\, y_{i-1} + \varepsilon_i \qquad i = 1, 2, \ldots n$$

$$E[y_i \varepsilon_i] = \beta E[y_{i-1} \varepsilon_i] + E(\varepsilon_i^2)$$
$$= E[\varepsilon_i^2] \neq 0$$

So the regressor is not orthogonal to the past error term (y_i is the regressor for observation $i+1$)

$$\hat{\gamma}_j \equiv \frac{1}{n} \sum_{t=j+1}^{n} \hat{\varepsilon}_t \hat{\varepsilon}_{t-j}$$

$$= \frac{1}{n} \sum_{t=j+1}^{n} \left[\varepsilon_t - x_t'(\hat{\beta} - \beta)\right]\left[\varepsilon_{t-j} - x_{t-j}'(\hat{\beta} - \beta)\right]$$

$$= \gamma_j - \frac{1}{n} \sum_{t=j+1}^{n} (x_{t-j}\, \varepsilon_t + x_t\, \varepsilon_{t-j})'(\hat{\beta} - \beta)$$

$$+ (\hat{\beta} - \beta)\left(\frac{1}{n} \sum_{j+1}^{n} x_t x_{t-j}'\right)(\hat{\beta} - \beta)$$

If $E(x_t\, \varepsilon_{t-j})$, $E(x_{t-j}\, \varepsilon_t)$, and $E(x_t x_{t-j}')$ are all finite, then because $\hat{\beta} - \beta \xrightarrow[p]{} 0$,

we have

$$\hat{\gamma}_j - \gamma_j \xrightarrow{p} 0.$$

However $\sqrt{n}\,(\hat{\gamma}_j - \gamma_j) \not\to 0.$

$$\sqrt{n}\,\hat{\gamma}_j = \sqrt{n}\,\gamma_j - \overbrace{\frac{1}{n}\sum\left(x_{t-j}\varepsilon_t + x_t\varepsilon_{t-j}\right)'}^{??}\;\overbrace{\sqrt{n}(\hat{\beta}-\beta)}^{O_p(1)}$$

$$+\underbrace{\underbrace{\sqrt{n}\,(\hat{\beta}-\beta)'}_{O_p(1)}\;\underbrace{\left(\frac{1}{T}\sum_{j+1}^{T} x_t x_{t-j}'\right)}_{O_p(1)}\;\underbrace{(\hat{\beta}-\beta)}_{o_p(1)}}_{o_p(1)}$$

$$\frac{1}{n}\sum_{t=1}^{n}\left(x_{t-j}\,\varepsilon_t + x_t\,\varepsilon_{t-j}\right) \xrightarrow[p]{} E\left(x_{t-j}\,\varepsilon_t\right) + \underbrace{E(x_t\varepsilon_{t-j})}_{=\,0}$$

if all the regressors are strictly exogenous

What if the regressors are predetermined but not strictly exogenous?

Predetermined regressors: $E(x_{ik}\,\varepsilon_i) = 0 \quad \forall i$, $k = 1, 2, \ldots K$

The AR(1) satisfies this assumption.

When the regressors are not strictly exogenous, we need to modify the Q statistic to restore its asymptotic distribution. For this purpose we will impose two restrictions

(stronger form of predeterminedness)

$$E(\varepsilon_t \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, x_t, x_{t-1} \ldots) = 0$$

(stronger form of ~~homo~~skedasticity)

$$E(\varepsilon_t^2 \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, x_t, x_{t-1} \ldots) = \sigma^2 > 0$$

Proposition (testing for serial correlation with predetermined regressors)

Suppose $\underset{1\times 1}{y_t} = \underset{1\times k}{x_t'} \underset{k\times 1}{\beta} + \varepsilon_t \quad (t = 1, 2, \ldots T)$

with
- $\underset{k+1}{\{y_t \; x_t\}}$ jointly stationary and ergodic
- $E(x_t x_t') = \Sigma_{xx}$ full rank
- $E(\varepsilon_t \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots x_t, x_{t-1}, \ldots) = 0$
- $E(\varepsilon_t^2 \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots x_t, x_{t-1}, \ldots) = \sigma^2 > 0$

Then

$$\sqrt{n}\, \hat{\gamma} \xrightarrow{d} N(0, \sigma^4 (I_p - \Phi)) \quad \text{and}$$

$$\sqrt{n}\, \hat{\rho} \xrightarrow[d]{} N(0, I_p - \Phi)$$

where

$$\hat{\gamma} = (\hat{\gamma}_1, \ldots \hat{\gamma}_p)', \quad \hat{\rho} = (\hat{\rho}_1, \hat{\rho}_2, \ldots \hat{\rho}_p)'$$

and

ϕ_{jk} ($\equiv$ (j,k) element of the $p\times p$ matrix Φ)

is given by

$$\phi_{jk} = E(x_t \varepsilon_{t-j})' E(x_t x_t')^{-1} E(x_t \varepsilon_{t-k}) / \sigma^2$$

Proof: Hayashi (pg 165)

By the Ergodic Theorem, matrix Φ is consistently estimated by its sample counterpart

$$\hat{\Phi} \equiv (\hat{\phi}_{j,k}), \quad \hat{\phi}_{jk} \equiv \hat{\mu}_j' S_{xx}^{-1} \hat{\mu}_k / s^2$$

$$(j,k = 1,2,\ldots p).$$

where

$$s^2 \equiv \frac{1}{n-K} \sum_{t=1}^{T} \hat{\varepsilon}_t^2 \; ; \quad \hat{\mu}_j \equiv \frac{1}{n} \sum_{t=j+1}^{T} x_t \hat{\varepsilon}_{t-j}$$

From last proposition we have

$$\left[\text{modified Box-Pierce } Q \equiv n \hat{\rho}' (I_p - \hat{\Phi})^{-1} \hat{\rho} \xrightarrow{d} \chi^2_{(p)} \right]$$

<u>An Auxiliary Regression-Based Test</u>

Regress $\hat{\varepsilon}_t$ on $x_t, \hat{\varepsilon}_{t-1}, \hat{\varepsilon}_{t-2}, \ldots \hat{\varepsilon}_{t-p}$

$$pF \sim \chi^2(p)$$

The F statistic for the hypothesis that the p coefficients of $\hat{\varepsilon}_{t-1}, \hat{\varepsilon}_{t-2}, \ldots \hat{\varepsilon}_{t-p}$ are all zero.

an

$$nR^2 \sim \chi^2(p) \quad \text{too.}$$