Questions and Answers on Unit Roots, Cointegration, VARs and VECMs

L. Magee Winter, 2012

---

1. Let ϵ_t, $t = 1, \ldots, T$ be a series of independent draws from a N[0,1] distribution. Let w_t, $t = 1, \ldots, T$, be another series of independent draws from a N[0,1] distribution. Each w_t is distributed independently of all of the ϵ_t's. Let $x_t = \sum_{s=1}^{t} \epsilon_s$ and let $z_t = x_t + w_t$.

(a) Express x_t as a function of x_{t-1} and ϵ_t.

(b) Is x_t I(0) or I(1)? Explain briefly.

(c) Is z_t I(0) or I(1)?

(d) Are x_t and z_t cointegrated?

(e) Convert the formula $z_t = x_t + w_t$ to an error correction model, with Dz_t on the left hand side; and Dx_t and the lagged error correction term $z_{t-1} - x_{t-1}$ on the right hand side.

2. Let $y_t = \alpha + \beta x_t + \epsilon_t$. What can be said about the order of integration of ϵ_t in the following cases? Can we tell if ϵ_t is I(0) or I(1)?) Assume that ϵ_t is not correlated with x_t.

(a) x_t is I(0) and y_t is I(0)

(b) x_t is I(1) and y_t is I(0)

(c) x_t is I(0) and y_t is I(1)

(d) x_t is I(1) and y_t is I(1)

3. Let y_t and x_t be I(1) random variables, where

$$y_t = 3 + .8x_t + \epsilon_t$$

and ϵ_t is I(0), and let the following relationship hold

$$y_t = 5.8 + .4y_{t-1} + 1.3x_t - .82x_{t-1} + u_t$$

where u_t is I(0) and $Eu_t = 0$. Write an error correction model implied by these two equations.

4. Given the 2-variable regression model

$$y_t = \alpha + \beta x_t + u_t \quad \text{where } Eu_t = 0$$

(a) describe the important difference between the statistical properties of the t-statistic for testing $H_0 : \beta = 0$ in the following two cases.
Case A: x_t and y_t are I(0)
Case B: x_t and y_t are random walks, and are statistically independent of each other

(b) describe the important difference between the statistical properties of the OLS estimator of β in the following two cases.
Case A: x_t and y_t are I(0)
Case C: x_t and y_t are I(1) and cointegrated.

5. Suppose that $\hat{\beta}$ is a superconsistent estimator of β, and that $\hat{\alpha}$ is merely a "usual" $\sqrt{n}$-consistent estimator of α. What, if anything, does this imply about

(a) which variance is larger, $\text{Var}(\hat{\alpha})$ or $\text{Var}(\hat{\beta})$, if both are based on a sample of size $n = 50$?

(b) the ratio $\text{Var}(\hat{\alpha})/\text{Var}(\hat{\beta})$ as $n \to \infty$?

6. Give a short answer for each of the following.

(a) Describe the (Granger) causality test procedure for investigating whether x_t causes y_t.

(b) Describe the spurious regression problem that could be encountered when estimating the model $y_t = \alpha + \beta x_t + u_t$ by ordinary least squares.

(c) What properties are implied by the statement "y_t is covariance stationary"?

(d) What is meant by the statement "y_t is I(2)"?

(e) What is meant by the statement "x_t and y_t are cointegrated"?

7. Describe an augmented Dickey-Fuller test procedure for testing for the presence of a unit root in the scalar time series process y_t.

8. y_t and x_t are cointegrated I(1) random variables, with cointegrating equation

$$y_t = 1 + 3x_t + \epsilon_t$$

and ϵ_t is I(0). They also are related according to the autoregressive distributed lag (ARDL) model

$$y_t = 4 + .5y_{t-1} + 2x_t + \theta x_{t-1} + u_t$$

where u_t is I(0) and $Eu_t = 0$. Determine the numerical value of θ.

9. Let y_t be an m-element vector of $I(1)$ time series variables and let ϵ_t be an $I(0)$ white noise scalar error term in the vector error correction model

$$Dy_t = \mu + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i^* Dy_{t-i} + \epsilon_t$$

(a) Explain how the rank of the $m \times m$ coefficient matrix Π relates to the number of cointegrating relationships among the elements of y_t.

(b) For the case $m = 3$, write a numerical value of Π that implies the existence of two cointegrating relationships.

(c) Suppose that Π has full rank. What does this imply about the order of integration of the elements of y_t, and/or its cointegrating relationships?

10. (a) Briefly describe the assumptions of a linear regression model with time series data for which the ordinary least squares estimator of a coefficient is

(i) consistent

(ii) superconsistent

(iii) inconsistent

(b) What is the difference in the behaviour of the variance of a consistent estimator and a superconsistent estimator as $n \to \infty$?

Answers

1. (a) $x_t = x_{t-1} + \epsilon_t$

(b) x_t is I(1), since $\rho = 1$ when x_t is expressed as $x_t = \mu + \rho x_{t-1} + \epsilon_t$

(c) z_t is I(1). $z_t = x_t + w_t$. x_t is I(1) and w_t is I(0), so $x_t + w_t$ must be I(1).

(d) Yes, x_t and z_t are cointegrated since there is a linear combination: $z_t - x_t = w_t$, that is I(0).

(e)

$$\begin{aligned} z_t &= x_t + w_t \\ z_t - z_{t-1} &= x_t - z_{t-1} + w_t \\ Dz_t &= x_t - x_{t-1} - (z_{t-1} - x_{t-1}) + w_t \\ Dz_t &= Dx_t - (z_{t-1} - x_{t-1}) + w_t \end{aligned}$$

2. (a) ϵ_t is I(0). A weighted sum of a fixed number of I(0) series is I(0).

(b) $\beta = 0$ and ϵ_t is I(0). We cannot have ϵ_t being I(1) and cancelling the I(1) part of βx_t since ϵ_t is not correlated with x_t, so β must be equal to 0 for y_t to be I(0).

(c) ϵ_t is I(1). Something must be I(1) on the right-hand side since y_t is I(1).

(d) ϵ_t could be I(1) or I(0). It can be I(0) if y_t and x_t are cointegrated and β is set at the value where the I(1) parts of y_t and βx_t are equal. Otherwise ϵ_t is I(1).

3. For an ECM, convert the second equation to one involving only I(0) variables. Here, these are Dy_t, Dx_t, and the error correction term $y_{t-1} - (3 + .8x_{t-1})$.

$$\begin{aligned}
y_t &= 5.8 + .4y_{t-1} + 1.3x_t - .82x_{t-1} + u_t \\
Dy_t &= 5.8 - .6y_{t-1} + 1.3x_t - 1.3x_{t-1} + 1.3x_{t-1} - .82x_{t-1} + u_t \\
Dy_t &= 5.8 - .6y_{t-1} + 1.3Dx_t + .48x_{t-1} + u_t \\
Dy_t &= 5.8 + 1.3Dx_t - .6(y_{t-1} - .8x_{t-1}) + u_t \quad \text{, or} \\
Dy_t &= 4 + 1.3Dx_t - .6(y_{t-1} - (3 + .8x_{t-1})) + u_t
\end{aligned}$$

4. (a) case A: If x_t and y_t are I(0) and the u_t's are normally distributed with no auto or het then the t-stat has a t distribution under the null $\beta = 0$. Even if the u_t's are not normally distributed, as long as there is no auto or het and it has a finite mean and variance, then it will have this t-distribution property as the sample size grows large (which is the same as saying that t converges in distribution to a standard normal).

case B: If x_t and y_t are independent random walks, the t-distribution property given above does not hold. The t-stat tends to be larger in magnitude as the sample size grows, even when $\beta = 0$. This is the spurious regression problem.

(b) In case C, when x_t and y_t are I(1) and cointegrated, the OLS estimator of β is "superconsistent", meaning that as the sample size n grows, the OLS variance approaches zero at the rate $1/n^2$, which is faster than in case A, where the variance of OLS approaches zero at the rate $1/n$.

5. (a) Superconsistency is a property about the convergence rate (to zero) of the variance as the sample size goes to infinity. We cannot say which variance is larger at any fixed sample size.

(b) As $n \to \infty$, $\text{Var}(\hat{\alpha})$ approaches zero at a rate $1/n$ whereas $\text{Var}(\hat{\beta})$ approaches zero at a rate $1/n^2$. Therefore the ratio $\text{Var}(\hat{\alpha})/\text{Var}(\hat{\beta})$ increases at a rate n.

6. (a)
- regress y_t on p lags of y_t and p lags of x_t (and possibly other regressors, such as p lags of another variable, z_t).
- test the joint null hypothesis that the coefficients of all the lagged x_t variables equal zero.
- this null hypothesis in words is "x_t does not Granger-cause y_t" and the alternative hypothesis (which is that at least one of the coefficients of the lagged x_t variables does not equal zero) in words is "x_t Granger-causes y_t".

- use an F-test.

(b) If x_t and y_t have unit roots, then a t-test of the null hypothesis $\beta = 0$ is rejected with probability approaching one as $n \to \infty$, even when it is true.

(c) Ey_t, $\text{Var}(y_t)$ and $\text{Cov}(y_t, y_{t-s})$ do not depend on t.

(d) y_t is not stationary, Dy_t is not stationary, and $D^2 y_t$ is stationary.

(e) x_t and y_t are I(1), but there is a value of β for which $y_t - \beta x_t$ is I(0).

7.
- Estimate by OLS the equation

$$Dy_t = \mu + \gamma^* y_{t-1} + \beta t + \sum_{i=1}^{p} \phi_i D y_{t-i} + \epsilon_t$$

- The βt term could be removed and it would produce a different version of an augmented Dickey-Fuller (ADF) test. If the $\sum_{i=1}^{p} \phi_i D y_{t-i}$ terms are removed, it would be called a Dickey-Fuller (i.e. not augmented) test.
- The ADF test statistic is the usual t-statistic for testing $H_0 : \gamma^* = 0$, which is the hypothesis that y_t has a unit root. The alternative hypothesis is one-sided, $H_a : \gamma^* < 0$, which is interpreted as "y_t does not have a unit root".
- Accept H_0 if $t > t_{crit}$ and reject H_0 if $t < t_{crit}$, where t_{crit} is a critical value satisfying $\text{Prob}(ADF < t_{crit}) = \alpha$, where ADF is a random variable having the "null distribution" (the distribution that the ADF statistic has when H_0 is true) and α is the significance level.
- The null distribution is not Student's t, but another distribution that can be looked up in special tables or supplied by econometric software.

8. Convert the ARDL model to an error correction model.

$$\begin{aligned}
y_t &= 4 + .5y_{t-1} + 2x_t + \theta x_{t-1} + u_t \\
y_t - y_{t-1} &= 4 + .5y_{t-1} - y_{t-1} + 2x_t + \theta x_{t-1} + u_t \\
Dy_t &= 4 - .5y_{t-1} + 2x_t + \theta x_{t-1} + u_t \\
&= 4 - .5y_{t-1} + 2x_t - 2xt - 1 + 2xt - 1 + \theta x_{t-1} + u_t \\
&= 4 - .5y_{t-1} + 2Dx_t + (2 + \theta)x_{t-1} + u_t \\
&= 4 + 2Dx_t + (-.5y_{t-1} + (2 + \theta)x_{t-1}) + u_t \\
&= 4 + 2Dx_t + -.5(y_{t-1} - (4 + 2\theta)x_{t-1}) + u_t
\end{aligned}$$

There is only one value of θ which will give an I(0) error correction term, implied by the cointegrating equation $y_t = 1 + 3x_t + \epsilon_t$. We need $4 + 2\theta = 3$, therefore $\theta = -\frac{1}{2}$.

9. (a) The rank of Π equals the number of cointegrating relationships among the elements of y_t, unless rank(Π) $= m$. For that case, see part (c).

(b) Write any 3×3 matrix for which two rows (or columns) are linearly independent, and the remaining row (or column) is a linear combination of the other two. Two examples are

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 0 & 2 \\ 3 & 0 & 3 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 5 & 8 \end{bmatrix}$$

(c) If Π has full rank, then the variables in Πy_{t-1} span the same vector space as the variables in y_{t-1} itself. (See Appendix A.3 of Greene's 5^{th} edition for background material about vector spaces.) Therefore at least one element of Πy_{t-1} is I(1). But this is not possible given the assumptions since none of the other terms in the VECM model are I(1). Therefore if rank(Π) $= m$, the assumptions of the VECM must be modified in one of two ways. Either every element of y_t is I(0), or else at least one element of ϵ_t is I(1).

10. (a) (i) an e.g. of a regression model in which the OLS estimator of β is consistent: $y_t = \alpha + x_t\beta + \epsilon_t$, where x_t and ϵ_t are stationary, $E\epsilon_t = 0$ and $Ex_t\epsilon_t = 0$

(ii) an e.g. where the OLS estimator of β is superconsistent: $y_t = \alpha + x_t\beta + \epsilon_t$, where x_t is I(1) and ϵ_t is I(0), $E\epsilon_t = 0$ and $Ex_t\epsilon_t = 0$

(iii) an e.g. where the OLS estimator of β is inconsistent: $y_t = \alpha + x_t\beta + \epsilon_t$, where x_t and ϵ_t are stationary, and $Ex_t\epsilon_t \neq 0$

(b) As $n \to \infty$, the variance of a consistent estimator approaches zero at a rate $1/n$ whereas the variance of a superconsistent estimator approaches zero at a rate $1/n^2$.